Commitment Ambiguity and Prudence in Climate Pledges

Vegard Tørstad [*‡]

Vegard Wiborg [†‡]

December 6, 2022

Abstract

This article evaluates how ambiguity affects cooperation under the Paris Agreement on climate change, which embodies a pledge-and-review system wherein governments unilaterally decide on the depth of commitments. We challenge the claim that ambiguity undermines compliance in international institutions. Instead, we theorize that exogenous ambiguity—uncertainty originating in governments' inability to perfectly gauge their commitment potential—undermines ambition. This prudence effect of ambiguity is driven by compliance concern: the anticipated international and domestic audience costs arising from non-compliance induce policymakers to adjust ambition downwards. Our empirical analysis of all climate pledges under the Paris Agreement demonstrates that ambiguous pledges are less ambitious than precise pledges, in line with the prudence conjecture. We also show that democracies are more prudent than autocracies, which we attribute to systemic variations in domestic audience costs. Our findings bear implications for how ambiguity relates to ambition and compliance under self-reporting regimes in international cooperation.

Keywords: Ambiguity; compliance; international environmental agreements; Paris Agreement; self-reporting; transparency.

[*]Department of Political Science, University of Oslo. Email: v.h.torstad@stv.uio.no

[†]Department of Economics, University of Oslo.

[‡]We are grateful to Sarah Anderson, Xun Cao, Christina Davis, Alexander Gard-Murray, Federica Genovese, Jon Hovi, Bob Keohane, Vally Koubi, Sam Rowan, Ken Scheve, Håkon Sælen, Manuel Teodoro, Dustin Tingley, Mike Tomz, Azusa Uji, Francesca Vantaggiato, Tore Wig, Noah Zucker, and participants at the 2020 EPG Conference, 2021 MPSA Annual Conference, 2021 PEIO Conference, and 2022 Norwegian Political Science Conference for helpful comments and discussions. We also thank Andreas Kokkvoll Tveit for contributing to the conjoint experiment; and Yann Robiou du Pont, Nicolas Lancesseur, and Peiran Liu for sharing data. This work was supported by Research Council of Norway (projects no. 261491 and 223274) and BI Norwegian Business School.

1 Introduction

The Paris Agreement established a pledge-and-review system under which governments are supposed to undertake progressively ambitious climate policies. Acknowledging the infeasibility of a top-down approach with internationally negotiated mitigation targets, the Agreement allows states to individually self-determine the ambition level of mitigation commitments (Falkner 2016). Every fifth year, states are required to pledge new targets and climate actions through the submission of Nationally Determined Contributions (NDCs). After an implementation period, states' compliance with their self-determined pledges is subject to technical expert review and scrutiny by the Agreement's compliance committee. Despite these independent monitoring mechanisms, the mitigation commitments states pledge are not legally binding and the compliance committee does not have authority to induce material sanctions on states that renege on their commitments (UNFCCC 2015).

Under a bottom-up treaty with weakly sanctioned compliance review, how credible are states' commitments? Although the Paris Agreement's pledge-and-review system is in several respects a unique invention in global governance, its compliance provisions face similar credible commitment problems as many comparable regimes that are based on self-reported efforts (Hafner-Burton et al. 2017; Creamer and Simmons 2019; Karlas 2021; Raiser et al. 2022).[1] Ideally, self-reporting regimes induce states to provide reliable and precise information about their commitments and implementation performance, which in turn can generate benign reciprocity effects for international cooperation (Mitchell 1998; Keohane and Oppenheimer 2016). Alternatively, however, self-reporting regimes can also incentivize cheap talk and false promise-making (Koremenos 2016), especially if policymakers value the benefits accrued from strong commitments in the short term over the potential adverse consequences of non-compliance in the long term (Hafner-Burton et al. 2017).

Congruous with the expectation that transparency generates regime effectiveness (Mitchell 1998), the Paris Agreement aims to build 'mutual trust' and promote compliance with NDCs through a logic of mutual assessment (UNFCCC 2015, art.13). Yet, the efficacy of this review mechanism presumably depends on states' readiness to provide precise information about their cooperative performance (Aldy et al. 2016; Keohane and Oppenheimer 2016). One obstacle to mutual assessment under the Paris Agreement is

[1] Examples include the OECD's Development Assistance Committee Peer Reviews; UN Convention Against Corruption's Implementation Review Mechanism; World Trade Organization's Trade Policy Review Mechanism; the Universal Periodic Review of the Human Rights Council; and the International Labour Organization's Complaint Procedure (Raiser et al. 2022).

commitment ambiguity: While some climate pledges contain precise information about mitigation targets and their implementation trajectories, many lack essential technical clarifications that leave their mitigation assumptions open to interpretation (Rogelj et al. 2017; Pauw et al. 2018; Rowan 2019). The widespread ambiguity in NDCs hampers review of individual commitments, which can undermine cooperation because states are unsure whether their peers are undertaking comparable efforts (Keohane and Victor 2011).

This article introduces a novel theoretical conjecture on how ambiguity in self-determined targets is related to the ambitiousness of commitments. Commitment ambiguity increases the flexibility of climate pledges, much like escape clauses in international trade agreements (Rosendorff and Milner 2001). However, whereas traditional top-down treaties require mutual coordination of policy obligations between states, the lenient rules of the Paris regime leave governments in complete control over their individual commitments. In this context, where the depth of cooperation is unilaterally decided, we theorize the ambiguity-ambition nexus as a credible commitment problem. Given that unambitious targets are easier to comply with (Downs et al. 1996; Barrett 1999; Johns 2014), we argue that the ambiguity-ambition nexus matters for understanding the potential effectiveness of self-reporting regimes, including their compliance prospects. From a compliance perspective, one principal question is whether ambiguous pledges are equally credible signals as precise pledges. We explain why ambiguity both can induce overly ambitious pledging ('imprudence') and overly careful pledging ('prudence'), before empirically testing whether ambiguous commitments differ systematically in ambition from precise commitments.

We formulate a theoretical model that delineates the relationship between commitment ambiguity and ambition, and empirically probe its plausibility by evaluating how mitigation ambiguity and ambition are related in states' climate pledges under the Paris Agreement. Our main argument is that exogenous ambiguity—uncertainty originating in governments' inability to perfectly gauge their commitment potential—induces prudent pledging. Prudence is motivated by compliance concern: under imperfect information about commitment potential, the anticipated international and domestic audience costs that arise in case of non-compliance deter states from pledging ambitiously (Fearon 1994; Koremenos 2005; Guzman 2008; Hafner-Burton et al. 2017). Our empirical analysis of all Paris Agreement NDCs points to an overall negative correlation between ambiguity and ambition. We proceed to show that this prudence effect of ambiguity is more pronounced for democracies, which we attribute to their stronger accountability mechanisms in case of non-compliance with pledges. Finally, a conjoint experiment fielded in

five democracies shows that compliance likelihood exerts a stronger causal effect on the general public's support for climate agreements than ambition—which helps explain why governments in democracies are rational to be prudent in the face of domestic audience costs.

Broadly, this article contributes a model of how ambiguity relates to ambition and compliance under self-reporting review systems in international institutions. Whereas the ambiguity-compliance relationship has received extensive attention in previous literature (e.g., Chayes and Chayes 1993; Mitchell 1998; Koremenos 2016; Keohane and Oppenheimer 2016), our model shows how the source of ambiguity (exogenous versus endogenous), the ambition level of commitments, and degree of compliance concern all matter in distinct ways for how ambiguity relates to compliance. By analyzing a bottom-up regime where the depth of cooperation is self-determined, our model complements existing work on institutional flexibility in top-down agreements (Rosendorff and Milner 2001; Kucik and Reinhardt 2008; Koremenos 2016; Hafner-Burton et al. 2017). Finally, our findings add to the empirical literature on international climate cooperation (Keohane and Victor 2011; Aldy et al. 2016; Keohane and Oppenheimer 2017; Rowan 2019; Victor et al. 2022) by providing evidence on the ambiguity-ambition nexus in states' pledges under the Paris Agreement, with implications for the effectiveness prospects of the treaty (Dimitrov et al. 2019; Tørstad 2020).

2 Ambiguity and Prudence under Pledge-and-Review

2.1 Credibility of Ambiguous Pledges

Under the Paris Agreement, Parties self-determine their mitigation targets by submitting NDCs. Article 4.2 of the Agreement requires that 'Each Party shall prepare, communicate and maintain successive nationally determined contributions that it intends to achieve' (UNFCCC 2015). However, the Agreement sets few requirements for the type of information that NDCs should contain, which has thus far led to substantial variation in the precision of NDC targets (Keohane and Oppenheimer 2016; Pauw et al. 2016). For example, the initial NDCs are based on different types of mitigation targets: 32 NDCs contain absolute emission targets, 78 contain baseline targets relative to business-as-usual, 9 have intensity targets, and 35 only outline 'policies and actions' (Pauw et al. 2016). Of these four target types, only absolute targets have clear global warming impacts absent the reliance on significant socio-economic assumptions and projections (Rogelj et al. 2017). Moreover, the NDCs cover different sets of greenhouse gases; in-

clude varying numbers of mitigation sectors in the targets; provide varying precision in mitigation cost estimations (if any); and specify different conditions such as finance or technology transfers for mitigation targets to be met.

Overall, the Paris Agreement's lenient requirements for the content of NDCs have engendered substantial *commitment ambiguity*—that is, limited available knowledge about the probability distribution of mitigation outcomes—in states' climate pledges. Although the flexibility allowed in the precision of NDCs offered an easy opt-in that initially contributed to the Paris Agreement's broad participation (Falkner 2016), the resulting widespread ambiguity in mitigation targets also renders the pledges incomplete and potentially unverifiable contracts of emissions reductions. Ambiguity induces doubt about the way in which pledges are to be executed (Koremenos 2016). Over the longer run, this information deficiency in states' pledges can undermine reciprocal collective action and effective cooperation. A key function of international institutions is to provide information about states' cooperative intent and performance (Keohane 1984); and both enforcement theorists and managerialists in the international negotiations literature agree that the provision of precise information can facilitate reciprocal commitments and spur increased compliance (Chayes and Chayes 1993; Tallberg 2002; Dai 2005; Aldy 2014). Crucially, tit-for-tat-like strategies can only generate cooperative equilibria in repeated prisoner's dilemma games if players can perfectly observe each others' behavior, allowing reciprocity (Axelrod 1984; Simmons 1998). Keohane and Oppenheimer (2016) thus propose that pledge-and-review under the Paris Agreement will 'only work if there is transparency' and Aldy et al. (2016) argue that transparency enhances the credibility of targets and the likelihood that Parties will comply with their NDCs.

However, positing a straightforward relationship between information provision and compliance, existing work on the ambiguity-compliance nexus has not formally accounted for states' varying levels of commitment ambition—which is a central feature of bottom-up regimes like the Paris Agreement. Whether information precision in the pledge phase of a pledge-and-review system can generate enhanced compliance rates presumably depends on the ambition level of mitigation commitments. The ambitiousness of mitigation commitments is here understood as deviations from what states would do in the absence of the Paris Agreement (i.e., 'business-as-usual' emissions scenarios), corresponding also to what others call the 'depth' of an international agreement (Downs et al. 1996; Barrett 1999; Johns 2014). Following Tørstad (2020) and Victor et al. (2022), we operationalize mitigation ambition as the implied temperature rise of NDCs[2]; and compliance as adequate implementation of the NDCs' mitigation components.

[2]See empirical strategy section for a detailed explanation of how this is calculated.

A trade-off exists between ambition (or 'depth') and compliance: Since unambitious commitments are easier to comply with, lower ambition should generate higher compliance rates (Downs et al. 1996; Barrett 1999; Johns 2014; Dimitrov et al. 2019; Victor et al. 2022).[3] Given that unambitious pledges are easier to comply with, we propose that the relationship between ambiguity and compliance depends on ambition. Based on the compliance-ambition trade-off, we assess whether states that have ambiguous mitigation targets in their NDCs have taken on systematically different levels of ambition than states with precise pledges. Understanding the relationship between commitment ambiguity and ambition can shed light on the credibility of states' climate pledges. If states with high ambiguity in pledged targets are more ambitious than states with low commitment ambiguity, ceteris paribus, we can infer that the targets of ambiguous pledges are inflated—and compliance will hence likely be lower than for pledges with precise targets. Conversely, if ambiguous pledges are less ambitious than pledges with precise targets, the ambiguous pledges are conservative—and compliance more easily achievable.

2.2 Balancing Ambition and Compliance

When formulating mitigation targets in a climate pledge, policy-makers have partly conflicting incentives for deciding on the ambition level of commitments. On one hand, there are strategic reasons to pledge ambitious targets. Following a logic of reciprocity, states can signal ambitious mitigation targets with the aim to spur other states to do the same (Tingley and Tomz 2014; Weikmans et al. 2019). Hence, signalling ambition can lead states to obtain specific mitigation benefits from others. The potential benefits of high ambition can also include more diffuse objectives such as enhanced international reputation or willingness by other states to cooperate in other institutions (Keohane and Oppenheimer 2016). Finally, pledging ambitious targets can also be a mechanism to please or attract domestic constituencies such as environmental interest groups or voters more broadly (Keohane and Oppenheimer 2016).

On the other hand, pledging unambitious targets can help states easily achieve compliance and hence avoid political repercussions internationally or domestically (Dai 2005; Keohane and Oppenheimer 2016; Hafner-Burton et al. 2017). The Paris Agreement does not set any minimum requirements for the ambition level of pledges, and policy-makers have to consider that their country's implementation performance will be extensively reviewed both by civil society, voters, political opponents, and other states. Illustratively, Hafner-Burton et al. (2017) show through experiments that elite decision-makers in the

[3] Victor et al. (2022) show empirically that this trade-off manifests empirically in the Paris Agreement NDCs.

US are reluctant to make false compliance promises even in the absence of formal enforcement mechanisms. This finding suggests that states may formulate relatively more prudent mitigation targets in the face of a compliance review mechanism such as the transparency framework under the Paris Agreement. From a compliance perspective, pledging unambitious NDC targets is a particularly appealing strategy for states that have low capacities to formulate precise and detailed targets, in order to maximize the likelihood that the pledged targets will be achieved.

Overall, these conflicting sets of incentives lead to a more general question of how states balance the potential benefits of ambitious commitments with the potential adverse consequences of non-compliance under self-reporting regimes (Koremenos 2016; Hafner-Burton et al. 2017). From a 'cheap talk' perspective, states will be unconcerned by the prospect of making imprudently ambitious promises; but from a 'costly signal' perspective, states will rarely make commitments they do not intend to implement (Hafner-Burton et al. 2017). Whereas existing literature on the Paris Agreement has argued that the upside of ambitious pledging prompts states to take on imprudently high mitigation targets (Victor et al. 2017; Brown et al. 2019; Rowan 2019), we propose that the inability of states to perfectly gauge their mitigation potential exerts the opposite effect—prudence—on the ambition level of pledges under the Paris Agreement. The reason is that, while ambitious and unambitious pledging each offers appealing benefits, the two strategies differ in the likelihood that the different sets of associated benefits will materialize. Notably, the posited international and domestic benefits of pledging ambitiously depend on other actors perceiving that a pledge is ambitious. However, no ambition level is likely sufficiently high to *guarantee* widespread acclaim. Further, although high ambition may spur reciprocal ambition among other states, this outcome is only likely to ensue if a critical number of states pledge ambitiously (Nyborg, 2018). Hence, the likelihood that the benefits of ambitious pledging will materialize is uncertain. In contrast, all states that submit an NDC are aware that their performance in implementing the pledge will be scrutinized by the Paris Agreement's review mechanism, other states, environmental NGOs, interest groups, and potentially domestic courts.[4] The certainty of review (writ large) constitutes a tangible shadow of the future—comprising both international and domestic audience costs (Fearon 1994)—that gives states reason to pledge prudently if any doubt prevails on their ability to comply. Consequently, the key driving force of prudence under our framework is *compliance concern.*[5]

[4]The potential for judicial review of climate targets was recently illustrated in the *State of the Netherlands v. Urgenda Foundation* case of 2019, where the Dutch Supreme Court ruled that the government must meet an emissions goal of 25% reductions from 1990 levels by 2020.

[5]Our concept of compliance concern is motivated by (but not equal to) what Hafner-Burton et al.

In a recent survey, government representatives were asked about the major obstacles to enhance the ambition of their countries' NDCs (UNFCCC 2020). The two dominant obstacles to ambition were perceived to be anticipated financial resources for implementation and government capacity to implement. This uncertainty about implementation prospects illustrates our theoretical notion of compliance concern. The argument that uncertainty in compliance prospects induces governments to take on prudent commitments has been established by the rational design literature in the context of international trade agreements (Milner and Rosendorff 2001; Kucik and Reinhardt 2008). Milner and Rosendorff (2001) and Kucik and Reinhardt (2008) demonstrate that the greater the uncertainty that political leaders face about their ability to maintain compliance with international agreements in the future, the more likely agreements are to contain flexibility provisions such as escape clauses. Similarly, we posit that states facing fundamental uncertainty about compliance prospects seek a safety valve for the degree of ambitiousness in their mitigation obligations under climate cooperation.

One likely source of compliance concern is domestic audience costs, which vary across political systems (Fearon 1994; Hafner-Burton et al. 2017; Creamer and Simmons 2019). While the international repercussions of non-compliance under the Paris Agreement are limited to naming and shaming, the domestic audience costs of non-compliance are likely substantially higher in democracies than autocracies. There is ample evidence that the stronger accountability mechanisms in democracies—including elections, independent courts, free media, and NGOs—render democracies more conducive to comply with their international obligations (Fearon 1994; Dai 2005; Simmons 2009; Creamer and Simmons 2019; Koliev et al. 2021). Simmons (2009), for example, shows how NGOs use domestic courts to hold governments accountable for the human rights practices to which they had agreed internationally. More recently, Koliev et al. (2021) demonstrate that the International Labor Organization's self-reporting review mechanism induces more compliance among democracies—which they attribute to the higher political and legal pressure that accrue domestically on democratic policy-makers. Finally, based on a survey of climate negotiators and scientists from across the world, Victor et al. (2022) find that pressure from civil society constitutes one of the foremost motivations for countries to comply with their NDCs. Owing to systemic differences in domestic audience costs, we hence expect that democracies will be more compliance concerned than autocracies.

(2017) call 'patience'.

2.3 Exogenous and Endogenous Ambiguity

The prudence motive in ambition emerges when states have imperfect information about their mitigation potential. In the following, we use the term 'ambiguity' to describe limited knowledge about the probability distribution of mitigation for a given ambition level. To capture the difference in ambiguity resulting from a *lack of information* necessary to formulate climate policy, on one hand, and *strategically induced* ambiguity, on the other, we distinguish between what we call *exogenous* and *endogenous* ambiguity in climate pledges. While exogenous ambiguity is uncertainty in mitigation targets that originates in policy-makers' inability to perfectly gauge their country's commitment potential (e.g. due to lack of technical capacity), endogenous ambiguity is uncertainty that is strategically determined by policy-makers. The distinction between exogenous and endogenous ambiguity helps delineate the causal relationship between ambiguity and ambition—that is, how policy-makers determine the levels of ambition and total ambiguity in the formulation of pledges.

This article outlines a sequential formulation process of climate pledges, wherein the determination of ambition and ambiguity depends upon each other. Specifically, we posit a (stylized) two-stage process wherein, first, (1) exogenous ambiguity affects ambition, followed by (2) ambition affecting endogenous ambiguity. In the first stage, the (exogenous) availability of mitigation-related information constrains policy-makers in the determination of ambition. In the second stage, policy-makers can choose to strategically add ambiguity to their determined mitigation targets based on their preferences towards the overall level of ambiguity of their pledges. While the two ambiguity sources are hard to disentangle empirically, our decision-theoretic model produces a framework for interpreting the relationship between total ambiguity (exogenous and endogenous) and ambition.

Our choice of analyzing the relationship between ambiguity and ambition through a simple choice model—as opposed to a dynamic model of cooperation—owes to the bottom-up structure of the Paris Agreement. In contrast to previous top-down climate agreements (e.g. the Kyoto Protocol), the Paris Agreement's pledge-and-review system does not involve mutual coordination of mitigation policies at the pledge stage. While in the following, we focus on countries' individual decisions in determining ambition when formulating pledges, Supplementary Material H provides a game-theoretical intuition of how exogenous ambiguity can undermine cooperation in a coordination game.

To illustrate the two-stage process of our model, consider the following simplified example wherein a policy-maker from Country A is formulating a climate pledge. In order

to decide the ambition level of the pledge—that is, how much greenhouse gases Country A pledges to cut within a given timeframe—the policy-maker needs to evaluate Country A's mitigation potential. To assess the country's mitigation potential, the policy-maker relies on information about a range of country-specific characteristics—such as Country A's current and past emissions, its projected emissions under different socio-economic scenarios, the share of fossil fuels in the country's energy mix, the costs of mitigation, the energy efficiency of industrial sectors, and so on. The extent to which the policy-maker can access precise information about these and any other relevant characteristics determines the exogenous ambiguity of a pledge. Previous literature has shown that states' capacity to specify precise mitigation targets varies considerably (Khan et al. 2019), and exogenous ambiguity could originate in factors such as a country's scientific and technical capacity, bureaucratic capability, and fiscal resources (Chayes and Chayes 1993; Röser et al. 2020; Karlas 2021). In our model, this ambiguity is exogenously given because the policy-maker cannot affect the precision of this information made available to her. In the first stage of our model, the policy-maker uses the information available about Country A's mitigation potential to determine the ambition level of Country A's pledge.

The second stage of our model is when a policy-maker endogenously adjusts ambiguity in light of the ambition level of a pledge. This stage describes a mechanism whereby—after setting the pledge's ambition level—the policy-maker strategically adds ambiguity to the ambition level of a pledge, in order to obtain beneficial policy-objectives such as enhanced implementation leeway or to signal high ambition. If, for example, the policy-maker decides that a 10% cut is a realistic mitigation potential, the policy-maker could add ambiguity to that target by pledging an interval (e.g. "10-20%") instead of a fixed target, by introducing conditional statements (e.g. "we will cut 20% if we receive financial support"), or by otherwise intentionally adding uncertainty to the mitigation target.

The distinction between exogenous and endogenous ambiguity matters because the two types of ambiguity have different consequences in the formulation of climate pledges. We argue that exogenous ambiguity leads to prudent ambition in pledges, while endogenous ambiguity blurs the relationship between ambition and implementation capacity. In what follows, we use standard economic risk theory to model states' formulation of pledges. [6]

[6] We rely on economic risk theory as states are likely to have at least *some* knowledge about the likelihood of their own mitigation scenarios.

2.4 Formal Model

In our model, states determine the optimal combination of their pledges and endogenous ambiguity in two stages. Both stages are conditioned by the extent of exogenous ambiguity a state faces. In stage one, states pledge a mitigation commitment by balancing ambitiousness and compliance probability. Given the optimal pledge, states seek to maximize the function G(Total ambiguity; Ambition) in stage two, to determine the optimal amount of endogenous ambiguity which affects total ambiguity communicated in the NDC. We model a state's incentive to signal the potential for high ambition as a concave function of total ambiguity. This choice reflects that too much ambiguity—such as introducing endless numbers of conditional statements—may decrease the credibility of pledges. Hence, states weigh compliance prospects and signalling potential outcomes of their mitigation efforts in two stages. Beginning with the first stage, the utility of a state over exogenous ambiguity and pledged ambition is:

$$U = -\Omega(p)u(q(X) - p) \tag{1}$$

where p denotes a states' pledged ambition and $q(X) - p$ is the discrepancy between the latent mitigation variable, $q(X)$, reported in the review stage, and the pledge, p. For notational simplicity we simply write q throughout. q is the estimated mitigation conditional on country characteristics, X, including factors such as mitigation capacity, vulnerability to climate change, and fossil fuels endowments (Tørstad et al. 2020). Ω is a scalar that potentially depends on p. This parameter reflects that states may value discrepancies between q and p differently depending on the size of p. Pledging ambitiously could be attractive because states hope to incentivize other states to invest in mitigation. However, states have to weigh the benefits of ambitious pledging against the feasibility of compliance. The utility function u in (1) has a symmetric U-shape, say quadratic, where the unique minimum (and maximum of (1)) reflects the optimal pledge, p^*. This bliss point may be bigger or smaller than q depending on the benefits states attach to ambitious pledges relative to compliance.

Since we focus on the determination of pledges, we treat mitigation, q, as an exogenous, random variable which—conditional on different country characteristics—has the following distribution: $q \sim D(\mu_q, \sigma_q^2)$, where the probability density function is symmetric about μ_q and independent of other country characteristics than X. A state with 'full' control over its own mitigation, $\sigma_q^2 \to 0$, will pledge to mitigate approximately p^*, the optimal pledge when there is no uncertainty about q, and obtain utility $U = -(p^*)u(q - p^*)$ where $q \approx \mu_q$. As σ_q^2 increases, so does the chances of severe com-

pliance and non-compliance. Hence, the state is compelled to balance these concerns.

In the following we define, for notational simplicity, the variable $C := q - p$ as the discrepancy between q and p. C is then distributed with expectation $\mu_C = \mu_q - p$ and variance σ_q^2, and $C^* = \mu_q - p^*$ is the optimal realization of C. Finally, f(C) is the probability density function of C. To summarize, a given state cannot affect the probability of achieving C^*, but it can determine its pledge to increase the probability of compliance, $C > C^*$, by reducing its pledge and thus increasing μ_C; or accept a higher probability of non-compliance by setting p such that $\mu_C < C^*$.

We now follow Waud (1976) to analyze how states determine C relative to C^* when facing uncertainty about the realization of q. Suppose for simplicity, and without loss of generality, that $C^* = 0$, meaning that the optimal pledge equals the expected mitigation. Formally, we assume the following:

$$\begin{aligned} u(C) \qquad & \text{if} \quad C > 0 \\ \Omega u(C) \qquad & \text{if} \quad C < 0 \end{aligned} \tag{2}$$

$\Omega > 1$ reflects severe compliance concern for instance because the benefits of compliance are more tangible than the benefits of pledging more ambitiously. Hence, for a discrepancy of b, a positive number, a state would experience a greater loss if $\bar{C}$, the realized C, was $\bar{C} = -b$ than if $\bar{C} = b$. If, however, $\Omega < 1$ these relative losses are reversed such that overshooting is considered less attractive then undershooting. The state then maximizes (1) by choosing the μ_C that minimizes the expected loss:

$$\underset{\mu_C}{\text{minimize}} \quad -E(U) = \Omega \int_{-\infty}^{0} u(C) f(C; \mu_C) dC + \int_{0}^{\infty} u(C) f(C; \mu_C) dC \tag{3}$$

Result 1: *If $\Omega > 1$ and $\sigma_q^2 > 0$ the state will choose $\mu_C > C^*$. If $\Omega < 1$ and $\sigma_q^2 > 0$ the state will choose $\mu_C < C^*$.*

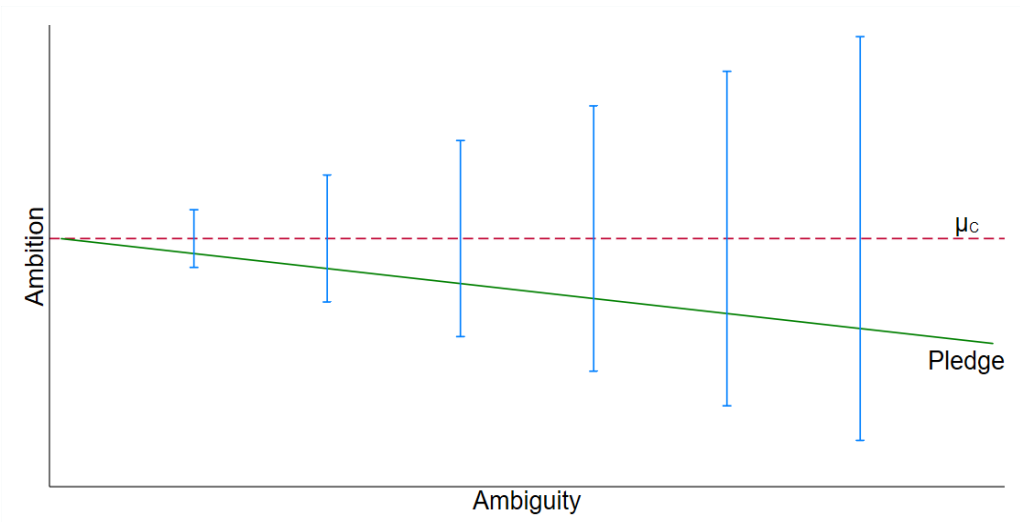
Proof. See proof of Proposition 2 in Waud (1976). We provide a proof with our notation in Supplementary Material F ■

It immediately follows from Result 1 that if $\Omega > 1$ states' pledges will be lower than p^*. Moreover, as the variability of q increases, $\sigma_q^2 \uparrow$, states decrease the ambition level of their pledges (Waud, 1976). The intuition is that as the σ_q^2 increases, a mean preserving spread, the loss associated with undershooting relative to overshooting is given more weight. The increased probability of low realizations of q disincentivizes states to pledge

ambitiously. We call this behavior prudence. While states' actual mitigation may be higher or lower than pledged, they focus on the possible down-side risk—which is more tangible than the upside risk due to the anticipation that compliance will be reviewed. Prudence implies more conservative pledging relative to the case where mitigation levels are more certain, as illustrated in Figure 1. Contrastingly, imprudence manifests when states' pledges, p, increase in σ_q^2, meaning that they weigh losses associated with overshooting relatively higher. States that are unconcerned about the prospect of compliance review could seek to reap cooperative benefits of appearing ambitious by pledging imprudently.

The extent to which the prudence motive will manifest for a given state depends on how concerned the state is about compliance (Hafner-Burton et al. 2017), which is governed by the size of Ω in our model. The negative effect of exogenous ambiguity on ambition (Result 1) is thus magnified by increased compliance concern since the disutility of not reaching a given mitigation target increases (Waud, 1976). In contrast, a reduction in Ω would be associated with less prudent behavior and ultimately imprudence if $\Omega < 1$. Thus, increasing Ω will continuously reduce p for a given σ^2.

Figure 1: The prudence effect of exogenous ambiguity



Notes: This figure illustrates the predicted relationship between exogenous ambiguity and ambition as measured in states' climate pledges, conditional on $C^* = 0$ and $\Omega > 1$. The blue, vertical line segments represent exogenous ambiguity. The dashed line shows expected mitigation level conditional on country characteristics. The green line represents ambition as a function of ambiguity.

Until this point, our model has presumed that ambiguity in NDCs is exogenous (represented by σ_q^2), meaning that ambiguity originates in factors beyond the control

of policy-makers. Such exogenous factors could, for example, be scientific and technical capacity, bureaucratic capability, and fiscal resources (Chayes and Chayes 1993). Previous literature on NDC formulation processes indicates that this type of ambiguity is prevalent, as the preparation of NDCs '(...)requires substantive amounts of data, knowledge and capacity to assess the potential outcomes of various policy options' (Röser et al. 2020). Some states may, however, have incentives to also intentionally introduce ambiguity in their targets (Keohane and Oppenheimer 2016; Rowan 2019). We call this endogenous ambiguity and henceforth denote it A_{En}.

Endogenous ambiguity has been shown to feature in a wide range of domestic and international institutions. For example, endogenous ambiguity can be winning strategy for candidates and political parties trying to attract voters (Tomz and Van Houweling 2009; Bräuninger and Giger 2018); for a small state engaging in an arms race with a big power (Baliga and Sjöström 2008); for leaders and staff of international organizations (Best 2012); and for judges that seek to pre-empt defiance of judicial rulings (Staton and Vanberg 2008). Under self-reporting review systems, there are at least two incentives for endogenous ambiguity. First, a given state may introduce ambiguity in order to obtain a degree of flexibility in the review process, effectively obfuscating whether the state is in compliance with its targets or not (Simmons 2010; Keohane and Oppenheimer 2016). Second, a state could introduce more ambiguity in order to signal a higher commitment level than it actually intends to pursue, in order to obtain reciprocity benefits from other states (Keohane and Oppenheimer 2016). In either of these scenarios, the ambition-compliance nexus will be blurred.

In our model, we depict the formulation of pledges as a two-stage process where states have so far decided their optimal pledges (under exogenous information constraints) and next adjust the ambiguity regarding q to signal prospects for ambitious or unambitious commitments in the second stage. Suppose that the endogenous ambiguity of a given state's mitigation, q, is decided by adjusting A^{+}_{En} and A^{-}_{En}. By setting $A^{+}_{En} > 0$, the state gives the impression that higher values of q are achievable than implied by the exogenous ambiguity. Hence, if states' primary concern is to signal their potential to implement large emission cuts, we would expect them to increase A^{+}_{En}. For example, states could introduce conditional statements in their targets that imply higher levels of mitigation if certain conditions are fulfilled, e.g. financial support from other states. Conversely, $A^{-}_{En} > 0$ would portray lower mitigation levels as more probable. In our model, states do not have an incentive to do this as a means of achieving greater leeway. The possibility of low realizations of q is captured by the exogenous ambiguity and if states want to hedge against the possibility of not reaching their target, this concern

would be captured by their choice of p in expression (1). One could, however, imagine that states wanted to set $A^{-}_{En} > 0$ and exert little effort to reach q, but this would affect the choice of μ_q in the first place since this is considered the optimal mitigation level given country characteristics. Unambitious states would rather adhere to their optimal level of mitigation, set their pledges optimally in the first stage, and increase A^{+}_{En} to reflect potentially large emission reductions. Overall, the introduction of endogenous ambiguity does not alter the directional effect outlined in Result 1 since the level of p relative to μ_q is unresponsive.

We now return to the function G, representing the second-stage maximization problem, and model a state's incentive to signal the potential for high mitigation levels as a concave function of total ambiguity. States care about the level of total ambiguity in their NDCs, as this is what is observed by others. Hence, the A^{+}_{En} is a function of σ_q^2. The degree of endogenous ambiguity also depends on p since prudent states reduce their ambitions as σ_q^2 increases, thus expanding the room for realization of q above p and ultimately rendering A^{+}_{En} less useful. We end up with the following problem:

$$\underset{A^{+}_{En}}{\text{maximize}} \quad G(\mu_q + A^{+}_{En} + \sigma_q - \beta p(q, \mu_q, \sigma_q^2)) \tag{4}$$

where β represents a state's perception of the optimal degree of A^{+}_{En}, which we assume is independent of σ_q^2. Combined, the three leftmost terms within G form a measure of the spread of total ambiguity above the expected mitigation level. The concavity of G reflects that too much ambiguity—such as introducing endless numbers of conditional statements—may decrease the credibility of targets. The first order condition of (4), $G'(\mu_q + A_{Ex}(p, \mu_q, \sigma_q) + \sigma_q - \beta p(q, \mu_q, \sigma_q^2)) = 0$, pins down the optimal level of endogenous ambiguity as a function of the exogenous ambiguity. Differentiating with respect to σ_q generates the following result:

$$\frac{\partial A_{En}}{\partial \sigma_q} = \beta \underbrace{\frac{\partial p}{\partial \sigma_q}}_{<0} - 1 \tag{5}$$

Result 2: *Endogenous ambiguity is negatively related to exogenous ambiguity if* $\Omega > 1$ *and if* $\Omega < 1$ *&* $\beta \frac{\partial p}{\partial \sigma_q} < 1$.

Hence, there is only a positive correlation between endogenous and exogenous ambiguity if states are sufficiently imprudent. Furthermore $\frac{\partial A_{En}}{\partial \sigma_q \partial \Omega} = \beta \frac{\partial p}{\partial \sigma_q \partial \Omega}$, which is negative. The relationship in Result 2 is therefore magnified by Ω, meaning that, for a

given β, more compliance-concerned states reduce A_{En} at a faster rate than states that are less so, if exogenous ambiguity increases. The reason is that states that behave more prudently have a lesser need to top up exogenous with endogenous ambiguity.

2.5 Linking Theory to Data

We cannot observe each state's optimal pledge in the case of precision, $\sigma_q^2 \to 0$, or μ_q. Hence, our identification strategy of risk behavior relies on cross-country observations of how states act on NDC mitigation ambiguity in the mapping from $p \to q$ and control variables that are likely to affect ambition and ambiguity. While the inclusion of endogenous ambiguity does not alter a state's pledge, p, relative to expected mitigation μ_q, it may affect our interpretation of the relationship between ambiguity and ambition as measured in the NDCs because it is hard to disentangle exogenous and endogenous ambiguity empirically. As in Figure 2, we will in our regressions array states' pledges based on total ambiguity—i.e. the sum of exogenous and endogenous ambiguity. Now suppose we were to estimate the following OLS regression to find the overall, linear relationship between *ambition* and *pledge* in Figure 2:

$$p = \mu_q + \gamma(\sigma_q^2 + A_{En})$$

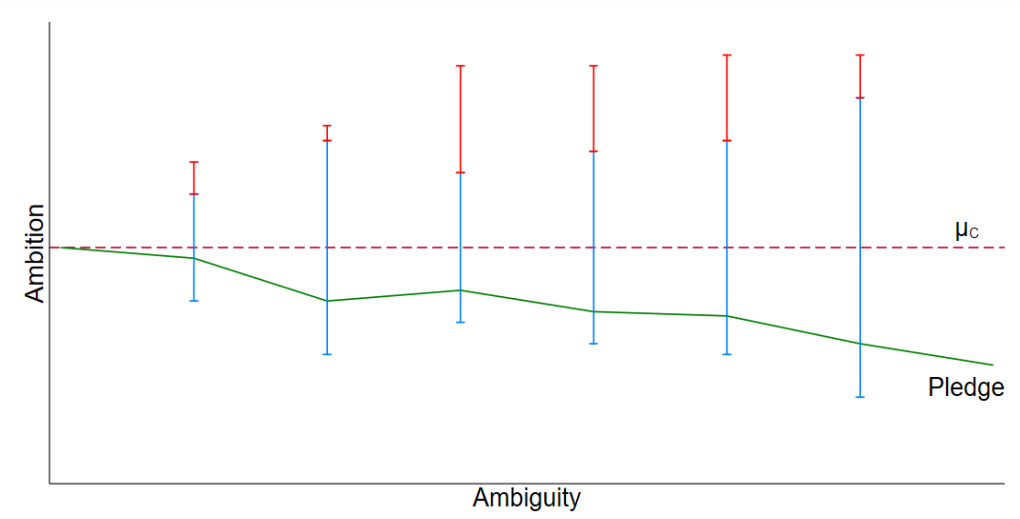
where μ_q is expected mitigation and the slope is denoted γ and its sign depends on the following:

$$\gamma = \frac{cov(\sigma_q^2 + A_{En}^+, p)}{var(\sigma_q^2 + A_{En})} = \frac{cov(\sigma_q^2, p)}{var(\sigma_q^2 + A_{En})} + \frac{cov(A_{En}, p)}{var(\sigma_q^2 + A_{En})} \tag{6}$$

Whether the causal exogenous or endogenous ambiguity determines the sign of γ depends on the signs and magnitudes of the two rightmost terms in (6). Suppose first that states are prudent. Based on our theoretical considerations—and given that these relationships have linear tendencies—these have opposite signs. From comparative statics on Result 1, we have that $cov(\sigma_q^2, p) < 0$, and it is immediately clear from 4 that $cov(A_{En}^+, p) > 0$ since increases p need to be offset by increased A_{En}^+. This is independent of prudence. A negative γ would thus indicate the existence of prudence. If states were imprudent, both $cov(\sigma_q^2, p)$ and $cov(A_{En}^+, p)$ would be positive, which means that $\gamma > 0$ implies one of two scenarios: 1) that states are prudent but $cov(A_{En}^+, p) > 0$ dominates, or 2) that states are imprudent. Thus, it is only $\gamma < 0$ that unambiguously identifies states' risk behavior. Since $cov(\sigma_q^2, p)$ would be increasing in magnitude by increased compliance concern (Ω), differences in β across states with different Ω values

would clarify the role of prudence.

Figure 2: Ambition as a function of endogenous and exogenous ambiguity



Note: This figure illustrates the predicted relationship between ambiguity and ambition as measured in states' pledges, conditional on $C^* = 0$. The blue, vertical line segments represent exogenous ambiguity, which is beyond states' control, while the orange line segments portray potential strategic ambiguity. The dashed line shows the expected mitigation level of states conditional on controls. As the combined level of ambiguity increases, the pledges decrease relative to the expected mitigation, which reflects prudent behavior by states.

In sum, the discussion above showed why states have an incentive to pledge prudently in the face of ambiguity. The prudence motive is driven by the anticipation of external review: If states are concerned about future compliance review, we propose that ambiguity and mitigation ambition are negatively related in states' climate pledges. Importantly, our model is indifferent to whether the review is in effect undertaken by international or domestic actors; the driving force of Result 1 is that states' concern for achieving compliance ($\Omega > 1$)—for whichever reasons—results in prudent behavior.

To probe this prediction empirically, we indirectly infer whether states are prudent, on average, by assessing the correlation between ambiguity and ambition in states' climate pledges under the Paris Agreement. A negative correlation between ambiguity and ambition in these pledges indicates that states pledge prudently in the face of ambiguity, and that ambiguous targets are deflated compared to precise targets. The opposite tendency—a positive correlation between ambiguity and ambition—would suggest imprudent pledging. In order to capture the role of Ω, we also test whether the relationship between ambition and ambiguity varies with form of governance. As explained

above, we expect that the prudence effect will manifest more strongly for democracies because policy-makers in democracies face higher domestic audience costs in case of non-compliance than their autocratic counterparts. This prediction constitutes a crucial test of the model's mechanics.

3 Data and empirical strategy

In the empirical analysis we conduct a cross-sectional statistical analysis of ambiguity and ambition in states' NDCs under the Paris Agreement, coupled with a conjoint experiment. Following our theoretical model, the two main variables in the statistical analysis are *NDC ambition* and *NDC ambiguity*.

NDC ambition: First, our dependent variable *NDC ambition* is based on Robiou du Pont & Meinshausen (2018), who apply a hybrid allocation approach to estimate the global temperature impact (measured in °C) consistent with each state's NDC. The metric, based on a sophisticated modelling approach that compares countries' NDCs to different global emissions scenarios, provides an assessment of global warming impact if all states adopted the ambition level of a given NDC. The data scores range from 1.2°C warming (most ambitious) to over 5.1°C (least ambitious). We invert the scale to facilitate interpretation, so that higher scores mean higher ambition, with a range from 0 to 3.9. Robiou du Pont & Meinshausen's (2018) assessment takes three effort-sharing principles into account—capability to pay (GDP per capita), historical responsibility (convergence to equal cumulative per capita emissions), and equality (convergence to equal per capita emissions)—and the global warming consistency of a given NDC is calculated based on the principle most lenient for the given state. The three effort-sharing principles are grounded in the IPCC's fifth assessment report. In contrast to other ambition assessments[7], Robiou du Pont & Meinshausen's (2018) variable is peer-reviewed, covers nearly all NDCs, minimizes the normative choices made, has an intuitive interpretation (global warming impact measured in °C), and avoids making counterfactual assumptions about business-as-usual emissions (Tørstad et al. 2020). For full information about this ambition metric, including country rankings and correlates of ambition, see Robiou du Pont & Meinshausen (2018)[8], Tørstad et al. (2020), and Victor et al. (2022). In the main text, we analyze the ambition of the *initial* set of NDCs (due in

[7]See Sælen et al. (2019) for an overview.

[8]An interactive map of the ambition data is also available at http://paris-equity-check.org/warming-check.html

2015), most of which came into force in 2020 and apply to 2030. We focus on the initial NDCs because data are far more widely available for these than for the updated NDCs (due in 2020-21). Moreover, only the ambition data for the initial NDCs have previously undergone peer review. Nonetheless, we also present analyses of the updated NDCs in Supplementary Material C. Finally, Supplementary Material C also tests the robustness of our results using two alternative ambition metrics as dependent variable (Burck et al. 2018; Lancesseur et al. 2021). The analyses in Supplementary Material C indicate that all our main results hold for two different alternative operationalizations of NDC ambition, as well as for the updated NDCs. However, statistical power is significantly reduced in these analyses owing to lower sample sizes.

NDC ambiguity: Second, to measure *NDC ambiguity*, we collect information on the precision of all NDCs from Pauw et al's (2016) *NDC explorer* database. We code the ambiguity of 20 different NDC mitigation target characteristics, e.g., which gases the NDCs cover; what types of mitigation targets the NDCs set; and whether targets are conditional on financial or technological support. Descriptions of all ambiguity variables and their coding are provided in Supplementary Material A. We measure two main types of ambiguity in these NDC target characteristics. *Impact precision* is the degree to which global warming consequences of mitigation targets can be derived with certainty from the NDCs. For instance, absolute mitigation targets (i.e. emission reductions relative to a specified base year) have clearer global warming implications than emission intensity targets (i.e. emission reductions relative to economic indicators such as GDP)—as the latter depend on the future socio-economic development trends of a given country (Rogelj et al. 2017). *Information completeness* refers to the breadth of policy sectors and tools included in the formulation of the NDC: For example, whether an NDC covers policy sectors such as transport or agriculture and whether it covers policy tools such as carbon capture and storage or renewable energy generation. Impact precision is closely related to the ambition level of NDC targets, in the sense that higher impact precision renders ambition more straightforward to evaluate. Information completeness, on the other hand, can be understood as the level of details in the implementation trajectory of an NDC—and is hence more related to the credibility that a country will achieve its stated target than the actual ambition of the target. We recode all mitigation-related variables from Pauw et al.'s (2016) database such that higher variable values indicate higher degrees of NDC ambiguity.

Since we are not interested in the impact of each ambiguity dimension but rather a unified measure of ambiguity, we construct three ambiguity indices based on three

different approaches. In our additive, unweighted index all dimensions are weighted equally regardless of correlational patterns. The two other indices are reflective, gearing to a potentially latent concept of ambiguity by taking into account the variation of the 20 ambiguity dimensions. The *Regression index* is an additive index in which weights are proportional to the size of the coefficients associated with each ambiguity dimension in explaining cross country variation in the variable 'type of target', which we consider to be a particularly valid proxy for the concept of NDC ambiguity (based on Rogelj et al. 2017). The *PCA index* uses principal component analysis to evaluate what variables capture the same latent concept among the 20 ambiguity variables we have selected from the NDCs. We only use the first component for simplicity. This component explains (22%) of the variation in the dataset; about twice as much as the second component. Supplementary Material D provides details about the construction of the indices.

Control variables: To reduce potential bias in our regression estimates, we also include a set of control variables that Tørstad et al. (2020) show correlate with our NDC ambition variable. The control variables we include are GDP per capita (World Bank 2016), fossil fuels rents (World Bank 2015a; 2015b; 2015c), vulnerability to climate change (ND-GAIN 2015), and form of governance (Coppedge et al 2017).[9] To measure form of governance—which is the only control variable we have theoretical predictions for—we use V-Dem's multiplicative polyarchy index (Coppedge et al. 2017). This index ranges between 0-1, with higher values for more democratic countries. We use 2015 data for all controls because that is the year the NDCs were formulated. To select only the relevant controls—that correlate with both ambition and ambiguity—we implement the double-lasso procedure outlined by Urminsky, Hansen, and Chernozhukov (2016) in several of our regressions. Double-lasso selection allows us to include both controls and their interactions to better fit the data while at the same time avoiding overfitting the models. In some of the regression tables in the main analysis we only indicate whether controls are included or not. The full models are reported in Supplementary material I.

Conjoint experiment: Following the main analysis of ambiguity and ambition among states, we implement a randomized conjoint experiment (Hainmueller et al. 2014) to test whether compliance concern (Ω in our formal model) manifests among the general public in five large democracies—Germany, Mexico, South Africa, the United Kingdom,

[9]Some of the control variables are missing for some states inhibiting the inclusion of all countries in the regression analyses we employ. States with missing values are slightly more ambiguous and ambitious on average. However, point estimates are close to zero and correlations are weak (Table E.2 in Supplementary material).

and the United States. The experiment implements simple ambition and compliance treatments to gauge the degree to which people are compliance concerned when selecting between different climate cooperation arrangements that their government can partake in. Since we argue that democracies have reason to be more concerned with achieving compliance with international commitments than autocracies, eliciting people's compliance preferences in five democracies provides a most-likely test of whether our theoretical compliance concern construct translates into real-world preferences of citizens, who impose domestic audience costs on governmental leaders. Additional methodological information and results are provided in Supplementary Material G.

4 Empirical analysis

The empirical analysis proceeds as follows. We first provide a descriptive overview of NDC ambiguity among states. We then assess the aggregate relationship between ambiguity and ambition. Subsequently, we evaluate how form of governance conditions the relationship between ambiguity and ambition. Finally, we report the main results of our conjoint experiment testing the causal effect of compliance likelihood among the general public.

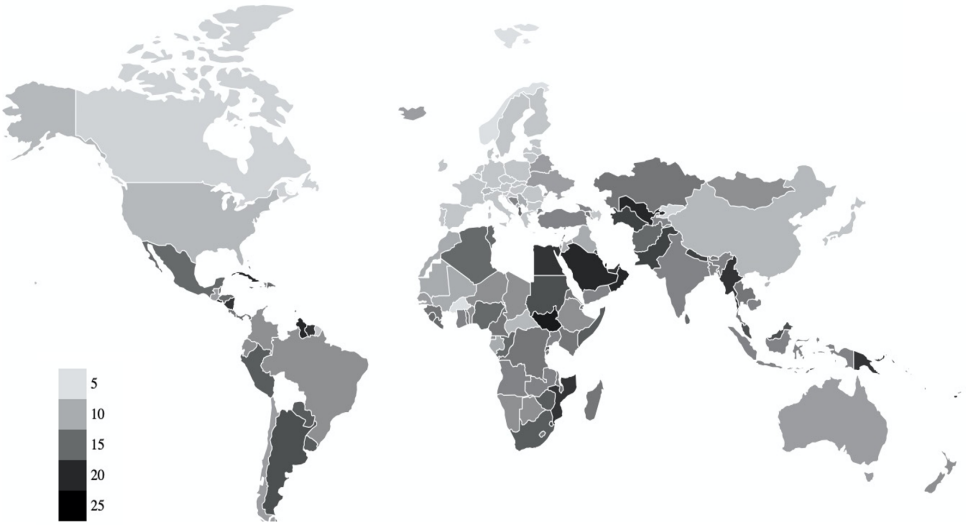
4.1 Descriptive statistics

Overall, we identify substantial cross-country variation in NDC ambiguity. Figure 3 displays country scores on a mitigation ambiguity index with equal weights for all 20 NDC characteristics that we code. Darker gray indicates higher ambiguity in NDCs.[10] Countries such as Syria, Bahrain, Bolivia, Qatar, South Sudan, Cuba, and Saudi Arabia have some of the most ambiguous mitigation targets in NDCs. Inversely, Norway, Canada, Japan, and the EU countries have highly precise NDCs. On a regional level, countries in Western Europe, East Asia, and North America have overall more precise NDCs than countries in the Middle East, South America, and Africa. The correlates of NDC ambiguity include form of governance, climate change vulnerability, and fossil fuels rents. [11]

[10] White color (e.g. Libya and the Philippines) indicates missing data.

[11] Democracies have more precise NDCs, while vulnerability and fossil fuels rents are associated with more ambiguous NDCs. See Table E.1 in the Supplementary Material.

Figure 3: Country scores on the unweighted mitigation ambiguity index



4.2 Ambiguity, Ambition, and Compliance Concern

We first evaluate the aggregate effects of ambiguity on NDC ambition, using three ambiguity indices that are weighted differently. Table 1 shows the relationships between the different indices and NDC ambition.

Table 1: The effects of three NDC ambiguity indices on ambition

	No weights		Regression weights		PCA	
Dep Var: NDC Ambition	b/se	b/se	b/se	b/se	b/se	b/se
Ambiguity Index	-0.0684**	-0.0448	-0.610*	-0.315	-0.173	-0.0408
	(0.0306)	(0.0313)	(0.338)	(0.335)	(0.147)	(0.161)
All controls included	Yes	Yes	Yes	Yes	Yes	Yes
Double Lasso	No	Yes	No	Yes	No	Yes
Observations	148	148	148	148	148	148
R^2	0.552		0.546		0.540	

Notes: This table displays the effects of three ambiguity indices on countries' NDC ambition using OLS regression. In columns 2, 4 and 6, control variables are selected using the double-lasso variable selection procedure. Control variables are described in detail in Table A.1, Panel B. Robust standard errors are in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels.

Columns 1-6 above respectively show regressions of a summative unweighted index,

the regression index, and a principal component index. All coefficients are negative—indicating prudence—yet, the effect sizes are small and only the unweighted index is statistically significant. This weak relationship might reflect that the models do not account for heterogenous effects of states' compliance concern (Ω in our model): ambiguity does not give states any particular reason to pledge prudently if they are not invested in achieving compliance.

The next stage in our analysis evaluates whether the relationship between ambiguity and ambition is conditioned by form of governance. Theoretically, we previously argued that exogenous ambiguity is likely to result in prudence *only if* states weigh the downside risk of non-compliance more heavily than the upside risk of ambitious pledging. We thus expect compliance concern—the sensitivity parameter Ω in our theoretical model—to condition the relationship between ambiguity and ambition. While compliance concern is fundamentally unobservable, we previously argued that form of governance is a reasonable proxy for the concept due to the higher domestic audience costs that democracies face in case of non-compliance with their pledges. Substantiating this argument, Supplementary Material B demonstrates that democracies are significantly more compliant with the procedural requirements of the Paris Agreement than autocracies.

Table 2: Conditional Effect of Democracy on the Ambiguity-Ambition Nexus

	No weights	Regression weights	PCA
Dep Var: NDC Ambition	b/se	b/se	b/se
Ambiguity index	0.0252	0.661	0.480***
	(0.0464)	(0.452)	(0.178)
Democracy	3.639***	0.605	1.033**
	(1.183)	(0.532)	(0.433)
Ambiguity index x	-0.209**	-2.953**	-1.509***
Democracy	(0.104)	(1.249)	(0.369)
Controls	Yes	Yes	Yes
Double-lasso	Yes	Yes	Yes
Observations	148	148	148

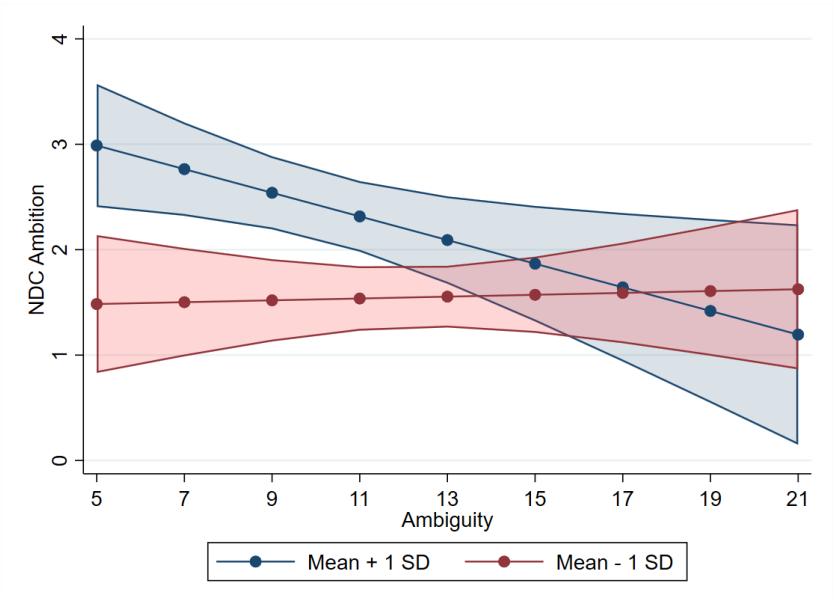
Notes: This table displays the effects of the three ambiguity indices on states' NDC ambition using OLS regression. *Democracy* is a continuous measure of states' level of democracy and a proxy for compliance concern. *Ambiguity index x Democracy* is its interaction with the ambiguity indices. Control variables are selected using the double-lasso variable selection procedure. Robust standard errors are in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels.

The regression models in Table 2 display the direct effects of ambiguity and form of governance on ambition, and the interaction term between ambiguity and form of governance. These models corroborate the importance of democracy in three ways. First, the direct effect of ambiguity changes from negative to positive when we include the interaction terms and hence control for the heterogenous effects of form of governance. Second, the models also show that ambition and democracy are positively related when holding ambiguity constant. Third, and most significant for our purposes, the interaction term between ambiguity and democracy is negative: the prudence effect of ambiguity is more pronounced for democratic countries. This finding is illustrated in Figure 4, which plots the estimated model in column 1 of Table 2. The estimated slope is positive for countries that score low on the democracy index but negative for countries that are highly democratic.[12] In other words, ambiguity induces prudence among the most democratic

[12]For states with relatively low score on the democracy index, such as Cc=0.079 (the blue line in Figure 4), the direct, positive effect of ambiguity slightly dominates the negative interaction term. If we divide

countries, while the least democratic countries are slightly imprudent. Overall, the negative interaction effect indicates that higher levels of democracy correlates with more prudence.

Figure 4: Interaction effect between form of governance and NDC ambiguity



Notes: This figure plots the estimated model (No weights) in column 1, Table 2, with ambiguity on the X axis and ambition on the Y axis. The two curves illustrate this predicted relationship for highly democratic countries (blue line) and undemocratic countries (red line), operationalized as one standard deviation above/below average cross-country scores on the V-Dem multiplicative polyarchy index.

4.3 Ambition and Compliance Concern among the General Public

As a final stage in our empirical analysis, we test whether compliance concern manifests among the general public. Since compliance concern forms a crucial component in our theoretical justification for the prudence effect of ambiguity, the overall validity of our model depends on this concept actually holding explanatory power in real-world applications. The reason we now shift the unit of analysis from states to the general public is to ensure a controlled setting suitable to establish the causal effect of compliance concern on the attractiveness of climate agreements. Moreover, domestic audience costs—which can be induced by regular citizens on their governments—potentially play a central role

the absolute value of the coefficient on *Ambiguity index* by its interaction, using the No weight-model in Table 2, the resulting ratio of 0.121 reflects the democracy level at which states would be behave neither prudently nor imprudently. Our measure suggests that 147 countries have a higher compliance concern than this value.

in the determination of states' compliance concern.

To test whether citizens take the probability of compliance into account when evaluating whether to support a climate agreement, we implemented a conjoint experiment (Hainmueller et al. 2014) where participants were asked to choose between two hypothetical climate agreements. The agreements had three attributes that were randomized—participation, ambition, and implementation likelihood—with two levels each. The participation attribute varied whether one's own country participated in the agreement or not. The ambition attribute varied whether the agreement demanded 20% cuts in greenhouse gas emissions or 40%.[13]

Finally, the compliance attribute varied between 20% and 50% probability that the mitigation target would be implemented.[14] We recruited a balanced sample of 757 participants from Germany, Mexico, South Africa, the United Kingdom, and the United States. Below, we present Average Marginal Component Effects (AMCEs)[15] of the ambition and compliance likelihood attributes.[16]

[13]These figures—20 and 40%—where selected because they are among the most common NDC headline targets (Rowan 2019, supplementary material). 40% emissions cuts is the mode of headline targets, while 20% is the second most common target (tied with 15%). 20% (as of 1990 emissions) is also the mean of all NDC targets (Rowan 2019).

[14]Liu and Raferty (2021) project the probability of NDC target achievement for 122 countries. They find that the median probability of compliance with NDCs is 35% (Liu and Raferty 2021). Our two scenarios are thus ± 15% from the projected median probability of full compliance.

[15]AMCEs are the marginal effect of a given attribute averaged over the joint distribution of the remaining attributes (Hainmueller et al. 2014).

[16]We do not report the effects of the participation attribute here because it is not directly relevant for the ambition-compliance nexus. The results reported here are nonetheless unchanged if participation is also included in the analysis; see Supplementary Material G.

Figure 5: Average Marginal Component Effects (AMCEs) of Ambition and Compliance likelihood

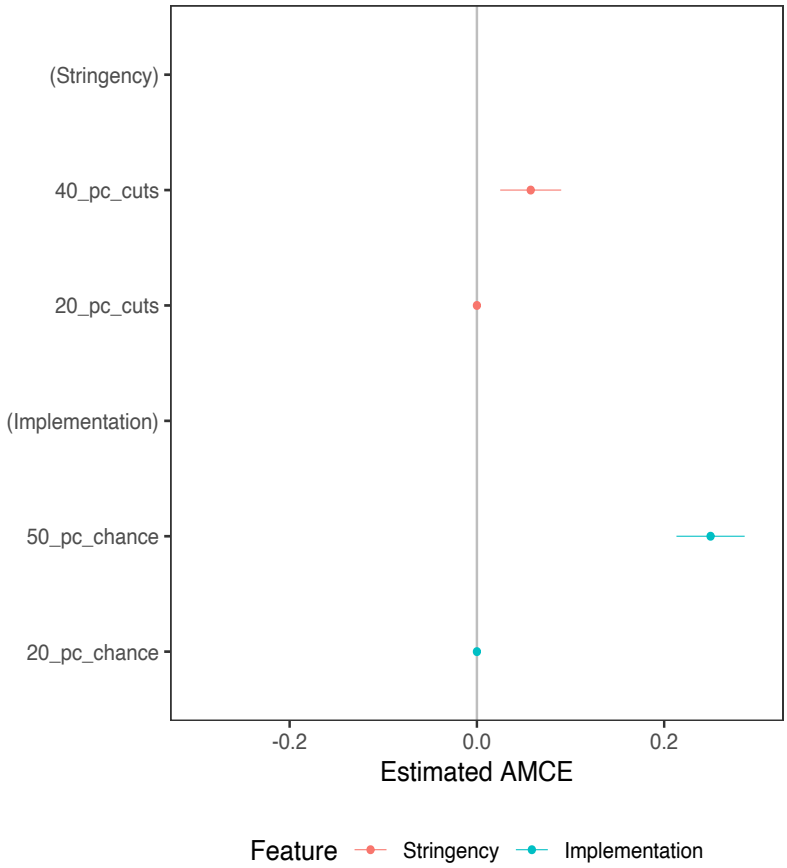


Figure 5 shows that both ambition and compliance likelihood positively affect people's selection of hypothetical climate agreements. However, compliance likelihood has a substantively stronger effect than ambition. Increasing the compliance likelihood from low likelihood (20%) to medium likelihood (50%) raises the probability that a respondent prefers a given climate agreement by 25%. The corresponding effect of increasing ambition from low ambition (20% cuts) to medium ambition (40% cuts) is 6%. While this controlled setting is decidedly stylized, the positive causal effect of compliance likelihood—and its relative importance compared to ambition—complements the previously identified prudence effect on the state level. Citizens in the five democracies we analyze react positively to compliance likelihood: if states anticipate domestic audience costs, the prudence effect of ambiguity can thus be understood as a rational accommodation to public opinion preferences. We underline here that our non-representative pool

of respondents are sampled from five democracies only, which is a most-likely setting for compliance concern to manifest. Yet, we view the positive causal effect of compliance likelihood among the general public—and the relative importance the public attaches to compliance compared to ambition—as a first basic validation of the compliance concern concept.

4.4 Discussion

Our analysis identifies a negative correlation between ambiguity and ambition in countries' climate pledges under the Paris Agreement. In light of our model, we interpret the overall negative relationship to mean that policymakers are prudent when they are faced with imperfect information concerning their country's true commitment potential. The prudence effect applies to approximately 70% of the states in our sample, but is amplified in democratic countries.[17] Contrastingly, the only imprudence effect we identify is among the subset of states that are highly autocratic. The heterogenous effects of form of governance suggest that exogenous ambiguity is an important factor in the determination of NDC ambition; and also that the anticipation of (domestic) compliance review is a likely motive for policy-makers' prudence. In support of the latter argument, our conjoint experiment shows that citizens in five democracies react more positively to higher compliance likelihood than to higher ambition. Hence, democratic policy-makers are rational to be compliance concerned in the face of domestic audience costs.

The findings bear implications for how ambiguity relates to cooperation under self-reporting regimes. First, a broader implication of the prudence effect is that ambiguous pledges are equally credible compliance signals as precise pledges. If ambiguity in climate pledges were primarily strategically determined to reap positive cooperative benefits (Aldy et al. 2016; Keohane and Oppenheimer 2016), we should have observed a positive correlation between ambiguity and ambition. Instead, our finding that countries with ambiguous pledges have adopted less ambitious targets suggests that these are more concerned about target achievement than signalling high ambition through audacious pledges. This finding corresponds with previous literature arguing that states care about their compliance record to such an extent that they will comply with international obligations even in the absence of effective formal enforcement mechanisms (Guzman 2008; Hafner-Burton et al. 2017). If states' willingness to adhere to the Paris Agreement depends on widespread compliance, the prudence effect can indicate that ambiguity is not necessarily detrimental to future cooperation under the Agreement. While

[17]The calculation is based on the share of states for which their level of democracy implies a negative relationship between ambition and ambiguity, across the three models in Table 2.

ambiguity reduces ambition, it does not necessarily undermine compliance—because less ambitious pledges are easier to comply with (Barrett 1999; Dimitrov et al. 2019).

Second, however, the prudence effect can prove harmful to cooperation by itself. If countries reciprocate ambition, ambiguity can reduce the prospect of positive cooperative cycles (see Supplementary Material H). Consequently, although the transparency-compliance nexus is not straightforward under self-reporting regimes, our findings are compatible with the notion that ambiguity can undermine reciprocal cooperation (Chayes and Chayes 1993; Mitchell 1998; Keohane and Victor 2011; Keohane and Oppenheimer 2016). We propose, however, that the detrimental effect of ambiguity manifests primarily through the ambiguity-ambition nexus, and not the ambiguity-compliance nexus.

Aside from the potentially negative cooperation effects, prudent ambition also reduces the environmental effectiveness of the Paris Agreement. In that respect, our analysis points to an untapped ambition potential for states with ambiguous pledges, which could allow substantial improvements in the global warming impact of their commitments (Rogelj et al. 2017). Given the prevalence of exogenous ambiguity, enhanced capacity building in low-income countries—through institutions such as the Paris Committee on Capacity-building and the Green Climate Fund—could be an efficient way to decrease NDC ambiguity and thereby potentially increase ambition (Chayes and Chayes 1993; Stender et al. 2019; Weikmans et al. 2019). Decreasing NDC ambiguity also has the added benefit that the collective goal achievement of the Agreement will be easier to assess, which has been shown to have beneficial effects on reciprocal cooperation (Barrett and Dannenberg 2012). Finally, however, capacity building is not a panacea to eliminate ambiguity, as our analysis also identifies a group of states that have pledged imprudently with endogenous ambiguity. Although endogenous ambiguity is not highly prevalent in states' pledges, our theoretical model suggests that the subset of states that have pledged ambiguously and face low risk of compliance review domestically are unlikely to fully comply with their pledged commitments.

5 Conclusion

This paper analyzes the relationship between ambiguity and ambition in climate pledges both theoretically and empirically. Theoretically, the paper constructs a model of ambiguity and risk behavior in climate pledges. The formal model takes into account both exogenous and endogenous commitment ambiguity, and describes why—in the face of compliance review—states have reason to exercise prudence when determining the ambition level of commitments. Empirically, the paper tests whether ambiguity in states'

climate pledges under the Paris Agreement is related to the ambition levels of their mitigation commitments. Based on our formal model, we indirectly identify whether states that have more ambiguous pledges have set more prudent or imprudent mitigation commitments than states with precise pledges.

Overall, we see the main contributions of this paper as twofold. First, our theoretical modelling of ambition, ambiguity, and compliance under (unenforced) self-reporting systems speaks to the literatures on institutional design (Rosendorff and Milner 2001; Koremenos 2005; Koremenos 2016; Creamer and Simmons 2019) and compliance under uncertainty (Chayes and Chayes 1993; Mitchell 1998; Koremenos 2005; Guzman 2008; Hafner-Burton et al. 2017) in international relations. The theoretical model, based on a trade-off between ambitious pledging and achievable compliance (Downs et al. 1996; Barrett 1999; Johns 2014; Dimitrov et al. 2019), shows how ambition can mediate the relationship between precise information provision and compliance with commitments. While existing literature posits a straightforward relationship between ambiguity and non-compliance (Chayes and Chayes 1993; Mitchell 1998; Simmons 1998; Keohane and Oppenheimer 2016), we show that ambiguity does not necessarily undermine compliance under a bottom-up treaty where governments unilaterally decide on the depth of cooperation. Instead, we argue that ambiguity incentivizes states to pledge prudent targets—which should translate to *higher* compliance rates, ceteris paribus. Drawing on Hafner-Burton et al. (2017), we propose that compliance concern is a key factor in inducing prudent commitments among states. We argue that form of governance is a useful proxy for compliance concern, and empirically demonstrate that democracies are more prudent than autocracies. Beyond the prudence effect, our model also highlights that the source of ambiguity (exogenous or endogenous) conditions how ambiguity is related to ambition and compliance, which gives rise to an analytically important distinction that has hitherto not been extensively discussed in the international relations compliance literature.

Our generalizable theory could extend to self-reporting regimes in areas such as international trade, human rights, or security (Creamer and Simmons 2019; Karlas 2021). Given the scope condition of a relevant compliance constituency, the prudence effect of exogenous ambiguity could in principle apply to any international self-reporting system in which policy-makers pledge a given political goal that is subsequently evaluated by either domestic audiences or other states. In this paper, we demonstrate that domestic audiences may impose non-compliance costs on governments in case of imprudent pledging: our conjoint experiment shows that citizens in five democracies value compliance likelihood over ambition. While the democracies are likely to be more compliance

concerned across different international institutions, the prudence effect we identify also depends on the issue area subject to cooperation. Notably, the prudence effect in our theoretical model originates in states' inability to perfectly gauge their commitment potential; and climate cooperation is an issue area with particularly high exogenous ambiguity (Rogelj et al. 2017). In other self-reporting regimes where states have better information about their commitment potential, endogenous ambiguity may be relatively more pronounced—which could generate false impressions of prudence and deficient compliance (e.g., human rights; see Hafner-Burton and Tsutsui 2005). On the other hand, compliance concern likely varies with institutional design: the Paris Agreement's review mechanism is largely a 'soft law' regime (Abbott and Snidal 2000), and systems with stronger enforcement mechanisms could incentivize even higher prudence than we identify here.

Second, our empirical analysis contributes to the literature on the effectiveness of the Paris Agreement (Barrett and Dannenberg 2016; Keohane and Oppenheimer 2016; Dimitrov et al. 2019; Tørstad 2020). Overall, we identify a negative relationship between ambiguity and ambition in states' climate pledges under the Paris Agreement, which suggests that countries are prudent when faced with ambiguity regarding their future emission reductions. That finding offers implications for the effectiveness prospects of the Paris Agreement, the success of which depends on both ambitious commitments and widespread compliance. An empirical implication of our model is that ambiguous mitigation commitments are unlikely to undermine compliance. Instead of pledging unrealistically high targets, our analysis suggests that the pledge-and-review system incentivizes states that face ambiguity to formulate targets they can realistically comply with. This finding aligns with the rational design literature highlighting the cooperative-inducing effects of flexibility (Milner and Rosendorff 2001; Koremenos 2005; Kucik and Reinhardt 2008). On the other hand, in a bottom-up setting where the ambition of targets is self-determined rather than mutually coordinated, states unilaterally lower the ambition of commitments in response to uncertainty about compliance prospects—leading to a negative cooperation effect of flexibility on depth of cooperation. The ambition level of ambiguous pledges are (on average) deflated compared to precise pledges: hence, states with ambiguous targets have more leeway to further enhance the ambition level of their commitments in the next round of pledging under the Paris Agreement. In sum, our analysis suggest that a subset of states that currently have ambiguous targets would pledge more ambitiously under a counterfactual agreement with lower levels of ambiguity. Presuming that states are inclined to reciprocate ambitious commitments, less ambiguity hence results in a more environmentally effective climate agreement.

Finally, we draw attention to three limitations of our analysis. First, our empirical strategy is unable to categorically determine the origins of ambiguity in pledges, including what amount of the observed ambiguity is exogenous or endogenous. Future research could better isolate the two concepts empirically and explore their causal effects more systematically. Relatedly, qualitative research on how state representatives formulate pledges could be helpful for understanding how ambiguity arises in climate pledges. Second, our point estimates of the effects of ambiguity may be biased by omitted variables. Hence, future research should identify other exogenous sources of ambiguity variation. Third, the current analysis has relied on the relationship between ambiguity and ambition to discuss the compliance prospects of the pledges. A caveat to this discussion is that it assumes that states with ambiguous commitments are equally capable of implementing their commitments as those with precise commitments. Whereas we have no reason to believe otherwise—e.g., capability does not correlate with ambiguity (Supplementary Material E)—the direct relationship between ambiguity and compliance should be tested when NDC implementation data become available.

References

[1] Abbott, Kenneth W., and Duncan Snidal. 2000. "Hard and Soft Law in International Governance." *International Organization* 54 (3): 421–56.

[2] Aldy, Joseph E. 2014. "The Crucial Role of Policy Surveillance in International Climate Policy." *Climatic Change* 126 (3–4): 279–92.

[3] Aldy, Joseph, William Pizer, Massimo Tavoni, Lara Aleluia Reis, Keigo Akimoto, Geoffrey Blanford, Carlo Carraro, et al. 2016. "Economic Tools to Promote Transparency and Comparability in the Paris Agreement". *Nature Climate Change* 6 (11): 1000–1004.

[4] Axelrod, Robert. 1984. *The Evolution of Cooperation.* New York, NY: Basic Books.

[5] Baliga, Sandeep, & Tomas Sjöström. 2008. "Strategic Ambiguity and Arms Proliferation." *Journal of Political Economy* 116 (6): 1023–57.

[6] Barrett, Scott. 1999. "A theory of full international cooperation." *Journal of Theoretical Politics* 11: 519–554.

[7] Barrett, Scott, & Astrid Dannenberg. 2012. "Climate negotiations under scientific uncertainty." *Proceedings of the National Academy of Sciences* 109 (43): 17372–17376.

[8] Barrett, Scott. (2013). Climate treaties and approaching catastrophes. *Journal of Environmental Economics and Management* 66(2): 235–250.

[9] Barrett, Scott, & Astrid Dannenberg. 2016. "An experimental investigation into "pledge and review" in climate negotiations." *Climatic Change* 138 (1-2): 339–351

[10] Berger, Loïc, and Massimo Marinacci. 2020. "Model uncertainty in climate change economics: A review and proposed framework for future research." *Environmental and Resource Economics* 77: 1-27.

[11] Best, Jacqueline. 2012. "Ambiguity and Uncertainty in International Organizations: A History of Debating IMF Conditionality." *International Studies Quarterly* 56 (4): 674–88.

[12] Bräuninger, Thomas, & Nathalie Giger. 2018. "Strategic Ambiguity of Party Positions in Multi-Party Competition." *Political Science Research and Methods* 6 (3): 527–48.

[13] Brown, Calum, Peter Alexander, Almut Arneth, Ian Holman, and Mark Rounsevell. 2019. "Achievement of Paris Climate Goals Unlikely Due to Time Lags in the Land System." *Nature Climate Change* 9 (3): 203–8.

[14] Burck Jan, Hagen Ursula, Franziska Marten, Niklas Höhne, & Christoph Bals. 2018. "The Climate Change Performance Index: Results 2019." Germanwatch, Bonn. Available at: `https://climate-change-performance-index.org/sites/default/files/documents/ccpi2019_results.pdf` (Accessed 25 March 2021).

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

[16] Climatewatch. 2022. *NDC Enhancement Tracker*. Available at: `https://www.climatewatchdata.org/2020-ndc-tracker` (Accessed 5 July 2022).

[17] Coppedge Michael, et al. 2017. V-Dem Dataset v7.1. Varieties of Democracy (V-Dem) Project. Available at: `https://www.v-dem.net/en/data/archive/previous-reference-materials/reference-material-v7/`. (Accessed 25 March 2021).

[18] Creamer, Cosette D., and Beth A. Simmons. 2019. "Do self-reporting regimes matter? evidence from the convention against torture." *International Studies Quarterly* 63 (4): 1051-1064.

[19] Dai, Xinyuan. 2005. "Why Comply? The Domestic Constituency Mechanism." *International Organization* 59 (2): 363–398.

[20] Dannenberg, A., Löschel, A., Paolacci, G., Reif, C., & Tavoni, A. (2015). "On the Provision of Public Goods with Probabilistic and Ambiguous Thresholds." *IEnvironmental and Resource Economics* 61(3): 365–383.

[21] Dimitrov, Radoslav, Jon Hovi, Detlef Sprinz, Håkon Sælen, and Arild Underdal. 2019. "Institutional and Environmental Effectiveness: Will the Paris Agreement Work?" *Wiley Interdisciplinary Reviews: Climate Change* 10 (4), e583.

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

[23] Falkner, Robert. 2016. "The Paris Agreement and the New Logic of International Climate Politics." *International Affairs* 92 (5): 1107–25.

[24] Guzman, Andrew. 2008. *How International Law Works: A Rational Choice Theory.* New York: Oxford University Press.

[25] Hafner-Burton, Emilie M., and Kiyoteru Tsutsui. 2005. "Human Rights in a Globalizing World: The Paradox of Empty Promises." *American Journal of Sociology* 110 (5): 1373–1411.

[26] Hafner-Burton, Emilie M., Zachary C. Steinert-Threlkeld, and David G. Victor. 2016. "Predictability Versus Flexibility: Secrecy In International Investment Arbitration." *World Politics* 68 (3): 413–53.

[27] Hafner-Burton, Emilie M., Brad L. LeVeck, and David G. Victor. 2017. "No False Promises: How the Prospect of Non-Compliance Affects Elite Preferences for International Cooperation." *International Studies Quarterly* 61 (1): 136–49.

[28] Hainmueller, Jens, Daniel J. Hopkins, and Teppei Yamamoto. 2014. "Causal Inference in Conjoint Analysis: Understanding Multidimensional Choices via Stated Preference Experiments." *Political Analysis* 22 (1): 1–30.

[29] Johns, Leslie. 2014. "Depth versus Rigidity in the Design of International Trade Agreements." *Journal of Theoretical Politics* 26 (3): 468–95.

[30] Karlas, Jan. 2021. "Why states inform: Compliance with self-reporting obligations in universal treaty regimes." *International Political Science Review.* DOI: https://doi.org/10.1177/0192512121990575.

[31] Keohane, Robert O. *After Hegemony: Cooperation and Discord in the World Political Economy.* Princeton, NJ: Princeton University Press.

[32] Keohane, Robert O., and Michael Oppenheimer. 2016. "Paris: Beyond the Climate Dead End through Pledge and Review?" *Politics and Governance* 4 (3): 142-151.

[33] Keohane, Robert O., and David G. Victor. 2011. "The Regime Complex for Climate Change." *Perspectives on Politics* 9 (01): 7–23.

[34] Khan, Mizan, David Mfitumukiza, and Saleemul Huq. 2019. "Capacity building for implementation of nationally determined contributions under the Paris Agreement." *Climate Policy* 20 (4): 499-510.

[35] Kolenikov, S., & Angeles, G. (2009). Socioeconomic status measurement with discrete proxy variables: Is principal component analysis a reliable answer?. *Review of Income and Wealth, 55*(1), 128-165.

[36] Koliev, Faradj, Thomas Sommerer, and Jonas Tallberg. "Compliance without Coercion: Effects of Reporting on International Labor Rights." *Journal of Peace Research* 58(3): 494–509.

[37] Koremenos, Barbara. 2005. "Contracting around International Uncertainty." *American Political Science Review* 99 (4): 549–65.

[38] Koremenos, Barbara. 2016. *The Continent of International Law.* Cambridge: Cambridge University Press.

[39] Kucik, Jeffrey, and Eric Reinhardt. 2008. "Does Flexibility Promote Cooperation? An Application to the Global Trade Regime." *International Organization* 62(3)): 477–505.

[40] Lancesseur, Nicolas, Colin Subtil, Camille Cros, and Jaakko Kooroshy. 2021. The COP26 Net Zero Atlas. `https://content.ftserussell.com/sites/default/files/research/the_COP26_Net_Zero_Atlas_FINAL.pdf` (Accessed: 1 July 2022).

[41] Liu, Peiran R., and Adrian E. Raftery. 2021. "Country-Based Rate of Emissions Reductions Should Increase by 80% beyond Nationally Determined Contributions to Meet the 2C Target." *Communications Earth Environment* 2(29). https://doi.org/10.1038/s43247-021-00097-8.

[42] Mitchell, Ronald B. 1998. "Sources of Transparency: Information Systems in International Regimes." *International Studies Quarterly* 42 (1): 109–30.

[43] ND–GAIN (Notre Dame Global Adaptation Index). 2015. ND–GAIN Vulnerability scores. Available at: `http://index.gain.org/ranking/vulnerability` (Accessed: 9 January 2018).

[44] Nyborg, Karine. 2018. "Reciprocal climate negotiators." *Journal of Environmental Economics and Management* 92: 707-725.

[45] Pauw, W. P., Richard J. T. Klein, Kennedy Mbeva, Adis Dzebo, Davide Cassanmagnago, and Anna Rudloff. 2018. "Beyond Headline Mitigation Numbers: We Need More Transparent and Comparable NDCs to Achieve the Paris Agreement on Climate Change." *Climatic Change 147* (1–2): 23–29.

[46] Pauw, W.P, Cassanmagnano, D., Mbeva, K., Hein, J., Guarin, A., Brandi, C., Dzebo, A., Canales, N., Adams, K.M., Atteridge, A., Bock, T., Helms, J., Zalewski, A., Frommé. E., Lindener, A., Muhammad, D. 2016. NDC Explorer. German Development Institute / Deutsches Institut für Entwicklungspolitik, African Centre for Technology Studies, Stockholm Environment Institute. `https://doi.org/10.23661/ndc_explorer_2017_2.0`

[47] Raiser, Kilian, Başak Çalı, and Christian Flachsland. 2022. "Understanding Pledge and Review: Learning from Analogies to the Paris Agreement Review Mechanisms." *Climate Policy* 0 (0): 1–17.

[48] Ray, Amal Kanti. 2008. "Measurement of social development: an international comparison." *Social Indicators Research, 86*, 1–46.

[49] Robiou du Pont, Yann, and Malte Meinshausen. 2018. "Warming Assessment of the Bottom-up Paris Agreement Emissions Pledges." *Nature Communications 9* (1).

[50] Robiou du Pont, Yann. 2022. "Pledged Warming Map." `http://paris-equity-check.org/warming-check.html` (Accessed 6 July 2022).

[51] Rogelj, Joeri, Oliver Fricko, Malte Meinshausen, Volker Krey, Johanna J. J. Zilliacus, and Keywan Riahi. 2017. "Understanding the Origin of Paris Agreement Emission Uncertainties." *Nature Communications* 8(1): 1–12.

[52] Rosendorff, B. Peter, and Helen V. Milner. 2001. "The Optimal Design of International Trade Institutions: Uncertainty and Escape." *International Organization* 55 (4): 829–57.

[53] Rowan, Sam S. 2019. "Pitfalls in Comparing Paris Pledges." *Climatic Change* 155 (4): 455–67.

[54] Simmons, Beth. 1998. "Compliance with International Agreements." *Annual Review of Political Science* 1: 75-93.

[55] Simmons, Beth. 2009. *Mobilizing for Human Rights: International Law in Domestic Politics*. Cambridge: Cambridge University Press.

[56] Simmons, Beth A. 2010. "Treaty Compliance and Violation." *Annual Review of Political Science* 13: 273-296.

[57] Staton, Jeffrey K., and Georg Vanberg. 2008. "The Value of Vagueness: Delegation, Defiance, and Judicial Opinions." *American Journal of Political Science 52* (3): 504–19.

[58] Stender, Frederik, Ulf Moslener, and Willem P. Pauw. 2019. "More than Money: Does Climate Finance Support Capacity Building?" *Applied Economics Letters 0* (0), 1–5.

[59] Sælen, Håkon, Vegard Tørstad, Christian Holz, and Tobias Dan Nielsen. 2019. "Fairness Conceptions and Self-Determined Mitigation Ambition under the Paris Agreement: Is There a Relationship?" *Environmental Science & Policy* 101: 245–254.

[60] Tallberg, Jonas. 2002. "Paths to Compliance: Enforcement, Management, and the European Union." *International Organization* 56 (3): 609–43.

[61] Tingley, Dustin, and Michael Tomz. 2014. "Conditional Cooperation and Climate Change." *Comparative Political Studies* 47 (3): 344–68.

[62] Tomz, Michael, and Robert P. Van Houweling. 2009. "The Electoral Implications of Candidate Ambiguity." *American Political Science Review 103* (01): 83–98.

[63] Tørstad, Vegard H. 2020. "Participation, Ambition and Compliance: Can the Paris Agreement Solve the Effectiveness Trilemma?" *Environmental Politics* 29(5): 761-80.

[64] Tørstad, Vegard, and Håkon Sælen. 2018. "Fairness in the Climate Negotiations: What Explains Variation in Parties' Expressed Conceptions?" *Climate Policy* 18 (5): 642–54.

[65] Tørstad, Vegard, Håkon Sælen, and Live Standal Bøyum. 2020. "The Domestic Politics of International Climate Commitments: Which Factors Explain Cross-Country Variation in NDC Ambition?" *Environmental Research Letters 15* (2), 024021.

[66] UNFCCC. 2015. Paris Agreement. Available at: `https://unfccc.int/sites/default/files/english_paris_agreement.pdf` (Accessed 25 March 2021).

[67] UNFCCC 2020. Survey on NDCs. Available at: `https://unfccc.int/sites/default/files/resource/Survey%20on%20NDC_Sharing%20%283%29.pdf` (Accessed 29 March 2022).

[68] UNFCCC. 2021. ADP bodies page. Available at `https://unfccc.int/adp-bodies-page` (Accessed 28 September 2021).

[69] Urminsky, Oleg, Christian Hansen, and Victor Chernozhukov. 2016. "Using double-lasso regression for principled variable selection." Available at SSRN: `http://dx.doi.org/10.2139/ssrn.2733374`.

[70] Victor, David G., Keigo Akimoto, Yoichi Kaya, Mitsutsune Yamaguchi, Danny Cullenward, and Cameron Hepburn. 2017. "Prove Paris Was More than Paper Promises." *Nature* 548 (7665): 25–27.

[71] Victor, David G., Marcel Lumkowsky, and Astrid Dannenberg. "Determining the Credibility of Commitments in International Climate Policy." *Nature Climate Change*, September 1, 2022, 1–8. https://doi.org/10.1038/s41558-022-01454-x.

[72] Waud, Roger N. 1976. "Asymmetric policymaker utility functions and optimal policy under uncertainty." *Econometrica* 41(1): 53-66.

[73] Weikmans, Romain, Harro van Asselt, and J. Timmons Roberts. 2019. "Transparency Requirements under the Paris Agreement and Their (Un)Likely Impact on Strengthening the Ambition of Nationally Determined Contributions (NDCs)." *Climate Policy* 20 (4): 511-526.

[74] World Bank. 2015a. Coal rents (% of GDP)

[75] World Bank. 2015b. Natural gas rents (% of GDP)

[76] World Bank. 2015c. Oil rents (% of GDP)

[77] World Bank. 2016. GDP per capita.

[78] World Resources Institute. 2022. Climate Watch database. Washington, D.C.: World Resources Institute. https://www.climatewatchdata.org/2020-ndc-tracker (Accessed 21 March 2022).

Supplementary material for ‘Commitment Ambiguity and Prudence in Climate Pledges’

Table of Contents

A Variables Included in the Analysis SM - 2

B Democracy and Procedural Compliance with the Paris Agreement SM - 5

C Robustness tests SM - 8

D Weights in the Composite Indices SM - 17

E Correlates of NDC Ambiguity SM - 20

F Proof of Result 1 (Waud 1976) SM - 21

G Conjoint experiment: Methodological information and Ethical Considerations SM - 23

H Game-theoretical Analysis of Exogenous Ambiguity and Cooperation SM - 27

I Regression tables with full models SM - 30

A Variables Included in the Analysis

Table A.1: Ambiguity and control variables

Panel A: Ambiguity dimensions

Variables	Description
Type of target	Discrete variable measuring whether a country's NDC mitigation target is formulated as an absolute target (compared to a base year); a target relative to a business-as-usual trajectory; an intensity target (e.g. GHG emissions per unit of GDP); a peaking target (specifying a date by which GHG emissions will peak; a 'policy and actions' target (which does not say anything particular about emissions trajectories); or an 'adaption with mitigation co-benefits' target. The specificity of target type is listed in the order of ambiguity, with absolute targets being least ambiguous and policy and actions targets being most ambiguous. Values: 0/5.
Mitigation costs (ccm)	Dummy variable indicating whether countries' NDCs include cost estimations of the pledged mitigation target. If costs are estimated, the NDC is coded as more precise. Values: 0/1.
Renewable energy	Dummy variable indicating whether renewable energy is considered in order to reach the pledged mitigation target. If renewable energy is considered, the NDC is coded as more precise. Values:0/1.
Energy efficiency	Dummy variable indicating whether energy efficiency is considered in order to reach the pledged mitigation target. If energy efficiency is considered, the NDC is coded as more precise. Values: 0/1.
Transport	Dummy variable indicating whether transport sector is considered in order to reach the pledged mitigation target. Value: 0/1.

Carbon capture and storage	Dummy variable indicating whether carbon capture and storage is considered in order to reach the pledged mitigation target. Values: 0/1.
Agriculture	Dummy variable indicating whether agriculture is considered in order to reach the pledged mitigation target. Values: 0/1.
Land use and forestry	Dummy variable indicating reference to land use and forestry and whether emissions and mitigation potential are quantified. Values: 0/1.
Mitigation documents	Dummy variable indicating reference to domestic or international (or both) mitigation plans and strategies. Values: 0/1.
Reducing non-co2 gases	Dummy variable indicating whether a country has considered the reduction of non-Co_2 gases in their pledged mitigation target. Values: 0/1.
Land use change	Discrete variable indicating reference to land-use change in order to reach the pledged mitigation target. Values: 0/3.
Conditionality of finance	Discrete variable indicating whether the NDC mentions and the extent to which pledged mitigation target is conditional on international financial support. Values: 0/3.
Technology needs	Discrete variable indicating reference to (specific) technologies to use for adaption or mitigation. Values: 0/2.
Conditionality of technology transfers	Discrete variable indicating whether achievement of the pledged mitigation target is conditional on technology transfers. Values: 0/2.
Conditionality of capacity building	Discrete variable indicating whether achievement of the pledged mitigation target is conditional on capacity-building measures. Values: 0/2.
Planning of NDC formulation	Dummy variable capturing whether the NDC includes references to the planning process of the NDC. Values: 0/1.

Stakeholder consultation	Dummy variable indicating whether stakeholders were consulted in the NDC formulation process. Values: 0/1.
Planning of NDC implementation	Discrete variable indicating mentioning of how NDC targets are to be implemented and whether references are made to domestic laws and policies. Values: 0/2.
Monitoring and review	Dummy variable indicating reference to national or international (or both) assessments and review of NDCs. Values: 0/1.
Waste	Dummy variable indicating whether waste sector is considered in order to reach the pledged mitigation target. Values: 0/1.

Panel B: Controls included in the empirical analyses	
GDP	Logarithm of PPP-adjusted GDP per capita (internationals dollars, 2015). Higher scores=higher GDP per capita (World Bank 2016).
Democracy index	Country scores on the 2015 V-Dem multiplicative polyarchy index (Coppedge et al. 2017). The index measures a country's degree of freedom of association, clean elections, freedom of expression, elected executives and suffrage. Higher scores=higher level of democracy.
Climate change vulnerability	ND-GAIN Vulnerability index (ND-GAIN 2015). Measures a country's exposure, sensitivity and capacity to adapt to the negative effects of climate change. The index (which is GDP-adjusted) captures overall vulnerability by considering six life-supporting sectors – food, water, health, ecosystem service, human habitat, and infrastructure. Higher scores=higher vulnerability to climate change.

Coal rents	The difference between the value of both hard and soft coal production at world prices and their total costs of production (World Bank 2015a). Measured as % of GDP. Higher scores=higher coal rents..
Oil rents	The difference between the value of crude oil production at regional prices and total costs of production (World Bank 2015c). Measured as % of GDP. Higher scores=higher oil rents.
Natural gas rents	The difference between the value of natural gas production at regional prices and total costs of production (World Bank 2015b). Measured as % of GDP. Higher scores=higher natural gas rents.

Notes: Panel A in this table shows a list of all ambiguity variables that are extracted from the NDCs. For each variable, higher values indicate more less precision and more ambiguity. Panel B lists the variables that are noted as 'control' in our empirical analyses. The second columns provide short descriptions of the variables.

B Democracy and Procedural Compliance with the Paris Agreement

A well-established finding in the international compliance literature is that the stronger accountability mechanisms in democracies render them more conducive to comply with their international obligations than autocracies (Dai 2005; Simmons 2009; Creamer and Simmons 2019). We therefore conjecture that policy-makers in democracies are more *compliance concerned* than those in autocracies. As explained in the main text, our concept of compliance concern is theoretically based on Hafner-Burton et al. (2017), who argue that decision-making elites vary in the rate with which they discount the future downside risk of non-compliance. While the authors of that study do not systematically theorize what explains differences in compliance concern, they do speculate that:

> (...) leaders in autocracies may have different attitudes about compliance risk than those from democracies. Perhaps democratic leaders, for example, are much more aware of the many ways that national political processes can yield involuntary defection—and also political pressures for compliance. This would make them more sensitive to how such outcomes harm the prospects for international cooperation. (Hafner-Burton et al. 2017, p. 147)

This expectation also corresponds with Fearon (1994), who suggests that democracies have higher domestic audience costs than autocracies, as foreign policy in democracies is made by an agent (government officials) on behalf of principals (voters) who have the power to sanction the agent electorally or through the workings of public opinion.

Does the relationship between form of governance and actual compliance also hold for climate cooperation under the Paris Agreement? Since we use democracy as a proxy for compliance concern in our empirical analysis, a crucial test for the validity of our analysis is whether democracies actually are more likely to comply with the requirements of the Paris Agreement than autocracies. While it is too early to measure the actual implementation trajectories of countries' NDCs (most mitigation targets in NDCs are due in either 2025 or 2030), an early indicator of actual compliance is whether countries have *updated* their NDCs. One of the few legally binding provisions in the Paris Agreement is that countries update their NDCs every fifth year (UNFCCC 2015). The original due date for the first NDC update was in early 2020, but was later extended to December 31, 2020 due to COVID-19. As of June 2022, 155 countries (including the EU) have submitted updated NDC targets, while 44 countries have not yet updated their NDCs. Of the 155 countries that have updated, 63 did so by the agreed deadline and 92 belatedly. As the updating of NDCs is a legally binding requirement under the Paris Agreement, the countries that have not updated their NDCs are in non-compliance with a key procedural provision of the Agreement. We leverage the variation in which countries that have updated their NDCs to test whether democracy predicts actual (procedural) compliance under the Agreement. We distinguish between three categories of procedural compliance: 1) updated on time (full compliance); 2) updated belatedly (partial compliance), and 3) no updating (non-compliance). If democracies are more compliance concerned, we should observe a positive correlation between NDC updating and democracy.

Table B2 below shows an ordered logistic regression model with our procedural compliance measure as dependent variable. The control variables are the same as in the main analyses, including form of governance. The model indicates that democracy is only country characteristic that predicts procedural compliance, in line with our theoretical expectation that democracies are more compliance concerned.

Table B.1: Correlation between NDC Uspdating and Country Characteristics

Dep Var: Updated NDC	(1) b/se
Democracy	1.814**
	(0.783)
Coal rents	-0.0394
	(0.663)
Oil rents	-0.0344
	(0.0230)
Natural gas rents	-0.334*
	(0.177)
Climate change vulnerability	-1.354
	(3.101)
GDP (log)	0.298
	(0.264)
cutpoint 1	0.710
	(3.532)
cutpoint 2	3.509
	(3.543)
Observations	157
Psudo R^2	0.13

Notes: This table displays an ordered logistic regression model with timing of *Updated NDC*, as the dependent variable. The dependent variable is equal to 0 if a country's NDC has not yet been updated; 1 if the NDC was updated after the deadline; and 2 if the NDC was updated before the deadline. The control variables are country characteristics described in Table A.1, Panel B. Standard errors are in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels.

C Robustness Tests

In this section we present six tables, four of which regard the robustness of correlation between prudence and compliance concern as tested in Table 2 and two exploiting updated NDC data. In Table C.1 we reproduce Table 2 without using the double lasso procedure, while in Table C.2 we use countries type of target as a measure of ambiguity (see Table A.1 for definition) in addition to an index where all 20 ambiguity dimensions were standardized before added together. In Tables C.3 and C.4 we use two alternative measures of ambition and, finally, Table C.5 and C.6 display analyses of updated NDCs.

Table C.1: Conditional Effect of Democracy on the Ambiguity-Ambition Nexus: Without lasso-selection of controls

Dep Var: NDC Ambition	No weights b/se	Regression weights b/se	PCA b/se
Ambiguity index	0.0362	0.500	0.519***
	(0.0457)	(0.465)	(0.181)
Democracy	4.306***	1.161**	1.100***
	(1.099)	(0.469)	(0.405)
Ambiguity index x Democracy	-0.255***	-2.978**	-1.639***
	(0.0975)	(1.194)	(0.336)
GDP (log)	-0.811***	-0.776***	-0.818***
	(0.183)	(0.185)	(0.179)
Climate change vulnerability	5.801**	5.692**	5.834***
	(2.262)	(2.250)	(2.090)
Coal rents	-0.850***	-0.969***	-0.795***
	(0.210)	(0.236)	(0.210)
Oil rents	0.0144	0.0128	0.0203
	(0.0207)	(0.0206)	(0.0205)
Natural gas rents	-0.0414	-0.0569	-0.0440
	(0.0544)	(0.0521)	(0.0470)
Observations	148	148	148
R^2	0.574	0.567	0.603

Notes: This table displays the effects of the three ambiguity indices on states' NDC ambition using OLS regression. *Democracy* is a continuous measure of states' level of democracy and a proxy for concern for compliance. *Ambiguity index x Democracy* is its interaction with the ambiguity indices. Control variables are country characteristics that are described in Table 2, Panel B. Robust standard errors are in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels.

Table C.2: Democracy, Ambiguity and Ambition: Using Type of Target and Index with standardized components

	(1)	(2)
Dep Var: NDC Ambition	b/se	b/se
Index (std)	0.00716	
	(0.0265)	
Type of target		0.274**
		(0.122)
Index (std) x	-0.122**	
Democracy	(0.0607)	
Type of target x		-0.834***
Democracy		(0.301)
Democracy	0.647	1.849***
	(0.505)	(0.632)
Controls	Yes	Yes
Double-lasso	Yes	Yes
Observations	148	149

Notes: This table displays the effects of the two alternative ambiguity measures on states' NDC ambition using OLS regression. In column 1 we use an additive index of where each of the 20 ambiguity dimensions is standardized. In column two we use states' type of target as the ambiguity measure. *Democracy* is a continuous measure of states' level of democracy and a proxy for concern for compliance. *[Ambiguity] x Democracy* is its interaction with the ambiguity indices. Control variables are selected using the double-lasso selection procedure. Robust standard errors are in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels.

C.1 Alternative Dependent Variables

Table C3 and C4 re-runs the main regressions reported in Table 2 (main text) with two alternative climate ambition metrics: the CLAIM model (Lancesseur et al. 2021) and the Climate Change Performance Index (CCPI) (Burck et al. 2018). The CLAIM model is most similar to the NDC ambition variable that we use in the main analysis (Robiou du Pont and Meinshausen 2018), in that it also measures the implied temperature rise of NDC targets. The CCPI, on the other hand, measures climate policy more broadly–and includes national progress on GHG emissions policies, renewable energy policies, energy use, in addition to a country's pledged international efforts. For both of the alternative metrics, the main results are very similar to the results we report in the main text. However, statistical power is lower due to much lower numbers of countries covered.

Table C.3: Main Results with Alternative Dependent Variable: CLAIM Model

Dep Var: NDC Ambition (CLAIM)	No weights b/se	Regression weights b/se	PCA b/se
Ambiguity index	0.149***	0.568	0.483*
	(0.0448)	(0.440)	(0.257)
Democracy	4.358***	0.735	0.853**
	(0.943)	(0.501)	(0.412)
Ambiguity index x Democracy	-0.298***	-2.276*	-0.634
	(0.0922)	(1.166)	(0.452)
GDP (log)	-0.536***	-0.572***	-0.365**
	(0.180)	(0.189)	(0.155)
Climate change vulnerability	0.802	0.591	0.107
	(2.495)	(2.261)	(2.520)
Coal rents	-0.533	-0.700	-0.708
	(0.428)	(0.430)	(0.536)
Oil rents	-0.105***	-0.0582***	-0.0441***
	(0.0213)	(0.0188)	(0.0133)
Natural gas rents	0.0589	-0.128	0.00925
	(0.135)	(0.150)	(0.165)
Observations	46	46	46
R^2	0.526	0.537	0.489

Note: This table displays three OLS models regressing the three ambiguity indices on states' NDC ambition. The dependent variable is an alternative measure of NDC ambition derived from the CLAIM model (Lancesseur et al. 2021). *Democracy* is a continuous measure of states' level of democracy and a proxy for compliance concern. *Ambiguity index x Democracy* is its interaction with the ambiguity indices. Control variables are described in Table A1. Robust standard errors are in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels.

Table C.4: Main Results with Alternative Dependent Variable: CCPI

Dep Var: Climate Ambition (CCPI)	No weights b/se	Regression weights b/se	PCA b/se
Ambiguity index	0.343	3.691	0.460
	(0.680)	(5.377)	(2.203)
Democracy	43.20**	6.766	15.26**
	(17.34)	(5.072)	(5.958)
Ambiguity index x Democracy	-2.734*	-39.56**	-4.382
	(1.461)	(15.64)	(5.523)
GDP (log)	-18.49***	-18.00***	-17.39***
	(5.072)	(4.415)	(5.016)
Climate change vulnerability	-58.59	-49.38	-54.91
	(54.11)	(47.60)	(54.94)
Coal rents	-9.975*	-13.35***	-13.98**
	(5.714)	(4.106)	(5.708)
Oil rents	-1.085***	-1.049***	-1.044**
	(0.377)	(0.337)	(0.399)
Natural gas rents	0.303	-0.668	0.774
	(2.230)	(1.861)	(2.482)
Observations	53	53	53
R^2	0.526	0.537	0.489

Note: This table displays three OLS models regressing the three ambiguity indices on states' climate ambition. The dependent variable is an alternative measure of NDC ambition derived from the Climate Change Performance Index (CCPI) (Burck et al. 2018). *Democracy* is a continuous measure of states' level of democracy and a proxy for compliance concern. *Ambiguity index x Democracy* is its interaction with the ambiguity indices. Control variables are described in Table A1. Robust standard errors are in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels.

C.2 Analyses of Updated NDCs

In this section we present two analyses using data on the updated NDCs, which were due by the end of 2020. First, Table C.5 shows the correlation between updating of information—which makes an NDC more precise—and increased ambition in the updated NDCs' mitigation targets. These information and ambition variables are based on data from Climatewatch (2022), and are different to the information and ambition metrics that we use in the main text. The battery of controls, however, is the same as in Table 2 in the main text. In line with our prudence conjecture, we find that increased information provision in the enhanced NDCs strongly correlates with increased NDC ambition. This finding resonates with our argument that reducing exogenous ambiguity in NDCs should lead to higher ambition (ceteris paribus).

Table C.5: Correlation between Increased Ambition and Increased Information in the Updated NDCs

	(1)
Dep Var: NDC increased ambition	b/se
NDC increased information	0.475***
	(0.0751)
All controls included	Yes
Double Lasso	Yes
Observations	157

Notes: This table shows the correlation between updated information and updated ambition in the second round of NDCs. Increased ambition is equal to 1 if a state increased its ambition and 0 otherwise. Increased information is 1 if a state updated its NDC and increased its information and 0 otherwise. Control variables are described in detail in Table A.1, Panel B. Robust standard errors are in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels.

Table C.6 reproduces the main analyses from Table 2, but uses data on the updated NDCs instead of the initial NDCs. We again collect ambiguity data from Pauw et al. (2016)'s *NDC Explorer*, but this time for the updated NDCs. The updated NDC ambition data are from Robiou Du Pont (2022), using the same method as for our

main NDC ambition metric but for a much smaller subset of countries. The preliminary updated NDC ambition data are available at: `http://paris-equity-check.org/warming-check.html`. Unfortunately, there is only a very limited number of countries for which the updated NDC ambiguity data *and* updated NDC ambition data is available thus far. Hence—while Table C.6 shows the same exact prudence effects as the main analysis of initial NDCs (Table 2)—the limited sample size precludes any firm inference.

Table C.6: Conditional Effect of Democracy on the Ambiguity-Ambition Nexus: Using Data from Updated NDCs

Dep Var: NDC Ambition (updated)	No weights b/se	PCA b/se
Amiguity index	0.897***	1.154***
	(0.204)	(0.320)
Democracy	7.903***	1.584*
	(1.340)	(0.801)
Ambiguity index x Democracy	-1.037***	-0.914*
	(0.306)	(0.444)
GDP (log)	0.104	0.525
	(0.455)	(0.535)
Climate change vulnerability	10.72*	9.826*
	(5.287)	(4.850)
Coal rents	-0.589	-0.640
	(0.451)	(0.371)
Oil rents	-0.0294	-0.198
	(0.219)	(0.245)
Natural gas rents	0.637	0.808
	(0.429)	(0.498)
Observations	26	26
R^2	0.727	0.760

Notes: This table displays OLS regressions with two ambiguity indices (unweighted additive index and PCA index) and NDC ambition. Both the ambiguity indices and ambition levels are based on data from the updated NDCs. *Democracy* is a continuous measure of states' level of democracy and a proxy for compliance concern. *Ambiguity index x Democracy* is its interaction with the ambiguity indices. Control variables are country characteristics that are described in Table A1, Panel B. Robust standard errors are in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels.

D Weights in the Composite Indices

This section gives information regarding the weighting procedures in the three different types of ambiguity indices. While the regression and PCA indices take into account correlations, our unweighted index is an alternative which treats all dimensions equally although they are at slightly different scales. The resulting weights are displayed in Table D.1. In the column named 'Equal', variables are weighted equally. To construct the regression index (rightmost column) we compute weights based on the relative explanatory power of variables in predicting the variable 'type of target' (see Ray (2008) for an overview of composite indices). First, we run an OLS regression on 'type of target'. The weight of a variable is the size of the absolute value of the regression coefficient relative to the sum of the absolute values of all coefficients.[18]

The PCA analysis evaluate what variables capture the same latent concept among the 20 ambiguity variables we have selected from the NDCs. After performing the PCA analysis, we rotate the factor loading matrices producing orthogonal components. We use the first principal component which, by far explains most of the variance compared to the other components, 22%. Ideally, continuous variables are used in PCA analyses (Kolenikov and Angeles, 2009). Although suboptimal, we use both dummies and ordinal level categorical variables. We treat the ordinal variables as continuous avoiding problems with dependence between dummies created from categories (see Kolenikov and Angeles, 2009).

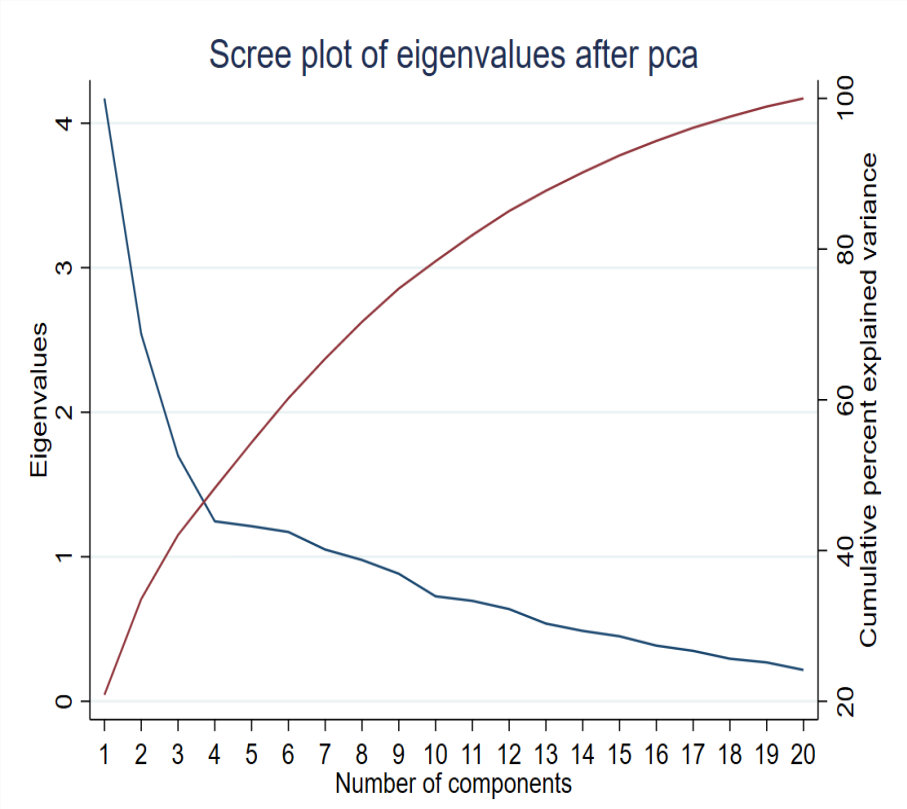
[18]In SM C we reproduce Table 2 in the main text using an unweighted index with standardized dimension and type of target as measures of ambiguity.

Table D.1: Weights of variables for different indices

	Equal	Reg
Mitigation costs (ccm)	1	0.051
Type of target	1	0
Renewable energy	1	0.007
Energy efficiency	1	0.069
Transport	1	0.027
Carbon capture and storage	1	0.080
Agriculture	1	0.007
Land use and forestry	1	0.022
Mitigation documents	1	0.050
Reducing non-co2 gases	1	0.171
Land use change	1	0.011
Conditionality of finance	1	0.072
Technology needs	1	0.027
Conditionality of technology transfer	1	0.168
Conditionality of capacity building	1	0.008
Planning of NDC formulation	1	0.039
Stakeholder consultation	1	0.082
Planning of NDC implementation	1	0.039
Monitoring and review	1	0.054
Waste	1	0.021

Notes: This table shows the weights that we use to construct our three ambiguity indices. While weights are displayed with three decimals, we used nine decimals in the analysis. See further details in the text above.

Figure D.1: Scree plot of eigenvalues and variance after PCA



Notes: This figure shows the relationship between the number of components, eigenvalues and the cumulative explained variance The blue curve display eigenvalues with corresponding on the y-axis on the left. The red curve is the cumulative explained variance and has corresponding values on the y-axis on the right.

E Correlates of NDC Ambiguity

Table E.1: Correlates of NDC Ambiguity

	No weights	Regression weights	PCA
Dep Var: Ambiguity index	b/se	b/se	b/se
GDP (log)	0.489	0.00823	-0.292**
	(0.514)	(0.0580)	(0.116)
Democracy	-3.757***	-0.244*	-0.307
	(1.240)	(0.129)	(0.275)
Climate change vulnerability	17.89***	0.832	2.801**
	(6.019)	(0.648)	(1.327)
Coal rents	0.915	-0.0409	0.345***
	(0.789)	(0.0442)	(0.130)
Oil rents	0.00955	0.00294	0.0104
	(0.0627)	(0.00425)	(0.0104)
Natural gas rents	0.416***	0.0243**	0.00154
	(0.136)	(0.0121)	(0.0177)
Constant	0.779	-0.413	1.468
	(6.921)	(0.772)	(1.588)
Observations	156	156	156
R^2	0.324	0.180	0.432

Notes: This table displays OLS regressions with the three ambiguity indices as dependent variables and country characteristics as independent variables. These country characteristics are described in Table 2, Panel B. Robust standard errors are in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels.

E.1 Ambiguity and Missing Values on the Control Variables

Table E.2: Correlation between Missing Values, Ambiguity and Ambition

Dep. Var:	Ambiguity (No weights) b/se	Ambiguity (Regression weights) b/se	Ambigutiy (PCA) b/se	Ambition b/se
Missing	2.994***	0.235***	0.309*	0.202
	(0.719)	(0.0653)	(0.186)	(0.386)
Constant	12.01***	-0.0505*	-0.0606	2.035***
	(0.310)	(0.0263)	(0.0787)	(0.133)
Observations	194	194	194	168
R^2	0.086	0.072	0.015	0.002

Notes: This table displays OLS regressions with the three NDC ambiguity indices and ambition as dependent variables. *Missing* takes the value 1 if missing values disallow the inclusion of a state in the regressions (with controls) in our empirical analyses in Table 1 and 2 in the main text, and 0 otherwise. Robust standard errors are in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels.

F Proof of Result 1 (Waud 1976)

The optimal value of μ_C, is analogous to Proposition 2 in Waud (1976), page 56. The difference is that we use C and Ω instead of y and λ, respectively. We use Ω to avoid confusion with a parameter in the empirical section of our paper. The following closely follows the proofs of Proposition 1 and 2 in Waud's (1976).

Proof of $p < \mu_q$ *if* $\sigma_q^2 > 0$ *and* $C^* = \mu_q - p = 0$.

The aim of a given state is to minimize the following expression:

$$\underset{\mu_C}{\text{minimize}} \quad -E[U] = \Omega \int_{-\infty}^{0} u(C) f(C;\mu_C) dC + \int_{0}^{\infty} u(C) f(C;\mu_C) dC$$

First define $\Phi(\mu_C;\Omega,\sigma_q^2) = -E[U]$. Suppose first that $\Omega = 1$ and that $C^* = 0$. Then we have:

$$-E(U) = \int_{-\infty}^{\mu_C} u(C) f(C,\mu_C) dC + \int_{\mu_C}^{\infty} u(C) f(C;\mu_C) dC \tag{7}$$

Given the assumptions about u and f we have the following equivalences

$$\int_{-\infty}^{\mu_C} u(C)f(C,\mu_C)dC = \int_{-\infty}^{0} u(C+\mu_C)f(C;0)dC \tag{8}$$

which is due the right side of the equality being shifted leftwards. By symmetry

$$\int_{\mu_C}^{\infty} u(C)f(C,\mu_C)dC = \int_{-\infty}^{0} u(C-\mu_C)f(C;0)dC \tag{9}$$

Also have by symmetry that

$$\int_{0}^{\infty} u(C+\mu_C)f(C,0)dC = \int_{-\infty}^{0} u(C-\mu_C)f(C;0)dC \tag{10}$$

By inserting (9) and (10) into (8), we get:

$$\begin{aligned}
-E(U) &= \int_{-\infty}^{0} u(C+\mu_C)f(C,0)dC + \int_{-\infty}^{0} u(C-\mu_C)f(C;0)dC \\
&= \int_{-\infty}^{0} u(C+\mu_C) + u(C-\mu_C)f(C;0)dC \\
&> 2 * \int_{-\infty}^{0} u(C)f(C,0)dC = \Phi(0;1,\sigma_q^2)
\end{aligned}$$

Hence, $\mu_C = 0$ minimizes $-E[U]$ for $\Omega = 1$. Now, suppose $\Omega > 1$ and that $-E[U]$ is differentiable at $C = 0$. Low realizations of C are now given large weight. Hence, the expected loss becomes

$$-E(U) = \Omega \int_{-\infty}^{0} u(C)f(C,\mu_C)dC + \int_{0}^{\infty} u(C)f(C;\mu_C)dC$$

Add and subtract $\int_{-\infty}^{0} u(C)f(C,\mu_C)dC$ on the right side to get

$$-E(U) = (\Omega - 1)\int_{-\infty}^{0} u(C)f(C,\mu_C)dC + \Phi(\mu_C;1,\sigma_q^2)$$

By evaluating the last expression we see that integral term is decreasing μ_C as the probability mass moves further beyond $C = 0$, and it goes towards 0 in the limit. Conversely, it increases as μ_C decreases. As showed above $\Phi(\mu_C;1,\sigma_q^2)$ attains its minimum at $\mu_C = 0$ and it goes to infinity as $|\mu_C|$ is increasing. At $\mu_C = 0$ this expression is not changing as it is at its minimum (The derivative is zero). Hence, at $\mu_C = 0$, we have that $\int_{-\infty}^{0} u(C)f(C,\mu_C)dC$ is decreasing and $\Phi(\mu_C;1,\sigma_q^2)$ is constant, meaning that there

is scope for decreasing $-E(U)$. As μ_C increases the former decreases while the latter increases, meaning that there exists a minimum at $\mu_C > 0$. Finally, since in $C = q - p$ in our framework this implies that $p < q$ to hedge against low realizations of q. One can use the same argument as above to show that the reverse is true if $\Omega < 1$.

We refer to Waud (1976) pages 56-58 for discussion and comparative statics on how the the optimal value of C changes depending on the value of σ_q^2 and Ω using a quadratic utility function.

G Conjoint experiment: Methodological information and ethical considerations

We implemented a conjoint experiment where participants were asked to choose between two hypothetical climate agreements. Participants were provided the following information:

> Please read the following hypothetical scenario:
>
> *The government of your country is participating in negotiations of a major climate agreement that is aimed to reduce global greenhouse gas emissions. All countries in the world are taking part in the negotiations.*
>
> We would now like to show you a pair of different climate agreements that your country could sign. We will then ask you to choose which of these agreements you would prefer.

As described in the main text of this article, the pairs of climate agreements had three attributes—participation, stringency, and implementation—with two levels each. The values of the levels (participate vs not participate; 40% cuts versus 20% cuts; 20% compliance likelihood versus 50%) were randomized. Moreover, each participant was given the task to select a preferred agreement twice in order to improve the precision of our estimates. We recruited 757 participants through Prolific (www.prolific.co), and the conjoint was administered through Qualtrics (www.qualtrics.com). The gender balance and nationality of participants were balanced (50% men and women. 150 participants from Germany; 155 from Mexico; 151 from South Africa; 151 from the UK; 150 from the US) but the respondent samples were otherwise not representative of the respective national populations.

The conjoint experiment opened with a consent form that the participants had to read and actively consent to in order to proceed. The consent form followed the recommendations of the General Data Protection Regulation and Personal Data Act. The consent form contained information about the purpose of the project, who were responsible for data collection and storage, how the data would be stored, and a privacy statement regarding the collection and use of personal data. The consent form also explicitly stated that participation was voluntary and that participants had the option to withdraw from the experiment at any point. There was no deception involved.

The experiment took approximately 15 minutes to complete, and participants were compensated £3.50 for their participation. This compensation was deemed 'good' by the company that recruited our participants (Prolific). We did not ask participants about personal information such as name or contact details, and did not collect email addresses or IP addresses. The responses could not be traced back to individuals. Prolific's anonymization system ensured that no one—neither we (the researchers) nor Prolific—could access data that could be directly linked to individual participants. As per Norwegian higher education guidelines, our conjoint experiment was exempt from review by relevant ethics boards because the data collection procedure was fully anonymized.

Figure G.1 shows that there were no significant spillover effects resulting from asking respondents to rate two climate agreements. Figure G.2 shows the full results, including the participation attribute. It shows that participation exerts an almost equally strong positive effect on people's preferred climate agreement as compliance likelihood. Finally, Figure G.3 shows the full results grouped by respondent nationality. The figure reveals some heterogeneity in causal effects by nationality for all treatments.

Figure G.1: Spillover effects test

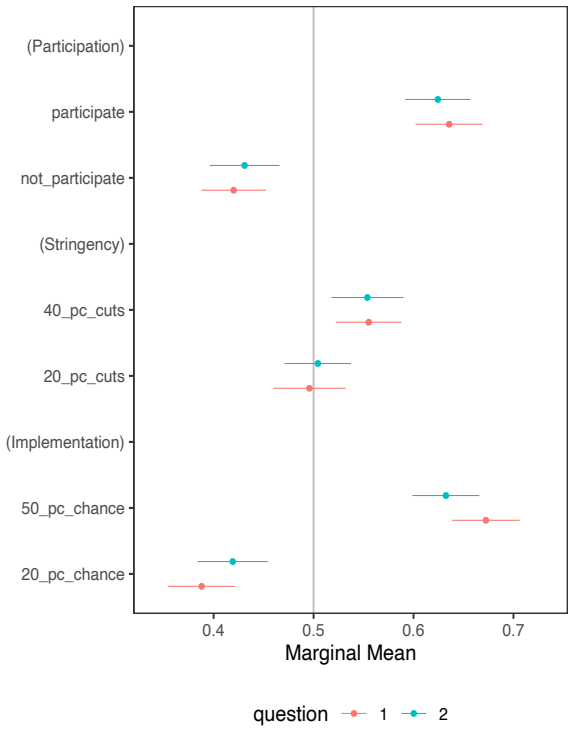


Figure G.2: Full results

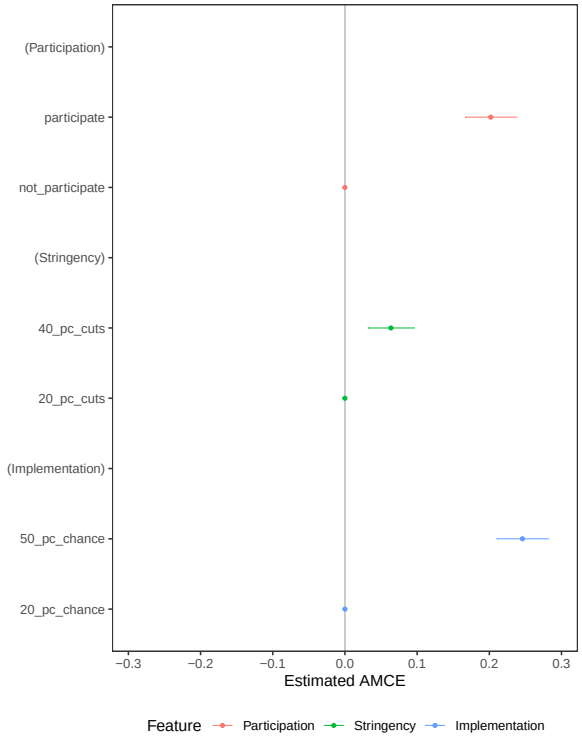
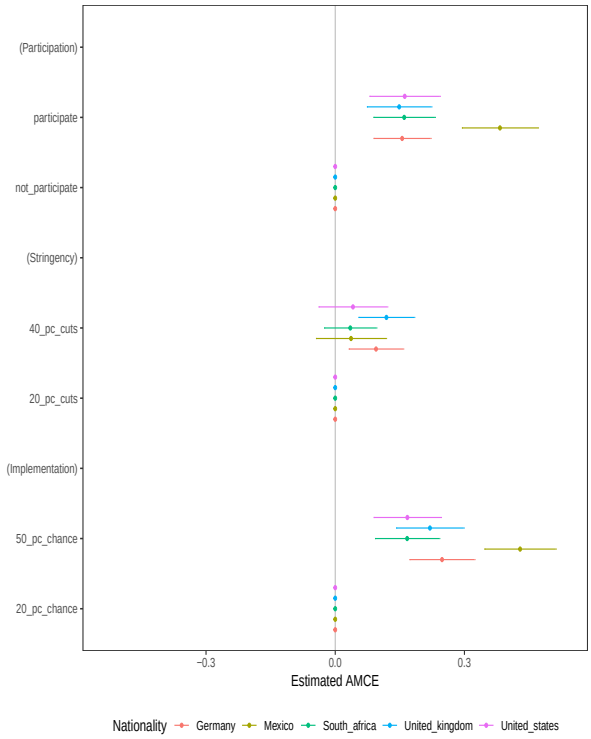


Figure G.3: Full results by nationality



H Game-theoretical Analysis of Exogenous Ambiguity and Cooperation

Traditional approaches to modelling environmental agreement, based on non-cooperative game theory, may be central for explaining the relationship between NDC ambiguity and ambitions, especially over time. Hence, in both our theoretical and empirical analysis we cannot rule out that a country's level of ambiguity has further impacts on the ambitions of others (e.g. reciprocity), thus impacting total ambitions.

To illustrate the potentially detrimental effects of ambiguity on cooperation, we present a simple game-theoretical model based on Barrett's and Danneberg's (2012) analyses of a Threshold Public Goods (TPG) game. As opposed to the regular Public Goods Game, the TPG game features a threshold that represents a minimum level of cooperation. Usually in the games, players experience a loss if they do not provide enough resources to reach the threshold value of total resources, or they gain significant return on their contributions if they do reach it.

In the regular Public Goods game, the players' dominant—but pareto-inefficient—strategy is usually to defer from cooperation since the net marginal returns to investment are neg-

ative. In the TPG game, however, multiple equilibria often exist and several of these are Pareto-efficient. While one equilibrium is mutual deference, the other equilibria are located at the threshold - for instance, collectively reduce emissions to reach a specific target - the existence of which induces large losses if cooperation fails.

Barrett and Dannenberg (2012) show experimentally that communication (a 'treaty') virtually guarantees coordination on or above the threshold in a one-shot game. They also find that uncertainty about impact if the threshold is not reached, has no effect on cooperation as long as the expected value is sufficiently high. Uncertainty in the precise location of the threshold, however, has an adverse effect on coordination: Nearly every group that faced an uncertain threshold failed at reaching the investment threshold they had agreed upon. Barrett (2013) shows that the dramatic fall in success rate occurs because threshold uncertainty transforms the Coordination game into a Cooperation game (Prisoner's dilemma). We illustrate that the same thing can happen if there is uncertainty in the nexus between countries' mitigation efforts and actual mitigation levels.

In the model, N symmetric countries contribute to a public good (mitigation) to avoid a climate catastrophe. Each country, i, have an endowment of W resources and choose their mitigation level, q_i, in order to reach or not reach a mitigation threshold, $\bar{Q}$. In the context of the Paris Agreement, the threshold can represent the collective target of limiting global warming to 2 °C). We denote aggregate mitigation Q, the sum of each country's mitigation level, q_i. There is no return to mitigating; however, in reaching the threshold countries get to keep the remainder of their uninvested endowments. If the threshold is not reached, countries lose X% of their resources not invested in mitigation. Countries gain nothing by overshooting the threshold.

We assume that any one country cannot mitigate sufficiently to reach the mitigation threshold on their own. We also assume that the payoff of contributing at least $q_i = \bar{Q}/N$ is higher than unilaterally deviating at the threshold and experiencing a loss of X% of remaining endowments.

$$W - q_i \geq W * (1 - X) \quad \text{if} \sum q_i \geq \bar{Q} \tag{11}$$

As it stands, this is a coordination game with two sets of equilibria. One equilibrium, suboptimal, entails that every country contributes 0 to the public good. The other set of equilibria is that all countries contribute such that $\sum q_i = \bar{Q}$ and (11) is satisfied.

In the following we separate between investments q_i and the actual mitigation, κ_i, called contributions. Suppose that the investments of $n < N$ countries are realized as

contributions, κ_i, to the public good, according to some probability distribution $F(\kappa; q)$. These countries invest in mitigation technology, but the actual mitigation may be smaller or higher than the investment. In terms of the topic of this paper, one can think of investments as the NDC ambition level, contributions as the actual mitigation level reported in the review phase, and $F(\kappa; q)$ as the ambiguity reflected by NDCs.

To put som structure on $F(.)$, suppose for an investment q_i, that the lower bound of κ_i is $q_{i,min}$ and upper bound $q_{i,max}$ and that $q_{max} - q = q - q_{min}$ equals a constant D for investment $q_i \geq D$. An investment $q_i \leq D$, κ_i is distributed between $[0, 2q]$. The rest of the countries, $N - n$, will make investments in the same way as the N countries above: an investment of q_i is realized as $\kappa = q_i$ with probability 1. The assumptions above imply that the aggregate contribution made by the N countries is a random variable, K, where Q_{min} and Q_{max} are the bounds on the uncertainty interval.

Consider the situation where all N countries invest such that $\sum q_i = \bar{Q}$ and $q_i > D \forall i$. K may take any value in the range $[\bar{Q}_{min} = \bar{Q} - n \cdot D, \bar{Q} + n \cdot D = \bar{Q}_{max}]$. Hence, there is a positive probability of reaching and not reaching the threshold. In effect, this type of uncertainty could turn the game into a Prisoner's Dilemma around the threshold depending on countries' beliefs about $F(K; Q) = Pr(K \geq \bar{Q})$. While investing $\bar{Q}$ was an equilibrium in the certainty case, it may be non-existent in the case with uncertain contributions. At $\bar{Q}$ each country may have an incentive to unilaterally deviate by reducing investments slightly since there is no longer an abrupt impact, X, of doing so. If the following condition is met, country i has an incentive to reduce mitigation by ϵ:

$$\begin{aligned}(W - (q_i - \epsilon))F(K - \epsilon \geq \bar{Q}) + (W - (q_i - \epsilon))(1 - X)(1 - F(K - \epsilon \geq \bar{(}Q))) \geq \\ (W - q_i)F(K \geq \bar{Q}) + (W - q_i)(1 - X)(1 - F(K \geq \bar{Q}))\end{aligned} \tag{12}$$

Where the left side of the inequality is the expected wealth if investment is $q_i - \epsilon$ and the right side q_i. (12) can be simplified to

$$\frac{\epsilon(1 - X)}{X} \geq F(K \geq \bar{Q})(W - q_i) - F(K - \epsilon \geq \bar{Q})(W - (q_i - \epsilon)) \tag{13}$$

The right side expresses the difference in wealth left over after investestments in the two scenarios are made, weighted by the corresponding probabilities of reaching the threshold. Suppose, for instance, that the perceived reduction in the probability of reaching the threshold is sufficiently small such that $F(K \geq \bar{Q})(W - q_i) < F(K - \epsilon \geq \bar{Q})(W - (q_i - \epsilon))$. Then the inequality strictly holds since the left side of the inequality

is always bigger or equal to zero.

I Regression tables with full models

In this section we reproduce Table 1, Table 2, Table C.2 and Table C.5, but we display the entire list of controls selected by the double lasso procedure.

Table I.1: Reproduction of Table 1, displaying all control variables

	No weights		Regression weights		PCA	
Dep Var: NDC Ambition	b/se	b/se	b/se	b/se	b/se	b/se
Ambiguity Index	-0.0684**	-0.0448	-0.610*	-0.315	-0.173	-0.0408
	(0.0306)	(0.0313)	(0.338)	(0.335)	(0.147)	(0.161)
GDP (log)	-0.729***	0.0746	-0.754***	0.215	-0.802***	0.305
	(0.181)	(0.617)	(0.181)	(0.608)	(0.200)	(0.689)
Democracy	1.372***		1.482***		1.547***	
	(0.423)		(0.414)		(0.420)	
Climate change	5.974***	20.89*	5.280**	23.56**	5.256**	25.18**
vulnerability	(2.239)	(11.37)	(2.195)	(11.22)	(2.242)	(12.16)
Coal rents	-0.927***		-1.008***		-0.929***	
	(0.254)		(0.246)		(0.262)	
Oil rents	0.0141		0.0144		0.0148	
	(0.0221)		(0.0203)		(0.0211)	
Natural gas rents	-0.0269	-0.902	-0.0389	-0.520	-0.0511	
	(0.0440)	(1.041)	(0.0431)	(0.994)	(0.0463)	
GDP (log) x Coal rents		-0.119***		-0.104***		-0.114***
		(0.0274)		(0.0225)		(0.0341)
GDP (log) x		0.0842		0.0443		-0.0105**
Natural gas rents		(0.110)		(0.104)		(0.00426)
GDP (log) x Democracy		0.102**		0.109**		0.113**
		(0.0460)		(0.0427)		(0.0437)
GDP (log) x Democracy x		-0.0519				
Natural gas rents x		(0.232)				
Coal rents x						
Oil rents						
GDP (log) x Democracy x		0.00219***		0.00192***		0.00166***
Natural gas rents x		(0.000685)		(0.000701)		(0.000396)
Oil rents						
GDP (log) x Climate change		-1.728		-2.079		-2.254
vulnerability		(1.302)		(1.266)		(1.378)
Democracy x Climate change		7.477***				
vulnerability x		(2.272)				
Natural gas rents x						
Coal rents						
Climate change						0.290
vulnerability x						(0.507)
Coal rents x Oil rents						
Observations	148	148	148	148	148	148
R^2	0.552	0.583	0.546	0.576	0.540	0.572

Note: This table reproduces Table 1 but displays the coefficients of all control variables included. The controls in column 2, 4 and 6 are selected using the double lasso procedure. Robust standard errors are in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels.

Table I.2: Reproduction of Table 2, displaying all control variables

Dep Var: NDC Ambition	No weights b/se	Regression weights b/se	PCA b/se
Ambiguity index	0.0252	0.661	0.480***
	(0.0464)	(0.452)	(0.178)
Democracy	3.639***	0.605	1.033**
	(1.183)	(0.532)	(0.433)
Ambiguity index x	-0.209**	-2.953**	-1.509***
Democracy	(0.104)	(1.249)	(0.369)
GDP (log)	-0.302	0.0405	-0.554
	(0.650)	(0.628)	(0.741)
GDP (log) x	-0.105***	-0.108***	-0.118***
Coal rents	(0.0210)	(0.0229)	(0.0296)
GDP (log) x	0.00108	-0.000177	-0.0121
Oil rents	(0.00190)	(0.00165)	(0.00919)
GDP (log) x	-0.996	-1.685	-0.567
Climate change vulnerability	(1.367)	(1.314)	(1.489)
Climate change vulnerability x	0.832		
Coal rents x	(0.562)		
Natural gas rents			
Natural gas rents	-0.0389	-0.845***	
	(0.0572)	(0.232)	
Climate change vulnerability	14.84	19.79*	10.43
	(11.91)	(11.45)	(12.93)
Climate change vulnerability x		0.129	
Coal rents x Oil rents x		(0.296)	
Natural gas rents			
Climate change vulnerability x		1.707***	
Natural gas rents		(0.465)	
GDP (log) x			0.0301
Climate change vulnerability x			(0.0192)
Oil rents			
Climate change vulnerability x			0.726
Coal rents x Oil rents			(0.547)
Observations	148	148	148
R^2	0.579	0.591	0.606

Note: This table reproduces Table 2 but displays the coefficients of all control variables included. The controls are selected using the double lasso procedure. Robust standard errors are in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels.

Table I.3: Reproduction of Table C.2, displaying all control variables

Dep Var: NDC Ambition	Index (std) b/se	Type of target b/se
Ambiguity measure	0.00716	0.274**
	(0.0265)	(0.122)
Democracy	0.647	1.849***
	(0.505)	(0.632)
Ambiguity measure x Democracy	-0.122**	-0.834***
	(0.0607)	(0.301)
Gdp (log)	0.145	-0.0415
	(0.611)	(0.648)
GDP (log) x Coal rents	-0.117***	-0.330*
	(0.0225)	(0.199)
GDP (log) x Coal rents x Natural gas rents	0.0539	
	(0.0546)	
GDP (log) x Oil rents	-0.000295	-0.000815
	(0.00175)	(0.00198)
GDP (log) x Climate change	-1.813	-1.600
vulnerability	(1.285)	(1.339)
Climate change vulnerability x	-0.142	
Coal rents x Oil rents x	(0.343)	
Natural gas rents		
Climate change vulnerability x	1.581***	0.254
Natural gas rents	(0.455)	(3.083)
Natural gas rents	-0.773***	-0.283
	(0.224)	(1.228)
Climate change vulnerability	21.15*	18.29
	(11.27)	(11.69)
Coal rents x Oil rents x		0.0269
Natural gas rents		(0.261)
Climate change vulnerability x		4.323
Coal rents		(4.685)
Climate change vulnerability x		-0.449
Coal rents x Natural gas rents		(2.041)
Climate change vulnerability x		0.891
Coal rents x Oil rents		(0.644)
Climate change vulnerability x		0.0201
Oil rents x Natural gas rents		(0.0368)
Observations	148	149
R^2	0.597	0.591

Note: This table reproduces Table C.2 but displays the coefficients of all control variables included. The controls are selected using the double lasso procedure. Robust standard errors are in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels.

Table I.4: Reproduction of Table C.5, displaying all control variables

	(1)
Dep Var: NDC increased ambition	b/se
NDC increased information	0.475***
	(0.0751)
GDP (log) x Democracy	0.0132
	(0.0115)
GDP (log) x Climate change vulnerability	-0.340***
	(0.0724)
Climate change vulnerability x	0.000206
Natural gas rents	(0.0200)
Observations	157
R^2	0.351

Note: This table reproduces Table C.5 but displays the coefficients of all control variables included. The controls are selected using the double lasso procedure. Robust standard errors are in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels.