# Deadlines and Commitments: The Effectiveness of and Compliance with Agreements on Air Pollution in Europe

Andreas Kokkvoll Tveit


PhD Dissertation
Department of Political Science
Faculty of Social Sciences

UNIVERSITY OF OSLO

*Series of dissertations submitted to the*
*Faculty of Social Sciences, University of Oslo*
*No. 754*

ISSN 1564-3991



Cover: Hanne Baadsgaard Utigard.
Print production: Reprosentralen, University of Oslo.

# Deadlines and Commitments: The Effectiveness of and Compliance with Agreements on Air Pollution in Europe

Andreas Kokkvoll Tveit

# Acknowledgements

I could not have written this dissertation without the support of a host of people.

First, I thank my supervisor, Jon Hovi, for generously sharing his razor-sharp thoughts in countless meetings over the last four-and-a-half years. Whenever I wanted to discuss anything related to my research (or, occasionally, football), I just walked down the hall and knocked on Jon's door. Having a top international scholar as an able and willing discussion partner has been invaluable.

Second, I had the pleasure of working with a truly exceptional group of PhD candidates at the Department of Political Science at the University of Oslo (UiO). Besides being talented researchers (the future of Norwegian political science is bright!), these young scholars are all genuinely good people who care about each other's professional development and general well-being. That is truly praiseworthy, considering that we are future competitors for the same academic positions. In particular, I thank Atle Haugsgjerd for being my primary go-to guy for writing sessions and discussions about academic and non-academic aspects of life since we started working on our PhD dissertations in 2014. Besides being a really good squash player and a mediocre weightlifter, Atle is an impressive political scientist and an even better friend. I would also like to thank Øyvind Bugge Solheim at the Institute for Social Research for being my brilliant friend and colleague since we started studying together in 2011. I thank Peter Egge Langsæther for all the tea breaks that lasted much longer than the intended ten minutes, and for accepting the role as my private STATA teacher without protest.

Third, I have benefitted greatly from the working conditions of the Department of Political Science and the Faculty of the Social Sciences at UiO. The value of working in the same department as so many impressive scholars can hardly be overstated. In particular, I thank Arild Underdal, Håkon Sælen, and Olav Schram Stokke for their useful comments to draft versions of the introductory chapter to this dissertation. Moreover, to spend research and teaching time efficiently, scholars rely on highly knowledgeable administrative staff to solve many major and minor administrative issues. For instance, having heard many parents complain about the incomprehensible application forms of the Norwegian Labour and Welfare Administration (NAV), I feared that it would take days or weeks just to fill out the

application for funding for my first paternity leave. Great was my relief when I discovered that the Faculty of Social Sciences has a team that helps fix such stuff! Additionally, generous funding from the Department of Political Science has enabled me to attend several conferences as well as to visit the International Institute for Applied Systems Analysis (IIASA) in Austria in May 2016.

Last, but certainly not least, I want to thank my family for all their care and support. I thank my mother for always having my back and being prepared to drop whatever she is doing in order to help me. Through his curiosity and eagerness always to learn more about the world we live in, my dad has been instrumental in my developing an interest in the social sciences. I thank Kirsten for being such an incredible wife, my best friend, and my most important discussion partner. Most importantly, however, I want to thank Kirsten for being such an outstanding mother for our sons, Olav and Thorvald. There is no denying that it has been challenging to be a parent while writing a PhD dissertation (or, more accurately, trying to write a PhD dissertation while parenting). Children always remind you, however, that there are much more important things in life than work. If I ever have a bad day in the office (perhaps because some reviewer just told me that my research stinks), it immediately turns good when I go into the kindergarten and a two-year-old runs towards me shouting "Pappaaaaaaaa!" My days get even better when I come home and find a baby smiling at me from his mother's lap. So, finally, I thank Olav and Thorvald for being the most wonderful distractions I could ever have.

Andreas Kokkvoll Tveit
Blindern, Oslo, Norway
February 2019

# Summary

Since the 1970s, a group of mainly European states have sought to reduce emissions of air pollutants through cooperation under the Convention on Long-range Transboundary Air Pollution (CLRTAP). Focusing on these efforts, this PhD dissertation aims to answer two overarching questions: 1) What explains states' (non)compliance with international agreements on air pollution? 2) How might states best design institutions and treaties targeted at reducing such pollution? While my theoretical point of departure is the two main theories in the international cooperation literature, the management school and the enforcement school, I also develop and test hypotheses derived from other theories and perspectives. I rely on a variety of mutually reinforcing analytical tools and data-collection methods to answer my research questions. One of my four articles is an in-depth case study; another uses a multi-methods approach; and two use statistical tools only.

My main findings may be summarized in five points.

First, the management school cannot explain much of the rather widespread noncompliance with CLRTAP protocols. My evidence lends no support to the propositions that treaty ambiguity and lack of state capacity explain noncompliance. However, unexpected changes between commitment and deadline may explain some of the noncompliance with the 1999 Gothenburg Protocol.

Second, neither does my evidence lend much support to the enforcement school's explanation of noncompliance. Specifically, my in-depth study of Norway's noncompliance with its 2010 target for nitrogen oxides ($NO_x$) under the Gothenburg Protocol shows that Norway from 2007 onwards implemented costly policies that did not maximize Norway's net private benefit.

Third, the explanatory power increases substantially if we allow some synthesizing of the management and enforcement schools. Building my case study of Norway's noncompliance with its $NO_x$ target under the Gothenburg Protocol, I develop the "deadline-pressure hypothesis," which can be seen as a hybrid of the management and enforcement schools. The evidence suggests that a logic of consequences was the more important driver of Norway's $NO_x$ policies as long as the protocol's deadline was distant; in contrast, a logic of

appropriateness grew stronger as the deadline was approaching.

Fourth, other perspectives contribute to explaining compliance with CLRTAP protocols. Arguing that the preferences of office-holding politicians determine policy strength, the "office-incumbent" hypothesis is consistent with much of the evidence from the case study of Norway's $NO_x$ noncompliance. It was not until an environmentalist party gained considerable influence over $NO_x$ policies that stricter $NO_x$ policies were enacted. Cross-national emissions data seem, however, to contradict the office-incumbent hypothesis: After 2005, $NO_x$ emissions reductions were even stronger in *all* the 10 other Northern and Western European Gothenburg parties than they were in Norway. Moreover, state capacity may explain (non)compliance with CLRTAP protocols, but not in the (positive) way hypothesized by managerialists. Using several measures of state capacity, I find a negative effect on compliance with five protocols under the CLRTAP. I argue that such a negative effect is less counterintuitive than it may seem; indeed, it may actually be expected among reluctant states that pursue policy goals correlating negatively with compliance.

Finally, the influence of European Union (EU) law on air pollution is substantially stronger than that of CLRTAP regulations. My assessments show that the 1988 Sofia Protocol had little, if any, impact on $NO_x$ emissions. In contrast, a series of EU directives caused considerable emissions reductions in EU member states in the post-agreement period (1989–1996). Because the EU relies on both managerial and enforcement measures to enhance the effectiveness of its legislation, this finding is consistent with the "hybrid" theory.

# Contents

# Publication status

Article 1 was published in *International Environmental Agreements: Politics, Law and Economics* (2018: 4). Article 2 was published in *Global Environmental Politics* (2018: 1). As of June 2019, neither article 3 nor article 4 have been published.

# Introductory Chapter: Background, Purpose, and Contributions

## Introduction and motivation

Despite recent decades' technological improvements and declining emissions, a range of air pollutants continue to impose large costs on local communities, countries, and whole regions. Since the late 1970s, a group of mainly European states has sought to alleviate these problems through cooperation under the Convention on Long-range Transboundary Air Pollution (CLRTAP).[1] Focusing on these emissions-reducing efforts, this dissertation's four articles aim to answer two overarching questions:

1. What explains states' (non)compliance with international agreements on air pollution?

2. How might states best design institutions and treaties targeted at reducing such pollution?

Being transboundary in nature, air pollution problems likely require international cooperation to be solved. The same is true for many other of today's most pressing environmental problems. Among numerous examples we find damage to commons such as the high seas, lakes, rivers, fish stocks, and most importantly, the global climate. However, evidence of suboptimal and sometimes entirely flawed international environmental cooperation abounds (Young 1989; Breitmeier et al. 2006; Hønneland et al. 2012; Underdal 2008). At the same time, despite substantial progress over recent decades (Young 2011), scholarly knowledge on the sources of (non)compliance with and the effectiveness of international environmental agreements (IEAs) is still well below the level required for providing decision makers with policy advice at the desired level of confidence.

[1] A few years later, the European Union (EU) started enacting potent legislation targeting the same environmental issues as the CLRTAP does.

This is not to say that international cooperation has been ignored by social scientists. Indeed, at least since the 1970s, many international regimes and institutions have been under intense scholarly scrutiny (Krasner 1985; Strange 1982). Initially, analysts were preoccupied with the formation of international cooperation.[2] After many IEAs had been in effect for some time, and their influence could be detectable, scholars began focusing on the degree to which states behaved in accordance with internationally agreed-upon rules. A motivation for these studies was the expectation that the degree of compliance could indicate the agreement's effect on state behavior. However, because this effect can be zero even with perfect compliance, researchers soon began looking for complementary ways to assess regime and treaty *effectiveness* (Mitchell 2008, 82–84; Victor et al. 1998b; Barrett 2008; Zürn 1998).

That researchers changed their primary focus in such a step-by-step process does not imply that they already knew everything there was to know about the subjects they abandoned. As my literature review below shows, a number of research gaps and limitations remain in the empirical IEA compliance-and-effectiveness literature. For now, it suffices to list only three. First, despite several large volumes having assessed various explanations of noncompliance, empirical knowledge of *state-level* causes of noncompliance remains limited. Second, bridging the main theories in the IEA literature will still require substantial theoretical and empirical work. Finally, even though many environmental problems are targeted by multiple international cooperative efforts, we have very limited knowledge of the relative effectiveness of such overlapping international institutions. For instance, a number of existing CLRTAP protocol studies have failed to consider that the same emissions have been targeted by parallel EU law.

Perhaps somewhat roughly put, scholarly debates on international environmental cooperation have been framed by the views of optimists and pessimists concerning states' ability to solve collective-action problems through cooperation. The former position is usually termed the management school (Chayes and Chayes 1993; Chayes, Chayes, and Mitchell 1995; Young

[2] For an account of the varying success in establishing regimes across environmental issues, see Young (1989).

1979). Although labels on the latter position vary,[3] the enforcement school is the most common (Downs et al. 1996). Claiming that states usually try to honor their international commitments, managerialists argue that the causes of noncompliance lie beyond the reach of states.[4] In contrast, scholars in the enforcement camp view lack of incentives as the main explanation of noncompliant behavior. I aim to contribute to IEA scholarship by empirically assessing explanations offered by these existing theories, and by developing and examining further theoretical conjectures or hypotheses concerning compliance and effectiveness. I also seek to contribute to a promising, yet so far limited IEA research effort: to bridge the management and enforcement schools by deriving and assessing synthesized hypotheses.

I focus on the efforts to solve three interrelated environmental problems under one convention – CLRTAP. Studying CLRTAP protocols has at least two major advantages. First, the five protocols I study specify deadlines and emissions targets for their parties. In contrast, the commitments made under many other regimes and agreements can hardly be quantified or measured.[5] Second, thanks to the substantial scientific collaboration underpinning CLRTAP cooperation, we have substantial knowledge about states' expectations of future conditions for compliance. Such projections enable assessments of a central, but understudied, managerial hypothesis – that noncompliance might be explained by unexpected developments from the time commitments are made (i.e., when agreements are adopted) to the agreed-upon deadline.

The remainder of this introductory chapter proceeds as follows. First, a background section presents air pollution and the (mainly European) efforts to abate it. It also provides some crucial definitions. Second, I present my research questions. Third, based on a review of the theoretical and empirical literature concerning the compliance with and the effectiveness of international (environmental) cooperation, I identify a number of research gaps. Fourth, I show how my methods and research design help answer my research questions. Fifth, I review each of my four articles and summarize the main findings of the dissertation as a whole. Finally, I offer

[3] For instance, Breitmeier et al. (2006) sometimes use the term "the incentive perspective."
[4] I elaborate further on these factors – treaty ambiguity, lack of state capacity, and unexpected changes between commitment and deadline – in my literature review below.
[5] For examples, see Breitmeier et al.'s (2006, 46–57) overview of the units of their comprehensive empirical assessments.

suggestions for future research on the compliance with and the effectiveness of international environmental agreements.

# Background and core concepts

## *Air pollution and its international solutions*

Although most human-induced air pollution is a consequence of burning fossil fuels and thus has been present since the beginning of industrialization, only in the 1970s did a consensus emerge that emissions such as sulfur caused harm far away from the immediate surroundings. This strong transboundary character of air pollution prompted a group of mainly European states to seek international solutions. The CLRTAP was adopted in 1979, institutionally localized under the United Nations Economic Commission for Europe (UNECE, or simply ECE). When the "first wave" of states entering the convention was over by the end of 1983, 25 states plus the European Union (EU)[6] had ratified, accepted, or approved the CLRTAP. As of 2018, the CLRTAP has 51 parties. Although air pollution is mainly a regional (not global) environmental problem – meaning that European states' emissions of air pollutants mainly affect other European states – the United States, Canada, and a handful of Asian states neighboring on or proximate to Europe (e.g., CLRTAP members such as Kazakhstan, Azerbaijan, and Georgia bordering on the European part of Russia) have taken part in the collaboration under CLRTAP (EMEP 2009).[7]

I study five CLRTAP protocols that seek to solve three interconnected environmental problems

[6] Before the EU was established by the Maastrict Treaty in 1993, it was known at the European Economic Community (EEC). However, for convenience's sake, I use only the present name, the European Union (EU).
[7] Several of these countries remain "fringe actors" in CLRTAP cooperation, and many have chosen not to ratify CLRTAP protocols. While the CLRTAP has 51 parties, the 1999 Gothenburg Protocol has only 27 (as of November 2018). In some cases, CLRTAP members have not even been assigned emissions targets. For instance, the original Gothenburg Protocol does not specify any targets for Canada and the United States (US). Wettestad (2012, 35) argues that "the underlying East-West détente encouraged the US and Canada to join the air regime as slightly exotic partners," thereby suggesting that environmental concerns were not the major reason these countries became CLRTAP parties.

by setting national emissions targets for four pollutants (Table 1). Sulfur and nitrogen oxides ($NO_x$) emissions *acidify* water and soil, thereby harming humans, animals, and vegetation. While combustion of fossil fuels in large industrial and power plants causes both sulfur and $NO_x$ emissions, road and sea transport also constitute main sources of $NO_x$. Important measures to reduce sulfur emissions include substituting low-sulfur coal for high-sulfur coal, substituting gas for coal, and traditional end-of-pipe abatement measures such as scrubbers on industrial smokestacks. The most effective emissions-reducing measures against $NO_x$ have been industrial end-of-pipe measures similar to those used against sulfur, as well as car catalysts and other measures cutting vehicle emissions. In combination with non-methane volatile organic compounds (NMVOC),[8] $NO_x$ causes *ground-level ozone*, which harms humans and other organisms. As in the case of $NO_x$, car catalysts have been among the most important emissions-reducing measures against NMVOC.

Table 1: Protocols and substances

| Protocol | Year of adoption | Deadline year(s) | Regulated substance(s) |
|---|---|---|---|
| Helsinki | 1985 | 1993 | Sulfur |
| Sofia | 1988 | 1994 | $NO_x$ |
| Geneva | 1991 | 1999 | NMVOC |
| Oslo | 1994 | 2000 (2005, 2010) | Sulfur |
| Gothenburg | 1999 | 2010 | Sulfur, $NO_x$, NMVOC, ammonia |

*Eutrophication* is a process in which the nutrition content of a water body becomes too high, leading to excessive plant and algal growth. Such growth further reduces the water's oxygen content, thereby changing the ecosystem to the benefit of dominant over vulnerable species. Unsurprisingly, agricultural fertilizers are among the most important causes of eutrophication.

[8] NMVOC comprises a collection of different compounds that stem from many sources (road transport vehicles being an important example) and display similar behavior after emission.

Ammonia emissions levels depend largely on fertilizer consumption and storage (Wagner et al. 2017; Norwegian Environmental Agency 2018).

Although its adoption was considered a major step forward to reduce air pollution, the CLRTAP entailed few obligations requiring emissions reductions. Instead, the national emissions targets outlined in a series of protocols are considered as the main regulatory elements under the CLRTAP (Di Primio 1998).[9] Table 1 provides an overview of the five CLRTAP protocols with specific national emissions targets, whose deadlines have already passed.[10]

Because sulfur was the first long-range air pollutant to catch the attention of natural scientists and political decision makers, the 1985 Helsinki Protocol included only sulfur emissions targets. By 1993, a 30 percent emissions reduction (base year 1980) was to be reached by all parties.[11] The 1988 Sofia Protocol followed suit, binding parties to stabilize emissions of $NO_x$ at their respective 1987 levels no later than 1994. The first national emissions targets for NMVOC were specified in the 1991 Geneva Protocol (deadline year 1999). Likely prompted by Helsinki's deadline passing the previous year (1993), a second sulfur protocol was adopted in Oslo in 1994, requiring parties to reach national emissions targets for 2000.[12] In 1999, the Gothenburg Protocol was adopted, which includes specified national emissions targets for sulfur, $NO_x$, NMVOC, and ammonia (deadline year 2010).[13] The rationale for negotiating such a multi-pollutant protocol was the interlinkage between the four pollutants and the problems they cause. For instance, considerable sulfur and $NO_x$ emissions stem from the same sources (i.e., power and industrial plants). Similarly, the transport sector is a major emitter of both $NO_x$ and NMVOC. Moreover, interactions between the pollutants, such as NMVOC forming ozone when reacting with $NO_x$,

[9] In addition to those included in Table 1, three more CLRTAP protocols exist: Indeed, the first CLRTAP protocol was the 1984 Protocol on Long-term Financing of the Cooperative Programme for Monitoring and Evaluation of the Long-range Transmission of Air Pollutants in Europe (EMEP). In 1998, two protocols, targeting heavy metals and persistent organic pollutants (POPs), respectively, were adopted. However, none of these protocols include national emissions targets similar to those of the five protocols shown in Table 1.
[10] The targets of these five protocols are the units of this dissertation's article 3. Articles 1 and 2 study compliance with the 1999 Gothenburg Protocol, and article 4 assesses the effectiveness of the 1988 Sofia Protocol.
[11] Although the term "sulfur emissions" in the CLRTAP context refers to several sulfur compounds, sulfur dioxide ($SO_2$) is usually used to express national emissions targets.
[12] For some parties, the Oslo Protocol also included national emissions targets for 2005 and 2010.
[13] An amended version of the Gothenburg Protocol was adopted in 2012.

underpin the joint regulation in one protocol.

Efforts to reduce emissions of long-range transboundary air pollutants are often touted as something of a success story. No doubt, as is evident from Figure 1, emissions have declined. However, the speed of the decline has varied widely across the targeted emissions.

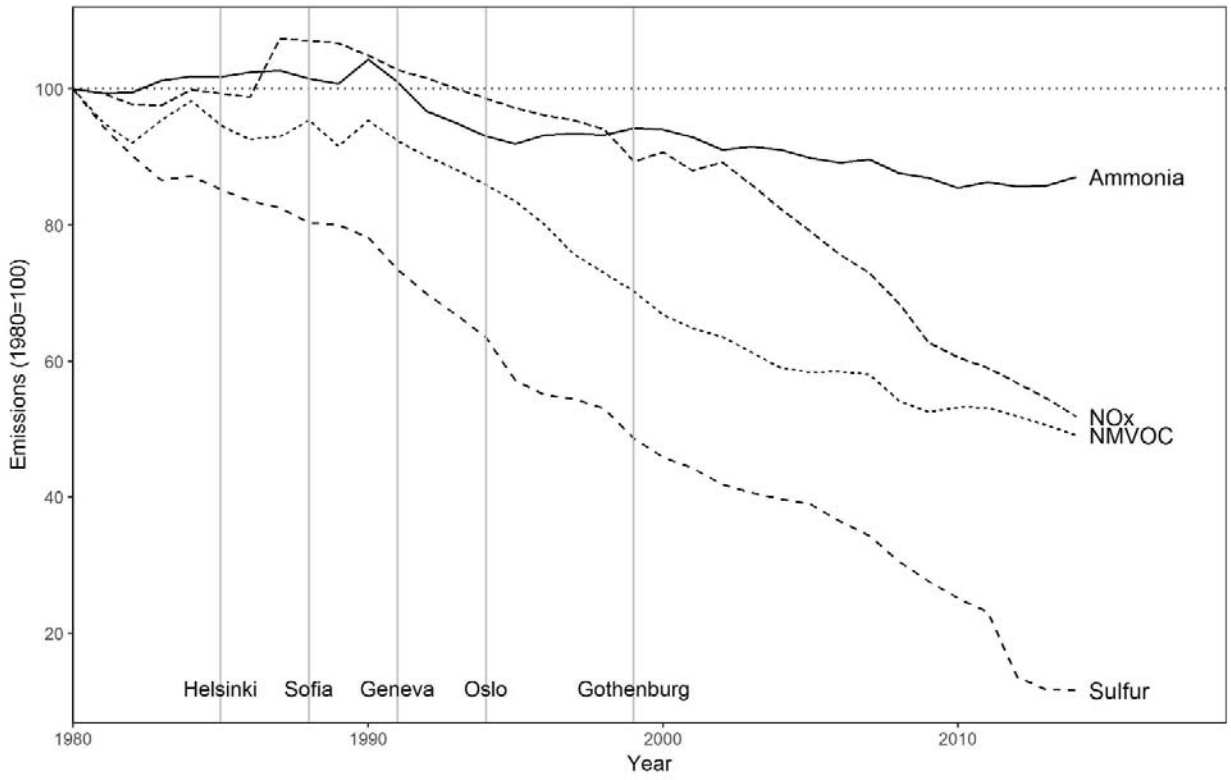


*Figure 1: Total sulfur, NMVOC, $NO_x$, and ammonia emissions of CLRTAP members 1980–2014 (as reported to the CLRTAP in 2017). Including only countries with complete emissions data for the entire period 1980–2014. The five vertical lines mark the adoption year of each protocol.*

Interestingly, at the beginning of the CLRTAP cooperation, parties saw sulfur emissions reductions as their most urgent task. Today, things are the other way around: Although the

amended version of the Gothenburg Protocol of 2012 includes sulfur targets (for 2020), sulfur is considered less troublesome than the other three pollutants regulated by Gothenburg. That comes as no surprise if we compare the total emissions trajectories shown in Figure 1.[14] Whereas sulfur emissions have declined steadily since 1980, it has taken decades to reduce $NO_x$ emissions substantially. As late as in 2002, total $NO_x$ emissions equaled roughly 90 percent of 1980 levels. NMVOC emissions were rather stable until 1990, when they started declining considerably. After a decade or so, the downward trend weakened somewhat, yet the total emissions have been more than halved since 1980. The emissions curve of ammonia is the flattest of the four, with 2014 emissions equaling almost 90 percent of 1980 levels.

To some extent, the relatively strong decline in sulfur emissions might be ascribed to their being regulated domestically and internationally for a longer time than the three other pollutants (Table 1). It is, however, beyond doubt that reducing sulfur emissions has been considerably simpler than reducing emissions of $NO_x$, NMVOC, and ammonia. Many of the sulfur emissions reductions Europe has witnessed since 1980 have structural explanations, and came rather cheaply (Rafaj et al. 2014; Wettestad 2012). In contrast, many potent measures against $NO_x$, NMVOC, and ammonia emissions have been so costly that implementation has made little sense to individual countries.

Thus, despite the observed emissions reductions, the environmental issues targeted by CLRTAP have yet to be solved. Monetizing various impacts of air pollution, and, even more so, impacts of *border-crossing* air pollution, is a challenging task. Estimates of total air pollution damage costs make it undoubtedly clear, however, that local communities, countries, and regions still suffer the damages caused by the emissions targeted by CLRTAP and other European policy initiatives. In the EU area alone, 2010 health-related costs of air pollution were estimated to be EUR 330 to 940 billion (Maas and Grennfelt 2016; European Commission 2013).[15]

[14] I end my time series in 2014 because national emissions data (as reported to the CLRTAP) often are incomplete for the 2–4 most recent years.

[15] All costs are in 2005 Euros. The substantial uncertainty of this estimate bears witness to the challenges associated with quantifying the costs of air pollution. The estimate includes damage costs of several pollutants, not only those targeted by the protocols I study.

## *Effectiveness, compliance, and depth of commitments*

Three of my articles seek to explain patterns of *compliance* – defined as the degree to which a state conforms to a provision in an international agreement (Börzel et al. 2010; Young 1979). The fourth article assesses the *effectiveness* of international cooperation, that is, the extent to which the agreement causes changes in the behavior of the targeted states (Victor et al. 1998b, 7).[16] Observed state behavior is compared to the hypothetical situation in which the agreement did not exist. This no-agreement counterfactual is here called business as usual (BAU). Agreements that are merely codifications of BAU are by definition *shallow*, while agreements requiring states to deviate from BAU are *deep*.

Implementation, compliance, and effectiveness constitute interlinked yet analytically distinct phenomena. Because compliance with deep commitments requires implementation of some policy, compliance might sometimes be seen as a product of implementation. Similarly, because effective agreements must prompt states to commission policies beyond BAU, effective international cooperation cannot exist without some form of policy being implemented. However, implementation is often studied as a process (Hill and Hupe 2014; Dai 2007, 15) without necessarily including any measures of compliance or effectiveness. Compliance, on the other hand, is often viewed as a necessary (yet not sufficient) condition for international cooperation to be effective: If compliance is high and commitments are deep, but self-selection or other mechanisms have made only a small fraction of the necessary countries participate, effectiveness will be low. Nor are broad participation and deep commitments alone enough to ensure effectiveness: If compliance with those commitments is low, effectiveness might be zero. Thus, treaty effectiveness has three necessary but individually insufficient conditions – high

[16] Effectiveness can also be measured as the impact on environmental outcomes. However, given that such outcomes have many other causes besides international cooperative efforts, my dependent variable in article 4 is state behavior, thereby reducing the length of the causal chain under assessment. For further discussions of definitions and aspects of effectiveness, see Mitchell (2002), Young (2011), and Young and Levy (1999).

compliance, broad participation, and deep commitments (Barrett 2008).[17,18]

# Research questions

This dissertation is organized around three research questions related to the compliance with and the effectiveness of CLRTAP protocols:

1 Can the main IEA compliance theories – the management and enforcement schools – account for (non)compliance with the emissions targets of CLRTAP protocols?

2 What alternative perspectives, if any, can provide better explanations than the management and enforcement schools do?

3 What is the impact of CLRTAP regulations on air pollution relative to that of EU law?

My dissertation's first article, *Can the Management School Explain Noncompliance with International Environmental Agreements?* focuses exclusively on research question 1. Specifically, it develops a framework for assessing the explanatory power of the management school's three explanations of noncompliance, and conducts a rigorous empirical test in the context of a well-suited case – the 1999 Gothenburg Protocol.

Article 2, entitled *Norms, Incentives, or Deadlines? Explaining Norway's Noncompliance with the Gothenburg Protocol*, also addresses research question 1. In contrast to article 1, it tests both the management school's and the enforcement school's explanatory power in an in-depth case study of one particular case of noncompliance with the 1999 Gothenburg Protocol. Because neither school can explain very much of Norway's noncompliance with its $NO_x$ target in

[17] This implies that IEA compliance studies somehow must consider depth analytically. Hence, such IEA studies differ from studies in another international compliance field: assessments of states' adherence to rulings of international courts (e.g., Grewal and Voeten 2015). Given that a dispute is admitted to a court, we might assume that any ruling in favor of one party will entail deviance from BAU by the other party. In contrast, when states themselves set the behavioral standard (as they do when treaties are negotiated in non-hierarchical forums), it is more likely that the behavioral standard (i.e., the treaty) is merely codifying BAU.
[18] Granted, if commitments are highly demanding (i.e., very deep), imperfect compliance is consistent with nonzero effectiveness, as long as (some) state behavior deviates from BAU.

Gothenburg, article 2 specifies and tests two additional hypotheses, concluding that the available data lend more support to what I call the "deadline-pressure" hypothesis than to the "office-incumbent" hypothesis.[19] Hence, article 2 also contributes to answering research question 2.

Article 3, *Does Capacity Increase Compliance? Examining Evidence from European Cooperation Against Air Pollution* addresses research questions 1 and 2. Using a data set including all emissions targets for parties to five CLRTAP protocols, I find that state capacity negatively affects compliance. To explain this finding, article 3 develops a novel conjecture of the capacity–compliance relationship, arguing that intention to comply is a crucial intervening variable whose effect has largely been ignored by previous IEA compliance research.

Finally, article 4, *The Relative Effectiveness of Overlapping International Institutions: EU versus UN Regulations of Air Pollution*, addresses research question 3. I demonstrate that the negative effect on $NO_x$ emissions ascribed by previous research to the Sofia Protocol should rather be ascribed to a series of EU directives targeting large combustion plants and road vehicle emissions.

# Literature review

Representatives of the main theories of international environmental cooperation, the management and enforcement schools, agree that international compliance is generally high. However, they sharply disagree on the roots of this high compliance. Claiming that states maximize net private benefits, enforcement scholars argue that states usually comply because international agreements tend to be shallow (Barrett 2008; Downs et al. 1996; Aakre et al. 2016). According to managerialists, however, states often negotiate treaties that require considerable deviations from BAU and sincerely try to act in accordance with their commitments. States have

[19] Because they describe some general principle(s) that can explain a phenomenon (state behavior toward international commitments), these hypotheses and their logical underpinnings could have been labeled theories. However, I reserve the term "theory" for even larger bodies of principles that are richer in empirical implications than the deadline-pressure and office-incumbent hypotheses are.

such a "propensity to comply" because their actions are guided by norms,[20] including the norm that commitments should be honored (Henkin 1968; Franck 1988; Chayes and Chayes 1993; Chayes, Chayes, and Mitchell 1995).

While concurring regarding the level of compliance, the two schools disagree concerning the level of effectiveness of international agreements. The enforcement school argues that shallowness makes many agreements ineffective. Moreover, should states commit deeply, the enforcement school would expect low compliance (absent potent enforcement).[21] The management school, in contrast, argues that unless compliance is inhibited by treaty ambiguity, lack of state capacity, or unexpected and uncontrollable changes, international agreements will usually be effective.[22]

Considerable energy has been devoted to empirical analyses of international environmental cooperation (among numerous examples, we find larger volumes such as Brown Weiss and Jacobson 1998; Victor et al. 1998a; Breitmeier et al. 2006; Miles et al. 2002; Young 1999).[23] Nonetheless, much work remains to be done concerning IEA compliance and effectiveness (Young 2011). In the next section, I outline the main research gaps motivating my dissertation.

[20] According to Chayes and Chayes (1993), the propensity to comply is also caused by considerations of efficiency (constant recalculation of one's own interest is inefficient, and decisions in international forums are costly) and interests (when sovereign states commit, the commitments likely align with state interests). Chayes and Chayes' (1993) emphasis on norms as drivers of state behavior suggests that they have much in common with constructivists such as Kratochwil (1984). Other scholars, for instance Keohane (1984), share Chayes and Chayes' (1993) rather optimistic view of the possibilities for solving international collective-action problems, but lean more towards the rationalist camp in the social sciences (see also Hasenclever et al. 1997).

[21] Interestingly, the claim that IEAs have little causal effect on behavior also comes from within the constructivist or cognitivist camp: As Underdal (2008) shows, scholars such as Ruggie (1982) and Conca (2006) distinguish between the "deep, normative structure of a system" and specific regimes or agreements. Because regimes or agreements are embedded in the normative structure, they should be seen as products of that structure, and not as entities with a causal influence of their own.

[22] Regime formation research also includes views similar to those of managerialist and enforcement scholars. The "pessimists" in regime-formation research argue that regimes were largely epiphenomenal, that is, mere reflections of the underlying power structures that would eventually determine outcomes anyway (see for instance Waltz 1979; Krasner 1985; Strange 1982; Mearsheimer 1994–1995). In contrast, the "optimists" argue that regimes are formed to solve common problems, and that power does not provide a satisfactory explanation of the formation, contents, and outcomes of regimes (Keohane 1984).

[23] While Brown Weiss and Jacobson (1998) study compliance and implementation, Victor et al. (1998a) study effectiveness and implementation, and Breitmeier et al. (2006) seek to explain patterns of compliance and effectiveness, Miles et al. (2002) and Young (1999) focus on regime effectiveness only.

## *Research gaps and the limitations of previous IEA studies*

In the remainder of this literature review, I first present previous studies of CLRTAP and its protocols, highlighting the research gaps in that CLRTAP literature. Then I discuss five more general research gaps or limitations of previous IEA compliance and effectiveness research. I do not suggest that previous research does not live up to scientific standards – quite the opposite. Most, if not all, of the gaps and limitations I identify might be explained by the fact that researchers are usually forced to make choices entailing trade-offs. For instance, I argue below that Breitmeier et al. (2006) teach us little about state-level causes of noncompliance. That is, however, an unavoidable consequence of their choice to focus on regime-level factors. Assessing the effects of state-level factors as well would have required considerable amounts of additional data. Assuming constant resources (which is more the rule than the exception in academic life), collecting such state-level data would have required them to spread their resources thinly, likely at the expense of data and research quality.

Similarly, many of the causal inferences drawn by Breitmeier et al. (2006) ultimately rely on their coders' ability to construct valid no-regime or no-agreement counterfactuals and to compare those counterfactuals to observed state behavior. Moreover, a common problem (whose magnitude is usually unknown) for researchers relying on the judgments of many different coders is that coding rules might not have been applied consistently. However, the sheer size of the International Regimes Database (IRD),[24] the database used by Breitmeier et al. (2006), suggests that these potential threats to validity and reliability were unavoidable.

Similarly, as I explain below, the contributions in the volume edited by Brown Weiss and Jacobson (1998) do little to discriminate between the relative explanatory power of the factors included in their analytical framework. A plausible explanation is their choice of a narrative research style, which is quite common in case studies. Surely, isolating the effects of various explanatory variables is by no means impossible in such studies (Goertz and Mahoney 2012). It

[24] The IRD includes data from 23 environmental regimes. Not only descriptives (such as the existence of sanctioning provisions or capacity-building mechanisms under the regimes or sub-regime units) are coded by experts; even causal assessments such as judging the regimes' impact on state behavior are left to the coders.

may, however, be highly challenging, because much information must be conveyed and analyzed *as text*. In contrast, quantitative studies analyze information that has been compressed into variable scores. In the case of Brown Weiss and Jacobson (1998), it might seem like some of the understanding of causal forces was lost in the thick description. Nonetheless, Brown Weiss and Jacobson (1998) deserve much credit for the considerable amount of data collected by their case researchers.

*Research gap 1: Several studies have focused on CLRTAP protocols; however, much data remains unassessed*

First, my dissertation is motivated by the fact that much empirical research remains to be done in the literature concerning CLRTAP protocols. Note, however, that several aspects of CLRTAP cooperation have received substantial scholarly attention (see for instance Zürn 1998). Among several compliance studies, we find Underdal and Hanf's (2000) assessment of nine countries' compliance with the first two CLRTAP protocols,[25] as well as Wettestad's (1998) analysis of three states' compliance with and implementation of the same protocols. Moreover, a series of qualitative and quantitative studies have assessed CLRTAP protocols' contributions to environmental problem solving and their effect on state behavior (examples include Levy 1993; Wettestad 2002; Ringquist and Kostadinova 2005).

These studies notwithstanding, a large untapped potential for empirical research exists in the CLRTAP context. First, the scholarly interest in CLRTAP and its protocols was considerably stronger in the 1990s than it has been since then. Hence, much data from the last couple of decades have yet to be analyzed. Second, because many previous contributions are case studies of one or a handful of countries' or of one or more CLRTAP protocols, much data from the early days of CLRTAP cooperation also remain to be analyzed. I contribute to filling these gaps in several ways, including (1) by publishing the first study of noncompliance with a particular

[25] Underdal and Hanf (2000) also assess parties' policies in relation to the third protocol (Geneva). However, their research was carried out prior to the 1999 deadline, thereby limiting the opportunities for measuring compliance.

target under the 1999 Gothenburg Protocol (article 2) and (2) by providing the first assessment of the compliance of *all* parties to the five protocols included in Table 1 (article 3).

*Research gap 2: Limited empirical knowledge of state-level causes of (non)compliance*

Second, my dissertation is motivated by the limited empirical[26] knowledge on state-level IEA noncompliance explanations – such as state capacity. The main aim of the compliance study with the largest N thus far, Breitmeier et al. (2006), was to assess *regime-level* causes of (non)compliance, rather than state-level causes. Because that volume's units of analysis are regimes, the authors' data set includes far more regime-level information than information on individual states.[27] Moreover, certain features of the study by Brown Weiss and Jacobson (1998), the most comprehensive existing assessment of state-level factors and compliance, inhibit clear and general conclusions. Their introductory chapter presents a 30-variable analytic framework which is to be used in the subsequent case-study chapters. When the editors' summary lists "the most important factors that affect compliance," *all* those 30 variables are included (Jacobson and Brown Weiss 1998, 536). Hence, except from stating that "the strength and health of national political-economic systems and a deep public commitment are the most important ingredients in compliance," Jacobson and Brown Weiss (1998, 542) provide few take-home messages for readers seeking a ranking of the factors most conducive to compliance.

Also limiting our empirical knowledge on state-level sources of IEA noncompliance is the fact that readers have difficulties finding measurements of Brown Weiss and Jacobson's (1998) explanatory variables. Neither the editors' summary nor the individual chapters include a transparent overview of the observational units' scores on the explanatory variables.[28] Brown

[26] Theoretical advances have, however, been made by scholars such as Dai (2005; 2007), who shows how international institutions may increase compliance through domestic political actors.
[27] Note, however, that some state-level information is collected. For instance, coders are asked to identify individual states that were particularly important actors under the regime, and then to code those states as pushers, laggards, or neutrals (Breitmeier et al. 2006, 262). Nonetheless, IRD coders did not systematically gather information on most potentially important state-level variables (e.g., capacity).
[28] In contrast, Miles et al. (2002, 487–491) include an appendix showing the variable scores of all units. I do not suggest that qualitative researchers should let spreadsheets replace nuance; however, variable score overviews

Weiss and Jacobson (1998) provide their readers with much information related to variables such as capacity, but not with the observational units' scores on those variables.

What lessons concerning state-level causes of noncompliance can be drawn from other larger empirical IEA studies? The main explanatory variables considered by Miles et al. (2002) are on the regime level,[29] and their outcome variable of interest is regime effectiveness. Similarly, Victor et al. (1998a) study implementation and effectiveness, and assess other explanatory factors than those specified by the enforcement and management schools.

Although several stand-alone articles consider state-level explanations of IEA (non)compliance, some of these articles are mostly theoretical. For example, Dai (2005) acknowledges (already in the abstract) that the article's empirical contents (drawn from the case of the 1985 Helsinki Protocol) serve only "illustrative" purposes. Likewise, when arguing that soft law might sometimes produce better outcomes than hard law does, Abbott and Snidal (2000, 424) draw on a number of examples, also from the environmental field. However, the authors acknowledge that "these examples do not provide a true empirical test of our arguments."

Hence, my dissertation seeks to provide empirical knowledge on state-level causes of noncompliance by empirically assessing various state-level factors' ability to explain the fairly widespread noncompliance with CLRTAP protocols (articles 1–3).[30]

offer readers some leverage when trying to keep track of authors' causal assessments. Conversely, in studies without such overviews, much understanding of causal forces might be lost in the thick description.

[29] Note, however, that some state-level factors are part of Miles et al.'s (2002) models, for instance the distribution of power among the participants in a regime.

[30] My assessment of the relative effectiveness of the Sofia Protocol and EU law also includes measures of state-level variables and controls for unobserved country-fixed factors.

*Research gap 3: Existing studies' inconclusiveness concerning the effect of capacity*

Third, as I show in article 3's literature review, existing empirical research is strikingly inconclusive concerning the effect of state capacity. For instance, Breitmeier et al. (2006, 111) find that regimes that include horizontal sanctioning mechanisms outperform regimes relying only on capacity building. Hence, they argue that "[our data] do not confirm expectations about the role of capacity building." In contrast, Jacobson and Brown Weiss (1998) claim that state capacity is important for compliance. These diverging findings suggests a need to provide more robust conclusions about the effects of capacity, for instance by specifying scope conditions under which one should expect capacity to influence compliance positively. This is exactly what I do in articles 1–3.

*Research gap 4: The need for testing the three "Chayesian" hypotheses*

Fourth, my work on compliance is inspired by the fact that a proper assessment of Chayes and Chayes' (1993) ability to explain a *given instance* of noncompliance requires testing three hypotheses simultaneously. Few, if any, previous studies have done so. Chayes and Chayes (1993) argue that noncompliance is caused by ambiguity, lack of state capacity, *or* unexpected changes between commitment and deadline. Therefore, demonstrating that *one* of those explanations fails to account for a given case of noncompliance does not suffice to show that Chayes and Chayes' (1993) theory fails to explain that case. For instance, even if a given breach of obligations cannot be explained by ambiguity, it might be explained by lack of capacity or by unanticipated change. This observation has important policy implications. If one of the management school's three explanations is able to account for a given instance of noncompliance, that instance could likely have been addressed by managerial measures (such as monitoring or capacity building). In contrast, if none of the management school's explanations can account for a given instance of noncompliance, the management strategy would likely fail in that instance. Hence, my article 2 aims at testing all three explanations' ability to account for one particular instance of noncompliance.

*Research gap 5: The need to assess the relative effects of overlapping international institutions*

Fifth, while some qualitative studies consider the effects of EU law targeting the same emissions as CLRTAP protocols do (see for instance Levy 1993; Wettestad 2002), none of the previous econometric studies (Ringquist and Kostadinova 2005; Aakvik and Tjøtta 2011; Vollenweider 2013; Bratberg et al. 2005) do so. Given the potency of the EU regulatory apparatus, EU law should indeed be taken into account by such medium-N or large-N effectiveness assessments. However, assessing the relative effect of EU law and CLRTAP protocols targeting the same emissions requires overcoming two challenges that might have hampered empirical progress in this field. First, such assessments face the general difficulties related to assessing international agreements' impact on state behavior (see for instance Hovi et al. 2003; Young 2003; Mitchell 2008; Vollenweider 2013). Second, given that outcomes such as emission levels are often over-determined, isolating the effects of multiple efforts to solve the same problem is notoriously difficult. Nonetheless, this dissertation's article 4 aims to overcome these challenges in order to compare the impact of EU law on $NO_x$ emissions to that of the 1988 Sofia Protocol. Because scholarly knowledge on such relative effectiveness of overlapping international institutions remains scarce, studies such as my article 4 are badly needed.

*Research gap 6: How can management and enforcement be merged?*

Finally, several scholars have argued that the management and enforcement schools may be viewed as complementary rather than antithetical (Tallberg 2002; Underdal 2008). Notwithstanding some notable efforts, such as Börzel et al. (2010),[31] much theoretical and empirical work remains to realize the full potential of this cross-fertilization. Underdal (2008, 68) points specifically at one such challenge that remains unsolved: translating "into specific form the vague proposition that actor behavior is driven by some combination of the logic of consequence and the logic of appropriateness." This is exactly what I attempt in my article 2,

[31] Inter alia, Börzel et al. (2010) test the effects of state capacity contingent on state power and argue why such interactions should sometimes be expected.

where I propose the deadline-pressure hypothesis, a hybrid of the management and enforcement schools. Moreover, I show that the deadline-pressure hypothesis helps explain the substantial tightening of Norwegian $NO_x$ policies from 2007 onwards – a policy shift that is difficult to attribute to changed incentives or changed norms. Finally, my article 4 adds to the theoretical underpinnings of a "hybrid" strategy for designing IEAs as well as to our empirical knowledge of its effectiveness.

# Research strategy

I rely on a variety of mutually reinforcing analytical tools and data-collection methods to answer my research questions. I combine cross-case and within-case analysis to make the most of the analytical power the data provide.

Research question 1 – whether the main IEA compliance theories can explain the noncompliance I observe – entails a number of sub-questions. Answering them requires various methods. Assessing the validity of Chayes and Chayes' (1993) ambiguity explanation requires knowledge of states' interpretations of the contents of the agreement. Such knowledge may be acquired by examining public documents, or by asking core actors how they view their commitments following accession to the agreement. If parties have mutually consistent interpretations of the agreement, ambiguity is unlikely to have caused noncompliance. Hence, using interview data and public documents from various countries, I am able to assess whether ambiguity constituted a barrier to compliance with the emissions targets in the 1999 Gothenburg Protocol.

Variations in N between my articles suggest that I assess the explanatory power of capacity – a second sub-question under research question 1 – by examining different kinds of evidence. In my medium-N analyses,[32] I assess the effect of capacity on compliance using OLS and logistic regression analysis. If Chayes and Chayes (1993) are right when they claim that capacity increases compliance, measures of state capacity should, after control for confounding variables, display a positive and statistically significant effect on compliance in my cross-sectional data. In

[32] These analyses are presented in articles 1 and 3. The N is 92 and 176, respectively.

particular, my statistical analyses enable me to estimate the effect of capacity *controlled for the ambition level* of the target (i.e., the size of the required emissions reductions).

Examining the explanatory power of Chayes and Chayes' (1993) third factor – unexpected social and economic changes between the adoption and deadline of commitments – is data-intensive. Ideally, it requires knowing how challenging (or how simple) states thought reaching compliance would be when making their commitments. In the case of the 1999 Gothenburg Protocol, we would want to know what states thought in 1999 about the prospects for reaching the emissions targets by 2010. Thanks to the extensive cooperative research efforts under the CLRTAP, a number of projections of the development of factors crucial to compliance were available to states prior to committing to the protocol's targets. Thus, by comparing the projections of those factors – for instance energy consumption – with the observed developments, articles 1 and 2 are able to assess the explanatory power of Chayes and Chayes' third explanation. For instance, if the 2010 energy consumption of noncompliant Gothenburg participants proved significantly higher than projected back in 1999, that change might explain noncompliance. Determining whether such unexpected changes can explain all or just some of the noncompliance requires that I quantify both the size of the change (typically measured in emitted tons of the substance) and the size of the noncompliance.

While article 1 develops a framework for assessing Chayes and Chayes' three explanations in the context of a medium-N data set, article 2 builds on and develops this framework. It does so in an in-depth case study of one instance of noncompliance with the Gothenburg Protocol: Norway's failure to reach its $NO_x$ target by 2010. While sacrificing empirical breadth, such a single-case study evidently has several strengths compared to the medium-N analyses of article 1. For instance, I assess several aspects of Norway's environmental state capacity that differ from the operationalizations of state capacity in article 1. Examples include my examination of public documents over time, which reveals that Norwegian authorities' knowledge of technological solutions to $NO_x$ emissions has been adequate for decades.

Much of the analytical insight gained from my case study of Norway relies on a series of

interviews with top bureaucrats and politicians. For instance, interview data substantially strengthen my basis for concluding that capacity was not a barrier to compliance in Norway's case. My strategy for assessing the capacity explanation was to find out what level of capacity was needed to comply and to compare what was needed with the capacity Norway actually had. That strategy sufficed to conclude that financial resources were not a barrier to compliance (full compliance would require only a very small fraction of one year's government spending on environmental issues). Assessment of other aspects of capacity was, however, more challenging. How many skilled bureaucrats were needed to ensure compliance with Norway's $NO_x$ target? Here, my interview data proved crucial. If factors such as bureaucratic resources were in short supply, it would likely have been pointed out by my interviewees. Despite having ample opportunities to do so, none of my interviewees said anything even remotely suggesting that lack of bureaucratic resources or competence caused Norway's noncompliance.[33]

Research question 1 also encompasses examining if the enforcement school can account for the noncompliance I observe in my data. Arguably, such investigations are challenging – and definitely more so than assessments of managerial hypotheses (such as asking if capacity sufficed for compliance). However, given the core views of the enforcement school, we may investigate whether state behavior conforms with the dictates of maximizing net private benefits. If we observe such a consistency between private interests and behavior, the enforcement school can be said to help explain (non)compliance and other behavioral outcomes. That is particularly true if such interests–behavior consistency occurs even though some norm (such as pacta sunt servanda – that agreements must be kept) dictates behaving differently.

If my interview data were important to some of my conclusions concerning the management school's explanatory power, they were outright indispensable to my assessment of the enforcement school. Consider my claim that Norway commissioned a potent $NO_x$ policy package from 2007 onwards despite not having any (material) incentives to do so. If I had to rely only on publicly available information (in news media or documents from authorities), it would

[33] The former Minister of the Environment and other politicians did so despite having good reasons to suggest that lack of bureaucratic resources was a problem, which would have allowed them to avoid some blame for the noncompliance.

be harder to rule out the possibility that Norwegian authorities anticipated sanctions or other reactions from the EU: The inclusion of the EU's National Emissions Ceilings (NEC) Directive in the EEA Agreement in 2009 made Norway's $NO_x$ emissions target of 156,000 tons binding. Because NEC's inclusion in the EEA perhaps could have been anticipated, one might hypothesize that Norwegian decision makers enacted strong $NO_x$ policies to avoid future reactions from the EU.[34] However, my interviews consistently suggest otherwise: Independently of each other, several interviewees revealed that political decision makers were not even aware of the NEC Directive, much less that it could lead to Norway's suffering punitive actions. In short, interviews allowed me to examine relevant processes much more closely than any other available data did, and my assessment of the enforcement school's hypothesis could hardly have been conducted using other methods.

The methods used to research additional explanations[35] – my research question 2 – are similar to those used to assess managerial and enforcement hypotheses. Having found that neither of the main theories (in their original forms) provide good explanations of Norway's $NO_x$ policies, I inductively developed two additional hypotheses from the data already collected. Thereafter, I collected further evidence in order to conduct a rigorous assessment of the novel hypotheses, labeled the office-incumbent and the deadline-pressure hypotheses. Together, interview, document, and emissions data seem to lend more support to the latter hypothesis.

In article 3, a series of OLS and logistic regressions consistently suggest that Chayes and Chayes' (1993) capacity explanation is unable to account for member countries' noncompliance with CLRTAP protocols. Interestingly, the empirical assessments leading to that conclusion contribute to explaining noncompliance with CLRTAP protocols, but not in the way expected by managerialists: I find a negative effect of capacity on compliance. Developing a novel conjecture concerning the capacity–compliance relationship, I argue that we should not be surprised that high state capacity may in some instances lead to noncompliance. The key is to understand the

[34] Indeed, that hypothesis was suggested in one of the reviewer reports secured by *Global Environmental Politics*, the journal that published my article 2.
[35] That is, hypotheses different from those I derive from the enforcement and management schools in their original forms.

circumstances under which capacity increases compliance, and the circumstances under which it has the opposite effect. Narrative evidence from one case of noncompliance supplements my statistical empirical findings and novel conjecture. By showing how Norway's high capacity likely aided its success as a petroleum extractor,[36] I support my argument that state capacity may have a negative effect on compliance among reluctant states that pursue policy goals detrimental to international compliance.

Finally, I assess the relative effectiveness of UN and EU law targeting air pollution (article 4) using OLS regression analysis. Specifically, deploying the difference-in-differences (DID) estimator, I use pre- and post-agreement emissions of three groups of states that are all CLRTAP members. First, I compare the emissions trends of non-parties and parties to the 1988 Sofia Protocol, finding no pre-agreement differences between the two groups. Second, because some but far from all Sofia parties are also members of the EU, I assess the post-1988 emissions trends of three groups: non-parties to Sofia, Sofia parties that are not members of the EU, and EU members. Only in the third group – the EU members – do post-1988 emissions trends differ from pre-1988 trends. Because the EU adopted a series of directives targeting combustion plant and road vehicle emissions from 1988 onwards, I ascribe the observed emissions slope change to these directives. Hence, I argue that EU law likely was far more effective in targeting $NO_x$ emissions than Sofia was. I also offer an explanatory typology that accounts for the observed pattern of effectiveness. I show that regimes targeting groups of states that vary in capacity and cooperative intent are more likely to succeed if they combine soft "managerial" and hard "enforcement" measures instead of relying on either type of measure alone.

Endogeneity poses analytical challenges to my compliance (articles 1–3) and effectiveness (article 4) assessments. Under a decentralized international order, states first decide what commitments to include in a treaty, and then whether to become a party. Hence, scholars commonly – and rightly – warn against interpreting considerable behavioral change among treaty participants as caused by the treaty: If treaty participants already had considerable changes (such as emissions reductions) underway, the treatment effect of participation might be zero. In my

[36] Extracting oil and gas is emissions intensive, given its reliance on energy produced by gas turbines at the extraction sites.

effectiveness study (article 4), I address this challenge by using the DID estimator. Given that certain conditions hold, DID studies reveal effects that are likely causal (Angrist and Pischke 2009).

Concerning my studies of compliance, endogeneity poses the most serious challenge to my statistical assessments of the effects of state capacity (articles 1 and 3). If high-capacity states accept deeper emissions targets than low-capacity states do (i.e., if capacity is positively correlated with depth), depth might be an intervening variable between capacity and compliance. Indeed, if that correlation is strong, it may account for the negative effect of capacity on compliance. I aim to address this challenge by including the "ambition-level" variable, because doing so takes into account the (relative) size of the required emissions reductions. Nonetheless, even my ambition-level variable might raise endogeneity concerns; hence, target depth might remain unobserved in my models. It may thus be hard to entirely escape the suspicion that the negative effect of state capacity found in papers 1 and 3 might be due to omitted variable bias. Although such a cautionary note is necessary, my ambition-level variable seems to constitute an innovative (if not first-best) solution to endogeneity issues raised in the IEA compliance literature. Moreover, Chayes and Chayes (1993; 1995) do not expect correlation between state capacity and depth to entail such substantial noncompliance among high-capacity states that my analyses reveal.

# Article summaries

## *Article 1: Can the Management School Explain Noncompliance with International Environmental Agreements?*

The dissertation's first article starts by establishing that despite the considerable attention the works of Chayes and Chayes (1993) have received, their three explanations of noncompliance remain understudied. In particular, previous studies have paid scant attention to their third explanation – unexpected changes between commitment and deadline. Hence, this article develops a framework for assessing the explanatory power of the factors specified by Chayes and Chayes and uses that framework in the context of a case well suited for the task: the 1999 Gothenburg Protocol. A careful reading shows that the language of the protocol is clear and unambiguous; indeed, there has been no disagreement over the treaty's content. Hence, Chayes and Chayes' ambiguity explanation is dismissed in the case of Gothenburg. A series of OLS and logistic regressions reveal that the effect of capacity on compliance is negative (not positive, as Chayes and Chayes would expect). This result holds under a number of conditions; thus, it is highly robust. Therefore, I also dismiss the capacity explanation. Finally, I show that parties had adequate time to meet their obligations, and unexpected developments explain only a small part of the observed noncompliance. In only two of 21 instances of noncompliance, unexpectedly high energy consumption may fully explain the gap between observed 2010 emissions and the emissions target for that year. Previous underestimations of emissions can also explain some of the observed noncompliance. In several instances of noncompliance, however, unexpected developments between commitment and deadline made the emissions targets *easier* to reach. Taking all the assessed developments into account, I conclude that unexpected changes explain some noncompliance, but leave most of it unaccounted for. Hence, albeit a considerably better explanation than the first two, Chayes and Chayes' (1993) "temporal dimension" provides far from a sufficient account of the noncompliance with the 1999 Gothenburg protocol. These findings pose a serious challenge to Chayes and Chayes' three explanations of noncompliance – at least as far as the Gothenburg Protocol is concerned.

## *Article 2: Norms, Incentives, or Deadlines? Explaining Norway's Noncompliance with the Gothenburg Protocol*

My second article expands the framework developed in article 1 by deploying it in an in-depth case study: Norway's failure to reach its $NO_x$ emissions target under the 1999 Gothenburg Protocol. This second article is partly motivated by the negative findings from article 1. If Chayes and Chayes (1993) cannot account for Gothenburg noncompliance, what theories can? As in article 1, little evidence suggests that the three explanations posed by the management school can account for Norway's noncompliance. Norwegian authorities' perceptions of Norway's obligations are consistent with those of other parties, and my interview data decisively reject that any significant actors have been in doubt about Norway's obligation to reach the emissions targets included in the protocol. Similarly, I show that Norway's issue-specific capacity to reduce $NO_x$ emissions was high and sufficient to reach compliance by 2010. Finally, although some unexpected changes of conditions for compliance made the target harder to reach, I argue that Chayes and Chayes' (1993) third explanation cannot explain much of Norway's noncompliance. Indeed, my interviews with top politicians and bureaucrats suggest that such changes were not important barriers to compliance.

Article 2 also examines the enforcement school's ability to account for Norway's $NO_x$ noncompliance. Surely, the main outcome – Norway's noncompliance with its $NO_x$ target – is consistent with the enforcement school. However, a closer assessment reveals that the enforcement school provides at best an insufficient explanation in the Norwegian case. Albeit too late to reach compliance by the 2010 deadline, a $NO_x$ tax was commissioned in 2007. And even though no enforcement mechanisms were in place, the ensuing emissions reductions were clearly deeper than in a BAU scenario. Since both of the main theories in the IEA compliance field thus fail to (fully) explain Norway's noncompliance, I derive and assess two additional hypotheses. Some evidence supports an office-incumbent hypothesis. The $NO_x$ tax was not introduced until after the 2005 elections, when an environmentalist party gained considerable influence over $NO_x$

policies. Hence, the sudden and considerable change from lax to strong $NO_x$ policies may be due to environmentalists replacing business-oriented politicians in office. However, the fact that $NO_x$ emissions declined substantially across Northern and Western Europe after 2005, and that several other Gothenburg parties achieved even larger emissions reductions than Norway did, suggests that the explanation for the policy shift is structural rather than particular. One such structural explanation is the deadline-pressure hypothesis: As the 2010 deadline neared, most parties considered action to cut emissions to be more urgent than before.

## *Article 3: Does Capacity Increase Compliance? Examining Evidence from European Cooperation Against Air Pollution*

In article 3, I conduct analyses similar to article 1's assessment of Chayes and Chayes' capacity explanation. The empirical scope is, however, much broader: Article 3 assesses the effect of capacity on compliance with all five CLRTAP protocols that include national emissions targets for sulfur, $NO_x$, NMVOC, and/or ammonia. My data set thus covers 31 states and three decades, thereby including all the emissions targets of the international society's most mature and institutionalized international effort to solve transboundary air pollution. As article 1 does, article 3 finds a negative effect of capacity on compliance. Crucially, this finding holds when I control for how ambitious the targets were, that is, for the size of the required emissions reductions. My empirical findings add to a previously unexplained inconsistency of previous studies concerning the effect of capacity. Using that inconsistency as a backdrop, I argue that intention to comply constitutes a crucial intervening variable, whose effect has largely been overlooked by previous compliance research. I argue that among reluctant[37] states that satisfy two criteria, state capacity may have a negative effect on compliance. First, the state must pursue one or more policy goal(s) that correlate negatively with compliance, that is, the more the state succeeds in realizing that policy goal, the lower the compliance. Emissions-intensive economic growth may be such a policy goal conflicting with CLRTAP protocol compliance. Second, the

[37] As opposed to *enthusiastic* states (Victor 2011)

realization of that policy goal must be facilitated by state capacity.

Using Norway's noncompliance with the $NO_x$ target in the Gothenburg Protocol to illustrate my novel conjecture, I show that Norway's $NO_x$ policies were weak or non-existent in a majority of the years between Gothenburg's adoption (1999) and its deadline (2010). Simultaneously, Norway pursued a strategy aiming at being an efficient petroleum exporter. Because extraction of oil and gas is emissions intensive, the goal of maximizing economic growth through petroleum exports conflicted with Gothenburg compliance. Norway's ability to spur economic growth through petroleum extraction was likely aided by its high state capacity. Hence, I argue that its high capacity may well have been a contributing factor to Norway's noncompliance with its $NO_x$ target in the Gothenburg Protocol.

## *Article 4: The Relative Effectiveness of Overlapping International Institutions: EU versus UN Regulations of Air Pollution*

Article 4 assesses how effective the 1988 Sofia Protocol has been compared to a series of EU directives targeting $NO_x$ emissions from large combustion plants and road vehicles. It thereby contributes to a thin scholarly field: Even though problems requiring international cooperation are often targeted by multiple cooperative efforts, we have little knowledge of the relative effectiveness of such overlapping agreements, regimes, or organizations. Using the DID estimator, I find that the emissions-reducing effect ascribed by previous research to the Sofia Protocol should rather be attributed to EU law. While the Sofia Protocol had little (if any) impact on $NO_x$ emissions, the EU directives seem to have had a substantially and statistically significant effect. Specifically, EU members reduced their $NO_x$ emissions 3.9 percent more per year than they would have in the counterfactual scenario. Taking into account that decoupling $NO_x$ emissions from economic activities has proven to entail substantial costs,[38] a 3.9 percent annual

[38] As I show in article 2, analyses from 1998 and 1999 suggested that compliance with the Gothenburg $NO_x$ target would cost Norway NOK 200–300 million annually (i.e., compliance implied substantial deviation from the expected emissions trajectory, and that deviation would require additional emissions-reducing measures with total annual costs of NOK 200–300 million.)

deviation from BAU is considerable. That is even truer if we take into account that air pollution is not an imminent threat to society (in contrast to climate forcers such as $CO_2$), and that cooperation on $NO_x$ pollution was in its early days when this impressive emissions reduction took place.

Developing a novel explanatory typology, I provide a theoretical account of this observed pattern of effectiveness. I show that when an international regime or agreement targets states with varying capacities and intentions to cooperate, it should influence the behavior of three types of states: enthusiastic low-capacity states, reluctant low-capacity states, and reluctant high-capacity states. While strictly managerial or enforcement strategies will influence the behavior of only one state type each (enthusiastic low-capacity and reluctant high-capacity states, respectively), a combination of managerial and enforcement measures affects the behavior of all three types. The empirical results in article 4 support this hybrid theory concerning the compliance with and effectiveness of international regimes. Thus, in a group of states where capacities and intentions to cooperate vary, regimes that combine soft (managerial) and hard (enforcement) measures – as the EU does – are more likely to succeed than are regimes relying on only one type of measure – as Sofia does.

# Main findings

The main findings of this thesis may be summarized in five points.

First, the management school cannot explain much of the rather widespread noncompliance with CLRTAP protocols. The explanatory power of Chayes and Chayes' (1993) three factors varies, however. While my evidence lends no support to treaty ambiguity and lack of state capacity, the explanatory power of Chayes and Chayes' (1993) third factor (unexpected social or economic changes between commitment and deadline) is somewhat higher. Nonetheless, most of the observed noncompliance with CLRTAP protocols is left unaccounted for by the management school.

Second, my compliance assessments also lend little support to the enforcement school. Because Norway's $NO_x$ policies from 2007 onwards in fact deviated substantially from BAU, the enforcement school cannot explain very much of Norway's Gothenburg noncompliance, or of Norway's $NO_x$ policies as a whole. As article 2 demonstrates, assessing if state behavior conforms with a BAU scenario requires large amounts of data.[39] Thus, this dissertation's testing of the enforcement school's ability to account for noncompliance is confined to the Norwegian case analyzed in article 2. In other words, my compliance assessments focus far more on the management school than on the enforcement school. Readers should therefore note that although managerial explanations of noncompliance receive little support in articles 1–3, this does not in any way suggest that the enforcement school explains more of the observed (non)compliance than the management school does. This asymmetry results from the management school's being substantially easier to test empirically than the enforcement school is.

Third, if we allow some synthetization of the management and enforcement schools, the ability to account for (non)compliance increases substantially. Indeed, the explanatory power of such a "hybrid" theory is even higher than that of the two schools combined. Norwegian $NO_x$ policies were lax for years, but a policy package (including a $NO_x$ emissions tax) came into force in January 2007 and reduced $NO_x$ emissions well beyond BAU. While Norway's pre-2007 policies were consistent with the enforcement school and the stricter policies thereafter were consistent with the management school, neither theory is capable of explaining the policy *change*.[40] In contrast, the deadline-pressure hypothesis, which may be seen as a hybrid of the two established theories, can explain Norway's policies throughout the period between its commitment to Gothenburg and the deadline for compliance. Simply put, the argument is that while the driving forces specified by managerialists and enforcement scholars – norms and (material) self-interest – may be present in authorities' calculations simultaneously, their relative influence on policies and behavior varies over time. Norms are the stronger force when an international compliance

[39] Unsurprisingly, previous empirical testing of the enforcement school has often been confined to less demanding hypotheses, such as the effect of regime-level sanctioning powers on compliance (see Breitmeier et al. 2006).
[40] In the case of Norway, both the enforcement school and the management school hypothesize $NO_x$ policy stability. Unless some factor shifts a state's cost/benefit calculus, the former predicts status quo policies. Similarly, the latter predicts little change unless some norm is triggered.

deadline is approaching, and self-interest is stronger when the deadline is distant.

The deadline-pressure hypothesis is a *hybrid explanation of state behavior* and has a close relative in the *hybrid strategy of increasing regime compliance and effectiveness*. My novel typology developed in article 4 helps explain why regimes designed in accordance with this hybrid strategy might be particularly effective. The typology shows that for cooperation between states that vary in terms of capacity and cooperative intent, combining soft "managerial" and hard "enforcement" measures will likely be most effective.

Fourth, several perspectives other than those of the enforcement and management schools contribute to explaining compliance with CLRTAP protocols. As shown above, article 2 demonstrates that both the office-incumbent and the deadline-pressure hypotheses are consistent with much of the evidence examined in the case of Norway's $NO_x$ policies, but only the latter is consistent with the fact that post-2005 emissions declined across Northern and Western Europe, and that Norway's emissions reductions were relatively modest.

In addition to the deadline-pressure and "office-incumbent" hypotheses, state capacity may explain compliance with CLRTAP protocols. The effect of capacity seems, however, to be the opposite of what managerialists expect. My medium-N analyses show a nearly consistent negative effect of capacity on compliance, regardless of model specification. In article 3, I develop the argument that such a negative effect is not as counterintuitive at it might seem, and that it indeed should be expected among reluctant states that pursue policy goals negatively correlated with compliance.

Fifth, and related to the third point above, the impact of EU law on air pollution is substantially stronger than that of CLRTAP regulations. Specifically, my assessments show that the 1988 Sofia Protocol had little, if any, impact on $NO_x$ emissions. In contrast, the evidence suggests that a series of EU directives caused an annual emissions reduction of 3.9 percent in EU member states. Given that previous assessments have found little effect of participation in CLRTAP

protocols,[41] that is an important finding, because it shows that international cooperative efforts are indeed important for reducing harmful air pollution. Rather obviously, EU law is important because it has a strong and direct effect on emissions. That finding also suggests a more subtle point – that also CLRTAP and its protocols might have made (indirect) contributions to emissions reductions. Indeed, the introduction to the Large Combustion Plants (LCP) Directive – an early directive aiming to reduce air pollution in the EU – explicitly refers to the EU's CLRTAP membership as a reason for enacting the directive (Council Directive 88/609/EEC). Such an indirect effect of CLRTAP hinges, however, on the assumption that EU law targeting air pollutants would have been substantially weaker absent CLRTAP, or that it would have been enacted later.

# Suggestions for future research

This dissertation provides several suggestions for future research on the compliance with and the effectiveness of international agreements. First, it may inspire future research by showing how analysts can empirically assess the explanatory power of independent variables suggested by existing theories. Second, it suggests explanatory variables and interactions that, while not necessarily being entirely new to researchers of international cooperation, prove to have a different impact than suggested by previous studies, or have received very limited attention by such studies.

Future compliance studies may be inspired by my assessments of the explanatory power of unexpected changes between commitments and deadlines (articles 1 and 2). Previous studies have paid scant attention to Chayes and Chayes' third explanation. However, using scenarios of emission drivers as well as historical emissions estimates, article 1 demonstrates that such assessments are indeed possible. In particular, future research should examine similar projections of key determinants of the conditions for compliance. Although the emissions-driver scenarios provided by the scientific cooperation under CLRTAP might be particularly detailed and

[41] Bratberg et al. (2005) being the exception.

scientifically well-founded, such projections are provided in the context of other international cooperation forums as well. Examining them might add to the knowledge provided here, and might lead to some more general conclusions concerning Chayes and Chayes' third explanation's ability to account for noncompliance.

Although assessing the enforcement school's ability to account for noncompliance is highly challenging, article 2 suggests how such an assessment may be conducted. In societies with effective bureaucracies, the costs and benefits of many policy options are often estimated by authorities (or by consultancies contracted by government bodies). Surely, such studies are faced with the challenging task of constructing valid BAU scenarios. Nonetheless, comparing observed state behavior with the behavior suggested by those cost-benefit analyses is one promising way forward for improving our knowledge on the enforcement school's explanatory power.[42]

Because my novel conjecture of the capacity–compliance relationship (see article 3) is not only consistent with patterns of (non)compliance with CLRTAP protocols, but also helps explain the inconsistency of previous studies' findings concerning the effects of capacity, future research should dig deeper into the effects of capacity. To the extent that this account is correct, common expectations concerning high-capacity states' compliance are imprecise. Of course, revealing the intentions of governments or authorities is a data-intensive task, and intentions often remain unobserved (Neumayer 2002). However, my second article demonstrates that an in-depth case study can tell much about a state's intention to comply. I show that Norwegian authorities enacted no potent $NO_x$ policies (even though such policies were indeed available and viable) despite being well aware that Norway was heading towards substantial noncompliance. Granted, it is by no means straightforward to use that evidence to score Norway's exact pre-2005 value on a scaled "intention-to-comply" variable, but it suffices to argue that Norway's intention to comply in those years was rather low. Moreover, future research should look into the sources of

[42] Interestingly, the reasoning underlying my evaluation of the enforcement school's explanatory power (in article 2) is somewhat similar to Dai's (2007, 93) assessment of Norway's policies against sulfur emissions in the late 1980s. She shows that the costs of Norway's sulfur policies by far outweighed the domestic benefits (i.e., that they were beyond-BAU policies), and argues that the policies make little sense unless we consider the increased strength of pro-environment domestic actors and their use of international monitoring mechanisms to increase compliance.

variations in states' intention to comply. Two starting points might be works emphasizing the role of domestic constituencies (Dai 2005; 2007) or domestic policy-making procedures (Meckling and Nahm 2018) in shaping governments' compliance preferences or policy outcomes.

My novel theoretical conjecture concerning the capacity–compliance relationship is developed in a specific empirical context – the cooperation under CLRTAP. However, the conjecture is in principle general, that is, it might be applicable in environmental issue areas different from air pollution, and even in non-environmental issue areas. Hence, in cases of states pursuing policy goals that conflict with compliance, and realization of these goals is facilitated by state capacity, state capacity could cause noncompliance with agreements on human rights[43] or trade. It remains, however, a task for future empirical research to assess the extent to which my theoretical argument is consistent with data from cooperation in those issue areas.

The deadline-pressure hypothesis presented and assessed in article 2 also suggests ways forward for future compliance research. Quantitative studies may assess when state behavior begins to conform with international commitments. For instance, using national emissions data over time (such as the data set I use in article 4), researchers could estimate whether (and if so, when) emissions trends have any breaking points. In contrast, qualitative studies may trace the political processes that might (or might not) lead to beyond-BAU policies being implemented.

Having shown that the enforcement and management schools may be synthesized, the present dissertation suggests that future IEA scholarship should continue the theoretical cross-fertilization previously advocated by authors such as Tallberg (2002), Underdal (2008), and Börzel et al. (2010). One such contribution could be to further develop the deadline-pressure hypothesis, most importantly to dig deeper into the behavioral foundation for the hypothesized change from weak to strong(er) policies. Such theoretical improvements would hopefully aid in developing precise expectations of when the shift from interest-driven to norm-driven behavior

[43] Risse and Ropp (2013, 15) argue that human rights scholars have paid too little attention to state capacity as an explanation of (non)compliance. Their main capacity-related hypothesis seems, however, to align with that of management scholars such as Chayes and Chayes (1993): That increased state capacity increases compliance.

should occur.[44]

Such theoretical bridging attempts should be accompanied by empirical studies. One promising path forward is to examine existing and novel data searching for interaction effects. For instance, it would be highly interesting to revisit the International Regimes Database and assess the effectiveness of or compliance with regimes that rely on *both* capacity-building and incentives, and to contrast the results with those of Breitmeier et al. (2006).

Suggesting that EU law constituted an important driver of $NO_x$ emissions, article 4 should inspire at least two ways forward for similar studies. First, as shown above, several studies find no effect of CLRTAP protocols such as Helsinki, Oslo, and Gothenburg (Ringquist and Kostadinova 2005; Aakvik and Tjøtta 2011; Vollenweider 2013); however, none of these studies take into account the fact that overlapping EU directives existed alongside all these protocols. The negative findings of the abovementioned studies do not rule out that EU law had substantially and significantly negative effects on the emissions of EU members. Indeed, increased or stable emissions among non-EU CLRTAP members might have concealed emissions reductions in EU member states. A rather simple starting point for this re-examination would be to replicate the abovementioned studies and assess the effects of EU law on emissions.

Second, my fourth article provides a blueprint for studies aiming to reveal the relative effectiveness of overlapping international regimes and agreements. Variation in the membership and/or timing of commitments might provide the analytical leverage researchers need to assess which of two (or more) overlapping institutions is more effective.

[44] The deadline-pressure hypothesis has at another notable limitation: It does not account for *variation in compliance*. That such variations exist in the empirical context I study, both across states and substances, is evident from Table 1 in article 1. However, this is hardly a major concern. Only under ceteris paribus conditions does the deadline-pressure hypothesis predict uniform compliance. Because everything else is usually *not* equal, it comes as no surprise that the deadline-pressure hypothesis should be complemented by other explanations.

# References:

Aakre, Stine, Leif Helland, and Jon Hovi. 2016. “When Does Informal Enforcement Work?” *Journal of Conflict Resolution* 60, no. 7: 1312–40.

Aakvik, Arild, and Sigve Tjøtta. 2011. “Do collective actions clear common air? The effect of international environmental protocols on sulphur emissions.” *European Journal of Political Economy* 27, no. 2: 343–51.
doi: 10.1016/j.ejpoleco.2010.12.001

Abbott, Kenneth W., and Snidal, Duncan. 2000. “Hard and Soft Law in International Governance,” *International Organization* 54, no. 3: 421–56.
doi: 10.1162/002081800551280

Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist’s Companion.* London: Princeton University Press.

Barrett, Scott. 2008. “Climate treaties and the imperative of enforcement,” *Oxford Review of Economic Policy* 24, no. 2: 239–58.
doi: 10.1093/oxrep/grn015

Börzel, Tanja A., Tobias Hofmann, Diana Panke, and Carina Sprungk. 2010. “Obstinate and Inefficient: Why Member States Do Not Comply With European Law.” *Comparative Political Studies*, 43, no. 11: 1363–90.
doi: 10.1177/0010414010376910

Brown Weiss, Edith, and Harold K. Jacobson (eds.). 1998. *Engaging Countries: Strengthening Compliance with International Environmental Accords*, Cambridge: MIT Press.

Bratberg, Espen, Sigve Tjøtta, and Torgeir Øines. 2005. “Do voluntary international environmental agreements work?” *Journal of Environmental Economics and Management*, 50, no. 3: 583–97.

Breitmeier, Helmut, Oran R. Young, and Michael Zürn. 2006. *Analyzing International Environmental Regimes. From Case Study to Database.* London: MIT Press.

Chayes, Abram, and Antonia Handler Chayes. 1993. “On Compliance.” *International Organization* 47, no. 2: 175–205.
doi: 10.1017/S0020818300027910

Chayes, Abram, and Antonia Handler Chayes. 1995. *The New Sovereignty.* London: Harvard University Press.

Chayes, Antonia Handler, Abram Chayes, and Ronald B. Mitchell. 1995. “Active Compliance Management in Environmental Treaties.” In Winfried Lang, ed., *Sustainable Development and International Law*, 75–89. London: Graham & Trotman.

Conca, Ken. 2006. *Contentious Transnational Politics and Global Institution Building.* London: MIT Press.

Council Directive 88/609/EEC of 24 November 1988 on the limitation of emissions of certain pollutants into the air from large combustion plants. https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:31988L0609&qid=1539599691494&from=EN

Dai, Xinyuan. 2005. “Why Comply? The Domestic Constituency Mechanism.” *International Organization*, 59, no. 2, 363–98.
www.jstor.org/stable/3877908

Dai, Xinyuan. 2007. *International Institutions and National Policies.* Cambridge: Cambridge

University Press.

Downs, George W., David M. Rocke, and Peter N. Barsoom. 1996. “Is the good news about compliance good news about cooperation?” *International Organization* 50, no. 3: 379–406. doi: 10.1017/S0020818300033427

Di Primio, Juan Carlos. 1998. “Data Quality and Compliance Control in the European Air Pollution Regime.” In *The Implementation and Effectiveness of International Environmental Commitments: Theory and Practice*, edited by David G. Victor, Kal Raustiala, and Eugene B. Skolnikoff, 283–302. London: MIT Press.

EMEP. 2009. “EMEP Status Report 1/09. Transboundary acidification, eutrophication and ground level ozone in Europe in 2007.”
Available at: http://emep.int/publ/common_publications.html#2009

European Commission. 2013. “Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. A Clean Air Programme for Europe.”
https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52013DC0918&from=EN
The full impact assessment is available from (also in file with author):
https://tinyurl.com/ybt4pxdf

Franck, Thomas M. 1988. “Legitimacy in the International System.” *American Journal of International Law* 8, no. 4: 705–59.

Goertz, Gary, and James Mahoney. 2012. *A tale of two cultures: Qualitative and quantitative research in the social sciences*. Princeton University Press.

Grewal, Sharanbir, and Erik Voeten. 2015. “Are New Democracies Better Human Rights Compliers?,” *International Organization* 69, no. 2: 497–518.

doi: 10.1017/S0020818314000435

Hasenclever, Andreas, Peter Mayer, and Volker Rittberger. 1997. *Theories of International Regimes*. Cambridge: Cambridge University Press.

Henkin, Louis. 1968. *How Nations Behave: Law and Foreign Policy*. London: Pall Mall Press.

Hill, Michael, and Peter Hupe. 2014. *Implementing Public Policy: An Introduction to the Study of Operational Governance.* London: Sage Publications.

Hovi, Jon, Detlef F. Sprinz, and Arild Underdal. 2003. "The Oslo-Potsdam Solution to Measuring Regime Effectiveness: Critique, Response, and the Road Ahead." *Global Environmental Politics* 3, no. 3: 74–96.

Hønneland, Geir, Steinar Andresen, and Elin Lerum Boasson. 2012. "Introduction." In *International Environmental Agreements*, edited by Steinar Andresen, Elin Lerum Boasson, and Jørgen Wettestad, 23–37. New York: Routledge.

Jacobson, Harold K., and Edith Brown Weiss. 1998. "Assessing the Record and Designing Strategies to Engage Countries." In *Engaging Countries: Strengthening Compliance with International Environmental Accords*, edited by Edith Brown Weiss and Harold K. Jacobson, 511–54. Cambridge: MIT Press.

Keohane, Robert. 1984. *After Hegemony: Cooperation and Discord in the World Political Economy*. Princeton University Press.

Krasner, Stephen D. 1985. *Structural conflict. The third world against global liberalism*. Berkeley: University of California Press.

Kratochwil, Friedrich. 1984. "The force of prescriptions." *International Organization* 38, no. 4:

685–708.
doi: 10.1017/S0020818300026916

Levy, Marc A. 1993. “European Acid Rain: The Power of *Tote-Board Diplomacy*.” In *Institutions for the Earth. Sources of Effective International Environmental Protection,* edited by Peter M. Haas, Robert O. Keohane, and Marc A. Levy, 75–132. London: MIT Press.

Maas, Rob, and Per Inge Grennfelt (eds). 2016. *Towards Cleaner Air. Scientific Assessment Report 2016. EMEP Steering Body and Working Group on Effects of the Convention on Long-Range Transboundary Air Pollution*. Oslo.
Available at:
https://www.unece.org/index.php?id=42861

Mearsheimer, John J. 1994–1995. “The false promise of international institutions.” *International Security* 19, no. 3: 5–49.

Meckling, Jonas, and Jonas Nahm. 2018. “The power of process: State capacity and climate policy.” *Governance* 31: 741–57.
doi: 10.1111/gove.12338

Miles, Edward L., Arild Underdal, Steinar Andresen, Jørgen Wettestad, Jon Birger Skjærseth, and Elaine M. Carlin. 2002. *Environmental Regime Effectiveness: Confronting Theory with Evidence.* London: MIT Press.

Mitchell, Ronald B. 2002. “A Quantitative Approach to Evaluating International Environmental Regimes.” *Global Environmental Politics* 2, no. 4: 58–83.

Mitchell, Ronald B. 2008. “Evaluating the performance of environmental institutions: what to evaluate and how to evaluate it?” In *Institutions and Environmental Change: Principal Findings, Applications, and Research Frontiers,* edited by Oran R. Young, Leslie A. King, and Heike

Schroeder, 79–114. London: MIT Press.

Norwegian Environmental Agency. 2018. “Acid Rain.”
Available at:
https://tinyurl.com/y75qpzgw
Accessed February 1st 2019

Neumayer, Eric. 2002. “Do Democracies Exhibit Stronger International Environmental Commitment? A Cross-country Analysis.” *Journal of Peace Research* 39, no. 2: 139–64.
doi.org: 10.1177/0022343302039002001

Rafaj, Peter, Marcus Amann, José Siri, and Henning Wuester. 2014. “Changes in European greenhouse gas and air pollutant emissions 1960–2010: Decomposition of determining factors.” *Climatic Change* 124, no. 3: 477–504.
doi.org: 10.1007/s10584-013-0826-0

Ringquist, Evan. J., and Tatiana Kostadinova. 2005. “Assessing the Effectiveness of International Environmental Agreements: The Case of the 1985 Helsinki Protocol.” *American Journal of Political Science*, 49, no. 1: 86–102.
doi: 10.1111/j.0092-5853.2005.00112.x

Risse, Thomas, and Stephen C. Ropp. 2013. “Introduction and Overview.” In *The Persistent Power of Human Rights: From Commitment to Compliance*, edited by Thomas Risse, Stephen C. Ropp, and Kathryn Sikkink, 3–25. Cambridge: Cambridge University Press.

Ruggie, John Gerard. 1982. “International Regimes, Transactions, and Change: Embedded Liberalism in the Postwar Economic Order.” *International Organization*, 36, no. 2: 379–415.
www.jstor.org/stable/2706527

Strange, Susan. 1982. “Cave! Hic Dragones: A Critique of Regime Analysis.” *International*

*Organization* 36, no. 2: 479–96.
www.jstor.org/stable/2706530

Tallberg, Jonas. 2002. “Paths to Compliance: Enforcement, Management, and the European Union.” *International Organization,* 56, no. 3, 609–43.
doi: 10.1162/002081802760199908

Underdal, Arild and Kenneth Hanf (eds.). 2000. *International Environmental Agreements and Domestic Politics: The Case of Acid Rain*. Aldershot: Ashgate.

Underdal, Arild. 2008. “Determining the Causal Significance of Institutions: Accomplishments and Challenges.” In *Institutions and Environmental Change: Principal Findings, Applications, and Research Frontiers,* edited by Oran R. Young, Leslie A. King, and Heike Schroeder, 49–78. London: MIT Press.

Victor, David G. 2011. *Global Warming Gridlock: Creating More Effective Strategies for Protecting the Planet.* Cambridge: Cambridge University Press.

Victor, David, Kal Raustiala, and Eugene B. Skolnikoff (eds.). 1998a. *The Implementation and Effectiveness of International Environmental Commitments: Theory and Practice*. London: MIT Press.

Victor, David, Kal Raustiala, and Eugene B. Skolnikoff. 1998b. “Introduction and overview.” In *The Implementation and Effectiveness of International Environmental Commitments: Theory and Practice*, edited by David G. Victor, Kal Raustiala, and Eugene B. Skolnikoff, 1–46. London: MIT Press.

Vollenweider, Jürg. 2013. “The Effectiveness of International Environmental Agreements.” *International Environmental Agreements: Politics, Law and Economics* 13, no. 3: 343–67.

Wagner, Susanne, Elisabeth Angenendt, Olga Beletskaya, Jürgen Zeddies. 2017. “Assessing ammonia emission abatement measures in agriculture: Farmers' costs and society's benefits – A case study for Lower Saxony, Germany.” *Agricultural Systems*, 157: 70–80. doi: 10.1016/j.agsy.2017.06.008

Waltz, Kenneth. 1979. *Theory of International Politics.* Reading, Massachusetts: Addison-Wesley.

Wettestad, Jørgen. 1998. “Participation in $NO_x$ Policy-Making and Implementation in the Netherlands, UK, and Norway: Different Approaches but Similar Results?” In *The Implementation and Effectiveness of International Environmental Commitments: Theory and Practice*, edited by David G. Victor, Kal Raustiala, and Eugene B. Skolnikoff, 381–430. London: MIT Press.

Wettestad, Jørgen. 2002. “The Convention on Long-Range Transboundary Air Pollution (CLRTAP).” In *Environmental Regime Effectiveness. Confronting Theory with Evidence,* Edward L. Miles, Arild Underdal, Steinar Andresen, Jørgen Wettestad, Jon Birger Skjærseth, and Elaine M. Carlin, 197–222. London: MIT Press.

Wettestad, Jørgen. 2012. “Reducing Long-range Transport of Air Pollutants in Europe.” In Steinar Andresen, Elin Lerum Boasson, and Jørgen Wettestad, eds., *International Environmental Agreements*, 23–37. New York: Routledge.

Young, Oran R. 1979. *Compliance and Public Authority: A Theory with International Applications.* London: Johns Hopkins University Press.
Young, Oran R. 1989. “The Politics of International Regime Formation: Managing Natural Resources and the Environment” *International Organization* 43, no. 3: 349–75.

Young, Oran R. 2003. “Determining Regime Effectiveness: A Commentary on the Oslo-Potsdam Solution.” *Global Environmental Politics* 3, no. 3: 97–104.

Young, Oran R. 2011. “Effectiveness of International Environmental Regimes: Existing Knowledge, Cutting-Edge Themes, and Research Strategies.” *Proceedings of the National Academy of Sciences of the United States of America* 108, no. 50: 19853–60.

Young, Oran R. (ed.). 1999. *The effectiveness of International Environmental Regimes. Causal Connections and Behavioral Mechanisms.* Cambridge, Mass.: MIT Press.

Young, Oran R., and Marc A. Levy. 1999. “The Effectiveness of International Environmental Regimes.” In *The effectiveness of International Environmental Regimes. Causal Connections and Behavioral Mechanisms*, edited by Oran R. Young. Cambridge, Mass.: MIT Press.

Zürn, Michael. 1998. “The Rise of International Environmental Politics: A Review of Current Research.” *World Politics* 50, no. 4: 617–49.
doi: 10.1017/S0043887100007383

Int Environ Agreements (2018) 18:491–512
https://doi.org/10.1007/s10784-018-9400-6

ORIGINAL PAPER

# Can the management school explain noncompliance with international environmental agreements?

**Andreas Kokkvoll Tveit**[1]

Accepted: 8 May 2018 / Published online: 5 June 2018


**Abstract** Although the management school has been highly influential in the international cooperation literature, the explanatory power of Chayes and Chayes' three explanations of noncompliance with international environmental treaties remain understudied. Having developed a framework for examining the explanatory power of treaty ambiguity, lack of state capacity, and unexpected social or economic developments, this paper conducts a rigorous empirical test in the context of a well-suited case—the 1999 Gothenburg Protocol. A careful reading shows that the language of the protocol is clear and unambiguous; indeed, there has been no disagreement over the treaty's content. Furthermore, statistical analyses show no positive effect of political capacity on compliance. Finally, parties had adequate time to meet their obligations, and unexpected developments explain only a small part of the observed noncompliance. These findings pose a serious challenge to Chayes and Chayes' three explanations of noncompliance—at least as far as the Gothenburg Protocol is concerned.

**Keywords** International agreements · Compliance · International environmental cooperation · Norms · State capacity

I am most grateful to Jon Hovi, two anonymous reviewers, participants at a panel at the ISA Annual Conference in Baltimore, February 2017, and participants at CICEP's annual research conference, Oslo, September 2016, for constructive comments and suggestions. My research has benefited greatly from my stay at the International Institute for Applied Systems Analysis (IIASA) in May 2016. Parts of my research were carried out while at the Fridtjof Nansen Institute, Lysaker, Norway. Any remaining errors are the author's responsibility.

✉ Andreas Kokkvoll Tveit
a.k.tveit@stv.uio.no

[1] Department of Political Science/Institutt for Statsvitenskap, University of Oslo, Boks 1097 Blindern, 0317 Oslo, Norway

## 1 Introduction

How can we account for noncompliance with international environmental agreements (IEAs)? In their seminal[1] article "On Compliance," Chayes and Chayes (1993) formulate their version of the management school.[2] They argue that "compliance problems often do not reflect a deliberate decision to violate an international undertaking on the basis of a calculation of interests" (Chayes and Chayes 1993: 176). Rather, noncompliance is usually caused by (one or more of) three factors beyond the control of national authorities: Treaty ambiguity, lack of state capacity, and what Chayes and Chayes refer to as "the temporal dimension"—unexpected changes of conditions for compliance following social and economic developments between commitment and deadline.

The compliance debate gained momentum during the 1990s and early 2000s; however, this progress was driven more by theoretical contributions than by empirical advances. Raustiala and Slaughter (2002: 548) argue that "compliance remains a relatively young field" and that "empirical testing of compliance theories is limited". Raustiala and Slaughter's statements still ring true.[3]

The present paper makes several contributions to the literature on compliance with IEAs. First, I develop a framework for assessing the explanatory power of treaty ambiguity, lack of state capacity, and unexpected social or economic developments. In particular, Chayes and Chayes' third explanation of noncompliance remains severely understudied. Based on this framework, I conduct a set of rigorous empirical tests in the context of a well-suited case—the 1999 Gothenburg Protocol.

Second, the present paper differs from previous studies in that I statistically control for the ambitiousness of the participating countries' commitments (i.e. the size of the required emissions reductions). Failing to control for ambitiousness entails a risk of biased results—a risk that is often overlooked (Raustiala 2005; Downs et al. 1996).

Third, unlike both Jacobson and Brown Weiss (1998) and Breitmeier et al. (2006), I measure compliance on the ratio level. Exceeding an emissions target by only 1% is indeed less problematic than exceeding it by 10 or 20%, and the compliance variable should indeed reflect such variance.[4] Finally, my study further differs from Breitmeier et al. (2006) in that I measure *each state's* compliance level (rather than the general compliance with a regime at large).

The case of the 1999 Gothenburg Protocol is well suited for the development of a framework to empirically assess the explanatory power of the management school. First, because the protocol includes national emissions targets for four pollutants, compliance can be measured precisely. Precise measurement of the dependent variable is a prerequisite for the kind of statistical analyses I conduct when I assess the effect of state capacity on compliance. Second, assessing whether compliance has been affected by unexpected social and/or economic developments between commitment and implementation (Chayes and Chayes' third explanation) is certainly challenging: Ideally, it requires data on how states believed the future would look like when they entered the protocol. However, Gothenburg is part of

[1] As of February 2018, Google Scholar counts 1328 citations of Chayes and Chayes' 1993 article.

[2] Scholars such as Young (1979) and Mitchell (1994, 2010) have formulated positions sharing several similarities with those of Chayes and Chayes (1993).

[3] Empirical studies of IEAshave grown in number, but scholars have focused more on effectiveness (for instance Miles et al. 2002; Victor et al. 1998) than on compliance.

[4] Granted, we may also conceive of compliance as a dichotomous concept. Hence, I also use a binary compliance variable.

an international cooperative effort with a strong scientific basis (Castells and Ravetz 2001; Rensvik 2017; Tuinstra 2008). Much energy has been devoted to modelling past and future environmental quality, emissions, and emissions drivers. Such projections were important when Gothenburg's emissions targets were agreed (see Kelly et al. 2010). Therefore, I use projections of future emissions drivers to assess whether compliance proved to be more difficult to reach than the member states expected when they entered Gothenburg.

The remainder of this article proceeds as follows. First, I briefly present the Gothenburg Protocol and the environmental problems it seeks to alleviate. Second, I elaborate on the debate between Chayes and Chayes' management school and its counterpart, the enforcement school. I also review previous attempts at testing these two schools' hypotheses against empirical evidence, and develop a set of hypotheses. Finally, focusing on (non) compliance with the Gothenburg Protocol, I assess the explanatory power of the management school. I show that states have mutually consistent interpretations of the contents of the agreement, and that the protocol's language is unequivocal. Hence, in the case of Gothenburg, there is no ambiguity at all. Chayes and Chayes' "ambiguity explanation" is thus clearly incapable of explaining noncompliance with the Gothenburg Protocol. A series of regressions show a negative relationship between state capacity and compliance, thereby suggesting that the "capacity explanation" cannot account for Gothenburg noncompliance. Finally, although Chayes and Chayes' third explanation can account for some noncompliance, in only four cases can (unexpected) social and economic developments explain all of the gap between targets and observed emissions in 2010. Thus, Chayes and Chayes' theory can account for only a small fraction of the noncompliance with Gothenburg.

## 2 Compliance, effectiveness, and the Gothenburg protocol

The Gothenburg Protocol[5] was adopted in 1999 under the Convention on Long-range Transboundary Air Pollution (CLRTAP). Gothenburg includes national emissions targets for four different pollutants—sulphur oxides ($SO_2$), nitrogen oxides ($NO_x$), non-methane volatile organic compounds (NMVOC), and ammonia. Together, these pollutants cause the three interconnected environmental problems: Acidification,[6] eutrophication,[7] and ground-level ozone.[8]

As these are regional environmental problems, most participants in the cooperation under CLRTAP are European states. The protocol came into force in 2005, and 2010 was chosen as the deadline for reaching the national emissions targets.

Table 1 presents all European states that became parties to Gothenburg no later than 2009 (i.e., before the deadline year), the emissions targets, and the compliance rates. The criterion for being compliant is straightforward: As my analysis below shows, Gothenburg includes emissions targets for four substances for each party, and no provisions that can relieve a party of its obligation to reach its target by 2010. Thus, a state complied with a

[5] I refer to the Gothenburg Protocol of 1999, not the amended protocol of 2012.

[6] Acidification is largely caused by sulphur and NOx, and affects life in water and soil (Miljødirektoratet 2015).

[7] Eutrophication, which increases algae growth and thereby harms other organisms, often stems from ammonia emissions.

[8] NOx reacting with non-methane volatile organic compounds (NMVOCs) causes harmful ground-level ozone.

target if and only if its emissions of the relevant substance in 2010 were below or equal to this target.[9] [10] Compliance thereby differs from effectiveness, since measuring IEA effectiveness typically involves measuring an IEA's ability to improve environmental quality or state behaviour compared to a no-agreement counterfactual. However, establishing such counterfactuals is notoriously difficult (Helm and Sprinz 2000; Hovi et al. 2003; Young 2003).

As demonstrated by Table 1, 21 national emissions targets were not reached by the deadline, 2010. Ten of them were targets for $NO_X$ emissions, eight for ammonia, and three for NMVOC. All $SO_2$ targets were reached by 2010. In seven instances of noncompliance, the target was exceeded by 10% or less. Six targets were exceeded by 10–20%, while another six were exceeded by 20–40%.

## 3 Theory, previous research, and hypotheses

### 3.1 Enforcement or management? Treaty design and sources of noncompliance

The enforcement school (Downs et al. 1996; Barrett 2003; Aakre et al. 2016) argues that states comply only if their expected marginal cost of complying are lower (or equal to) expected marginal revenue. The enforcement school thus views noncompliance as a rational, self-interested actor's reaction to a given material incentive structure. The generally high compliance with international agreements (Henkin 1968) is attributed to the shallowness of commitments (Downs et al. 1996: 382). Agreements are shallow if their commitments only codify what would happen even if the agreement did not exist.

Positive and negative incentives are enforcement scholars' main solution to malign collective action problems. In such cases, however, scholars in the enforcement camp are sceptical of the prospects of international cooperation. Since sanctions and rewards may entail high costs for the sender state, promises of carrots or threats of sticks are usually not credible. Unless material incentives are credibly altered, compliance beyond a business-as-usual (BAU) scenario cannot be expected.

In contrast, Chayes and Chayes (1993: 178) claim that sanctions are costly, inefficient, hard to sustain, and unnecessary. Their "managerial strategy" consists of softer measures: Monitoring and knowledge sharing, effective dispute settlement, building state capacity, and adjusting treaties in light of economic, technological, social, and political changes.

Managerialists argue that international society's anarchical structure is not as detrimental to cooperation as their opponents believe. The main reason is states' "general propensity" to comply—a tendency to sincerely try to act in accordance with international obligations: "In common experience, people, whether as a result of socialization or otherwise, accept that they are obligated to obey the law. So it is with states". In other words, states are largely norm-driven actors. And, in international relations, the norm is to do as agreed (Chayes and Chayes 1993: 178–185. See also Henkin 1968; Finnemore and Sikkink 1998; Simmons 1998, 2013).

[9] In accordance with Young's (1979) definition.

[10] Hence, Table 1 does not engage directly with the highly challenging task of distinguishing between Mitchell's (2010: 147) two kinds of noncompliant behaviour ("good-faith" and "intentional") or between his two kinds of compliant behaviour ("coincidental" and "treaty-induced" compliance).

**Table 1** Parties' compliance with Gothenburg targets (deadline year 2010)

| Party | $NO_X$ | NMVOC | Sulphur | Ammonia |
|---|---|---|---|---|
| Belgium | **139** | **107.9** | 57.1 | 88 |
| Bulgaria | 52.1 | 55.8 | 45.2 | 38.4 |
| Croatia | 74 | 61 | 49.6 | **129.4** |
| Cyprus | 80 | 71.3 | 56.3 | 62.2 |
| Czech Rep. | 77 | 78.4 | 56.6 | 67.1 |
| Denmark | **114** | **147.5** | 27.9 | **115.9** |
| Finland | 97.6 | 89.4 | 57.6 | **123.4** |
| France | **127.5** | 79.5 | 71.3 | 93.4 |
| Germany | **123.4** | **124.5** | 79 | **116.8** |
| Hungary | 77.8 | 91.3 | 5.7 | 86 |
| Latvia | 45.7 | 65.6 | 2.4 | 32.7 |
| Lithuania | 45 | 77.7 | 14.3 | 51.4 |
| Luxembourg | **358.7** | 94.3 | 43.9 | 67.5 |
| Netherlands | **103.1** | 82.7 | 68.2 | **112.3** |
| Norway | **113.6** | 71.6 | 89.5 | **119.3** |
| Portugal | 68.1 | 89 | 31.2 | 42.8 |
| Romania | 53.1 | 66.5 | 38.1 | 80 |
| Slovakia | 68.2 | 45.6 | 63.1 | 63.9 |
| Slovenia | **104.7** | 96 | 36.5 | 95 |
| Spain | **113** | 97.4 | 54.6 | **111** |
| Sweden | **101.1** | 79.5 | 47.7 | 90.6 |
| Switzerland | 98.3 | 62.6 | 46.7 | **101** |
| United Kingdom | 95.1 | 71.3 | 68.4 | 93.9 |

2010 emissions in % of targets. Targets that were not reached score values above 100, and are bolded (Source: Kokkvoll Tveit 2018)

Thus, whenever noncompliance occurs, the cause is usually not that cheating maximizes the individual state's private net benefit. Rather, the sources of noncompliance lie beyond the state's reach.

First, ambiguity may cause noncompliance. Chayes and Chayes (1993): 188–189) state that "Treaties (…) frequently do not provide determinate answers to specific disputed questions". Hence, "a zone of ambiguity within which it is difficult to say with precision what is permitted and what is forbidden" occurs.

Second, compliance might require more than parties can deliver. Scientific and technical competence, bureaucratic resources, and economy are the three constraining factors specified by Chayes and Chayes (1993, 1995).

Third, the "temporal dimension" might explain noncompliance. Chayes and Chayes (1993: 195) argue that "Significant changes in social or economic systems mandated by regulatory treaties take time to accomplish. Thus, a cross section at any particular moment in time may give a misleading picture of the state of compliance". The moment in time when compliance is assessed should therefore be chosen carefully. Furthermore, *conditions for compliance* may change between the moment when a commitment is made and the deadline for reaching the targets. If these changes are substantial, unexpected and difficult to control, they may affect states' compliance considerably.

### 3.2 Empirical studies of international environmental cooperation

Several large empirical studies have focused on international environmental cooperation and tested hypotheses derived from the management and enforcement schools. Breitmeier et al. (2006: 110–111) state that "neither the shallowness argument of Downs et al. (1996) nor the management school of Chayes and Chayes can explain patterns of compliance with international environmental regimes". Victor et al. (eds. 1998) focus on *implementation* and *effectiveness* of international environmental cooperation, and find that "some implementation failures are intentional" and that hard measures such as sanctions sometimes is necessary. Furthermore, they argue that "legally binding agreements often codify what is already under way," thereby supporting Downs et al.'s (1996) "shallowness claim" (Raustiala and Victor 1998: 662).[11]

Similarly, previous studies of cooperation to reduce long-range transboundary air pollution have mainly focused on effectiveness (Levy 1993; Böhmelt and Vollenweider 2015).[12] Wettestad (2012: 34) argues that much of the last decades' substantial emissions reductions are due to other factors than CLRTAP protocols. Helm and Sprinz (2000) find that the 1985 Helsinki and the 1988 Sofia protocols reduced emissions compared to the counterfactual scenario, although cooperation falls short of the collective optimum. That conclusion is supported by Bratberg et al.'s (2005) econometric analysis of Sofia participation. In contrast, Ringquist and Kostadinova (2005) find that Helsinki did not reduce participants' emissions.

What explanatory power have previous studies attributed to the three factors that, according to Chayes and Chayes, cause noncompliance? Concerning ambiguity and compliance, Breitmeier et al. (2006: 90–93, see also their Table 3.11) find that "the association between the precision of rules and compliance rates is positive but not strong". Jacobson and Brown Weiss (1998)[13] conclude similarly.

Jacobson and Brown Weiss (1998) and Breitmeier et al. (2006) offer divergent findings concerning capacity. The latter conclude that "[our data] do not confirm expectations about the role of capacity building," while the former find that administrative capacity is important. Their differing findings may to some extent be explained by differences in research design and observational units: While Breitmeier et al. (2006) study the general compliance with a treaty or regime, the case studies included in Brown Weiss and Jacobson's (998, eds.) assess individual states' compliance. Although four of the five treaties studied by Brown Weiss and Jacobson (998, eds.) are among the 23 regimes under scrutiny by Breitmeier et al. (2006), the latter's empirical focus is certainly the broadest of the two. Assessments of the explanatory power of Chayes and Chayes' third explanation of noncompliance—changed conditions for compliance following unexpected social or economic changes—are few and far between.[14]

[11] In contrast, Bernauer et al. (2013) find no support for the enforcement school's hypothesis of a trade-off between depth and participation.

[12] However, an assessment of previous CLRTAP protocols concluded that negotiation positions, implementation, and compliance (operationalized as emissions reductions) were reasonably well predicted by a model of states as unitary rational actors (Underdal 2000: 351–353).

[13] This anthology includes studies of eight states' (and the EU's) compliance with five international environmental treaties.

[14] See, however, Kokkvoll Tveit's (2018) recent in-depth case study.

### 3.3 Hypotheses and research design

In the face of noncompliance, the empirical expectations of Chayes and Chayes' management school are clear.

From the ambiguity explanation, the following hypotheses may be derived:

**H1a** The contents of the Gothenburg Protocol are open to interpretation.

**H1b** The parties have divergent views of their obligations under the protocol.

Likewise, if a lack of state capacity explains the observed noncompliance, we should find a positive effect of political capacity on compliance:

**H2** The higher a state's capacity, the higher the (likelihood of) compliance.

Chayes and Chayes' third explanation suggests that time was too short to reach the targets that were not complied with. I examine the developments of several conditions that are crucial for compliance. If these conditions have developed differently than the parties expected when the agreement was adopted—for instance if consumption of energy in 2010 was higher than projected in 1999—the temporal dimension may account for noncompliance. However, the difference between projections and what actually happened must be substantial, and large enough to account for the gap between the 2010 emissions and the 2010 target.

**H3** Unexpected changes between commitment and deadline made the emissions targets substantially harder to reach.

## 4 Analysis: Can treaty ambiguity explain noncompliance?

Gothenburg's Article 3, Paragraph 1, states that "Each party shall, as a minimum, control its annual emissions of polluting compounds in accordance with the obligations in annex II". Annex II specifies emissions ceilings for ammonia, $NO_X$, sulphur, and NMVOC—in thousand (metric) tonnes per year—for 36 states (as well as for the EU). The deadline year is 2010. Gothenburg includes no provision that could exempt parties from being obliged to reach the emissions ceilings by 2010—unless they withdraw from the agreement.

Thus, Gothenburg's language is clear and unequivocal: States that become parties to the agreement shall in 2010 and thereafter not exceed their designated annual emission ceilings.

Germany's environmental agency, the Umweltbundesamt (2017), writes that "After 2010, $NO_X$ emissions above 1081 thousand tonnes are not allowed [by the Gothenburg Protocol]."[15] Statements from Danish (Miljøstyrelsen 2002), Swedish (Naturvårdsverket 2016), and British (Department of Environment, Food and Rural Affairs 2015) authorities express views fully consistent with the statements from Germany's Umweltbundesamt.

[15] The original text is as follows: "Seit dem Jahr 2010 dürfen 1.081 Tausend Tonnen $NO_x$ nicht mehr überschritten werden.".

It seems clear that ambiguity did not cause the noncompliance with the Gothenburg Protocol. The fact that compliant and noncompliant parties alike have mutually consistent interpretations of their obligations strengthens this conclusion.

## 5 Analysis: Can lack of capacity explain noncompliance?

In this section, I examine the effect of state capacity on compliance.

### 5.1 Operationalization

Being a highly contested concept, state capacity is challenging to measure (Hanson and Sigman 2013; Jänicke 1997). For want of a generally accepted operationalization, I use two operationalizations[16] that were suggested by theorists of the management school, have high face validity, and allow comparison across states. First, I operationalize capacity as states' scores on one of the World Bank's Worldwide Governance Indicators (WGI).[17] In the words of the World Bank (2017), the Government Effectiveness indicator "reflects perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies".

Second, I operationalize capacity as GDP per capita (log-transformed).[18] According to Chayes and Chayes (1993: 194), economic wealth increases states' capacity for compliance. Moreover, scholars seem to agree that states' bureaucratic resources and capabilities strongly depend on their general level of economic development (Chayes and Chayes 1995; Jänicke 1997; Jacobson and Brown Weiss 1998: 531).

I operationalize ambition level as the deviation of the 2010 emissions target from the corresponding emissions in 1999, the year Gothenburg was adopted. Each country's 1999 emissions of a given substance[19] are divided by the country's 2010 emissions target for that substance. For instance, because the UK's 1999 emissions were 58% above the emissions target for 2010, the UK $NO_X$ target unit scores 1.58 on ambition level.

Compliance is operationalized in two ways. First, a continuous compliance variable measures the 2010 emissions' deviance from the 2010 target. Values above 0 indicate that emissions were below the target (the state concerned was thus in compliance), while targets that were not reached score below 0. For instance, the UK $NO_X$ target unit scores 0.049 on the compliance variable, because the 2010 UK $NO_X$ emissions were 4.9% below the target (see also Table 1). Second, because we may conceptualize compliance as dichotomous, I also use a binary compliance variable. If the 2010 emissions were higher than the target, the unit scores 0. Conversely, units with emissions below or equal to the target score 1.[20]

[16] In their study of compliance with EU law, Börzel et al. (2010) operationalize state capacity as GDP per capita and scores on a government effectiveness index.

[17] WGI scores are based on surveyed views of experts, citizens and enterprise respondents.

[18] I log-transform GDP per capita because its relationship to political capacity is likely nonlinear.

[19] Unless I state otherwise, all emissions are in metric tonnes, and as reported to UNECE in 2015.

[20] Using the binary compliance variable is also warranted by the considerable over-compliance by several parties shown in Table 1. Such over-compliance *may* suggest that the emissions levels were not primarily a result of deliberate efforts to reach the target. Regressions using the binary compliance variable do not estimate on that potentially irrelevant information.

### 5.2 Data and estimation

The observational unit of my regressions is a given emissions target concerning a particular substance for a given party. All emissions targets shown in Table 1 thus correspond to a unit in my data set. Every party has four obligations, one for each regulated substance. Consequently, standard errors are clustered on states.

Because all my variables are measured on the interval scale or are dichotomous, I use OLS and logistic regression to estimate the causal effects of my independent variables.[21]

### 5.3 Results

Table 2 shows the results of six OLS regressions. In three regressions (Models 1–3), I operationalize capacity as Government Effectiveness. In the other three (Models 4–6), I operationalize capacity as (log) GDP per capita.

Models 1 and 4, which include capacity as the only independent variable, show a negative and statistically significant effect of capacity on compliance. When I control for ambition level (Models 2 and 5), the effect of capacity on compliance remains negative (yet statistically significant only in model 5.) Models 3 and 6 add an Eastern Europe dummy variable that controls for geographical, historical, political, and economic ties between countries in Europe. When this dummy is added, the estimates for capacity and ambition level are similar to those of Models 2 and 5.

Even though the models' explained variance is not crucial for the purpose of this paper, it is interesting to note that $R^2$ increases substantially when ambition level is included.

Table 3 shows the results of six regressions corresponding to those in Table 2, except that the dependent variable is dichotomous in Table 3's models. The effects of capacity reported in Table 3 are consistently negative, although statistically insignificant in models 9 and 12.

Thus, I do not find the positive effect of capacity on compliance with the Gothenburg Protocol expected by the management school (H2). The sensitivity checks reported in the appendix (Tables 9 and 10) show that this conclusion holds under a number of conditions. The effect of capacity is consistently negative in models using a third operationalization of capacity as well as in models that include substance-specific dummies. Hence, the conclusion that capacity does *not* have a positive effect on compliance seems highly robust.

[21] A multilevel model is infeasible because of few (4) units on the state level. Likewise, estimating causal effects by using instrumental variables (Angrist and Pischke 2009) is infeasible since it is highly doubtful that any valid instrument Z exists for my variables (see Angrist and Pischke's (2009: 117) discussion of criteria for valid instrumental variables). Bratberg et al. (2005) estimate the effect of participation in CLRTAP agreements on emissions by employing the difference-in-differences (DID) estimator, thus comparing participants to non-participants. The DID technique is, however, less feasible when compliance is the dependent variable, since only states that participate in the agreement may comply (or defect).

**Table 2** OLS regressions. Dependent: compliance (continuous)

| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
|---|---|---|---|---|---|---|
| Constant | 0.508 | 0.809 | 0.756 | 4.719 | 3.015 | 3.488 |
| WGI_GovtEff | −0.239*** | −0.100 | −0.073 | | | |
| (log) GDP/capita | | | | −1.01*** | −0.533** | −0.634*** |
| Ambition level | | −0.395 | −0.392 | | −0.354 | −0.356 |
| Eastern Europe | | | 0.44 | | | −0.053 |
| $R^2$ | 0.149 | 0.365 | 0.366 | 0.229 | 0.393 | 0.394 |
| N | 92 | 92 | 92 | 92 | 92 | 92 |

*Coefficient is significant at the 10% level

**Coefficient is significant at the 5% level

***Coefficient is significant at the 1% level

Standard errors are clustered on states

**Table 3** Logistic regressions. Dependent: Compliance (dichotomous)

| | Model 7 | Model 8 | Model 9 | Model 10 | Model 11 | Model 12 |
|---|---|---|---|---|---|---|
| Constant | 3.807 | 5.297 | 4.786 | 22.83 | 20.28 | 10.89 |
| WGI_GovtEff | −1.760*** | −1.571*** | −1.302 | | | |
| (log) GDP/capita | | | | −4.789** | −3.83* | −1.83 |
| Ambition level | | −1.288** | −1.285** | | −1.32 | −1.28** |
| Eastern Europe | | | 0.471 | | | 1.34 |
| Pseudo–$R^2$ | 0.143 | 0.205 | 0.206 | 0.115 | 0.175 | 0.193 |
| N | 92 | 92 | 92 | 92 | 92 | 92 |

*Coefficient is significant at the 10% level

**Coefficient is significant at the 5% level

***Coefficient is significant at the 1% level

Standard errors are clustered on states

## 6 Can the temporal dimension explain noncompliance?

This section reviews compliance-relevant changes from 1999 (when Gothenburg was adopted) to 2010 (Gothenburg's deadline year) and asks if they were sufficiently significant to explain instances of noncompliance with the Gothenburg Protocol.

### 6.1 Can energy consumption developments explain noncompliance with $NO_X$ targets?

Amann et al. (1999)[22] identify population size, GDP per capita, the number of vehicles, and energy consumption as major determinants of $NO_X$ emissions. However, because GDP

[22] This report was written by scientists at the International Institute for Applied Systems Analysis (IIASA) to make the scientific background for Gothenburg's commitments available to the wider public. Projections and other analyses from IIASA are considered as important inputs in the process deciding emissions targets (Castells and Ravetz 2001; Rensvik 2017; Tuinstra 2008).

**Table 4** Energy consumption and $NO_X$ noncompliance (all numbers in 1000 tonnes of $NO_X$)

| Party | Emissions attributable to the difference between projected and observed 2010 energy consumption | Noncompliance | Noncompliance minus emissions attributable to the difference between projected and observed 2010 energy consumption |
|---|---|---|---|
| Belgium | −3.26 | 39.7 | 42.96 |
| Denmark | 5.6 | 1.8 | −3.8 |
| Germany | −57.8 | 241.9 | 299.7 |
| Spain | 1.6 | 136.9 | 135.3 |
| France | −43.6 | 220.3 | 263.9 |
| Luxembourg | 9.3 | 35.2 | 25.9 |
| Netherlands | −14.1 | 9.9 | 23.99 |
| Sweden | 1.3 | 13.4 | 12.1 |
| Norway | 30.8 | 28.3 | −2.5 |

A state's noncompliance is calculated by subtracting its 2010 emissions target in the Gothenburg protocol from its observed 2010 emissions (as reported to UNECE in 2015)

per capita, population size, and the number of vehicles largely influence emissions through energy consumption, I examine only energy consumption developments.[23]

Evidence presented in Table 4 may be used to examine (1) whether the observed 2010 energy consumption deviated from projections, and (2) if such unexpected developments can explain the observed $NO_X$ noncompliance. First, I calculate the difference between projected (from Amann et al. 1999) and observed[24] 2010 energy consumption for each noncompliant party (results not reported here). Second, I multiply that difference with the $NO_X$ emissions per energy unit consumed in 2010, thereby calculating the amount of $NO_X$ emissions the unexpected energy consumption development can account for. Finally, I subtract that amount from each state's noncompliance.

The result is presented in the far-right column of Table 4 ("Emissions attributable to the difference between projected and observed 2010 energy consumption"). Values below zero suggest that unexpectedly high energy consumption can fully explain the instance of noncompliance under consideration. Values above zero indicate that it cannot. For instance, unexpectedly high energy consumption accounts for 1300 tonnes of Sweden's $NO_X$ emissions in 2010. However, Sweden's 2010 $NO_X$ emissions were 13,400 tonnes above the target. At best, therefore, unexpectedly high energy consumption explains only a small fraction of Sweden's noncompliance with its $NO_X$ target. In contrast, Denmark's unexpectedly high energy consumption accounts for 5600 tonnes of $NO_X$, thereby outweighing its noncompliance of 1800 tonnes. I therefore conclude that deviance between projected and observed 2010 energy consumption explains Denmark's $NO_X$ noncompliance.

Thus, this analysis suggests that only two of nine $NO_X$ noncompliance cases may be explained by unexpectedly high energy consumption (the second case being Norway).

[23] Since emission coefficients vary considerably among sources of energy, *aggregate* energy consumption is not my first-best data. However, this is the only projection concerning energy consumption included by Amann et al. (1999).

[24] Based on data from Eurostat (2017).

### 6.2 Previous underestimation of $NO_X$ emissions

If national authorities wrongfully believe that they are on an emissions trajectory consistent with compliance (or that compliance already has been reached), they may not commission policies that otherwise would have been put in place. Hence, underestimation of emissions may be a barrier to compliance.

Over the last couple of decades, it has been discovered repeatedly that diesel vehicles emit more $NO_X$ than previously thought (UNECE 2003, European Commission 2017). For instance, the minutes from a 2003 meeting in a CLRTAP[25] science and advisory body state that "the Task Force had noted that several countries were reviewing $NO_X$ emission data from heavy-duty vehicles (HDVs).[26] The findings seemed to suggest that $NO_X$ emissions from HDVs following the EURO 2 and 3 specifications were in reality higher than assumed in previous estimates" (UNECE 2003). The EURO 2 and 3 standards are two of a series of European Union (EU) emissions standards for road vehicles. The actual $NO_X$ emissions of diesel vehicles have been found to exceed the limits set by several of these standards (European Commission 2017).

If diesel vehicle noncompliance with EU standards (or other sources of incorrect estimation) has misled national authorities, it must have done so by deflating estimates of aggregate national emissions. Table 5 shows how estimates of the 1999 $NO_X$ emissions (as reported by national authorities to UNECE) of noncompliant states have varied over time.[27] These data allow comparison of *what states believed* were their $NO_X$ emissions in a given year to *what the actual*[28] *emissions were.* Additionally, the appendix includes estimates of 2003, 2006, 2008 and 2010 emissions over time.

In 2015, Belgium's 1999 $NO_X$ emissions were estimated at 312,700 tonnes. In 2001–2007, the Belgian 1999 emissions were reported at 292,000 tonnes—20,700 tonnes below the 2015 estimate. 20,700 tonnes equal roughly 50% of Belgium's noncompliance of 39,700 tonnes. However, the 2008 estimate of the 1999 emissions was only 400 tonnes below the 2015 estimate, suggesting that Belgian authorities became aware of the "real" $NO_X$ emissions in time to introduce additional policies. Moreover, Table 11 (see the Appendix) shows that Belgium's 2003 emissions were consistently *overestimated*. Thus, it seems that variations in emissions estimates can explain little (perhaps even nothing) of Belgium's noncompliance.

Denmark's 1999 emissions were consistently underestimated until 2010 (varying from 9200 to 26,800 tonnes below the 2015 estimate). Since Denmark's noncompliance equalled only 1800 tonnes, it seems fair to conclude that underestimation of emissions may explain Denmark's $NO_X$ noncompliance.

France's 1999 $NO_X$ emissions estimates have varied considerably. Table 5 shows that the estimates from 2005 to 2006 are approximately 191,000 tonnes below the 2015 estimate, a difference that corresponds to approximately 87% of France's noncompliance of 220,300 tonnes. However, in 2008, the estimates were only 45,300 tonnes below the estimate from 2015. As in the case of Belgium, it thus seems that the knowledge of the "actual" $NO_X$

[25] Except for Spain, all parties that did not reach their NOx targets were represented by national experts at the meeting.

[26] These vehicles often use diesel fuel.

[27] I include only states that were noncompliant with their 2010 NOx Gothenburg target.

[28] Actual emissions are here defined as the estimates reported in 2015. Obviously, there is an artificiality to this classification, since even recent emissions estimates may subject to change because of new scientific evidence. However, since estimates from 2015 are derived from the presently best available scientific knowledge, I use 2015 estimates as baseline.

**Table 5** Estimates of 1999 $NO_X$ emissions of noncompliant states (thousand tonnes)

| Party | 2001 | 2003 | 2005 | 2006 | 2007 | 2008 | 2010 | 2015 |
|---|---|---|---|---|---|---|---|---|
| Belgium | 292 | 292 | 292 | 292 | 292 | 312.3 | 324.1 | 312.7 |
| | *20.7* | *20.7* | *20.7* | *20.7* | *20.7* | *0.4* | *− 11.4* | |
| Denmark | 210.2 | 227.8 | 225.4 | 215.1 | 222.5 | 220.9 | 215.9 | 237 |
| | *26.8* | *9.2* | *11.6* | *21.9* | *14.5* | *16.1* | *21.1* | |
| France | 1530 | 1516.9 | 1462.3 | 1462.4 | 1473.4 | 1608 | 1675.5 | 1653.3 |
| | *123.3* | *136.4* | *191* | *190.9* | *179.9* | *45.3* | *− 22.2* | |
| Germany | 1637 | 1619 | 1717.5 | 1915.7 | 1913 | 1887.6 | 1914.5 | 1981 |
| | *344* | *362* | *263.5* | *65.3* | *68* | *93.4* | *66.5* | |
| Luxembourg | 16.1 | 16.1 | 16.1 | 16.1 | 16.1 | 16.1 | 17 | 37 |
| | *20.9* | *20.9* | *20.9* | *20.9* | *20.9* | *20.9* | *20* | |
| Netherlands | 408 | 429.2 | 429.2 | 429.2 | 410.4 | 389.6 | 404.1 | 413 |
| | *5* | *− 16.2* | *− 16.2* | *− 16.2* | *2.6* | *23.4* | *8.9* | |
| Norway | 230 | 237.7 | 238 | 238 | 228.7 | 223.3 | 215.5 | 213.7 |
| | | *− 24* | *− 24.3* | *− 24.3* | *− 15* | *− 9.6* | *− 1.8* | |
| Slovenia | 58 | 58 | 58 | 58 | 58 | 58 | 49.3 | 52.1 |
| | | *− 5.9* | *− 5.9* | *− 5.9* | *− 5.9* | *− 5.9* | *2.8* | |
| Spain | N/A | 1412.3 | 1446.9 | 1431.4 | 1437 | 1440.2 | 1372.7 | 1385.1 |
| | | *− 27.2* | *− 61.8* | *− 46.3* | *− 51.9* | *− 55.1* | *12.4* | |
| Sweden | 261 | 258.6 | 231.6 | 230.1 | 241 | 242 | 222.7 | 214.9 |
| | *− 46.1* | *− 43.7* | *− 16.7* | *− 15.2* | *− 26.1* | *− 27.1* | *− 7.8* | |

Numbers in italics are the differences between the 2015 estimate and the estimate from the year at the column header

emissions came early enough to enable the French authorities to avoid a substantial part of its noncompliance. Furthermore, the estimates from 2001 to 2003 are closer to the "correct" estimate than those of 2005, 2006, and 2007. Considering France's unexpectedly low energy consumption (Table 4), it seems reasonable to conclude that the temporal dimension can explain a significant share of France's noncompliance, yet far from all of it.

The 2003 estimate of Germany's 1999 $NO_X$ emissions is 362,000 tonnes lower than the 2015 estimate. This gap outweighs the total German noncompliance (241,900 tonnes). However, the estimate of 1999 emissions increased significantly already in 2006, and was then only 65,300 tonnes below the 2015 estimate, a total that amounts to 27% of the total noncompliance, and the 2010 deadline was still 4 years away. Nonetheless, that Germany's 1999 emissions were somewhat underestimated in every year from 2001 to 2010 suggests that underestimation may explain part of Germany's noncompliance.

Luxembourg stands out in terms of the relative size of the deviance between recent and older emissions estimates. The 2015 estimate of the 1999 emissions is between 20,000 and 20,900 tonnes higher than the estimates from 2001 through 2010. These gaps are smaller than Luxembourg's 2010 noncompliance (35,200 tonnes), even when we consider the unexpectedly high energy consumption (Table 4). However, Table 11 (Appendix) shows that the underestimation of Luxembourg's 2003 emissions consistently exceeds 30,000 tonnes. Hence, unexpectedly high energy consumption and increased emissions estimates may account for all of Luxembourg's noncompliance.

The various estimates of the Netherlands' 1999 emissions are fairly consistent, varying from slightly below to somewhat above the estimate from 2015, thereby suggesting that Dutch authorities have not been misguided by underestimations.

In the cases of Norway and Sweden, Table 5 suggests that underestimation of emissions cannot explain noncompliance, since their 1999 emissions was consistently overestimated from 2001 through 2010. The same conclusion holds for Spain and Slovenia, since their estimates from 2015 are lower than all other estimates, except those from 2010.

The appendix includes tables with estimates of $NO_X$ emissions for other years than 1999. Except from the cases of Belgium and Luxembourg (discussed above), Tables 11, 12, 13 and 14 lead to the same conclusions as Table 5.

### 6.3 Change in the drivers of NMVOC emissions

As Tables 1 and 6 show, three states have not complied with their NMVOC targets—Belgium, Denmark, and Germany. The number of registered vehicles is the most important driver of NMVOC emissions for which Amann et al. (1999) include projections.

The entries in Table 6 were arrived at in a manner similar to that used for Table 4. First, I find the difference between the projected number of vehicles from Amann et al. (1999) and the observed number of vehicles (from European Commission 2012). Next, I calculate the average NMVOC emissions per vehicle in 2010. By multiplying the gap between observed and projected vehicle numbers by the average NMVOC emissions per vehicle, I derive the numbers shown in the second column from the left in Table 6.

As shown by the far-right column in Table 6, none of the three instances of NMVOC noncompliance can be explained by the temporal dimension. Belgium and Germany had fewer vehicles in 2010 than projected, and Denmark's noncompliance (40,400 tonnes) far exceeds the emissions attributable to unexpectedly high vehicle numbers.

### 6.4 Change in the drivers of ammonia emissions

Table 7 shows projected (from Amann et al. 1999) and observed fertilizer use in 2010 of the eight parties that did not comply with their 2010 ammonia targets. Except for Switzerland, all parties consumed less nitrogen fertilizer in 2010 than projected. Thus, the evidence suggests that the temporal dimension cannot explain these cases of noncompliance. In contrast, for Switzerland, the observed consumption exceeds the projection by almost 70%, and Switzerland's ammonia emissions were only 1% above the target (see also Table 1). Thus, the temporal dimension appears to be a plausible explanation of Switzerland's noncompliance.

**Table 6** Road transport vehicles and NMVOC compliance (thousand tonnes)

| Party | Emissions attributable to the difference between projected and observed 2010 vehicle numbers | Noncompliance | Noncompliance minus emissions attributable to the difference between projected and observed 2010 vehicle numbers |
|---|---|---|---|
| Belgium | − 0.54 | 11.4 | 11.94 |
| Denmark | 2.3 | 40.4 | 38.1 |
| Germany | − 12.67 | 2343.8 | 2356.47 |

States' noncompliance is calculated using their 2010 emissions as reported in 2015 (CEIP 2015). Road transport data from CEIP 2016

**Table 7** Projected and observed 2010 fertilizer use (thousand tonnes)

| Party | Projected | Observed |
|---|---|---|
| Croatia | 190 | 117.4 |
| Denmark | 261 | 187.1 |
| Finland | 180 | 151.3 |
| Germany | 1801 | 1499.1 |
| Netherlands | 291 | 219.5 |
| Norway | 92 | 85.4 |
| Spain | 1052 | 941 |
| Switzerland | 30 | 50.8 |

Data on observed nitrogen fertilizer consumption from EEA 2012, except Norway, Switzerland, and Croatia (from FAO 2016)

**Table 8** Summary of findings concerning the temporal dimension

| Substance | Noncompliant party | Can changed conditions explain (some of) the noncompliance? |
|---|---|---|
| $NO_X$ | Belgium | No |
| | Denmark | Yes |
| | France | Some, yet far from all |
| | Germany | Some, yet most of the noncompliance is unaccounted for |
| | Luxembourg | Yes |
| | Netherlands | No |
| | Norway | Yes |
| | Slovenia | No |
| | Spain | No |
| | Sweden | No |
| NMVOC | Belgium | No |
| | Denmark | No |
| | Germany | No |
| Ammonia | Croatia | No |
| | Denmark | No |
| | Finland | No |
| | Germany | No |
| | Netherlands | No |
| | Norway | No |
| | Spain | No |
| | Switzerland | Yes |

## 6.5 Conclusion: The temporal dimension's explanatory power

Table 8 summarizes my conclusions concerning the temporal dimension's ability to explain the noncompliance with the Gothenburg Protocol. Of the 21 targets that were not complied with, four are fully explained by unexpected developments between Gothenburg's adoption

and deadline. Some of the noncompliance of two large $NO_X$ emitters, France and Germany, is explained, yet five other cases of noncompliance with $NO_X$ targets are not accounted for at all. Thus, although Chayes and Chayes' third factor has more explanatory power than the first two, it leaves most of the observed noncompliance unaccounted for.

## 7 Conclusion

This article has demonstrated that the three factors specified by Chayes and Chayes cannot explain much of the noncompliance with the Gothenburg Protocol.

The evidence examined to test the ambiguity explanation is clear: Doubt or disagreement over obligations has not caused the quite widespread noncompliance with the Gothenburg Protocol. The analysis of the capacity explanation is also unambiguous, as the hypothesis derived from Chayes and Chayes received no support in a series of regressions under various conditions.

The findings are somewhat less clear concerning the temporal dimension. Worsened conditions may *fully* explain four cases of noncompliance, and two cases *partly*. However, the majority of the cases are far from being explained by such unexpected developments. Overall, then, the management school does not provide good explanations for the noncompliance with the Gothenburg Protocol.

My framework for assessing the management school's explanations should prove useful for future compliance studies. Of the rather few existing examinations of the capacity-compliance relationship, few or none have used operationalizations that allow studies of *degrees of (non)compliance.* The present paper demonstrates that this indeed is possible and desirable. I have also shown how projections may be used to assess whether reaching compliance proved to be more difficult than expected when member states entered the agreement. To my knowledge, this is the first systematic and rigorous assessment beyond a single-case study of Chayes and Chayes' third explanation of noncompliance.

Given the limited explanatory power of the management school, do the data I have presented suggest any alternative explanations? Because the enforcement school does not expect states to deviate from BAU, it expects no positive effect of capacity on compliance when ambition level is included as a control. That I do not find a positive effect of capacity on compliance is thus consistent with the enforcement school. A rigorous test of the enforcement school would, however, require large amounts of additional data on the costs and benefits of different emissions levels for all the states included in my dataset. Only if we can provide solid evidence suggesting that calculations of net private benefits can account for state actions, can we claim that the enforcement school provides a better explanation of (non)compliance than the management school does.

## Appendix

### Sensitivity check: statistical assessments of the effect of capacity on compliance

Table 9 shows the results of additional OLS regressions using another measure from Worldwide Government Indicators (WGI) to operationalize capacity. According to the World Bank's description, Regulatory Quality "reflects perceptions of the ability of the

**Table 9** OLS regressions, alternative capacity operationalization. Dependent: compliance

| | Model 13 | Model 14 | Model 15 |
|---|---|---|---|
| Constant | 0.612 | 0.884 | 0.791 |
| WGI_RegulatoryQuality | −0.329*** | −0.15 | −0.102 |
| Ambition level | | −0.408* | −0.399* |
| Eastern Europe | | | 0.581 |
| $R^2$ | 0.113 | 0.365 | 0.339 |
| N | 92 | 92 | 92 |

*Coefficient is significant at the 10% level

**Coefficient is significant at the 5% level

***Coefficient is significant at the 1% level

Standard errors are clustered on states

**Table 10** OLS regressions, incl. substance dummies. Dependent: compliance

| | Model 16 |
|---|---|
| Constant | 3.541 |
| (log) GDP/cap. | −0.56** |
| Ambition level | −0.415*** |
| Eastern Europe | −0.063 |
| $NO_X$ | −0.424*** |
| NMVOC | −0.369*** |
| Ammonia | −0.423*** |
| $R^2$ | 0.229 |
| *N* | 92 |

*Coefficient is significant at the 10% level

**Coefficient is significant at the 5% level

***Coefficient is significant at the 1% level

Standard errors are clustered on states

government to formulate and implement sound policies and regulations that permit and promote private sector development".

Once again, I find no positive relationship between capacity and compliance.

Table 10 shows the results of a final robustness check (Model 16). Here, I have included dummies for each substance that Gothenburg regulates. Again, the effect of capacity is negative and statistically significant. Since Model 16 includes dummies for all regulated substances except sulphur, the substance dummy estimates can be interpreted as the difference in compliance between the substance concerned and sulphur. As all Gothenburg parties complied with their sulphur targets (see Table 1 in the main document), it comes as no surprise that all substance dummy estimates shown in Model 16 are negative.[29]

[29] Yet another analysis shows that the estimate of capacity in Model 16 is not sensitive to operationalizing capacity as Government Effectiveness (see Tables 2 and 3). I have also run this full model using the dichotomous compliance variable, and the effect of capacity remains negative (not reported here, on file with author).

## Additional comparisons of emissions estimates over time

See Tables 11, 12, 13 and 14.

**Table 11** Estimates of 2003 $NO_X$ emissions of noncompliant parties (thousand tonnes)

| Party | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2015 |
|---|---|---|---|---|---|---|---|
| Belgium | 297.2 | 297.5 | 297.5 | 297.5 | 296.3 | 297 | 293.8 |
| | *−3.4* | *−3.7* | *−3.7* | *−3.7* | *−2.5* | *−3.2* | |
| Denmark | 207.8 | 197.9 | 210.3 | 208.3 | 203.7 | 203.9 | 225.8 |
| | *18* | *27.9* | *15.5* | *17.5* | *22.1* | *21.9* | |
| France | 1220.3 | 1244.1 | 1257.3 | 1450.1 | 1496.3 | 1529.4 | 1502.9 |
| | *282.6* | *258.8* | *245.6* | *52.8* | *6.6* | *−26.5* | |
| Germany | 1428 | 1604.7 | 1625.5 | 1580 | 1541.3 | 1613.8 | 1715.1 |
| | *287.1* | *110.4* | *89.6* | *135.1* | *173.8* | *101.3* | |
| Luxembourg | N/A | 17.5 | 17.5 | 17.5 | 16 | 16 | 47.5 |
| | | 30 | 30 | 30 | 31.5 | 31.5 | |
| Netherlands | 363.8 | 367.2 | 373.1 | 357.8 | 357.6 | 371.2 | 369.2 |
| | *5.4* | *2* | *−3.9* | *11.4* | *11.6* | *−2* | |
| Norway | 220.2 | 214.8 | 199.2 | 196.7 | 194 | 190.5 | 194.6 |
| | *−25.6* | *−20.2* | *−4.6* | *−2.1* | *0.6* | *4.1* | |
| Slovenia | 56 | 56 | 55.3 | 48.2 | 48.2 | 49.5 | 52.1 |
| | *−3.9* | *−3.9* | *−3.2* | *3.9* | *3.9* | *2.6* | |
| Spain | 1518.6 | 1493.2 | 1492.6 | 1500 | 1490.2 | 1401.4 | 1402 |
| | *−116.6* | *−91.2* | *−90.6* | *−98* | *−88.2* | *0.6* | |
| Sweden | 206 | 202.7 | 215.4 | 197.9 | 191.7 | 190 | 186 |
| | *−20* | *−16.7* | *−29.4* | *−11.9* | *−5.7* | *−4* | |

Numbers in italics are the differences between the 2015 estimate and the estimate from the year at the column header

**Table 12** Estimates of 2006 $NO_X$ emissions (thousand tonnes)

| | 2008 | 2009 | 2010 | 2015 |
|---|---|---|---|---|
| Belgium | 277.7 | 268.3 | 266.4 | 277.4 |
| | *−0.3* | *9.1* | *11* | |
| Denmark | 185.3 | 180.7 | 182.1 | 201.2 |
| | *15.9* | *20.5* | *19.1* | |
| France | 1351.2 | 1397.5 | 1414 | 1359.1 |
| | *7.9* | *−38.4* | *−54.9* | |
| Germany | 1394.3 | 1353.9 | 1520.5 | 1557.1 |
| | 162.8 | 203.2 | 36.6 | |
| Luxembourg | N/A | 14.4 | 14.4 | 53.6 |
| | | *39.2* | *39.2* | |
| Netherlands | 310.8 | 307.2 | 324.1 | 327.2 |
| | *16.4* | *20* | *3.1* | |
| Norway | 190.8 | 198 | 185.5 | 194.3 |

**Table 12** (continued)

| | 2008 | 2009 | 2010 | 2015 |
|---|---|---|---|---|
| | *3.5* | *– 3.7* | *8.8* | |
| Slovenia | 46.9 | 46.8 | 46 | 50.4 |
| | *3.5* | *3.6* | *4.4* | |
| Spain | 1481.2 | 1465.1 | 1401 | 1366.9 |
| | *– 114.3* | *– 98.2* | *– 34.1* | |
| Sweden | 174 | 170.4 | 169.1 | 172.2 |
| | *– 1.8* | *1.8* | *3.1* | |

Numbers in italics are the differences between the 2015 estimate and the estimate from the column header

**Table 13** Estimations of 2008 $NO_X$ emissions (thousand tonnes)

| Party | 2010 | 2012 | 2015 |
|---|---|---|---|
| Belgium | 240.5 | 238.7 | 236.5 |
| | *– 4* | *– 2.2* | |
| Denmark | 151.7 | 150.5 | 170.1 |
| | *18.4* | *19.6* | |
| France | 1272.5 | 1194.4 | 1197.6 |
| | *– 74.9* | *3.2* | |
| Germany | 1393.3 | 1417.5 | 1410.8 |
| | *17.5* | *– 6.7* | |
| Luxembourg | N/A | 50.2 | 44.9 |
| | | *– 5.3* | |
| Netherlands | 292.7 | 308.9 | 299.3 |
| | *6.6* | *– 9.6* | |
| Norway | 173.7 | 189.1 | 185.2 |
| | *11.5* | *– 3.9* | |
| Slovenia | 52.9 | 53.1 | 55.5 |
| | *2.6* | *2.4* | |
| Spain | 1236.3 | 1175.6 | 1170.9 |
| | *– 65.4* | *– 4.7* | |
| Sweden | 154.4 | 158 | 157 |
| | *2.6* | *– 1* | |

Numbers in italics are the differences between the 2015 estimate and the estimate from the year at the column header

**Table 14** Estimates of 2010 $NO_X$ emissions (thousand tonnes)

| Party | 2012 | 2015 |
|---|---|---|
| Belgium | 220.7 | 252.3 |
| | *31.6* | |
| Denmark | 128.8 | 145.3 |
| | *16.5* | |
| France | 1080.3 | 1096.4 |
| | *16.1* | |
| Germany | 1322.9 | 1333.7 |
| | *10.8* | |
| Luxembourg | 46.2 | 39.45 |
| | *−6.75* | |
| Netherlands | 275.9 | 274.2 |
| | *−1.7* | |
| Norway | 184.3 | 177.2 |
| | *−7.1* | |
| Slovenia | 44.7 | 47.1 |
| | *2.4* | |
| Spain | 983.9 | 959.7 |
| | *−24.2* | |
| Sweden | 161.4 | 149.6 |
| | *−11.8* | |

Numbers in italics are the differences between the 2015 and 2010 estimates

## References


Aakre, S., Helland, L., & Hovi, J. (2016). When does informal enforcement work? *Journal of Conflict Resolution, 60*(7), 1312–1340. https://doi.org/10.1177/0022002714560349.

Amann, M., Bertok, I., Cofala, J., Gyarfas, F., Heyes, C., Klimont, Z., et al. (1999). *Integrated assessment modelling for the protocol to abate acidification, eutrophication and ground-level ozone in Europe*. Haag: Ministry of Housing, Spatial Planning and the Environment.

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. London: Princeton University Press.

Barrett, S. (2003). *Environment and statecraft. The strategy of environmental treaty-making*. Oxford: Oxford University Press.

Bernauer, T., Kalbhenn, A., Koubi, V., & Spilker, G. (2013). "Is there a 'depth versus participation' dilemma in international cooperation? *The Review of International Organizations, 8*(4), 477–497. https://doi.org/10.1007/s11558-013-9165-1.

Böhmelt, T., & Vollenweider, J. (2015). Information flows and social capital through linkages: The effectiveness of the CLRTAP network. *International Environmental Agreements: Politics, Law and Economics, 15*(2), 105–123. https://doi.org/10.1007/s10784-013-9218-1.

Börzel, T. A., Hofmann, T., Panke, D., & Sprungk, C. (2010). Obstinate and inefficient: Why member states do not comply with European Law. *Comparative Political Studies, 43*(11), 1363–1390. https://doi.org/10.1177/0010414010376910.

Bratberg, E., Tjøtta, S., & Øines, T. (2005). Do voluntary international environmental agreements work? *Journal of Environmental Economics and Management, 50*(3), 583–597. https://doi.org/10.1016/j.jeem.2005.03.002.

Breitmeier, H., Young, O. R., & Zürn, Michael. (2006). *Analyzing international environmental regimes. From case study to database*. London: MIT Press.

Brown Weiss, E., & Jacobson, H. K. (Eds.). (1998). *Engaging countries: Strengthening compliance with international environmental accords*. Cambridge: MIT Press.

Castells, N., & Ravetz, J. (2001). Science and policy in international environmental agreements: Lessons from the European experience on transboundary air pollution. *International Environmental Agreements: Politics, Law and Economics, 1*(1), 405–425. https://doi.org/10.1023/A:1013322222903.

CEIP. 2015. *Trend tables 2015*. At http://www.ceip.at/ms/ceip_home1/ceip_home/status_reporting/2015_submissions/. Last Accessed 14 Feb 2018.

CEIP. 2016. *Officially reported emissions data. Road transport*. At http://webdab1.umweltbundesamt.at/cgi-bin/wedb2_controller.pl?pollutants=NMVOC&countries=%3AALL&State=official&years=2010&sectordefinitions=NFR2014_L1&reportyear=2015&sectors=556&datatype=national_html&database=web&horizontal=area&vertical=year&unit=Gg&info=on. Accessed 14 Feb 2018.

Chayes, A., & Chayes, A. H. (1993). On compliance. *International Organization, 47*(2), 175–205. https://doi.org/10.1017/S0020818300027910.

Chayes, A., & Chayes, A. H. (1995). *The new sovereignty*. London: Harvard University Press.

Department for Environment, Food and Rural Affairs. 2015. *Emissions of air quality pollutants*. At http://naei.defra.gov.uk/documents/AQPI_Summary_1990-2013_Issue_v1.1.pdf. Accessed 15 Feb 2018. Also on file with author.

Downs, G. W., Rocke, D. M., & Barsoom, P. N. (1996). Is the good news about compliance good news about cooperation? *International Organization, 50*(3), 379–406. https://doi.org/10.1017/S0020818300033427.

European Commission. 2012. *EU transport in figures: Statistical pocket book 2012*.

European Commission. 2017. *Transport emissions: Air pollutants from road transport*. At https://www.eea.europa.eu/data-and-maps/indicators/transport-emissions-of-air-pollutants-8/transport-emissions-of-air-pollutants-5. Last Accessed 12 Feb 2018.

Eurostat. 2017. *Simplified energy balances—annual data* At http://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do. Accessed 26 Apr 2017.

FAO. 2016. *Country profiles*. At http://www.fao.org/countryprofiles/en/. Accessed 14 Feb 2017.

Finnemore, M., & Sikkink, K. (1998). International norm dynamics and political change. *International Organization, 52*(4), 887–917.

Hanson, J. K., & Sigman, R. 2013. *Leviathan's latent dimensions: Measuring state capacity for comparative political research*. APSA 2011 Annual Meeting Paper. Available at SSRN: https://ssrn.com/abstract=1899933.

Helm, C., & Sprinz, D. (2000). Measuring the effectiveness of international environmental regimes. *Journal of Conflict Resolution, 44*(5), 630–652. https://doi.org/10.1177/0022002700044005004.

Henkin, L. (1968). *How nations behave: Law and foreign policy*. London: Pall Mall Press.

Hovi, J., Sprinz, D. F., & Underdal, A. (2003). The Oslo-Potsdam solution to measuring regime effectiveness: Critique, response, and the road ahead. *Global Environmental Politics, 3*(3), 74–96.

Jacobson, H. K., & Brown Weiss, E. (1998). Assessing the record and designing strategies to engage countries. In E. Brown Weiss & H. K. Jacobson (Eds.), *Engaging countries: Strengthening compliance with international environmental accords* (pp. 511–554). Cambridge: MIT Press.

Jänicke, M. (1997). The political system's capacity for environmental policy. In M. Jänicke (Eds.), *National environmental policies. A comparative study of capacity-building* (pp. 1–24). Berlin: Springer.

Kelly, A., Lumbreras, J., Maas, R., Pignatelli, T., Ferreira, F., & Engleryd, A. (2010). Setting national emission ceilings for air pollutants: Policy lessons from an ex-post evaluation of the Gothenburg Protocol. *Environmental Science & Policy*. https://doi.org/10.1016/j.envsci.2009.09.003.

Kokkvoll Tveit, A. (2018). Norms, incentives, or deadlines? Explaining Norway's noncompliance with the gothenburg protocol. *Global Environmental Politics, 18*(1), 76–98.

Levy, M. A. (1993). European acid rain: The power of tote-board diplomacy. In P. M. Haas, R. O. Keohane & M. Levy (Eds.), *Institutions for the earth. Sources of effective international environmental protection* (pp. 75–132). London: MIT Press.

Miles, E. L., Underdal, A., Andresen, S., Wettestad, J., Skjærseth, J. B., & Carlin, E. M. (2002). *Environmental regime effectiveness. Confronting theory with evidence*. London: MIT Press.

Miljødirektoratet. 2015. *Air pollution*. At http://www.environment.no/topics/air-pollution/. Accessed 14 Feb 2017.

Miljøstyrelsen. 2002. *Indikatorrapport*. At http://www2.mst.dk/common/Udgivramme/Frame.asp?. http://www2.mst.dk/udgiv/publikationer/2002/87-7972-277-6/html/kap06.htm. Accessed 3 Apr 2017. Also on file with author.

Mitchell, R. B. (1994). Regime design matters: Intentional oil pollution and treaty compliance. *International Organization, 48*(3), 425–458. https://doi.org/10.1017/S0020818300028253.
Mitchell, R. B. (2010). *International politics and the environment*. London: SAGE Publishing.
Naturvårdsverket. 2016. *FN:s luftvårdskonvention (CLRTAP)*. At http://www.naturvardsverket.se/Miljoarbete-i-samhallet/EU-och-internationellt/Internationellt-miljoarbete/miljokonventioner/Luftvard/. Accessed 14 Feb 2018.
Protocol to The 1979 Convention on Long-Range Transboundary Air Pollution to Abate Acidification, Eutrophication and Ground-level Ozone At www.unece.org/fileadmin/DAM/env/lrtap/full%20text/1999%20Multi.E.Amended.2005.pdf. Last Accessed 14 Feb 2018.
Raustiala, K. (2005). Form and substance in international agreements. *American Journal of International Law, 99*(3), 581–614.
Raustiala, K., & Slaughter, A.-M. 2002. International law, international relations, and compliance. In W. Carlnaes, T. Risse, B. Simmons, (Eds.), *The handbook of international relations*, London: Sage Publications. Also available at SSRN: https://ssrn.com/abstract=347260.
Raustiala, K., & Victor, D. G. (1998). Conclusions. In K. Raustiala, D. G. Victor (Eds.), *The implementation and effectiveness of international environmental commitments: Theory and practice*. IIASA: MIT Press.
Rensvik, H. 2017. Author's interview with Harald Rensvik, Secretary General in Norway's Ministry of the Environment 1996–2011, Oslo, January 2017.
Ringquist, E. J., & Kostadinova, T. (2005). Assessing the effectiveness of international environmental agreements: The case of the 1985 Helsinki protocol. *American Journal of Political Science, 49,* 86–102. https://doi.org/10.1111/j.0092-5853.2005.00112.x.
Simmons, B. A. (1998). Compliance with international agreements. *The Annual Review of Political Science, 1,* 75–93.
Simmons, B. A. (2013). From ratification to compliance: Quantitative evidence on the spiral model. In T. Risse, S. C. Ropp & K. Sikkink (Eds.), *The persistent power of human rights: From commitment to compliance* (pp. 43–59). Cambridge: Cambridge University Press.
Tuinstra, W. 2008. "European air pollution assessments: co-production of science and policy," *International Environmental Agreements: Politics, Law and Economics 8*(35). https://doi.org/10.1007/s10784-008-9064-8.
Umweltbundesamt. 2017. $NO_X$ *emissions.* At http://www.umweltbundesamt.de/daten/luftbelastung/luftschadstoff-emissionen-in-deutschland/stickstoffoxid-emissionen#textpart-1. Accessed 16 Apr 2017.
Underdal, A. 2000. Comparative analysis: accounting for variance in actor behaviour. Chapter 13. In A. Underdal & K. Hanf (Eds.), *International environmental agreements and domestic politics.* Aldershot: Ashgate.
UNECE. 2003. *Integrated assessment modelling*" At www.unece.org/fileadmin/DAM/env/documents/2003/eb/wg5/eb.air.wg.5.2003.5.e.pdf. Accessed 19 Apr 2017.
Victor, D., Raustiala, K., & Skolnikoff, E. B. (Eds.). (1998). *The implementation and effectiveness of international environmental commitments: Theory and practice*. London: MIT Press.
Wettestad, J. (2012). Reducing long-range transport of air pollutants in Europe. In S. Andresen, E. L. Boasson & J. Wettestad (Eds.), *International environmental agreements* (pp. 23–37). New York: Routledge.
World Bank. 2017. *Worldwide Governance Indicators*. At http://info.worldbank.org/governance/wgi/index.aspx#home. Accessed 20 Apr 2017.
Young, O. R. (1979). *Compliance and public authority: A theory with international applications*. London: Johns Hopkins University Press.
Young, O. R. (2003). Determining regime effectiveness: A commentary on the Oslo-Potsdam solution. *Global Environmental Politics, 3*(3), 97–104.

# Norms, Incentives, or Deadlines? Explaining Norway's Noncompliance with the Gothenburg Protocol

*Andreas Kokkvoll Tveit**

## Abstract

Norway, previously an international frontrunner concerning reductions of transboundary air pollution, fell far short of its 2010 target for nitrogen oxides ($NO_x$) under the 1999 Gothenburg Protocol. In this article I show that leading international compliance theories cannot explain much of this noncompliance. While little evidence supports the management school's explanations, Norwegian policies are also inconsistent with the enforcement school. Albeit too late to meet the deadline, Norway imposed a $NO_x$ tax in 2007. Moreover, the resulting emissions reductions were deeper than in a business-as-usual scenario, despite no international enforcement. That the $NO_x$ tax was imposed only after an environmentalist party gained considerable influence over $NO_x$ policies in 2005 supports an office-incumbent hypothesis. However, as emissions also declined significantly in many other European countries after 2005, the explanation is likely structural. One possibility is the deadline-pressure hypothesis: As the deadline approached, decision-makers across Northern and Western Europe considered emissions reductions to be more urgent than before.

When scientific and public awareness of acid rain rose in the 1970s, Scandinavian countries served as front runners in international efforts to address the problem (Wettestad 2012, 25–26). As cooperation grew increasingly institutionalized over the following decades, Norway was eager to keep this position. Therefore it may seem surprising that Norway failed to comply with its 2010 target for nitrogen oxides ($NO_x$) under the 1999 Gothenburg Protocol.[1]

* Parts of the research were carried out at the Fritjof Nansen Institute, Lysaker, Norway. I thank Jon Hovi, participants at CICEP's annual research conference in September 2016, and three anonymous reviewers for many comments that helped improve the article substantially. Finally, my research has benefited greatly from my stay at the International Institute for Applied Systems Analysis in May 2016.

1. Protocol to the 1979 Convention on Long-Range Transboundary Air Pollution to Abate Acidification, Eutrophication, and Ground-Level Ozone, usually referred to as the 1999 Gothenburg Protocol. An amended Gothenburg Protocol was adopted in 2012, which included emissions targets for 2020. I study the original protocol from 1999, not the amended protocol. Hence, in the present article, the "Gothenburg Protocol" refers to the 1999 version.

*Global Environmental Politics* 18:1, February 2018, doi:10.1162/GLEP_a_00446

Using an in-depth case study, I aim to explain this noncompliance. I find that neither the enforcement school (Downs et al. 1996; Barrett 2003) nor the management school as formulated by Chayes and Chayes (1993, 1995; see also Chayes et al. 1995; Young 1979) explains Norway's noncompliance and $NO_x$ policies well. I thus turn to two alternative explanations, the *office-incumbent hypothesis* and a *deadline-pressure hypothesis*. I find that both of these alternative explanations are consistent with the Norwegian case; however, only the latter is also consistent with the emissions trajectories of other Gothenburg countries. Because the deadline-pressure hypothesis shares some features with both the management and enforcement schools, I argue that it may be viewed as a hybrid of the two.

This article contributes to the international compliance literature in four ways. First, it provides the first study of noncompliance with the Gothenburg Protocol. Because Norway is an important player in regional environmental cooperation to reduce long-range air pollution,[2] understanding the causes of its noncompliance should be interesting to scholars, to Norwegian authorities, and to other Gothenburg parties. In particular, findings concerning the Norwegian case may be useful for explaining other states' noncompliance with Gothenburg targets. Second, and perhaps surprisingly, few (if any) scholars have done what the present article aims to do: derive and empirically assess precise hypotheses concerning the causes of one particular case of noncompliance. Third, although sacrificing breadth, the present analysis is deeper than any previous study of compliance with a protocol under the 1979 Convention on Long-Range Transboundary Air Pollution to Abate Acidification, Eutrophication and Ground-Level Ozone (CLRTAP).[3] Finally, my assessment of hypotheses derived from the management and enforcement schools may contribute to theory development. Much of the compliance literature has revolved around these two schools. Generally, the more important a school or theory is to a field of research, the more interesting it is if hypotheses derived from it prove inconsistent with evidence.

The remainder of the article proceeds as follows. The next section describes the Gothenburg Protocol and its goals. The following section reviews relevant research and develops a set of hypotheses. Thereafter, I confront the management school's three explanations of noncompliance with empirical evidence for Norway. I find that neither ambiguity, incapacity, nor the "temporal dimension"[4] can account for Norway's noncompliance. Next, I show that Norwegian policies are also inconsistent with the enforcement school's expectations. Albeit too late to reach compliance by the 2010 deadline, a Norwegian $NO_x$ tax was commissioned in 2007. The ensuing emissions reductions were clearly deeper

2. Levy (1993, 16) describes Norway as a "hardcore environmentalist countr[y]."
3. Available online at https://tinyurl.com/ybg72xt6, last accessed November 27, 2017.
4. Changed conditions for compliance owing to social and economic changes between commitment and deadline.

than in a business-as-usual (BAU) scenario, despite no enforcement measures being in place to incentivize such deep reductions.

I then turn to two alternative theories. I show that some evidence suggests that Norway's $NO_x$ policies are consistent with an office-incumbent hypothesis. Despite being widely considered the most effective measure, an emissions tax was not introduced until after the 2005 elections, when an environmentalist party gained substantial influence over $NO_x$ policies. However, that several other Gothenburg parties conducted even larger $NO_x$ emissions reductions from 2007 onward suggests that we should look for a structural explanation rather than a particular one. One such structural explanation is the deadline-pressure hypothesis developed in this article: as the 2010 deadline came closer, the Gothenburg parties (including Norway) considered actions to cut emissions as increasingly urgent.

## The Gothenburg Protocol and Norway's $NO_x$ Target

The 1999 Gothenburg Protocol was the eighth CLRTAP protocol. Because transboundary air pollution is largely a regional problem, most parties to the convention are European states. These states chose the UN Economic Commission for Europe (UNECE) as the institutional foundation for the collaboration. Gothenburg seeks to solve three interconnected environmental problems. The first problem is acidification, which harms life in water and soil and is largely caused by sulfur and nitrogen oxide emissions. The second problem is eutrophication – enrichment of water by nutrients (such as agricultural emissions of ammonia, $NH_3$). Eutrophication may change ecosystems, for instance by increasing algae growth and depleting fish stocks. Harmful ground-level ozone is the third environmental challenge targeted by Gothenburg; ground-level ozone stems from $NO_x$ reacting with volatile organic compounds (VOCs).[5]

Gothenburg includes four emissions target for each state: one target each for $NO_x$,[6] sulfur,[7] VOCs, and ammonia. A majority of the national emissions targets in the Gothenburg Protocol were reached by the 2010 deadline (Table 1). However, twenty-one of ninety-two targets were not reached by 2010, ten of which were for $NO_x$ emissions, eight for ammonia, and three for VOCs. All $SO_2$ targets were reached by 2010.

Table 1 includes all twenty-three European countries that became parties to Gothenburg before the 2010 deadline. Thirteen of these countries failed to meet at least one target by the deadline.

Norway did not comply with its annual emissions target of 156,000 metric tons of $NO_x$ by 2010. Norway's 2010 emissions[8] were 177,200 metric tons—approximately

5. See http://www.environment.no/topics/air-pollution/, last accessed November 27, 2017.
6. $NO_x$ emissions targets are expressed as nitrogen dioxide ($NO_2$).
7. Sulphur emissions targets are expressed as sulphur dioxide ($SO_2$).
8. Unless otherwise stated, all emissions are in metric tons as reported to UNECE in 2015.

**Table 1**
Compliance With Targets (Deadline Year 2010) in the Gothenburg Protocol

| *Party* | *$NO_x$* | *VOC* | *Sulfur Dioxide* | *Ammonia* |
|---|---|---|---|---|
| Belgium | *139.0* | *107.9* | 57.1 | 88.0 |
| Bulgaria | 52.1 | 55.8 | 45.2 | 38.4 |
| Croatia | 74.0 | 61.0 | 49.6 | *129.4* |
| Cyprus | 80.0 | 71.3 | 56.3 | 62.2 |
| Czech Rep. | 77.0 | 78.4 | 56.6 | 67.1 |
| Denmark | *114.0* | *147.5* | 27.9 | *115.9* |
| Finland | 97.6 | 89.4 | 57.6 | *123.4* |
| France | *127.5* | 79.5 | 71.3 | 93.4 |
| Germany | *123.4* | *124.5* | 79.0 | *116.8* |
| Hungary | 77.8 | 91.3 | 5.7 | 86.0 |
| Latvia | 45.7 | 65.6 | 2.4 | 32.7 |
| Lithuania | 45.0 | 77.7 | 14.3 | 51.4 |
| Luxembourg | *358.7* | 94.3 | 43.9 | 67.5 |
| Netherlands | *103.1* | 82.7 | 68.2 | *112.3* |
| Norway | *113.6* | 71.6 | 89.5 | *119.3* |
| Portugal | 68.1 | 89.0 | 31.2 | 42.8 |
| Romania | 53.1 | 66.5 | 38.1 | 80.0 |
| Slovakia | 68.2 | 45.6 | 63.1 | 63.9 |
| Slovenia | *104.7* | 96.0 | 36.5 | 95.0 |
| Spain | *113.0* | 97.4 | 54.6 | *111.0* |
| Sweden | *101.1* | 79.5 | 47.7 | 90.6 |
| Switzerland | 98.3 | 62.6 | 46.7 | *101.0* |
| United Kingdom | 95.1 | 71.3 | 68.4 | 93.9 |

Emissions are for 2010 in percentage of targets. Targets that were not reached are in italics. Emissions data are from Centre on Emission Inventories and Projections (CEIP) trends tables for 2015, available online at https://tinyurl.com/y8zvadz5, last accessed March 27, 2017.

13.6 percent above the target (Table 2). Also, its $NO_x$ emissions were declining for most of the period after 1999 (Table 2). However, the downward trend became substantially steeper after 2007. The low 2009 emissions were likely caused by reduced economic activity during the financial crisis.[9]

9. See Statistics Norway, Nedgang i klimagassutslippene, men er det varig? Available online at https://tinyurl.com/yazhzc8k, last accessed November 27, 2017.

**Table 2**
Norwegian $NO_x$ Emissions, 1999–2013

| *Year* | *Emissions (metric tons)* |
|---|---|
| 1999 | 213,700 |
| 2000 | 202,000 |
| 2001 | 200,300 |
| 2002 | 195,100 |
| 2003 | 194,600 |
| 2004 | 195,700 |
| 2005 | 196,100 |
| 2006 | 194,300 |
| 2007 | 195,800 |
| 2008 | 185,200 |
| 2009 | 175,100 |
| 2010 | 177,200 |
| 2011 | 170,000 |
| 2012 | 163,100 |
| 2013 | 154,400 |

Data are from CEIP trends tables for 2015, available online at https://tinyurl.com/y8zvadz5, last accessed March 27, 2017.

## Previous Research and Hypotheses

In this section, I present the two main theoretical perspectives in the literature on international compliance and then some recent contributions. I also develop a set of hypotheses.

Granted, the enforcement and management schools are not the only explanatory perspectives I could have tested. For example, Franck (1988) argues that equitability is a crucial determinant of compliance (see also Breitmeier et al.'s 2006 "legitimacy" perspective as well as Kim et al. 2017). Nonetheless, Breitmeier et al.'s (2006, 110–111) summary of findings concerning compliance points specifically at the management and enforcement schools. They thereby suggest that those two schools have sparked more debate than other perspectives.

### *The Enforcement School*

According to the enforcement school's model, states act like unitary, rational actors. Evaluating its options according to its (private) costs and benefits, each

state chooses the action that maximizes its net (private) benefit (Aakre et al. 2016, 1317). Unless (marginal) abatement costs are outweighed by (marginal) abatement benefits, unilateral emissions reductions are economically irrational. Thus defection constitutes each state's *dominant strategy*: each state will be better off by not contributing to problem solving, regardless of other states' actions. If each state pursues this dominant strategy, the outcome entails suboptimal public goods provision.

Therefore, successful treaties restructure states' incentives, by ensuring credible punishment of noncompliers or rewards to compliant states. Because international third-party enforcement is rare, such credibility usually requires that other parties to the agreement have incentives to implement punishment or rewards. Hence prospects for solving malign collective action problems are gloomy.[10] Wettestad (2002, 205–208) characterizes long-range transboundary air pollution as a malign problem.

### *The Management School*

Managerialists argue that nothing inherent or structural in the international system warrants pessimism concerning cooperative efforts. The cornerstone of this reasoning is the claim that "states have a general propensity for compliance" that makes calculated, intentional noncompliance rare (Chayes and Chayes 1993, 175–178).[11] Thus enforcement measures are not only expensive and inefficient but also unnecessary. Managerialists advocate "softer" mechanisms, such as monitoring, increasing states' capacities, and sharing knowledge and information (Chayes et al. 1995, 84–85).

The alleged propensity to comply originates in interests (states negotiate and sign treaties aligned with their interests, and noncompliance means jeopardizing (1) your reputation as a reliable partner [Chayes and Chayes 1993, 177, 183–184]); (2) efficiency (constant recalculation of interests is inefficient, while acting in accordance with agreements reduces costs); and (3) international norms (in international relations, a core norm is to do as promised [Chayes and Chayes 1993, 185; Henkin 1968; Finnemore and Sikkink 1998]). Thus, violations of international agreements typically have causes beyond the noncompliant state's control: treaty ambiguity, insufficient state capacity, and what Chayes and Chayes (1993) label the "temporal dimension." According to the ambiguity explanation, legal documents may be open to different interpretations. The state capacity explanation argues that financial constraints or insufficient bureaucratic and technical competence may impede goal achievement even in wealthier states (Chayes and Chayes 1993, 194).

10. Political malignancy depends on asymmetries, cleavages, and "the "incentives of the underlying game" (Underdal 2002, 15–18; see also Mitchell 2006, 78).
11. Still, Chayes and Chayes (1993, 176) certainly do not deny that deliberate noncompliance sometimes occurs.

The temporal dimension contends that immediate compliance often cannot be expected because policies must be implemented and then acted on by polluters (Chayes and Chayes 1993, 195). During the time between commitments and their implementation, social and economic changes may alter the conditions for compliance (Aakre et al. 2016, 1317).

### *Two Decades of Empirical Research*

Since the mid-1990s, several large empirical studies have been conducted; however, the jury is still out concerning which theory has more explanatory power (Perkins and Neumayer 2007). Summarizing their analyses concerning compliance,[12] Breitmeier et al. (2006, 110–111) state that "neither the shallowness argument of Downs, Rocke and Barsoom (1996) nor the management school of Chayes and Chayes can explain patterns of compliance with international environmental regimes."

Neither do Jacobson and Brown Weiss (1998) provide a clear summary of which theoretical perspective gets more support from their study.[13] Nonetheless, both schools receive some support from different findings; their Figure 15.2 summarizes the findings of their case studies. It lists thirty variables that the authors "believe are the most important factors that affect compliance" (Jacobson and Brown Weiss 1998, 534–536). Among them are sanctions, in keeping with the enforcement school's expectations. Scholars of the enforcement camp would, however, not expect most of these factors (e.g., equity, reporting requirements, NGOs) to increase compliance with deep commitments absent enforcement.[14] Unfortunately, Jacobson and Brown Weiss (1998) do not systematically assess depth. They support the management school by suggesting that factors like administrative capacity and monitoring increase compliance. It is, however, difficult to draw clear conclusions concerning which school receives more support because Jacobson and Brown Weiss (1998) do little to distinguish between their thirty explanatory factors' relative importance.[15]

Although *implementation* and *effectiveness* are the main dependent variables in Victor et al.'s (1998) fourteen case studies of regimes, their findings also shed light on compliance. They argue, "We find that some implementation failures are intentional and that 'harder' measures, such as sanctions, are available and sometimes necessary" (Victor et al. 1998, x).

12. Breitmeier et al. (2006) studied the twenty-three international environmental regimes included in the International Regimes Database (Young and Zürn 2006).
13. Jacobson and Brown Weiss (1998) are the authors of the concluding chapter of *Engaging Countries* (Brown Weiss and Jacobson 1998, eds.), an anthology that includes studies of compliance by eight states and the EU with five international environmental treaties.
14. A commitment is deep to the extent that it requires a party to do more than it would do *in the absence of the commitment* (Downs et al. 1996, 382).
15. The exception is a statement that "the strength and health of national political-economic systems and a deep public commitment are the most important ingredients in compliance" (Jacobson and Brown Weiss 1998, 542).

Underdal and Hanf (2000) provide the most comprehensive study concerning CLRTAP protocols (not including Gothenburg). According to Underdal (2000, 351–353), a model of states as unitary rational actors predicts patterns of compliance (operationalized as emissions reductions), negotiation positions, and implementation reasonably well.

### *Hypotheses and Research Design*

Given Norway's noncompliance, the management school would expect the Gothenburg Protocol to be ambiguous, Norway's capacity to be inadequate, or time to have been too short to enable Norwegian compliance. I assess the ambiguity explanation by asking if there has been any doubt concerning what Gothenburg obliges Norway to do. If not, the ambiguity explanation is unable to account for Norway's noncompliance. Likewise, I assess the capacity explanation by asking if Norway's capacity was adequate to reach compliance by 2010. If it was, then capacity cannot account for Norway's noncompliance. Finally, I assess Chayes and Chayes' temporal dimension by asking if compliance was realistically within reach, given the time frame and developments between the protocol's adoption and its deadline.

Empirically assessing the enforcement school's explanatory power is less straightforward. Two questions must be answered. First, was the target shallow or deep? Second, were any enforcement mechanisms in place?

Conducting counterfactual judgments is notoriously challenging. However, Norway's $NO_x$ target was arguably deep and thus deviates from a BAU scenario: Cost analyses conducted by Norwegian authorities in 1998 and 1999 suggested that compliance with the $NO_x$ target in the Gothenburg Protocol would amount to 200–300 million Norwegian kroner (NOK) annually *compared to the expected emissions trajectory*.[16]

Like other UN agreements, CLRTAP protocols have no significant enforcement mechanisms (Wettestad 2012, 35). Essentially, Gothenburg consists of emissions targets and timetables (Kokkvoll Tveit, 2017).

Likewise, no regulation following Norway's membership in the European Economic Area (EEA) has provided incentives for Norway to comply with its $NO_x$ target. Although the inclusion of the EU's National Emissions Ceilings (NEC) Directive (Directive 2001/81/EC) in the EEA Agreement in 2009 made Norway's $NO_x$ target of 156,000 metric tons binding under the EEA Agreement, evidence suggests that Norway's policies were unaffected by the NEC Directive. Norway strengthened its $NO_x$ policies well before 2009. Throughout the 2000s, Norwegian authorities rarely, if ever, refer to the NEC Directive when $NO_x$ policies are discussed. For instance, the government's budget proposal from October 2009 only mentions the Gothenburg Protocol (Prop. 1 S [2009–2010],

16. Miljøverndepartementet, Om samtykke til ratifikasjon, st. prp. 87 (1999–2000).

129–130). Erik Solheim, Norway's minister of the environment from 2005 to 2012, states that he never heard anyone suggest that noncompliance with Gothenburg targets could result in punitive actions.[17] Geir Axelsen, state secretary in the Ministry of Finance from 2005 to 2009, states that the NEC Directive was "not in his mind at all" while the $NO_x$ tax was prepared and implemented.[18] Moreover, the NEC Directive was not brought up in $NO_x$ policy discussions between state authorities and the business sector.[19] Neither does Harald Rensvik, secretary general[20] in the Ministry of the Environment from 1996 to 2011, suggest that Norway's $NO_x$ policies were affected by any anticipation of sanctions following Norway's EEA membership.[21] Additionally, as of December 2017, the EFTA (European Free Trade Association) Surveillance Authority had taken no action following Norway's delayed compliance with the $NO_x$ target.[22]

Considering the deep target and no enforcement, the enforcement school would predict Norwegian noncompliance with its $NO_x$ target and that Norway would not commission policies to reduce $NO_x$ emissions beyond a BAU scenario.

## Empirical Analysis I: The Management School

I first consider the management school's ambiguity explanation, then the capacity explanation, and finally the temporal dimension.

### *Can Ambiguity Explain Norway's Noncompliance?*

Gothenburg states, "Each party shall, as a minimum, control its annual emissions of polluting compounds in accordance with the obligations in annex II" (Article 3, paragraph 1). The protocol includes no provision that may relieve Norway of the obligation to reach the target—unless it withdraws from the protocol. Thus the protocol seems unambiguous concerning Norway's $NO_x$ obligations.

Public statements from Norwegian authorities suggest that they share this interpretation: under the headline "Did not comply with $NO_x$ obligation," the

17. Author's interview with Erik Solheim, Paris, May 2014.
18. Author's interview with Geir Axelsen, Oslo, February 2016. The state secretary is the political second-in-command in Norwegian Ministries. According to Kristin Halvorsen, minister of finance from 2005 to 2009, Axelsen was following the implementation of the $NO_x$ policies very closely. Author's e-mail correspondence with Erik Tollefsen, Kristin Halvorsen's adviser, October 2015.
19. Author's e-mail correspondence with Geir Høibye, January 6, 2016. Høibye is former assistant director at the Confederation of Norwegian Enterprise and former manager of the $NO_x$ fund.
20. The secretary general is the highest-ranking permanent bureaucrat of Norwegian Ministries.
21. Author's interview with Harald Rensvik, Oslo, January 2017.
22. Author's e-mail correspondence with Eli Marie Åsen, senior adviser at the Norwegian Ministry of Climate and Environment, December 2017.

Norwegian Environment Agency stated that, "in 2010 [the $NO_x$ emissions] were 19 per cent above Norway's emissions target in the Gothenburg Protocol."[23] When I asked former minister Solheim if there ever were doubts over the protocol's implications for Norway, he replied, "I cannot remember that anyone ever suggested that the protocol's content was unclear.... The focus was on two questions: what time frames are achievable, and what kinds of costs are we willing to impose on the affected businesses?"[24]

If ambiguity influenced Norwegian $NO_x$ policies, it should be known by the then minister in charge. We have little reason to distrust Solheim's statement. Generally, scholars must be careful when using information from political actors who might want to give audiences a certain impression. However, when actors present facts or opinions that may be conceived of as unfavorable to themselves, despite having the possibility to frame them differently, they seem trustworthy. In short, ambiguity was not a barrier to Norwegian compliance.

### *Can Lack of Capacity Explain Norway's Noncompliance?*

In 1999[25] and 2006,[26] Norwegian authorities published cost analyses of $NO_x$ emissions reductions. According to the 1999 study, reaching the 2010 $NO_x$ target would require implementation of all measures with abatement costs up to NOK 20 per kilogram. The estimated *total* compliance cost was NOK 200–300 million annually. In the 2006 study, only the first type of estimate was included: compliance would require implementation of all measures with costs up to NOK 60 per kilogram.

Thus we do not know how high the total costs of reaching compliance would be. However, given that the first estimate tripled from NOK 20 to NOK 60 per kilogram, we cannot rule out the possibility that the total costs might have tripled as well. Thus, although the numbers are uncertain, total costs *may* have been in the range of NOK 600–900 million annually.

Compared to the Norwegian government's total spending on environmental measures, NOK 600–900 million is substantial: the Ministry of the Environment's total 2002 budget was approximately NOK 2.8 billion,[27] and it was NOK 5.4 billion in 2013.[28] It seems, however, safe to conclude that Norway did have the funds to cover its compliance costs. When the 2010 national budget was

23. Miljødirektoratet, Klarte ikke innfri $NO_x$-forpliktelsene. Available online at http://tinyurl.com/zgqy6og, last accessed November 27, 2017. As shown by Table 2, more recent reports of Norway's 2010 emissions suggest that Norway's 2010 noncompliance was 13.6 percent.
24. Author's interview with Erik Solheim, Oslo, February 2013.
25. See Miljøverndepartementet, Om samtykke til ratifikasjon, st. prp. 87 (1999–2000).
26. http://tinyurl.com/jf4rk5h, last accessed November 27, 2017.
27. Miljøverndepartementet. 2001. Available online at: tinyurl.com/hber222, last accessed December 10, 2017.
28. Miljøverndepartementet. 2012. Available online at https://tinyurl.com/zfh324g, last accessed November 27, 2017.

presented, total incomes were estimated at NOK 974 billion, with a budget surplus of NOK 67 billion.[29]

Even more importantly, none of the individuals I interviewed indicated that lack of resources or increased costs caused Norway's noncompliance. When I asked former minister Solheim why Norway did not reach its 2010 $NO_x$ target, he replied, "This was all about one thing: mobilizing the political will. Technological barriers and similar factors were negligible."

If increased costs explain some or all of Norway's noncompliance, Solheim would likely have mentioned it. He was minister of the environment until 2012—well after Gothenburg's 2010 deadline. Hence Solheim should have every reason to point at factors that might excuse the noncompliance. Neither did former state secretary Axelsen nor former secretary general Rensvik direct our attention to (unexpectedly high) compliance costs, despite ample opportunities to do so during my interviews with them.

What about lack of knowledge? Three important documents concerning Norwegian $NO_x$ policies largely agree on (1) how emissions can be reduced and (2) which sources' emissions should be cut. White papers from 1994–1995,[30] 2004–2005,[31] and 2016–2017[32] all point to emission limits for road vehicles, using low-$NO_x$ technology on the petroleum industry's diesel turbines, international regulation of shipping emissions, retrofitting of modern technology on small coastal vessels, and corresponding measures for land-based industry. This continuity indicates that knowledge concerning sources of and solutions to $NO_x$ emissions was reasonably mature already in the mid-1990s.

Finally, can lack of bureaucratic resources explain Norwegian noncompliance? In April 2008, the Office of the Auditor General (OAG) of Norway presented a report on Norwegian authorities' efforts to reduce $NO_x$ emissions in accordance with national goals. The OAG concludes that "the authorities control relevant measures, but the implementation of these measures overall has not contributed to significant emissions reductions." The OAG also finds that, "judging by the measures implemented by December 2007, we find it very likely that Norway will not be able to reduce its $NO_x$ emissions in accordance with its obligations in the Gothenburg Protocol by 2010" (Riksrevisjonen 2008, 89).

The ninety-four-page OAG report was the outcome of a thorough review process. Five ministries provided detailed comments to draft versions (Riksrevisjonen 2008, 15, 18, 65). Thus the OAG's conclusions were likely based on the best available information. If bureaucratic resources were in

29. Finansdepartementet. Available online at http://www.statsbudsjettet.no/Statsbudsjettet-2010, last accessed November 27, 2017.
30. Miljøverndepartementet, Norsk politikk mot klimaendringer og utslipp av nitrogenoksider, st. meld. 41 (1994–1995).
31. Miljøverndepartementet, Om regjeringens miljøvernpolitikk og rikets miljøtilstand, st. meld. 21 (2004–2005).
32. Finansdepartementet, Perspektivmeldingen 2017, meld. st. 29 (2016–2017).

short supply, or the bureaucratic institutions in any way were incapable of carrying out governmental instructions, the auditor general—and former minister Solheim—would almost certainly have directed attention to this problem.

Thus the evidence suggests that lack of capacity was not a problem.

### *The Management School: The Temporal Dimension*

Setting an emissions target and deadline is a decision made under incomplete information. Matching the parties' information and expectations in 1999 with what actually happened can tell us more about the temporal dimension's explanatory power. If the target proved significantly harder to reach than Norwegian authorities expected when the protocol was adopted, Chayes and Chayes' "temporal dimension" may fully or partly explain Norway's noncompliance.

Several sources (UNECE 2003; European Commission 2015, 2016) suggest that during the last ten to fifteen years, scientists have several times increased estimations of diesel vehicles' $NO_x$ emissions because emissions under real-life conditions have proven to be higher than emissions under tests. Hence diesel vehicles have failed to live up to a number of EU emissions standards, thereby (potentially) increasing countries' total emissions. Underestimation may make compliance with a quantified emissions target less attainable. For example, if Norwegian $NO_x$ emissions in 1999 were significantly underestimated, Norway's authorities may have believed that reaching the 156,000 metric tons target was easier than what proved to be true.

Table 3 shows how estimates of Norway's $NO_x$ emissions in 1999, 2005, and 2010 have varied over time.[33] This procedure allows comparison of *what Norwegian authorities believed* were the $NO_x$ emissions in those years to *the emissions levels that were verified later*.[34] Evidently, Norway's total $NO_x$ emissions have largely been overestimated. For instance, Norway's emissions in 1999 were estimated at 213,700 metric tons in 2015 but at 239,000 metric tons in 2001. Thus the 1999 emissions reported in 2001 were 25,300 metric tons higher than they were in 2015.

The only instance of underestimation shown in my tables is the 2010 estimate of the 2005 emissions. This estimate is 9,200 metric tons lower than the estimate from 2015. If anything, the significant overestimation could have made compliance more attainable, because it may have (mis)led Norwegian authorities to believe that compliance required even stronger efforts than what proved to be true. Furthermore, new studies suggesting that some diesel

33. As reported to UNECE.
34. Since the most recent estimates are based on the best scientific knowledge available today, I use the estimates reported in 2015 as my baseline. It is, of course, true that any estimate may prove incorrect. Basing my judgments on the best data presently available is, however, the only viable solution.

**Table 3**
Estimations of Norway's $NO_x$ Emissions

| *Year* | *Emissions (metric tons)* |
|---|---|
| 1999 | |
| 2001 | 239,000 |
| 2005 | 238,000 |
| 2010 | 215,500 |
| 2015 | 213,700 |
| 2005 | |
| 2007 | 196,900 |
| 2010 | 186,900 |
| 2015 | 196,100 |
| 2010 | |
| 2012 | 184,300 |
| 2015 | 177,200 |

vehicles' $NO_x$ emissions in fact were higher than allowed by the EURO 2 emissions standard were presented as early as 2003,[35] at a meeting in an advisory body under CLRTAP. Experts from most parties, including Norway, participated (UNECE 2003). Thus, six and a half years before the 2010 deadline expired, state authorities got an "early warning" about diesel vehicles' violations of EU emissions standards.

## Empirical Analysis II: The Enforcement School

Because Gothenburg includes no enforcement mechanism, the enforcement school would expect Norway *not to implement any policies to cut $NO_x$ emissions beyond BAU*. This section argues that although it happened too late to reach compliance by 2010, a policy package introduced in 2007 and 2008 led to emissions reductions well beyond a BAU scenario.

### *Norway's $NO_x$ Policies, 1999–2010: From Weak to Strong*

After the Storting (Norway's parliament) consented to ratifying Gothenburg (December 2000), an expert group examined the prospects for cutting emissions.

35. Vehicle noncompliance with more recent emissions standards due to inconsistency between real-life and test-cycle emissions have been discovered (European Commission 2015).

Its 2004 report included no specific advice concerning measures. However, it stressed that BAU would not suffice to reach compliance: "Significantly stronger measures to reduce $NO_x$ emissions are required" (Riksrevisjonen 2008, 7, 47–51). Similarly, as shown earlier, the OAG concluded that Norwegian $NO_x$ policies had not significantly reduced emissions by December 2007. Emissions data (Table 2) support the OAG's assessment: between 2000 and 2007, emissions were relatively stable, and the weak downward trend was not sufficient for reaching the 2010 target.

After 2007, however, the downward trend of the emissions curve became steeper. This change coincided with the introduction of (1) the $NO_x$ tax that came into force on January 1, 2007, and (2) the so-called $NO_x$ agreement between the Norwegian government and several sector organizations that are members of the Confederation of Norwegian Enterprise. Most significant $NO_x$ emitters are subject to the tax, which covers approximately 55 percent of Norwegian emissions. Companies entering the $NO_x$ agreement are exempted from the tax and pay only a lower rate to the so-called $NO_x$ fund, which supports $NO_x$-reducing investment. Thus, rather than being collected by the treasury, the revenue is redistributed to emitters able and willing to reduce emissions.[36] The agreement now covers more than 95 percent of taxable emissions.[37]

According to the first $NO_x$ agreement, the total emissions reductions from the affiliated enterprises should amount to 18,000 metric tons from 2008 to 2010. Certification and consultancy foundation Det Norske Veritas (DNV) verified these reductions. In 2012, having reviewed projects supported by the $NO_x$ fund, the Norwegian Environment Agency concluded that the 2008–2010 target was reached and that "the $NO_x$ emissions have been reduced by 21,211 metric tons between 2008 and 2011" (Miljødirektoratet 2012).

Of course, DNV's verifications might be exaggerated or otherwise incorrect. Consultants may be reluctant to draw negative attention to their customers' prestige projects. Likewise, it would be naive to rule out the possibility that the Norwegian Environment Agency might be influenced by the ministry's need to show results.

Conversely, both DNV and the Norwegian Environment Agency are staffed with highly qualified personnel trained to adhere to strict scientific norms. Moreover, their findings are supported by Norway's decreasing emissions. Norway experienced a total $NO_x$ emissions reduction of 15,200 metric tons in 2008 and 2011 (Table 2). Albeit lower than the 21,211 metric tons reduction from projects verified by DNV, this experienced reduction suggests that DNV's and the Environment Agency's claims correspond reasonably well to reported emissions.

36. Miljøavtale om reduksjon av $NO_x$-utslipp for perioden 2011–2017. Available online at https://tinyurl.com/hnvo8t7, last accessed November 27, 2017; NOU, Sett pris på miljøet, pp. 15, 98. Available online at https://tinyurl.com/y7yolhwy, last accessed November 27, 2017.
37. Author's e-mail correspondence with $NO_x$ fund manager Tommy Johnsen, April 2016.

### *Can Domestic Benefits Explain the Strict Post-2006 Policies?*

Thus far, I have argued that Norway's policies from 2007 onward reduced emissions beyond BAU. An important foundation for this claim is the fact that Norway was in compliance by 2013 despite an analysis from 1999 that suggested that reaching Gothenburg's $NO_x$ target would be NOK 200–300 million annually *compared to an expected emissions scenario*. Could it be, however, that the reductions were caused by a reevaluation of the domestic costs and/or benefits of $NO_x$ emissions reductions?

All the evidence I have collected consistently suggests otherwise. Between 1999 and 2010, Norwegian authorities conducted no new analyses of the total costs of complying with the $NO_x$ target. As mentioned previously, a Norwegian Environment Agency report from 2006 argued that compliance required implementation of significantly more expensive measures than were found in 1999. Moreover, in the national budget for 2007, "[the Government proposes] a tax on $NO_x$ emissions in order to fulfill the obligations in the Gothenburg Protocol of 1999."[38] Domestic advantages of stricter national $NO_x$ regulation are not even mentioned.

Harald Rensvik, former secretary general in the Ministry of the Environment, argues that "the abatement costs of stricter $NO_x$ policies were considerable, at least in the short run. I do not think domestic benefits of Norway's emissions reductions can explain why these policies were introduced. As I see it, the crucial determinant was the wish to reach Gothenburg's $NO_x$ target." Rensvik's claim is consistent with statements from former minister Solheim and manager of the $NO_x$ fund Tommy Johnsen. None of these individuals have incentives to avoid emphasizing the domestic benefits of strict $NO_x$ regulation. If anything, we would expect politicians like Solheim to (over-)emphasize domestic benefits of environmental policies, because it could increase voter support.

Moreover, Norway's $NO_x$ policies since January 2007 mainly reduce emissions from sea vessels and petroleum installations at the continental shelf (Table 4). If domestic damage costs were decision-makers' primary concern, we would rather expect that they target urban emissions.

It thus seems unlikely that the emissions reductions since 2007 are maximizing Norway's net private benefit. Hence Norway's emissions trajectory deviates from BAU and thereby runs contrary to the enforcement school's expectation.

## Empirical Analysis III: An Opportunity Lost?

Thus far, this article has found that the two main theories in the compliance literature cannot explain very much of Norway's noncompliance. Granted, the lax policies until 2007 are consistent with the enforcement school. Similarly, the

38. Finansdepartementet. st. prp. nr. 1 (2006–2007). Available online at: http://www.statsbudsjettet.no/Upload/Statsbudsjett_2007/dokumenter/pdf/gulbok.pdf, last accessed December 4, 2017.

**Table 4**
Emissions Reductions Supported by the $NO_x$ Fund, 2006–2017

| *Source* | *Emissions Reduction (%)* |
|---|---|
| Offshore Service Vessels | 33 |
| Offshore Petroleum Installations | 16 |
| Fishing Vessels | 13 |
| Ferries/Passenger Vessels | 12 |
| Land-Based Industry | 12 |
| Cargo/Tank Vessels | 11 |
| Drilling Rigs | 3 |
| Total | 100 |

Data are from https://tinyurl.com/yd4tmmcy, last accessed November 27, 2017. Includes emissions-reducing projects that were completed and measures that were applied for by 2014.

stringent policies in force since 2007 are consistent with the management school. Nonetheless, the enforcement school expects no emissions reductions beyond BAU *throughout the period*, whereas managerialists would expect Norway to tighten policies *early enough to reach compliance*. Neither school can explain the change of policy stringency from 2007 onward. Hence, below I develop and consider two explanations that may account for Norway's behavior throughout the period from Gothenburg's adoption in 1999 until its 2010 deadline.

When I asked about his opinion concerning why the $NO_x$ tax and the $NO_x$ agreement were not introduced earlier, former $NO_x$ fund manager Geir Høibye answered, "Even though it is difficult to verify this information, several reliable sources have suggested that Kristin Halvorsen [see Table 5] was the fourth minister of finance that handled the proposal of a $NO_x$ tax and the first that did not turn it down." Høibye's account suggests that characteristics of politicians in office—or their parties—explain Norway's noncompliance and its $NO_x$ policies between 1999 and 2010. Between 2001 and 2005 Øystein Børmer was state secretary and thus political second-in-command in the Ministry of Finance.[39] Having stressed that it is challenging to recall details about specific proposals more than a decade later, he stated, "$NO_x$-reducing measures were discussed throughout the 1990s and into the 2000s.... A $NO_x$ tax must have been a part of our assessments continuously throughout our time in office."[40]

39. Author's e-mail correspondence with former minister of finance Per-Kristian Foss, October 2015. Foss advised the author to interview former state secretary Børmer, due to Børmer's in-depth knowledge of environmental tax policies between 2001 and 2005.
40. Author's e-mail correspondence with Øystein Børmer, October 2015, April 2016.

**Table 5**
Politicians in Key Positions, 1999–2010

| *Appointed* | *Minister of the Environment* | *Minister of Finance* | *Prime Minister* |
|---|---|---|---|
| 1997 | Fjellanger (Liberals) | Restad (Centre) | Bondevik (Christian Dem.) |
| 2000 | Bjerke (Labor) | Schjødt-Pedersen (Labor) | Stoltenberg (Labor) |
| 2001 | Brende (Conservatives) | Foss (Conservatives) | Bondevik |
| 2004 | Hareide (Christian Dem.) | Foss | (Bondevik) |
| 2005 | Bjørnøy (Socialist Left) | Halvorsen (Socialist Left) | Stoltenberg |
| 2007 | Solheim (Socialist Left) | (Halvorsen) | (Stoltenberg) |
| 2009 | (Solheim) | Johnsen (Labor) | (Stoltenberg) |

Børmer thus largely confirms that a $NO_x$ tax was indeed considered several years before 2007.

When asked about the bureaucracy's opinions concerning a $NO_x$ tax versus a $NO_x$ agreement, Geir Axelsen, one of Børmer's successors as state secretary in the Ministry of Finance, replied, "Most economists, as well as academic literature on environmental economics, would argue that emissions taxes are more effective than green technology subsidies. Thus, it is rather simple to imagine what advice the bureaucracy gave."

Together, the statements from Børmer and Axelsen strengthen the impression that an emissions tax was high on the agenda in $NO_x$ policy discussions both before and after 1999 and that such a tax was proposed several times to the Ministry of Finance's political leadership.

Moreover, politicians and bureaucrats alike were well aware that Norway was heading toward noncompliance. As shown earlier, an expert group argued in 2004 that "significantly stronger measures" were required to reach the $NO_x$ target. Their conclusion echoes findings of another expert committee (appointed by the Ministry of Finance) more than a decade earlier: assessing Norway's chances of fulfilling a nonbinding 1988 declaration of a 30 percent $NO_x$ emissions reduction by 1998 (see Wettestad 2012, 29),[41] the committee wrote that "far-reaching measures in addition to current policies are needed."[42]

The government's awareness of the need for additional policies is shown by the budget proposal presented in October 2001: "The Government is currently assessing what measures should be implemented to meet Norway's obligations

41. Implying Norwegian emissions not exceeding 153.000 metric tons.
42. NOU 1992: 3, «Mot en mer konstnadseffektiv miljøpolitikk i 1990-årene», p. 28. Available online at http://tinyurl.com/h5oun44, last accessed December 10, 2017.

under the [Gothenburg] protocol, and will return [to the Storting] with its views."[43] One year later, this assessment was still ongoing.[44]

In summary, it had long been known what measures could cut emissions and that additional policies were indeed needed. Moreover, the 2010 deadline had been known since 1999. The question is why the $NO_x$ policy shift came only around 2007. What changed?

One factor that did change was who was in charge. After the parliamentary elections in 2005, a cabinet consisting of the Labor Party, the Agrarians (the Centre Party), and the Socialist Left Party replaced another coalition consisting of the Conservative Party, the Liberal Party, and the Christian Democrats. The latter (minority) coalition won the 2001 elections, whereas the former retained its majority in the 2009 election.

Table 5 shows that between 2001 and 2005, the Conservatives had key influence over $NO_x$ policies. Although Knut Arild Hareide, a Christian Democrat, was minister of the environment for sixteen months in 2004 and 2005, his time was outweighed by the Conservative Børge Brende's two and a half years. Furthermore, the Conservatives' Per-Kristian Foss was minister of finance from 2001 to 2005. From 2005 to 2013, all ministers of the environment represented the Socialist Left Party. Equally important, from 2005 to 2009, Kristin Halvorsen was minister of finance, thereby giving the Socialist Left Party major influence on both environmental and tax policies.

Arguably, the 2005 change of government increased the influence of environmental parties over $NO_x$ policies: the Socialist Left Party is categorized as a typical ecosocialist European party (Arter 2008, 111; see also Heidar 2001, 69). As shown by Table 6, at the time of the elections in 2001, 2005, and 2009, voter support for the Socialist Left Party's environmental policies was strong and stable. For instance, in 2001, 36 percent of Norwegian voters thought that the Socialist Left Party had the best climate and environmental policies among Norwegian parties.[45] In contrast, the Norwegian Conservatives, like many of their sister parties in Europe, traditionally appeal strongly to business interests (Heidar 2008, 46). Thus the 2007 $NO_x$ policy shift may be explained by the change of government after the 2005 elections: politicians with seemingly good reasons to be reluctant to implement costly yet environmentally effective policies were replaced by politicians giving higher priority to environmental issues.

Evidence from my interview with former secretary general Rensvik suggests that Halvorsen's role is key to understanding the $NO_x$ tax introduction. According to Rensvik, "any minister of finance's political maneuvering space is restricted by the ministry's mainstream reasoning on economic policies. However,

43. Miljøverndepartementet, st. prp. 1 (2001–2002), p. 28. Available online at http://tinyurl.com/hber222, last accessed December 2010, 2017.
44. Miljøverndepartementet, st. prp. 1 (2002–2003), p. 67. Available online at: www.regjeringen.no/no/dokumenter/stprp-nr-1–2002–2003-/id295908/, last accessed April 3, 2017.
45. The respondents are asked, "Concerning climate and the environment, which party in your opinion has the best policies?"

**Table 6**
Voter Support for Parties' Environmental Policies

| | *2001* | *2005* | *2009* |
|---|---|---|---|
| Socialist Left | 36 | 34 | 35 |
| Labor | 10 | 15 | 15 |
| Liberal | 16 | 13 | 15 |
| Christian Democrat | 7 | 3 | 1 |
| Centre | 5 | 8 | 5 |
| Conservative | 5 | 5 | 8 |
| Progress | 1 | 2 | 4 |

Data are from Karlsen and Aardal (2007, 123; 2011, 140).

I believe that Kristin Halvorsen was important in the process that got the $NO_x$ tax proposal up and running, for example to support that a first draft was presented to the government."[46] However, although the Socialist Left Party held crucial positions when Norway's $NO_x$ regulations were tightened, it is still possible that the same policies would have been implemented by other parties—had these other parties remained in power after 2005. After all, the closer one gets to a deadline, the more urgent it might seem to act to meet it.

Thus the increased willingness to reduce $NO_x$ emissions from around 2007 may have been conditioned by an increased general awareness of the upcoming 2010 time limit. If the deadline-pressure hypothesis is correct, Norway's emissions should not develop differently from those of other states, because the approaching deadline would affect politicians and bureaucrats in all member countries more or less equally. In contrast, the office-incumbent hypothesis suggests that Norway's emissions trajectory should deviate from the trajectories of other Gothenburg parties.

Although Norway's emissions clearly went down after 2006, the downward trend was even steeper in all other Northern and Western European Gothenburg Protocol parties (Table 7). Furthermore, the shape of Norway's emissions curve is roughly similar to those of other Northern and Western European states. Both curves in Figure 1 show a steeper downward trend in the latter half of the 2000s.

The verdict seems clear: Table 7 and Figure 1 support the deadline-pressure hypothesis but not the office-incumbent hypothesis.

## Conclusions

This article has shown that the management and enforcement schools fail to give a convincing account of Norway's breach of its 2010 $NO_x$ target under the Gothenburg Protocol. The management school is unable to explain Norway's

46. Author's interview with Harald Rensvik, Oslo, January 2017.

**Table 7**
$NO_x$ Emissions Reductions, 2006–2010

| | *Emissions Reductions (%)* |
|---|---|
| Denmark | –27.7 |
| United Kingdom | –27.2 |
| Luxembourg | –26.9 |
| France | –19.3 |
| Netherlands | –16.2 |
| Germany | –14.3 |
| Switzerland | –14.0 |
| Sweden | –13.1 |
| Finland | –11.3 |
| Belgium | –9.1 |
| Norway | –8.8 |

*Note.* Data are from CEIP trends tables for 2015, available online at https://tinyurl.com/y8zvadz5, last accessed March 27, 2017.

noncompliance: there were no doubts about Gothenburg's contents. Although compliance costs were higher than expected when the protocol was adopted, Norway's capacity to comply seems high nonetheless. Interviews with top politicians and bureaucrats support this conclusion. Furthermore, Norway had adequate time to reach compliance.

The enforcement school's expectation that states—unless incentivized—will not implement costly emissions reductions is clearly inconsistent with Norway's behavior after 2007. Although Norway's material interests concerning $NO_x$ emissions did not change significantly around 2007, effective policies were

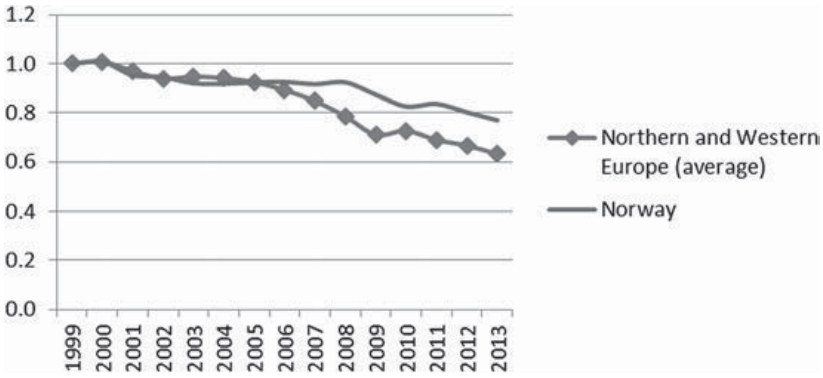


**Figure 1**
Domestic emissions levels as share of 1999 emissions

implemented. Thus the lack of action until the introduction of the $NO_x$ tax is consistent with the enforcement school, but the increased policy stringency is not.

I thereafter developed inductively, and assessed, two more hypotheses. In accordance with the office-incumbent hypothesis, rapid emissions reductions followed the 2007 introduction of a $NO_x$ tax. Such a tax had long been considered; however, it was not imposed until after the 2005 elections, when an environmentalist party gained major influence over fiscal and environmental policies. However, this theory cannot explain why other Gothenburg states also reduced their emissions significantly after 2006. These simultaneous reductions support the deadline-pressure hypothesis: only when the 2010 deadline got close did action to reduce $NO_x$ emissions begin to seem urgent to the Gothenburg member countries.

One may well hypothesize that norms, specifically, the *pacta sunt servanda* norm was the driver of the increased efforts. However, the norm was evidently not strong enough to induce Norwegian emissions reductions early enough to reach compliance by 2010. Similarly, because the lack of strong $NO_x$ policies until 2007 is consistent with the enforcement school, one may argue that Norway's policies followed a logic of consequences until 2007, and a logic of appropriateness from then onward (see March and Olsen 1998). The Norwegian case may suggest that norms and incentives affect state behavior simultaneously but that their relative influence varies over time. Whereas the logic of consequences is the more important driver when the deadline is distant, the logic of appropriateness influence grows stronger when the deadline draws near.

Thus the deadline-pressure hypothesis may be seen as a hybrid of the management and enforcement schools. In contrast with the two theories in their original forms, the synthesized hypothesis is able to account for Norway's policies throughout the period between 1999 and 2010 and for the change from weak to strong $NO_x$ policies.

**Andreas Kokkvoll Tveit** is a PhD candidate in the Department of Political Science, University of Oslo. His main research interests include international environmental governance and comparative environmental politics.

## References

Aakre, Stine, Leif Helland, and Jon Hovi. 2016. When Does Informal Enforcement Work? *Journal of Conflict Resolution* 60 (7): 1312–1340.

Arter, David. 2008. *Scandinavian Politics Today*. Manchester, UK: Manchester University Press.

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

Breitmeier, Helmut, Oran R. Young, and Michael Zürn. 2006. *Analyzing International Environmental Regimes: From Case Study to Database*. London: MIT Press.

Brown Weiss, Edith, and Harold K. Jacobson, eds. 1998. *Engaging Countries: Strengthening Compliance With International Environmental Accords*. Cambridge, MA: MIT Press.

Chayes, Abram, and Antonia Handler Chayes. 1993. On Compliance. *International Organization* 47 (2): 175–205.

Chayes, Abram, and Antonia Handler Chayes. 1995. *The New Sovereignty*. London: Harvard University Press.

Chayes, Antonia Handler, Abram Chayes, and Ronald B. Mitchell. 1995. Active Compliance Management in Environmental Treaties. In *Sustainable Development and International Law*, edited by Winfried Lang, 75–89. London: Graham and Trotman.

Downs, George W., David M. Rocke, and Peter N. Barsoom. 1996. Is the Good News About Compliance Good News About Cooperation? *International Organization* 50 (3): 379–406.

European Commission. 2015. Fact Sheet: Air Pollutant Emissions Standards. Available online at: http://europa.eu/rapid/press-release_MEMO-15–5705_en.htm, last accessed November 27, 2017.

European Commission. 2016. Transport Emissions: Air Pollutants From Road Transport. Available online at: http://ec.europa.eu/environment/air/transport/road.htm, last accessed November 27, 2017.

Finnemore, Martha, and Kathryn Sikkink. 1998. International Norm Dynamics and Political Change. *International Organization* 52 (4): 887–917.

Franck, Thomas M. 1988. Legitimacy in the International System. *American Journal of International Law* 82 (4): 705–759.

Heidar, Knut. 2001. *Norway: Elites on Trial*. Boulder, CO: Westview Press.

Heidar, Knut. 2008. Norge. In *Politikk i Europa*, edited by Knut Heidar, Einar Berntzen, and Elisabeth Bakke, 42–57. Oslo: Universitetsforlaget.

Henkin, Louis. 1968. *How Nations Behave: Law and Foreign Policy*. London: Pall Mall Press.

Jacobson, Harold K., and Edith Brown Weiss. 1998. Assessing the Record and Designing Strategies to Engage Countries. In *Engaging Countries: Strengthening Compliance With International Environmental Accords*, edited by Edith Brown Weiss and Harold K. Jacobson, 511–554. Cambridge, MA: MIT Press.

Karlsen, Rune, and Bernt Aardal. 2007. Politisk dagsorden og sakseierskap. In *Norske velgere: En studie av stortingsvalget 2005*, edited by Bernt Aardal, 111–132. Oslo: Damm og Søn.

Karlsen, Rune, and Bernt Aardal. 2011. Kamp om dagsorden og sakseierskap. In *Det politiske landskap: En studie av stortingsvalget 2009*, edited by Bernt Aardal, 131–160. Oslo: Damm og Søn.

Kim, Yoomi, Katsuya Tanaka, and Shunji Matsuoka. 2017. Institutional Mechanisms and the Consequences of International Environmental Agreements. *Global Environmental Politics* 17 (1): 77–98.

Kokkvoll Tveit, Andreas. 2017. Can the Management School Explain Noncompliance? Paper presented at the 58th Annual Convention of the International Studies Association (ISA). Baltimore, Maryland, February 22–25, 2017. Available at: http://www.sv.uio.no/isv/personer/vit/andreakt/isa-2017_can-the-management-school-explain-noncompliance.pdf.

Levy, Marc A. 1993. European Acid Rain: The Power of Tote-Board Diplomacy. In *Institutions for the Earth: Sources of Effective International Environmental Protection*, edited by Peter M. Haas, Robert O. Keohane, and Marc A. Levy, 75–132. Cambridge, MA: MIT Press.

March, James G., and Johan P. Olsen. 1998. The Institutional Dynamics of International Political Orders. *International Organization* 52 (4): 943–969.

Miljødirektoratet. 2012. *Årsrapport 2011—oppfyllelse av forpliktelsene i reduksjon av $NO_x$-utslipp for 2011 i miljøavtalen om $NO_x$*. Kommentarer til rapporter fra $NO_x$-fondet. On file with author.

Mitchell, Ronald B. 2006. Problem Structure, Institutional Design, and the Relative Effectiveness of International Environmental Agreements. *Global Environmental Politics* 6 (3): 72–89.

Perkins, Richard, and Eric Neumayer. 2007. Implementing Multilateral Environmental Agreements: An Analysis of EU Directives. *Global Environmental Politics* 7 (3): 13–41.

Riksrevisjonen. 2008. Riksrevisjonens undersøkelse av myndighetsutøvelse for reduksjon av $NO_x$-utslipp. Dokument 3:9 (2007–2008), Riksrevisjonen.

Underdal, Arild. 2000. Comparative Analysis: Accounting for Variance in Actor Behaviour. In *International Environmental Agreements and Domestic Politics*, edited by Arild Underdal and Kenneth Hanf, 341–384. Aldershot, UK: Ashgate.

Underdal, Arild. 2002. One Question, Two Answers. In *Environmental Regime Effectiveness: Confronting Theory With Evidence*, edited by Edward L. Miles, Arild Underdal, Steinar Andresen, Jørgen Wettestad, Jon Birger Skjærseth, and Elaine M. Carlin, 433–465. London: MIT Press.

Underdal, Arild, and Kenneth Hanf. 2000. *International Environmental Agreements and Domestic Politics: The Case of Acid Rain*. Aldershot, UK: Ashgate.

UNECE. 2003. *Progress Report*. Available online at https://tinyurl.com/y9rnvzl6, last accessed November 27, 2017.

Victor, David, Kal Raustiala, and Eugene B. Skolnikoff, eds. 1998. *The Implementation and Effectiveness of International Environmental Commitments: Theory and Practice*. Cambridge, MA: MIT Press.

Wettestad, Jørgen. 2002. The Convention on Long-Range Transboundary Air Pollution (CLRTAP). In *Environmental Regime Effectiveness: Confronting Theory With Evidence*, edited by Edward L. Miles, Arild Underdal, Steinar Andresen, Jørgen Wettestad, Jon Birger Skjærseth, and Elaine M. Carlin, 197–222. London: MIT Press.

Wettestad, Jørgen. 2012. Reducing Long-Range Transport of Air Pollutants in Europe. In *International Environmental Agreements: An Introduction*, edited by Steinar Andresen, Elin Lerum Boasson, and Geir Hønneland, 23–37. New York: Routledge.

Young, Oran. 1979. *Compliance and Public Authority:* A Theory *W*ith International Applications. London: Johns Hopkins University Press.

Young, Oran, and Michael Zürn. 2006. The International Regimes Database: Designing and Using a Sophisticated Tool for Institutional Analysis. *Global Environmental Politics* 6 (3): 121–143.