A Quantitative Assessment on the Effectiveness of the International Climate Change Regime

Master's thesis in Political Science

Department of Sociology and Political Science

NTNU, Trondheim (Spring 2011)

Hanne Seter

Acknowledgements

This thesis would not have been possible without help and advice from several people. I'm therefore indebted to many.

First, I would like to thank my supervisor Gunnar Fermann for guiding, revising, and encouraging me in the process. He has had faith in my thesis from the very first time we spoke, he has always shown great interest and engagement, and he has always provided me with excellent guidance. I am also very grateful for the help and advice provided by my second supervisor Karin Dyrstad, for answering my many questions.

I am also grateful for comments from Anna Barnwell, Gina Olafsson, and Beate Øien. Participants at the VIP-seminar at ISS, NTNU also deserve a thank for their comments. I also want to thank my class mates, kull 17, for making this process fun. My room mates at Brakka, Lillian Hansen and Walther Sommer, deserve an additional mentioning for listening to my questions and discussions during the last year.

Last, I would like to thank Terje Bjørnås. For all your encouraging words, patience, language corrections, and for always taking the time to listen to me.

Any remaining errors are my responsibility alone.

Trondheim, 08.06.2011

Hanne Seter

Contents

List of Tables and Figures

TABLES

FIGURES

List of Abbreviations and Acronyms

CH_4: Methane

CO_2: Carbon Dioxide

CDM: Clean Development Mechanism

CFC: Chlorofluorocarbons

COP: Conference of the Parties

EiT: Economies in Transition

GHG: Greenhouse Gas

HFC: Hydrofluorcarbon

ICCR: International Climate Change Regime

IPCC: Intergovernmental Panel on Climate Change

N_2O: Nitrous Oxide

OLS: Ordinary Least Square

PFC: Perfluorcarbon

PCSE: Panel Corrected Standard Errors

PPM: Parts per Million

TSCS: Time-Series Cross-Section

UNFCCC: United Nations Framework Convention on Climate Change

1. Introduction

That so many of us are here today is a recognition that the threat from climate change is serious, it is urgent, and it is growing. Our generation's response to this challenge will be judged by history, for if we fail to meet it - boldly, swiftly, and together - we risk consigning future generations to an irreversible catastrophe.

Barack Obama, September 22, 2009, UN climate change speech[1]

This characterization of the challenge of human-induced climate change by Barack Obama is a common apprehension today. Climate change has been recognized as one of the great global challenges of the 21th century. Many factors have contributed to this recognition: Some contributions have come from Hollywood, producing disaster movies such as *The Day After Tomorrow*. Al Gore's documentary movie *An Inconvenient Truth* gained much attention, especially after he, together with the United Nation's Intergovernmental Panel on Climate Change (IPCC) earned the Nobel Peace Price for their contributions. Perhaps most notably; severe weather disasters such as extreme heat waves, floods, drought - and perhaps the best known weather disaster from the last decade, the hurricane Katrina - have shown the world that human vulnerability to climate-related extreme events in the future is potentially very high, perhaps even higher than previously expected.

Human-induced climate change comes from the increase of greenhouse gases (GHGs) in the atmosphere. As the introductory quote by Obama indicates, there is a broad acceptance of human-induced climate change today. However, this was a controversial claim to make only a few decades ago. After a period of discussion among researchers in the 1970s and -80s, there was an increased scientific recognition of the fact that enhanced emissions of GHGs were causing global warming (Børsting and Ferman 1997: 53). However, the concept of the *natural* greenhouse gas effect has been known for over one hundred years. The effect of the GHGs in the atmosphere is that the average temperature of the Earth's surface is about 33 C warmer than it would be if the atmosphere obtained no heat-absorbing GHGs (Kanchberry and Victor 1995: 29). The greenhouse effect occurs because not all of the thermal radiation from the sun, which falls on our planet's surface, gets reflected back into space (Cowie 2007: 3). Growth in the greenhouse effect will result in increased global average air and ocean temperature, widespread melting of snow and ice, and rising global average sea level (IPCC 2007a: 5). These changes are the irreversible catastrophe which Obama

1 The entire speech can be found online (Obama 2009a).

referred to in the introductory quote. The challenge of human-induced climate change is met by the international community with an international regime: The International Climate Change Regime (ICCR). The research question of this thesis concerns the request that Obama made - is this ICCR a bold, swift and common approach to the challenge faced by climate change? As the first period of agreement is coming to an end, an evaluation of the success of the ICCR is important. The Kyoto Protocol is important for *future* commitment periods. I therefore pose the following research question for this thesis:

To what extent is the international climate change regime an effective regime?

The most obvious contribution of this thesis to the research field and the society in general is the information on the empirical case of the ICCR. Little research has been done on the effectiveness of the ICCR, because it is a rather new regime (the formal treaty was ratified in 1994). As the introduction suggests, climate change is one of the largest challenges in the 21st century, the ability of the international regime to meet these challenges is therefore crucial.

In addition to addressing the empirical case, there are two other contributions of this thesis. The thesis contributes to the research field by i) providing and ii) applying an analytical framework to investigate the level of regime effectiveness of the ICCR. First, I will argue how international regimes should be evaluated. Most of the theoretical framework is dedicated to this purpose. The research field of regime effectiveness can utilize a large theoretical and methodological repertoire, since regime effectiveness can be explored at different stages on the chain of causality. These different empirical stages of development are measured by different theoretical indicators of regime effectiveness. The analytical framework I provide is a review of the research field of international regimes, with an emphasis on regime effectiveness. Since the dependent variable of this thesis is regime effectiveness, I attach importance to the operationalization of this theoretical and empirical concept. Particular weight is put on the different indicators of effectiveness. Which theoretical indicator the researcher should choose, I will argue, is based on two considerations: the validity of the indicator of effectiveness must be as high as possible, and the data of this indicator needs to be available. I make use of a case, the ICCR, and show how this evaluation of effectiveness indicator should be conducted. My second contribution to the research field, therefore, is that I will employ my empirical case to the methodological approach that I have suggested. Many scholars have characterized the ICCR, with primarily the Kyoto Protocol in mind, as a failure (see Böringer and Vogt 2003, Harris 2007, Victor 2001). However, these studies are not employing the indicators of effectiveness (as this study does) on the regime effectiveness of the ICCR. Most of the research on regime effectiveness have a qualitative orientation - often because the international regime is observed at an early stage of the causality chain (for examples, see Barnwell 2011; Jakobsen 2007). This study takes a new approach, methodologically, by applying quantitative methods, and more precise, by regression analysis, to a field that has been characterized by qualitative methods.

When I employ my analytical framework to my case, the ICCR, I conclude that the ICCR is not an effective regime. The common emissions of the members of the regime have decreased, however, these changes, as we shall see, are not a consequence of the ICCR. Before the outline of the thesis is explained, a brief review of the background of the ICCR is necessary to clarify the object that will be analyzed in this thesis: the ICCR.

1.1 The Background of the ICCR

The breakthrough of climate change as a political issue came at the 1988 Toronto Conference, where the IPCC was established. In the beginning of the 1990s, there was a phase of interaction between scientific research,[2] strong public demand for action, and a stronger visualization of the damaging effect of global warming (Andersen and Agrawala 2002: 44). These events together resulted in negotiations about an international agreement on climate change.

The formal treaty-making process began in December 1990, and was ready for signature in 1992. 196 countries met at Rio De Janeiro, Brazil, and negotiated the *United Nations Framework Convention on Climate Change* in June 1992 (UNFCCC) (Grubb, Vrolijk, and Brack 1999: 37). The Convention says the following concerning the aims of UNFCCC: “The ultimate objective of this Convention, is stabilization of greenhouse gas concentrations in the atmosphere at a level that would prevent dangerous anthropogenic interference with the climate system” (UNFCCC 1992, Article 2). The Framework Convention entered into force in 1994, and has been ratified by nearly 190 countries (Dessler and Parson 2006: 13). This action marks the formal beginning of the ICCR, and is the foundation of ICCR. The Convention divided countries into three main groups according to their different commitments. First, the Annex I Parties include the industrialized countries that were members of the OECD in 1992, plus countries with economies in transition (the EiTs), such as Russia, the Baltic States, and several Central and Eastern European States (UNFCCC 2011a). The

2 The establishment of the IPCC has meant a great deal to the remedy the scientific uncertainty concerning human-induced climate change. Particularly important has been IPCC's assessments on human-induced climate change research to policy-makers (Børsting and Fermann 1997).

Annex I countries have “quantified limitation and reduction commitments” (Kyoto Protocol, Article 2). Second, the Annex II Parties of the Convention consists of the OECD members of the Annex I, but not the EiT parties (UNFCCC 2011a). The Annex II countries are identified to have an additional financial and technology transfer commitment (Kyoto Protocol, Article 11). Last, the non-Annex I Parties are mostly developing countries (UNFCCC 2011a). Both the time-frame and the targets remain unspecified in the Framework Convention, it was therefore necessary to develop a Protocol that was more definite. Thus, the Parties of the UNFCCC decided in 1995 to negotiate a protocol consisting of binding, quantified emission limitations for Annex I countries (UNFCCC 1995). The third Convention of the Parties (CoP-3) was held in Kyoto, Japan in December 1997. The Kyoto Protocol set the overall framework for intergovernmental efforts to mitigate the threat of human-induced climate change (Børsting and Fermann 1997: 80). Unlike the Framework Convention, the Kyoto Protocol specifies obligations for reductions of GHGs for the Parties: The parties agreed to make quantifiable emission limitations on the six GHGs specified in Annex A: Carbon Dioxide (CO_2), Methane (CH_4), Nitrous Oxide (N_2O), Hydrofluorocarbons (HFCs), Perfluorcarbons (PFCs) and Sulphur Hexafluoride (SF_6) (Kyoto Protocol 1997). The assigned amounts of GHG emissions should aggregately be reduced by “at least 5 per cent below 1990 levels in the commitment period 2008 to 2012” (Kyoto Protocol 1997). The year 1990 was chosen as the baseline year because it was the year in which the UN initially launched negotiations on climate change (Gunawansa 2010: 477).

There are two different strategies the parties of the Protocol can make use of to reach their emission targets; the parties can either cut emissions at home, or they can employ one of the three flexibility mechanisms. The flexibility mechanisms are included in the Protocol to guarantee cost effective emissions reductions: First, emission trading allows any two Parties to the Kyoto Protocol to exchange part of their emission commitment. Second, the Kyoto Protocol also enables emission savings or sink enhancement arising from cross-border investments between Annex I Parties (Grubb et al. 1999: 128). Third, by employing the clean development mechanism (CDM), the Annex I countries may meet their emission ceilings by carrying through certified projects that reduce emissions in developing countries (Hackl and Pruckner 2001: 208).

For the Kyoto Protocol to enter into force the ratifying countries had to account for 55 per cent of the total CO_2 emissions for 1990 by the Annex I countries. The Kyoto Protocol went into force on February 15, 2005, after the Russian Federation ratified the Protocol (Henry and Sundstrom 2007: 47). The proponents of the Kyoto Protocol celebrated it as a breakthrough in international climate policies because i) it is the first international environmental agreement to impose GHG emission limitations for industrialized member countries, and ii) it established an international mechanism for extending and deepening the cooperation on climate change (Böhringer and Finus 2005: 253; Böhringer and Rutherford 2009: 177).

There has been much debate concerning the weak aspects of the Kyoto Protocol. The most common critiques will be briefly reviewed here. One of the most common critiques of the Kyoto Protocol is the phenomenon referred to as *hot air*. It is expected that some members of the Kyoto Protocol (most notably the Eastern European economies) have been assigned emission limitations that are too mild. This makes it possible to sell large amounts of their abundant emission rights (Böhringer 2000: 779). Another criticism of the agreement is that it does not include all major emitters of GHG. First, the US is not a part of the agreement. The US withdrawal from the Kyoto Protocol has a strong influence especially on the emissions trading market because the US is the largest emitter of GHG (Springer 2003: 528). Second, the Protocol does not limit the emissions of developing countries. Developing countries have, during the negotiation process, fought hard for their right to develop, including the right to burn cheap fossil fuel, and converting forests into agricultural or industrial fields (Parks and Roberts 2010: 135). It is also argued that many developing countries lack the capacity to develop a regulatory program that could achieve a limitation in emissions (Gerrard 2007: 53) Major emitters such as China and India therefore have no emission limitations under the Kyoto Protocol. The Kyoto Protocol has also been characterized as having low ambitions. This is reflected in the goals of the Protocol. The latest estimates from the IPCC show that the concentration of GHGs in the atmosphere has already reached the maximum recommended limit,[3] therefore the IPCC urge that the discharge of GHGs must peak as soon as possible, no later than in 2015. Furthermore, the IPCC recommends that GHG emissions must be reduced by 50-85 per cent by 2050, with the year 2000 as baseline (IPCC 2007b). Hence, in relation to these recommendations, the Kyoto Protocol's aims, of at least 5 per cent decrease in emissions, are not particularly ambitious due to the limited time-frame. In addition, there are no penalties for countries that have ratified the Protocol, but fail to meet the reduction requirements. The Kyoto Protocol has therefore often been characterized as a “soft instrument”, which can be easily ignored by the Annex II countries (Gunawansa 2010: 480). Before I present the theoretical framework, I want to give the

3 The estimates of what is considered “dangerous” levels of GHG emission is uncertain. These estimation techniques is influenced by the time-lag in climate change adaption, modeling decisions, and definition of the concept of “danger” for the Earth's ecosystems and climate (Barnwell 2011). However, the current estimations by the IPCC have recommended the CO_2 equivalents in the atmosphere below between 445 and 490 particles per million (ppm) (IPCC 2007b).

reader a brief introduction to the main chapters of this thesis.

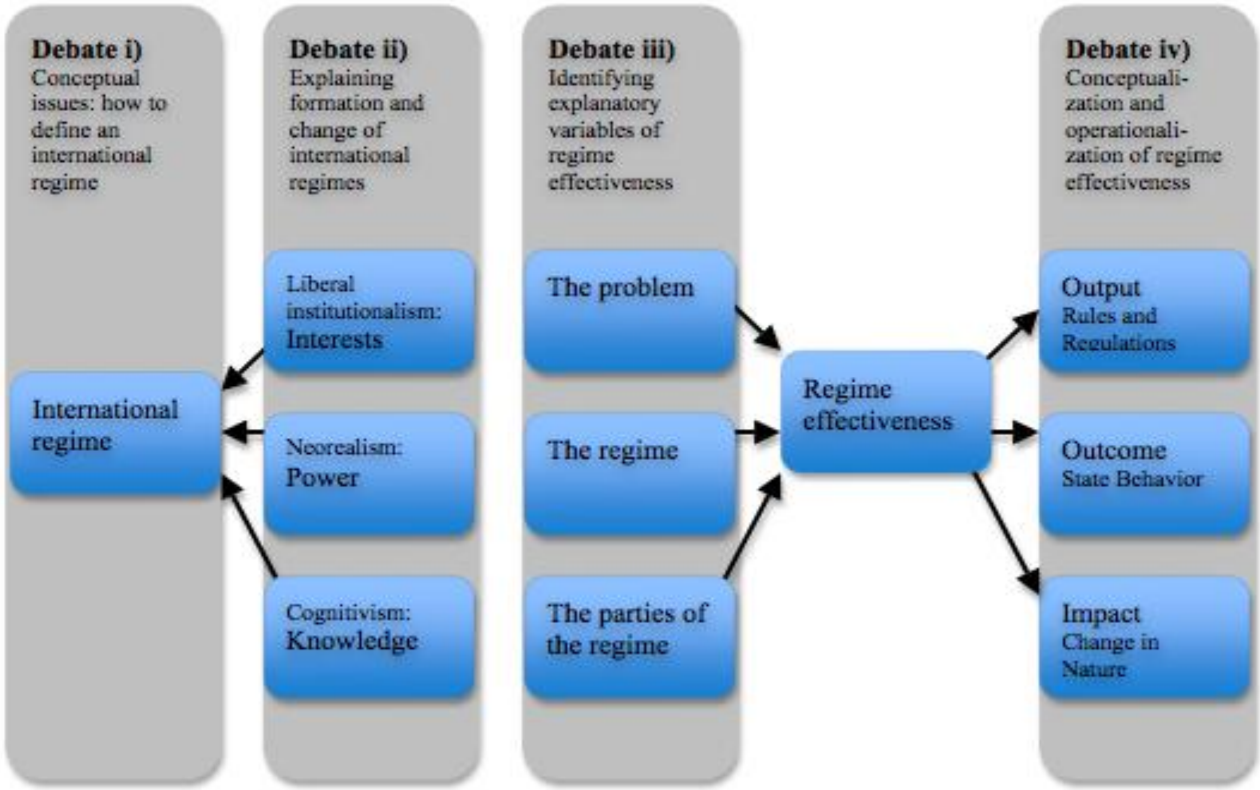
1.2 Outline of the Thesis

This introductory chapter is followed by five chapters. Chapter 2 presents the theoretical framework of the thesis, which illustrates the development of the research field of international regimes. In the research field of international regimes I have identified four different debates. The first debate is concerned with the operationalization of the concept of international regime. The second debate examined why international regimes are formed, and why they change. The third debate considers why some international regimes are more successful than others. The fourth debate, which is the main focus of this thesis, is the debate concerning the operationalization of the concept of regime effectiveness. Based on my systematic review of the research field, I propose a methodological approach to measure regime effectiveness in the case of the ICCR in the end of Chapter 2. Chapter 3 introduces the research design and methods of the thesis. Here I will discuss some general methodological considerations, and present the model for evaluating the effectiveness of the ICCR. The control variables of the analysis will be presented here. Chapter 4 presents the methods for collecting and analyzing data. First, I introduce my data set by evaluating it according to the data quality. Second, I present the data operationalization and summary to make the reader more familiar with the data. Third, the method for analyzing data is presented. Then, I introduce some potential problems that arise by using this method. Afterwards, I will present some solutions to these problems. In Chapter 5, I present and discuss the results of the analysis. This chapter is divided into three parts; first, a descriptive analysis, second, a regression analysis, third, a sensitivity analysis to investigate the robustness of the findings in the regression analysis. In chapter 6 the main conclusions will be presented. I will also suggest how future research can proceed for further development of the theoretical and empirical aspects of regime effectiveness.

2. Theoretical Framework

A research question that invokes the study of the effectiveness of the ICCR, requires a theoretical framework that sheds light on the development of the scholarly debate of regime effectiveness. Even though the research field of regime effectiveness is a rather new field within international politics, the literature has grown considerably during the past decades. After investigating the research field, I have identified four debates. First there were discussions on the concept of international regime (see Krasner 1982a). The research field then gravitated toward a second debate. The development of explanatory models of the establishment and change of international regimes, where the discussions between liberal institutionalism, neorealism and cognitivism has resulted in an ecumenical debate (see Hasenclever, Mayer, and Rittberger 2000). The third debate evolved into a debate of how to explain variations in effectiveness between different regimes (see Underdal 2002). More recently the regime literature has focused on a fourth debate, regarding the conceptualization and operationalization of regime-effectiveness (see Mitchell 2010). The relationship between these four debates is illustrated in figure.1.

Figure 1: The relationship between the four debates within the research field of international regimes



The four debates are probably familiar, as individual debates, to most scholars within the international regime field; however, the schematic interpretation of the literature of international regimes is likely to be new. The purpose of this schematic interpretation of the literature is twofold: first, to provide a clearly set out guide for this chapter on the theoretical framework. Second, to naturally lead the reader toward the main focus of the theoretical framework; namely the fourth debate of the conceptualization and operationalization of regime effectiveness. Note that I have switched the position of debate 3 and 4. Naturally, the debate of how to define regime effectiveness must occur before the debate concerning how to explain regime effectiveness. However, in this thesis I introduce the explanatory project first, to lead the reader to the main focus of the thesis, which is the conceptual debate. The theoretical framework results in an evaluation of which indicator is best fitted for the evaluation of the effectiveness of the ICCR. This evaluation illustrates which considerations the researcher has to make when deciding on which indicator of effectiveness she wants to employ. However, before I elaborate on the debate of the conceptualization of regime effectiveness, the other three debates have to be introduced, starting with the conceptual issues of defining an international regime.

2.1 Conceptual Issues: Defining International Regimes

The first debate identified within the research field is the task of defining an international regime. Naturally, the first debate was a debate on the conceptional issues, because "for theorist […] everything depends on the concepts learned" (Haas 1982: 209). Before one can do scientific research in the social sciences, one needs a definition on the object that will be investigated. The unclarity of the concept international regime was noticed by Susan Strange in 1982. She directly criticized the validity and usefulness of the concept "international regime". Strange used words as "imprecise" and "woolly" to describe this field of research (Strange 1982: 479). In the aftermath of Strange's criticism, on a conference to prepare the special issue of *International Organization*, 1982, a so-called "consensus definition" on the concept international regime was made, as a response to the question: What is an international regime? (Hasenclever et al. 2000a: 8) Stephen D. Krasner says that:

> Regimes can be defined as sets of implicit or explicit principles, norms, rules, and decision-making procedures around which actors' expectations converge in a given area of international relations. Principles are beliefs of fact, causation, and rectitude. Norms are standards of behavior defined in terms of rights and obligations. Rules are specific prescriptions or proscriptions for action. Decision-making procedures are prevailing practices for making and implementing collective choice (Krasner 1982a: 2).

Krasner's “consensus definition” was subsequently attacked by several scholars, but the definition is still one of the most employed definitions of an international regime. Two particular aspects of the consensus definition have been particularly criticized. The first aspect concerns the precise meaning of, and the relationship between the four components of a regime: What distinguish the “principles”, “norms”, “rules”, and “procedures” from one another? The second concern arises from the phrase “around which actors' expectations converge”. How can we know when a regime exists within an issue-area (Hasenclever, Mayer, and Rittberger 1996: 179)?

There seems, however, to be a broad agreement about two elements of Krasner's consensus definition. First, the consensus definition treats international regimes as *social institutions*. Second, the consensus definition characterizes regimes as *issue-specific* (Levy, Young, and Zürn 1994: xi). These two elements are important in several authors definition of international regime: “Regimes are social institutions governing the actions of those interested in specifiable activities (or meaningful sets of activities)” (Young 1980: 332). “International regimes are social institutions created to respond to the demand for governance relating to specific issues [...]” (Breitmeier, Young, and Zürn 2006: 3). The first debate of international regimes has not resulted in an absolute consensus on what an international regime is, despite numerous attempts from several scholars. However, it is an important debate because the definitions need to be in place before one can begin to explain something. When the conceptual issues of international regimes had been addressed, this led to a discussion regarding what can explain the formation and change of international regimes.

2.2 Explaining Formation and Change of International Regimes

The second debate of the research field of international regimes, which is a discussion of the establishment and change of international regimes, has been shaped by three schools of thought.[4] These three schools are the liberal institutionalists, the neorealists, and the cognitivists. Each of these three schools have declared and defended their view on why formation and change of international regimes happens. As will be argued in the following section, the research field has been going toward an ecumenical or synthetically discussion (Hasenclever et al. 2000b: 3). However, the driving force behind the research field of environmental regimes is the school of liberal institutionalism (Vogler 2003: 36). Liberal institutionalism founds its analysis on interest-based theories, which emphasize the role which international regimes play in helping states to

4 Schools of thought here refer to sets of ideas or theories. These theories share certain assumptions and emphases in making sense of international regimes (Hasenclever et al. 1997: 6).

realize common interests (Hasenclever et al. 1996: 183). Nevertheless, the research field of international regimes owes much to both the neorealist- and the cognitive school of thought as well. Neorealism, in particular, provides some crucial assumptions for the research field of international regimes. The liberal institutionalists and neorealists also share a commitment to rationalism,[5] which portrays states as self-interested and utility maximizing (Hasenclever et al. 2000a). Liberal institutionalism and neorealism therefore share much of the assumptions about the international system, but disagree on whether states are most interested in absolute or relative gains (O'Neill, Balsinger, and VanDever 2004: 153). It is argued that liberal institutionalism is "using realist assumptions to derive institutionalist conclusions" (Hasenclever et al. 1996: 184). Cognitivism, however, emphasizes actors' social and causal knowledge (Hasenclever et al. 2000b: 5). This school of thought has received much attention in the discussion of international environmental regimes. There have been, as Hasenclever et al. (2000a: 211) state, debates between the different schools of thought regarding the *conditions* for the establishment and change of international regimes. An understanding that the different schools can explain different aspects of the establishment and change of international regime has evolved, leading to the development of an ecumenical research field.[6] As Hasenclever et al. (2000b: 6) emphasize, a "synthesis or division of labor among the three schools of thought" has developed (Hasenclever et al. 2000b: 6). It is therefore true, as Krasner said in 1982: "The notion of rules of the game or international regimes – of principles, norms, rules, and decision-making procedures – that guide international behavior can be endorsed by the [three] schools of thought" (Krasner 1982b: viii). To further explain the establishment and change of international regimes, a more thorough review is needed. The neorealist argument will be presented first, the liberal institutionalist second, and the cognitive argument will be presented third. To illustrate the points of the different theoretical schools, and how they work synthetical, some examples from the ICCR will be made use of.

2.2.1 Neorealism: The Power-based Argument

Even though liberal institutionalism has guided the research field of international regimes, the research field has depended heavily on some crucial neorealist assumptions. It is therefore sensible to introduce the neorealist argument first. The neorealist assumptions, which are crucial to the

5 In the research field of international relations, rationalism portrays states as "self-interested, goal-seeking actors whose behavior can be accounted for in terms of the maximization of individual utility" (Hasenclever et al. 2000b: 7).

6 An example of the ecumenical research on regime formation and change is *Polar Politics* from 1993, where Gail Osherenko, Oran Young, and their colleagues document major examples of successful establishment of international regimes. In their introductory chapter they present different hypothesis' founded on the power-based, interest-based and knowledge-based argument (Osherenko and Young 1993a).

theoretical framework of international regimes, can be illustrated by looking toward neorealist Kenneth Waltz, who emphasizes the *structure* of the international system (see Waltz 1959 and 1979). First, he argues that in a world with sovereign states as the main actors in international politics, with no over-national authority, conflict is bound to occur (Waltz 1959: 159). The assumption that states are the main actor in the international community is also reflected in the legal framework of the ICCR:

> States have, in accordance with the Charter of the United Nations and the principles of international law, the sovereign right to exploit their own resources pursuant to their own environmental and developmental policies, and the responsibility to ensure that activities within their jurisdiction or control do not cause damage to the environment of other States or of areas beyond the limits of national jurisdiction" (Rio Declaration on Environment and Development 1992, article 2).

The second assumption that Waltz presents is that the international system is characterized by an international anarchy, which leads to a system that is characterized by self-help (Waltz 1959: 159). This self-help system lacks central authority and hence has a problem of collective action (Wendt 1992: 392). States will simply leave an international regime when they no longer perceive benefit from continued participation (Parks and Roberts 2010: 138).

The neorealist approach to international regimes regards regimes as at best "actors used by powerful states to transform their state interests into internationally agreed upon governance systems" (Breitmeier 1997: 91). This reflects the theory of hegemonic stability, which is a classical example of a neorealistic theory of international regimes. The theory of hegemonic stability originates from Charles Kindleberger's book from 1973 on the depression in the 1920s and 1930s. He argued that "the 1929 depression was so wide, so deep, and so long because the international economic system was rendered unstable by British inability and U.S. unwillingness to assume responsibility for stabilizing it [...]" (Kindleberger 1973: 289). Kindleberger here implies that international stability was dependent upon Britain and the U.S. (hegemonic powers). The theory of hegemonic stability claims that i) the presence of a dominant state leads to more stability in the international system, and ii) that this stability benefits all states in the system (Snidal 1985: 582). Thus, the theory of hegemonic stability assumes that the hegemon determines the international response to a case such as climate change (Rowlands 2001: 44). It has been argued that low levels of cooperation in the case of climate change will occur because of the lack of a hegemon, especially in the economic sphere. The accuracy of this argument will not be further discussed here.

2.2.2 Liberal Institutionalism: The Interest-based Argument

While the neorealist school of thought is skeptical toward the extent to which international cooperation is possible, the liberal institutionalist school is much more optimistic about the prospects for inter-state cooperation. They see international regimes as tools to help states realize common interests. Mancur Olson's research on public goods[7] is a point of departure for the interest-based argument. He tried figuring out why groups of rational actors that have a common interest or objective not always act to achieve that objective (Olson 1965: 2). The conclusion he came to was that these common interests or objectives may be public goods. This is much of the same conclusion that Garrett Hardin made in his famous theory of *the tragedy of the commons.* Hardin stated that resources held in common, such as air, rivers, the sea, or common grazing-land would be subject to massive degradation, due to an overload on the carrying capacity (Hardin 1968, Hardin 1998).

Liberal institutionalism makes use of economic theories to explain why failure to provide a collective good appears in the international system, and why an international regime may be the solution to this problem. An answer to these questions is provided by the theory of *market failure*. "Market failure refers to situations in which the outcomes of market-mediated interaction are suboptimal" (Keohane 1982: 335). When market failure appears, the problems are not attributed to the rational actors themselves (the sovereign states), but rather to the *structure* of the system (the international anarchy) and the existing institutions, or the lack of institutions (the international regime). These attributes of the system and the existing institutions impose transaction and information costs which are responsible for the barriers to effective cooperation among the actors. Liberal institutionalists argue that just like imperfect markets, the international system is characterized by institutional deficiencies that create barriers to effective cooperation (Keohane 1982: 335). The concept of negative externalities is particularly important to explain why an international regime may be the solution to the market failure. Externalities exists "whenever an acting unit does not bear all of the costs, or fail to reap all of the benefits, that results from its behavior" (Keohane 1982: 322). Negative externalities are common in the international system because there is no global authoritative institution. However, states may seek to reduce negative externalities by coordinating their behavior (ibid: 326). In this perspective, the function of a regime is therefore to coordinate the behavior of the states to achieve a particular outcome in a special field

7 A public good is recognized as "having benefits that cannot easily be confined to a single "buyer" (or set of buyers). Yet once they are provided, many can enjoy them for free" (Kaul, Grunberg, and Stern 1999a: xx).

(Krasner 1982: 7).[8] The constraints on international cooperation appear because of the characteristics of the international system according to Keohane (1984). Because an international anarchy exists, uncertainty follows. But international regimes may reduce this uncertainty, thereby facilitating cooperation between states. Thus, even though liberal institutionalism makes use of the same assumptions as neorealism, the former arrives at conclusions that are much more optimistic about the prospects for international cooperation, the realization of public goods, and for the possibility that international regimes may change state-behavior within a specific issue-area.

The view that international regimes are established to secure a collective good is not generally accepted (Hasenclever et al. 2000a: 95). It has also been argued that most public goods can only be defined in a specific group. Some goods that are private to some groups are public goods to other groups (Olson 1965: 14)[9]. However, both Hovi, Sprinz, and Underdal (2009: 30) and Sebenius[10] (1991: 120) state that to mitigate climate change is a global public good. Nevertheless, other authors, do not share this view. Climate change does alter the balance of nature (sea rise, loss of biodiversity, etc.), but this is a relatively slow process that will take decades, perhaps centuries to unfold. Even when applying a worse-case scenario of global climate change, it is not the equivalent of the Earth being hit by a mega-asteroid, which would have been devastating for all the states in the world immediately. This leads to a second point; different countries will be affected in different ways by climate change. Third, to mitigate climate change will be a costly affair, and will most likely require diversion of resources from other good causes. This is a problem that usually characterizes global public goods, but to a different extent. Finally, the success of reducing the worlds GHG emissions depend on the aggregate effort of all the states in the world. One country, or a group of countries, can not stabilize the level of GHG in the atmosphere alone. Free-riding – the inclination to withdraw from the costs of realizing the public good while enjoying the fruits of other states' efforts - is likely to be a problem in the mitigation of climate change (Barret 2007: 5). This is illustrated by the classic “prisoner dilemma”.[11] Some have argued that the Kyoto Protocol “only disguised the incentives to free ride; it did not correct them” (Barret 2007: 6). It has been argued that the US' refusal to ratify the agreement and the fact that the agreement only includes a few

8 This apprehension reflects Krasners consensus definition of international regimes as social institutions that are issue-specific.

9 The example employed by Olson (1965: 14) is a parade where the parade is a public good to all those living tall buildings around the parade era, while it is a private good for to those who cannot see it if they do not buy a ticket.

10 Sebenius (1991) argues that “the global atmosphere is a true commons in that any greenhouse gas emissions by a single country eventually mix and adversely affect the entire world (Sebenius 1991: 121).

11 The following situation is illustrative: All states have interest in reducing harmful environmental degradation, and all states are better off with international cooperation, but each of the states also has an incentive to free-ride and thereby to enjoy the benefits of abatement without have to bear the costs of abatement (Neumayer 2001: 124).

industrialized countries, are both characteristics of free riding (ibid).

2.2.3 Cognitivism: The Knowledge-based Argument

The school of cognitivism stresses ideas, knowledge and learning as explanatory variables of the establishment and change of international regimes. This school is skeptical to the rationalist theories (both neoliberalism and neorealism) of international politics. The big flaw of rationalist theories, cognitivists argue, is that they treat interests as exogenously given. A significant source of explanation of the establishment, and certainly the change of an international regime, is lost by treating the process of states' self-understanding as a black box (Hasenclever 2000a: 136). "Where functional theories see regimes as more or less efficient responses to fixed needs, cognitive theories see them as conditioned by ideology and consensual knowledge and evolving as actors learn" (Haggard and Simmons 1987: 499). In the cognitive perspective, the preferences of the actors are not singularly assumed, but rather treated as an empirical question. The particular answer may change over time due to the process of learning in connection with interaction between actors (Haas 1982: 236). The members of the international regime will gradually evolve a culture; This culture will create a social system that subsumes the traditional political world (Parks and Roberts 2010: 138). This can be seen in the negotiations during the Conference of the Parties (CoP), where the interests and identities of the climate change actors (individuals, sub-state actors, states, and international organizations) interact, and therefore bring forward a greater understanding of climate change policies (Pettenger 2007: 7).

One of the hallmarks of the scientific study of climate change has been high levels of uncertainty. This uncertainty has been concentrated not so much around the validity of the human-induced climate change, but rather focused on the questions of how much of climate change can be explained by human activities, how rapid and how far human-induced climate change may go, the distribution of vulnerability between regions and countries, and what measures of abatement and adoption are the most cost-effective (Vogler 2000: 134). Correct scientific information is widely acknowledged to be a key condition for environmental policy making. Three different components of information can be identified: i) the extent of the problem, ii) its causes, and iii) its consequences. Trustworthy information on these three components makes utility calculations possible, thus, it facilitates decisions regarding collective action (Dimitrov 2003: 124). Science has attracted much attention as an important force to promote international cooperation on human-induced climate change. Scientific expertise is without doubt important for other international regimes as well, but it has a particular central role in understanding the magnitude, causes, and solutions to environmental questions (Mitchell 2010: 119). The cognitive approach is often used to explain why the content of regime rules change and why they evolve, because cognitive theories can better explain dynamic elements of international regimes, than the other theoretical approaches (Haas 1982: 510).

As this elaboration shows, different aspects of regime formation and change can be explained by the different schools of thought. Power, interest, and knowledge based theories can all explain different aspects of the formation and change of international regimes. After this discussion the research field started to address regime effectiveness: How well do regimes function? What are the conditions that need to be satisfied for an regime to be successful? Are there certain factors that can determine the success and stability of an international regime? The next section will examine how one can explain variations in effectiveness between regimes.

2.3 Explaining Regime Effectiveness

The previous section elaborated on the establishment and change of international regimes (regime-formation), and how the different schools of thought explain these phenomena. The third debate that evolved within the research field of international regimes was a debate on how to explain variations in effectiveness between different regimes. Previous research has proposed a wide variance of independent variables for explaining regime effectiveness. One of the most explicit and rigorous efforts is referred to as the "Oslo-Potsdam solution" (Hovi, Sprinz, and Underdal 2003: 74). The Oslo-Potsdam solution suggests the most promising explanatory factors to be identified into three main clusters: i) the nature of the problem, ii) characteristics of the group of parties, and iii) properties of the regime itself (Underdal 2004: 40). First, some problems are harder to solve than others. Arild Underdal (2004: 21) makes the distinction between a benign and a malign problem. There are two dimensions the classification of problem severity can rest on, an intellectual dimension (scientific uncertainty) and a political dimension (deep-rooted conflicts) (Andersen, Walløe, and Rosendal 2005: 40). A benign problem is therefore characterized by scientific certainty, coordination, symmetry, and cross-cutting cleavages. A malign problem, on the other hand, is characterized by scientific uncertainty, incongruity, asymmetry and cumulative cleavages. The argument is as following: International regimes that address direct and visible environmental threats (a benign problem) can be solved relatively inexpensively are likely to be easier to solve (Mitchell 2010: 173).

Second, some groups of actors have greater capacity for collective action than other groups (Underdal 2004: 41). There has been some discussion within the research field regarding what this capacity constitutes. Elinor Ostrom (1995) argues that social capital influence the capability of collective action.[12] More specific aspects have also been suggested, such as the distribution of power between the parties of the regime. Some countries can use various types of resources to make other states comply with the regime (Aggarwal 1983: 620).[13] Others have emphasized aspects like instrumental leadership. Underdal (2002: 35) states that instrumental leadership may come from several factors such as officers of intergovernmental organizations, conference or working groups, national delegates, transnational organizations or informal networks.

Third, it is argued that effectiveness also depends on the problem-solving capacity of the regime. The argument of this explanatory variable of regime effectiveness is that some problems are solved more effectively because they are dealt with by "stronger" institutions. The main determinants of problem-solving capacity are i) the institutional setting and, ii) the skill and energy available for the political engineering of cooperative solutions (Underdal 2003: 23). The institutional setting refers to institutions as arenas. More specifically, institutions are regarded as "the framework within which politics takes place" (March and Olsen 1989: 18). The framework of the institution specifies the questions of *who*, *what*, *when* and *where*? Underdal (2003: 25) argues that decision rules and procedures are the most important determinant of the institutional setting. Some regimes contribute to regime effectiveness through providing procedures, arenas or other facilities that make it possible for the parties of the regime to develop a base of consensus, knowledge and shared beliefs (Underdal 2004: 41).[14]

This section has elaborated on the causes of variation of regime effectiveness between different regimes. This leads to the fourth debate, the main focus of this thesis, which is the conceptualization and operationalization of regime effectiveness.

---

12 The literature on social capital is an extensive literature. One of the best know definitions within social sciences is Putnam's definition: "Social capital here refers to features of social organisation, such as trust, norms and networks, that can improve the efficiency of society by facilitating coordinated actions (Putnam 1992: 167).

13 An example of this is the efforts of the U.S. to establish an ozone regime in the 1980s. It has been argued that one of the most important reasons (in addition to technological innovations) for the success of the Ozone Regime was the strive of the U.S. to end all use of CFCs (Sprinz and Vaahtoranta 1994: 94; Wettestad 2003: 160).

14 For instance, Wettestad (2003: 165) classifies the institutional capacity of the Ozone Regime as intermediate due to i) the restrictions on trade with nonparties and ii) the establishment of the Multilateral Fund.

2.4 Conceptualization and Operationalization of Regime Effectiveness

The fourth debate of international regimes is a discussion of conceptualization and operationalization of the concept of regime effectiveness. This debate is originally the third debate that originates within the research field of regime effectiveness. A concept such as regime effectiveness naturally has to be operationalized before it can be explained. However, this thesis' main focus is to investigate whether the ICCR is an effective regime or not. Therefore, it follows naturally that the conceptualization and operationalization of regime effectiveness is the last debate to be reviewed. The following paragraphs will give a brief introduction to what regime effectiveness is, and how it can be evaluated.

In its most basic form, the research field of regime effectiveness assesses regimes in how well they function. The notion of effectiveness therefore implies the idea of international regimes as tools to perform a particular task. Like other political tools, regime effectiveness can therefore be evaluated in terms of their usefulness.[15] This instrumental perspective on regime effectiveness only pays attention to a subset of consequences, notably those consequences that are meaningful to the purpose assigned to the regime. Other consequences, such as side effects, are only of importance if they have direct or indirect influence on this task or purpose. Therefore, when assessing the regime effectiveness, the costs attached to the operation of the regime is generally taken out of the equation. Regime effectiveness is therefore an evaluation of the gross achievements, not net achievements, and should not be confused with regime efficiency,[16] which directs the attention towards the net-effects of a particular regime (Underdal 2004: 27). To evaluate an international regime's effectiveness, one must first ask: what should be evaluated, and where should we look for regime effectiveness? It is crucial to use an indicator which is a reasonably valid measure of regime effectiveness, and at the same time practical in terms of gathering of data. As will become evident, these two requirements are not always easy to harmonize. There are three potential indicators that have been employed by the public policy trichotomy: outputs, outcomes, and impacts (Mitchell 2010: 148). This division reflects the different stages of development of the international regime: most regimes produce a chain of consequences, and regime effectiveness can be measured at different points in this chain of causality (Underdal 2004: 34). The three different indicators of effectiveness therefore represent three different operationalizations of the dependent variable, regime effectiveness. Output refers to the regime formation, outcome refers to the regime

15 The fourth debate of operationalization and conceptualization of regime effectiveness is largely influenced by the literature of public policy evaluation (e.g. Mohr 1995).

16 The research field of regime efficiency identifies two techniques of evaluation of efficiency; cost-benefit analysis and cost-effectiveness analysis, where cost-effectiveness is the most employed method of analysis (Day 2002: 181).

implementation, and impact refers to the state of the environmental quality (Underdal 2002: 7). The chain of regime effectiveness is illustrated in the following figure:

Figure 2: The chain of regime effectiveness indicators

The key distinctions between these three indicators of regime effectiveness are elaborated below, due to its importance for the subsequent empirical analysis.

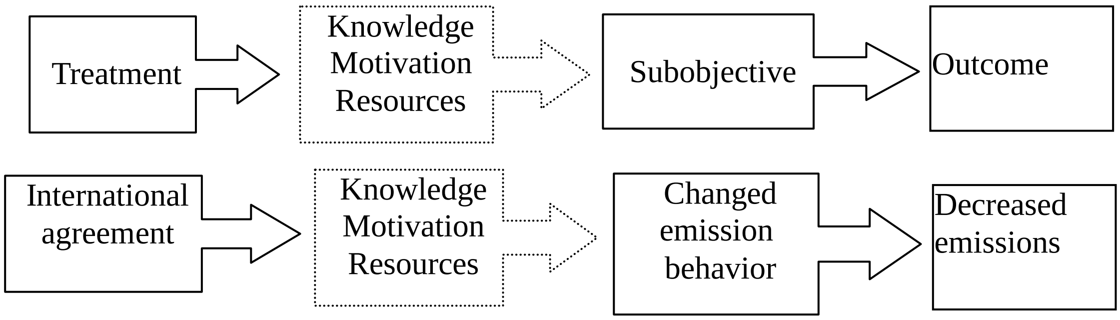
2.4.1 Output: The Rules and Regulations in the Body of Policies

Output refers to the new set of rules and regulations applied by the international regime. Output is measured by the laws, policies, and regulations that are adopted at the state level as a response to the requirements of the international regime (Mitchell 2010: 148). The study of output is typically concerned with the matter of the formation stage of an international regime (Young 2004: 12). The focus is often on “measures that states take to make international accords effective in their domestic law” (Jacobson and Brown Weiss 2000). An analysis of the output indicator of regime effectiveness is often a case-study of the policies performed at the national level designed to adjust to the new rules of the international regime. Earlier studies of regimes have used the output indicator by evaluating criterion such as the stringency of its rules and regulations, the extent to which the system of activities targeted is in fact brought under its jurisdiction or domain, and the level of collaboration established (Underdal 2002: 6). However, as Mitchell (2010: 148) states, even though output is easily available information, such as in legislative laws, new laws and policies, it is often incomplete as a measure of regime effectiveness. Laws and policies are not evidence of meaningful behavioral change, even though new laws and policies are necessary.

2.4.2 Outcome: The Member's Practical Adherence to Policies and Regulations

Most scholars in the field of regime effectiveness look for effectiveness in outcomes (the implementation of the regime). This signifies an evaluation of governmental or sub-state actor behavior: Through the use of the outcome indicator it is possible to “capture the vast variety of events and conditions in the environment as they pertain to the persistence of a political system” (Easton 1965: 26). Changing behavior is valuable as an indicator of effectiveness since international environmental regimes almost always identify behavioral changes that must occur to achieve the agreement goals. Whatever the ultimate purpose for the formation of the regime is, human behavior is always the immediate target of regulation (Underdal 2004: 34). However, it is important to notice that often changed human behavior is a subobjective to a desired outcome. This means that another outcome must be achieved before and in order for some further outcome can be achieved as a result of the rules and regulations of the international regime. Figure 3 illustrates how the international regime is supposed to cause changes in emissions behavior, which is the subobjective, to further lead to decreased emissions, which is the outcome.

Figure 3: The connection between the treatment, behavior prerequisites, subobjective and the outcome



In addition, there are some prerequisites necessary to channel change in behavior: First, the units must know what is to be done. Second, the units must be motivated or have the incentive to do it. Third, the units must have the ability and other resources necessary. These are referred to as knowledge, motivation, and resources (Mohr 1995: 31). When altered human behavior is the main objective, the behavior prerequisites are pertinent. The relationship between the behavioral prerequisites and changed human behavior is illustrated in figure 3.

Ronald B. Mitchell (see Mitchell 2004: 2010) has made considerable contributions to the research of the outcome indicator of regime effectiveness. He argues that one has to identify observable behavior for each country for each year of membership. By making use of country/years as the main unit of analysis, the analysis will often generate a sufficiently large number of observations to support the use of various statistical analysis (Young 2003: 101). Mitchell (2004: 125) argues that the best way to choose a dependent variable for an analysis of the effectiveness of an international regime is to employ a dependent variable that corresponds to the goals identified in the agreements that form the legal basis for the regime. Using behavior has three advantages: i) to change behavior is necessary to increase environmental quality, ii) behavior is easier to model accurately than environmental quality indicators, and iii) behavioral data is more easily available and of better quality than environmental quality data. These points are presented not to argue that modeling emission behavior of countries is easy, but to argue that modeling behavior is easier than modeling environmental quality (ibid: 126). However, as we will see in the following section, there is without doubt, some indisputable advantages to employing the impact indicator.

2.4.3 Impact: The Tangible Results of Altered Behavior in Terms of Changes in Environmental Quality

Impact refers to changes in environmental quality. As argued in the previous section; it is difficult to measure a regime's direct effect on environmental quality. In addition, the process of improvement of the environmental quality can be a slow process (Mitchell 2010: 148). Nature can often take years, sometimes even decades or centuries, to respond to changes in human behavior (Underdal 2004: 35; Helm and Sprinz 2000: 632). The goals of environmental quality often change as scientific understanding improves (Mitchell 2003: 445). This makes it unfit as an indicator of regime effectiveness as most indicators of environmental quality have not changed yet. Underdal (2004: 34) argues that assessing impact will often require substantial expertise in other disciplines such as the natural sciences, and impact can therefore often be too complex to analyze.

This section has elaborated on conceptual issues on regime effectiveness, and how regime effectiveness can be evaluated. The main focus has been on the three indicators of regime effectiveness. These indicators are different operationalizations of the dependent variable, regime effectiveness. Which of the dependent variables is chosen depends on what developmental stage the regime is. The next section examines how the effectiveness of the ICCR should be evaluated, by discussing which indicator of effectiveness is best fitted for the ICCR.

2.5 Bridging the Gap between the Theoretical and the Empirical Universe: Regime-effectiveness as Concept and Phenomenon

The review of the three indicators of effectiveness indicates that the theoretical concepts appear along a chain of causality. Any attempt to measure regime effectiveness will have to refer to the state of affairs at one particular point in time. This chapter emphasizes the importance of identifying the chain of consequences which an international regime produces when evaluating the effectiveness of a regime. Hence, regime effectiveness can be measured at different stages of causality (or development). In 2003, Underdal stated that "at several junctures we have seen that progress in dealing with methodological challenges will require progress in developing more precise conceptual tools and a better understanding of complex causal pathways" (Underdal 2003: 44). One way of interpreting this statement is that the research field of international regimes needs bridge-building between the theory (conceptual tools), the empirical universe (causal pathways), and the methodology. I therefore argue that the indicator of effectiveness must reflect the connection between the regime effectiveness theory and the empirical universe of the regime under evaluation, and that these two should be connected through methodology (by emphasizing validity and availability of data). Underdal himself has discussed some of these links between theory, the empirically universe and methodology indirectly in his work on regime effectiveness (see Underdal 2002; Underdal 2003).[17] However, Underdal has characterized these links as methodological approaches. No schematic illustration of the connection between the regime effectiveness theory, the empirical universe, and the methodology has been composed explicitly within the research field before. Below, an attempt is made to build a bridge between the world of concepts and that of empirical phenomena, by linking the operationalized, tricotomic regime effectiveness scale with the historic development of the ICCR.

Figure 4 is based on a time-line which illustrates the empirical development of the ICCR. Underneath the time-line, the empirical observations are divided into six phases. Above the time-line the theoretical approaches of international regimes are presented (introduced earlier in this chapter). The debates are not debates that are separated, they may very well overlap at several occasions. The intent behind this time-line is to make the reader aware of how each phase of the development of the ICCR can be explained by the theoretical framework presented above. It is therefore of great importance for the scholars of regime effectiveness to be aware of this time-line to be able to detect the chain of consequences that an international regime produces. This argument is illustrated further in a stepwise review of the three main indicators of regime effectiveness of the ICCR: output, outcome and impact.

---

17 Underdal addresses the questions of what precisely is the *object* to be evaluated, against which *standard* is this object to be evaluated, and how one can *compare* the object to the standard (Underdal 2002 and 2003).

Figure 4: Illustration of the relationship between regime effectiveness theory and the empirical universe of the ICCR - With an emphasis on the Kyoto Protocol as the main objective of evaluation.

<table>
<tr><td colspan="3" rowspan="3">Debate 1: The definitional content of the international regime concept.

Debate 2: Explaining establishment and change of international regimes
i) Interests
ii) Power
iii) Knowledge</td><td colspan="3">Debate 3: Explaining differences in regime effectiveness
i) The problem
ii) The properties of the regime itself
iii) Characteristics of the groups of parties</td></tr>
<tr><td colspan="3">Debate 4: Conceptualization and operationalization of regime effectiveness</td></tr>
<tr><td>i) Output: Rules and Regulations</td><td>ii) Outcome: Emission Behavior</td><td>iii) Impact: Change in Nature</td></tr>
<tr><td>1970-1987</td><td>1988-1991</td><td>1992-1996</td><td>1997-2007</td><td>2008-2012</td><td>The future</td></tr>
<tr><td>Scientific recognition and consensus-building of human-induced climate change</td><td>Issue ascending to the political agenda: Preparing the Framework Convention</td><td>The Framework Convention is signed and ratified (UNFCCC)</td><td>A binding quantified agreement: The Kyoto Protocol is signed</td><td>Implementation period of the Kyoto Protocol and renegotiations</td><td>A new period of agreement</td></tr>
</table>

2.5.1 from One Extreme... Why Impact is the Ideal Indicator of ICCR Effectiveness, and why it is a Premature Measure

The last stage of development of the ICCR (starting in 2013) naturally lies in the future, and therefore has no empirical observations. The regime effectiveness of the last phase may be measured empirically by change in nature, with an indicator such as concentrations of GHGs in the atmosphere and global mean temperature. International environmental regimes are established to protect the environment from environmental degradation - making impact in nature the ultimate interest and the most valid measure of regime effectiveness (Underdal 2002: 6). However, the impact of an environmental regime on nature is complicated by several reasons. One major reason is due to the absence of data, which is a result of two factors: i) technical and scientific problems of measuring environmental indicators, and ii) time lag between implementation and consequences. The time-lag is particularly severe for problems of pollution, because the impact of pollutive activities is felt only after a long time (Helm and Sprinz 2000: 632). The latter is a serious limitation in the case of human induced climate change due to the long time for decomposition of CO_2 in the atmosphere, which is about 120 years, and even longer for more complex GHGs (Rennenberg, Wassmann, Papen, and Seiler 1995: 61). The time-lag is troublesome because policy makers and scholars usually want to assess the effectiveness of an international regime before impact on the environmental problem can be determined (Underdal 2004: 35). Hence, the effectiveness indicator of impact is not an option (even though it is the most valid measure of regime effectiveness) for an analysis of regime effectiveness of the ICCR because of the lack of data.

2.5.2 ...to Another. Why Output Alone is an Unsatisfactory Indicator of ICCR Effectiveness

Only when the institutional framework of the international regime is established, is it relevant to consider the analysis of regime effectiveness. The institutional framework is established during the first phases of the ICCR – and particularly so in the two phases stretching from 1992-1996 and 1997-2005. In this phase of the ICCR a quantified agreement is signed (the 1997 Kyoto Protocol), and an analysis of regime effectiveness based on the output indicator is therefore possible. An analysis of regime effectiveness with output as indicator has to evaluate the regime at two different levels of analysis: i) the norms, principles, and rules of the international regime, and ii) the member states' *measures* to comply with the international regime. The appropriate method for investigating the output of an international regime is therefore the case study, a case study of the international agreement and a case study of the measures of the member states' (Helm and Sprinz 2000: 632).[18] The data needed to conduct an output analysis of the international level is easily available in the texts of the ICCR (most importantly the UNFCCC and the Kyoto Protocol). While at the national level, the first natural step is to examine the adoption of national laws and reduction programs (Wettestad 1999: 10). The data is therefore readily available at both the international and the national level of analysis.

However, even though changes in political output is necessary to change the behavior of the member-states, it does not guarantee the desired effects because rules may be ineffective or simply be neglected (Helm and Sprinz 2000: 633). This argumentation calls attention to the fact that the output indicator is not as valid as the two other measures of regime effectiveness. That is why most scholars recommend an analysis of regime effectiveness to be based at outcome, and not on output (see Mitchell 2004; Underdal 2004, etc.).

2.5.3 Why Outcome is the Optimal Indicator of ICCR Effectiveness, Balancing Validity and Availability of Data

I argue that the outcome indicator is the best indicator of regime effectiveness of the ICCR. This argumentation consists of two factors. First, outcome is a more valid measure of regime effectiveness than output. Second, the data of the outcome indicator is available (although not for the whole period of commitment), as opposed to the impact indicator of regime effectiveness. The

18 For more information on how to measure the output indicator see Andersen, Skjærseth, and Wettestad (1995) and Wettestad (1999).

reasoning for this argument should be clear after the following section.

An analysis of the outcome indicator necessarily has to be empirically based on emission behavior of states. The indicator of outcome must be closely related to the main goals of the international regime. Recall that according to the Kyoto Protocol, the main objective of the member states of the regime is to achieve "quantified emission limitations and reduction commitments [...]" (The Kyoto Protocol 1997, Article 2). Changed behavior in the case of the ICCR shall therefore result in reduced GHG emissions from the member-states in accordance with their differentiated commitments as identified by the Kyoto Protocol. Because the main goal of the Kyoto Protocol is to limit GHG emissions, it is natural to employ GHGs as the dependent variable in an analysis of the effectiveness of the outcome indicator. If one apply 5 per cent reductions of GHG emissions as the threshold to call the ICCR an effective regime, this would be in accordance with the line that the compliance literature suggests.[19]

An analysis of GHG emissions naturally calls for quantitative methods, and more specifically, regression analysis.[20] To employ quantitative methods one has to ask the question: Is the data both reliable and available? Usually, extensive country-year data on pollutants exist because the international environmental regimes require annual reports (Mitchell 2004: 139), this is also the case of the ICCR (see the Kyoto Protocol 1997, Article 3). This makes it easier to access reliable data. Because 2012 is the end of the period of agreement, an analysis of the effectiveness with the indicator of outcome might seem somewhat premature.[21] It has been argued that the "actual change in behavior and environmental impact can be determined only at a late stage (usually after several years of cooperation)" (Underdal 2002: 6). One of the key questions of this analysis, due to availability of data, is therefore: Is the development of the ICCR at the stage that allows to employ input as an indicator of regime effectiveness, given that approximately two years remain of the implementation-period of the Kyoto Protocol? First, the Kyoto Protocol states that "[e]ach party included in Annex I shall, by 2005, have made demonstrable progress in achieving its commitments under this Protocol" (Kyoto Protocol 1997, Article 3). The Kyoto Protocol does not specify any criteria or a process for how this "demonstrable progress" may be fulfilled. However, failure to take early action is one of the largest threats of noncompliance with the Kyoto Protocol. This is mainly

19 For an introduction to the compliance literature, see Simmons (1998).
20 Mitchell has, on several occasions, raised a concern about "the almost-exclusively use of case studies" within the research field (Mitchell 2007: 512).
21 Ringquist and Kostadinova (2005: 86) state that premature analyses is one of the largest obstacles to the research field of regime effectiveness.

due to economics; it makes sense neither to reduce GHG emissions over night to meet the Kyoto obligation, nor to delay reductions until the beginning of the period of agreement (2008) (Michaelowa and Rofle 2001: 284). In addition, it is argued by several scholars that early action is required for the member states to be able to meet their commitments (see Kennedy 2002; Pan and Regemorter 2004). Lack of early action to decrease emissions is likely to increase the reliance of the flexibility mechanisms of the Protocol. However, there are transaction costs attached to these mechanisms (such as search and information costs, negotiation costs, approval costs, monitoring and enforcement costs, e.g.). Nevertheless, these transactions costs may be suppressed by early action (Pan and Regemorter 2004: 1477). It is therefore required that the member-states of the ICCR take early action to be able to meet their requirements.

In table 1 I briefly summarize the argumentation of validity and data availability. Notice that the outcome indicator is the only indicator that I don't evaluate to have low on either of the two criteria.

Table 1: Evaluation of the three indicators of regime effectiveness

	Validity	**Availability of data**
Output	*Low*: No guarantee that laws, policies, regulations will change behavior	*High*: Due to a variety of reports and legal documents
Outcome	*High*: Changed emission behavior is necessary to decrease emissions	*Medium*: The end of the period of agreement is not until 2012
Impact	*Very high*: Impact is the ultimate interest	*Low:* Due to i) Technical and scientific problems, and ii) Time-lag

To sum up, this chapter has served two purposes: It has guided the reader through the research field of international regimes, by assessing the four different debates. Further, the chapter has evaluated how effectiveness of international regimes should be evaluated, and how the effectiveness of the ICCR is best evaluated. As a result of the compromise between validity and data availability, the outcome indicator was chosen to measure regime effectiveness of the ICCR. The next section will elaborate on the research design of this thesis, the main focus will be how to model the ICCR with the outcome indicator of regime effectiveness.

3. Research Design

In the previous chapter the theoretical framework of this thesis was presented. The four debates within the research field of international regimes was introduced, and the relationship between these four debates was explored. However, the main focus of the chapter was to explore the operationalization and conceptualization of regime effectiveness, with emphasis on the three indicators of effectiveness. The chapter resulted in the choice to employ the outcome indicator as the dependent variable of the analysis in this thesis. These discussions and considerations are an important part of the methodology of this thesis, the research design and the methods of this thesis will therefore follow as a natural consequence. The aim of methodology is “to describe and analyze the methods, throwing light on their limitations and resources, clarifying their presuppositions and consequences, relating their potentialities to the twilight zone at the frontiers of knowledge” (Kaplan 2004: 23). Methodology shall help us to understand not just the products of scientific inquiry, but the process itself. Methods, on the other hand, refer to specific procedures for gathering and analyzing (or manipulating) data (Brannen 2004: 312).[22] Data in social science generally involves either qualitative or quantitative formats, or both. Qualitative methods for collection of data include interviews, observation, and document analysis. Quantitative methods for collection of data largely refer to i) collecting your own data (primary data), or ii) use data collected by someone else (secondary data). These two paradigms have an ongoing discussion.[23] However, it is customary to acknowledge that strengths and weaknesses are to be found in both the qualitative and quantitative approach (see King, Keohane, and Verba 1994; Brady and Collier 2004).

A research question which concerns the level of regime effectiveness of the ICCR demands a complementary research design. The research question is merely a diagnostic research question of a case-study: is the ICCR effective, or not? To answer this research question however, a concept-clarification of regime effectiveness was needed to explore the dependent variable of this thesis: regime effectiveness. After the concept-clarification, the outcome indicator of regime effectiveness was chosen, making GHG emissions the dependent variable of a regression analysis. In addition, this project is explanatory because I investigate whether the ICCR is the cause of the changed emission behavior. Even though this project has a design that is complementary, the main method of analysis is the regression analysis. The main focus of Chapter 4 on methods is therefore the regression analysis. As will be clear after the following sections, the regression analysis of this

22 The concept of “data” in scientific research refers to systematically collected elements of information about the world (King et al. 1994).

23 This discussion has at times been characterized as a “paradigm war” (see Tashakkori and Teddlie 1998: 3).

thesis tries to imitate a quasi-experiment. A quasi-experiment is an experiment which lacks random assignment of units to conditions but that otherwise have similar purposes and structural attributes to randomized experiments (Shadish 2002: 104). The results from quasi-experiments need to be interpreted with caution because the design includes a nonequivalent control group. In designs with nonequivalent groups, the researcher does not control for the assignment of the individuals to groups because the groups already exist (Morris 2007: 365). The groups of membership are already decided in this thesis. The analysis of this quasi-experiment of the ICCR includes a control group, control variables and a pretest. The purpose of these three design elements will be explained in the following.

3.1 Modeling the ICCR

As the previous chapter emphasized, a range of methods are available in the field of regime effectiveness. The choice of method in this thesis depends on the indicator of effectiveness that is chosen. The outcome indicator was chosen after an evaluation of validity and data availability. As a consequence of the choice of indicator of effectiveness, quantitative methods, and more specific, regression analysis was presented as the method of analysis. The dependent variable was identified as GHG emissions. The threshold of successful emission behavior according to the Kyoto Protocol is an aggregated 5 per cent reduction in relation to 1990 level. However, because this thesis make use of a quantitative approach, it is possible to investigate not just if the member countries meet their emission limitations, but also whether there are other causes to the change in emission behavior, other than the regime itself. This thesis is therefore concerned with *regime effects*, which are changes in the dependent variable that can be attributed to the regime (Helm and Sprinz 2000: 636).

The ICCR regulates the emissions of six GHGs. In this thesis CO_2 will be employed as the dependent variable. CO_2 is recognized as the most important among the six GHGs regulated by the Kyoto Protocol (IPCC 2007a). The next step of the analysis is to identify the equation of the regression analysis. For the time being, the following equation might be proposed to model emission behavior:

$$EMISSIONS_{i,t} = \beta_0 + \beta_1 * MEMBER\ OF\ REGIME_{i,t} + \beta_2 * AFTER\ KYOTO_{i,t} + \beta_3 * MEMBER\ OF\ REGIME_{i,t} * AFTER\ KYOTO_{i,t} + \varepsilon_{i,t} \quad [1]$$

Where EMISSIONS is annual emissions of CO_2 for country *i* in year *t*. β_0 is the Y-intercept.24 MEMBER OF REGIME is a dummy variable, that is, a variable which is coded as 0 for those countries that are not members of the regime, and 1 for those countries that are members of the regime. Those states that are coded 1 are the treatment group, and those countries which are coded 0 are the control group (Mohr 1995: 110). The variable AFTER KYOTO is also a dummy variable, coded 0 in the years before the Kyoto Protocol was ratified, and 1 for the years after the Kyoto Protocol was ratified. β_2 is the posttest, where the years before the Protocol was ratified is the pretest, and the years after is the posttest. The goal of a pre-posttest design is to evaluate the influence of the intervening treatment by comparing observations before the treatment with observations after the treatment (Gravetter and Forzano 2008: 296). The interaction term MEMBER OF REGIME*AFTER KYOTO gives us the effect of the regime for member-states after the Protocol was ratified. β_3 is therefore the effect on the treatment group after the treatment was conducted - the information whose discovery is the purpose of the evaluation. Last, ε is the disturbance term.[25]

This model can tell us whether the emission behavior changes when a country ratifies the Kyoto Protocol, while holding other variables constant. The effect of the interaction term is predicted to be negative, with the assumption that membership in the ICCR actually leads to desired changes in emission behavior in terms of reduction or slowing the growth of GHG emissions. In addition, the t-statistics makes is possible to evaluate whether the findings are due to statistically significant patterns, or due to coincidence (Mitchell 2004: 126-127). Therefore, in this quasi-experimental design, the units (countries) are assigned to the treatment or control group, and each unit is measured both before and after the treatment has been assigned (Morris 2007: 356). However, even though there are changes in the treatment group's emissions in the posttest, it is not evident that these changes are due to membership in the international regime. There may be other factors influencing the the treatment group other than the treatment itself (Mohr 1995: 112). Any behavior which can be influenced by an international regime may just as well be influenced by other factors. Indeed, research shows that reduction of GHGs may happen over long periods of time even without the existence of an international climate change regime and commitments to reduce CO_2 emissions (Fermann 1995). Therefore, control variables must be included in the model.

---

24 Mohr describes the Y-intercept as “the height at which the regression line crosses the vertical axis, or the value of Y when X=0” (Mohr 1995: 109).

25 It has been argued that quasi-control group is a better expression for the control group in statistical analysis. The term quasi is added because the researcher sorts subjects into naturally occurring groups based on the subjects' values on the independent variable (member of regime) (Buttolph 2008: 155).

3.2 Control Variables: Other Causes of Changes in CO_2 emissions

Several control variables have been proposed. However, economic variables are often considered the most important ones. A large literature argues that increased GDP per capita influences environmental degradation. The argument is simple: the more economic activity, ceteris paribus, the higher the level of pollution is likely to be, due to increased resource use and waste generation (Panayotou 1997: 469; Stern 2002: 203).[26] However, a much-discussed theory exists which claims that increased GDP does not necessarily lead to increased CO_2 emissions. The theory of the Kuznets Curve was developed by Simon Kuznets (1955).[27] In the 1990s the Kuznets Curve was applied to the relationship between environmental degradation and per capita income. The inverted-U curve between economic growth and environmental degradation is known as the Environmental Kuznets Curve (EKC): Environmental degradation will rise with increased GDP per capita only up to a certain point, and then environmental degradation will decrease with GDP per capita (Dinda 2004: 433). The most common argument for the EKC is income elasticity of environmental quality demand. As per capita income grows, people achieve a higher standard of living and start to care more for the quality of the environment they live in (Nordstrom and Vaughan 1999: 29). Although the empirical existence of the EKC is still debated among scholars, most quantitative models of environmental degradation account for GDP per capita and its squared term. If the EKC exists, the coefficient of GDP per capita will be positive, and the coefficient of GDP per capita squared will be negative (Li and Reuveny 2006: 942).

One may also expect the composition of the economic activity to influence GHG emissions because some sectors are more resource-intensive than others. There are three different sectors: the primary sector (agriculture, fisheries, and mining), the secondary sector (industry), and the tertiary sector (services), where it is argued that industry is the most polluting sector, and services the least polluting. This argument is employed by several scholars (see Nordstrom and Vaughan 1999: 29; Panayotou 1997: 470; Stern 2002: 204; Stern, Young, and Druckman 1992: 81). In the negotiations of the quantified emissions target in the Kyoto Protocol, a group of countries (Australia, New

26 Even though the argument is simple, it is desirable to investigate the argument further: What is put into production and consumption has to come out- either as goods or services, or as waste materials. The conservation of input to useful output can never be 100 per cent, and it is therefore fair to say that increased economic activity inevitably stresses the environment because of the increased waste materials (Stern et al 1992: 80).

27 Kuznets' theory was originally applied to income equality. He predicted that the changing relationship between per capita income and income inequality would look like an inverted-U shape. Therefore, as per capita income rises, income inequality would rise at first, and then start to decline after a turning point (Kuznets 1955).

Zealand, Russia, Ukraine, Norway and Iceland) were allowed to increase their GHG emissions due to what the Kyoto Protocol identifies as "their common, but differentiated responsibilities and their specific national and regional development priorities and circumstances" (The Kyoto Protocol 1997, Article 10). These six countries are faced with special circumstances within the energy and industry sector, and have therefore been granted larger emission quotas than the other countries of the agreement. Countries with energy-efficient economies tend to suffer relatively higher marginal costs on abatement costs (Fermann 1997: 182). Hence, to abate to climate change is a more cost-effective to an economy based at non-renewable energy (such as coal or natural gas) than an economy based on renewable energy sources (such as hydroelectric- or wind power). Examples of this are countries such as New Zealand and Norway, which have economies based on renewable hydroelectric power (Grubb et al. 1999: 33). Norway's hydroelectric power represents about 99 per cent of the country's electricity production, making emission reductions more difficult than the historical coal-based EU economies (Christiansen 2002: 236).

Population growth has, since Thomas Malthus' *An Essay on the Principle of Population* (1798), been debated as a driving force behind environmental degradation (Carson 1996: 134). Ehrlich and Holdren (1971) argues that population growth causes "a disproportionate negative impact on the environment" (Ehrlich and Holdren 1971: 1212). The argument is simple: Large populations can produce and consume more than small populations (Dietz and Rosa 1994). There are disagreements between scholars, but most studies have found that population, among other factors, is an important explanatory variable of environmental degradation (see Meyerson 1998; Li and Reuveny 2006: 943; Shi 2002; Stern 1996).

Trade has also been proposed as a control variable in studies of environmental degradation. There has been discussion regarding the direction of the relationship. First, some scholars have argued that increased trade results in increased environmental degradation. In its simplest form, trade is contributing to increased CO_2 emissions when goods are shipped between different parts of the world (Nordstrom and Vaughan 1999: 20). Most widely discussed is the *race to the bottom hypothesis* "which says that open countries in general adopt looser standards of environmental regulation, out of fear of a loss in international competitiveness" (Frankel and Rose 2005: 85). Several studies have demonstrated that growth in trade increases the emission levels of CO_2 (see Frankel and Rose 2005: 89). Second, there are scholars that argue that increased trade result in reduced CO_2 emission. Matthew A. Cole (2004: 72) states that as countries face greater competitive pressure they become more efficient in resource use. Other explanations for why trade may influence the environment have also been put forward: “Openness and competition will tend to increase investment in new technology which embodies cleaner processes to meet the higher environmental standards of technology exporting countries” (Suri and Chapman 1998: 196).

It is also plausible that the price of energy, and more specific, the price of engine fuel, influences the CO_2 emissions. The more expensive the engine fuel is, the less CO_2 will be emitted, simply because it costs more.

Improvements in the state of technology may affect the global environment in three ways. First, new technologies lead to new ways to discover and exploit natural resources. Second, technology may change the efficiency of production and consumption production. Third, different kinds of technologies may have different effects on the environment (for instance, fossil fuels and nuclear energy have different wastes (Stern et al. 1992: 83). Stern et al. argue that “modern technology is seen as a much more significant contribution to environmental degradation than either population or economic growth” (Stern et al. 1992: 83). It is difficult to control for technologies because of the lack of data, this variable will therefore not be included in the analysis, even though improvements in the state of technology may very well affect emissions of CO_2.

As seen in the argumentation above, both country specifics and characteristics of the international context can explain major shifts in environmental practices (Mitchell 2003: 449). By making use of control variables, the argument that the international regime caused observed changes in behavior becomes more credible because one can control for other factors for which it is reasonable to assume also influence emission patterns. The control variables can also serve as comparators, providing a basis for declaring a regimes influence as “high” or “low”. In addition, the control variables can be included in interaction terms, clarifying the influence of regime-related variables on the values of non-regimes variables (Mitchell 2004: 125). The following model might be proposed to model emission behavior including control variables:

$$\begin{aligned} \text{EMISSIONS}_{i,t} = {} & \beta_0 + \beta_1\text{*MEMBER OF REGIME}_{i,t} + \beta_2\text{*AFTER KYOTO}_{i,t} + \\ & \beta_3\text{*MEMBER OF REGIME}_{i,t}\text{*AFTER KYOTO}_{i,t} + \beta_4\text{*GDPPC}_{i,t} + \beta_5\text{*GDPPC}^2_{i,t} + \\ & \beta_6\text{*INDUSTRY}_{i,t} + \beta_6\text{*AGRICULTURE}_{i,t} + \beta_7\text{*POPULATION}_{i,t} + \beta_8\text{*TRADE}_{i,t} + \\ & \beta_9\text{*PRICE OF ENERGY}_{i,t} + \varepsilon_{i,t} \qquad [2] \end{aligned}$$

What can this model tell us? β_1 represents the predicted difference in CO_2 emissions of the members of the ICCR, relative to the non-members. This variable measures the effect of being in the treatment group, relative to the effect of being in the control group, before the treatment is executed. β_2 is the predicted difference between the period after the ratification of the Protocol, relative to the period before. This variable represents the posttest. β_3 is the interaction term between membership of the regime and the period after the Kyoto Protocol, which measures the effect on the member countries after the Protocol was ratified. This variable measures the effect of the treatment. It is predicted that β_3 is negative, on the assumption that membership in the ICCR leads countries to reduce their emissions. Second, the coefficients of the control variables, β_4 through β_9 coincide with the estimated increase or decrease in CO_2 emissions that would arise from one unit increase in that independent variable (Mitchell 2004: 127). By including control variables in the analysis, the results are open to alternative explanations which may influence the effect the international regime has on emission behavior. There may, however, also be uncontrolled extraneous variables (Jackson 2008: 128). The functions of the different variables are summarized in table 2.

Table 2: The function of the variables included in the regression analysis

Variable	Function
Member of regime	Treatment vs. control group
After Kyoto	Posttest
Member of regime*After Kyoto	Effect of treatment
GDPPC, industry, agriculture, population, trade energy price	Control variables

This chapter's main focus has been to examine how the ICCR should be modeled in a regression analysis. However, before the regression analysis is executed, a closer examination of the collection and processing of data is necessary.

4. Methods for Collecting and Analyzing Data

This chapter investigates the methods for collecting and analyzing data. First, I introduce the data I have collected by evaluating the quality of these data. Here, I make use of five guidelines introduced by King et al. (1994). Second, the variable operationalization and data summary is presented. The variable operationalization is based at the evaluation of the theoretical concept and the data set that I have collected. Third, I introduce the methods of analyzing the data - the regression analysis, and more precise, OLS (ordinary least squares) regression. Here, the reader is briefly introduced to the time-series cross-section (TSCS) design. Fourth, I present the possible problems of a OLS analysis with TSCS structure. The emphasis is on the fact that the assumptions of regression often is violated due to the TSCS structure. Fifth, I introduce some possible solutions to the problems related to TSCS models. These solutions will be followed in the empirical analysis, to show whether the results of the TSCS model can be trusted.

4.1 Methods for Collecting Data

Some guidelines exist for improving the data quality when one collects data. The following five guidelines will be discussed in relation to the data collection of this thesis, to ensure the quality of the data. First, one needs to record and report the process by which the data are generated. If one does not have this information, it is difficult to know whether using standard procedures in analyzing the data will lead to biased estimations. In quantitative analysis reporting data-generation refers to identifying where the data originates from, and what has been done with the data by the researcher (King et al. 1994: 23). The secondary data of this thesis is based at two different data sources; the International Energy Agency (IEA) and the World Bank's World Development Indicators (WDI). The data from the IEA contains estimates of CO_2 emissions by country from 1971 to 2008.[28] The estimates on CO_2 emissions have been calculated by the IEA by using the IEA energy database and the default methods and emissions factors from the *Revised 1996 IPCC Guidelines for National Greenhouse Gas Inventories* (IEA 2011a). The data on the control variables which is collected from the World Bank's WDIs covers 209 countries from 1960 to 2009 (World Bank 2011). The data from the World Bank follows the *Data Quality Assessment Framework (DQAF)* which is a methodology for assessing data quality developed by the International Monetary

28 Estimates of CO_2 emissions are available from 1960 to 2008 for the Annex B countries (International Energy Agency 2010: 3). However, the information from 1960 to 1970 could not be used due to the lack of data for countries not a part of Annex B countries.

Fund (IMF) and the World Bank.[29] However, not all of the time-series data can be used due to the shorter time-interval on the data of the dependent variable. In addition to these two sources of secondary data, I have generated two variables; one measuring membership to the Kyoto Protocol, and one measuring the period after the Kyoto Protocol. The information on the variable which measures membership to the regime is collected from the text of the Kyoto Protocol, where the members of the regime are identified in Annex B (see Kyoto Protocol 1997, Annex B). The Annex B-countries Monaco and Liechtenstein are excluded from this analysis because of lack of data on the dependent variable. I do not regard these two missing countries as problematic because they account for a small amount of the emissions of the Annex B countries. The variable that measures the Kyoto Protocol is a dichotomous variable where the period before the ratification is coded as 0 and the period after is coded as 1. The period before the Kyoto Protocol is defined as 1971 to 2004. The period after the Kyoto Protocol is defined as 2005 to 2008. This distinction is made because even though the Kyoto Protocol was adopted in 1997, it was not ratified until 16 February 2005. Hence, the Protocol was not legally binding until 2005 (UNFCCC 2011b).

This leads us to the second guideline to improve the data quality; to collect data on as many of its observable implications as possible. There are two different approaches to include more observable implications, one can i) collect more observations on the dependent variable, or ii) record additional dependent variables (King et al. 1994: 24). To collect more information on the dependent variable is of course preferable, both in time and space. Most importantly, to expand the time-series of this thesis beyond 2008 would have strengthened the conclusions of the analysis. However, it was not possible to obtain data from 2009 and 2010. It was possible to obtain data from the IEA which include 2008. In light of the fact that the member states are supposed to make demonstrable progress within 2005 and the fact that early action is required for the countries to be able to meet their commitments, data which include 2008 is regarded as adequate for an analysis of GHG emissions. Regarding additional dependent variables, alternatives such as the output (the rules and regulations)- and the impact indicator (change in nature) have been discussed earlier in the thesis.

The third guideline one has to account for when collecting data is to maximize the validity of the measures (King et al. 1994: 25). Validity, in its most basic form, refers to the approximate truth of a knowledge claim. A judgement of validity is therefore concerned with evaluating to what extent the relevant evidence supports that knowledge claim as being true or correct. No absolute consensus

29 For more information on the DQAF see IMF 2011.

exists on what validity consists of. However, some attempts to create a typology of validity has been made. These typologies have been made with the purpose to identify plausible alternative explanations (Farrington 2003: 51). Cook and Campbell (1979) developed one of the best known validity typologies, which is comprised by four related types of validity: statistical conclusion validity, internal validity, construct validity, and external validity. *The statistical conclusion validity* is concerned with whether the presumed cause and effect covary, and how strongly it varies. The most common way to address whether the cause and effect varies is null hypothesis significance testing (Shadish, Cook, and Campbell 2002: 42). Statistical conclusion validity is particularly important for the analysis in this thesis, due to its importance, I will therefore return to it later. *Internal validity* refers to "the correctness of the key question about whether the intervention really did cause the change in the outcome" (Farrington 2003: 52). Internal validity is concerned with whether A preceded B in time, that A covaries with B (however, this is already covered under statistical conclusion validity), and that no other explanations for the relationship are plausible (Shadish et al. 2002: 53). The statistical method equips the researcher with correlations (Holland 1986: 945), which does not say anything about causation. Therefore, the researcher needs to make sound causal inferences (Andersen and Wettestad 2004: 51). To be able to make sound causal inferences, it is important for the researcher not only focus at the statistical significance. The results of this analysis is therefore not only investigated through statistical significance, but also through a more substantial manner. *Construct validity* (also known as measurement validity) refers to "the adequacy of the operational definition and measurement of the theoretical constructs that underlie the intervention and the outcome" (Farrington 2003: 53). The researcher does empirical studies with specific instances of units, treatments, observations, and settings. These instances are of interest only because they can be defended as empirical measures of general theoretical constructs. To favor construct validity one has to be explicit about the theoretical constructs and the empirical instances of the study, and the match between these two (Shadish et al. 2002: 66). In this thesis it is important to be aware of the construct validity of the dependent variable: regime effectiveness. Many of the considerations concerning the validity of CO_2 emissions as the dependent variable of this thesis were explored in Chapter 3. *External validity* refers to "the generalizability of causal relationships across different persons, places, times, and operational definitions of interventions and outcomes" (Farrington 2003: 53). The ICCR is a case of an international regime. In general case studies have weak external validity because it is difficult to generalize from the findings of a case study (Barrington 2009: 15). For the findings of this study to be regarded as being generalizable, the representativity of ICCR as an international regime must be evaluated. I do not elaborate further on this evaluation; the point is to make the reader aware of external validity.

Fourth, one needs to ensure that data-collection methods are reliable. The reliability and consistency of the data are largely dependent upon the experience of the collecting staff and the reporting agents (IEA 2007: 101). Both the IEA and the World Bank have long experience with collection of data. The data from IEA are originally collected from official bodies (generally by national statistical offices) to meet national needs, and are then reported to the EIA using special questionnaires to ensure international consistency (IEA 2011b). A recent comparison of the data from IEA with the inventories submitted to the UNFCCC showed that the calculations were within 5 per cent (IEA 2010: 29).[30] Much of the WDI data also comes from the statistical systems of the member countries; therefore, the quality of the data will depend on how well these national systems perform (World Bank 2011).[31] The data of the WDI 2009 are generally consistent with those in WDI 2008. The data has been revisited and updated wherever new information has been available (World Bank 2009). Missing data also affects the reliability of the data. However, if the observations are missing at random, it is assumed that that missing data does not pose a problem for the analysis. Missing completely at random is satisfied if "the probability of missing data on Y is unrelated to the value of Y itself or to the values of any other variables in the data set" (Allison 2002: 3). The assumption of missing completely at random is often violated, especially when the units are countries: It is highly doubtful that different countries have the same standards for the national statistical systems, and that the quality of the statistical system is unrelated to variables such as income. However, data can be said to be missing at random if "the probability of missing data on Y is unrelated to the value of Y, after controlling for other variables in the analysis" (Allison 2002: 4). There is no reason to expect that the missing data on the dependent variable is due to their rate of CO_2 emissions. This means that the observations in this thesis can be considered as missing on random, which signify that no special procedures are requisite to make up for the missing data.

Fifth, all data and analysis should, insofar as possible, be replicable (King et al. 1994: 26). Basically, replicability comes down to giving as much information when presenting the data as someone else would need to reproduce the data and the analysis (Hancké 2009: 91). I have carefully described the encodes that I have done, so that replicability is possible. The encodes are described in the next chapter which deal with the data operationalization and summary of data.

---

30 For some EITs and non-Annex I countries the differences between the data of IEA and the data submitted to the UNFCCC were larger (IEA 2010: 29).

31 Especially, the statistical systems in many developing countries are still weak. However, the World Bank works to help developing countries improve the capacity, efficiency and effectiveness of their national statistical systems (World Bank 2011).

4.2 Variable Operationalization and Data Summary

The operationalization of the variables provides the necessary link between the theoretical concepts and the data to be tested. I start by explaining the operationalization of the variables, then I present a brief summary of the data.

CO_2 emissions, which is the dependent variable in this analysis, only includes data for energy-related CO_2. These data do not account for the other GHGs (emissions of CH_4, N_2O, HFC, PFC and SF_6). The IEA data does not include CO_2 emissions from non-energy-related sources and gas-flaring. However, it is estimated that combustion of fossil fuels alone account for 80 per cent of the total anthropogenic CO_2 emissions (Schmalensee, Stoker, and Judson1998: 15). The estimates do not take CO_2 emissions/removals from land use, land use change and forestry (LULUCF) into account (IEA 2010: 31).

The variable that measures membership of the ICCR is a dummy variable with four categories. This dummy set is a further development of the equations presented earlier in the thesis, where a distinction was made only between members and non-members. The first category in this dummy set is *Annex II countries* which include those of the Annex B countries which have ratified the Kyoto Protocol, and are not undergoing the process of transition to marked economy (EiT). The Annex II countries are mainly OECD countries, consisting of both leaders and laggards of the international climate policy. The EU has been the major industrialized leader, while the informal coalition JUSSCANNZ (Japan, the US, Switzerland, Canada, Australia, Norway, and New Zealand) was the major opponent (Oberthür and Ott 1999: 16). The Annex II category is therefore not a very uniform group in their view on climate change. However, they have all determined to reduce their CO_2 emissions. Note that the US was originally a part of the Annex II countries, however, due to the withdrawal of the US from the Kyoto Protocol, it is not a part of the Annex II category. The second category include Annex B countries which are identified as the *EiTs*. The EiTs are previous members of the Soviet union. The EiTs are characterized by some common features. First, they all experienced an economic collapse in the beginning of the 1990s, causing the GHG emissions to rapidly decrease. Second, the EiTs have a high level of industrial production, making it easy to apply efficient and highly cost-effective climate measures (Huang, Lee, and Chih 2008: 242). Even with steady economic growth, which will increase their GHG emissions again, the emissions are unlikely to reach “Soviet” levels again, due to competition, energy prices and efficiency improvements (Oberthür and Ott 1999: 22). Some of the EiTs use a different base year than 1990 because they were allowed to choose the base year themselves (An overview of the different base years can be found in appendix, table 13). The third category include the *United States*, which has signed the Kyoto Protocol but not ratified it. President George W. Bush announced in March 2001 that the US would not ratify the Kyoto Protocol due to three objectives: First, the science of climate change was considered as uncertain, second, that the protocol did not impose emission limitations on developing countries, third, that the agreement would be too costly to the American economy (Hovi, Skodvin, and Andresen 2003: 1). The US therefore has no legal obligations in the Kyoto Protocol. However, because the US is so different from the rest of the countries in the non-Annex I group, it is most definitely not a developing country, I have chosen to let the US be a category of its own. The fourth category, which functions as the reference category in the analysis, is the *non-Annex I* countries. The non-Annex I countries are developing countries without any legal obligations in the Kyoto Protocol. The central concern of this group has been equity - their main concern is that their economic and social development should not be slowed down due to a problem caused mainly by the developed countries (Oberhür and Ott 1999: 27). The countries in the four categories are listed in table 3.

Table 3: The four groups of members to the Kyoto Protocol

Category	Countries
Annex II	Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Japan, Luxembourg, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom
EiTs	Bulgaria[a], Croatia, Czech Republic, Estonia, Hungary[a], Latvia, Lithuania, Poland[a], Romania[a], Russia, Slovak Republic, Slovenia[a], Ukraine
US	US
Non-Annex I	Others

[a] base year other than 1990

The variable *after Kyoto* is a dichotomous variable, coded 1 for the period before the Kyoto Protocol was ratified, and 0 for the period after the Kyoto Protocol was ratified. The period before the Kyoto Protocol is defined as 1971 to 2004, this is the pretest. The period after the Kyoto Protocol is defined as 2005 to 2008, which is the posttest. This distinction is made because even though the Kyoto Protocol was adopted in 1997, it was not ratified until 16 February 2005. Hence, the Protocol was not legally binding until 2005 (UNFCCC 2011b). However, I control for the effect of a different posttest. Then I make use of 1997 to 2008 as the posttest.

To control for whether there have been changes in the CO_2 emissions of the members of the ICCR after the Protocol was ratified, an *interaction term* between the variable that measures membership and the variable that measures after Kyoto is introduced.

GDP per capita is gross domestic product divided by midyear population (World Bank 2010: 35). GDP is “the sum of value added by all residents producers plus any product taxes (less subsidies) not included in the valuation of output” (World Bank 2010: 35). Data are in constant 2000 US dollars. This variable is include to control for the level of economic activity. Because the data consists of time-series, in addition to cross-sections, economic growth over time is also measured by this variable.

The industrial origin of value added is determined by the International Standard Industrial Classification (ISIC) revision 3.[32] *Agriculture* value added corresponds to ISIC divisions 1-5 and includes forestry and fishing. *Industry* value added comprises mining, manufacturing, construction, electricity, water, and gas (ISIC divisions 10-45). *Services* value added corresponds to ISIC divisions 50-99 (World Bank 2009: 365). These variables control for the composition of the economic activity. In the analysis, services will function as the reference category.

The *population* variable counts all residents, regardless of legal status or citizenship, except refugees not permanently settled in the country in the country of asylum. These refugees are generally considered as a part of the population of the country of origin (World Bank 2009). This variable controls for the size of the population. Because the data consists of time-series, in addition to cross-sections, population growth over time is also measured by this variable.

Trade is “the sum of exports and imports of goods and services measured as a share of gross domestic product” (World Bank 2010: 357).

To account for energy prices, the variable *pump price* for diesel fuel is included in the analysis. Fuel prices refers to the pump price of the most widely sold grade of diesel fuel. Prices are converted from the local currency to U.S. dollars (World Bank 2010). The data on diesel pump price was only available for every second year, the variable was therefore extrapolated.

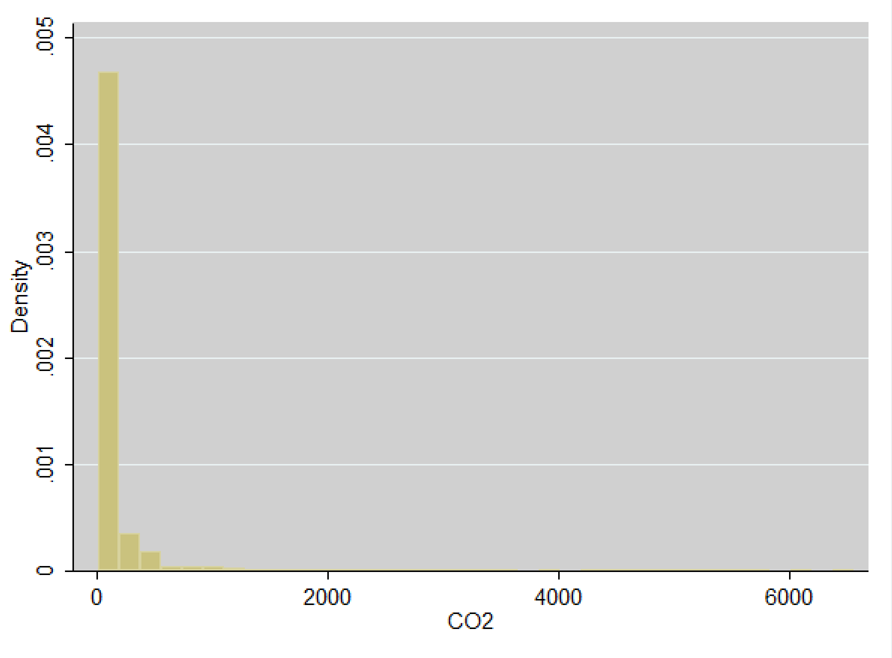
32 The specifics of the classification of agriculture, industry and services that WDI employes can be found at the homepage to United Nations Statistics Division (UNSD) (2011). WDI 2009 make use of revision 3 of the ISIC.

Table 4: Data Summary, 1971-2008

	Observations	Mean	St.Dev	Min	Max	Percent coded 1
CO_2 emissions	4514	158,13	570,41	0,19	6550,5	-
Annex II	4978	-	-	0	1	16,79
EiT	4978	-	-	0	1	9,92
US	4978	-	-	0	1	0,76
Non-Annex I	4978	-	-	0	1	72,52
After Kyoto	5978	-	-	0	1	10,53
Annex II* after Kyoto	4978	-	-	0	1	1,77
EiT* after Kyoto	4978	-	-	0	1	1,08
US* after Kyoto	4978	-	-	0	1	0,08
Non-Annex I* after Kyoto	4978	-	-	0	1	7,63
GDP per. capita	4258	6899,5	9147,17	80,62	56624,73	-
GDP per capita^2	4258	131000000	294000000	6500,33	3210000000	-
Industry	3910	33,14	12,2	90,51	90,51	-
Agriculture	3912	15,15	13,2	0,06	71,76	-
Services	3910	51,14	13,16	8,15	92,24	-
Population size	4956	3,85	3,85	0,01	132,47	-
Trade	4172	74,49	45,72	0,31	438,09	-
Pump price	1891	0,56	0,36	0,01	1,2	-

Figure 5 below shows the distribution of the dependent variable, CO_2 emissions. The distribution indicates that there are many outliers with high CO_2 emissions.

Figure 5: The distribution of CO_2 emissions



The data has now been presented through an evaluation of the data quality and operationalization of the variables, the next section will therefore elaborate on the methods for analyzing these data.

4.3 Methods for Analyzing Data

A regression analysis of CO_2 emissions across countries, over time, is commonly referred to as pooled time-series cross-section (TSCS) data in the research field of political science. TSCS data is characterized by having repeated observations on fixed units (Stimson 1985: 918). TSCS data is only one of several ways to refer to clustered data.[33] TSCS is characterized as data with multi-level structure, where the higher level is the individual- or as in this thesis, the country. The lower level consists of the observations for the unit at each point in time (Petersen 2004: 344). Therefore, TSCS data are clustered in both the time and spatial dimensions.

There are several advantages of employing TSCS data, in comparison to pure cross-sectional or time-series analysis. The two most important advantages are the ability to combine the study of dynamic relationships (due to time-series), with the ability to model the differences, or heterogeneity, among the subjects (due to cross-sections) (Frees 2006: 5). First, being able to model dynamic relationships is important for the political science because much of the research conducted is devoted to understanding the causes and consequences of events, opinions, behavior and

33 The descriptors *panel data* and *longitudinal data* is also common within the social sciences.

institutional change as they unfold *over time* (De Boef and Keele 2008: 184).[34] Second, TSCS data has the advantage of being able of include heterogeneity. Heterogeneity means that “observations from different subjects tend to be dissimilar whereas observations from the same subject tend to be similar” (Frees 2006: 243).

A linear regression analysis estimates expected Y as a linear function of X (Hamilton 1992: 66). This technique is commonly referred to as ordinary least squares (OLS). By making use of OLS regression it is possible to test whether and how the independent variables (X) influence the dependent variable (Y). OLS is commonly described by the following equation:

$$y_{it} = \beta_{0T} + \beta_{1T} x_{it} + \beta_{2T} Z_i + \varepsilon_{i,t} \qquad [3]$$

where y_{it} denotes a continuous dependent variable for individual i in period t. β_{0T} is the constant term. It is necessary to distinguish between those variables that are time-varying (x_{it}) and those that are time-constant (Z_i). Each observation on an unit in equation 3 is treated as an independent observation without acknowledging that the information comes from the same unit at different points in time. This may cause some problems for the error-structure of my regression analysis. These potential problems will be discussed in the following.

4.4 Problems related to TSCS Models

Data which are clustered often have problems with the error structure because of the cross-sectional and time-series features of the data. Both the political science and econometric research field have proposed solutions to how to model the heterogeneity and the dynamic relationships which occurs because of the space and time dimension. Before the problems related to TSCS models are presented, a brief review of the assumptions of OLS is necessary, because the most common problems of TSCS models occur because some of the core assumptions of a linear regression model may not be satisfied.

There are five assumptions that needs to be satisfied for employing OLS. These five assumptions will be presented in the following, with an emphasis on the two most important assumptions for analysis of TSCS data. First, one assumes a fixed X. Second, one assumes the errors to have zero

34 Dynamics refers to “any process where some variable has an impact that is distributed over time (Beck and Katz 2011: 332).

mean. These two assumptions together ensure unbiased estimation of all parameters.[35] These two assumptions cannot be tested in a regression analysis, and is therefore not the main focus of this discussion. Third, it is assumed that errors have constant variance (homoscedasticity). One may expect the errors in TSCS data to show "panel heteroscedasticity", where the variances of the errors differs from unit to unit. In my data this problem could arise because the scale of the dependent variable, CO_2 emissions, may differ between countries (Beck and Katz 1995: 636, Kittel 1999: 229). In unpooled data sets such differential variability is usually of modest concern. In pooled data, however, variability is likely to affect whole sets (e.g. all years for one country), hence, having a considerably greater potential for mischief (Stimson 1985: 919). In the presence of heteroscedasticity, OLS estimates will be unbiased, but the usual significance tests are often inappropriate and their use can lead to incorrect inferences (Long and Ervin 2000: 217). Fourth, one assumes that errors are uncorrelated with each other. Autocorrelation is often described as the most crucial assumption to meet in TSCS analyses because it is so easily broken (Wilson and Butler 2007: 102). The assumption of no autocorrelation can be split into two. First, errors for a particular unit at one time are unrelated to errors for that unit at all other times (no serial correlation). Serial correlation appears because of the dynamic relationships in TSCS data: One can assume that within one country, an observation in 2010, is correlated with an observation in year 2009. This is a plausible argument. One expect the Norwegian GDP of 2010 to be correlated with its GDP of 2009. The second form for autocorrelation is spatial correlation. Spatial autocorrelation appears because of heterogeneity, when errors for one unit are related to the errors for other units (Beck and Katz 1995: 636). One expects the economies of neighboring countries, such as Germany and France, to be linked. Autocorrelation is likely to lead to an upward bias in the estimated coefficients (Plümber and Neumayer 2010: 440). Under assumption three and four, standard errors are unbiased, and OLS is the most efficient estimator. The fifth assumption, normally distributed errors, is unnecessary for OLS to be BLUE (best linear unbiased estimator). However, this assumption justifies the theoretical use of the t and F distributions as hypothesis tests, which is important for an analysis based at regression analysis.[36] Various solutions have been proposed by the literature of TSCS to account for these problems in the error structure. Some of these solutions are presented in the following, with a focus on the problems which occurs because of dynamic relationships and heterogeneity (which cause autocorrelation to appear).

---

35 Unbiasedness means that "over the long run, sample estimates (b_k) center on the true parameter values (β_k)" (Hamilton 1992: 110).

36 In addition, given the errors are normally distributed, OLS is more efficient than any other unbiased estimator, linear or not (Hamilton 1992: 112).

4.5 Some Solutions to the Problems of TSCS

Today it is a common view within the literature to "regard these violations as interesting features to be modeled and not swept under the rug" (Beck 2001: 274). This approach is called sensitivity testing (Wilson and Butler 2007: 119). Sensitivity testing is actually a method of testing the statistical conclusion validity - one evaluates the sensitivity to detect a real contrast or difference between experimental conditions on some characteristic of interest. By investigating the relative variability in the baseline model, one can evaluate how reliable the estimates are (Lipsey 1990: 14). To investigate the reliability and robustness of the results is the main reason why an extensive sensitivity analysis is conducted in this thesis. Sensitivity testing is continually getting more attention within the TSCS methodologies, since "[a] fragile inference is not worth taking seriously" (Leamer 1985). It is important for researchers to carefully consider a variety of specifications, alternative models, appropriate diagnostics, and the long-established pitfalls of regression analysis. The main interests of my sensitivity analysis is what happens to the sign, the magnitude, and the statistical significance of the estimates.[37] To evaluate the sensitivity of the results, several methodological alternatives can be employed. It is "critical that we learn to assess the properties of complicated estimation strategies" (Katz and Beck 1995: 645). I therefore put great emphasis on describing *what* the more technical approaches actually do with the data, *why* I employ this technique, and *how* to interpret the results. The different techniques, which I will apply to the data in the sensitivity analysis, will be introduced in the following.

4.5.1 Modeling Heterogeneity

The simplest way to model unit heterogeneity is to allow the intercepts to vary by unit, in this case countries (Beck 2008: 482). By letting the intercepts vary by unit one deals with spatial autocorrelation. Two models have been suggested to allow for unit to unit variation in the model parameters: The fixed-effect estimator and the random-effects estimator. However, even though the random-effect estimator is frequently applied to TSCS data, it is often an unfit estimator, because the random-effect model is based on some crucial assumptions that often is violated.[38] Petersen (2004) argues these are strong assumptions to make, and are therefore often violated. Because the assumptions of the random-effect model are violated, the random-effect estimator will not be employed in this analysis.

37 The sensitivity analysis is largely based at the same approach as Wilson and Butler (2007).

38 First, the unobserved individual effects should be randomly drawn from a population (Hsiao 2003: 43). When the units of analysis are countries, as in the case of this analysis, the assumption of the random units are difficult to rationalize (Wilson and Butler 2004: 9). Second, the individual-specific and idiosyncratic error terms v_i and e_{it} must be independent of each other and of the measured variables χ_{it} and Z_{it}.

The fixed effect model (also called the within or the least-squares dummy variables estimator) can be described as:

$$y_{it} = \beta_{0W} + \beta_{iW}X_{it} + \beta_{2W}Z_i + \beta_{3W}Z_i^u + \varepsilon_{it} \quad [4a]$$

$$= \beta_{0W} + \beta_{1W}X_{it} + \alpha_i D_i + \varepsilon_{it} \quad [4b]$$

This model is easy to estimate with OLS. In the fixed effect equation [4.3b] one includes a dummy variable D_i for each individual i and then estimates its effect $\alpha_{i,}$ one per person. Thus, $\alpha_i D_i$ accounts for the grouped nature of the data. The variable D_i captures the effects of the measured and unmeasured time-constant variables Z_i and $Z_i^{u,}$, as shown in equation [4.3a] (Petersen 2004: 333). The fixed-effect estimator has been proposed as a solution to the problem of unmeasured heterogeneity, also known as omitted-variable bias (Green, Kim and Yoon 2001; Plümper, Troeger and Manow 2005: 330). Omitted-variable bias arises from failure to measure all relevant variables. However, there are two drawbacks connected to the use of FE models. First, there is the problem of estimating time-invariant variables, and it severely biases the estimation of variables that have very little within variance. Because the FE model only uses the within variance for the estimation, and therefore disregard the between variation, it does not allow for the estimation of time-invariant variables (Plümper and Troeger 2007: 124). Therefore, a drawback of the FE estimator in this thesis is that it fails to estimate the time-invariant variable that measures membership to the Kyoto Protocol (Annex II, EiTs, US, and non-Annex I). Second, the fixed-effect estimator, because it only make use of the within-country variation, eliminates any common trends and external shocks to which all countries are equally exposed (Kittel and Winner 2005).

4.5.2 Modeling Dynamics

There are two common ways in which one can model the dynamic relationships in the data, to account for serial correlation. First, a lagged dependent variable has been suggested by several scholars to model dynamic relationships (see Beck and Katz 1995; Kittel 1999). In addition to having a methodological purpose, there are also good theoretical reasons for including the lagged dependent variable in this thesis. It is plausible to assume that this year's CO_2 emissions are dependent on the CO_2 emissions of previous years. By introducing a lagged dependent variable, CO_2 emissions is a function of past CO_2 emissions, as modified by GDP and control variables (Rabe-Hesketh and Skrondal 2005). When one introduces a lagged dependent variable in the regression analysis, the estimates have a different interpretation, as "the estimated effects of the covariates on the response *after controlling for the previous response"* (Rabe-Hesketh and Skrondal 2005). Therefore, a model with a lagged dependent variable does not treat serial correlation as a problem, but rather tries to explain the level of the dependent variable with the previous level of the dependent variable (Kittel 1999: 230). There is no consensus within the literature that a lagged dependent variable corrects autocorrelation (see Achen 2000; Keele and Kelly 2006; Wilson and Butler 2007).[39] It has been argued that a lagged dependent variable will absorb large parts of the trend, making the coefficients of the lagged dependent variable biased upwards, while the coefficients of the other independent variables are likely to be biased downward (Plümper, Troeger, and Manow 2005: 335). However, it is desirable to investigate whether the results changes.

Second, one can account for serial correlation in the error term by employing year dummies (Driscoll and Kraay 1998: 549). Remember that the error term (ε) is simply all the omitted variables, and by adding $y_{i,t-1}$ to the specification, one includes the common omitted variables at time *t-1* (Beck and Katz 2011: 339). This means that the year dummies capture time-series variations that are common to all countries, such as external shocks or new technology (Bassanini and Scarpetta 2001: 402). It is very likely that an international event, such as the financial crisis, would affect the CO_2 emissions of the world. However, controlling for these dynamics may cause the opposite problem to occur: downward bias of the estimated coefficient (Plümber and Neumayer 2010: 440).

The most common way to account for autocorrelation today is Beck and Katz's Panel Corrected Standard Errors (PCSE). PCSE was put forward as a (better) alternative to the pooled TSCS than the generalized least square (GLS), which had been the most common way to account for the error structure. The PCSE corrects the inaccurate standard errors of the OLS estimates, by taking into account autocorrelation and heteroscedasticity (Beck and Katz 1995: 638). PCSE is not suitable for this analysis because the PCSE functions only in temporally dominated TSCS models, "where a limited number of units are observed for a relatively long period of time" (Beck and Katz 1995: 686). With 38 years, and 128 countries, this analysis is not dominated temporally. Hence, the PCSE is not suited. I therefore want to check for the effect of robust standard errors. Robust standard errors are "a general means of empirically correcting variance-covariance estimates in the presence

39 Models with lagged dependent variables have previously been estimated with great frequency (Keele and Kelly 2005). Recently the research field has taken on the discussion that a lagged dependent variable can lead to biased coefficient estimates (Achen 2005). The current state of the research concerning a lagged dependent variable is that it is appropriate under certain conditions, and inappropriate under others (Keele and Kelly 2005).

of heteroscedasticity, clustering, and other forms of conditional dependence" (Zorn 2006: 329). It is argued that the robust standard errors are consistently close to the true standard errors (Petersen 2007).

4.5.3 Influential Cases

Another potential problem is influential cases. "A case is influential if its deletion substantially changes the regression results" (Hamilton 1992: 125). Influence results from a particular combination of values on the variables in the regression, not necessarily from irregular values on one or two variables. Figure 5 indicated that there are several observations in the data set which have high values on the dependent variable. However, it is possible to test for influential cases by making use of techniques such as dfbetas and Cook's D. Dfbetas are case statistics, while Cook's D measures influence on the model as a whole (ibid). By inspecting these measures of influence, one can investigate whether some cases exist with potentially high influence. Another, less technical approach, is to evaluate the influential cases in a more qualitatively manner. I also want to employ this non-technical approach to investigate possible influence of cases.

4.5.4 Alternative Specifications

I want to explore several alternative specifications to test how sensitive the results are to changes in the specification. Employing alternative specifications can help to reveal endogenous change such as maturational change (Baker, Pistrang, and Elliot 2002: 148), pseudoeffects masquerading as causal evidence (Campbell 1996: 377). Several alternative specifications will be explored in the sensitivity analysis: alternative pretest, alternative time-frame, alternative dependent variable, and alternative independent variables. These alternative specifications are intended to investigate the construct validity of the analysis. Does the collected empirical information accurately capture the concepts or variables contained in the theoretical model (Mitchell and Bernauer 2004: 84)?

This chapter has elaborated on methods for collecting and analyzing data. When methods for analyzing data was discussed I attached importance to the problems and potential solutions to the problems related to analyses of TSCS data. In the next chapter the empirical analysis will be conducted, with emphasis on discovering potential problems with the analysis.

5. Empirical Analysis: Descriptive, Regression and Sensitivity

The previous chapter investigated the methods for collecting and analyzing data. In this section, I analyze the data that I have collected. The empirical analysis carries out the methodology that the three previous chapters have argued is the most valid method possible with the present data availability. This section will give an answer to whether member countries have changed their emissions behavior after the Kyoto Protocol was ratified, and whether the changes in behavior are due to the ICCR. The empirical analysis is divided into three main sections; descriptive, regression, and sensitivity analysis. First comes the descriptive analysis. Here, the mean and standard deviation of the pre- and posttest for the members and the non-members of the Kyoto Protocol is compared. This will answer how the emission behavior to the different groups has changed after the Protocol was ratified. Second, I investigate further the two different groups of members that I have identified, the Annex II and the EiTs - to look for patterns of emission behavior in the data. Then I turn to the regression analysis. First, the baseline model is presented, and the results are discussed. The regression analysis makes it possible to control for the effect of other variables, and to test for statistical significance. Afterwards, I turn to the sensitivity testing, to test the baseline model for different specifications and to test the robustness of the findings. The sensitivity analysis shows whether the findings made in the baseline model can be trusted. Last, I will summarize the main findings of the empirical analysis.

5.1 Descriptive analysis

The main aim of the descriptive analysis is to investigate whether and to what extent the emissions behavior of the members of the Kyoto Protocol has changed after the Protocol was ratified. First I present a simple overview of the CO_2 emissions for the different groups of members, before and after the Protocol was ratified. Table 5 shows changes in CO_2 emissions between the pretest and the posttest in mean and standard deviations for the treatment group and the control group. The pretest represents the period of time from 1971 to 2004, while the posttest represents the period of time from 2005 to 2008. The main purpose of this table is to observe whether the emission behavior to the members and the non-members of the Kyoto Protocol have changed after the Protocol was ratified. In other words: Does the treatment, the Kyoto Protocol, make the treatment group change their emissions behavior?

Table 5: CO_2 emissions for the different groups of members, before and after the Kyoto Protocol was ratified

		Pretest (1971-2004)		Posttest (2005-2008)	
		Mean	St.dev	Mean	St.dev
Treatment group	Annex II	223.2	287.3	249.4	303.8
	EiT	264.5	608.3	86.5	104.1
Control group	US	4930.3	448.4	5703.8	81.8
	Non-Annex I	66.9	264.7	138.3	612.6

The mean for Annex II countries has increased from the pretest to the posttest, the standard deviation has also increased. Because the mean has increased, this indicates that the Annex II countries as a group have not managed to decrease their emissions after the Kyoto Protocol was ratified. The increased standard deviation means that the spread has increased within the emissions of the Annex II countries. The opposite emission trend occurs for the EiTs. Both the mean and standard deviation have decreased for the EiTs. The large difference between the standard deviation in the pretest and the posttest is not surprising - due to the considerably longer time-frame of the pretest in proportion to the posttest. In addition, one may suspect that some of the most extreme outliers have decreased their CO_2 emissions considerably due to the economic collapse in the previous Soviet states, causing both the mean and the standard deviation to decrease. The US has also increased its emissions in the posttest, relative to the pretest. Both the mean and the standard deviation have increased for the non-Annex I countries. This means that the non-Annex I countries is becoming a less homogeneous group, which is probably due to the rapid economic development in countries such as China, India and Brazil. The findings in table 5 indicate that the effect of the treatment on the treatment group is twofold: First, the Annex II countries have not decreased their emissions, second, the EiTs have decreased their emissions. Based on these findings, the effect of the ICCR on its members emission behavior is in doubt because the CO_2 emissions of the Annex II countries have not decreased.

However, it is still not clear whether the emission levels presented in table 5 is due to the ICCR, or whether other variables that can explain the emission behavior of the members of the regime exist. Before the regression analysis is conducted, with control variables, a further investigation of the CO_2 emissions of Annex II countries and EiTs are required to explore how the trends of CO_2 emissions for each member appear.

Table 6: Emissions trends for Annex II members, per 2008

	Emissions in base year	Emissions in 2008	Reduction commitment as stated by the Kyoto Protocol (%)[a]	Difference between the commitment and 2008 emissions (%)
Australia	260,1	397,5	8	34,6
Austria	56,5	69,3	-8	18,5
Belgium	107,9	111	-8	2,8
Canada	432,3	550,9	-6	21,5
Denmark	50,4	48,4	-8	-4
Finland	54,4	56,6	-8	3,9
France	352,3	368,3	-8	4,3
Germany	950,4	803,9	-8	-15,4
Greece	70,1	93,4	-8	24,9
Iceland	1,9	2,2	10	15,4
Ireland	29,8	43,8	-8	32
Italy	397,4	430,1	-8	6,6
Japan	1064,4	1151,1	-6	7,5
Luxembourg	10,5	10,4	-8	-0,6
Netherlands	155,8	177,9	-8	12,4
New Zealand	22	33,3	0	33,9
Norway	28,3	37,6	1	24,7
Portugal	39,3	52,4	-8	25
Spain	205,8	317,6	-8	35,2
Sweden	52,8	45,9	-8	-13,1
Switzerland	40,7	43,7	-8	6,9
UK	549,3	510,6	-8	-7,1
In average	224,2	243,4	-5,5	11,8

[a] The aggregate emission goals for the Kyoto Protocol will in this analysis be slightly different from those stated in the Kyoto Protocol due to the lack of data on Monaco and Liechtenstein.

Table 6 shows that the Annex II countries, as a common unit, in 2008 increased their emissions with 11,8 per cent in proportion to 1990-levels. This makes the difference between the collective goal of -5,5 per cent reduction set by the Kyoto Protocol, and the actual emissions, 17,3 per cent at the aggregate level (2008). These results support the findings made in table 5 - the Annex II countries have increased their aggregate emissions from 1990 to 2008. Thus, the Annex II countries are facing an underachievement of their Kyoto goals as the period of agreement is coming to an end. However, there are some exceptions to the general trend towards emissions-reduction failure among the Annex II countries; Germany, Sweden and the UK have all managed to decrease their emissions in proportion to 1990 levels. Because the gap between the collective emission reduction target and the actual collective emission trend this close to the deadline of 2012 is so prominent, the possibility that the Annex II countries can meet their commitments in the Protocol without making use of the flexibility mechanisms is low.

Table 7: Emissions trends for EiT members, per 2008

	Emissions in base year	Emissions in 2008	Reduction commitment as stated by the Kyoto Protocol (%)	Difference between the commitment and 2008 emissions (%)
Bulgaria[a]	82,1	48,8	-8	-40,6
Croatia	21,6	20,9	-5	-3,2
Czech Republic	155,1	116,8	-8	-24,7
Estonia	36	17,6	-8	-51,1
Hungary[a]	80,1	53	-6	-33,8
Latvia	18,6	7,9	-8	-57,5
Lithuania	33,1	14,2	-8	-57,1
Poland[a]	343,8	298,7	-6	-13,1
Romania[a]	187,8	89,9	-8	-52,1
Russia	2813,2[b]	1513,1[c]	0	-46,2
Slovak Republic	56,7	36,2	-8	-36,2
Slovenia	14,4	16,7	-8	13,8
Ukraine	687,8	309,6	-8	-65
In average	348,5	195,6	-6,8	-35,9

[a] Baseyear other than 1990
[b] The estimate of Russia is based at a interpolation[40], due to missing on that particular year
[c] This estimate is from 2004, due to missing for Russia in the time period 2005-2008

Table 7 shows that the EiTs, as a common unit, has decreased their CO_2 emissions by 35,9 per cent in proportion to 1990-levels. This means that the CO_2 emissions in 2008 lies 29,1 per cent under the emission targets set by the Kyoto Protocol, thus amounting to an overachievement of the emission-targets for this group of countries. This also supports the findings made in table 5 - the EiTs have decreased their emissions from the period before the Protocol was ratified (the pretest) in comparison with the period after the Kyoto Protocol was ratified (the posttest). These findings are not very surprising due to the economic collapse of the Eastern European economy after the Soviet

40 Interpolation refers to “the estimation of a figure or value within the given limits of the available data” (Aggrawal and Khurana 2010: 3).

Union was dissolved. I will return to this argument during the regression analysis.

Table 8: The percentage changes from 1990 to 2008 in emissions among the treatment and control group

		Goal (%)	Change in emissions (%)	Difference between Kyoto and actual emissions (%)
Treatment group	Annex II countries	-5,5	11,8	17,3
	EiTs	-6,8	-35,9	-29,1
	Total for members	-6,1	-12,1	-5,9
Control group	US	-	13	-
	non-Annex I	-	53,6	-
	Total for non-members	-	33,3	-

In table 8 the overall change in percentage emissions of both the treatment - and the control group are investigated for the time-period 1990-2008. There are several conclusions that can be drawn from the information presented in table 8: The control group consisting of the US and developing countries have increased their emissions. The increase in CO_2 emissions is particularly large for the developing countries, which most likely is a reflection of the economic development in countries such as China and India. The Annex II countries have not decreased their emissions. The EiT countries have significantly decreased their emissions. Annex II countries are facing an underachievement, while the EiTs are facing an overachievement of the emissions goals as identified in the Kyoto Porotocol. Therefore, the Annex II countries can most likely only meet their emissions requirements by the end of the period of agreement if they make use of the flexibility mechanisms of the Kyoto Protocol. By making use of the emission trading system, the Annex II countries can purchase part of the EiT's emission allocation. It is not realistic that the Annex II countries can decrease their emissions by 17,3 per cent in the period 2009-2012, which they would have to if they were to comply to their common emission target without making use of the flexibility mechanisms.

Before the regression analysis is conducted, a brief review of the main findings made in the descriptive analysis is necessary. The descriptive analysis has shown that Annex II countries have increased their emissions, while the EiTs have decreased their emissions. This means that the Annex II countries underachieve, while the EiTs overachieve their commitments to the Kyoto Protocol. Together, the members of the Kyoto Protocol, in the end of 2008, have decreased their common CO_2 emissions. This means that it is probable that the members of the Kyoto Protocol can meet their emission limitations in the end of 2012 if they make use of the flexibility mechanism. However, the descriptive analysis does not account for the fact that the changes in the emission behavior can be due to coincidence or spuriousness. The most common way to check for coincident is by employing significance testing. Spuriousness can be checked for by making use of control variables. A regression analysis of CO_2 emissions can answer both of these questions.

5.2 Regression analysis

In this section I will present the baseline OLS regression model, and interpret the results. This model accounts for coincidence by testing for statistical significance, and spuriousness by employing control variables. The baseline model is presented in table 9, where model 0 is the baseline model without the interaction terms. The focus in the description of the results will be model 1.

From table 9 we can determine whether membership in the ICCR significantly affects emission behavior, after controlling for other variables. First, the dummy set consisting of Annex II countries, EiTs, the US, and the reference category non-Annex I can tell how membership in one of these groups affects emission behavior. The coefficient must always be seen in relation to the reference category, non-Annex I. Due to the positive coefficient of Annex II countries of the Kyoto Protocol one can see that in the period 1970 to 2008 this group of countries had emitted more CO_2 than non-Annex I counties. The EiTs have also emitted more CO_2 than the non-Annex I countries, in addition the EiTs have emitted more than the Annex II countries. The US has emitted considerably more than any of the other groups of members. The variable after Kyoto shows whether the emissions behavior of the world has changed after the Kyoto Protocol was ratified. After Kyoto shows that the world as a whole has emitted more CO_2 in the time-period after the Kyoto Protocol was ratified than the period before. The interaction term between membership and after Kyoto shows the actual effect the Kyoto Protocol has on the different groups of members. Because the interaction term is negative for both the Annex II countries and the EiTs , one can see that they have decreased their CO_2 emissions relative to the non-Annex I countries after the agreement was ratified. The US has continued to increase its CO_2 emissions after the protocol was ratified. Table 9 therefore indicates that both the Kyoto Protocol causes changes in emission behavior among its members. The main descriptive findings of the effect of the ICCR from table 5-7 is reflected in the findings of the regression analysis in table 9. However, there is one large, and important difference; the Annex II countries have, according to the regression analysis, decreased their emissions after the Kyoto Protocol was ratified. This is the opposite finding that was made in the descriptive analysis.

Possible answers to these two contradicting findings will be investigated later.

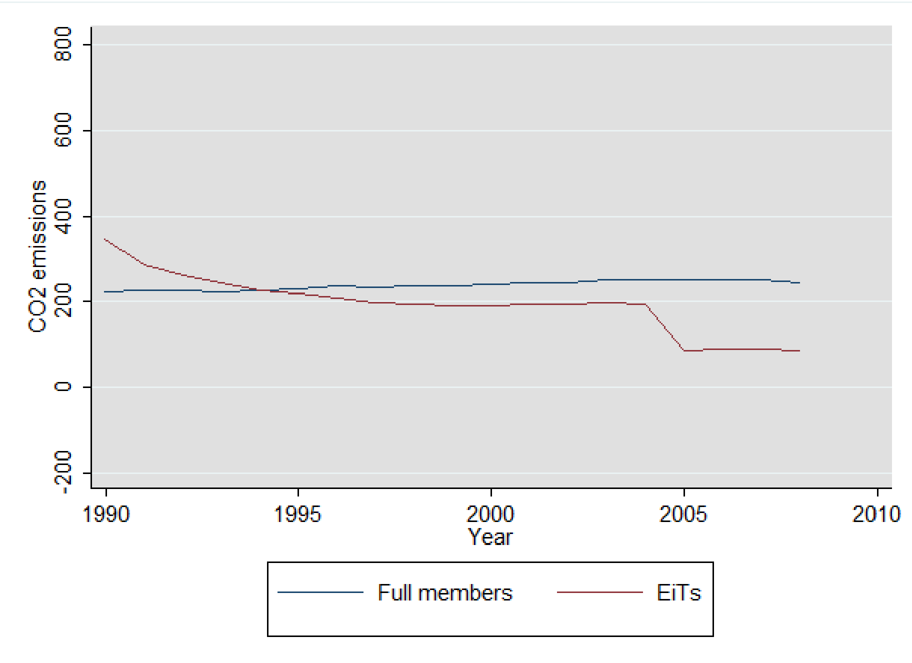
Table 9: Baseline model of CO_2 emissions

	Model 0	Model 1
Annex II[a]	143.960*** (50.587)	97.573** (46.261)
EiT[a]	225.552*** (60.989)	222.705*** (52.112)
US[a]	3985.615*** (207.100)	3929.431*** (177.548)
After Kyoto	10.887 (6.999)	19.251** (8.151)
Annex II*After Kyoto		-44.353** (17.843)
EiT*After Kyoto		-57.597*** (21.793)
US*After Kyoto		521.107*** (67.568)
GDPPC	0.004*** (0.001)	0.010*** (0.002)
GDPPC^2		-0.000*** (0.000)
Agriculture[b]	-0.823* (0.480)	-0.810 (0.500)
Industry[b]	0.099 (0.361)	0.365 (0.365)
Population	39.978*** (0.744)	37.625*** (0.714)
Trade	0.047 (0.118)	0.093 (0.093)
Constant	-106.919*** (29.924)	-113.781*** (29.006)
R^2	76.1	78.0
Countries	128	128
Observations	3710	3710

Note: Standard errors in parentheses, * significant at 10 %, ** significant at 5%, *** significant at 1%
[a] Reference category is non-Annex I
[b] Reference category is services

Even though the results are diverging for the Annex II countries, the conclusions so far are much more certain for EiT countries. The results so far show that the EiTs are facing a large overachievement of their Kyoto targets. As discussed earlier, there has been a fear of the Eastern European excess emissions (hot-air) since the US withdrew from the Kyoto Protocol. So far, both the descriptive and the regression analysis support the hot-air hypothesis. The Eastern European economies have a large amount of emissions rights to spare - most notably a result of the economic collapse, not a result of the Kyoto Protocol. The hot-air hypothesis is illustrated in figure 6: The Annex II countries have increased their emissions in the period 1990-2008. The EiTs have decreased their emissions dramatically, from over 300 to around 100 in the period 1990-2008. This downward trend of the EiTs from 1990 coincides with the shut down of heavy industries in the former Soviet states (see appendix, figure 8). Making the economic collapse, not the Kyoto Protocol, the most probable cause of the decrease in the CO_2 emissions of the EiTs. The trend is clear - in the start of the 1990s the economic activity collapsed, making the CO_2 emissions drop considerably. The downward jump in the estimates from 2004 to 2005 in figure 6 is due to missing on the estimates for Russia in the period 2005-2008. The data employed in figure 6 are interpolated due to missing on the variable that measures CO_2 emissions for some of the EiTs.

Figure 6: Emission trends for Annex II countries and EiTs in the period 1990-2008



Earlier it was pointed out that the flexibility mechanisms might lead to free-riding. The results from the analysis so far confirm that this may very well happen. The supply of hot air would lower the international permit price, hence making the international market price of emissions quotas low (Böhringer and Rutherford 2009: 182). With that, free-riding appears.

The findings on emission behavior from the baseline model are presented in table 10. This table shows the predicted mean values of the pretest and the posttest for the treatment group (Annex II countries and EiTs) are compared to the mean values of the control group (the US and non-Annex

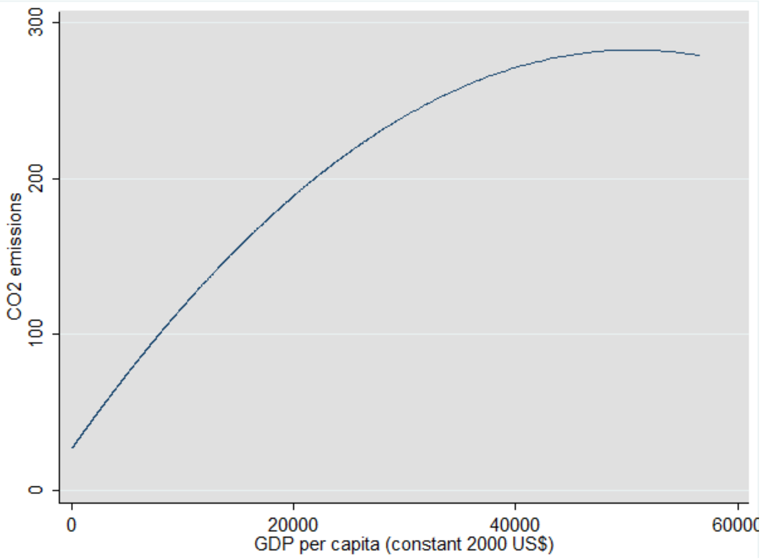
I). These estimates are predicted values founded on the results from baseline model, which is different from the actual mean that was presented in table 5. Table 10 illustrates the findings made in the regression analysis. Both treatment groups have decreased their emissions - the Annex II countries from 204,4 to 179,3, and the EiTs from 329,5 to 291,2. This finding indicates that after the Kyoto Protocol was ratified in 2005, both of the member groups of the Protocol have decreased their CO_2 emissions. The two groups identified as non-members have continued to increase their emissions after the Kyoto Protocol was ratified. Remember that these effects were all statistically significant in the baseline model.

Table 10: Predicted mean values for Annex II countries, EiTs, the US, and non-Annex I countries, before and after the Kyoto Protocol was ratified

	Treatment group		Control group	
	Annex II countries	EiTs	US	Non-Annex I countries
Pretest	204,4	329,5	4036,2	106,8
Posttest	179,3	291,2	4576,6	126,1

The findings from the baseline model together indicate that the Kyoto Protocol causes its members to decrease its emissions, hence the ICCR does affect emission behavior. This would mean that the ICCR is an effective regime. However, this finding is in contradiction with the results from the descriptive analysis. Possible reasons for this, and how it affects the conclusions that can be made from these analyses, will be investigated later. Before a further discussion on the results on the variables that measures membership to the Kyoto Protocol is examined through the sensitivity analysis, the effect of the control variables needs to be investigated. The control variables also show some interesting results. GDP per capita is positive, while its squared term is negative, both are statistical significant. The effect of GDP per capita on CO_2 emissions is illustrated in figure 7. The other variables are at their mean values in this figure.

Figure 7: Predicted CO_2 emissions by GDP per capita



This figure reflects the theory of the EKC; CO_2 emissions increase with GDP per capita only up to a certain level, where the emissions will start to decrease with GDP per capita. As studies within the EKC literature have shown earlier, the turning point is at extremely high income levels (Holtz-Eakin and Selden 1995, Stern 2004), in this analysis the turning point is at approximately US$ 50 000 GDP per capita. This indicates that even though the EKC hypothesis is supported in the analysis, the variable that measures the quadratic term of GDP has little substantial effect on CO_2 emissions. The dummy set that accounts for the structure of the economy is not significant in the baseline model. This indicates that the structure of the economy does not affect CO_2 emissions. Structure of the economy is the variable with most missing values in the dataset, I therefore ran the baseline model without this variable to see whether the results of the other variables altered. The results did not change, and the variable is therefore not excluded in the analysis (appendix, table 14, model ii,). Population size is positive and significant, meaning that larger populations emits more CO_2 than small ones. Trade is insignificant. This means that the level of trade in a country does not affect the level of CO_2 emissions. The variable diesel pump-price was also explored as a control variable. However, because the only time-series available for diesel pump-price is from 1995 to 2008, and the variable is insignificant, the diesel-price variable is excluded from further analysis. The baseline model with the diesel-price variable included can be found in the appendix (table 14, model i). I also ran the baseline model with an interpolated dependent variable, to be sure that the missing on the dependent variable did not influence the findings. The results did not alter (see appendix table 14, model iii).

To sum up the findings so far; the descriptive analysis of table 5-8 all indicate two diverging trends among the members of the Kyoto Protocol: First, the Annex II countries have increased their CO_2 emissions after the Kyoto Protocol was ratified. The second trend observed in the descriptive analysis is that the EiTs have decreased their CO_2 emissions tremendously in absolute terms, and also hugely in relation to previous emissions and the Kyoto-targets. The regression analysis confirmed the downward emission trend for the EiTs. However, the regression analysis did not confirm the results from the descriptive analysis for Annex II countries: The regression results for the Annex II countries showed that the Annex II countries have decreased their emissions. This means that the findings from the descriptive analysis and the regression analysis are not in accordance with each other. An explanation for these two diverging findings will be elaborated on below.

The most probable explanation for these two different findings is that there may be a problem with the statistical conclusion validity in the regression analysis: the conclusions from the regression estimates may not be accurate reflections of the real world (Adèr and Mellenbergh 1999: 326). These estimates may be making the researcher draw false conclusions. This implies that the analysis of the baseline model may be subject to a Type I error; one incorrectly conclude that cause and effect covary when they do not (Shadish et al. 2002: 42). There are several threats to statistical conclusion validity (see Shadish et al. 2002 for a thorough review). I evaluate the most probable cause to be that the assumptions of statistical tests are violated. As argued earlier, the data of this analysis are particularly sensitive to violated assumptions because of the TSCS structure. The violation of assumptions for regression analysis may result in biased estimates. The results from the regression analysis therefore need an extensive sensitivity testing to investigate the sensitivity of the estimates to a variety of specifications, alternative models, and appropriate diagnostics. The different methodological approaches to investigate the assumptions of regression analysis will be examined below to answer the following question: Can the estimates from the regression analysis be trusted?

5.3 Sensitivity testing

In the regression analysis control variables were employed to evaluate whether there are variables other than membership to the ICCR that can explain changing emissions-behavior. In addition to test for control variables it is desirable to investigate the robustness of the findings made in the baseline model to be sure that the results can be trusted. The credibility of the estimates from the regression analysis is in doubt due to the diverging results from the descriptive to the regression analysis. The main focus during the sensitivity testing is the main independent variables of this thesis: The different groups of members, the variable that measure the period after the ratification of the Kyoto Protocol, and the interaction term between these variables. The control variables will therefore only be mentioned if changes occur in proportion to the baseline model. The sensitivity analysis consist of several tests: alternative specifications, removal of influential cases, testing for dynamics and heterogeneity, and testing for the assumptions for OLS. The focal point of the sensitivity analysis is what happens with the sign, magnitude, and statistical significance of the variables. The more these changes, the less robust the findings are. After the sensitivity analysis has been conducted, a brief summary of the main findings of the robustness test is presented, to conclude whether the results from the baseline model can be trusted.

5.3.1 Alternative specifications

I want to explore whether an alternative pretest changes the results in the baseline model. An alternative pretest may help to detect long-term trends that are endogenous. When employing 1997-2008 as the posttest, and 1971-1996 as the pretest, two notable changes occur (appendix, table 15, model i): First, the variable that measures after Kyoto (the posttest) is no longer significant. Second, the interaction term is no longer significant for the Annex II countries. What may these two important changes tell us? It is possible that the changes which appear in the posttest of 2005-2008 only appear in this time frame, not before. Hence, the after Kyoto variable and the interaction term become insignificant when an alternative pretest is employed. A reason for this finding is that the Annex II countries may have started to take action only after the Protocol was ratified in 2005, since the Kyoto Protocol was not legally binding until then. Hence, the alternative pretest does not show any changes in emission behavior among the Annex II countries. The fact that the effect of being an EiT after the Kyoto Protocol was signed actually becomes larger with an alternative pretest, supports the observation that the emission reductions observed within the EiTs are not due to membership in the ICCR. This decrease is most likely due to structural changes in the economy, which are unrelated to the Kyoto Protocol.

Another desirable alternative specification to test is to use an alternative time-frame on the analysis. The advantage of making use of the time-frame 1990 to 2008 in the regression analysis is that the longer time frame may not be very interesting, and just function as “noise”. Remember that the main interest of this analysis is whether the emissions of the member countries of the Kyoto Protocol have changed relative to the goals stated in the Protocol, which is a percentage decrease in regard to the 1990-emission level. The disadvantages are that the time-frame is considerably shortened, and much of the dynamics in the data is excluded. The analysis of the data from 1990 to

2008 shows much of the same results as the baseline model (see appendix, table 15, model ii). The largest difference is that the estimate of Annex II is no longer statistical significant. This means that when making use of the time frame 1990 to 2008, Annex II countries of the Kyoto Protocol have not emitted more CO_2 than non-Annex I countries. This change is due to trends in the global economy: The economic growth in large developing countries such as China and India has escalated in the 1990s and through the 2000s, making their CO_2 emissions increase. This finding is an example of a historical trend - the emission behavior of the countries is not decided by the ICCR, but rather by endogenous trends such as rapid economic growth within large developing countries. However, the analysis still shows that Annex II countries have decreased their emissions after the Kyoto Protocol was ratified - meaning that this model shows that the ICCR does affect emission behavior among its members (both Annex II countries and EiTs).

An alternative to total CO_2 emissions as the dependent variable is CO_2 per capita (appendix, table 15, model iii). The advantage of employing CO_2 per capita is that this dependent variable accounts for the countries' various size and population. Even though the analysis with this dependent variable may not say that much about the outcome indicator of regime effectiveness, it does say something about the responsibility - which is a very important consideration in the climate change issue. In addition, CO_2 per capita as the dependent variable can detect a selection-maturation threat - which indicates that the treatment effect is due to differential rates of normal growth across the treatment and control group. The members of the Kyoto Protocol (the Annex II countries and the EiTs) have larger per capita emissions than non-Annex I countries, and the US has the largest emissions, in the time-frame covered in the data set. The per capita emissions of the world have increased after the Kyoto Protocol was ratified. Now for the most interesting finding in this specification: The Annex II countries have increased their per capita emissions, after the Kyoto Protocol was ratified. This observation is in accordance with the findings made in the descriptive analysis: The descriptive analysis showed an increase in both the mean and the standard deviation for the Annex II countries - which is a reflection of increased per capita emissions for some of the Annex II countries.

Alternative independent variables may reveal whether the results found in the baseline model are due to the classification of the membership variable. To employ alternative independent variables is a way of investigating the construct validity of the thesis. Will the outcome of the analysis change if the operational definition is changed? Remember that assignment to the Kyoto Protocol (the treatment) is not controlled by the researcher, therefore, I have identified four different member groups and justified why I employ exactly these four groups. The alternative classification of membership that I want to employ is a twofold division between members and non-members of the Kyoto Protocol. Annex II countries and EiTs are both included in the member category, the US and the group earlier characterized as non-Annex I countries are now included in the non-member category. The regression analysis with the alternative classification shows that there overall is no significant differences in CO_2 emissions between member and non-members (see appendix, table 15, model iv). This finding is not very surprising because the non-member group now includes the US, which has high emissions. The after Kyoto variable is significant and positive, meaning that the emissions in the world have increased since the Protocol was ratified. The interaction term between members and after Kyoto is significant and negative. This finding means that the members of the Kyoto Protocol have, after the Protocol was ratified, decreased their emissions in relation to the non-members. We know, due to the results from the baseline model, that the US has not decreased its emissions. This classification of the main independent variable is therefore a more general, but possibly less correct approach than the approach chosen in the baseline model. None of the estimates of the control variables were altered. R^2 drops from over 70 per cent in the baseline model to 33,4 per cent. This implies that much of the power of explanation of the baseline model is due to the comparatively detailed membership variable, which allows for much heterogeneity to be modeled in the regression analysis: This means that there are large differences *between* the countries identified as being a part of the treatment group, and the countries identified as being a part of the control group.

5.3.2 Influential Cases

Another potential problem to the analysis is influential cases. By investigating influential cases, one can evaluate whether the coefficients in the baseline model is heavy influenced by only some few cases. Figure 9 in appendix shows that there are three countries that stand out in their size of CO_2 emissions: The US, Russia and China. The US is already a separate category in the analysis, making it less possible that the extreme values influence the analysis.

I believe that Germany and the UK may be affected by their local history, and that the estimates of the Annex II variable may be heavy influenced by these two countries. Both Germany and the UK benefited from the choice of 1990 as the base year: 1990 was the year of German unification, which was followed by the collapse of the East German economy, resulting in a substantial windfall for German GHG emissions. Between 1990 and 1995, 20 per cent of the fossil fuel-fired capacity in the UK was switched from coal to gas. In combination with modern combined cycle technology this resulted in 60 per cent lower CO_2 emissions for the same electricity output (Kellow 1999: 277).

These events may cause the history threat to internal validity - events occurring concurrently with treatment could cause the observed effect (Shadish et al. 2002: 55). I therefore want to investigate whether these events for Germany and the UK can explain the changes that occur in emission behavior among the Annex II countries. This suspicion is confirmed (see appendix, table 16, model i): When Germany and the UK are introduced as separate dummy-categories in the baseline model, the dummy that measures Annex II membership to the Kyoto Protocol becomes insignificant. The significance of the interaction term between Annex II countries and after Kyoto has dropped. The dummy for both the UK and Germany is significant and positive. The interaction term between UK and after Kyoto is not significant. The interaction term between Germany and after Kyoto, however, is significant and positive. These results confirm the finding made in the descriptive analysis, where it became evident that only three Annex II countries have decreased their emission: The UK, Germany, and Sweden.

I also employ more technical approaches to identifying potential influential cases. I utilize the natural logarithm of the dependent variable to test for potential outliers on the dependent variable. The results from the analysis with the natural logarithm of the dependent variable is similar to the results found in the baseline model. The only change that appears in this model is that the interaction term between the US and the after Kyoto variable is insignificant (appendix, table 16, model ii). The most obvious reason for this is that the US is an outlier, and that the results of the analysis change because of this. Because the US dummy is stable across the different specifications during the sensitivity analysis, I chose not to put too much emphasis on this finding.

I make use of Cook's D to measure the *i*th observation's influence on all coefficients in the *whole* model (Hamilton 2009: 221). After a visual inspection of the box-plot for Cook's D, I identified the influential cases to appear after 0,05 (appendix, figure 10). I therefore did the analysis of the baseline with a limit on the Cook's D value on 0,05. This resulted in 7 excluded observations. The major change that occurs is that after Kyoto is no longer significant. The significance of the interaction term between Annex II countries, EiT and after Kyoto has decreased. The exact same results occur in the baseline model when I made use of the other measure of influence; df-betas (see appendix, figure 11 and 12). Df-betas indicate how much each observation influences *each* regression coefficient (Hamilton 2009: 223). According to the box-plot for df-betas the influential cases appear after 0,5. The analysis of the baseline model was therefore conducted once again with a limit on the df-beta's value to 0,5. This lead to the exclusion of 15 observations. Together these two measures of influence show that the positive, and significant finding in the baseline model for after Kyoto most likely is due to influential cases on this particular variable. Therefore, the world as a whole has not emitted more CO_2 after the Kyoto Protocol was ratified, when one controls for influential cases.

5.3.3 Testing for Dynamics and Heterogeneity

As emphasized in the previous chapter, dynamics and heterogeneity cause autocorrelation (serial and spatial) to appear in the estimates. The Durbin-Watson test (sig = 0.000) revealed that there is autocorrelation in the baseline model. I therefore employ a lagged dependent and year dummies to try to model the dynamics, and I employ a fixed-effect estimator to model the heterogeneity (see appendix, table 17, model i). When I include year dummies, the most notable changes that happen is that the Annex II dummy is no longer significant, neither is the after Kyoto variable. One possible explanation for this is that the findings made in the baseline model, which were not significant in the model with year dummies, may be due to dynamics in the international economic conjunctions. The interaction term between Annex II countries and after Kyoto, is still statistically significant. This means that the negative emission trend in the baseline model for the member countries is not due to external influence which affects the whole world.

The model with the lagged dependent variable is the model where the estimates change the most from the baseline model (see appendix, table 17, model ii). As I emphasized in the description of a lagged dependent variable, this variable may absorb large parts of the trend, causing upward estimation of the lagged dependent and downward estimation of the other variables in the analysis. Several of the variables that are clearly significant in the baseline model become insignificant in the model with the lagged dependent variable. The only member group that is significant in this model is the US. After Kyoto is still significant, and positive. However, none of the interaction terms are significant anymore. This means that there is no effect of the ICCR, after controlling for last year's level of CO_2 emissions. Neither is GDP per capita significant in the model with the lagged dependent. Because the standard errors of the variables from the baseline model have decreased considerable it is plausible that the lagged dependent causes downward estimation of the other variables in this model. The lagged dependent is statistically significant, meaning that much of this years CO_2 emissions can be explained by last years CO_2 emissions. The R^2 increases from 78.0 to 97.6 per cent. This increase is not surprising because it is expected that much of the current response on the dependent variable can be explained by the previous response. Remember, there is no consensus in the literature that a lagged dependent actually do help against autocorrelation (see Achen 2000).

As emphasized earlier, the fixed-effect estimator has one big flaw, it cannot estimate variables that change little or not at all. This problem occurs when I employ the fixed-effect estimator to analyze the data (appendix, table 17, model iii). The estimates of the dummy set with member groups is therefore omitted because they have no within group variation, which is problematic because these variables are among the main focus of this thesis. Turning to the estimates of the other variables, they confirm the findings made in the baseline model. Because this model only estimates the changes that happens within the units, the R^2 decreases to 33,3 per cent. This analysis once again confirms that much of the differences in CO_2 emissions are due to large differences between the countries.

5.3.4 Testing for Assumptions of OLS

It is common to see panel heteroscedasticity in TSCS data. The Cook-Weisberg test for heteroscedasticity shows a significant (sig=0,000) level of heteroscedasticity in the baseline model. This finding implies that the standard errors and the significance testing may be invalid (Hamilton 2009: 213). If there is heteroscedasticity, the coefficients will be correct, but the standard errors will be biased downwards (Beck and Katz 1995: 645). The values of the significance testing will then be upwardly biased. To account for this, I ran a model with robust standard errors (appendix, table 18, model i). The robust standard errors alter many of the findings made earlier: The variable that measure the effect of being an Annex II country is no longer significant. The estimates for the EiTs and the US have not changed. The variable that measures the period after the Kyoto Protocol was signed, is no longer significant. The interaction term between Annex II countries and after Kyoto is neither significant. Thus, according to this model, Annex II countries have not decreased their CO_2 emissions after the Kyoto Protocol was ratified. However, the EiTs have decreased their emissions (p-value 0,055). The US has increased its emissions. GDP per capita is still statistically significant, however, the quadratic term of GDP per capita is not. The statistical significance is generally lower when employing robust standard errors, making it plausible that the statistical significance tests in the baseline model is upward biased. This means that it is probable that the estimates in the baseline model will show a Type I error, making the researcher conclude that cause and effect covary when they do not.

Multicollinearity may lead to low t-statistics, unstable results and insignificant coefficients despite high R^2. I therefore tested for variance inflation factor (VIF) to test for multicollinearity. Values of VIF which is greater than 10 is considered as too much variance influence (Hamilton 2009: 224).

There are two variables that have VIF values greater than 10; GDP per capita, and its squared term GDP per capita^2. This is not surprising because GDP per capita and GDP per capita^2 are closely related. As can be seen in the summary of the sensitivity analysis low t-statistics is not a problem, neither is insignificant coefficients despite high R^2. Unstable results are, however, a problem, but most likely it does not appear because of multicollinearity in the baseline model.

5.3.5 Summary of the Sensitivity Analysis

The purpose of the sensitivity analysis was to investigate what happens to the sign, magnitude, and the statistical significance of the estimates of the baseline model when one applies different specifications, remove influential cases, testing for dynamics and heterogeneity, and testing for the assumptions for OLS. In this section I try to summarize what main conclusions can be drawn from the sensitivity analysis that I have conducted in the previous. First, I evaluate the sign and statistical significance of the main independent variables in this thesis: The status of membership, after Kyoto, and the interaction term between these two variables. Second, I evaluate the magnitude of the estimates. The main findings are presented in table 11 and 12.

Table 11: Summary of findings from the sensitivity analysis for statistical significance

Test for	Model	Annex II	EiT	US	After Kyoto	Annex II*After Kyoto	EiT* After Kyoto	US* After Kyoto
	Baseline	**	***	***	**	**	***	***
Alternative specifications	Alternative pretest	*	***	***	-	-	***	***
	Alternative time-frame	-	***	***	***	***	***	***
	Alternative dep. variable	**	***	***	***	***	***	**
	Alternative indep. variables	-	-	-	***	***	***	***
Influential cases	UK and Germany as dummies	-	***	***	**	*	***	***
	Ln dependent	***	***	-	***	***	***	***
	Cooks D	**	***	***	-	*	**	***
	Df-betas	***	***	***	-	**	**	***
Dynamics	Year dummies	-	***	***	-	***	***	***
	Lagged dependent	-	-	***	**	-	-	-
Heterogeneity	Fixed effect	Omitted[a]	Omitted[a]	Omitted[a]	**	**	*	***
	Robust standard errors	-	**	***	-	-	*	***

Note: * significant at 10 %, ** significant at 5%, *** significant at 1%, - not significant
[a] Omitted because there is no within variation

Table 11 shows a summary of the findings from the sensitivity testing. This sensitivity analysis reveals that simple methodological alternatives can alter the estimates of the analysis. Many of the results were robust against alternative specifications. First, the estimates show large consistency in the sign, except in two instances.[41] The estimates of the significance testing do not show the same robustness: The findings for the dummies of EiTs, the US and the interaction term of the EiTs, the US and after Kyoto show great stability in the significance during the sensitivity analysis. The estimates of the Annex II countries and the interaction term between Annex II countries and after Kyoto show instability. Particularly unstable is the variable that measures after Kyoto, which also is unstable in its prediction of the sign.

Table 12: Summary of findings from the sensitivity analysis of the magnitude of the coefficients

Test for	Model	Annex II	EiT	US	After Kyoto	Annex II*After Kyoto	EiT* After Kyoto	US* After Kyoto
Alternative specifications	Alternative pretest	No change	Increase	Decrease	Decrease	Decrease	Increase	Decrease
	Alternative time-frame	Decrease	Decrease	Increase	No change	No change	No change	Decrease
	Alternative dep. variable	Decrease	Decrease	Decrease	Decrease	Decrease	Decrease	Decrease
	Alternative indep. variables	Decrease	Decrease	Decrease	No change	No change	No change	No change
Influential cases	UK and Germany as dummies	Decrease	No change	No change	No change	Decrease	No change	No change
	Ln dependent	Decrease	Decrease	Decrease	Decrease	Decrease	Decrease	Decrease
	Cooks D	Decrease	No change	No change	Decrease	Decrease	Decrease	No change
	Df-betas	No change	Decrease	Decrease	Decrease	Decrease	Decrease	Decrease
Dynamics	Year dummies	No change	No change	No change	Decrease	No change	No change	No change
	Lagged dependent	Decrease	Decrease	Decrease	Decrease	Decrease	Decrease	Decrease
Heterogeneity	Fixed effect	Omitted	Omitted	Omitted	No change	No change	Decrease	Decrease

Note: Estimates that increase in magnitude refer to those variables where the magnitude of the coefficient increased by least half of the initial standard error (from the baseline model). If the estimates did not change by more than half the initial standard errors, the findings were classified as unchanged. The last category is the estimates that fall in magnitude. This category include those variables where the magnitude of the coefficient decreased by least half of the initial standard error.

Table 12 shows that the baseline model is sensitive to model specifications in its estimates of

41 The exception is for the variable after Kyoto in the model which make use of 1997 to 2008 as the posttest, and when the values of the df-betas are included (see appendix, table 8.3, model i, and table 8.4, model iv).

coefficient magnitude. The magnitude of the coefficients decrease in 46 of the total 74 instances (62,2 per cent). There is no change larger than half the initial standard errors in 24 instances (32,4 per cent). The magnitude increase in 3 instances (4,1 per cent). None of the variables stands out as being more robust than others in the magnitude of the estimates. These findings indicate that the baseline model underestimate variance, and thus indicates stronger correlations than is true. This is probably due to a substantial level of autocorrelation and heteroscedasticity in the baseline model. As suggested earlier, the analysis of some of the estimates in the baseline model has a higher probability to a Type I error. Most notably the estimates of Annex II, after Kyoto and Annex II*after Kyoto from the baseline model cannot be trusted.

5.3.6 Conclusions from the Sensitivity Analysis

The sensitivity analysis has shown that the baseline model is sensitive to alternative model specifications. Table 11 showed that the significance testing was stable in some estimates, while it was unstable in others. The most stable estimates in the sensitivity testing were the estimates of EiTs, the US, and the interaction term between these two groups of members and the after Kyoto variable. These four variables were consistently strongly significant during the different model specifications, making the findings from the baseline model robust and credible. The variable that measures the effect of being an Annex II country is unstable in the sensitivity analysis. The interaction term and after Kyoto show more stability than the Annex II variable, however, the estimates are still not robust enough to make trustworthy conclusions based on these estimates.

In general, the findings decrease in magnitude when alternative specifications are added. None of the estimates of the variables stand out as being more robust than others against alternative specifications. Autocorrelation and heteroscedasticity are known to cause upward bias in the estimates, which makes these two violated assumptions the main suspects for the unstable estimates of Annex II countries and after Kyoto. Therefore, if the baseline model had been the only estimation technique, the probability of making a type I error in this analysis would have increased. Hence increasing the probability to incorrectly and conclude a covariation between the cause and effect.

Previously in this chapter, the descriptive analysis showed that the Annex II countries had increased their CO_2 emissions in the period after 2005, while the baseline regression model showed that the Annex II countries had decreased their emissions after 2005. After the sensitivity analysis, it is clear that the results from the baseline model are not robust. The conclusion I draw from these findings is that there cannot be made reliable statistical conclusions concerning the CO_2 emissions of the

Annex II countries: First, the results from the descriptive analysis and regression analysis show different results. Second, the findings from the regression analysis are not robust. There is no reliable statistical significant difference between the Annex II countries and the reference category. The results for the EiTs are much more uniform, and therefore more trustworthy. Both the descriptive analysis and the regression analysis show that the EiTs have decreased their emissions relative to the non-Annex I countries.

6. Conclusions

In this chapter I will summarize the thesis' research question, theoretical framework, methodological approach and the main findings. I also suggest a strategy for further development of the research field. Last, I make some concluding remarks regarding the effectiveness of the ICCR.

6.1 Summary and Main Findings

This thesis' research question was: *To what extent is the international climate change regime an effective regime?* The theoretical framework presented the development of the research field of international regimes. I pointed out two major contributions of my thesis to the research field in the introduction of this thesis to both provide and apply an analytical framework to investigate the level of regime effectiveness of the ICCR. I therefore dedicated most of my theoretical framework to this purpose. The theoretical framework was divided into four main debates, whose purpose was to function as an introduction to the research field, and to lead the reader into the main focus of the thesis: Regime effectiveness. Naturally, the first debate in the research field was concerned with the conceptual issues of an international regime. Second, a discussion concerned with how to explain formation and change of international regimes appeared. The third debate was concerned with which factors can explain regime effectiveness. Fourth, there was a discussion concerning the conceptualization and operationalization of regime effectiveness, which has been the main focus of this thesis. Three different indicators of effectiveness were presented: output, outcome, and impact. Which one of these indicators one chose, depends on where in the chain of causality one wants to measure effectiveness. After a discussion concerning the validity and data availability of the three different indicators, the outcome indicator was chosen as the indicator of regime effectiveness. This indicator makes use of the behavior of the members of the regime as the dependent variable. Changed behavior is a subobjective to the goal of the ICCR; reduction in GHG emissions. CO_2 emissions were therefore chosen as the unit of analysis, and regression analysis was chosen as the main method to answer the research question. These considerations on regime effectiveness were meant to bridge the gap between the theoretical and the empirical universe with methodology and to provide an analytical framework to evaluate the effectiveness of the ICCR.

The analysis of the effectiveness of the ICCR consisted of three elements; first, a descriptive analysis to investigate the actual emissions behavior of members to the Kyoto Protocol. Second, a regression analysis, which main purpose was to test for statistical significance, and to investigate the effect of control variables. Third, a sensitivity analysis of the findings from the baseline model was conducted to examine the robustness and trustworthiness of the results. The analysis of this thesis is where I apply the theoretical framework, to my empirical case: The ICCR. The descriptive analysis showed that the Annex II countries had increased their emissions, while the EiTs had decreased their emissions after the Kyoto Protocol was ratified. The two groups identified as the control group, the US and the non-Annex I countries, have both increased their emissions after the Protocol was ratified. However, the descriptive analysis cannot answer whether these differences are due to interfering variables or whether the findings are statistical significant. Therefore, a regression analysis was conducted. The baseline model showed that Annex II countries and the EiTs have both reduced their emissions after the Protocol was ratified, relative to the non-Annex I countries. The US has increased their emissions. The finding from the regression analysis indicated that Annex II countries have decreased their emissions after 2005 was in contradiction to the results from the descriptive analysis. The results from the sensitivity analysis showed that this finding was not robust to different specifications. Hence, the estimation of Annex II after the Kyoto Protocol was ratified (the interaction term) could not be trusted. The finding for EiTs and the US was, on the other hand, strongly robust to different specifications.

This leads to the answer of the research question - to what extent is the ICCR an effective regime? Based on the findings made in the analyses of this thesis, I argue that the ICCR is not an effective regime because the changed emissions behavior is not a consequence of the ICCR. First, the descriptive analysis showed that the Annex II countries have not reduced their emissions. The regression analysis showed us the opposite finding, namely that the Annex II countries have decreased their emissions. However, this finding is not robust across different specifications, making the regression coefficient not reliable. Second, both the descriptive and the regression analysis showed that the EiTs have decreased their emissions. These results were robust during the sensitivity analysis, making the result credible. Nevertheless, I have argued that much of this reduction in emissions in the EiTs is not due to the ICCR, but due to the economic collapse of the former Soviet-states. Had this thesis' main focus been compliance, the conclusion would have been different. The Parties of the Kyoto Protocol will most likely be able to meet their common commitments in the end of 2012, if they make use of the flexibility mechanisms. Barret (2007) describes this as “a loophole that ultimately fail to reduce global emissions” (Barret 2007: 6). The conclusions of this thesis are in accordance with this characterization: The existence of under-achievements for the Annex II countries, and an over-achievements for the EiTs, makes it probable that the Annex II countries will buy emission quotas from the EiTs. The problem with this is that the over-achievement of the emission targets for the EiTs is not due to the ICCR. This leads to an ineffective regime, because it ultimately does not change the emissions behavior of its members.

This thesis has given an answer to the level of effectiveness for the ICCR. There are, however, some additional questions that have not been addressed in this thesis, which naturally appear as a consequence of the findings of this thesis. Why is the ICCR not effective? How can one compare the effectiveness of the ICCR to other environmental regimes? I will briefly review my suggestions to further research for how to answer these questions in the following section.

6.2 Further Development

For future research I will suggest two approaches to further development of the themes addressed in this thesis: A theoretical and an empirical development. First, a theoretical development concerning how to measure effectiveness across different regimes. In the following, I will argue how one can apply the analytical framework I have developed in this thesis to compare the effectiveness convincingly across different regimes. Second, I will suggest an empirical development: Why is the ICCR an ineffective regime? I suggest that an analysis of this research question can be based on the behavioral prerequisites introduced in the theoretical framework of this thesis.

5.2.1 Theoretical Development: How to Measure Effectiveness Across Regimes?

This thesis has been concerned with providing and applying an analytical framework to investigate the level of effectiveness of the ICCR. The analysis of this thesis is based on *one* case. I will in the following give a brief introduction to how the analytical framework of this thesis can be applied to other international regimes, so that comparing effectiveness across regimes becomes possible with the outcome indicator.

One of the largest questions within the research field of international regimes today concerns the possibility to compare convincingly the effectiveness across different regimes (Hovi et al. 2003). Quantitative methods have been proposed as a solution to this question: First, one can calculate percentage changes. Second, one can employ per unit effort, measured as abatement costs (Mitchell 2004). These two approaches have some considerable weaknesses. The calculation of percentage changes (as was conducted in paragraph 5.1 in this thesis) does not say anything about inference of variables, or statistical significance. The abatement criterion is difficult to measure convincingly. I want to propose a different solution to this problem of comparative regime effectiveness, namely the use of impact analysis (note that impact analysis must not be confused with the impact indicator), which is frequently used within the research field of policy analysis (see Easton 1965, Mohr 1995). A short introduction into the theoretical framework of policy analysis' impact analysis is needed to illustrate how the regime effectiveness field also can benefit from this framework.

The literature of policy analysis first divides the different designs into experimental, quasi-experimental, and ex post facto - where each of these may be further divided into subtypes, depending on whether there are estimates of i) the subjects at one or more previous time periods (pretest), and/or ii) a group of comparable subjects (control group) (Mohr 1995: 61). In this study, and in most pollution regimes in general, a pretest and a control group exists. There are observations on the dependent variable before the regime was established, and not all the countries in the world are members. This leads to a quasi-experimental design. By employing the diagram proposed by Mohr (1995: 62), one can calculate a score of effectiveness. The diagram consists of standardized measures, making it possible to compare the scores of different regimes. The scores may prove valuable when the goal is to attach a level of effectiveness to an international regime, and compare this to the effectiveness of other regimes. This approach rests on much of the same techniques as the Oslo-Potsdam solution does.[42] However, the approach I suggest would be closely connected to the analytical framework the policy analysis field has proposed. The research field of policy analysis has developed an advanced analytical framework, which can be applied to evaluation of international regimes - remember, there is no need to carry the coals to Newcastle.

6.2.2 Empirical Development: Why is the ICCR an Ineffective Regime?

Based on the findings of this thesis, an interesting question to answer is why the ICCR is an ineffective regime? To elaborate on the potential reasons for the findings of this thesis: It is likely that one of the behavioral prerequisites of the outcome indicator is not present in the case of the Kyoto Protocol. The three behavioral prerequisites presented in the theoretical framework of this thesis were knowledge, motivation, and resources. I believe the most probable cause of the ineffectiveness of the ICCR is that the motivation prerequisite is not present. It has been noticed that the failure of governmental officers and institutions to deliver on policy promises concerning climate change might be a reflection on the fact that such policies are merely rhetorical rather than genuine commitments (Eckersley 2007: 310). Rhetorical, there are several examples of politicians who promise to keep their commitments, and broaden them as well: Andrea Merkel, the German Chancellor, announced in 2007 that the G8 nations committed to “taking strong and early action to tackle climate change” (The Times 2007). The Norwegian Prime Minister, Jens Stoltenberg, said in 2007 that “Norway will sharpen our emission cuts obligations under the Kyoto Protocol by 10 % in the period up to 2012” (Stoltenberg 2007). The British Prime Minister David Cameron said on 17th of May this year that “we will position the UK as a leading player in the global low-carbon

42 For a brief introduction to the Oslo-Potsdam solution see Hovi et al. 2003.

economy [...]" (Reuters 2011). These examples are illustrative of how governmental officers refer to their commitments to mitigate human-induced climate change.

On the other hand, there are many examples of how the actual dealing with human-induced climate change is not the top priority of the member countries. Canada's federal minister of environment, John Baird, announced in April 2007 that Canada no longer attempts to meet its emission reduction commitment as stated by the Kyoto Protocol. This decision was justified with the argument that compliance to the Kyoto targets would drive Canada into a deep economic recession (Böhringer and Rutherford 2009: 181). Under the CoP 16 in Cancún, 2011, the negotiations threatened to break down after the Japanese refusal to extend Kyoto emissions commitments without the active participation of the two biggest emitters, China and the US (Guardian 2011). These examples illustrates what Tony Blair said in 2005: "The blunt truth about the politics of climate change is that no country will want to sacrifice its economy in order to meet this challenge" (Guardian 2005).

What conclusions can be drawn from these empirical examples? Most important, there seems to be an imbalance between the actual domestic policies and international politics. The interaction between domestic and international politics has been addressed by Putnam (1988) in his well-known two-level game. He argued that international agreements are a result of a two-level game between domestic and international politics, where governmental officers seek to accommodate domestic and international demands. By analyzing the international and domestic level through a two-level game, one would be able to investigate the gap between rhetoric and behavior in the case of climate change.

6.3 Concluding Remarks on the ICCR

The road ahead for the ICCR is unclear and unknown: Should the Kyoto Protocol be extended into a new period of commitment? Or should a new agreement with more extensive coverage, including both developed and developing countries, be adopted? Will there be a new, binding international agreement at all after 2012? The answers are few, and the questions are many. The Copenhagen Conference in 2009 was intended as the deadline to resolving the remaining issues concerning the post-2012 climate change regime (Bodansky 2010). The Conference received massive attention; 120 Heads of State met in the negotiations, and about 35 000 people travelled to Copenhagen. As the Conference developed, the large gap between the anticipations of the industrialized- and the developing countries became obvious (Grubb 2010: 128). On the closing day of the Conference, President Obama summarized the Copenhagen experience with the following words:

The question, then, before us is no longer the nature of the challenge - the question is our capacity to meet it. For while the reality of climate change is not in doubt, I have to be honest, as the world watches us today, I think our ability to take collection action is in doubt right now, and it hangs in the balance.

Barack Obama, speech at the Copenhagen Convention, 2009[43]

The Copenhagen negotiations concluded without a binding agreement - casting doubt on the Parties' ability to take collective action against human-induced climate change. As Obama says, the future of a quantified, binding international agreement on emission reductions hangs in the balance as the first period of agreement is coming to an end.

However, the Copenhagen Conference did lead to an agreement known as the Copenhagen Accord. The Copenhagen Accord pledge that "the increase in global temperature should be below 2 degrees" (UNFCCC 2009). Many scholars have criticized the accord for being too weak (see Dimitrov 2010, Nordhaus 2010). The accuracy of this critique will not be further investigated here. However, I wish to turn the attention of the reader to the standard of evaluation of this accord. What is striking about the aim of the Copenhagen Accord, is that the standard of evaluation has changed in proportion to the Kyoto Protocol. To be able to illustrate this point, I will apply this example to the theoretical framework of this thesis. I have made use of the outcome indicator of regime effectiveness to evaluate the effectiveness of the ICCR. This indicator was chosen because it balances the validity and the availability of data. The dependent variable of the analysis therefore was CO_2 emissions. Change in average temperature as the dependent variable, such as the Copenhagen Accord suggests, is a study of change in nature, and therefore naturally belongs within the impact indicator of regime effectiveness. To change the ultimate goal of the ICCR from goals concerning emissions behavior to goals concerning temperature is questionable because of two reasons: First, the the availability of data for the impact indicator is problematic because of technical and scientific measurement problems, and the long decomposition time of CO_2 in the atmosphere. Thus, to move further down the chain of causality is not something that should be taken easily even though the validity of the indicator is higher. Second, the analysis in this thesis has shown that the members of the ICCR as a common unit have decreased their CO_2 emissions relative to the non-Annex I countries after the ratification of the Kyoto Protocol. The descriptive and the regression analysis both indicate this. However, most of these reductions are due to the economic collapse of the former Soviet Union, and not a consequence of the international regime. The ICCR is therefore not an effective regime

43 The entire speech can be found online (Obama 2009b)

because it has not altered the emissions behavior of its members. Thus, if the outcome indicator shows us that the ICCR is not effective, why should the impact indicator, which is further down in the chain of causality, tell us otherwise?

7. Literature

Achen, Christopher H. (2000) “Why Lagged Dependent Variables Can Suppress the Explanatory Power of Other Independent Variables,” paper presented at the Annual Meeting of Political Methodology, Los Angeles.

Adèr, Herman J. and Gideon J. Mellenbergh (1999) *Research Methodology in the Social, Behavioral & Life Sciences*, London: Sage Publications Ltd.

Aggarwal, Vinod K. (1983) “The Unraveling of the Multi-Fiber Arrangement, 1981: An Examination of International Regime Change” in *International Organization*, 37 (4), pp. 617-645.

Aggarwal S.C., and S.K. Khurana (2010) *Business Statistics*, New Delhi: FK Publications.

Allison, Paul D. (2002) *Missing Data*, London: Sage Publications Ltd.

Andresen, Steinar, and Shardul Agrawala (2002) “Leaders, Pushers and Laggards in the Making of the Climate Regime,” in *Global Environmental Change*, 12 (1), pp. 41-51.

Andersen, Steinar, Lars Walløe and Kristin Rosendal (2005) “The Precautionary Principle: Knowledge Counts but Power Decides?” in Rosie Cooney and Barney Dickson (eds.) *Biodiversity & the Precautionary Principle*, London: Earthscan, pp. 39-54.

Andersen, Steinar, Jon Birger Skjærseth, Jørgen Wettestad (1995) Regime, the State and Society: Analyzing the Implementation of International Environmental Commitments, Laxenburg, IIASA Working Paper WP-95-43, June.

Baker, Chris, Nancy Pistrang, and Robert Elliot (2002) *Research Methods in Clinical Pshychology. An Introduction for Students and Practitioners*, 2nd ed., Chichester: John Wiley & Sons, Ltd.

Barnwell, Anna (2011) “Multiple Measurement of International Regime Effectiveness: Comparative Study of the International Ozone Depletion Regime and Climate Change Regime” Master Thesis, NTNU.

Barret, Scott (2007) *Why Cooperate? The Incentive to Supply Global Public Goods*, Oxford: Oxford Univerity Press.

Barrington, Lowell, Micheal J. Bosia, and Kathleen Bruhn (2009) *Comparative Politics. Structures & Choices*, Boston: Wadsworth.

Beck, Nathaniel (2001) “Time-Series-Cross-Section Data: What Have We Learned in the Past Few Years” in *Annual Review of Political Science*, 4, pp. 271-293.

Beck, Nathaniel (2008) “Time-Series Cross-Section Methods” in Janet M. Box-Steffensmeier, Henry E. Brady and David Collier, *The Oxford Handbook of Political Methodology*, Oxford: Oxford University Press, pp. 475-493.

Beck, Nathaniel and Jonathan N. Katz (1995) “What to do (and not to do) with Time-Series Cross-Section Data” in *American Political Science Review*, 89 (3), pp. 634-647.

Beck, Nathaniel and Jonathan N. Katz (2011) “Modeling Dynamics in Time-Series—Cross-Section Political Economy Data” in *Annual Review of Political Science*, 14, pp. 331-352.

Bodansky, Daniel (2010) “The Copenhagen Climate Change Conference: A Postmortem” in *American Journal of International Law*, 104 (2), pp. 230-240.

Brady, Henry E. and David Collier (2004) *Rethinking Social Inquiry: Diverse Tools, shared Standards*, Lanham: Rowman & Littlefield Publishers.

Brannen, Julia (2004) “Working Qualitatively and Quantitatively” in Clive Seale, Giampietro Gobo, Jaber F. Gubrium and David Silverman, *Qualitative Research Practice*, pp. 312-327, London: Sage Publications.

Breitmeier, Helmut (1997) “International Organizations and the Creation of Environmental Regimes,” in Oran R. Young (ed.) *Global Governance: Drawing Insights from the Environmental Experience*, Cambridge: The MIT Press, pp.87-114.

Breitmeier, Helmut, Oran R. Young and Michael Zürn (2006) *Analyzing International*

Environmental Regimes. From Case Study to Database, Cambridge: The MIT Press.
Buttolph, Janet J. (2008) *Political Science Research Methods*, 6th edition, Washington: CQ Press.
Böhringer, Christoph (2000) "Cooling Down Hot Air: A Global CGE Analysis of Post-Kyoto Carbon Abatement Strategies" in *Energy Policy,* 28, pp. 779-789.
Böhringer, Christoph and Michael Finus (2005) "The Kyoto Protocol: Success or Failure?" in Dieter Helm (ed.), *Climate Change Policy,* pp. 253-281, Oxford: Oxford University Press.
Böhringer, Christoph and Thomas F. Rutherford (2009) "The Costs of Compliance: A CGE Assessment of Canada's Policy Options under the Kyoto Protocol" in *The World Economy,* 33 (2), pp. 177-211.
Böhringer, Christoph and Carsten Vogt (2003) "Economic and Environmental Impacts of the Kyoto Protocol" in *The Canadian Journal of Economics,* 36 (2), pp. 475-494.
Børsting, Georg, and Gunnar Fermann (1997) "Climate Change Turning Political: Conference-Diplomacy and Institution-Building to Rio and Beyond," in Gunnar Fermann (ed.) *International Politics of Climate Change. Key Issues and Critical Actors,* Oslo: Scandinavian University Press, pp. 53-82.
Campbell, Donald T. (1996) "Regression Artifacts in Time-Series and Longitudinal Data" in *Evaluation and Program Planning,* 19 (4), pp. 377-389.
Carson, Rachel (1996) "Silent Spring" in Matthew Alan Cahn and Rory O'Brien (eds.), *Thinking about the Environment. Readings on Politics, Property, and the Physical World,* New York: M.E.Sharpe, Inc, pp. 150-155.
Christiansen (2002) "New Renewable Energy Developments and the Climate Change Issue: a Case Study of Norwegian Politics" in *Energy Policy,* 30, pp. 235-243.
Cole, Matthew A. (2004) "Trade, the Pollution Haven Hypothesis and the Environmental Kuznets Curve: Examining the Linkages" in *Ecological Economics,* 48, pp. 71-81.
Cook, Thomas D. and Donald Thomas Campbell (1979) *Quasi-Experimentation: Design & Analysis Issues for Field Settings,* Boston: Houghton Mifflin Co.
Cowie, Jonathan (2007) *Climate Change: Biological and Human Aspects,* Cambridge: University Press.
Dai, Xinyuan (2010) "Global Regime and National Change" in *Climate Policy,* 10 (6), pp. 622- 637.
Day, Shane (2002) "A Proposed Methodology for Measuring the Effectiveness and Efficiency of Environmental Policy Regimes" in Patric Hayden, Bryan Hilliard, Tom Lansford and J.L. Walsh (eds.) *Debating Environmental Regimes,* New York: Nova Science Publishers, pp. 171-190.
Dessler, Andrew E. and Edward A. Parson (2006) *The Science and Politics of Global Climate Change*, Cambridge: Cambridge University Press.
De Boef, Suzanna and Luke Keele (2008) "Taking Time Seriously" in *American Journal of Political Science,* 52 (1), pp. 184-200.
Dietz, Thomas and Eugene A, Rosa (1994) "Rethinking the Environmental Impacts of Population, Affluence and Technology" in *Human Ecology Review,* 1, pp. 277-300.
Dimitrov, Radoslav S. (2003) "Knowledge, Power, and Interests in Environmental Regime Formation" in *International Studies Quarterly,* 47 (1), pp. 123-150.
Dimitrov, Radoslav S. (2010) "Inside Copenhagen: The State of Climate Governance" in *Global Environmental Policy,* 10 (2), pp. 18-24.
Dinda, Soumyananda (2004) "Environmental Kuznets Curve Hypothesis: A Survey" in *Ecological Economics,* 49, pp. 431-455.
Driscoll, John C. and Aart C. Kraay (1998) "Consistent Covariance Matrix Estimation with Spatial Dependent Panel Data" in *The Review of Economics and Statistics,* 80 (4), pp. 549-560.
Easton, David (1965) *A Systems of Analysis of Political Life,* New York: John Wiley & Sons.
Eckersley, Robyn (2007) "Ambushed: The Kyoto Protocol, the Bush Administration's Climate Policy and the Erosion of Legitimacy" in *International Politics,* 44, pp. 306-324.

Ehrlich, Paul R. and John P. Holdren (1971) “Impact of Population Growth” in *Science*, 171, pp. 1212- 1217.
Farrington, David P. (2003) “Methodological Quality Standards for Evaluation Research” in *Annals of the American Academy of Political and Social Science*, 58, pp. 49-68.
Fermann, Gunnar (1995) “Climate Change Policy-Making in Japan: Vulnerabilities and Opportunities” in *Current Politics and Economics of Japan*, 4 (2), pp. 103-128.
Fermann, Gunnar (1997) “The Requirement of Political Legitimacy: Burden-Sharing Criteria and Competing Conceptions of responsibility” in Gunnar Fermann (ed.), *International Politics of Climate Change. Key Issues and Critical Actors*, Oslo: Scandinavian University Press, pp. 11-52.
Frankel, Jeffrey A. and Andrew K. Rose (2005) “Is Trade Good or Bad for the Environment? Sorting out the Causality” in *Review of Economics and Statistics*, 87 (1), pp. 85-91.
Frees, Edward W. (2006) *Longitudinal and Panel Data: Analysis and Applications in the Social Sciences*, Cambridge: Cambridge University Press.
Gerrard, Michael B. (2007) *Global Climate Change and U.S. Law*, Chicago: American Bar Association.
Gravetter, Frederick J. and Lori-Ann B. Forzano (2008) *Research Methods for the Behavioral Sciences*, Belmont: Thomson Wadsworth.
Green, Donald P., Soo Yeon Kim and David H. Yoon (2001) “Dirty Pool” in *International Organization*, 55 (2), pp. 441-168.
Grubb, Michael (2010) “Copenhagen: Back to the Future?” in *Climate Policy*, 10, pp. 127-130.
Grubb, Michael, Christian Vrolijk, and Duncan Brack (1999) *The Kyoto Protocol. A Guide and Assessment*, London: The Royal Institute of International Affairs.
Guardian (2005) “Blair Signals Shift over Climate Change” downloaded 22.05.2011 from http://www.guardian.co.uk/environment/2005/nov/02/greenpolitics.frontpagenews
Guardian (2011) “Cancún Climate Change Summit: Japan Refuses to Extend Kyoto Protocol” downloaded 19.05.2011 from http://www.guardian.co.uk/environment/2010/dec/01/cancun-climate-change-summit-japan-kyoto.
Gunawansa, Asanga (2010) “The Kyoto Protocol and Beyond: A South Asian Perspective” in Koh Kheng-Lian, Lye Lin-Heng, and Jolene Lin (eds.) *Crucial Issues in Climate Change and the Kyoto Protocol. Asia and the World*, Singapore: World Scientific Publishing, pp. 474-505.
Gössling, Stefan (2011) *Carbon Management in Tourism. Mitigating the Impacts on Climate Change*, New York: Routledge.
Haas, Ernst B. (1982) “Words Can Hurt You; Or, Who Said what to Whom about Regimes” in *International Organization*, 36 (2), pp. 207-243.
Hackl, Franz and Gerald J. Pruckner (2001) “The Economics of the Kyoto Protocol,” in Paul J.J. Welfens (ed.), *Internationalization of the Economy and Environmental Policy Options*, Berlin: Springer, pp. 205-222.
Haggard, Stephen and Beth A. Simmons (1987) “Theories of International Regimes” in *International Organizations*, 41 (3), pp. 491-517.
Hamilton, Lawrence C. (1992) *Regression with Graphics. A Second Course in Applied Statistics*, Belmont: Duxbury Press.
Hamilton, Lawrence C. (2009) *Statistics with STATA*, 10th ed., Belmont: Brooks/Cole.
Hancké, Bob (2009) *Intelligent Research Design. A Guide for Beginning Researchers in the Social Sciences*, Oxford: Oxford University Press.
Hardin, Garrett (1968) “The Tragedy of the Commons” in *Science*, 162, pp. 1243-1248.
Hardin, Garrett (1998) “Extensions of “The Tragedy of the Commons”” in *Science*, 280 (1), pp. 682-683.
Harris, Paul G. (2007) “Collective Action on Climate Change: The Logic of Regime Failure” in *Natural Resources Journal*, 47 (1), pp. 195-224.

Hasenclever, Andreas, Peter Mayer, Volker Rittberger (1996) "Interests, Power, Knowledge: The Study of International Regimes" in *Mershon International Studies Review*, 40 (2), pp. 177-228.
Hasenclever, Andreas, Peter Mayer, Volker Rittberger (2000a) *Theories of International Regimes*, Cambridge: Cambridge University Press.
Hasenclever, Andreas, Peter Mayer and Volker Rittberger (2000b) "Integrating Theories of International Regimes" in *Review of International Studies*, 26 (1), pp. 3-33.
Helm, Carsten and Detlef F. Sprinz (2000) "Measuring the Effectiveness of International Environmental Regimes," in *Journal of Conflict Resolution*, 44 (5), pp. 630-652.
Henry, Laura A. and Lisa McIntosh Sundstom (2007) "Russia and the Kyoto Protocol: Seeking an Alignment of Interests and Image" in *Global Environmental Politics*, 7 (4), pp. 47-69.
Holland, Paul W. (1986) "Statistics and Causal Inference" in *Journal of the American Statistical Association*, 81 (396), pp. 945-960.
Holtz-Eakin, D. and T. M. Selden (1995) "Stoking the Fires? CO_2 Emissions and Economic Growth" in *Journal of Public Economics*, 57 (1), pp. 85-101.
Hovi, Jon, Tora Skodvin and Steinar Andresen (2003) "The Persistence of the Kyoto Protocol: Why Other Annex I Countries Move on Without the United States" in *Global Environmental Politics*, 3 (4), pp. 1-22.
Hovi, Jon, Detlef F. Sprinz, and Arild Underdal (2003) "The Oslo-Potsdam Solution to Measuring Regime Effectiveness: Critique, Response, and the Road Ahead," in *Global Environmental Politics*, 3 (3), pp. 74-96.
Hovi, Jon, Detlef F. Sprinz, and Arild Underdal (2009) "Implementing Long-Term Climate Policy: Time Inconsistency, Domestic Politics, and International Anarchy" in *Global Environmental Politics*, 9 (3), pp. 20-39.
Hsiao, Cheng (2003) *Analysis of Panel Data*, Cambridge: Cambridge University Press.
Huang, Wei Ming, Grace W. M. Lee, and Chih Cheng Wu (2008) "GHG Emissions, GDP Growth and the Kyoto Protocol: A Revisit of Environmental Kuznets Curve Hypothesis" in *Energy Policy*, 36, pp. 239-247.
International Energy Agency (2007) *Renewable Energy. Market & Policy Trends*, Paris: International Energy Agency.
International Energy Agency (2010) *CO_2 Emissions from Fuel Combustion. Highlights*, downloaded 20.03.2010 from http://www.iea.org/co2highlights/.
International Energy Agency (2011a) downloaded 01.03.2010 from http://www.iea.org/publications/free_new_Desc.asp?PUBS_ID=2143.
International Energy Agency (2011b) *Energy Statistics Division*, downloaded 24.03.2011 from http://www.iea.org/about/esd.asp.
International Monetary Found (2011) *Data Quality Assessment Framework*, 4th edition, downloaded 24.03.2011 from http://www.imf.org/external/np/sta/dsbb/2001/071001.pdf.
International Panel on Climate Change (IPCC) (2007a) "Summary for Policymakers", in *Climate Change: The Physical Science Basis*, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, New York: Cambridge University Press.
International Panel on Climate Change (IPCC) (2007b) *Climate Change 2007: Synthesis Report*, Geneva: Intergovernmental Panel on Climate Change.
Jackson, Sherri L. (2008) *Research Methods: A Modular Approach*, Belmont: Thomson Wadsworth.
Jacobson Harold K. and Edith Brown Weiss (2000) "A Framework for Analysis" in Edith Brown Weiss and Harold K. Jacobson (eds.), *Engaging Countries. Strengthening Compliance with International Environmental Accords*, Cambridge: The MIT Press, pp. 1-18.
Jakobsen, Henrik B. (2007) *Environmental Regime Effectiveness: A Case Study of the Kyoto Protocol*, master's thesis, Oslo: Universitetet i Oslo.

Kaplan, Abraham (2008) [1964] *The Conduct of Inquiry. Methodology for Behavioral Science*, 4th edition, New Brunswick, Transaction Publishers.
Kaul, Inge, Isabelle Grunberg and Marc A. Stern (1999a) "introduction" in Kaul Inge, Isabelle Grunberg and Marc A. Stern (ed.) *Global Public Goods. International Cooperation in the 21st Century,* New York: The United Nations Development Programme.
Kaul, Inge, Isabelle Grunberg and Marc A. Stern (1999b) "Defining Global Public Good" in Kaul Inge, Isabelle Grunberg and Marc A. Stern (ed.) *Global Public Goods. International Cooperation in the 21st Century,* New York: The United Nations Development Programme, pp. 2-19.
Keele, Luke and Nathan J. Kelly (2006) "Dynamic Models for Dynamic Theories: The Ins and Outs of Lagged Dependent Variables" in *Political Analysis,* 14 (2), pp. 186-205.
Kellow, Aynsley (1999) "Australia in the Greenhouse: Science, Norms and Interests in the Kyoto Protocol" in *Energy & Environment,* 10 (3), pp. 275-291.
Kennedy, Peter W. (2002) "Optimal Early Action on Greenhouse Gas Emissions" in *The Canadian Journal of Economics,* 35 (1), pp. 16-35.
Keohane, Robert O. (1982) "The Demand for International Regimes," in *International Organization,* 36 (2), pp. 325-355.
Keohane, Robert O. (1984) *After Hegemony: Cooperation and Discord in the World Political Economy,* Princeton: Princeton University Press.
Kindleberger, Charles P. (1986) [1973] *The World in Depression, 1929-1939,* Los Angeles: University of California Press.
King,Gary, Robert O. Keohane and Sidney Verba (1994) *Designing Social Inquiry. Scientific Inference in Qualitative Research,* Princeton: Princeton University Press.
Kittel, Bernhard (1999) "Sense and Sensitivity in Pooled Analysis of Political Data" in *European Journal of Political Research*, 35, pp. 225-253.
Kittel, Bernhard and Hannes Winner (2005) "How Reliable is Pooled Analysis in Political Economy? The Globalization-Welfare State Nexus Revisited" in *European Journal of Political Research*, 44 (2), pp. 269-293.
Krasner, Stephen D. (1982a) "Structural Causes and Regime Consequences: Regimes as Intervening Variables," in Stephen D. Krasner (ed.) *International Regimes,* Itahaca: Cornell University Press, pp. 1-22.
Krasner, Stephen D. (1982b) "Preface," in *International Regimes*, Stephen D. Krasner (ed.), Itahaca: Cornell University Press, pp. vii-ix.
Kuznets, Simon (1955) "Economic Growth and Income Inequality" in *American Economic Review,* 45 (1), pp. 1-28.
Kyoto Protocol (1997) *Kyoto Protocol to the United Nations Framework Convention on Climate Change.*
Lanchberry, John and David Victor (1995) "The Role of Science in the Global Climate Negotiations" in Helge Ole Bergesen, Georg Parmann, and Øystein B. Thommessen (eds.), *Green Globe Yearbook of International Co-operation on Environment and Development 1995*, Oxford: Oxford University Press.
Leamer, Edward E. (1985) "Sensitivity Analyses Would Help" in *American Economic Review*, 75, pp. 308-313.
Li, Quan and Rafael Reuveny (2006) "Democracy and Environmental Degradation" in *International Studies Quarterly*, 50 (1), pp. 935-956.
Lipsey, Mark W. (1990) *Design Sensitivity. Statistical Power for Experimental Research,* London: Sage Publications.
Levy, Marc A., Oran R. Young and Michael Zürn (1994) *The Study of International Regimes,* Working paper 113, Laxenburg: International Institute for Applied System Analysis.
Long, Scott J. and Laurie H. Ervin (2000) "Using Heteroscedasticity Consistent Standard Errors in

the Linear Regression Model" in *The American Statistician,* 54 (3), pp. 217-224.
March, James G. and Johan P. Olsen (1989) *Rediscovering Institutions. The Organizational Basis of Politics,* New York: The Free Press.
Meyerson, Frederick A. B. (1998) "Carbon Emissions, and Global Warming: The Forgotten Relationship at Kyoto" in *Population and Development Review,* 24 (1), pp. 115-130.
Michaelowa, Axel and Chris Rofle (2001) "Early Action to Reduce Greenhouse Gas Emissions Before the Commitment Period of the Kyoto Protocol: Advantages and Disadvantages" in *Environmental Management,* 28 (3), pp. 281-291.
Mitchell, Ronald B. (2003) "International Environmental Agreements: A Survey of Their Features, Formation and Effects" in *Annual Review of Environment and Resources*, 28, pp. 429-461.
Mitchell, Ronald B. (2004) "A Qualitative Approach to Evaluating International Environmental Regimes," in *Regime Consequences. Methodological Challenges and Research Strategies,* Arild Underdal and Oran R. Young (eds), Dordrecht, Kluwer Academic Publishers, pp. 121-149.
Mitchell, Ronald B. (2010) *International Politics and the Environment*, London: Sage Publications.
Mitchell, Ronald B. and Thomas Bernauer (2004) "Beyond Story-Telling: Designing Case Study Research in International Environmental Policy" in Detlef Sprinz and Yeal Wolinsky-Nahmias (eds.), *Models, Numbers, and Cases: Methods for Studying International Relations*, pp. 81-106.
Mohr, Lawrence B. (1995) *Impact Analysis for Program Evaluation*, 2nd edition, London: Sage Publications.
Morris, Scott B. (2007) "Estimating Effect Sizes From Pretest-Posttest-Control Group Designs" in *Organizational Research Methods,* 11 (2), pp. 364-386.
Neumayer, Eric (2001) "How Regime Theory and the Economic Theory of International Environmental Cooperation Can Learn from Each Other," in *Global Environmental Politics,* 1 (1), pp. 122-147.
Nordhaus, William D. (2010) "Economic Aspects of Global Warming in a Post-Copenhagen Environment" in *Proceedings of the National Academy of Sciences of the United States of America*, 107 (26), pp. 721-726.
Nordstrom, Hakan and Scott Vaughan (1999) "Trade and the Environment" Special Studies 4, *WTO.*
Obama, Barack (2009a) "Barack Obama's UN climate change speech" in *Times Online,* accessed 30.04.2011, at http://www.timesonline.co.uk/tol/news/environment/article6844525.ece.
Obama, Barack (2009b) "COP15 Remarks of President Barack Obama" accessed 26.05.2011, at http://www.copenhagenclimatecouncil.com/get-informed/news/cop15-remarks-of-president-barack-obama/printview.html.
Oberthür, Sebastian and Herman E. Ott (1999) *The Kyoto Protocol. International Climate Policy for the 21st Century*, Berlin: Springer.
Olson, Mancur (1965) *Logic of Collective Action. Public Goods and the Theory of Groups,* Cambridge: Harvard University Press.
O'Neill, Kate, Jörg Balsinger and Stacy D. VanDeveer (2004) "Actors, Norms and Impact: Recent International Cooperation Theory and the Influence of the Agent-Structure Debate," in *Annual Review of Political Science*, 7 (1), pp. 149-175.
Osherenko, Gail and Oran R. Young (1993a) *Polar Politics,* Ithaca: Cornell University Press.
Ostrom, Elinor (1995) "Constituting Social Capital and Collective Action" in Robert O. Keohane and Elinor Ostrom (eds) *Local Commons and Global Interdependence. Heterogeneity and Cooperation in Two Domains,* London: SAGE Publications, pp. 125-160.
Pan, Haoran and Denise Van Regemorter (2004) "The Costs and Benefits of Early Action Before Kyoto Compliance" in *Energy Policy*, 32, pp. 1477-1486.
Panayotou, Theodore (1997) "Demystifying the Environmental Kuznets Curve: Turning a Black Box into a Policy Tool" in *Environment and Development Economics,* 2, pp. 465-484.

Parks, Bradley C. and J. Timmons Roberts (2010) "Climate Change, Social Theory and Justice" in *Theory, Culture & Society*, 27, pp. 134-166.
Petersen, Trond (2004) "Analyzing Panel Data: Fixed- and Random-Effects Models" in Melissa Hardy and Alan Bryman (eds.), *Handbook of Data Analysis*, London: SAGE Publications, pp. 331-346.
Petersen, Mitchell A. (2007) "Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches", Working Paper.
Pettenger, Mary E. (2007) "Introduction: Power, Knowledge and the Social Construction of Climate Change" in Mary E. Pettinger (ed.) *The social Construction of Climate Change. Power. Knowledge, Norms, Discourses*, Hampshire: Ashgate Publishing, pp. 1-22.
Plümper, Thomas and Eric Neumayer (2010) "Model Specification in the Analysis of Spatial Dependence" in *European Journal of Political Research*, 49 (3), pp. 418-442.
Plümper, Thomas and Vera E. Troeger (2007) "Efficient Estimation of Time-Invariant and Rarely Changing Variables in Finite Sample Panel Analyses with Unit Fixed Effects" in *Political Analysis*, 15, pp. 124-139.
Plümper, Thomas, Vera E. Troeger and Philip Manow (2005) "Panel Data Analysis in comparative Politics: Linking Method to Theory" in *European Journal of Political Research*, 44 (2), pp. 327-354.
Poe, Steven C., Neal Tate and Linda Camp Keith (1999) "Repression of the Human Right to Personal Integrity Revisited: A Global Cross-National Study Covering the Years 1976-1993" in *International Studies Quarterly*, 43 (2), pp. 291-313.
Putnam, Robert D. (1988) "Diplomacy and Domestic Politics: The Logic of Two-Level Games" in *International Organization*, 42 (3), pp. 427-460.
Putnam, Robert D. (1992) *Making Democracy Work. Civic Traditions in Modern Italy.* Princeton, Princeton University Press.
Rabe-Hesketh, Sophia and Anders Skrondal (2005) *Multilevel and Longitudinal Modeling Using Stata*, Texas: Stata Press.
Rennenberg, H., R. Wassmann, H Papen and W Seiler (1995) "The Role of Methane and Nitrous Oxide in Climate Change" in Shaobing Peng, Keith T. Ingram, Heinz-Ulrich Neue and Lewis H. Ziska (eds.), *Climate Change and Rice*, Berlin: Springer, pp. 60-68.
Reuters (2011) "UK to Halve Greenhouse Gases by 2025, if EU follows" accessed at http://uk.reuters.com/article/2011/05/17/uk-britain-emissions-target-idUKTRE74G66Z20110517, 05.06.2011.
Ringquist, Evan J. and Tatiana Kostadinova (2005) "Evaluating the Effectiveness of International Environmental Agreements: The Case of the 1985 Helsinki Protocol" in *American Journal of Political Science*, 49 (1), pp. 86-102.
Rio Declaration on Environment and Development (1992) *Article 2,* United Nations Organization (1972) Report of the United Nations Conference of the Human Environment, Stockholm: United Nations Organization.
Rowlands, Ian H. (2001) "Classical Theories of International Relations," in Urs Luterbacher and Detlef F. Sprinz (ed.) *International Relations and Global Climate Change,* Cambridge: The MIT Press.
Schmalensee, Richard, Thomas M. Stoker, and Ruth A. Judson (1998) "World Carbon Dioxide Emissions: 1950-2050" in *The Review of Economics and Statistics*, 80 (1), pp. 15-27.
Sebenius, James K. (1991) "Designing Negotiations Toward a New Regime: The Case of Global Warming," in *International Security*, 15 (4), pp. 110-148.
Shadish, William R., Thomas D. Cook and Donald T. Campbell (2002) *Experimental and Quasi-Experimental Designs for Generalized Causal Interference,* Boston: Houghton Mifflin Company.

Shi, Anqing (2002) “The Impact of Population Pressure in Global Carbon Dioxide Emissions, 1975-1996: Evidence from Pooled Cross-Country Data” in *Ecological Economics*, 44, pp. 29-42.
Simmons, Beth A. (1998) “Compliance with International Agreements” in *Annual Review Political Science*, 1, 75-93.
Snidal, Duncan (1985) “The Limits of Hegemonic Stability Theory,” in *International Organization*, 39 (4), pp. 579-614.
Springer, Urs (2003) “The Market for Tradable GHG Permits Under the Kyoto Protocol: A Survey of Model Studies” in *Energy Economics*, 25, pp. 527-551.
Sprinz, Detlef and Tapani Vaahtoranta (2002) “National Self-Interest: a Major Factor in International Environmental Policy Formulation” in Mostafa K. Tolba (ed.) *Responding to Global Environmental Change*, Chichester: John Wiley & Sons, pp. 323-328.
Stern, David I. (1996) “Progress on the Environmental Kuznets Curve?” The Australian National University: Center for Resource and Environmental Studies Ecological Economics Programs.
Stern, David I. (2002) “Explaining Changes in Global Sulfur Emissions: an Econometric decomposition approach” in *Ecological Economics*, 42, pp. 201-220.
Stern, David I. (2004) “The Rise and Fall of the Environmental Kuznets Curve” in *World Development*, 32 (8), pp. 1419-1439.
Stern, Paul C., Oran Young and Daniel Druckman (1992) *Global Environmental Change. Understanding the Human Dimensions*, Washington: National Academy Press.
Stimson, James A. (1985) “Regression in Space and Time: A Statistical Essay” in *American Journal of Political Science*, 29 (4), pp. 914-947.
Strange, SUSn (1982) “Cave Hic Dragones: A Critique of Regime Analysis,” in *International Organization*, 36 (2), pp. 479-496.
Stoltenberg, Jens (2007) “Speech to the Congress of the Labour Party” *Office of the Prime Minister*, downloaded from http://www.regjeringen.no/en/dep/smk/Whats-new/Speeches-and-articles/statsministeren/statsminister_jens_stoltenberg/Speech-at-Trafalgar-Square-London/Speech-to-the-congress-of-the-Labour-Par.html?id=463749, 22.05.2011.
Suri, Vivek and Duane Chapman (1998) “Economic Growth, Trade and Energy: Implications for the Environmental Kuznets Curve” in *Ecological Economics*, 25 (1), pp. 195- 208.
Swanson, Timothy and Robin Mason (2003) “The Impact of International Environmental Agreements: The Case of the Montreal Protocol,” in *Environmental Policy in an International Perspective*, ed: Laura Marsiliani, Michael Raucher and Cees Wtihagen, Dorerecht: Kluwer Academic Publishers, pp. 51-81.
Tashakkori, Abbas and Charles Teddlie (1998) *Mixed Methodology. Combining Qualitative and Quantitative Approaches*, London: Sage Publications.
The Times (2007) “G8 Commits to Substantial Greenhouse Gas Cuts” accessed at http://www.timesonline.co.uk/tol/news/world/europe/article1898154.ece, 05.06.2011.
Underdal, Arild (2002) “One Question, Two Answers,” in *Environmental Regime Effectiveness: Confronting Theory With Evidence*, ed: Edward L. Miles, Arild Underdal, Steinar Andersen, Jørgen Wettestad, Jon Birger Skjærseth and Elaine M. Carlin, Cambridge: The MIT Press, pp. 3-45.
Underdal, Arild (2004) “Methodological Challenges in the Study of Regime Effectiveness” in *Regime Consequences. Methodological Challenges and Research Strategies*, ed: Arild Underdal and Oran R. Young, Dordrecht: Kluwer Academic Publishers, pp. 27-49.
United Nation Statistics Division (2011) *International Standard Industrial Classification of All Economic Activities*, downloaded from http://unstats.un.org/unsd/cr/registry/regcst.asp?Cl=2&Lg=1, 16.03.2011.

UNFCCC (1992) ”Article 2: Objective” in *United Nations Framework Convention on Climate change*.
UNFCCC (1995) “Report of the Conference of the Parties on its First Session, Held at Berlin From 28 March to 7 April 1995,” *FCCC/CP/1995/7/Add.1*.
UNFCCC (2009) “Copenhagen Accord,” *FCCC/CP/2009/L.7*.
UNFCCC (2011a) “Parties and Observers” *http://unfccc.int/parties_and_observers/items/2704.php, downloaded 26.05.2011, UNFCCC Secretariat, Bonn.*
UNFCCC (2011b) “Status of Ratification of the Kyoto Protocol” http://unfccc.int/kyoto_protocol/status_of_ratification/items/2613.php, UNFCCC Secretariat, Bonn.
Victor, David G. (2001) *The Collapse of the Kyoto Protocol and the Struggle to Slow Global Warming*, Princeton: Princeton University Press.
Vogler, John (2000) *The Global Commons. Environmental and Technological Governance*, second edition, Chichester: John Wiley & Sons Ltd.
Vogler, John (2003) “Taking Institutions Seriously: How Regime Analysis can be Relevant to Multilevel Environmental Governance” in *Global Environmental Politics*, 3 (2), pp. 25-39.
Waltz, Kenneth N. (1959) *Man, the State and War. A Theoretical Analysis*, New York: Colombia University Press.
Waltz, Kenneth (1979) *Theory of the International Politics*, New York: McGraw Hill Press.
Wendt, Alexander (1992) “Anarchy is what States Make of it: The Social Construction of Power Politics” in *International Organization*, 46 (2), pp. 391-425.
Wettestad, Jørgen (1999) *Designing Effective Environmental Regimes*, Cheltenham: Edward Elgar Publishing Limited.
Wettestad, Jørgen (2003) “The Vienna Convention and Montreal Protocol on Ozone-Layer Depletion” in *Environmental Regime Effectiveness. Confronting Theory with Evidence*, Edward L. Miles, Arild Underdal, Steinar Andersen, Jørgen Wettestad, Jon Birger Skjærseth and Elaine M. Carlin (eds.) Cambridge: The MIT Press, pp. 149-172.
Wilson, Sven E. and Daniel M. Butler (2004) “A Lot More to do: The Promise and Peril of Panel Data in Political Science,” Working Paper: Brigham University.
Wilson, Sven E. and David M. Butler (2007) “A Lot More to Do: The Sensitivity of Time-Series Cross-Section Analyses to Simple Alternative Specifications” in *Political Analysis*, 15, pp. 101-123.
World Bank (2009) *World Development Report. Reshaping Economic Geography*, Washington: The World Bank.
World Bank (2010) *World Development Indicators 2010*, Washington: The World Bank.
World Bank (2011) The World Development Indicators, downloaded from http://data.worldbank.org/data-catalog, 27.02.2010.
Zorn, Christopher (2006) “Comparing GEE and Robust Standard Errors for Conditionally Dependent Data” in *Political Research Quarterly*, 59 (3), pp. 329-341.
Yamin, Farhana and Joanna Depledge (2004) *The International Climate Change Regime. A Guide to Rules, Institutions and Procedures*, Cambridge: Cambridge University Press.
Young, Oran R. (1980) “International Regimes: Problems of Concept Formation” in *World Politics*, 32 (3), pp. 331-156.
Young, Oran R. (2003) “Determining Regime Effectiveness. A Commentary on the Oslo-Potsdam Solution” in *Global Environmental Politics*, 3 (3), pp. 97-104.
Young, Oran R. (2004) “The Consequences of International Regimes” in *Regime Consequences. Methodological Challenges and Research Strategies*, Arild Underdal and Oran R. Young (eds), Dordrecht: Kluwer Academic Publishers, pp. 3-24.

8. Appendix

Table 13: Alternative base years of EiTs (Oberthür and Ott 1999: 36).

Country	Base year
Bulgaria	1988
Hungary	average 1985-1987
Poland	1988
Romania	1989
Slovenia	1986

Figure 8: The trend of industrial output in the EiTs

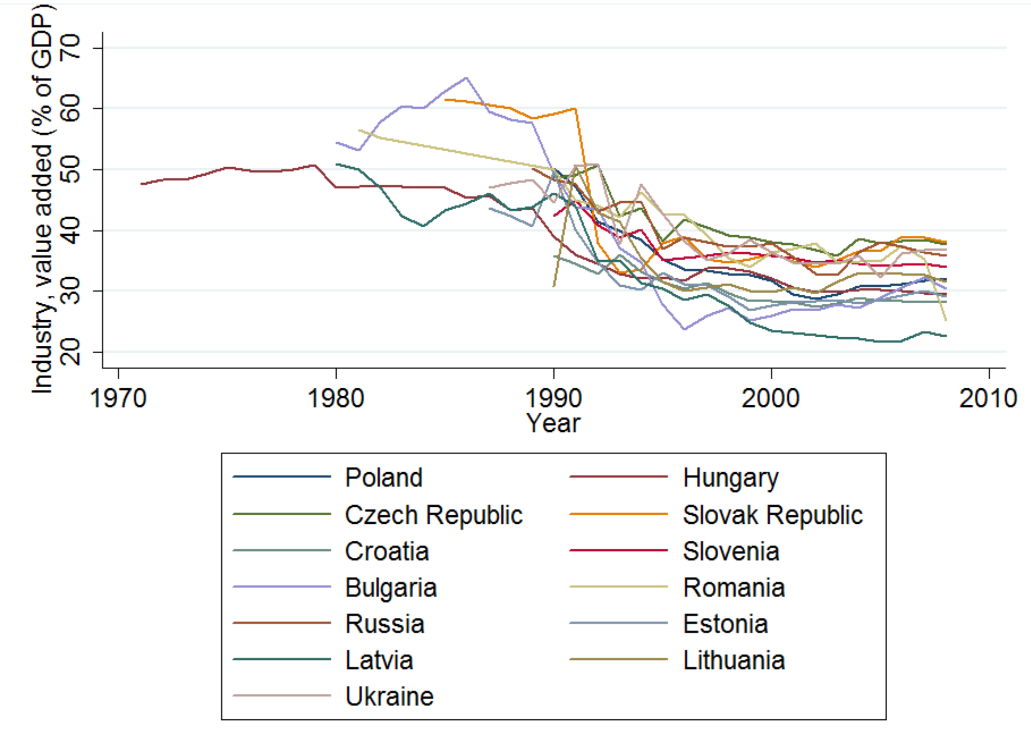


Table 14: Check for effect of variables

	Model (i) Baseline with diesel pump-price	Model (ii) Baseline without structure of economy	Model (iii) Baseline with interpolated dependent
Annex II[a]	2.651 (75.465)	107.986** (44.949)	93.205** (46.044)
EiT[a]	153.792** (64.464)	228.041*** (51.657)	221.074*** (51.642)
US[a]	4366.571*** (230.338)	3935.089*** (176.280)	3928.290*** (176.151)
After Kyoto	23.717** (10.338)	17.167** (7.202)	20.004** (8.255)
Annex II*After Kyoto	-37.985** (18.035)	-34.966** (16.300)	-44.789** (18.075)
EiT*After Kyoto	-50.459** (21.078)	-50.649** (20.534)	-57.608*** (22.079)
US*After Kyoto	179.602*** (64.650)	526.554*** (64.534)	522.493*** (68.455)
GDPPC	0.014*** (0.004)	0.010*** (0.002)	0.011*** (0.002)
GDPPC^2	-0.000** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)
Agriculture[b]	-1.088 (1.041)	-	-0.640 (0.505)
Industry[b]	-0.354 (0.761)	-	0.166 (0.369)
Population	30.219*** (1.193)	37.959*** (0.675)	37.414*** (0.718)
Trade	0.171 (0.197)	0.069 (0.104)	0.082 (0.118)
Diesel price	-4.995 (19.769)	-	-
Constant	-78.396	-124.497*** (19.961)	-121.770*** (29.086)
R^2	85.7	77.6	77.6
Countries	126	129	128
Observations	1734	4058	3718

Note: Standard errors in parentheses, * significant at 10 %, ** significant at 5%, *** significant at 1%
[a] Reference category is non-Annex I
[b] Reference category is services

Table 15: Sensitivity testing - Alternative specifications

	Model (iii) 1997-2008 as posttest	Model (iv) 1990-2008 as time-frame	Model (v) CO_2 per capita as dependent	Model (vi) Alternative independent variables
Annex II[a]	94.956* (53.197)	18.807 (63.236)	18.129** (8.1977)	-
EiT[a]	258.065*** (61.803)	194.899*** (59.813)	39.421*** (9.829)	-
US[a]	3733.481*** (208.920)	4243.946*** (211.308)	111.655*** (33.414)	-
Member[c]	-	-	-	122.453 (85.557)
After Kyoto	-3.921 (5.809)	21.739*** (7.844)	3.761*** (1.931)	22.581*** (7.832)
Annex II*After Kyoto	-14.349 (12.882)	-40.060** (17.526)	7.133*** (1.931)	-
EiT*After Kyoto	-78.135*** (16.657)	-60.270*** (20.118)	-15.600*** (2.357)	-
US*After Kyoto	696.789*** (43.537)	237.365*** (64.557)	14.338** (7.303)	-
Member*After Kyoto	-	-	-	-52.163*** (14.286)
GDPPC	0.012*** (0.002)	0.016*** (0.003)	0.007*** (0.000)	0.011*** (0.002)
GDPPC^2	-0.000 (0.000)	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)
Agriculture[b]	-0.994** (0.504)	-0.096 (0.643)	0.121** (0.056)	0.630 (0.497)
Industry[b]	-0.201 (0.369)	0.675 (0.512)	0.195*** (0.040)	0.150 (0.359)
Population	39.450*** (0.733)	33.243*** (1.105)	0.289*** (0.089)	45.749*** (0.815)
Trade	0.233** (0.115)	0.097 (0.158)	-0.009 (0.013)	-0.052 (0.117)
Constant	-120.473*** (31.120	-143.736*** (35.180)	6.908 (4.269)	-131.629*** (49.615)
R^2	77.0	83.9	51.5	33.4
Countries	128	128	128	128
Observations	3710	2246	3710	3710

Note: Standard errors in parentheses, * significant at 10 %, ** significant at 5%, *** significant at 1%
[a] Reference category is non-Annex I
[b] Reference category is services
[c] Reference category is non members

Table 16: Sensitivity testing - Influential cases

	Model (vii) Baseline with Germany and UK as dummies	Model (viii) Baseline with ln dependent	Model (?) Baseline without influential cases (Cooks D)	Model (?) Baseline without influential cases (Df-betas)
Annex II[a]	58.898 (45.225)	0.989*** (0.319)	98.407** (38.431)	102.138*** (37.412)
Germany[a]	620.775*** (165.177)	-	-	-
UK[a]	302.205* (165.225)	-	-	-
EiT[a]	219.471*** (48.948)	1.360*** (0.391)	201.956*** (45.789)	195.961*** (44.753)
US[a]	3949.74*** (166.819	4.105*** (1.328)	4071.938*** (155.708)	4084.769*** (152.176)
After Kyoto	19.882** (8.184)	0.241*** (0.033)	-4.227 (4.489)	-2.907 (4.199)
Annex II*After Kyoto	-32.516* (18.576)	-0.179*** (0.052)	-18.968* 9.789	-20.218** (9.119)
Germany*After Kyoto	-210.484*** (67.918)	-	-	-
UK*After Kyoto	-110.545 (67.396)	-	-	-
EiT*After Kyoto	-60.006*** (21.904)	-0.637*** (0.063)	-24.894** (11.946)	-25.983** (11.126)
US*After Kyoto	526.532*** (67.918)	-0.058 (0.196)	571.221*** (37.012)	571.101*** (34.470)
GDPPC	0.010*** (0.002)	0.000*** (0.000)	0.009*** (0.001)	0.008*** (0.001)
GDPPC^2	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)
Agriculture[b]	-0.922* (0.500)	-0.026*** (0.002)	-0.407 (0.281)	-0.393 (0.262)
Industry[b]	-0.038 (0.365)	0.002* (0.001)	-0.041 (0.203)	-0.103 (0.190)
Population	36.678*** (0.700)	0.037*** (0.002)	30.926*** (0.464)	30.574*** (0.471)
Trade	0.112 (0.117)	0.002*** (0.000)	-0.059 (0.066)	-0.049 (0.062)
Constant	-107.152*** (28.305)	2.384*** (0.153)	-75.066*** (20.504)	-72.266*** (19.708)
R^2	79.7	45.6	83.0	83.8
Countries	128	128	128	128
Observations	3710	3710	3703	3695

Note: Standard errors in parentheses, * significant at 10 %, ** significant at 5%, *** significant at 1%
[a] Reference category is non-Annex I
[b] Reference category is services

Figure 9: Outliers in the baseline model

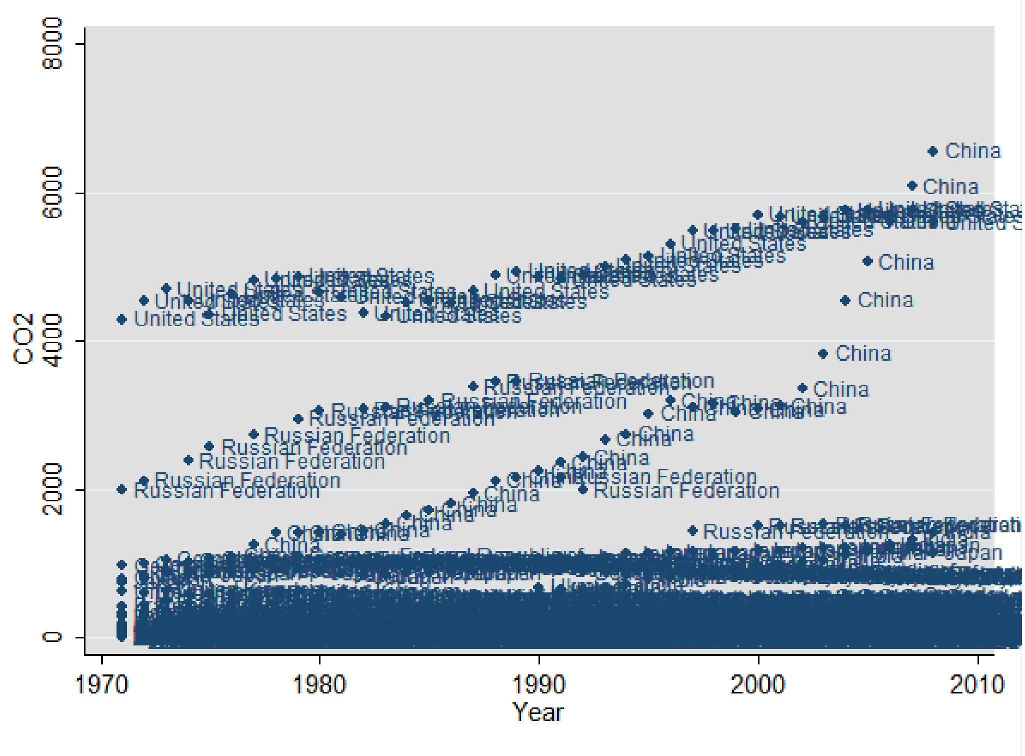


Figure 10: Cook's D for the baseline model

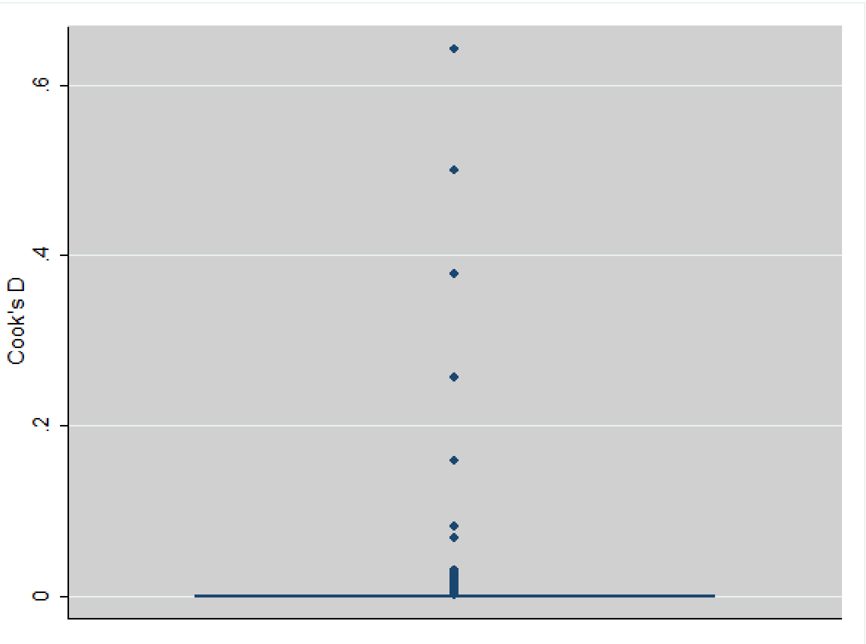


Figure 11: Df-betas for treatment variables for the baseline model

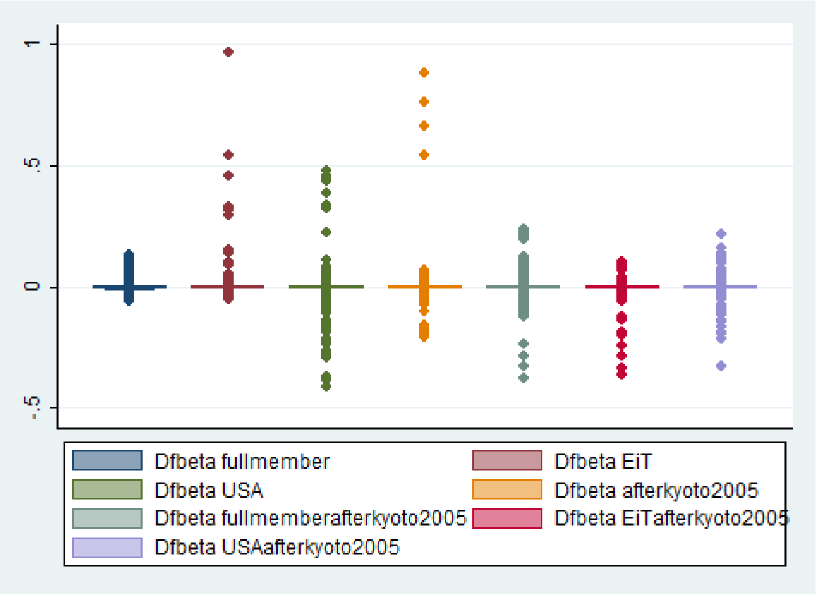


Figure 12: Df-beta for control variables for the baseline model

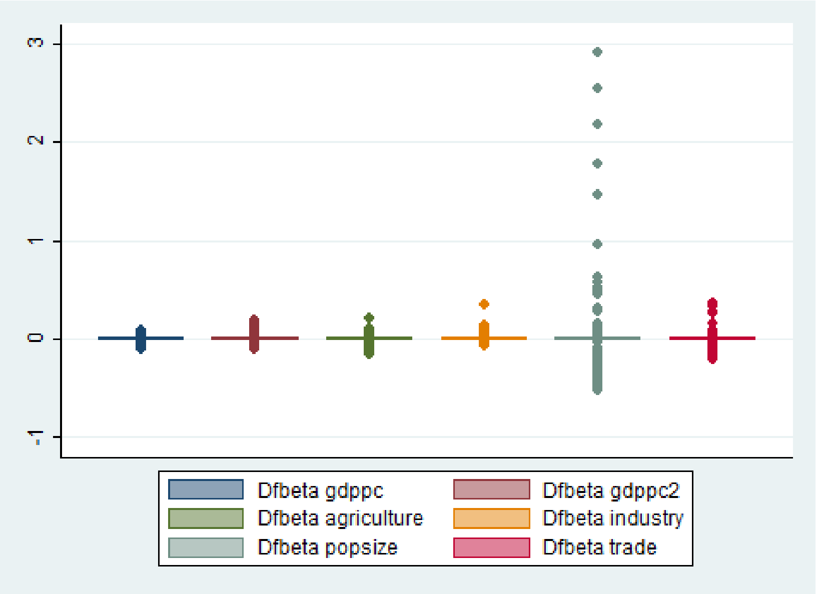


Table 17: Sensitivity testing - dynamics and heterogeneity

	Model (iix) Year dummies[c]	Model (ix) Lagged Dependent	Model (x) Fixed-effect
Annex II[a]	76.268 (46.752)	7.090 (7.139)	Omitted
EiT[a]	219.338** (51.055)	10.014 (6.958)	Omitted
US[a]	3904.882*** (174.721)	652.637*** (35.252)	Omitted
After Kyoto	1.164 (21.991)	11.765** (5.659)	15.212** (7.778)
Annex II*After Kyoto	-46.097*** (17.992)	-12.982 (12.718)	-37.239** (16.925)
EiT*After Kyoto	-62.937*** (22.133)	-15.760 (16.070)	-39.395* (20.651)
US*After Kyoto	521.387*** (68.011)	3.807 (50.484)	474.789*** (63.965)
GDPPC	0.011*** (0.002)	0.001 (0.001)	0.011*** (0.002)
GDPPC^2	-0.000 (0.000)	-0.000 (0.000)	-0.000*** (0.000)
Agriculture[b]	-1.316** (0.541)	-0.589*** (0.197)	0.898* (0.493)
Industry[b]	-0.228 (0.392)	0.231 (0.158)	0.269 (0.355)
Population	37.556*** (0.729)	3.704*** (0.165)	47.246*** (0.836)
Trade	0.131 (0.121)	-0.566 (0.039)	-0.018 (0.117)
Lagged dependent	-	0.853*** (0.007)	-
Constant	-84.745** (33.660)	-0.472 (9.094)	-127.078*** (23.542)
R^2	78.1	97.6	33.3
Countries	128	128	128
Observations	3710	3688	3710

Note: Standard errors in parentheses, * significant at 10 %, ** significant at 5%, *** significant at 1%
[a] Reference category is non-Annex I
[b] Reference category is services
[c] Year dummies not shown in table, 1971 as reference category

Table 18: Sensitivity testing - Robust standard errors

	(xi)
Annex II[a]	97.573 (65.941)
EiT[a]	222.705** (103.726)
US[a]	3929.431*** (293.184)
After Kyoto	19.251 (16.673)
Annex II*After Kyoto	-44.353 (31.485)
EiT*After Kyoto	-57.597* (30.030)
US*After Kyoto	521.107*** (83.900)
GDPPC	0.010** (0.004)
GDPPC^2	-0.000 (0.000)
Agriculture[b]	-0.810 (1.173)
Industry[b]	0.021 (0.620)
Population	37.624*** (13.755)
Trade	0.093 (0.274)
Constant	-113.781* (66.429)
R^2	78.0
Countries	128
Observations	3710

Note: Standard errors in parentheses, * significant at 10 %, ** significant at 5%, *** significant at 1%
[a] Reference category is non-Annex I
[b] Reference category is services

Table 19: Sensitivity testing - Test for multicollinearity

Variable	VIF
Annex II	3,64
EiT	1,3
US	1,48
After Kyoto	1,47
Annex II*After Kyoto	1,51
EiT*After Kyoto	1,39
US*After Kyoto	1,15
GDPPC	21,79
GDPPC^2	12,7
Agriculture	2,65
Industry	1,46
Population	1,1
Trade	1,41