



Does capacity increase compliance? Examining evidence from European cooperation against air pollution

Andreas Kokkvoll Tveit^{1,2}

Accepted: 12 August 2020
© The Author(s) 2020

Abstract

Scholars commonly hypothesize that enhanced capacity—improved ability to do as agreed—increases states' compliance with international agreements. In contrast, using a novel dataset that covers 31 states and three decades of cooperation, I find a negative effect of capacity on compliance. To help explain this seemingly counterintuitive finding, I offer a novel conjecture of the capacity–compliance relationship. In particular, I argue that the effect of capacity may vary substantially across states, because states' intention to comply constitutes a crucial intervening variable. Among reluctant states pursuing policy goals that affect compliance negatively, high capacity may in fact cause noncompliance. I exemplify the conjecture through evidence from a high-capacity noncompliant state (Norway).

Keywords International environmental agreements · Air pollution · Compliance · International relations · State capacity

1 Introduction

What factors determine states' compliance with international environmental agreements (IEAs)? Proponents of the management school such as Chayes and Chayes (1993) argue that lack of capacity—states' ability to do as agreed—constitute a crucial barrier to compliance. Specifically, they hypothesize that the likelihood of compliance depends on scientific, technical, bureaucratic, and fiscal resources (Chayes and Chayes 1993, 194).

Despite the considerable scholarly attention Chayes and Chayes have received, the capacity explanation of noncompliance remains understudied. Moreover, the findings of previous studies (both within and without the IEA literature) are surprisingly inconclusive: Although some studies find the expected positive effect, several others report a zero effect (Jacobson and Brown Weiss 1998; Breitmeier et al. 2006; Mastenbroek 2005; Börzel et al. 2010).

Using a series of regressions, I assess the effect of states' capacity on compliance with the emissions targets of five protocols under the Convention on Long-range Transboundary

✉ Andreas Kokkvoll Tveit
akt@toi.no

¹ Institute of Transport Economics – Norwegian Centre for Transport Research, Oslo, Norway

² Department of Political Science, University of Oslo, Oslo, Norway

Air Pollution (CLRTAP). My highly robust results do not reveal the expected positive effect of state capacity on compliance; rather, the effect is negative¹ across a number of model specifications. In contrast to most previous studies, I control for the ambitiousness of each emissions target (i.e., the size of the required emissions reductions). Acknowledging the challenges associated with measuring state capacity (Hendrix 2010), I use several alternative operationalizations. Because states may display different behavior toward different pollutants under the same regulatory regime (Murdoch et al. 1997), I report checks showing that my results are robust to including pollutant-fixed effects. Similarly, the results hold when my models include protocol-specific intercepts.

Building on a case study of a high-capacity noncompliant state, I provide a novel conjecture concerning the capacity–compliance relationship. I argue that the scope conditions of the capacity explanation of noncompliance have been underspecified. Among *reluctant* states² (Victor 2011) that pursue policy goals *negatively correlated with compliance*, we may expect general state capacity to influence compliance negatively. Such an effect is not as puzzling as it may seem. Put bluntly, my conjecture implies that enhanced capacity to save the environment (and therefore comply with an international agreement) may entail enhanced capacity to destroy it, too. Crucially, it is the intention to comply that determines whether the negative effect of capacity on compliance dominates the positive.

The remainder of this paper proceeds as follows. Section 2 briefly describes the protocols adopted by the CLRTAP parties. Section 3 presents the compliance literature’s main theories and expectations concerning the effect of capacity. It also shows that empirical assessments of those expectations have provided mixed results and presents a novel conjecture that may account for this inconclusiveness. Informed by an in-depth study of Norway’s noncompliance with the 1999 Gothenburg protocol, this conjecture also provides an explanation of the results of my statistical analyses, which are presented and discussed in Sects. 4 through 7.

2 Background: The problems, the solutions, and the agreements

Led by an avant-garde of Scandinavian countries, the international society’s focus on long-range transboundary air pollution (mainly acidification following emissions of nitrogen oxides and sulfur) gained momentum in the 1970s and early 1980s, when it became increasingly clear that such emissions were detrimental not only to their sources’ nearby surroundings,³ but also to distant environments across Europe. Although considered a victory for the activist Scandinavians, the CLRTAP—signed in Geneva in 1979—may be considered an “empty shell” (Wettestad 2012, 27). In contrast, CLRTAP’s protocols contain national emissions targets for the participating countries.

The Helsinki protocol (1985) included binding targets for emissions of sulfur (SO_x).⁴ A 30% reduction below 1980 levels was to be reached by all parties by 1993. Although the main focus in the early days of the CLRTAP cooperation was on SO_x emissions reductions,

¹ The estimated effects are negative across all regressions, but are statistically insignificant in one of my 12 main models as well as in two out of 12 robustness checks.

² That is, states having little or no intention to comply.

³ The main sources of sulfur emissions were large combustion plants in industry and energy production.

⁴ In CLRTAP protocols, the emissions targets for sulfur are usually expressed in thousand tons of sulfur dioxide (SO₂) equivalents. SO_x is often used as a collective term for a variety of sulfur emissions.

a nitrogen oxides (NO_x) protocol was adopted in 1988.⁵ This Sofia Protocol obliged parties to stabilize NO_x emissions at 1987 levels by 1994.

As became increasingly clear, regulation of other emissions was needed as well. Non-methane volatile organic compounds (NMVOC, or just VOC) stem from a large number of sources and processes, such as combustion, solvent usage, evaporation from fossil fuels, paint application, and dry cleaning. In combination with sunlight, NMVOC and NO_x create harmful ground-level ozone. Hence, the 1991 Geneva Protocol set NMVOC national emissions targets with 1999 as the deadline.

A second sulfur protocol was adopted in Oslo in 1994. Although the Oslo Protocol's main deadline year was 2000, it also included 2005 and 2010 emissions targets for some states. Finally, the 1999 Gothenburg protocol included targets not only for SO_x , NO_x , and NMVOC, but also for ammonia (NH_3). Originating mainly from agriculture, ammonia exacerbates acidification and creates over-fertilization (eutrophication) of water systems (Norwegian Environment Agency 2015).⁶

Two additional CLRTAP protocols regulate emissions. However, neither the 1998 Protocol on Persistent Organic Pollutants (POPs) nor the 1998 Protocol on Heavy Metals contains emissions targets like those of the five protocols mentioned previously. Consequently, I study states' compliance with national emissions targets specified in the following five protocols: Helsinki, Sofia, Geneva, Oslo, and Gothenburg.

Lacking potent mechanisms for inducing participation and compliance (Wettestad 2012, 35), the main contents of all five protocols are targets and timetables.

3 Previous research, hypotheses, and a novel conjecture on capacity and compliance

Arguing that intentional free-rider behavior is rare, Chayes and Chayes (1993, 176) contend that states have a "general propensity" for compliance with international agreements. That propensity is due to norms being a crucial driver of states' behavior: "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" (Chayes and Chayes 1993, 185).

According to Chayes and Chayes, the causes of noncompliance are usually beyond the reach of states. Most importantly, the actions necessary to comply often require "scientific and technical judgment, bureaucratic capability, and fiscal resources." Hence, lack of such capacities may entail noncompliance. Managerialists also argue that noncompliance can be caused by treaty ambiguity and unexpected changes between a treaty's adoption and its deadline. Chayes, Chayes and Mitchell (1995) argue, however, that capacity building is "[maybe] the most important part of active management," thereby suggesting that lack of state capacity is the most important source of noncompliance.

Unlike intentional defection, "good-faith noncompliance" may be addressed by the "managerial strategy"—ensuring transparency, review, capacity-building measures, and treaty adaption (Mitchell 2009; Young 2011).

The management school expects a positive effect of state capacity on compliance:

⁵ Major NO_x emitters include power plants and industry, in addition to transport vehicles.

⁶ Over-fertilization may increase algae growth, thereby disturbing aquatic ecosystems.

H1 State capacity influences (the likelihood of) compliance positively.

In contrast, scholars belonging to the enforcement school argue that maximization of (net) private benefit constitutes a crucial driver of state actions. According to them, the generally high compliance with international agreements is not caused by states changing their behavior in order to comply. Rather, it is due to international agreements being *shallow*, meaning that their targets would have been met *even if the agreements concerned did not exist*. Consequently, a shallow commitment is consistent with business as usual (BAU), while a deep commitment requires behavioral change and deviation from BAU (Downs et al. 1996; Barrett 2003; Aakre et al. 2016). Because states display BAU behavior, variations in target depth account for variations in compliance. While shallow targets will be reached, deep targets will not. In both cases, the enforcement school hypothesizes a zero effect of capacity.

Thus, the enforcement school expects that:

H2 State capacity has no effect on (the likelihood of) compliance.

3.1 Previous empirical research

Of the major empirical studies⁷ in the literature on international environmental cooperation, only two study the effects of state capacity or capacity-building measures on compliance. They offer diverging findings: While Jacobson and Brown Weiss (1998, 534–536) conclude that capacity is an important explanatory variable, the findings of Breitmeier et al. (2006, 111) “do not confirm expectations about the role of capacity building.”

In the comprehensive literature on (non)compliance with EU law, several studies find a positive effect of factors that might be seen as proxies for state capacity, such as finances (Demmke 2001) and administrative characteristics (Falkner et al. 2005; Börzel et al. 2010). However, even in this EU law literature the empirical knowledge concerning capacity’s explanatory power remains limited, perhaps because other explanations have received more attention (e.g., Knill and Lenschow 1998; see also Mastenbroek 2005). Moreover, among the studies that have assessed state capacity’s effect on compliance, results are rather inconsistent. Some find little or even no effect (examples include both quantitative studies such as König and Mäder (2014) and qualitative studies such as Zürn and Joerges (2005)). Despite finding a positive effect on compliance of a bureaucracy-effectiveness measure, Börzel et al. (2010) find no effect of their other two state capacity measures (GDP/cap and government autonomy from domestic veto players).

In summary, the literatures reviewed here remain inconclusive concerning the effect of capacity on compliance—although finding a positive effect is somewhat more common than finding no effect. Neither the management school nor the enforcement school can explain this inconclusiveness. Therefore, the next section develops a novel theoretical account of the possible effects of state capacity on compliance.

⁷ That is, edited volumes such as Victor et al. (1998), Miles et al. (2002), Brown Weiss and Jacobson (1998), as well as Breitmeier et al. (2006). Underdal and Hanf (2000) could also be included on that list, but is not as broad empirically as the aforementioned volumes, and its compliance assessments lack specific measures of capacity.

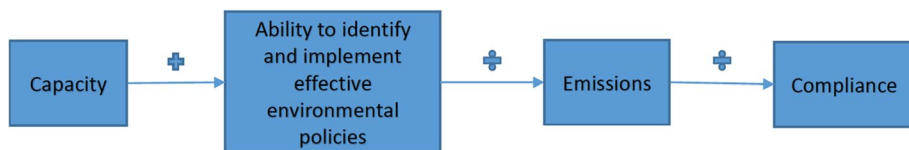


Fig. 1 Positive effect of state capacity on compliance

3.2 A novel conjecture concerning the capacity–compliance relationship

Previous research has paid little attention to the scope conditions of the hypothesized positive effect of state capacity on compliance. Under what circumstances (if any) should we expect capacity to increase compliance? In particular, does the effect depend on other state characteristics or on factors related to the problem concerned? When arguing that states have a “general propensity for compliance,” Chayes and Chayes (1993) implicitly point at a potentially crucial factor—the intention to comply. In a more recent work, Victor (2011) distinguishes between enthusiastic and reluctant states, the former being relatively keen to devote resources to the provision of international common goods such as clean air.⁸ Although readily admitting that exceptions exist, Chayes and Chayes (1993) argue that the intention to comply is generally strong (stronger than the intention to free ride). They then go on to describe the mechanisms through which high capacity may increase compliance. For instance, countries with competent and well-staffed bureaucracies will likely be able to identify effective strategies for fulfilling commitments. Moreover, wealthy countries will likely be able to pay for costly abatement.

Figure 1 provides a simple illustration of the proposed positive effect of state capacity on compliance with emissions control protocols: Increased state capacity may *reduce* emissions through increasing the ability to identify and implement optimal policies and technological solutions for abatement. Such a *negative effect on emissions* implies a *positive effect of capacity on compliance*.

Chayes and Chayes (1993) pay less attention to the possibility that alternative mechanisms may connect state capacity to compliance. Exploring such pathways, the next section discusses the case of a noncompliant high-capacity state, arguing that there might be important links between state capacities, intentions to comply, and compliance. Building on the case study, I formulate a general account of how high state capacity may cause non-compliance with IEAs.

3.3 Rich, able, and unwilling: Norway’s capacity and compliance

While often thought of as an international environmental leader, Norway failed to reach its 2010 NO_x emissions target under the 1999 Gothenburg protocol. Norway’s noncompliance may seem surprising, not least because Norwegian authorities received early and repeated warnings from expert groups that Norway would be far off its NO_x target unless stronger policies were implemented (Kokkvoll Tveit 2018). Equally important, politicians

⁸ Although developed in the context of climate politics, Victor’s (2011) concepts are generally applicable in cases where externalities render international cooperation necessary.

and bureaucrats alike knew that effective measures were available at nonzero, yet modest cost: Estimating annual costs of compliance with the NO_x target at 200–300 million NOK,⁹ a 1999 study made it clear that technology was unlikely to pose a barrier to compliance (Norwegian Pollution Control Authority 1999).

Because Norway's state capacity appears high and adequate, a closer look at the mechanisms at work might prove insightful. Until 2007,¹⁰ Norway's policies for reducing NO_x emissions were weak or even nonexistent (Wettstad 2012). Thus, NO_x policies were lax for almost a decade despite politicians and bureaucrats' being well aware that Norway was heading toward substantial noncompliance. Potent measures such as a NO_x tax were proposed several times to ministers of the environment as well as ministers of finance (Kokkvoll Tveit 2018).

According to one (presumably) particularly well-informed individual, Norway's non-compliance was rooted in patterns well known to practitioners of environmental politics: Harald Rensvik was secretary general¹¹ in the Ministry of the Environment from 1996 to 2011. Prior to that, he was director of the Norwegian Pollution Control Agency. When asked to elaborate on the main political and bureaucratic barriers to or drivers of compliance with ambitious agreements on air pollution, Rensvik replied as follows:

The most potent counter-forces are activated whenever environmental policies pose a threat to economic growth. Ministries, regardless of the sector they are responsible of, often seek to be their sector's best advocate and cheerleader. In recent decades, the interests of Norway's oil industry have been vigorously supported by the Ministry of Petroleum and Energy and other powerful ministries, such as the Ministry of Finance. Although exceptions exist, the political and bureaucratic forces backing the oil industry have tended to outweigh and overrun the forces backing interests such as environmental protection (Author's interview with Harald Rensvik, Oslo, March 2020).

As indicated by Rensvik, at the same time as its NO_x policies were passive, Norway pursued other (and highly emissions-intensive) goals: A number of domestic policies were in force aimed at exploring and extracting oil and gas from Norway's continental shelf. Among several examples of important political and bureaucratic inventions, we find measures taken in 2003 to speed up the process of granting production licenses. Stimulating exploration of new petroleum resources, a tax-incentive scheme was expanded to make more companies eligible for deductions. Moreover, Norwegian authorities removed restrictions on buying and selling of shares in existing projects (Tellmann 2012; Ryggvik et al. 2020).

Seemingly, Norway's production-enhancing policies were successful: During the 1990s and early 2000s, its petroleum production increased considerably (Statistics Norway 2018). So did calculations of the size of the native petroleum reserves. According to a Norwegian Petroleum Directorate (2017) report, estimates of Norway's total remaining petroleum resources went up by 40% between 1990 and 2017. Much of the upward adjustment of

⁹ Although dwarfed by Norway's total 1999 GDP of NOK 1265 billion (current prices), the annual compliance costs constitute a nonnegligible expenditure (Statistics Norway 2020).

¹⁰ On January 1, 2007, an NO_x emissions tax was imposed on the offshore oil and gas industry, on land-based industry, and on sea vessels. Together with other policies, this NO_x tax reduced Norway's emissions beyond BAU, but the reductions came all too late to reach the 2010 deadline.

¹¹ The secretary general is the highest ranking permanent bureaucrat in a Norwegian ministry.

remaining resources is due to the development and adoption of new technologies for petroleum discovery and extraction. Hence, at the same time as Norway extracted large amounts of oil and gas, it succeeded at discovering more and more petroleum, and at turning previously unextractable oil and gas into extractable resources.

What does the Norwegian case teach us about the effects of state capacity on compliance? That authorities used expert advice to monitor NOx emissions trajectories and abatement strategies resonates well with Chayes and Chayes' (1993) expectation: Being a high-capacity state, Norway used the knowledge of its trained bureaucrats and other experts to be sure to receive timely information if more stringent policies were needed to comply.

Surely, we have no counterfactual to measure the observed outcome against, but it seems plausible that Norwegian authorities were better informed about the prospects for compliance and alternative abatement measures than they would have been absent a strong and competent bureaucracy. So far, Norway's behavior is consistent with Fig. 1.

Norwegian authorities' actions upon the expert advice were, however, not in accordance with managerialists' expectations: The early warnings about the likely noncompliance were not followed by swift and firm policy change, but status quo—despite ample technological and economic opportunities to reduce emissions. Keeping in mind Fig. 1, Norway's bureaucratic resources were used to identify effective policies, but its fiscal resources, technological competence, and well-functioning state machinery was not put in motion to implement policies sufficient to reach compliance by the 2010 deadline.

Hence, the pathway from capacity to compliance illustrated in Fig. 1 did not fully materialize. Could there, however, be alternative mechanisms connecting capacity to compliance? Evidence from the Norwegian case indicates that such mechanisms indeed may exist: While proposals of potent NOx-reducing policies were repeatedly turned down, emissions-intensive activities were supported by the Norwegian state apparatus. Capable environmental bureaucrats investigated solutions to the NOx problem, but their ideas were never implemented. In contrast, the intellectual capabilities of bureaucrats in other sector ministries were used to aid the development of competing interests. As noted above, several policies were in force to incentivize extraction of oil and gas on the Norwegian continental shelf. Would these production-enhancing policies have been as successful if Norway was a medium-capacity or low-capacity state? Obviously, such counterfactual judgments are inherently uncertain—particularly so in a single case study. Considerable amounts of research, however, suggest that its high state capacity likely enhanced Norway's effectiveness as a petroleum extractor and its ability to make good use of the revenues.

In well-governed states (i.e., states with high scores on measures of bureaucratic quality and rule of law), natural resources increase growth (e.g., Robinson et al. 2006; Mehlum et al. 2006). Indeed, states with strong ex ante institutions tend to identify and implement policies that are enhancing their ability to transform natural resources into long-term economic wealth: As noted by Wright and Czelusta (2004), "Fears of impending [resource] scarcity have been overwhelmed by technological progress in exploration, extraction, and substitution over the past two centuries." Moreover, "returns to investments in country-specific minerals knowledge have stayed high in recent decades, so that production and reserve levels have continued to grow in well-managed resource economies." Such a simultaneous production and reserves increase is exactly what Norway achieved in the 1990s and early 2000s. Hence, Norway's development as a petroleum producer resonates well with the general pattern in countries with good institutions. Assessing the case of Norway specifically, Wright and Czelusta (2004, 22) concludes that "forecasts of impending depletion have been repeatedly overturned and reserve estimates adjusted" much due to the ingenuity in the domestic petroleum industry, the country's human resources, and its high-quality

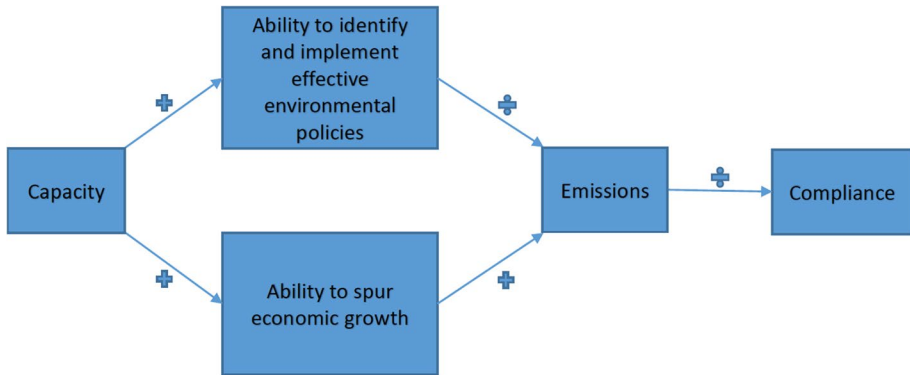


Fig. 2 Two effects of state capacity on compliance

institutions. Reviewing economic research on the case of Norway as a petroleum economy, Holden (2013) emphasizes the benevolent effects of its well-developed bureaucracy, arguing that Norway belongs to a group of countries in which “the [natural] resource has been used to the benefit of the country, leading to higher growth and income.”

The case study suggests that Fig. 1 should be supplemented by another pathway from state capacity to compliance. Building on Fig. 1, therefore, Fig. 2 shows *two* possible effects of capacity on compliance with emissions control protocols.

Figure 2’s top (positive) effect is identical to the pathway hypothesized by the management school. The bottom pathway, however, implies a negative effect of state capacity on compliance. While the core elements of state capacities—e.g., bureaucratic quality and fiscal resources—increase compliance in the top pathway, they take a different role in the bottom pathway: High state capacity now cause noncompliance. It does do by enhancing the state’s success in realizing a goal that affects compliance negatively. In the Norwegian case, that goal was petroleum-driven economic growth.

During our meeting, former secretary general Harald Rensvik reflected upon the various roles of bureaucracies in wealthy states with good institutions. “In my experience, the intellectual capacities of bureaucracies in countries such as Norway may be very high. Whenever tasked with finding the optimal solution to a problem, they often succeed. However, when the preferred policies from sectoral ministries conflict, the outcome is determined by the political forces backing each policy,” Rensvik said. In doing so, he illustrates some of the logic underpinning Fig. 2. Sector bureaucracies work to support the realization of various politically defined goals. Highly competent bureaucracies tend to provide better advice than lesser ones do. In cases where sector departments promote conflicting interests, the policy serving the interests with the strongest political backing is chosen. The policy promoted by the losing sector ministry never comes to fruition. In contrast, the policy promoted by the winner materializes. Because of the high quality of the bureaucracy, that policy is likely particularly well suited to realize the goals of the winning sector—in certain cases at the expense of the interests promoted by the losing side.

Given the two possible (and competing) effects of state capacity on compliance, what determines their relative strength? I argue that variations in states’ intentions to comply may be key to understand how high state capacity may cause noncompliance. Consider again the Norwegian case. Although intentions are challenging to measure, it seems safe to argue that Norway’s intention to comply was nonzero. Former secretary general Harald

Rensvik's account suggest that the Ministry of the Environment and its underlying bodies performed many tasks aiming at reducing NO_x emissions. Alternative policies were investigated, cost analyses were conducted, and projections were used to estimate how deep emissions cuts each policy alternative would produce. In short, Norway took every action needed to comply, except from the most important one—enacting stringent policies. Compliance was not reached by 2010 because competing goals were given a higher priority by political decision makers: “Attempts to reduce NO_x emissions ran again and again into the stumbling block of economic interests backed at the top political level,” former secretary general Rensvik said.

The Norwegian case therefore suggests that we may think of intention to comply as a relative, not absolute, concept: Authorities can ascribe some value to compliance, but if they ascribe even more value to competing goals, the latter wins. While not explicitly represented in Fig. 2, intention to comply determines whether the positive effect dominates the negative effect or vice versa. If intention to comply is stronger than the will to pursue competing goals, the positive effect will likely prevail. If intention to comply is (relatively) weak, the negative effect may well dominate, so that state capacity affects compliance negatively.

The discussions above imply that capacity may have a negative effect on compliance in a state satisfying three criteria. First, the state must pursue at least one policy goal that correlate negatively with compliance (i.e., the more successful the state is in realizing that goal, the lower the compliance). Second, the realization of that policy goal must be aided by state capacity. Third, the state must be reluctant to comply.

Conditions one and two likely hold for at least some CLRTAP members. Economic growth is but one of many policy goals that may conflict with emissions reductions. Virtually all emissions of SO_x and NO_x derive from fossil fuel combustion.¹² Unless decoupled from emissions, economic growth correlates negatively with compliance with SO_x and NO_x targets. In principle, such decoupling is possible. However, a reluctant state will unlikely implement emissions-reducing measures if the private costs of doing so outweigh the benefits. Notwithstanding some variation between the regulated substances (Murdoch et al. 1997; Wettstad 2012), compliance with CLRTAP protocol targets often entail considerable beyond BAU costs.

What about the third condition? What data can be used to score values on the intention-to-comply variable? In this regard, my case study of Norway provides two take-home messages. First, it is highly challenging to identify an operationalization of intention to comply that is valid and applicable across protocols and countries. Consider democracy (Neumayer 2002)¹³ or post-materialism (Inglehart 1981), two likely candidates. Despite Norway's status as one of the world's most stable and highly developed democracies, my case study suggests that Norway's intention to comply was relatively low most of the period between Gothenburg's adoption and deadline. Second, in-depth studies may reveal such intentions—but it requires considerable amounts of data. Indeed, in large-N or even medium-N studies, collecting in-depth data sufficient to distinguish the enthusiasts from the reluctant states could require more resources than many research projects have.

To summarize, state capacity may have positive, zero, or even negative effect on compliance, depending on the country's intention to comply—but revealing such intentions is very data-intensive. Lacking a better basis from which an empirical expectation

¹² Four of the five protocols I study include targets for SO_x, NO_x, or both pollutants.

¹³ Neumayer (2002) finds that on average, democracies show more “commitment” in environmental issues.

can be derived, I formulate the following hypothesis in light of the Norwegian case introduced above:

H3 State capacity influences (the probability of) compliance negatively.

4 Case selection and data

My dataset includes targets for all European states that are parties to one or more of my five CLRTAP protocols and who became parties to the protocol involved *no later than the year prior to the deadline*. These five CLRTAP protocols are well suited for statistical assessment of capacity's effect on compliance, for two reasons. First, because compliance can be measured precisely, I need not rely on coders' judgment of compliance (as some previous studies do). Second, as my operationalization of compliance (below) shows, CLRTAP protocols enable measuring the *degree* of compliance. Missing a target by 5% is less serious than missing it by 10%. Hence, such differences should be reflected in the data (Dai 2007, 15).

This study's units of observation are emissions targets rather than agreements or states. More precisely, each of my 176 units constitutes an emissions target for a given state specified by some CLRTAP protocol. Since the CLRTAP protocols include national emissions targets for four substances, my units are "protocol-country-substances." For example, because Germany became a party to the 1999 Gothenburg protocol in 2004, and Gothenburg specifies national emissions targets for four substances (NO_x, sulfur, NMVOC, and ammonia), four units in my dataset correspond to Germany's emissions targets in the Gothenburg protocol.

5 Definitions and operationalizations

5.1 Dependent variable

The criterion for being compliant is straightforward: The CLRTAP protocols include emissions targets for 1–4 substances for each party, and contain no provisions that can relieve parties from the obligation to reach their designated targets by the deadline—except by withdrawing from the protocol. Thus, for a given emissions target, a state is in compliance only if its deadline year emissions of the substance concerned were no higher than the target.

I operationalize compliance in two ways—as a dichotomy and as a continuous variable. The dichotomous variable scores 1 if the national emissions target was reached by the deadline and 0 otherwise. On the continuous variable, a unit's score equals its positive or negative deviance from its target when the deadline expires. For instance, Germany's deadline year (2010) NO_x emissions were 23.4% higher than allowed by the 1999 Gothenburg protocol. Consequently, the Gothenburg–Germany–NO_x unit scores –0.234 on the continuous compliance variable.

5.2 Independent variables

Being a contested concept, state capacity is challenging to measure (Jänicke 1997; Hendrix 2010). For want of a generally accepted operationalization, I use two main operationalizations that have been used in previous compliance studies, were suggested by theorists of the management school, have high face validity, and allow comparison across states. The first is the World Bank's (2017) Government Effectiveness indicator,¹⁴ and the second is GDP/cap (log-transformed).¹⁵ Moreover, "Appendix" includes robustness checks where I use a third alternative operationalization.

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." Performance estimates range from approximately -2.5 (weak) to 2.5 (strong). Because most countries' score varies between the adoption and deadline of each protocol, I use the average of all available scores within that period. For instance, for all units from the 1999 Gothenburg protocol, I use countries' average Government Effectiveness scores from 1999 to 2010.

My second main operationalization of state capacity, (log) GDP/cap, is supported by a widespread expectation in the literature that state capacity depends strongly on economic development (Jänicke 1997; Chayes and Chayes 1993, 194; Jacobson and Brown Weiss 1998, 536). Note, however, that in the literature seeking to explain emissions trajectories more generally (see Fiorino's 2011 review) it is often argued that such economic factors may be important *drivers* of emissions. Nonetheless, because I prefer to stay close to Chayes and Chayes' (1993, 195–204) own understanding of their key concepts, I use (log) GDP/cap as one of my two primary operationalizations of capacity.

As a robustness check, I have also run models operationalizing state capacity as the Varieties of Democracy (V-Dem) project's most valid measure of national public administrations' quality (Pemstein et al. 2018).¹⁶ Note also that some of my models control for governments' autonomy from domestic veto players (Henisz 2002), which some previous studies (e.g., Börzel et al. 2010; see also Meckling and Nahm 2018) use to operationalize state capacity.

The operationalization of ambition level is a function of the size of the target and the emissions level *when the protocol was adopted*. To illustrate, consider again Germany's NO_x target in the 1999 Gothenburg protocol. Germany's NO_x emissions in 1999 were 83% higher than the emissions target for 2010. Hence, the Gothenburg–Germany–NO_x unit scores 1.83 on the ambition-level variable.

Because EU directives targeting air pollutants have been in force since the 1980s (see Knill and Lenschow 1998; Wettestad 2012; Jahn 2016), some models include dummies separating states that were EU members (by the relevant protocol's deadline year) from nonmembers. My dataset also includes a measure of countries' GDP growth between each protocol's adoption and its deadline year (World Bank 2018).

On GDP/cap, the quality of public administrations, and government autonomy, values equal the average score between protocol adoption and deadline.

¹⁴ WGI scores are based on surveyed views of experts, citizens, and enterprise respondents.

¹⁵ I log-transform GDP/cap because its relationship to capacity is likely nonlinear.

¹⁶ This measure is labeled "Rigorous and impartial public administration."

Moreover, emissions export and import patterns might condition compliance (e.g., Volenweider 2013). Hence, some models control for domestic sources' share of the total depositions of a given substance in a country (EMEP 2009). Because equivalent export–import data exist only for two of my substances (sulfur and NO_x), this domestic depositions variable is included in only four of my 12 (main) models.

6 Estimation and research design

Assessing the causal relationship between independent variables such as state capacity and compliance is challenging: States self-select into treaty participation and may set targets that are (more or less) consistent with BAU. Moreover, the units of observation are not independent of each other.

Such challenges suggest considering one of the designs proposed by the “causal inference” literature (see Angrist and Pischke 2009)—for instance instrumental variables (IV), regression discontinuity design (RDD), and difference-in-differences (DID).

However, neither IV, RDD nor DID would be appropriate for the present paper's research question and data. First, no valid instrument Z seems to exist for my capacity variable.¹⁷ Second, using the fact that some thresholds as good as randomly assign units into treatment and control groups, RDD studies estimate the treatment effect by comparing units just above some threshold to units just below it. However, my capacity variable displays no such threshold. Finally, state capacity does not display the characteristics required by the DID estimator: In DID studies, the independent (treatment) variable of interest must somehow be “turned on” (typically, change from 0 to 1 among the treated units and remain 0 among nontreated units) at a point within a series of repeated observations. Hence, DID is less suited in studies such as mine, where the independent variable of interest (capacity) would have a nonzero value at all times for all units.

Arguably, there might be other ways to alleviate some of the above-mentioned challenges. First, my ambition-level variable reduces the risk of a biased estimate of capacity's effect on compliance. Second, dependence between observations for the same state can be addressed by clustering standard errors. Third, dependence between observations in the same protocol is accounted for by introducing protocol dummies.

7 Results

Table 1 shows the results of six models using the continuous compliance variable. In Models 1–3 the Government Effectiveness indicator operationalizes state capacity, while I use (log) GDP/cap in Models 4–6. Due to missing data, N decreases when I add the economic growth and domestic depositions variables. Hence, I add those variables to my models in a stepwise process.

Regardless of specification, the effects of my state capacity measures are consistently negative and statistically significant (albeit in Model 3 only at the 10% level). Considering that Model 3 includes six independent variables, while N equals only 99, the lack of strongly significant estimates in that model is unsurprising. In Model 1, the effect of Government Effectiveness is estimated at -0.108 . Hence, increasing Government Effectiveness by one scale unit decreases the expected compliance score by 10.8%. One scale unit on Government Effectiveness corresponds approximately to the difference between the

¹⁷ A valid instrument Z is an independent variable that affects the dependent variable exclusively through the original independent variable of interest (Angrist and Pischke 2009, 117).

average scores between 1999 and 2010 of Finland, a typical high-capacity state, and Portugal, often thought of as a medium- or low-capacity European state.

Unsurprisingly, the effect of ambition level is negative in all six models and statistically significant in four. The rest of my controls have no effect. The zero effect of EU membership (after controlling for ambition level) may seem somewhat surprising given the fact that EU law long has targeted air pollutants. The zero effect of government autonomy suggests that domestic veto players are unimportant for compliance with CLRTAP protocols. Finally, my results suggest that countries' compliance decisions are unaffected by the degree to which they have other countries' emissions deposited within their territories.

I now turn to the regressions using the dichotomous dependent variable (Table 2). Again, in three models (7–9) the Government Effectiveness indicator operationalizes state capacity, while I use (log) GDP/cap in Models 10–12. The effect of capacity remains negative across all regressions, and the effect is statistically significant in all but one model.¹⁸ Figure 3 (“Appendix”) visualizes the results of Model 7, showing graphically how the predicted probability of compliance = 1 decreases when scores on Government Effectiveness increase.

As in Table 1, ambition level is the only control that affects compliance, except that GDP growth has a positive and marginally significant effect in Model 12.

Several sensitivity checks support the results in Tables 1 and 2. First, I report checks in “Appendix” (Table 5) showing that an alternative state capacity measure (the quality of national public administrations) has a consistently negative effect on compliance, thereby echoing the results in Tables 1 and 2. Third, the effect of capacity is consistently negative

Table 1 OLS regressions; dependent: compliance (continuous)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Govt. Effectiv.	-0.108*	-0.132*	-0.119 ⁺			
	(0.042)	(0.048)	(0.059)			
(log) GDP/cap				-0.126**	-0.147**	-0.176**
				(0.037)	(0.038)	(0.060)
Ambition level	-0.330*	-0.372*	-0.367	-0.310*	-0.354*	-0.332
	(0.148)	(0.171)	(0.229)	(0.145)	(0.164)	(0.218)
Govt. autonomy	-0.251	-0.153	-0.225	-0.356 ⁺	-0.311	-0.550
	(0.164)	(0.357)	(0.469)	(0.193)	(0.410)	(0.529)
EU member	-0.030	0.000	-0.038	-0.028	-0.018	-0.034
	(0.054)	(0.069)	(0.063)	(0.052)	(0.060)	(0.060)
GDP growth		-0.299	-0.184		-0.283	-0.251
		(0.286)	(0.272)		(0.245)	(0.259)
Domest. depos.			0.204			0.071
			(0.333)			(0.299)
_cons	0.857**	0.971**	0.987**	2.001**	2.322**	2.707**
	(0.178)	(0.318)	(0.321)	(0.414)	(0.553)	(0.775)
N	176	161	99	174	161	99
R ²	0.230	0.256	0.242	0.246	0.277	0.271

Standard errors (clustered on states) in parentheses

⁺ $p < 0.10$; * $p < 0.05$; ** $p < 0.01$

¹⁸ Additionally, the effect is only weakly significant in Model 8.

Table 2 Logit regressions; dependent: compliance (dichotomous)

	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
Govt. Effectiv.	-1.096*	-0.926 ⁺	-0.656			
	(0.492)	(0.496)	(0.659)			
(log) GDP/cap				-1.088**	-1.005*	-1.479*
				(0.388)	(0.405)	(0.603)
Ambition level	-1.487**	-1.426**	-1.232*	-1.404*	-1.352**	-1.108*
	(0.534)	(0.499)	(0.569)	(0.573)	(0.500)	(0.505)
Govt. autonomy	4.344	3.515	8.562 ⁺	3.564	2.818	6.644
	(3.622)	(3.452)	(4.792)	(3.828)	(3.609)	(5.133)
EU member	0.138	0.114	-1.036	-0.034	-0.116	-1.253
	(0.532)	(0.608)	(1.280)	(0.548)	(0.592)	(1.040)
GDP growth		1.428	2.618		2.089	2.514 ⁺
		(1.452)	(1.828)		(1.302)	(1.286)
Domest. depos.			3.039			1.646
			(2.733)			(2.676)
_cons	3.949*	3.409*	2.140	13.817**	12.578**	17.480*
	(1.679)	(1.739)	(2.026)	(4.672)	(4.830)	(7.294)
<i>N</i>	176	161	99	174	161	99
pseudo <i>R</i> ²	0.163	0.147	0.173	0.174	0.162	0.202

Standard errors (clustered on states) in parentheses

⁺ $p < 0.10$; * $p < 0.05$; ** $p < 0.01$

even after adding protocol dummies (Table 6). Finally, also when I include *substance* dummies, the effect of capacity remains negative and statistically significant (Table 7).

Thus, none of my twelve main models and none of my sensitivity checks support the proposition that capacity increases compliance (H1). In contrast, because the effect of capacity is consistently negative, my results lend considerable support to H3. However, because the negative effect of capacity is statistically insignificant in one of my main models (and statistically significant only at the 10% level in another two models), H2 (a zero effect of capacity) cannot be entirely dismissed.

8 Conclusion

My regressions suggest that state capacity has a negative effect on compliance with the five CLRTAP protocols included in my dataset. This finding not only contradicts the hypothesis proposed by management scholars—that enhanced state capacity increases compliance.¹⁹ The consistently negative and largely statistically significant estimates appear puzzling because no previous studies have provided theoretical arguments that can explain such a finding. I have, however, provided a novel conjecture concerning the capacity–compliance relationship that may explain the observed pattern. Among reluctant states that pursue policy goals correlating

¹⁹ This finding does, however, not rule out that the other explanatory factors suggested by managerialists—treaty ambiguity and unexpected changes over time—can account for CLRTAP protocol noncompliance.

negatively with compliance, we should indeed expect a negative effect of capacity on compliance when the successful pursuit of these other goals is aided by state capacity.

To illustrate the mechanisms driving the negative effect of state capacity, I have argued that Norway's high capacity likely contributed to noncompliance with the Gothenburg protocol. The case study reveals that some of the mechanisms through which high state capacity could have enhanced compliance indeed were at work. However, their effects were repeatedly blocked at the political decision-making stage. Hence, the positive effect of state capacity hypothesized by managerialists failed to materialize. Norway's high state capacity was, however, affecting compliance through other causal pathways: Backed by powerful political actors, the highly competent petroleum bureaucracy aided the development of the oil and gas industry. Through bureaucratic inventions and ingenuity, ministries administering petroleum policies likely helped Norway develop and extract its native resources more effectively than it would have been able to absent its high capacity. Because oil and gas extraction is highly emissions-intensive, Norway's state capacity may indeed have caused noncompliance with its NO_x target under the Gothenburg protocol.

The present study has at least two important limitations. First, we can only draw limited conclusions based on a stand-alone case study such as my assessment of Norway's NO_x policies. Hence, future in-depth studies should assess the causal pathways through which I argue high state capacity may cause noncompliance. By using rigorous process-tracing methods (Bennett and Checkel 2014), future research might examine the proposed mechanisms even more closely than I have been able to do. Nonetheless, my case study likely provides useful starting points for such future studies: First, they should identify political goals that, to the extent they are reached, could affect compliance negatively. Second, researchers should assess whether high state capacity enhances the realization of the goals competing with compliance. Third, researchers should seek to identify specific points in political processes where authorities decide which of the competing goals they prioritize, thereby suggesting how strong their intentions to comply are.

The present paper's second and most important limitation is its failure to include any measure of intentions to comply in the statistical analyses. Accordingly, my novel conjecture is yet to pass a tough empirical test. The conjecture was developed as a response to my statistical analyses' strong suggestions of a negative effect of state capacity on compliance with CLRTAP protocols. Facing that puzzling finding, I explored plausible explanations using an in-depth case study. After indeed having identified causal pathways that could explain the statistical findings, I formalized and integrated them into a model (Fig. 2) that contrasts my conjecture with managerialist propositions of how state capacity affects compliance. Although my theoretical argument receives modest empirical support from the case study of Norway, it is up to future research to put it to proper empirical tests. The question is if there exist any valid measures of such intentions that can be used in future large-N studies, without requiring prohibitive amounts of resources for data collection. Keeping in mind that several candidates (such as scores on democracy or post-material values) appear crude and inaccurate in the case of Norway's NO_x policies, one could doubt the validity of many "usual suspects" for operationalizing such intentions. Hence, a good starting point for future research is to conduct in-depth analyses vetting the mechanisms implied by my theoretical argument. If any intentions to comply measures fit for large-N analysis do exist, they will likely be identified by such in-depth studies.

Although developed in the context of CLRTAP protocols, my theoretical conjecture is in principle general and might therefore be applicable in other contexts. Indeed, if states pursue policy goals that conflict with compliance, and the realization of these goals is facilitated by state capacity, state capacity may also cause noncompliance in other environmental issue

areas than air pollution. Such an effect might even be present in nonenvironmental issue areas, for instance human rights or trade. It remains, however, a task for future empirical research to assess whether my theoretical conjecture is consistent with data from those issue areas.

Although my statistical findings seem highly robust in the context of CLRTAP protocols, one particular feature of my units suggests a cautious approach to generalization: Most of the CLRTAP states are located in the middle or even the upper end of the global capacity distribution. This is not to say that the variance concerning capacity in my dataset is low. Indeed, between 1985 and 1993 the average GDP/cap of the least wealthy party to the Helsinki Protocol (Ukraine) equaled only 9.4% of the average GDP/cap of the wealthiest party (Luxembourg). Still, the variance on capacity would certainly be higher if I had studied a global agreement such as the 1997 Kyoto Protocol. However, global regimes that include clear-cut commitments (thereby enabling measurement of compliance) are few and far between. Hence, CLRTAP and its protocols provide a relatively good case for estimating the effect of capacity on compliance.

Acknowledgements I am most grateful to Jon Hovi, two anonymous reviewers, and participants at a panel at the Annual MPSA Conference, Chicago, April 2018, for constructive comments and suggestions. Any remaining errors are the author's responsibility.

Funding Open Access funding provided by Institute Of Transport Economics.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix: Does capacity increase compliance?

Table 3 shows the Pearson correlations between all the independent variables included. Unsurprisingly, my three alternative operationalizations of state capacity (Government Effectiveness, (log) GDP/cap, and Rigorous and impartial public administration) are highly correlated. The remaining independent variables correlate much less and can therefore be included in the same models. Table 4 shows descriptive statistics for the independent variables. Figure 3 visualizes Model 7's results (Table 2) by plotting the predicted probabilities of compliance = 1 for different Government Effectiveness scores. Table 5 shows OLS and logit regressions revealing a consistently negative and mainly statistically significant effect on compliance of the Rigorous and impartial public administration measure. Table 6 shows three OLS models (one for each of my operationalization of state capacity) including protocol dummies. The effects of the state capacity measures are consistently negative. Similarly, Table 7 shows that my results hold when I include substance dummies.

Table 3 Pearson correlations

	Government Effectiveness	(log) GDP/cap	Rigor./impart. publ. adm.	GDP growth	Govt. auton.	EU member	Ambition level	Domestic depositions
Government Effectiveness	1.0000							
(log)GDP_cap	0.9153	1.0000						
Rigor./imp. public adm.	0.9248	0.8389	1.0000					
GDP_growth	-0.3893	-0.3627	-0.3823	1.0000				
Govt. autonomy	-0.6090	-0.4996	-0.6318	0.2843	1.0000			
EU member	0.3643	0.3678	0.3828	0.0836	-0.3309	1.0000		
Ambition level	0.3462	0.4182	0.3518	-0.2075	-0.1958	0.3123	1.0000	
Domestic depositions	-0.1161	-0.1765	-0.0400	-0.1287	-0.0277	0.2146	0.1808	1.0000

Table 4 Descriptives

Variable	Obs	Mean	SD	Min	Max
Govt_Effectiveness	176	1.286152	0.6986915	-0.71	2.1300
(log) GDP_cap	174	10.15503	0.8052837	7.917536	11.51641
Rigor./impart. public adm.	174	2.385086	1.127337	-0.886	3.979
GDP_growth	161	0.2737379	0.1717864	-0.2515026	0.7141402
Govt. autonomy	176	0.245362	0.1223502	0.1074617	0.9885709
EU member	176	0.7329545	0.4436785	0	1
Ambition level	176	1.181534	0.4723542	0.2760654	3.499148
Domestic dep.	112	0.2189944	0.1830875	0	0.75

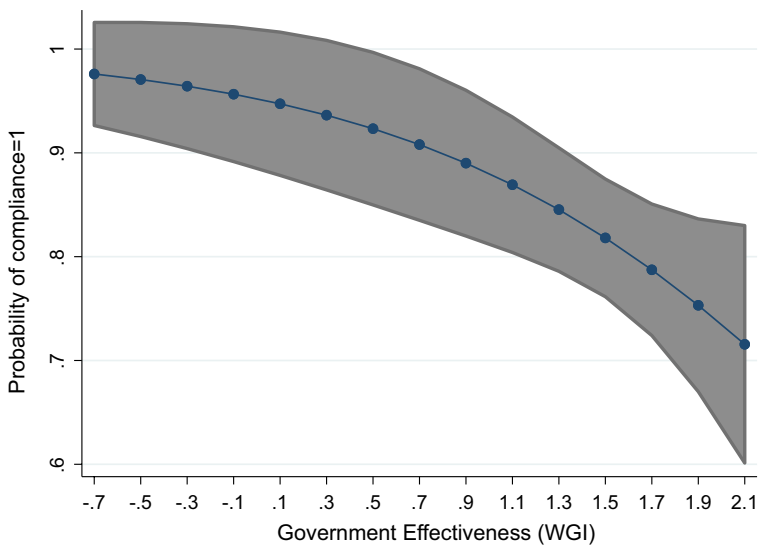
**Fig. 3** Predicted probabilities of compliance (95% confidence intervals)

Table 5 OLS and logit regressions; dependent: compliance

	Model 13	Model 14	Model 15	Model 16	Model 17	Model 18
Rigor./impart. public adm.	-0.063* (0.029)	-0.075* (0.032)	-0.043 (0.043)	-0.772* (0.305)	-0.665+ (0.343)	-0.156 (0.375)
Ambition level	-0.322* (0.155)	-0.367* (0.173)	-0.373 (0.232)	-1.327** (0.472)	-1.325** (0.466)	-1.215* (0.573)
Govt. autonomy	-0.236 (0.156)	-0.143 (0.348)	-0.061 (0.498)	2.731 (2.983)	2.412 (3.096)	8.980+ (4.749)
EU member	-0.014 (0.056)	0.011 (0.074)	-0.042 (0.065)	0.248 (0.528)	0.253 (0.623)	-1.072 (1.384)
GDP growth		-0.291 (0.293)	-0.122 (0.280)		0.738 (1.699)	2.970 (2.156)
Domestic depositions			0.267 (0.336)			3.497 (2.867)
_cons	0.842** (0.168)	0.961** (0.316)	0.874* (0.331)	4.484** (1.525)	4.019* (1.852)	1.275 (1.871)
<i>N</i>	174	161	99	174	161	99
<i>R</i> ²	0.220	0.247	0.229			
Pseudo <i>R</i> ²				0.171	0.154	0.166

Standard errors (clustered on states) in parentheses

+*p* < 0.10; **p* < 0.05; ***p* < 0.01

Table 6 OLS regressions including protocol dummies

	Model 19	Model 20	Model 21
Govt_Eff_WGI	-0.147** (0.053)		
(log) GDP/cap		-0.167** (0.040)	
Rigor./impart. public adm			-0.079* (0.035)
Ambition level	-0.372+ (0.199)	-0.353+ (0.189)	-0.371+ (0.203)
Govt. autonomy	-0.340 (0.388)	-0.535 (0.443)	-0.294 (0.371)
EU member	0.005 (0.069)	-0.018 (0.061)	0.014 (0.073)
GDP growth	-0.430 (0.362)	-0.435 (0.304)	-0.428 (0.373)
Sofia protocol	-0.411** (0.099)	-0.408** (0.095)	-0.410** (0.102)
Oslo protocol	0.093 (0.109)	0.114 (0.102)	0.089 (0.114)
Geneva protocol	-0.287+ (0.146)	-0.298* (0.141)	-0.281+ (0.152)
Gothenburg protocol	-0.148 (0.105)	-0.130 (0.099)	-0.132 (0.111)
_cons	1.201** (0.320)	2.743** (0.531)	1.172** (0.319)
<i>N</i>	161	161	161
<i>R</i> ²	0.372	0.402	0.359

Standard errors (clustered on states) in parentheses

Reference category (protocol dummies): the Helsinki protocol

+ $p < 0.10$; * $p < 0.05$; ** $p < 0.01$

Table 7 OLS regressions including substance dummies

	Model 22	Model 23	Model 24
Govt_Eff_WGI	-0.156** (0.055)		
(log) GDP/cap		-0.165** (0.041)	
Rig/imp publ adm			-0.081* (0.037)
Ambition level	-0.373* (0.148)	-0.352* (0.141)	-0.369* (0.151)
Govt. autonomy	-0.399 (0.421)	-0.551 (0.475)	-0.338 (0.408)
EU member	0.025 (0.067)	0.002 (0.060)	0.034 (0.072)
GDP growth	-0.330 (0.298)	-0.303 (0.245)	-0.305 (0.314)
NOx	-0.466** (0.089)	-0.468** (0.087)	-0.459** (0.090)
Ammonia	-0.414** (0.048)	-0.406** (0.046)	-0.402** (0.050)
NMVOc	-0.328** (0.060)	-0.335** (0.059)	-0.320** (0.062)
_cons	1.295** (0.319)	2.796** (0.590)	1.249** (0.322)
<i>N</i>	161	161	161
<i>R</i> ²	0.513	0.536	0.496

Standard errors (clustered on states) in parentheses

Reference category (substance dummies): Sulfur

+*p* < 0.10; **p* < 0.05; ***p* < 0.01

References

- Aakre, S., Helland, L., & Hovi, J. (2016). When does informal enforcement work? *Journal of Conflict Resolution*, 60(7), 1312–1340.
- 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.
- Bennett, A., & Checkel, J. (Eds.). (2014). *Process tracing: From metaphor to analytic tool*. Cambridge: Cambridge University Press.
- 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>.
- Breitmeier, H., Young, O. R., & Zürn, M. (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.
- Chayes, A., & Chayes, A. H. (1993). On compliance. *International Organization*, 47(2), 175–205. <https://doi.org/10.1017/S0020818300027910>.

- Chayes, A. H., Chayes, A., & Mitchell, R. B. (1995). Active compliance management in environmental treaties. In W. Lang (Ed.), *Sustainable development and international law* (pp. 75–89). London: Graham & Trotman.
- Dai, X. (2007). *International institutions and national policies*. Cambridge: Cambridge University Press.
- Demmke, C. (2001). *Towards effective environmental regulation: innovative approaches in implementing and enforcing European Environmental Law and Policy*. Jean Monnet Working Paper 5/01. Cambridge, MA: Harvard Law School.
- Downs, G. W., Rocke, D. M., & Barsboom, 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>.
- EMEP. (2009). *EMEP Status Report 1/09. Transboundary acidification, eutrophication and ground level ozone in Europe in 2007*. http://emep.int/publ/common_publications.html#2009.
- Fiorino, D. (2011). Explaining national environmental performance: Approaches, evidence, and implications. *Policy Sciences*, 44(4), 367–389.
- Hendrix, C. S. (2010). Measuring state capacity: Theoretical and empirical implications for the study of civil conflict. *Journal of Peace Research*, 47(3), 273–285.
- Henisz, W. J. (2002). The institutional environment for infrastructure investment. *Industrial and Corporate Change*, 11(2), 355–389. <https://doi.org/10.1093/icc/11.2.355>.
- Holden, S. (2013). Avoiding the resource curse the case Norway. *Energy Policy*, 63, 870–876. <https://doi.org/10.1016/j.enpol.2013.09.010>.
- Inglehart, R. (1981). Post-materialism in an environment of insecurity. *American Political Science Review*, 75(4), 880–900. <https://doi.org/10.2307/1962290>.
- Jacobson, H. K., & Brown Weiss, E. (1998). Assessing the record and designing strategies to engage countries. In E. B. Weiss & H. K. Jacobson (Eds.), *Engaging countries: Strengthening compliance with international environmental accords* (pp. 511–554). Cambridge: MIT Press.
- Jahn, D. (2016). *The politics of environmental performance. Institutions and preferences in industrialized democracies*. Cambridge: Cambridge University Press.
- Jänicke, M. (1997). The political system's capacity for environmental policy. In M. Jänicke & H. Weidner (Eds.), *National environmental policies. A comparative study of capacity-building* (pp. 1–24). Berlin: Springer.
- Knill, C., & Lenschow, A. (1998). Coping with Europe: The impact of British and German administrations on the implementation of EU environmental policy. *Journal of European Public Policy*, 5(4), 595–614.
- Kokkvoll Tveit, A. (2018). Norms, incentives, or deadlines? Explaining Norway's noncompliance with the Gothenburg protocol. *Global Environmental Politics*, 18(1), 76–98.
- König, T., & Mäder, L. (2014). The strategic nature of compliance: An empirical evaluation of law implementation in the central monitoring system of the European Union. *American Journal of Political Science*, 58, 246–263. <https://doi.org/10.1111/ajps.12038>.
- Mastenbroek, E. (2005). EU compliance: Still a 'black hole'? *Journal of European Public Policy*, 12(6), 1103–1120. <https://doi.org/10.1080/13501760500270869>.
- Meckling, J., & Nahm, J. (2018). The power of process: State capacity and climate policy. *Governance*, 31, 741–757. <https://doi.org/10.1111/gove.12338>.
- Mehlum, H., Moene, K., & Torvik, R. (2006). Institutions and the resource curse. *The Economic Journal*, 116(508), 1–20.
- 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.
- Mitchell, R. B. (2009). *International politics and the environment*. London: SAGE Publishing.
- Murdoch, J. C., Sandler, T., & Sargent, K. (1997). A tale of two collectives: Sulfur versus nitrogen oxides emission reduction in Europe. *Economica*, 64(254), 281–301.
- Neumayer, E. (2002). Do democracies exhibit stronger international environmental commitment? A cross-country analysis. *Journal of Peace Research*, 39(2), 139–164. <https://doi.org/10.1177/0022343302039002001>.
- Norwegian Environment Agency. (2015). *Air pollution*. www.environment.no/topics/air-pollution/. Accessed 2 March 2018.
- Norwegian Petroleum Directorate. (2017). *Annual resource report*. www.npd.no/en/facts/publications/reports2/resource-report/resource-report-2017/. Accessed September 2019.
- Norwegian Pollution Control Authority. (1999). *Reduksjon av NO_x-utslipp i Norge. Tiltaksanalyse for måååret 2010. SFT-rapport 99/13*.

- Pemstein, D., Marquardt, K. L., Tzelgov, E., Wang, Y., Krusell, J., & Miri, F. (2018). *The V-Dem measurement model: Latent variable analysis for cross-national and cross-temporal expert-coded data* (3 ed.). University of Gothenburg, Varieties of Democracy Institute: Working Paper No. 21.
- Robinson, J. A., Torvik, R., & Verdier, T. (2006). Political foundations of the resource curse. *Journal of Development Economics*, 79(2), 447–468.
- Ryggvik, H., Smith-Solbakken, M., & Tollaksen, T. G. (2020). «Norsk oljepolitikk». *Store norske leksikon*. Available from snl.no/Norsk_oljehistorie. Accessed 5 June 2020.
- Statistics Norway. (2018). *Net petroleum production*. www.ssb.no/statbank/table/09317/tableViewLayout1/. Accessed 18 September 2019.
- Statistics Norway. (2020). *Nasjonalregnskap*. www.ssb.no/statbank/table/09189/tableViewLayout1/. Accessed February 2020.
- Tellmann, S. M. (2012). The constrained influence of discourses: The case of Norwegian climate policy. *Environmental Politics*, 21(5), 734–752. <https://doi.org/10.1080/09644016.2012.692936>.
- Underdal, A., & Hanf, K. (Eds.). (2000). *International environmental agreements and domestic politics: The case of acid rain*. Aldershot: Ashgate.
- Victor, D. G. (2011). *Global warming gridlock: Creating more effective strategies for protecting the planet*. Cambridge: Cambridge University Press.
- Victor, D., Raustiala, K., & Skolnikoff, E. B. (Eds.). (1998). *The implementation and effectiveness of international environmental commitments: Theory and practice*. London: MIT Press.
- Vollenweider, J. (2013). The effectiveness of international environmental agreements. *International Environmental Agreements: Politics, Law and Economics*, 13(3), 343–367.
- 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*. <http://info.worldbank.org/governance/wgi/index.aspx#home>. Accessed 10 February 2018.
- World Bank. (2018). *World development indicators*. <http://databank.worldbank.org/data/source/world-development-indicators#>. Accessed 30 November 2018.
- Wright, G., & Czelusta, J. (2004). The myth of the resource curse. *Challenge*, 47(2), 6–38.
- Young, O. 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(50), 19853–19860.
- Zürn, M., & Joerges, C. (Eds.). (2005). *Law and governance in postnational Europe*. Cambridge: Cambridge University Press.