

Can the Management School Explain Noncompliance with International Environmental Agreements?

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 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.

Word count:

7,984

1 Introduction

How can we account for noncompliance with international environmental agreements (IEAs)? In their seminal¹ 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 implementation.

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).

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

² Other prominent managerialists include Young (1979) and Mitchell (1994, 2010).

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

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 an international cooperative effort with a strong scientific basis (Castells and Ravetz 2001, Rensvik 2017, Tuinstra 2008). Much energy has been devoted to modeling 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 if compliance proved 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

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

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 Protocol⁵ 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 (SO₂), 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 ozone⁸.

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 Gothenburg parties, 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 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

⁵ I refer to the Gothenburg Protocol of 1999, not the revised protocol of 2012.

⁶ Acidification is largely caused by sulphur and NO_x, 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 NO_x reacting with non-methane volatile organic compounds (NMVOCs) causes harmful ground-level ozone.

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

¹⁰ Hence, Table 1 does not engage with the highly challenging task of distinguishing between Mitchell's (2010: 147) two kinds of noncompliant behavior ("good-faith" and "intentional") or between his two kinds of

typically involves measuring an IEA's ability to improve environmental quality or state behavior 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 SO₂ 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%.

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 shown in grey (emissions data from CEIP 2015).

compliant behavior ("coincidental" and "treaty-induced" compliance).

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 skeptical 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, Simmons 2013).

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, Rocke, and Barsboom (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 (1998, eds.) assess individual states’ compliance. Although four of the five treaties studied by Brown Weiss and

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.

Jacobson (1998, 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

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 how conditions crucial for compliance have developed. 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.

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

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 1,081 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.

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.

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

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 operationalizations¹⁶ 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 (logtransformed).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 substance¹⁹ 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

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 logtransform 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.

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

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 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

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.

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).

(Models 2 and 5), the effect of capacity on compliance remains negative, but is no longer statistically significant. 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, except that the effect of (log) GDP per capita is once again significant (Model 6).

Table 2: OLS regressions. Dependent: Compliance (continuous)						
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Constant	0.509	0.886	1.074	4.709	1.985	3.552
WGI_GovtEff	-0.238***	-0.024	-0.107			
(log) GDP/capita				-1.008***	-0.263	-0.590**
Ambition level		-0.567***	-0.588***		-0.525***	-0.553***
Eastern Europe			-0.147			-0,189*
R ²	0.148	0.498	0.505	0,229	0,508	0,526
N	92	92	92	92	92	92
<p>* 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.</p>						

Even though the models' explained variance is not crucial for the purpose of this paper, it is interesting to note that R² 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.

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	18.13	13.38
WGI_GovtEff	-1.760***	-1.571**	-1.302			
(log) GDP/capita				-4.789***	-3.217**	-2.248
Ambition level		-1.288***	-1.285**		-1.883	-1.711**
Eastern Europe			0.471			0.658
Nagelkerke Pseudo-R ²	0.217	0.301	0.502	0.176	0.283	0.288
N	92	92	92	92	92	92
<p>* 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.</p>						

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 A1 and A2) 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.

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 NOx targets?

Amann et al. (1999)²² identify population size, GDP per capita, the number of vehicles, and

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).

energy consumption as major determinants of NO_x emissions. However, because GDP 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 observed²⁴ 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 1,300 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 5,600 tonnes of NO_x, thereby outweighing its noncompliance of 1,800 tonnes. I therefore conclude that deviance between projected and observed 2010 energy consumption explains Denmark's NO_x noncompliance.

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 on energy consumption included by Amann et al. (1999).

²⁴ Based on data from Eurostat (2017).

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).

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).

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 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 CLRTAP²⁵ 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

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

²⁶ These vehicles use diesel fuel.

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 A3 (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.

²⁷ I include only states that were noncompliant with their 2010 NO_x 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 be 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.

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

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

Denmark's 1999 emissions were consistently underestimated until 2010 (varying from 9,200 to 26,800 tonnes below the 2015 estimate). Since Denmark's noncompliance equalled only 1,800 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 and 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 emissions came early enough to enable the French authorities to avoid a substantial part of its noncompliance. Furthermore, the estimates from 2001 and 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 A3 (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 A3–A6 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 only 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. 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.

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.

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.

	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)

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.

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
NMVOC	Belgium	No
	Denmark	No
	Germany	No
Ammonia	Croatia	No
	Denmark	No
	Finland	No
	Germany	No
	Netherlands	No
	Norway	No
	Spain	No
	Switzerland	Yes

7 Conclusion

This article has demonstrated that the three factors specified by Chayes and Chayes cannot explain very 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 if 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.

References:

Aakre, Stine, Leif Helland, and Jon Hovi. 2016. "When Does Informal Enforcement Work?" *Journal of Conflict Resolution*, 60, no. 7: 1312–40.
doi: 10.1177/0022002714560349

Amann, Markus, Imrich Bertok, Janusz Cofala, Frantisek Gyarfas, Chris Heyes, Zbigniew Klimont, and Wolfgang Schöpp. 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, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. London: Princeton University Press.

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

Bernauer, Thomas, Anna Kalbhenn, Vally Koubi, and Gabrielle Spilker. 2013. "Is There a 'Depth versus Participation' dilemma in international cooperation?" *The Review of International Organizations* 8, no. 4: 477–97
doi: 10.1007/s11558-013-9165-1

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.
doi: 10.1016/j.jeem.2005.03.002

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

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: 1363–1390

doi: [10.1177/0010414010376910](https://doi.org/10.1177/0010414010376910)

Böhmelt, Tobias, and Jürg Vollenweider. 2015. “Information flows and social capital through linkages: the effectiveness of the CLRTAP network.” *International Environmental Agreements: Politics, Law and Economics*, 15, no. 2: 105–23.

doi:10.1007/s10784-013-9218-1

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

Castells, Nuria, and Jerry Ravetz. 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, no. 1: 405–25.

doi: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 February 14, 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§ordefinitions=NFR2014_L1&reportyear=2015§ors=556&datatype=national_html&database=web&horizontal=area&vertical=year&unit=Gg&info=on, accessed February 14, 2018.

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

Chayes, Abram, and Antonia Handler Chayes. 1993. On Compliance. *International*

Organization 47 (2): 175–205.

doi: 10.1017/S0020818300027910

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 February 15, 2018. Also on file with author.

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](https://doi.org/10.1017/S0020818300033427)

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 February 12, 2018.

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

Eurostat. 2017. “Simplified energy balances - annual data”

At <http://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do>, accessed April 26, 2017.

FAO. 2016. Country Profiles.

At <http://www.fao.org/countryprofiles/en/>, accessed February 14, 2017.

Finnemore, Martha, and Kathryn Sikkink. 1998. “International Norm Dynamics and Political Change.” *International Organization*, 52, no. 4: 887–917.

Miljødirektoratet. 2015. “Air pollution.”

At <http://www.environment.no/topics/air-pollution/>, accessed February 14, 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 April 3, 2017. Also on file with author.

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 February 14, 2018.

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*, Edith Brown Weiss and Harold K. Jacobson, eds., 511–54. Cambridge: MIT Press.

Hanson, Jonathan K., and Rachel Sigman. 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, Carsten, and Detlef Sprinz. 2000. "Measuring the Effectiveness of International Environmental Regimes." *Journal of Conflict Resolution* 44, no. 5: 630–52.
doi: 10.1177/0022002700044005004

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

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.

Jänicke, Martin. 1997. "The Political System's Capacity for Environmental Policy." In Martin Jänicke and Helmut Weidner, eds., *National Environmental Policies. A Comparative Study of Capacity-Building*, 1–24. Berlin: Springer.

Levy, Marc A. 1993. "European Acid Rain: The Power of *Tote-Board Diplomacy*." In Peter

M. Haas, Robert O. Keohane, and Marc A. Levy, *Institutions for the Earth. Sources of Effective International Environmental Protection*. 75–132. London: MIT Press.

Kelly, Andrew, Julio Lumbreras, Rob Maas, Tiziano Pignatelli, Francisco Ferreira, Anna Engleryd. 2010. “Setting national emission ceilings for air pollutants: policy lessons from an ex-post evaluation of the Gothenburg Protocol.” *Environmental Science and Policy*
[doi: 10.1016/j.envsci.2009.09.003](https://doi.org/10.1016/j.envsci.2009.09.003)

Kokkvoll Tveit, Andreas. 2018. “Norms, Incentives, or Deadlines? Explaining Norway’s Noncompliance with the Gothenburg Protocol.” *Global Environmental Politics* 18, no. 1: 76–98.

Mitchell, Ronald B. 1994. “Regime design matters: Intentional oil pollution and treaty compliance.” *International Organization*, 48(3), 425-458.
[doi:10.1017/S0020818300028253](https://doi.org/10.1017/S0020818300028253)

Mitchell, Ronald 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 February 14, 2018.

Raustiala, Kal, and David G. Victor. 1998. “Conclusions.” In *The Implementation and Effectiveness of International Environmental Commitments: Theory and Practice*, Kal Raustiala, David G. Victor, and Eugene B. Skolnikoff, eds., IIASA; MIT Press.

Raustiala, Kal, and Anne-Marie Slaughter. 2002. “International Law, International Relations, and Compliance.” In Walter Carlnaes, Thomas Risse, and Beth Simmons, eds., *The Handbook of International Relations*, London: Sage Publications. Also available at SSRN:
<https://ssrn.com/abstract=347260>

Raustiala, Kal. 2005. "Form and Substance in International Agreements." *American Journal of International Law* 99, no. 3: 581–614.

Rensvik, Harald. 2017. Author's interview with Harald Rensvik, Secretary General in Norway's Ministry of the Environment 1996–2011, Oslo, January 2017.

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: 86–102.

doi: 10.1111/j.0092-5853.2005.00112.x

Simmons, Beth A. 1998. "Compliance with International Agreements." *The Annual Review of Political Science*, 1: 75–93.

Simmons, Beth A. 2013. "From Ratification to Compliance: Quantitative Evidence on the Spiral Model." In Thomas Risse, Stephen C. Ropp and Kathryn Sikkink, eds., *The Persistent Power of Human Rights: From Commitment to Compliance*, 43–59. Cambridge: Cambridge University Press.

Tuinstra, Willemijn. 2008. "European air pollution assessments: co-production of science and policy," *International Environmental Agreements: Politics, Law and Economics* 8, no: 35.

doi: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 April 16, 2017.

Underdal, Arild. 2000. Comparative Analysis: Accounting for Variance in Actor Behaviour. Chapter 13 in Arild Underdal and Kenneth Hanf, eds., *International Environmental Agreements and Domestic Politics*. Aldershot: Ashgate.

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.

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 April 19, 2017.

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

Wettestad, Jørgen. 2012. "Reducing long-range transport of air pollutants in Europe." In Steinar Andresen, Elin Lerum Boasson, and Jørgen Wettestad: *International Environmental Agreements*, 23–37. New York: Routledge.

World Bank. 2017. "Worldwide Governance Indicators."

At <http://info.worldbank.org/governance/wgi/index.aspx#home>, accessed April 20, 2017.

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

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

Appendix to “Can the Management School Explain Noncompliance with International Environmental Agreements?”

Sensitivity check: Statistical assessments of the effect of capacity on compliance

Table A1 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 government to formulate and implement sound policies and regulations that permit and promote private sector development.”

Table A1: OLS regressions. Dependent: Compliance			
	Model 13	Model 14	Model 15
Constant	0.612	0.889	1.058
WGI_RegulatoryQuality	-0.329***	-0.03	-0.106
Ambition level		-0.572***	-0.595***
Eastern Europe			-0.098
R ²	0.113	0.498	0.502
N	92	92	92
* Coefficient is significant at the 10 percent level ** Coefficient is significant at the 5 percent level *** Coefficient is significant at the 1 percent level Standard errors are clustered on states.			

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

Table A2 shows the results of a final robustness check (Model 13). Here, I have included dummies for each substance that Gothenburg regulates. Again, the effect of capacity is negative and statistically significant. Since Model 13 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

Table A2: OLS regressions. Dependent: Compliance	
	Model 16
Constant	3.541
(log) GDP/cap.	-.56**
Ambition level	-.415***
Eastern Europe	-.063
NOx	-.424***
NMVOC	-.369***
Ammonia	-.423***
R ²	0,229
N	92
* Coefficient is significant at the 10 percent level. ** Coefficient is significant at the 5 percent level. *** Coefficient is significant at the 1 percent level Standard errors are clustered on states.	

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

Table A3: Estimates of 2003 NO _x emissions of noncompliant parties (thousand tonnes)							
Party	2005	2006	2007	2008	2009	2010	2015
Belgium	297.2 <i>-3.4</i>	297.5 <i>-3.7</i>	297.5 <i>-3.7</i>	297.5 <i>-3.7</i>	296.3 <i>-2.5</i>	297 <i>-3.2</i>	293.8
Denmark	207.8 <i>18</i>	197.9 <i>27.9</i>	210.3 <i>15.5</i>	208.3 <i>17.5</i>	203.7 <i>22.1</i>	203.9 <i>21.9</i>	225.8
France	1220.3 <i>282.6</i>	1244.1 <i>258.8</i>	1257.3 <i>245.6</i>	1450.1 <i>52.8</i>	1496.3 <i>6.6</i>	1529.4 <i>-26.5</i>	1502.9
Germany	1428 <i>287.1</i>	1604.7 <i>110.4</i>	1625.5 <i>89.6</i>	1580 <i>135.1</i>	1541.3 <i>173.8</i>	1613.8 <i>101.3</i>	1715.1
Luxembourg	N/A	17.5 <i>30</i>	17.5 <i>30</i>	17.5 <i>30</i>	16 <i>31.5</i>	16 <i>31.5</i>	47.5
Netherlands	363.8 <i>5.4</i>	367.2 <i>2</i>	373.1 <i>-3.9</i>	357.8 <i>11.4</i>	357.6 <i>11.6</i>	371.2 <i>-2</i>	369.2
Norway	220.2 <i>-25.6</i>	214.8 <i>-20.2</i>	199.2 <i>-4.6</i>	196.7 <i>-2.1</i>	194 <i>0.6</i>	190.5 <i>4.1</i>	194.6
Slovenia	56 <i>-3.9</i>	56 <i>-3.9</i>	55.3 <i>-3.2</i>	48.2 <i>3.9</i>	48.2 <i>3.9</i>	49.5 <i>2.6</i>	52.1
Spain	1518.6 <i>-116.6</i>	1493.2 <i>-91.2</i>	1492.6 <i>-90.6</i>	1500 <i>-98</i>	1490.2 <i>-88.2</i>	1401.4 <i>0.6</i>	1402
Sweden	206 <i>-20</i>	202.7 <i>-16.7</i>	215.4 <i>-29.4</i>	197.9 <i>-11.9</i>	191.7 <i>-5.7</i>	190 <i>-4</i>	186

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

Table A4: Estimates of 2006 NO _x emissions (thousand tonnes)				
	2008	2009	2010	2015
Belgium	277.7 <i>-0.3</i>	268.3 <i>9.1</i>	266.4 <i>11</i>	277.4
Denmark	185.3 <i>15.9</i>	180.7 <i>20.5</i>	182.1 <i>19.1</i>	201.2
France	1351.2 <i>7.9</i>	1397.5 <i>-38.4</i>	1414 <i>-54.9</i>	1359.1
Germany	1394.3 <i>162.8</i>	1353.9 <i>203.2</i>	1520.5 <i>36.6</i>	1557.1
Luxembourg	N/A	14.4 <i>39.2</i>	14.4 <i>39.2</i>	53.6
Netherlands	310.8 <i>16.4</i>	307.2 <i>20</i>	324.1 <i>3.1</i>	327.2
Norway	190.8 <i>3.5</i>	198 <i>-3.7</i>	185.5 <i>8.8</i>	194.3
Slovenia	46.9 <i>3.5</i>	46.8 <i>3.6</i>	46 <i>4.4</i>	50.4
Spain	1481.2 <i>-114.3</i>	1465.1 <i>-98.2</i>	1401 <i>-34.1</i>	1366.9
Sweden	174 <i>-1.8</i>	170.4 <i>1.8</i>	169.1 <i>3.1</i>	172.2
Numbers in italics are the differences between the 2015 estimate and the estimate from the column header.				

Table A5: Estimations of 2008 NO _x emissions (thousand tonnes)			
Party	2010	2012	2015
Belgium	240.5 <i>-4</i>	238.7 <i>-2.2</i>	236.5
Denmark	151.7 <i>18.4</i>	150.5 <i>19.6</i>	170.1
France	1272.5 <i>-74.9</i>	1194.4 <i>3.2</i>	1197.6
Germany	1393.3 <i>17.5</i>	1417.5 <i>-6.7</i>	1410.8
Luxembourg	N/A	50.2 <i>-5.3</i>	44.9
Netherlands	292.7 <i>6.6</i>	308.9 <i>-9.6</i>	299.3
Norway	173.7 <i>11.5</i>	189.1 <i>-3.9</i>	185.2
Slovenia	52.9 <i>2.6</i>	53.1 <i>2.4</i>	55.5
Spain	1236.3 <i>-65.4</i>	1175.6 <i>-4.7</i>	1170.9
Sweden	154.4 <i>2.6</i>	158 <i>-1</i>	157

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

Table A6: Estimates of 2010 NO _x emissions (thousand tonnes)		
Party	2012	2015
Belgium	220.7 <i>31.6</i>	252.3
Denmark	128.8 <i>16.5</i>	145.3
France	1080.3 <i>16.1</i>	1096.4
Germany	1322.9 <i>10.8</i>	1333.7
Luxembourg	46.2 <i>-6.75</i>	39.45
Netherlands	275.9 <i>-1.7</i>	274.2
Norway	184.3 <i>-7.1</i>	177.2
Slovenia	44.7 <i>2.4</i>	47.1
Spain	983.9 <i>-24.2</i>	959.7
Sweden	161.4 <i>-11.8</i>	149.6
Numbers in italics are the differences between the 2015 and 2010 estimates.		