Abstract

In recent decades, democratic countries have signed hundreds of international environmental agreements (IEAs). Most of these are weak and do not include effective enforcement. To study this puzzle, we propose a positive theory of IEAs in which the political incumbents negotiate in the shadow of re-elections. We show that incumbents are prone to negotiate treaties that are simultaneously overambitious and weak. The theory also provides a new perspective for understanding investments in green technologies, and predicts that countries are tempted to rely too much on technology instead of sanctions to make compliance credible. Preliminary evidence is consistent with our predictions.

*For useful comments and discussions we thank Stephen Coate, Eleonora Patachini, Torsten Persson, Ragnar Torvik, and seminar participants at Cornell, U. of Geneva, Harvard, Rice, Stanford GSB, the 2017 NICEP conference in Nottinham, the 2016 EEA meeting, and the 2016 Warwick political economy conference in Venice. We also thank Angela Cools, Torje Hegna and Esteban Mendez Chacon for outstanding research assistance.
1 Introduction

Over the past forty years, environmental issues have achieved increasing prominence in international politics. Both developed and developing countries have signed hundreds of international environmental agreements (IEAs). These agreements have targeted a wide range of goals, from forest preservation, to water management, to the regulation of transboundary pollution. This development has surprised economists. By reducing pollution, IEAs are designed to provide public goods; economic theory therefore suggests that countries should find it optimal to free ride. Why, then, do we see so much cooperation? In a survey on the “Economics of Climate Policy,” Kolstad and Toman [2005] referred to the rise of IEAs as the “Paradox of International Agreements.” A large literature has been devoted to highlighting and explaining this paradox.¹

Two features of IEAs, which have so far attracted little attention, suggest that the paradox should be qualified. The first is the fact that IEAs are typically very weak agreements: they generally do not include effective enforcement or monitoring mechanisms. The lack of enforcement is only partially explained by the lack of third party enforcement in global politics; after all, the countries could sign treaties where noncompliance is met by trade sanctions (as in trade and arms control treaties).² The second striking feature of IEAs is that many of them, including some of the most prominent, are generally seen as ineffective.³ These two facts suggest that the paradox may be that so many countries are negotiating and signing weak agreements, rather than the number of agreements itself. Negotiating treaties is an expensive and laborious process; signing treaties that are either not ratified (as was the case with the U.S. and the Kyoto agreement) or that are ratified and then reneged on (as was the case with Canada and the Kyoto agreement) is even more damaging. We may call this the “Paradox of Weak Agreements.”

In this paper, we present a positive theory of international environmental agreements to study this issue. We argue that, in the presence of reelection concerns, governments are biased toward signing weak agreements that leave the ultimate decision on compliance

¹See, for example, Carraro and Siniscalco [1993], Barrett [1994], Dixit and Olson [2000], and Battaglini and Harstad [2016]. We review this literature more extensively at the end of this section.

²The Montreal protocol of 1997 regulating chlorine emissions damaging the ozone layer, for instance, did indeed permit trade sanctions to be imposed on violators.

on the outcome of future elections. These agreements are characterized by enforcement mechanisms that are less effective than optimal, and that are indeed repudiated with positive probability. Interestingly, this is a general phenomenon that does not depend on the preferences of the incumbent government that negotiates the agreements: relatively “green” and “brown” governments alike are affected by it. It explains the underprovision of international cooperation by rationalizing weak agreements when strong agreements would be optimal. Even more surprisingly, it also explains how electoral concerns may induce governments to join agreements even when no agreement would be optimal. This explains the possibility of oversupply in ineffective agreements.

In our model, a political incumbent in the home country negotiates a treaty with a foreign country (or a group of foreign countries). The agreement is motivated by the fact that a country generates negative externalities on the other. Our mechanism permits—but does not require—there to be a symmetric externality from the foreign country onto the home country. The treaty specifies what the home country ought to do to reduce the externalities, as well as the consequence if it does not. After the negotiation, an election decides whether the incumbent party continues to be in charge or is replaced. At this stage, voters discern which party is best given the country commitments made in the first period: the green party, which has more environmentally friendly preferences than the median voter; or the brown party, which has less environmentally friendly preferences than the median voter. At the last stage of the game, the elected party decides whether or not to comply with the treaty, facing the options negotiated at the first stage of the game. We have a strong treaty if, no matter which party is in power in the following periods, the agreement is enforced. We have a weak treaty if it includes sanctions that are not sufficiently high to guarantee its implementation (and so it may be repudiated if the brown party is elected). We use this simple model to study how electoral incentives shape the type of agreement that is signed (weak vs. strong), the size and scope of the agreement, and the incentives to invest in green technologies.

Regarding the type of agreement, we first show that signing an IEA may or may not be optimal from a social point of view (depending on the preferences and the cost of the environmental policy); however, if the IEA is signed, it should always be strong. Nevertheless, when reelection incentives are sufficiently important, the equilibrium IEA is always weak and thus repudiated with positive probability, regardless of whether the first-period incumbent was green or brown. To understand the intuition behind the results, note that with no agreement or with a strong agreement, the incumbent and the challenger are identical (in

---

4 Examples of these type of negotiations are the protocols signed under the Convention on Long-Range Transboundary Pollution (CLTAP), which attempt to reduce Sulphur and other hazardous emissions with transboundary effects; or those signed under the United Nations Framework Convention on Climate Change (UNFCCC), which commit state parties to reduce greenhouse gas emissions.
this respect) from the median voter’s point of view, because they would behave in the same way after the election: in the first case, because there would be no agreement to implement; in the second, because both of them would implement the agreement. When the treaty is weak, however, the agreement is enforced only if the green party is elected and therefore the median voter’s preferences depend on whether they prefer to comply or face the sanction. The key insight of our analysis is that the median voter’s preferences depend on the details of the agreements: the median voter prefers to comply ex post if the sanction is sufficiently high, and to not comply if it is sufficiently low. Using this insight, we show that both the parties can design a weak agreement that gives them an advantage in the election. The green party designs a weak treaty in which the median voter wants implementation ex post and implementation is guaranteed only if the incumbent is reelected; the brown party designs a treaty in which the median voter does not want implementation and implementation can be avoided only if the incumbent is reelected.

Regarding the size of treaties, we show that electoral incentives induce a novel overshooting effect according to which the incumbent tends to make environmental commitments that, besides being weak as discussed above, are larger than what would be chosen without electoral incentives. This phenomenon, again, is remarkable because it characterizes both green and brown incumbents. As we will explain more extensively in Section 3, this phenomenon occurs because the incumbent, aware of the fact that he is signing a weak treaty, attempts to compensate with size for the fact that the treaty might not be fully complied with.

We also analyze investments in “green” technologies, such as abatement technology or renewable energy sources. The desire for a weak agreement may lead to either underinvestment or overinvestment in green technologies. By reducing the marginal cost of compliance, green technology makes the two parties similar ex post, making it easier for both to comply. This makes green technology similar to sanctions, and so it allows parties to use it to sustain a weak agreement. There is underinvestment when the parties limit the investments in efficient green technology in order to preserve a sufficient difference ex post between the two parties (i.e., to make sure that the brown party does not find it sufficiently easy to comply). We have overinvestment when an inefficient technology is used instead of sanctions: this may occur both when costly technology is necessary to separate oneself from the challenger, and when doing so is preferred to a socially optimal strong treaty enforced by sanctions.

Our work connects and contributes to two strands of literatures: the literature on environmental agreements, and the literature on the political economy of commitments. Traditionally, the first literature has studied the incentives of countries to join environmental agreements in the presence of free riding (Hoel [1992], Carraro and Siniscalco [1993], Barrett [1994], Dixit and Olson [2000], Battaglini and Harstad [2016]).\(^5\) While some recent work has

\(^5\)See Barrett [2003] for an extensive survey of this literature.
highlighted conditions under which large IEAs can be self-enforcing despite free riding problems, most of this literature has highlighted negative results, motivating the view that the rising number of IEAs is a paradoxical phenomenon. Two assumptions have characterized these analyses: first, that countries act as individual agents with no internal politics; and second, that once established, IEAs fully enforce their provisions. Some recent research has endogenized the government’s preferences, or linked the temptation to defect to domestic politics, but we are not aware of any work modeling the decision of weak vs. strong agreements, explaining the popularity of weak agreements, or its implications for welfare. In this paper, we attempt to shift the focus of the literature from simply explaining participation in a self enforcing agreement to analyze the very nature of the agreement. This analysis not only rationalizes the stylized facts mentioned above, but also opens a number of new questions that have not been studied to date.

There is naturally a large literature studying the relationship between international and national policies more generally. In economics, international cooperation has sometimes been viewed as collusion between incumbents, ruining beneficial tax competition (Rogoff [1985]; Kehoe [1989]), while elections allow voters to delegate strategically before policies are set or negotiated (Persson and Tabellini [1995] survey the early literature on such double-edged incentives). In political science, so-called two-level games have been analyzed in which nations negotiate before the treaty must be ratified domestically (Putnam [1988]; Evans et al. [1993]). Putnam also stressed that domestic conflicts between different parties are necessary for international agreements and their ratifications to succeed, since one party, often the minority, can then collude with the foreign country to get a policy implemented which neither of these two would have been able to succeed with alone. We add to this that

Two lines of research have been pursued. First, researchers have studied how voters (or a generic principal) choose the characteristics of the negotiator when bargaining over environmental protection in order to gain a bargaining advantage: see, for instance, Segendorff [1998] Buchholz et al. [2005], Eckert [2003], and Harstad [2008 and 2010]. Since the choice of these characteristics are also important when a treaty is renegotiated, Buisseret and Bernhardt (2017) discuss how a “hostile” incumbent may strategically negotiate a treaty that voters’ prefer to be renegotiated with (such) a strong hand. Second, researchers have studied how lobbying affects government preferences when bargaining for environmental protection with models a la Grossman and Helpman [1994] (see, for instance, Altamiano-Cabrera et al. [2007], Haffoudhi [2005], Dietz et al. [2012]).

A related line of work has been pursued by Fearon [1998a] who has studied arms control agreements as two-step processes in which first a deal is negotiated in a war of attrition, and then it is implemented in a repeated “enforcement game.” Rather than studying the strength of the resulting deals, Fearon focuses on the effect of the time horizon on the length of the negotiations. See also Fearon [1998b] for a general review of the literature on international relations.
even when all domestic parties find the policy costly, the agreement may still be signed—and designed in an inefficient way in order to influence future elections.

Because we study how self-interested governments strategically use IEAs to affect future governments’ behaviors and improve their electoral prospects, our work fits in a long tradition of political economy models studying the strategic role of commitment devices. Persson and Svensson [1989], Alesina and Tabellini [1990] and Aghion and Bolton [1990], for example, have highlighted how public debt can be used in this sense to limit expenditures of future governments; Milesi-Ferretti and Spolaore [1994] and Besley and Coate [1998] study how fiscal policy investments in public infrastructure can be used to affect the outcome of future elections; Biais and Perotti [2002] show how privatization can be used to manipulate the preferences of the median voter; Robinson and Torvik [2005] argue that inefficient local infrastructures often intend to influence elections, and Maggi and Rodriguez-Clare [2007] examine how trade agreements can be used as commitment devices to limit demands from lobbyists. Our contribution to this literature is to analyze how electoral concerns influence and explain the design of international treaties.

The paper is organized as follows. In Section 2 we provide further motivation for the analysis highlighting the role of domestic politics in shaping international negotiations in two recent examples, the Kyoto agreement and Paris Accord. In Section 3 we present a basic version of our model in which treaty and abatement decisions are zero-one variables, and we derive our main results in this simple setting in which the underlying intuition is most transparent. In Section 4 we extend this basic model in three directions: in Section 4.1 we allow for investments in green technology and relate their choice to the strength of the treaty and the choice of sanctions; in Section 4.2 we allow the countries to choose the depth and scope of the negotiation; and finally in Section 4.3 we show how our results on polarization are strengthened once we allow for uncertainty and stochastic compliance costs. In Section 5 we presents a first attempt to test some of the predictions of the theory with a large panel of countries over environmental treaties signed in the past 40 years. We conclude in Section 6. The Appendix presents important proofs, while an online appendix presents other proofs and details the data.

8Antras and Padro i Miquel [2011] analyze how a foreign country may try to influence domestic elections when domestic policies generate international externalities. In our model, it is instead the domestic incumbent that uses the international treaty to influence policies at home.
2 Domestic Politics and International Treaties: Two Examples

Are domestic electoral concerns important factors motivating and shaping international agreements? Evidence on the significance of domestic politics on international relations (and more specifically international agreements) has indeed long been discussed in the international relations literature (see Lantis [2006], Keleman and Vogel [2010], Hovi et al. [2012], for example). To motivate our model and introduce some themes that will prove relevant in the analysis below, we discuss here two recent examples in which the influence of domestic politics has been particularly evident.\(^9\)

**The Kyoto Protocol.** Consider first the case of the United States in the negotiations for the Kyoto Protocol of 1997. Until the final stages of its negotiations, the U.S. delegation was aiming for a modest target (GHG emissions in 2008-2012 equal to the 1990 levels). This reflected a long-standing cautious position taken by the previous administrations and the fact that the delegation expected resistance from the Senate, at the time controlled by the Republican Party. The stance of the U.S. delegation, however, changed abruptly when Vice President Gore took charge of the negotiations (see Hovi et al. [2012]). Gore pushed the delegation toward accepting a much more ambitious target of a 7% decrease in GHG. While this was widely seen as an unrealistic goal,\(^10\) the Clinton administration was looking forward to the incoming 2000 presidential election and congressional races. Lantis [2006:40] observed that “Clinton hoped that Democratic control of the House and Senate or even a Gore presidential victory in 2000 would create a better political climate for ratification.” According to a senior official participating to the negotiation, “Gore, planning to run for president in 2000, anticipated that climate-change policy would become a vote getting issue.”\(^11\) He therefore pre-positioned himself to take advantage of the negotiations, pushing for

---

\(^9\)We focus on recent IEAs to illustrate the importance of politics in the negotiations only for brevity. Just focusing on the U.S. experience, there is ample evidence on the effect of electoral incentives on policy makers’ decisions concerning international environmental commitments. See, for example, Hopgood [1998] for an in-depth discussion of the political calculus in the Nixon administration regarding the United Nations Conference on the Human Environment (UNCHE) held in Stockholm in 1972, and in the Bush administration regarding the United Nations Conference in the Environment and Development (the “Earth Summit”) held in Rio in 1992.

\(^10\)Bang et al. [2012:759] noted that “This target left little doubt that Kyoto would be unacceptable to the Senate.” Indeed, a few months after its proposal the Senate unanimously passed a resolution against it, the Byrd-Hagel resolution.

\(^11\)See Hovi et al. [2012:144]. Based on anonymous interviews with 26 participants in the negotiations.
an agreement that could be expected to be ratified only if he was elected to the presidency: a behavior that is in line with the logic of the model discussed in the introduction. Two features of this agreement appear are worth noting in light of the model that we will present below. First, the agreement pushed by Vice President Gore was overly ambitious given the political realities and it involved a fair amount of posturing. Second, it was weak and without explicit sanction. Shortly after the presidential election that brought the Republican George W. Bush to power, plans to ratify the agreement were abandoned. It reasonable to assume that this would not have happened if Gore had been elected, especially if the election came with a change in majority in the Senate.

A similar dynamic can be found in Canada, where the incumbent negotiating the agreements was also—in the terminology used above—a “green party”. In Canada, the Kyoto Agreement was signed and ratified by the liberal government of Jean Chretien, who committed his country to an ambitious reduction plan (6% reduction of GHG by 2012 from 1990 levels) but, notably, without making an attempt to generate domestic support for the treaty (Lantis [2006]). As noted by Lantis [2006:36], “Chretien rested on his political advantages rather than assuaging the concerns of his opponents.” This behavior appears consistent with an attempt to link the success of the treaty to the endurance of liberal governments. Indeed, as soon as the conservative prime minister Stephen Harper took office in 2006, a policy of deliberate indifference was pursued causing a sharp increase in GHG emissions. Canada invoked its withdrawal clause from the Kyoto agreement in 2011; see Austen [2011]. In the years since the agreement, Canadian emissions have risen by more than 30% above the 1990 target (Walsh [2011]).

The experience with the Kyoto agreement shows that incentives to sign weak agreements do not pertain only to left-leaning incumbent governments. In Japan, Australia and New Zealand, for example, the governments responsible for the negotiations were all supported by conservative parties unsympathetic to environmental issues (in the terminology of the model. “brown parties”). Despite this, all these countries signed the Kyoto protocol, although in weak forms that did not eventually survive. Ratifications of the signed agreements followed a pattern similar to the logic of the model. In Japan and Australia, the signature of the agreement was followed by conservative administrations that delayed or watered down its content as much as possible.12 In New Zealand, signature of the agreement was followed from the U.S. and Europe, Hovi et al. concluded that one of the most plausible reason for the failure at Kyoto was that the Clinton-Gore administration “essentially pushed for an agreement that would provide them a climate-friendly face.”

12The agreement was signed in Australia in 1998, but the conservative government of John Howard delayed ratification until the end of the mandate. Howard’s government also managed to negotiate extraordinarily lax targets that allowed emissions in GHG to increase by as much as 8% from the 1990 levels (Hamilton
in 1999 by the election of a “green party” that managed to ratify the agreement in 2002. The agreement however survived only for the period in which the Labor party remained in charge, and it was abandoned only in 2012 when the government shifted back to the National party, the very party that negotiated it.\(^{13}\)

**The Paris Accord.** While it is early to evaluate the success of the Paris Accord, it is clear that decisions surrounding this agreement were influenced by electoral considerations in the U.S.. Signed by the Obama Administration just one year before the 2016 Presidential elections, its ratification and implementation was debated in the presidential campaigns. Along with the negotiations, the Obama administration had committed to various measures incentivizing investments in green technologies: by attempting to reduce emissions from power plants using the regulatory power provided by the Clean Air Act; by tightening fuel economy standards for heavy-duty vehicles; and by developing standards to address methane emissions from landfills and the oil and gas sector.\(^{14}\) Our theory predicts that these investments should be sufficient to commit a Democratic candidate, but not a Republican. It is indeed the case that, after the election, the Republican President elect pledged “to rip up Paris Climate Agreement” (Sarlin [2016]) while the Democratic candidate vowed to uphold the U.S. commitment to climate actions signed by the Obama administration (Cohan [2016]). On August 4, 2017, the US state department submitted a notification to the UN that the administration intended to withdraw from the Paris climate Agreement.

\(^{13}\)The Kyoto agreement was officially ratified only in December 2007 after the Labor Party (with Kevin Rood as prime minister) assumed government control. A similar path has been followed by Japan, where the negotiating party in 1997 was the conservative Liberal Democratic Party party (LDP), which signed and ratified the Kyoto agreement. The agreement was not renegotiated in 2010, when the government repudiated the mandatory targets and opted for new voluntary targets. Despite watering down targets for cutting emissions by 2020, in 2013 Japan met its Kyoto Protocol obligations to lower greenhouse gas emissions only by buying carbon credits as actual emissions rose (Reuters [2013]).

\(^{14}\)New Zealand’s conservative Government announced in 2012 that it would not agree to the legally binding second Kyoto Protocol commitment period (Small [2012]). However, it said it would make a pledge to voluntarily reduce greenhouse gas emissions under the parallel “United Nation Convention Framework.” See the “Intended Nationally Determined Contribution” (INDC) submitted to the UN: http://newsroom.unfccc.int/unfccc-newsroom/united-states-submits-its-climate-action-plan-ahead-of-2015-paris-agreement/#downloads. Accessed on October 10th, 2016.
3 The Basic Model and Result

3.1 The Model

We begin our analysis by considering a basic model with two periods and two sets countries, the home country $H$ and the foreign country $F$ (alternatively, $F$ can be the set of other countries). Country $H$’s action, or "emission," generates an externality $e \geq 0$ on the foreign country. Country $H$, however, can abate pollution and eliminate the externality by incurring a cost. Although abatement may have some value also to $H$, we assume that the net cost of abating is positive for all citizens, so everyone in $H$ prefers to emit as long as there is no treaty. Section 4 generalizes the model and allows abatement to be nonbinary, among other things.

The two countries can negotiate a treaty in which $H$ is required to abate. To motivate compliance, the treaty also specifies some consequence or sanction that $F$ commits to impose on $H$ if $H$ does not comply.\footnote{In Section 4, we extend the model to allow the countries to negotiate on the possibility of investment in green technologies (that reduce the cost of compliance), and on the size of the abatement project (that allow abatement to be incomplete). In that section, we show that the theory generalizes to environments in which “sanctions” are exogenous (as in the case in which they comprise only a reputational cost) or even nonexistent.} The cost of the sanction to $H$ is $s \geq 0$ and $F$’s cost of imposing the sanction is $gs$. If $g > 0$, $F$ dislikes imposing the sanction (as, for example, when $s$ is imposed by restricting trade with $F$). If $g < 0$, $F$ benefits from imposing the sanction, perhaps because it takes the form of a monetary transfer.\footnote{Naturally, if the sanction is a pure monetary transfer, then we should expect $g = -1$.} We assume that $g \geq -1$, so that there is a deadweight loss $(1 + g) s \geq 0$ when the sanction is imposed.

Both when negotiating the treaty, and when deciding whether to comply, the home country’s decisions are made by one of two political parties. Parties and voters differ only in their perceived net cost of abatement: This cost is $c_G > 0$ for the party that is relatively "green," while it is $c_B > c_M$ for the political party that is relatively "brown." The cost for the median voter, $M$, is in between: $c_M \in (c_G, c_B)$.

The timing of the game is as follows. First, in period one, $F$ and $H$’s incumbent government $i \in \{B, G\}$ negotiate $s$. Second, an election determines whether the incumbent remains in power or is replaced. Third, in the second period, the winner of the election decides whether to comply or face the sanction $s$.

We will now explain each step in turn.

1. We make two important assumptions on the negotiations in period one. First, we assume that the two parties can use side transfers when negotiating the treaty. This implies...
that the equilibrium level of $s$ will simply be the $s$ that maximizes the two negotiators’ sum of expected payoffs. An advantage of this assumption is that, with side transfers, it is irrelevant whether there is also a symmetric problem where $F$ emits, harming $H$. As long as $F$ and $H$ can negotiate using side transfers, the two problems can be separated and thus can be considered independently. Second, we assume that $H$ and $F$ are fully committed to imposing the sanction if $H$ does not comply. The fact that countries can commit to a system of sanctions is demonstrated, for example, by the Montreal protocol. Countries may also be able to commit for reputational reasons, although we do not formalize the reasons for this commitment here. Section 4.1 proves that investments in technology is one way of facilitating commitment, while Section 4.4 argues that our results continue to hold also if $s$ can be renegotiated.

2. After the treaty has been negotiated, there is an election. The outcome of the election is determined by the median voter, $M$, who votes for the candidate delivering the highest expected payoff. Specifically, $M$ reelects the first-period incumbent $i \in \{B, G\}$ if $u^i_M - u^{-i}_M > \delta$, where $u^i_M$ (resp. $u^{-i}_M$) is $M$’s expected payoff if electing $i$ (resp. $-i$), while $\delta$ is some relative popularity shock in favor of the challenger $-i \in \{B, G\} \setminus i$. The popularity shock, realized after the treaty is signed, can refer to the importance of other policy differences, not explicitly modeled here. We assume $\delta$ to be uniformly distributed on $[-z/\sigma, (1-z)/\sigma]$, implying both that the density of the shock is $\sigma$, and that the incumbent wins with probability $z \geq 1/2$ if $u^i_M = u^{-i}_M$. The incumbency advantage is therefore measured by $z - 1/2 \geq 0$. We start by assuming that the support of the shock is sufficiently large so that reelection probabilities are interior in $(0,1)$. As will be shown below, this property is guaranteed if the density of the shock is so small that:

$$\sigma < \min \left\{ \frac{1-z}{c_B - c_M}, \frac{1-z}{c_M - c_G} \right\}. \quad (1)$$

Section 4.3 allows the popularity shock distribution to be arbitrary, and it allows the abatement costs to be stochastic.

3. At the final stage of the game, the newly elected policymaker $j \in \{B, G\}$ decides whether to comply with the treaty. By comparing the two costs, the second-period incumbent finds it optimal to comply if and only if the sanction $s$ is larger than the cost to $j$, $c_j$. If $s > \bar{s} \equiv c_B > c_G$, both of the parties will comply with the treaty, so we have what we call a strong treaty. If instead $s < \underline{s} \equiv c_G < c_B$, none of the parties will comply with the treaty, so we have an ineffective treaty. If $s \in [\underline{s}, \bar{s}]$, the treaty will be complied with if the second-period incumbent is $G$, but not if $B$ is in power. Since this treaty may or may not be complied with, we name it a weak treaty.\(^{18}\)

\(^{17}\)See Article 4 of the Protocol and, for a more extensive discussion, Barrett [2003].

\(^{18}\)Note that we assume that $G$ complies when indifferent, while $B$ does not comply when indifferent (i.e.,
Modulo the transfers that can be exchanged at the bargaining stage, the payoffs are in line with the discussion above. If $H$ complies, $F$ receives $e > 0$ while $i \in \{B, M, G\}$ pays the compliance cost $c_i > 0$. If $H$ does not comply, $F$ imposes the sanction at cost $g_s$, where $s > 0$ measures the cost for every individual in $H$. In addition, the second-period incumbent $j \in \{B, G\}$ enjoys the office rent $R \geq 0$ as the benefit of staying in office. (A similar office rent for the first period is sunk and would not influence the analysis.) The proofs in the Appendix allow the office rent to be conditioned on the identity of the second-period incumbent $j$, and the online appendix permits permits the office rent to be conditioned on whether $j$ complies. These contingencies do not influence the basic result and they are thus abstracted from here.

### 3.2 The Optimal Treaty

It is useful to start by describing a couple of relevant benchmarks. The first benchmark is the socially optimal solution, which we define as the allocation that maximizes the sum of payoffs for $F$ and the median voter in $H$. Obviously, it is optimal for $F$ and $H$ to commit to abatement if $e > c_M$, while it is optimal for $F$ and $H$ to abate if $e < c_M$. This outcome can be implemented if $H$ and $F$ sign a strong treaty when $e > c_M$ and no treaty otherwise. Note that a weak treaty is always dominated, and it is strictly dominated if $e \neq c_M$.

As a second benchmark, suppose the first-period incumbent $i \in \{B, G\}$ takes as exogenous the probability that the green party $G$ wins, $p_i$. In this situation, $i$ and $F$ jointly prefer that the second-period incumbent complies if $e > c_i$, but not if $e < c_i$. When the former condition holds, $i$ and $F$ sign a strong treaty. Otherwise, no treaty will be signed. Again, a weak treaty is always dominated.

We can summarize these considerations with the following result:

**Proposition 0.** In both the benchmark cases described above, a weak treaty is dominated:

(i) The optimal outcome is implemented by a strong treaty if $e > c_M$, and by no treaty if $e < c_M$.

(ii) If the first-period incumbent $i \in \{B, G\}$ takes $p_i$ as given, then $H$ and $F$ sign a strong treaty if $e > c_i$, and no treaty if $e < c_i$.

These tie-breaking rules just ensure that the set of weak treaties is closed, and it has no qualitative implication for the results.

---

19If, for example, the incumbent were a strong dictator, then we may have $p_i = 1$. Moreover, in the probabilistic voting model of democracy described above, we have $p_i = z$ (if $i = G$) or $p_i = 1 - z$ (if $i = B$) when $\sigma \to 0$, since the popularity shock will then dictate the electoral outcome.
3.3 The Equilibrium Treaty

Of course, the above benchmarks are for illustration only, since the probability of staying in power is endogenous in the model presented above and politicians do care about being in office. The next result shows that the endogeneity of the reelection probability changes the outcome dramatically if the office rent is sufficiently large. To shorten notation, we use $z_i = z$ if $i = G$, and $z_i = 1 - z$ otherwise. Thus, $z_i$ is the probability that $G$ wins when the first-period incumbent is $i$ and $u_i^M = u_M^i$.

**Proposition 1.** Let the first-period incumbent be $i \in \{B, G\}$:

(i) Suppose $R > R_i^*$, where $R_i^*$ is defined by:

$$R_i^*(e) = \begin{cases} 
\frac{(1+g)(1-z+\sigma c_M-c_i))(e-c_i)+(1+g)c_i)}{\sigma |c_M-c_i|} & \text{if } e \leq c_i, \\
\frac{(1-z_i+\sigma (c_M-c_i))(e-c_i)+(1+g)c_i)}{\sigma |c_M-c_i|} & \text{if } e > c_i.
\end{cases}$$

$H$ and $F$ always sign a treaty, and the treaty is always weak. Specifically, a brown first-period incumbent signs a treaty with sanction $s = s$, while a green first-period incumbent signs a treaty with sanction $s = \bar{s}$.

(ii) If $R < R_i^*$, $H$ and $F$ sign a strong treaty when $e > c_i$, and no treaty when $e < c_i$.

Figure 1 illustrates the type of treaty as a function of $R$ and $e$. While Proposition 1 is proven in the Appendix, it is instructive to outline the explanation for why it holds. At the election stage, the median voter anticipates that $u_M^G = u_M^B$ if the treaty is strong.
or ineffective, since then any second-period incumbent will take the same action regarding abatement. If the treaty is weak, however, the benefit of electing $G$ instead of $B$ is:

$$u^G_M - u^B_M = s - c_M \text{ for } s \in [\bar{s}, \overline{s}] .$$

Thus, $u^G_M - u^B_M > 0$ when $s \in (c_M, \overline{s}]$, and $u^G_M - u^B_M < 0$ if $s \in [\bar{s}, c_M)$, implying that the ex post benefit of the treaty for the median voter depends on $s$. Since an incumbent $i \in \{B, G\}$ is reelected if and only if $u^i_M - u^{-i}_M > \delta$, and $\delta$ is uniformly distributed on $[-z/\sigma, (1 - z)/\sigma]$, it follows that with a first-period incumbent $i$, $G$ is elected with probability:

$$p_i(s) = \begin{cases} 
    z_i & \text{if } s < \bar{s} \\
    z_i + \sigma (s - c_M) & \text{if } s \in [\bar{s}, \overline{s}] \\
    z_i & \text{if } s > \overline{s}
\end{cases} \quad (2)$$

Note that $p_i(s)$ is increasing in $s$ for $s \in [\bar{s}, \overline{s}]$ and $p_i(\overline{s}) = z_i + \sigma (c_B - c_M) > z_i$, so the probability that $G$ wins is maximized when $s = \overline{s}$ (see the left part of Figure 2). For such a large sanction, the median voter agrees with $G$ that it is better to comply, and he rationally expects that party $B$ will not comply. When the office rent is sufficiently large, the electoral gain is sufficiently important to compensate a green first-period incumbent for the possibility that the agreement is repudiated by the brown party if elected. In this case, the optimal $s$ is equal to $\overline{s}$. Intuitively, the green party wants to have the highest penalty consistent with a weak agreement in which $B$ alone would not comply; this is the best way to reduce the appeal of the brown party for the electorate, and thus maximize the reelection probability.

The case with a $B$ incumbent is surprisingly similar. In this case, the probability that $B$ is reelected, $1 - p_B(s)$, is declining in $s$ and maximized at $s = \bar{s}$ where we have: $1 - p_B(s) = z + \sigma (c_M - c_G) > z$, as shown in the left part of Figure 2. With such a small sanction, the median voter agrees ex post with $B$ that the cost of complying is too large, so it is better to get out of the agreement. Once again, if the office rent is sufficiently large, the preference for reelection trumps any other concern, and a weak treaty is signed, as shown in the right part of Figure 2.

In either case, both incumbents maximize the reelection probability by signing some kind of weak treaty. The weak treaty separates the incumbent from the challenger, while a strong or an ineffective treaty makes the two parties identical from the voter’s point of view.

Observe that $R^*_i(e)$ is a positive threshold, decreasing for $e \leq c_i$, increasing for $e > c_i$, reaching a minimum at $e = c_i$:

$$R_i = R^*_i(c_i) = \frac{(1 - z_i + \sigma (c_M - c_{-i})) (1 + g) c_{-i}}{\sigma |c_M - c_{-i}|}$$

as illustrated in Figure 1. Three factors therefore determine the region where weak agreements prevail. The first is the variance in the popularity shock. If $\sigma$ is small, the popularity
shock is likely to dictate the outcome of the election. Thus $R_i^*(e)$ increases when $\sigma$ falls, and a weak treaty is less likely for any given $R$. A weak treaty is signed only when $\sigma$ is large and the voters are substantially influenced by the payoffs they can expect. If $\sigma$ is so large that (1) is violated, then an incumbent can be reelected with probability one by strategically signing a weak treaty. Since this situation seems empirically unrealistic, we rule it out by assuming that (1) holds.\(^\text{20}\)

A second factor affecting the type of agreement is the deadweight cost of a sanction, $g$. As $g$ decreases, $R_i^*$ shifts downward uniformly, making the region in which weak agreements prevail larger. The presence of a distortionary sanction makes it more likely that a strong treaty is signed, since only then can we guarantee that no sanction will be imposed.

The third factor is the ideological bias of the opposition party with respect to the median voter, as measured by $|c_i - c_M|$ when $i$ is the incumbent. As this bias increases, the minimal point of $R_i^*$ at $c_i$ shifts downward, and $R_i^*$ becomes flatter both on the right and the left of $c_i$, so $R_i^*$ shifts down for any $e$. Intuitively, the larger the discrepancy between the opposition and the median voter, the more the incumbent can take advantage of it with a weak agreement.

The traditional literature on IEAs emphasizing free riding, as discussed in the Introduction, predicts that there is insufficient participation in IEAs. Proposition 1, in contrast, shows that in a political economy setting two phenomena may occur. When $e < c_M$, it is optimal with no agreement, but both parties will sign a weak agreement in equilibrium if just $R$ is large. Therefore, there can be an oversupply of IEAs. When $e > c_M$, on the contrary, it is optimal with a strong agreement. In equilibrium, however, there will be a weak agreement if $R$ is large. The problem here is not a lack of participation, but the quality

\(^{20}\)The historical examples in Section 2 justify the assumption that electoral incentives matter for the incumbent when negotiating an IEA (i.e., that $\sigma$ appears to be sufficiently high).
of the IEA. Both of these predictions appear to be consistent with the historical experience with IEAs, as discussed in the Introduction and, more extensively, in Section 5.

4 Negotiating Technology, Depth, and Risks

International treaties include many components in addition to sanctions. In fact, large portions of negotiations focus on aspects that we have deliberately ignored in the previous section, including the depth and scope of the treaty, the magnitude of the emission cuts, the number of industrial sectors that are to be regulated, and policy measures on green technologies. Furthermore, a country’s political or economic cost of abatement may vary with the business cycle, and it can be unknown at the negotiation stage.

These observations raise two sets of questions. First, does the fact that we rarely see sanctions mean that agreements are, in the terminology of the previous section, “ineffective” rather than “weak,” or can we have weak agreements (with all their strategic implications) even without sanctions? Second, how does the political economy mechanism above influence the equilibrium depth and choice of technology?

This section generalizes the basic model above to investigate the robustness of the basic result. The generalizations also allow us to address the two questions. Section 4.1 lets the compliance cost be endogenous and depending on how much the home country invests in green technology (for instance, renewable energy). In Section 4.2, where emission levels are nonbinary, the compliance cost depends on the size or the depth of the treaty, and we let this depth be endogenous. In both these cases, weak agreements naturally emerge even when the cost of failing to comply with an agreement (i.e., what is above referred to as the sanction) is exogenous and small/zero. In Section 4.3, when the compliance cost is stochastic, we show that the preferences for the treaty design remains polarized, even when the two parties’ compliance costs converge. In Section 4.4, we informally discuss how one could allow for renegotiation. While the basic result in Proposition 1 continues to hold, each extension sheds new light on the political economy of treaties.

4.1 Green Technologies and Compliance

Assume that the home country can invest in an abatement technology \( y \in [0, Y] \) at a cost \( qy \geq 0 \) as part of the negotiation. With investment \( y \), the abatement cost is reduced to \( c_i - y \) for all types \( i \in \{G, M, B\} \).\(^{21}\) We start by considering the situation where \( s \) is fixed,

\(^{21}\)It is natural to assume that, as \( y \) increases, the marginal benefit of the investment decreases. In this case, the green investment reduces the abatement cost to \( c_i - \phi(y) \) for some concave function \( \phi \). We assume
before letting both $y$ and $s$ be negotiated.

**Exogenous sanctions (or no sanctions at all).** We start by introducing two assumptions that are both relaxed below. First, to let technology be important, suppose an exogenous sanction satisfies $s < c_G$ and $Y + s > c_B$. The first condition ensures that with no green investment, we have an ineffective agreement with no compliance; the second ensures that with a sufficiently large investment, we have a strong agreement with full compliance. Second, suppose $q < 1$, so that the investment cost is smaller than the return. Then, signing an environmental agreement and complying is optimal for $F$ and the median voter in $H$ if and only if $e > c_M - (1 - q)Y$. If this condition holds, the first-best treaty is strong, i.e. it is never optimal to leave any uncertainty about compliance.

By reducing the cost of compliance, the green technology has two effects: first, obviously, a direct effect on welfare as it makes the agreement cheaper when implemented; but, secondly, a strategic effect determining when the agreement is implemented. A very high level of investment in green technology makes compliance optimal for both $B$ and $G$; similarly, a very low investment green technology makes compliance suboptimal for both $G$ and $B$. Incumbents may prefer to make compliance dependent on the winner, since that can boost their reelection probabilities, as described in Section 2. They can achieve this goal if:

$$c_G - s \equiv y \leq y \leq \overline{y} \equiv c_B - s.$$  

The first inequality guarantees that $G$ will comply with the treaty, and the second inequality guarantees that $B$ will not. By choosing $y = \overline{y} \equiv c_B - s$, a green incumbent achieves two goals: he ensures that compliance will be achieved if $G$ is reelected; and he ensures that this damages the reelection probability of $B$. To see the second point, note that $c_M < c_B$, so when $y = \overline{y}$, we have $s + y - c_M > 0$, implying that the median voter prefers compliance ex post, and that the probability that $G$ is reelected is maximized at $p_G^* \equiv z + \sigma (c_B - c_M)$.

Similarly, a B-incumbent can improve his electoral prospects by choosing $y = \underline{y} \equiv c_G - s$.

This level of investment guarantees that only party $G$ complies ex post, and that the median voter is more likely to prefer $B$, who does not comply. In fact, this level of technology minimizes the probability that $G$ is reelected and the probability becomes $p_B^* \equiv 1 - z - \sigma (c_M - c_G)$.

The following result characterizes the equilibrium with fixed exogenous $s$ when the green technology is efficient (i.e., $q < 1$):

**Proposition 2.** Let the first-period incumbent be $i \in \{B, G\}$ and $q < 1$. There exist thresholds $R_i^*$ such that:

above a linear $\phi$ only for simplicity; the results of this sections can be extended to allow for decreasing marginal returns of investments.
(i) If \( R > R^*_i \), the treaty is always weak, and only \( G \) will comply. If \( i = B \), investments are \( y = c_G - s \), while if \( i = G \), investments are \( \overline{y} = c_B - s \).

(ii) If \( R < R^*_i \), the treaty is never weak. If \( e > c_i - Y (1 - q) \), a strong treaty with \( y = Y \) is signed; otherwise no treaty is signed and \( y = 0 \).

The intuition for this result is similar to the intuition of Proposition 1. Politicians behave in the same way under a strong agreement and under no agreement, but they act differently once elected if the agreement is weak. If it is important to win office, there is a level of investment such that the green party will comply, the brown will not, and the median voter will prefer to stick with the incumbent regardless of the incumbent’s preferences. If the investment level is large, the median voter prefers compliance and party \( G \); if the investment level is low, the median voter is more likely to prefer party \( B \). If the office rent is sufficiently large, the electoral concerns outweigh other concerns, a weak treaty is always signed, and \( y \in \{ y, \overline{y} \} \).

There are three interesting implications of this proposition. First, we have a weak agreement even if the countries have no commitment power to impose sanctions (i.e., \( s = 0 \)). This occurs because the green investment is chosen by design to change the parties’ preferences.

Second, we can have a novel crowding-out effect of sanctions. Consider an increase in the exogenous cost of sanctions \( s \) that makes it more onerous for \( H \) to not comply.\(^{22}\) If \( R > R^*_i \), an increase in \( s \) does not translate into an increase in compliance when green investments are endogenous. To see this, note that if \( G \) is the incumbent, he chooses \( y = \overline{y} \) such that \( s + \overline{y} - c_B = 0 \): an increase in \( s \) will just reduce \( y \) with no effect on compliance. Similarly, if \( B \) is the incumbent, he chooses \( y = \underline{y} \) such that \( s + \underline{y} - c_G = 0 \): once more, an increase in \( s \) will just reduce \( y \) with no effect on compliance. In both cases, an increase in \( s \) has no impact whatsoever on the strength of the agreement.\(^{23}\)

Third, we have underinvestment in green technologies when a weak treaty is signed. This occurs because the incumbent does not want to make compliance a dominant strategy for

\(^{22}\) An example of this change is the recent design of the Paris Accord of 2015 that does not explicitly include monetary sanctions or enforcement agencies, but relies on the fact that the countries will not want to suffer “reputational costs” by missing the targets. The implementation of the accord is supposed to strengthen these costs by instituting a “name and shame” mechanism that exposes noncompliant countries, and the policy thus corresponds to an increase in \( s \) in our model.

\(^{23}\) An increase in \( s \) can influence the type of the treaty only if \( R \) is close to the thresholds \( R^*_i \) in Proposition 2. In this case, it becomes more costly to stick with a weak treaty when the sanctions are larger. If \( e > c_i - \max \{(1 - q) \overline{y}, Y (1 - q)\} \), a larger \( s \) makes it more likely that we move to a setting with a strong treaty. However, if \( e < c_i - \max \{(1 - q) \overline{y}, Y (1 - q)\} \), a larger \( s \) makes it more likely that we move to a setting with no treaty.
everyone, so he restricts technological investment despite its efficiency. This result, however, crucially depends on the fact that we have assumed green technologies are efficient. We will return to this aspect below, where we allow for investments in inefficient technologies and allow for both $s$ and $y$ to be endogenous.

**Endogenous sanctions and green investments.** In the previous section, we assumed $s$ to be exogenous. As a result, we could not study the equilibrium choice between sanctions and technology and its full implications for underinvestment or overinvestment. The next result characterizes the decision to adopt green technology in a political equilibrium in which both sanctions and green investments are endogenous.

**Proposition 3.** Let the first-period incumbent be $i \in \{B, G\}$. The equilibrium choice of IEAs is characterized by a threshold $R^*_i > 0$ such that:

(i) If $R > R^*_i$, $F$ and $H$ sign a weak treaty, and it is complied with at probability $p^*_i$. Specifically, $y = 0$ and $s = c_{-i}$ if $q \geq q^*_i \equiv 1 + g - gp^*_i$; and $s = 0$ and $y = c_{-i}$ if $q < q^*_i$.

(ii) If $R < R^*_i$, then $y = s = 0$ and no agreement is signed if $e < c_i - \max\{0, (1 - q)Y\}$; while otherwise $F$ and $H$ sign a strong agreement with $y = Y$ if $q < 1$, but $y = 0$ and $s > c_B$ if $q > 1$.

The proof and the definition of $R^*_i$ is in the Appendix. When $R$ is sufficiently small (i.e. $R < R^*_i$), electoral incentives are not sufficiently strong to lead to a weak agreement. In this case we either have no agreement or a strong agreement, as in Proposition 1. The possibility of green investments affects this decision only because it affects the cost of compliance. If $q > 1$, the investment is inefficient, the minimal investment $y = 0$ is chosen, and the final cost of compliance remains $c_i$. In this case, we have the strong agreement if and only if $e > c_i$. If $q < 1$, the investment is efficient, the maximal investment $y = Y$ is chosen and the cost of compliance is $c_i - (1 - q)Y$. In this case, we have a strong agreement if and only if $e > c_i - (1 - q)Y$.

The results change when electoral incentives are sufficiently strong to make a weak agreement optimal (i.e. $R \geq R^*_i$). In this case, two scenarios are possible, depending on whether $g < 0$, as when the sanction benefits $F$ (e.g., $H$ makes a transfer to $F$), or $g > 0$, so that the sanction hurts both $H$ and $F$ (e.g., when sanctions include trade restrictions). In the first case, inefficient technologies are never adopted; however, we may have underinvestment since not even an efficient technology is adopted if $q \in (q^*_G, 1)$. In the second case, an efficient technology is always adopted; but now we may have overinvestment since an inefficient level of investment is chosen when $q \in (1, q^*_i)$.

Interestingly, the brown party is the party that is more prone to invest in green technologies. To see this, note that $q^*_G < q^*_B$, so if the green party invests, then the brown party also finds it optimal to invest, but when $q \in (q^*_G, q^*_B)$, then only the brown party will invest.
The intuition behind these findings is as follows. Similarly to the analysis of the previous section, when \( R \) is large, the \( G \)-incumbent’s objective function is increasing in \( s + y \) in the region in which the agreement is weak, and the opposite is true for \( B \). In equilibrium we have a corner solution: either we have \( s + y = c_B \), if \( G \) is the incumbent, or \( s + y = c_G \), if \( B \) is the incumbent. This makes \( s \) and \( y \) strategic substitutes in weak agreements: an increase (resp., decrease) in \( y \) must be compensated by a reduction (resp., increase) in \( s \). So either we have sanctions or investments. If the treaty is complied with (and the technology is used) with probability \( p_i^* \), the net cost of investing is \( q - p_i^* \), which will be compared to the expected social cost of a unit of the sanction, \( (1 + g)(1 - p_i^*) \). By comparing the two, it is clear that partial compliance is better ensured by technology if \( q - p_i^* < (1 + g)(1 - p_i^*) \Rightarrow q < q_i^* \equiv 1 + g - gp_i^* \). Since a treaty negotiated by \( B \) is less likely to be complied with (since \( p_B^* < p_G^* \)), \( B \) is more likely to prefer (partial) compliance by technology than by sanctions compared to \( G \) when \( g > 0 \). A higher marginal cost of sanctions induces any incumbent to rely less on them and more on technology as a way to induce (partial) compliance.

Consistent with this prediction, right-wing republicans in the US have often been in favor of supporting green innovation and technology, while left-wing democrats have more often supported traditional abatement policies. In his 2008 speech on climate change, President George W. Busch said that "The right way [to address climate change] is to adopt policies that spur investments in the new technologies needed..." Even when President Trump notified the UN that the US would withdraw from the Paris Agreement, the White House stated: "We will continue to reduce our greenhouse gas emissions through innovation and technology breakthroughs."25

4.2 The Depth of the Treaty

Assume now that the home country’s level of abatement expenditure is a continuous variable, \( x \in [0, \infty) \). As before, different stakeholders in the home country disagree on the net benefit of such a policy. Thus, suppose the perceived net cost is \( c_jx \) for \( j \in \{B, G, M\} \), where \( c_G < c_M < c_B \), as before. To the foreign country, the benefit of these abatement expenditures is represented by the increasing and concave function \( e(x) \). The concavity assumption captures the fact that, as the size of the abatement expenditure increases, less and less efficient abatement opportunities are employed, inducing decreasing marginal returns to the expenditures, as measured by \( x \). The optimal level for \( F \) and the median voter in \( H \) is clearly to set \( x \) such that \( e'(x) = c_M \). We interpret \( x \) as the treaty’s size, scope, or depth.

24 The incumbents’ objective functions are qualitatively similar to the objective functions illustrated in Figure 2, with the only difference being that the horizontal axis is \( s + y \).

When both depth and the level of sanctions are negotiated, a treaty is defined by the associated target level of abatement \( x^* \) and sanction \( s_{x^*} : [0, x^*] \to \mathbb{R}_+ \) specifying a penalty \( s_{x^*}(x) \geq 0 \) for each abatement level \( x < x^* \). Just as before, the sanction can be either beneficial or costly for \( F \): the cost of imposing \( s \) is \( gs \) for \( F \), so the total social cost per sanction unit is \( 1 + g \geq 0 \).

Given the treaty, as represented by the target \( x^* \) and the function \( s_{x^*}(x) \), the second-period policymaker \( j \in \{ B, G \} \) prefers an abatement level that minimizes the total costs:

\[
x^*_j = \arg \min_x c_j x + s_{x^*}(x).
\]

Note that in equilibrium \( H \) and \( F \) always prefer to sign a treaty in which at least the green party fully complies with the treaty, so \( x^*_G = x^* \). In general, however, \( x^*_B \leq x^* \) so we can write \( x^*_B = x^*_G - \Delta^*_s \) for some \( \Delta^*_s \geq 0 \). We can therefore have two type of treaties: We have a strong treaty when \( \Delta^*_s = 0 \). In this case compliance is complete and the parties look equally good to the voters. For a strong treaty, it is necessary that the sanction is so large that any deviation is unattractive for every party. We have a weak treaty, instead, when \( \Delta^*_s > 0 \). In this case, abatement is contingent on the identity of the winner of the election. This is similar to what we found in the previous section. Now, however, instead of solely the dichotomy a weak vs. strong treaty, we have different degrees of weaknesses: the larger the value of \( \Delta^*_s \), the weaker the treaty.

For a weak treaty, where the two parties make different choices, we must have \( s_{x^*} \left( x^*_B \right) \in \left[ c_G \Delta^*_s, c_B \Delta^*_s \right] \), or \( S^*_s \in [c_G, c_B] \), where \( S^*_s \) is the average sanction per "unit of deviation":

\[
S^*_s \equiv \frac{s_{x^*} \left( x^*_B \right)}{\Delta^*_s}.
\]

The average sanction \( S^*_s \) relates to the median voter’s attitude toward \( B \): if \( S^*_s \in [c_G, c_M] \) the median voter likes the fact that \( B \) does not fully comply and prefers \( B \) to \( G \); if \( S^*_s \in [c_M, c_B] \) the median voter wants full compliance and prefers \( G \) to \( B \).

The next result provides a complete characterization of the equilibrium treaty with endogenous depth and sanction. We use starred superscripts to denote the equilibrium, and subscripts to denote the identity of the first-period incumbent negotiating the treaty.\(^{28}\)

---

\(^{26}\)It can be easily shown that weak agreements emerge in the case in which the emission level is endogenous but the total sanction \( s \) is exogenous. We omit a detailed discussion of this case here for brevity.  

\(^{27}\)To see this, suppose that \( x^*_G \leq x^* \). Then no matter who is elected, a positive sanction will be paid. By reducing \( x^* \) to \( x^*_G \), incumbent \( s \) can reduce the expected sanction by \( s \left( x^* \right) - s \left( x^*_G \right) \) without changing the probability of winning since it increases the utility provided by both parties by the same amount.  

\(^{28}\)Thus, when \( i \) is the first-period incumbent, \( x^*_i = x^*_G \) is the equilibrium size of the treaty, \( \Delta^*_s = \Delta^*_s^i \) is the equilibrium abatement gap, and \( S^*_i = S^*_s \) is the equilibrium average sanction.
guarantee interior solutions when \( x \) is continuous, condition (1) for the binary case should be strengthened to condition \( \sigma < \sigma \), where the threshold \( \sigma \) is derived and presented in the Appendix.

**Proposition 4.** Let the first-period incumbent \( i \in \{B, G\} \) negotiate the treaty, summarized as \( (x_i^*, \Delta_i^*, S_i^*) \) and suppose \( \sigma < \sigma \). In equilibrium, a green second-period incumbent complies in full by abating \( x_i^* \), a brown second-period incumbent abates \( x_i^* - \Delta_i^* \in [0, x_i^*] \), and the sanction satisfies \( S_i^* = c_i \) when \( \Delta_i^* > 0 \). We have two possible cases, which refer on the following thresholds:

\[
\hat{R}_G \equiv \frac{(1 - z)(1 + g) c_B}{\sigma (c_B - c_M)} \quad \text{and} \quad \hat{R}_B \equiv \frac{z(1 + g) c_G}{\sigma (c_M - c_G)}.
\]

(i) If \( R < \hat{R}_i \), the treaty is strong in that \( \Delta_i^* = 0 \), and the size is \( x_i^{**} \), defined as \( e'(x_i^{**}) \equiv c_i \).

(ii) If \( R > \hat{R}_i \), the size \( x_i^* \) is larger than if \( R \leq \hat{R}_i \), but the treaty is weak and \( x_i^* - \Delta_i^* < x_i^{**} < x_i^{**} \).

Similarly to the analysis of Section 3, the first-period incumbent is motivated to negotiate a weak treaty by the prospect of sufficiently large office rents. In the previous analysis, a weak agreement differed from a strong agreement only because it was associated with a positive probability of noncompliance; now we can instead distinguish two new phenomena.

The first phenomenon is the fact that the weakness of the agreement manifests itself as partial compliance, i.e. \( \Delta_i^* \in (0, x_i^*) \), for any \( R > \hat{R}_i \). This effect is explained by an intuition analogous to the intuition behind the weakness in the previous section. When \( \Delta_i^* = 0 \), the parties will behave identically in office, so the incumbent is reelected simply with probability \( z \). By choosing a weak treaty with \( \Delta_i^* > 0 \), the incumbent can improve his reelection probability by negotiating an appropriate sanction. The green party will choose a sanction sufficiently high so that the median voter but not the brown party wants to comply; the brown party will choose a sanction sufficiently small so that the green party but not the median voter wants to comply.

The second phenomenon is the **overshooting effect**. For \( R > \hat{R}_i \) the politically motivated incumbent \( i \) signs a treaty that is larger than the treaty that the same incumbent would have signed in the absence of electoral incentives, i.e. \( x_i^* > x_i^{**} \). This effect can be explained as follows. By an appropriate choice of the penalty \( S_i^* \), the incumbent can decouple the issue of the size of the treaty (i.e. \( x_i^* \)) from the issue of its strength (i.e. \( \Delta_i^* \)). Once the agreement is signed, what matters for the electoral competition is not \( x_i^* \), but \( \Delta_i^* \) and \( S_i^* \): that is, the difference in ex post behavior between the parties and its consequence. This implies that, given \( S_i^* \) and \( \Delta_i^* \), the incumbent can choose the “second best” depth that maximizes
If the office rent $R$ is large, then the treaty size $x$ is larger, but the treaty is also weaker.

his expected utility. In a strong agreement, the optimal size is $x_i^{**}$, the level at which the marginal benefit equals the marginal cost: $e'(x_i^{**}) = c_i$. Given the uncertainty of a weak treaty, it is optimal that the size is such that the expected marginal externality for $F$ equals the marginal cost for the first-period incumbent:

$$Ee' = p_i e'(x_i^*) + (1 - p_i) e'(x_i^* - \Delta_i^*) = c_i$$  \(5\)

Since party $B$ will not fully comply, the size must be larger so that the expected compliance stays at the right level. This implies that $G$ must abate more than the first-best level, and the size of the treaty is thus also larger than the first-best size. Formally, (5) implies that, when $\Delta_i^* > 0$, we have $e'(x_i^*) < c_i$, so $x_i^* > x_i^{**}$. Figure 3 illustrates all this.

The following result shows how the two effects described above evolve when we change the size of electoral incentives.

**Proposition 5.** Let the first-period incumbent $i \in \{B, G\}$ negotiate the treaty, summarized as $(x_i^*, \Delta_i^*, S_i^*)$.

(i) If $R > \hat{R}_i$ increases, then $\Delta_i^* > 0$ increases, the size $x_B^*$ increases monotonically and $x_B^* - \Delta_B^* \rightarrow e_i^{e-1}(c_B)$, while $x_G^* - \Delta_G^*$ declines monotonically and $x_G^* \rightarrow e_i^{e-1}(c_G)$.

(ii) If $R$ is sufficiently large, then we have $x_B^* > x_G^*$, unless preferences are so polarized that $\sigma(c_M - c_G)(x_G^{**} - x_B^{**}) > 1 - z$.

The main message of this result is that, as election incentives increase, so does the gap between what is promised by the incumbent (i.e., $x_i^*$) and what is actually done if the brown party wins the election; in other words, the potential for “disappointment” over the treaty implementation increases in $R$. This phenomenon, however, is not only due to the fact that the brown party chooses a low abatement level in absolute terms ex post if elected; it is also
driven by the fact that the incumbent, green or brown, becomes increasingly (and partly unrealistically) ambitious as $R$ increases.

To further understand part (i), note that if $R$ is very large, $\Delta_G^*$ is also very large and this increases the probability of getting reelected. When the first-period incumbent is $G$ and $p^*_G$ approaches one, $x_G^*$ must decline toward $x_G^{**}$ to satisfy (5). The intuition is that when it becomes almost certain that $G$ will win the election, then only $x_G^*$ is of importance and $x_G^*$ should be set optimally. The distortion that is necessary for the weak treaty (and the large $\Delta_G^*$) is better ensured by increasing $B$’s deviation $\Delta_G^*$, since $B$ is unlikely to be elected in any case.

The argument is similar when instead the first-period incumbent is $B$. When $R$ and $\Delta_B^*$ grow and $B$ becomes certain to stay in power, $x_B^* - \Delta_B^*$ should approach the optimal level, $x_B^{**}$. The large $\Delta_B^*$ is then better ensured by letting the promised level $x_B^*$ grow, while $x_B^* - \Delta_B^*$ stays close to $B$’s preferred level. The treaty is in any case unlikely to be fully complied with. For a sufficiently large $R$, the promised level $x_B^*$ is thus larger if the first-period incumbent is $B$, even though the expected abatement level is smaller. This is stated in part (ii) of the proposition, and, as specified, this possibility requires that the preferences are not too polarized and that the incumbency advantage is not too large.\footnote{The reason for this condition is that if the incumbency advantage is very large, then $B$ does not need to raise $\Delta_B^*$ and $x_B^*$ by very much in order to win for sure, and then $x_B^*$ will never need to increase to a level that is larger than $x_G^*$.}

### 4.3 Uncertainty and Robust Polarization

A strong assumption in the analysis presented above is that the parties’ cost and preference parameters are known in advance. With complete information on these parameters, a “weak” treaty implies that party $B$ never complies while party $G$ always complies. The outcome is not so clear-cut if the parameters are not fully known in the first period. For example, in a severe recession the actual cost of complying may be so large that none of the parties would like to do so; the reverse situation may occur in a boom. This section allows the cost of compliance to be a stochastic variable. While this extension shows that the basic insights from the simple model continue to hold, it also allows us to strengthen the results and derive new insights. In particular, we show that the two parties’ preferred types of weak treaties remain very different, even if the preferences converge.

To isolate the effects of this extension, let us return to the basic model of Section 3, without technology and with binary abatement levels. Assume that the net cost is $c_j = \hat{c}_j + c$, where $\hat{c}_j$ is a constant individual component for $j \in \{B, M, G\}$, while $c$ is a stochastic common variable distributed according to the cdf $F$ and pdf $f$. If the realization of $c$ is...
large, everyone’s \( c_j \) is large, although we always have that \( c_G < c_M < c_B \). It is realistic to maintain the ranking of preferences between the parties, in our view.

When \( s \) is the sanction, party \( j \in \{B, G\} \) complies with probability \( \rho_j (s) \), given by:

\[
\rho_j (s) = \Pr (c + \hat{c}_j < s) = F (s - \hat{c}_j).
\]

It follows that \( G \) is more likely to comply than \( B \) for any given \( s \): \( \rho_B (s) < \rho_G (s) \).

It is reasonable that \( c \) is unknown at the election stage as well as in the first period. We therefore assume that \( c \) is realized just before the second-period incumbent decides whether or not to comply. We also assume that \( f \) has the typical bell-shape. Thus, \( f \) is convex up to the inflection point \( \hat{c}_I \), where \( f'' (\hat{c}_I) = 0 \), \( f \) is concave for \( c \in (\hat{c}_I, \pi_I) \), where \( \hat{c}_I \) is the second inflection point at which \( f'' (\hat{c}_I) = 0 \), and \( f \) is convex for \( c > \hat{c}_I \).

At the election stage, the median voter understands that the election matters only if \( c \) happens to fall between the two parties’ thresholds, \( c \in (s - \hat{c}_B, s - \hat{c}_G) \). In expectations, the additional utility the median voter expects by electing \( G \) instead of \( B \) is:

\[
E (\Delta u_M) = \int_{s - \hat{c}_B}^{s - \hat{c}_G} (s - \hat{c}_M - c) dF (c).
\]

We continue to assume that the median voter elects party \( G \) if the additional expected utility for the median voter, \( E (\Delta u_M) \), is larger than some random popularity parameter favoring party \( B \). However, rather than requiring the popularity shock to be uniformly distributed, as above, we now allow it to be arbitrarily distributed according to some cdf \( H_i \), where \( i \) is the incumbent. The probability that \( G \) wins the election is then \( p_i (s) = H_i (E (\Delta u_M)) \).

To see how \( s \) influences \( p_i (s) \), note that \( \partial p_i (s) / \partial s > 0 \) if and only if \( \partial E (\Delta u_M) / \partial s > 0 \), regardless of \( H_i \). Furthermore, it is easy to show that \( \partial E (\Delta u_M) / \partial s > 0 \) if and only if:

\[
\int_{s - \hat{c}_B}^{s - \hat{c}_G} f (c) dc > \left( \frac{\hat{c}_B - \hat{c}_M}{\hat{c}_B - \hat{c}_G} \right) f (s - \hat{c}_B) + \left( \frac{\hat{c}_M - \hat{c}_G}{\hat{c}_B - \hat{c}_G} \right) f (s - \hat{c}_G)
\]

(6)

The left-hand side is the average density of the shock \( c \) over the interval in which \( G \) and \( B \) disagree on the policy. On the right-hand side, we have a (weighted) average of the levels that \( f \) takes at the two thresholds. The two weights are equal if the median voter is equally likely to agree with either candidate (i.e., if \( \hat{c}_M = (\hat{c}_B + \hat{c}_G) / 2 \)). Then, the inequality holds, and \( p_i \) increases in \( s \), if and only if \( f \) is (on average) concave over the disagreement interval. If \( f \) is (on average) convex over the disagreement interval, \( p_i \) decreases in \( s \). Since \( f \) is convex at the tails, this explains why \( p_i \) decreases in \( s \) to \( s_B^* \) before \( p_i \) increases to the peak when \( s = s_G^* > s_B^* \). Figure 4 illustrates the disagreement intervals and the equilibrium sanction levels.
Figure 4: Incumbent B is more likely to win if the median voter expects a high cost c, conditional on c falling within the disagreement interval. This implies that B prefers $f'$ to be large over the disagreement interval. Analogously, party G prefers $f'$ to be small and negative over the disagreement interval.

**Proposition 6.** Suppose the compliance cost is stochastic.

(i) There is a unique and finite $s_B^*$ minimizing $p_i(s)$, and there is a unique and finite $s_G^*$ maximizing $p_i(s)$. Both $s_B^*$ and $s_G^*$ are independent of $i$.

(ii) We have $s_B^* < s_G^*$, and $\rho_B(s_B^*) \in (0, \frac{1}{2})$ while $\rho_G(s_G^*) \in (\frac{1}{2}, 1)$.

Part (i) states that the $s_i^*$ that maximizes party $i$’s chance of winning is independent of the identity of the incumbent. Part (ii) says that the sanction level maximizing the chance that B wins is always smaller than the sanction level maximizing the chance that G wins. Part (ii) also says that, at these sanctions, party B would be more likely to not comply than to comply, while G would be more likely to comply than not comply.

Just as in the other extensions discussed above, our main result (from Proposition 1) is not only confirmed, but the extension also sheds new light on the equilibrium treaty.

**Proposition 7.** Suppose $\hat{c}_M = \frac{(\hat{c}_B + \hat{c}_G)}{2}$.

(i) Each disagreement interval covers an inflection point:

\[
\begin{align*}
  s_B^* - \hat{c}_B &< \frac{c}{2} \text{ and } s_B^* - \hat{c}_G, \\
  s_G^* - \hat{c}_B &< \frac{c}{2} \text{ and } s_G^* - \hat{c}_G.
\end{align*}
\]

(ii) Consequently, if $|\hat{c}_B - \hat{c}_G| \to 0$, then $s_B^* \to \frac{c}{2} + \hat{c}_M$ and $s_G^* \to \frac{c}{2} + \hat{c}_M$.

Proposition 7 states that the two thresholds are always close to (and the disagreement interval includes) an inflection point of $f$. This is intuitive, since $f$ is at its steepest at the inflection points. When $f$ is steep, there is a large difference in the probabilities that the median voter $M$ will disagree with $G$ and that $M$ will disagree with $B$. Party $B$ thus prefers to have the thresholds close to the point at which $f'$ is at the largest, while party $G$ prefers a sanction such that the thresholds are close the point at which $-f'$ is at the largest.
Consequently, if the objective is to win the election, the two parties continue to prefer very
different versions of the weak treaty even if their preferences are similar: the two policies \( s_B^* \) and \( s_G^* \) do not converge even if the parties’ preferences converge.\(^{30}\)

## 4.4 Other Extensions

Our basic model is simple and can be used as a workhorse for several other extensions. It
is beyond the scope of this paper to discuss them all, but three extensions are informally
discussed in the following.

**Renegotiating the treaty.** So far, we have made the assumption that country \( F \) commits
to impose the sanction on \( H \), if \( H \) does not comply. This assumption is useful but not
necessary for our main results. After all, Section 4.1 proved that the basic insight of our
model continued to hold if the countries did not negotiate sanctions but instead technologies
that were sunk (and thus committed to) before the compliance stage. In addition, one may
argue that our results hold even with sanctions that can be renegotiated. To see this, consider
the situation where \( H \) has failed to comply and \( F \) is ready to impose the sanction on \( H \). If
\( F \) has the upper hand in the subsequent renegotiation, then \( F \) may propose to \( H \) to drop
imposing the sanction in return for some other favors that could benefit \( F \). If this favor has
the cost \( \xi \geq 0 \) to \( H \) and the benefit \( \gamma \xi \geq 0 \) to \( F \), then \( H \) is willing to accept \( F \)’s offer for any
favor if size \( \xi \leq s \), and thus \( F \) proposes \( \xi = s \) and benefits \( \gamma s \). With this, all our formulae
above hold if just \( g \) is replaced by \(-\gamma \). The assumption \( g \geq -1 \) implies \( \gamma \leq 1 \), meaning that
the favor cannot be more beneficial to \( F \) than it is costly to \( H \). (If \( \gamma > 1 \), one would think
that the favor would have already been negotiated in another agreement).

**Salience of the treaty.** One may question whether compliance to international treaties
are sufficiently high on the political agenda to influence elections. If other policy differences
are much more important, then the popularity of these differences will dictate the election
outcome. This possibility is captured in our model by letting the popularity shock be drawn
from a large support (in that \( \sigma \) is small). In line with this intuition, the above equations do
imply that a weak treaty is less likely when \( \sigma \) is small.

At the same time, our main point is not that treaties will influence elections, but that the
prospects of elections will influence how treaties are designed. That is, if the environmental
policy/treaty is not very important compared to other political issues, then distorting the
policy/treaty may also have a small cost compared to the gains. There are two ways of

---

\(^{30}\)However, if uncertainty vanishes and \( f \) concentrates on a single value for \( c \), the two inflection points
converge. In this case, \( s_B^* \) and \( s_G^* \) will also converge when the preferences converge. This explains why the
sanction levels will converge in the basic model in Section 2, where we assumed that \( c \) was known in advance.
making this argument formally: First, note that for any given \( \sigma \), regardless of how small, the equilibrium treaty is always weak if just \( R \) is sufficiently large. Second, if the environmental policy/treaty is relatively unimportant, in that \( \sigma \) is small, then the compliance cost is arguably also relatively small. Thus, we may write \( \sigma = \epsilon \tilde{\sigma} \) and \( c_i = \epsilon \tilde{c} \), and reduce the salience of the issue by reducing \( \epsilon \). However, parameter \( \epsilon \) will cancel from the above formulae for \( R_i \) and \( \tilde{R}_i \), for example, implying that these thresholds are invariant in the salience \( \epsilon \).

**The political system.** We have emphasized above that while nondemocracies may be characterized by Proposition 0 (where \( p_i \) could be fixed/large), the weak treaties predicted by Proposition 1 are more likely for democratic countries, since democratic leaders are more accountable to the voters. The larger is the effect of utility on the probability for staying in power, the larger is parameter \( \sigma \), and thus the more likely it is that the equilibrium treaty is weak. Following this line of reasoning, one may also argue that the importance of utilities (\( \sigma \)) and the policy makers’ office rent (\( R \)) may systematically vary across political/electoral systems. For example, the office rent may be larger in presidential systems than in parliamentary systems in which power is shared among a larger number of legislators. Similarly, in (majoritarian) winner-takes-all electoral systems, the winner of the election may keep more of the office rent. If this translates into a larger \( R \), such systems should be more likely to sign weak treaties, according to our results. At the same time, electoral competition may be less intense if there are several electoral districts (as is typically the case in majoritarian electoral systems), particularly if gerrymandering has made the electoral outcomes predictable in many districts. Less competition can be translated into a smaller \( \sigma \) in our model, and that effect may reverse or cancel the effect of a larger \( R \). The combination of these arguments suggests that political systems can have nontrivial effects on the design of treaties, and that further research is necessary to sort out the effects in detail.

5 A First Take at the Data

While the extensions in Section 4 provide additional results, they also confirm that the basic results of Proposition 1 are quite robust. That basic model provides two predictions worth further discussion. (Taking the additional predictions of Propositions 2-6 to the data must await further research.) First, Proposition 1 suggests that democracies, i.e. countries in which governments have electoral concerns, are more likely to sign IEAs than nondemocracies. Indeed, when electoral concerns, as measured by \( R \), are sufficiently high, an agreement is always signed, even when it would not be efficient to do it. On the contrary, a policy maker who does not need to be reelected, i.e. an autocrat, signs an agreement only if \( \epsilon > c_A \), where \( c_A \) is the autocrat’s cost (see part (ii) in Proposition 0).
Second, Propositions 1 predicts the agreements signed by democracies to be weak, and thus less effective in reducing emissions. With a sufficiently high $R$, a democracy always signs an agreement, but the agreement is always weak. In an autocracy an agreement may or may not be signed, but it is strong if it is signed.

In this section, we present a preliminary quantitative evaluation of these predictions using a large panel data set on post-World War II environmental treaties.

Table 1 examines whether, as predicted by Propositions 0-1, democracies are more prone to sign international agreements. To this goal we have collected a data set of 151 countries on the major environmental treaties signed from 1976 to 2001. To select the treaties we refer to the list in Appendix 6.1 from Barrett [2003]. The data set includes 31 agreements. We estimate a logit model in which the dependent variable is a dummy variable equal to one if a country signs a treaty during the first five years that an agreement is open for signature and zero otherwise. The independent variables corresponds to characteristics of the country during the first year that the agreement was open for signature. Our key independent variable is a measurement of democracy.$^{31}$ We use two alternative measurement variables for democracy: $polity2_t$ from the Polity IV Project, which measures the country’s degree of democratization, for columns 1-4; or a dummy variable $democracy_t$, which is equal to one if and only if $polity2_t$ is larger than 0, for columns 5-8.$^{32}$ We consider alternative sets of control variables. Specifically, we include a set of geographical dummies, a variable qualifying the electoral regime and, importantly, country or treaty fixed effects to capture different types of unobservable factors. As can be seen from Table 1, in all specifications $polity2_t$ and $democracy_t$ appear positive and significant, suggesting that democratic regimes are indeed more prone to signing international environmental agreements even after controlling for other relevant characteristics. This finding provides support for our first theoretical prediction that regimes with larger electoral concerns are more prone to sign IEAs. This result is corroborated by previous empirical works that have also highlighted the fact that democracies are more prone to sign IEAs (see, for example, Congleton [1992], Midlarsky [1998] and Neumayer [2002]). The results in Table 1 extend these previous results by exploiting a more extensive data set and a larger set of controls.$^{33}$

---

$^{31}$The list of treaties and the description of the data sources for Table 1 and 2 is presented in the online appendix.

$^{32}$For the Polity IV Project see http://www.systemicpeace.org/polity/polity4.htm. To assess if a country is democratic we construct the $democracy$ variable following Persson and Tabellini [2006] and Besley et al. [2011].

The finding that democracies sign more IEAs is perhaps not surprising; the prediction that democracies are more prone to sign weak and less effective agreements appears more controversial. Is there evidence supporting it? As mentioned in the Introduction, there is certainly clear evidence that many IEAs signed or ratified by democracies are weak. The U.S., for instance, signed 11 agreements between 1989 and 2011, all of which have failed to achieve ratification (Bang et al. [2012]).

The specific question of whether democracies are better at dealing with environmental issues has been addressed by a large literature (see, for instance, Congleton [1992], Barrett and Graddy [2000]), Murdoch, J. T. Sandlerb and W. Vijverberga [2003]. Perhaps unsurprisingly, however, given the endogeneity of the political regime and the number of potential omitted variables affecting both the democratic regime and the environmental outcome, this literature has obtained mixed results.

In Table 2, we investigate the marginal effect of signing an agreement on reductions in CO₂ (the leading greenhouse gas). More importantly, we also examine how the political regime affects the marginal benefit of signing another treaty. To this goal, we have collected a large panel of 143 countries over 7 environmental treaties that belong to the Convention On Long-Range Transboundary Air Pollution lineage, which aims to control CO₂ or indirectly induce CO₂ reductions. The data cover the period 1960-2011. The dependent variable in Table 2 is the (log of) the level of CO₂ emissions per year (in kilotons). The target independent variables are as follows. First, #treaties₁₋₁ reports the number of greenhouse emission treaties on CO₂ emissions signed by a country up to period t – 1. Second, polity₂ₜ and democracyₜ measure democracy at t as described above. Third, and most importantly, we have interaction effects polity₂ₜ · #treaties₁₋₁ and demᵯ · #treaties₁₋₁.

Columns 1-4 report simple OLS estimates with various regional, economic, and institutional controls. By a superficial read, results here appear mixed, both in terms of the effect

Convention on the Prior Informed Consent Procedure for Certain Hazardous Chemicals and Pesticides in International Trade (the Rotterdam Convention), the Copenhagen Amendment to the Montreal Protocol, and the Cartagena Protocol on Biosafety to the Convention on Biological Diversity (Cartagena Protocol on Biosafety). Our dataset contains thirty-one agreements, and except for the Copenhagen Amendment, all the previous treaties are included.

Ratification of the Paris Accord of 2015 is an open political question. The administration claims it does not need a Senate vote since it sees the accord as an “executive agreement”, not a formal treaty.

For this analysis, we follow Slechten and Verardi [2014] who previously studied the effectiveness of treaties using CO₂ emissions. Slechten and Verardi [2014], however, did not study the effect of political institutions on the effect of treaties, which is the variable of interest for our work.

To select the treaties with effects on CO₂, we have followed Slechten and Verardi [2014]. The list of treaties is presented in the online appendix. As we show in the online appendix, the analysis is, however, robust to using the more comprehensive list used in Table 1.
of the number of treaties and in terms of democracy: \( \#\text{treaties}_{t-1} \) is significant at the 1% level in specifications 3-4; \( \text{polity2}_t \) and \( \text{democracy}_t \) are not significant; and, more importantly for us, the interaction effects are negative. These results would suggest that treaties have a larger effect on \( \text{CO}_2 \) when democracies sign them, a result that is in conflict with our previous findings. It is however the case that \( \text{democracy}_t \) and \( \text{polity2}_t \) are correlated with a number of other important variables that can determine the success of a treaty: the presence of a civil society, the history of the country, and the quality of the judicial system. Without controlling for these dependencies, we obtain only spurious results.

To control for these and other country specific variables, we perform the regression analysis with country fixed effects in columns 5-8. Results are then qualitatively very different and support our theoretical predictions: The variable \( \#\text{treaties}_{t-1} \) is now highly significant in all specifications; \( \text{polity2}_t \) and \( \text{democracy}_t \) remain insignificant and small; but the interaction effects are now positive and very significant. This suggests that treaties indeed have an impact on greenhouse gas (GHG) emissions, but signing a treaty has a smaller impact on \( \text{CO}_2 \) reductions for democratic regimes relative to other regimes—exactly as the theory would predict.

6 Conclusions

This paper sheds light on the connections between domestic and international politics. International treaties influence, and perhaps even limit, what domestic policymakers can do. The incentives provided by a treaty may affect different political candidates in different ways, and thus they might also influence domestic elections. Anticipating this, political incumbents will seek to negotiate and sign treaties strategically and in a way that both ties the hands of the next policymaker and improves the odds of staying in office. Our theory is built to deepen our understanding of these trade-offs and it results in a number of testable predictions.

First, political incumbents will be reluctant to sign “strong” treaties with which their countries must necessarily comply. A strong treaty will level the playing field since any future politician will behave in the same way. A “weak” treaty, in contrast, may or may not be upheld. A relatively green party is more likely to comply with the treaty than a relatively brown party, and the median voter’s preferred choice will depend on the negotiated consequence—or sanction—facing a country that does not comply. With a small sanction, the median voter prefers the brown party that does not comply; but with a somewhat larger sanction, he prefers that the green party is in power. Thus, some kind of weak treaty can maximize the reelection probability regardless of the identity of the incumbent.

Second, we show that treaties may also be too large in scope or depth. The explanation is that when the incumbent prefers a weak treaty that may not be fully complied with, there is
an “overshooting” effect that makes the treaty very large. Depth is helpful to the incumbent because the expected marginal externality to the foreign country can then stay at the right level, even when the treaty may not be fully complied with.

Third, countries might in equilibrium invest more in technology than what the first best would require. The reason is that, since a weak treaty may or may not be upheld, there is a fair chance of facing the sanction and the deadweight loss this involves. This deadweight loss can be avoided if one instead invests in technologies that raise the motivation to comply with the treaty. In this way, the probability of compliance may be increased to a moderate level (characterizing a weak treaty) without risking the deadweight loss that comes with sanctions.

To summarize, our theory predicts that political incumbents prefer treaties too often, and benefit from treaties that are too weak, too broad in scope, and are (partially) enforced by technology investments. These results hold even when a strong treaty enforced by sanctions is first best. The incumbents’ preferences are particularly strong when the perks from staying in office are large and there are many swing-voters who pay attention to the policy.

These predictions fit well with the preliminary evidence discussed in Section 5: democratic countries are more likely than others to sign international treaties, existing treaties are surprisingly weak, and treaties are enforced less by explicit sanctions than by countries’ investments in complementary technology. Our analysis has resulted in a large number of other testable predictions as well, and future research should aim to take the theory more carefully to the data.

Future research may also develop the theory in reasonable directions. To illustrate the results in a simple and intuitive way, we have limited attention to a three-stage model with only two sets of countries and two political candidates. We have also abstracted from asymmetric information and alternative ways in which the treaty may interact with domestic politics. However, our model is tractable enough to be used as a workhorse in analyzing a wide range of such extensions. And these extensions will be immensely important; in our view, the political economy of treaties must be better understood before we can successfully address the global challenges ahead.
7 Appendix

7.1 Proof of Proposition 1

The countries will reach an agreement that maximizes the surplus of the ruling parties in the two countries. Let \( U_i(s) \) be the utility generated in the domestic country for the incumbent \( i \) and \( U_F(s) \) for the incumbent in the foreign country. When the incumbent is \( i \), the equilibrium agreement \( s_i \) solves:

\[
\max_s \{U_i(s) + U_F(s)\}
\]  

(7)

Consider how the objective function \( W^i(s) = U_i(s) + U_F(s) \) changes with \( s \). There are two cases to consider: when the incumbent is a green party, and when it is a brown party. In the main text, we assumed that both candidates have the same office rent \( R \); in the following, for additional generality, we allow the office rents to be different for the two candidates: \( R_j \) for \( j = G, B \).

Case 1: The green party is the incumbent

If both \( G \) and \( B \) comply at \( t = 2 \), the objective function in (7) is: \( W^G_{BG}(s) = zR_G - c_G + e \).

If \( G \) only complies at \( t = 2 \):

\[
W^G_{G}(s) = p_G(s)(R_G - c_G + e) - (1 - p_G(s))(1 + g) s.
\]  

(8)

If there is no agreement or if there is an agreement and \( s < \bar{s} \):

\[
W^G_{0}(s) = zR_G - (1 + g) s.
\]

Note that, since \( p_G(s) \) increases in \( s \), \( W^G_{G}(s) \) is convex in \( s \). Using this fact and the formulas above, we have:

**Lemma 1.1.** The green party signs an agreement if \( e > e^*_G(R_G) \) with \( e^*_G(R_G) \) a nonnegative and nonincreasing function of \( R_G \).

**Proof.** The case with no agreement cannot occur if \( W^G_0(0) < W^G_{BG}(s) \) or if \( W^G_0(0) < W^G_G(s) \). Consider the first case first. The condition \( W^G_0(0) < W^G_{BG}(s) \) can be written as:

\[
zR_G - c_G + e = W^G_{BG}(s) > W^G_0(s) = zR_G \Rightarrow e > c_G.
\]

Consider now the second condition. Since \( W^G_G(s) \) is convex in \( s \) we have two cases: \( s = \bar{s} = c_B \) and \( s = \underline{s} = c_G \). We now show that it is never optimal to set \( s = \underline{s} = c_G \), since in this case it is better to have \( s \geq c_B \). With (2), we have \( W^G_G(s) > W^G_{BG}(s) \) only if:

\[
W^G_G(s) = (z + \sigma (\underline{s} - c_M))(R_G - c_G + e) - (1 - z - \sigma (\underline{s} - c_M))(1 + g) \underline{s} > zR_G - c_G + e.
\]
Since \( \bar{s} = c_G \), this condition holds only if:

\[
\sigma (c_G - c_M) R_G > (1 - z - \sigma (c_G - c_M)) (e + g c_G).
\]

But since \( c_G - c_M < 0 \) and \( e > c_G \), the previous inequality is impossible.

We must therefore have that, when the agreement is weak, \( s = \bar{s} = c_B \).\(^{37}\) Such an IEA is preferred to no IEA if:

\[
W^G_G(\bar{s}) = \left( (z + \sigma (\bar{s} - c_M))(R_G - c_G + e) - (1 - z - \sigma (\bar{s} - c_M))(1 + g) \bar{s} \right) > z R_G = W^G_\emptyset(0).
\]

So:

\[
[\sigma (c_B - c_M) R_G + [z + \sigma (c_B - c_M)] ((1 + g)c_B - c_G + e) - (1 + g) c_B] > 0.
\]

This is true if:

\[
e > \bar{e}^*_G(R_G) = \frac{(1 + g) c_B - (z + \sigma (c_B - c_M))((1 + g)c_B - c_G) - \sigma (c_B - c_M) R_G}{z + \sigma (c_B - c_M)}.
\]

where, we note, \( e^*_G(R_G) \) is decreasing in \( R_G \). Putting together the two conditions we have that party \( G \) chooses to sign an IEA if \( e > e^*_G(R_G) = \text{Min}\{c_G, \bar{e}^*_G(R_G)\} \). \( \blacksquare \)

We now prove the following result:

**Lemma 1.2.** There is a threshold \( e^*_G(R_G) \geq e^*_G(R_G) \) such that the green party finds it optimal to sign a weak agreement if \( e \in (e^*_G(R_G), e^*_G(R_G)) \), and a strong agreement if \( e > e^*_G(R_G) \).

**Proof.** Consider the green party first. For \( e < e^*_G(R_G) \) we have \( W^G_BG(s) < W^G_\emptyset(s) \) and \( W^G_G(s) < W^G_\emptyset(s) \), so no agreement is signed. For \( e \geq e^*_G(R_G) \), a strong agreement is signed if \( W^G_BG(s) < W^G_BG(s) \), that is:

\[
((z + \sigma (s - c_M))(R_G - c_G + e) - (1 - z - \sigma (s - c_M))(1 + g)s) < z R_G - c_G + e,
\]

where \( s = c_B \). This implies:

\[
e > \bar{e}^{**}_G(R_G) = \frac{(1 - z - \sigma (c_B - c_M)[c_G - (1 + g)c_B] + \sigma (c_B - c_M) R_G}{1 - z - \sigma (c_B - c_M)},
\]

where, we note, \( \bar{e}^{**}_G(R_G) \) is increasing in \( R_G \). For the result define \( e^{**}_G(R_G) = \text{max}\{e^*_G(R_G), \bar{e}^{**}_G(R_G)\} \). \( \blacksquare \)

\(^{37}\)Note that at \( s = c_B \), \( B \) is indifferent. There is however no loss of generality in assuming that when \( s = c_B \), \( B \) chooses not to comply since as it is easy to verify this is the unique behavior compatible with an equilibrium.
Let \( R_G \) be defined as \( e^*_G(R_G) = c_G \). It is easy to verify that:

\[
R_G = \frac{(1 + g)(1 - z - \sigma (c_B - c_M))c_B}{\sigma (c_B - c_M)}.
\]

Note that at the point \((c_G, R_G)\) we have \( W^G_G(s) = W^G_\emptyset(s) \) and \( W^G_{BG}(s) = W^G_G(s) \), implying that \( W^G_G(s) = W^G_{BG}(s) \) and so \( e^*_G(R_G) = c_G \): so the loci \( e^*_G(R_G), e^*_G(R_G) \) and \( c_G \) intersect at \((c_G, R_G)\).

Define \( R^*_G(e) \) to be equal to \([e^*_G]^{-1}(e)\) for \( e \leq c_G \) and to \([e^*_G]^{-1}(e)\) for \( e > c_G \), where \([e^*_G]^{-1}(e)\) and \([e^*_G]^{-1}(e)\) are the inverse of \( e^*_G(e) \) and \( e^*_G(e) \). So:

\[
R^*_G(e) = \begin{cases} 
\frac{(1+g)e_B-(z+\sigma (c_B-c_M))(e-c_G+(1+g)e_B)}{\sigma (c_B-c_M)} & e \leq c_G \\
\frac{(1-z-\sigma (c_B-c_M))(e-c_G+(1+g)e_B)}{\sigma (c_B-c_M)} & e > c_G
\end{cases}.
\]

The definition of \( R^*_G(e) \) implies that for \( R_G > R^*_G(e) \) we have \( e \in (e^*_G(R_G), e^*_G(R_G)) \), so by Lemma 1.2 we have that the green party finds it optimal to sign a weak agreement. If \( R_G < R^*_G(e) \) and \( e \geq e^*_G \), we have \( e > e^*_G(e) \) and \( e > e^*_G(e) \), Lemma 1.1 and A1.2 implies that the green party finds it optimal to sign a strong agreement. Finally, when \( R_G < R^*_G(e) \) and \( e < e^*_G \), we have \( e < e^*_G(e) \), and Lemma 1.1 implies that the green party finds it optimal to sign no strong agreement.

**Case 2: The brown party is the incumbent**

The welfare generated if both \( B \) and \( G \) comply is, for \( B \) and \( F \): \( W^B_{BG}(s) = zR_B - c_B + e \). If \( G \) only complies, then the sum of payoffs is:

\[
W^B_G(s) = [1 - z + \sigma (s - c_M)] (e - c_B) + [z - \sigma (s - c_M)] (R_B - (1 + g)s).
\]

Note that \( W^B_G(s) \) is convex in \( s \). We have:

**Lemma 1.3.** The brown party signs an agreement if \( e > e^*_B(R_B) \) with \( e^*_B(R_B) \) nonincreasing in \( R_B \).

**Proof.** The case with no agreement cannot occur if \( W^B_\emptyset(0) < W^B_{BG}(s) \), implying \( e > e^*_B \) = \( c_B \), or if \( W^B_\emptyset(0) < W^B_G(s) \). Since \( W^B_G(s) \) is convex in \( s \) we have two cases: \( s = s = c_B \) and \( s = s = c_G \), but it is easy to check that \( s \) is dominated, since \( W^B_G(s) > W^B_G(s) \) ⇒ \( W^B_{BG}(s) > W^B_G(s) \). So, for a weak IEA, \( s = s \). \( B \) and \( F \) prefer such a weak IEA to no IEA if \( W^B_G(s) > W^B_\emptyset(0) \), implying:

\[
(1 - z + \sigma (s - c_M))(e - c_B) - [z - \sigma (s - c_M)]((1 + g)s - R_B) > zR_B,
\]

which can be written as:

\[
e > \tilde{e}^*_B(R_B) \equiv \frac{[1 - z - \sigma (c_M - c_G)]c_B + [z + \sigma (c_M - c_G)](1 + g)c_G - \sigma (c_M - c_G)R_B}{1 - z - \sigma (c_M - c_G)},
\]

35
that, we note, is decreasing in $R_B$. Putting together the two conditions we have that party $G$ chooses to sign an IEA if $e > e^*_B(R_B) = \min\{e^*_B, \tilde{e}^*_B(R_B)\}$.

We now prove the following lemma:

**Lemma 1.4.** There is a threshold $e^*_B(R_B)$ such that the brown party signs a weak agreement if $e \in (e^*_B(R_B), e^*_B(R_B))$, and a strong agreement if $e > e^*_B(R_B)$.

**Proof.** For $e < e^*_B(R_B)$ we have $W^B_{BG}(s) < W^B_0(0)$ and $W^B_G(s) < W^B_0(0)$, so no agreement is signed. For $e \geq e^*_B(R_B)$, a strong agreement is preferred to a weak agreement if $W^B_G(s) < W^B_{BG}(s)$, that is:

$$(1 - z + \sigma (s - c_M))(e - c_B + (1 + g)s - R_B) - (1 + g)s + R_B < zR_B - c_B + e.$$ 

That is, if:

$$e > \tilde{e}^{**}_B(R_B) = \frac{z + \sigma (c_M - c_G)}{\sigma (c_M - c_G)} \left( c_B - (1 + g)c_G \right) + \frac{\sigma (c_M - c_G) R_B}{\sigma (c_M - c_G)};$$

which increases in $R_B$. For the result define $e^{**}_B(R_B) = \max\{e^*_B(R_B), \tilde{e}^{**}_B(R_B)\}$. 

As in the previous subsection, we can show that the loci $e^*_B(R_B), e^{**}_B(R_B)$ and $e^*_B$ intersect at the same point, $(c_B, R_B)$ with $R_B = \frac{\left[ z + \sigma (c_M - c_G) \right] [(1 + g)c_G]}{\sigma (c_M - c_G)}$. Define $R^*_B(e)$ to be equal to $[e^*_B]^{-1}(e)$ for $e \leq c_B$ and to $[e^{**}_B]^{-1}(e)$ for $e > c_G$, where $[e^*_B]^{-1}(e)$ and $[e^{**}_B]^{-1}(e)$ are the inverse of $e^*_B(e)$ and $e^{**}_B(e)$. So:

$$R^*_B(e) = \begin{cases} 
\frac{z + \sigma (c_M - c_G)\left[ c_B - (1 + g)c_G \right]}{\sigma (c_M - c_G)} & e \leq c_B \\
\frac{z + \sigma (c_M - c_G)\left[ c_B - (1 + g)c_G \right]}{\sigma (c_M - c_G)} & e > c_B
\end{cases}.$$ 

The definition of $R^*_B(e)$ implies that for $R_B > R^*_B(e)$ we have $e \in (e^*_B(R_B), e^{**}_B(R_B))$, so by Lemma 1.4 we have that the brown party finds it optimal to sign a weak agreement. If $R_B < R^*_B(e)$ and $e \geq e^*_B$, we have $e > e^*_B(e)$ and $e > e^{**}_B(e)$, Lemma 1.3 implies that the brown party finds it optimal to sign a strong agreement. Finally, when $R_B < R^*_B(e)$ and $e < e^*_B$, we have $e < e^*_B(e)$, Lemma 1.3 implies that the brown party finds it optimal to sign no strong agreement.

Restating the formulas of $R^*_B(e)$ and $R^*_B(e)$ in a unified notation we have the threshold stated in Proposition 1.

**7.2 Proof of Proposition 2**

See Online Appendix.
7.3 Proof of Proposition 3

See Online Appendix.

7.4 Proof of Proposition 4

As in Proposition 1 and 3, in the following, we allow the office rents to be different for the two candidates for additional generality: \( R_j \) for \( j = G, B \). We only consider the case in which the first-period incumbent is \( i = G \); the proof for a \( B \) incumbent is analogous and presented in the Online Appendix.

(i) As explained in the text, an equilibrium treaty can be summarized as the triplet \((x^*_G, \Delta^*_G, S^*_G)\). When \( p \) is the probability that \( G \) wins, and there is full compliance, the expected sum of payoffs for \( G \) and \( F \) is:

\[
p \left[ \frac{e(x^*_G) - e(x^*_G - \Delta^*_G)}{+ (1 + g)\Delta^*_G S^*_G - \Delta^*_G c_G + R_G} \right] + e(x^*_G - \Delta^*_G) - (x^*_G - \Delta^*_G) c_G - (1 + g)\Delta^*_G S^*_G,
\]

where \( p = z + \sigma (S^*_G - c_M) \Delta^*_G \). It is easy to see that this expression is convex in \( S^*_G \) and that the smallest \( S^*_G \) satisfying \( S^*_G \in [c_G, c_B] \) is dominated by either \( S^*_G = 0 \) or \( S^*_G > c_G \). Thus, if \( F \) and \( G \) implement a weak treaty, then in the equilibrium: \( S^*_G = c_B \). Given this \( S^*_G \), the first-order condition with respect to \( x^*_G \) is:

\[
p[e'(x^*_G) - e'(x^*_G - \Delta^*_G)] + e'(x^*_G - \Delta^*_G) - c_G = 0 \Rightarrow pe'(x^*_G) + (1 - p)e'(x^*_G - \Delta^*_G) = c_G,
\]

while the second-order condition trivially holds.

The first order condition with respect to \( \Delta^*_G \) is found by taking the derivative with respect to \( \Delta^*_G \) of the payoff sum and setting this derivative equal to zero. The derivative itself is:

\[
\sigma (c_B - c_M) [e(x^*_G) - e(x^*_G - \Delta^*_G) + (1 + g)\Delta^*_G S - \Delta^*_G c_G + R_G]
\]

\[
- (1 - p) [e'(x^*_G - \Delta^*_G) + (1 + g)c_B - c_G].
\]

The second-order condition is:

\[
\sigma (c_B - c_M) [e'(x^*_G - \Delta^*_G) + (1 + g)c_B - c_G]
\]

\[
+ \sigma (c_B - c_M) [e'(x^*_G - \Delta^*_G) + (1 + g)c_B - c_G]
\]

\[
+ (1 - p) e''(x^*_G - \Delta^*_G) < 0 \Rightarrow
\]

\[
\sigma < \sigma^*_G \equiv \frac{(1 - p) |e''(x^*_G - \Delta^*_G)|}{2(c_B - c_M) [e'(x^*_G - \Delta^*_G) + (1 + g)c_B - c_G]},
\]
which, for any $\sigma$, holds if $e$ is sufficiently concave. In the following, we assume that (11) holds. Then, when $\sigma$ increases, $\Delta^*_G$ must increase to ensure that (10) holds. To avoid that $p \to 1$, we must also assume that:

$$
\sigma < \bar{\sigma}_G^p,
$$

where $\bar{\sigma}_i^p$ is defined such that the inequality in (12) holds with equality. Combined with (11), we henceforth assume $\sigma < \bar{\sigma}_G \equiv \min \{\bar{\sigma}_G, \bar{\sigma}_G^*\}$. The Online Appendix derives the analogous threshold when $i = B$, so that we can define $\bar{\sigma} \equiv \min \{\bar{\sigma}_B, \bar{\sigma}_G\}$.

With this, note that $\Delta^*_G = 0$ is optimal if $\sigma < \bar{\sigma}_G \equiv \min \{\bar{\sigma}_G, \bar{\sigma}_G^*\}$. The Online Appendix derives the analogous threshold when $i = B$, so that we can define $\bar{\sigma} \equiv \min \{\bar{\sigma}_B, \bar{\sigma}_G\}$.

In this case, (9) boils down to $e'(x_G^*) - c_G = 0$. When this equality is substituted into the equation for $\tilde{R}_G$, we can rewrite it as:

$$
\tilde{R}_G \equiv \frac{(1-z) [(1+g)c_B]}{\sigma (c_B - c_M)}.
$$

From the above, it is clear that $\Delta^*_G > 0$ is optimal if $R_G < \tilde{R}_G$. A larger $R_G$ and thus $\Delta^*_G > 0$ implies that $e'(x_G^*) < c_G < e'(x_G^* - \Delta_G^*)$ for (9) to hold. And, when $\tilde{R}_G$ increases, $\Delta^*_G$ must increase for (10) to continue to equal zero, given that second order condition holds.

### 7.5 Proof of Proposition 5

(i) Assume $i = G$ (the case with $i = B$ is in the Online Appendix, which also contains the proof of part (ii) of the proposition). While $R_G$ does not influence (9) directly, (10) increases in $R_G$ so $\Delta^*_G$ must increase to ensure that the expression equals zero. Let $k_G = 0$. If $R_G$ and thus $\Delta^*_G$ increase, the larger $p_G$ reduces the left-hand side of (9), and, for the condition to continue to hold, $x_G^* - \Delta^*_G$ must decline. As $p_G^0 \to 1$, (9) also implies that $e'(x_G^*) \to c_G + k_G$, so $x_G^* \to x_G^{**}$.

### 7.6 Proof of Propositions 6 and 7

The proofs follow relatively straightforwardly from (6), and they are thus omitted.
References


