Participation and Duration of Environmental Agreements

Marco Battaglini

Cornell University and Einaudi Institute for Economics and Finance

Bård Harstad

University of Oslo

We analyze participation in international environmental agreements in a dynamic game in which countries pollute and invest in green technologies. If complete contracts are feasible, participants eliminate the holdup problem associated with their investments; however, most countries prefer to free ride rather than participate. If investments are noncontractible, countries face a holdup problem every time they negotiate; but the free-rider problem can be mitigated and significant participation is feasible. Participation becomes attractive because only large coalitions commit to long-term agreements that circumvent the holdup problem. Under well-specified conditions even the first-best outcome is possible when the contract is incomplete.

I. Introduction

A striking feature of the post–World War II period is the rise of international environmental agreements (henceforth, IEAs). More than

We have benefited from participants at the Climate and the Economy conference—Institute for International Economic Studies at Stockholm University; the Center for the Study of Industrial Organization/Institut d'Economie Industrielle workshop on Industrial Organization—Toulouse; the conference Instruments to Curb Global Warming: Recent Developments—Paris; the Global Climate Cooperation workshop—St. Gallen; the Contracts and Organizations workshop—Norwegian School of Economics; the University of Oslo; and the World Bank. We are particularly grateful for the comments from Tore Ellingsen, Lucas Maestri, Kjetil Storesletten, and Johannes Urpelainen. We thank Nemanja Antic for outstanding research assistance. Harstad's research received funding from the European Research Council under the European Union's 7th Framework Programme, ERC GA 283236.

Electronically published December 29, 2015
[*Journal of Political Economy*, 2016, vol. 124, no. 1]

350 treaties are currently in force, a number that has grown steadily during the last 60 years.[1] Three features seem to characterize such agreements. First, they are voluntary: no international organization can force sovereign countries to cooperate; even side payments to motivate countries to participate are rare.[2] Second, while agreements generally specify abatement levels or other related prohibitions, they leave the regulation of investments in green technology to the discretion of the member states.[3] Third, and despite the first two features, IEAs typically include many countries. The average number of countries in an IEA is 25 and more than three-quarters of them include more than five countries. Some agreements include well over a hundred countries. In the first commitment period of the Kyoto Protocol, 35 countries committed to an average emission reduction target of 5 percent compared to 1990 levels.

It is quite natural that countries may desire agreements in order to limit free riding, since a more healthy environment is a global public good. Participation in an IEA, however, is itself comparable to a public good contribution: besides the cost of the negotiation, it ultimately involves voluntary restrictions on economic activity that also benefit countries that do not participate. Participation, therefore, should be hindered by free-rider problems. Indeed, a recent influential literature has shown that there is no simple theoretical reason to expect that many countries will voluntarily accept to participate in an IEA, casting serious doubts on the efficiency of IEAs characterized by the three features mentioned above. Summarizing this literature, Kolstad and Toman (2005) describe its findings as the "Paradox of International Agreements": while IEAs seem to be ubiquitous, economic theory suggests that they should not exist, or at least they should not be effective in the form in which they are observed. Why, then, do we nevertheless observe a large number of countries that participate in IEAs? What are the consequences of the fact that such agreements are often "incomplete contracts" that specify

[1] This information and the information presented below are derived from the rich data set on IEAs presented in Barrett (2003).

[2] The Helsinki Protocol of 1985 and the Oslo Protocol of 1994 did not include side payments of any type. The Montreal and Kyoto Protocols admit side payments among participants, but only to finance new green investments, not as compensation for participation. In addition, these transfers are insignificant (at least compared to the size of the problems they are supposed to solve). The total budget for the Montreal accord for the 2012–14 triennium is US$450 million for an agreement involving 175 countries. The resources available to the Kyoto agreement for the 2010–12 period amounted to less than $434 million.

[3] For example, at the Durban summit in December 2011, when the European Union agreed to continue its commitments under the Kyoto Protocol, the importance of developing and transferring technology was recognized, but the "technology needs must be nationally determined, based on national circumstances and priorities," according to sec. 114 of the Cancun Agreement (UNFCCC 2011) and confirmed by the Durban Platform (UNFCCC 2012).

emissions but not investments in green technology? How should environmental agreements be designed to be more effective?

This paper presents a new dynamic theory to answer these questions. In our model, countries choose both emission levels and the amount of resources to invest in "green technologies," which are strategic substitutes for polluting activity.[4] Countries also decide whether to free ride or participate in an IEA. The length and depth of the cooperative agreement are endogenous: the coalition members negotiate the number of years for which the agreement holds and the abatement level for each participant. We consider both a "complete contracting" environment, in which the agreement can also specify the investments, and an "incomplete contracting" environment, in which such investments are not contractible. Confirming the previous literature, we show that very few countries find it optimal to cooperate in an environment with complete contracts—regardless of the discount factor and other parameters of the model. Surprisingly, the coalition may be much larger if contracts are incomplete. Under some conditions, even the first-best outcome may be feasible. Thus, our analysis shows that incomplete contracts can be beneficial and explains why environmental coalitions are often quite large.

An important part of our theory is the classic holdup problem. If a country has a large stock of green technology, it will be required to abate more in any efficient agreement or reasonable bargaining game. Anticipating this, countries have few incentives to invest in green technologies during a short-lasting agreement when the next bargaining round is just around the corner. While this observation is not new, the contribution of this paper is to integrate the holdup problem with a coalition formation model to show that an IEA may be successful precisely because it is plagued by a potential holdup problem.

To understand our results, we need to clarify how the duration of the contract depends on the size of the coalition. Suppose a country that is expected to participate instead chooses to deviate by not participating in a particular period. This generates two effects: First, it makes the agreement less ambitious since the policies are chosen to minimize only the externalities generated by the participating countries (it therefore reduces the "depth" of the agreement). Second, and more importantly, the deviation may reduce the duration of the agreement. Indeed, the remaining participants expect the deviator to return to the equilibrium strategy and thus the bargaining table next period, so they find it optimal to "wait" a period, by signing a short-term agreement, rather than to lock in an inefficient long-term agreement. With complete contracts, the

[4] Using terminology standard in the literature, we refer to technologies that reduce the cost of cutting pollution as "green technologies."

duration of the contract is not very important: the IEA will exploit the complete nature of contracts to ensure that countries invest. This is not possible when contracts are incomplete, and short-term agreements will then discourage investments thanks to the holdup problem. The holdup problem generated by a short-term agreement is thus a credible "threat" that reduces the incentives to free ride.

The key insights of the paper remain valid when we endogenize the contractual environment. Allowing the countries to choose whether to make investments contractible may or may not influence the details of the equilibrium—this depends on the exact timing of the decision process—but in any case, only incomplete contracts are signed in equilibrium. Our theory can thus explain why existing climate negotiations do not attempt to contract on investments.

Our positive analysis has important normative implications as well. First, the fact that the Kyoto Protocol is "incomplete" should not necessarily be seen as an accidental design flaw: an effort to closely monitor and control green investments may be counterproductive. Second, it is important to let the final coalition negotiate the duration of the agreement rather than announcing a length before countries have fully committed on whether or not to join. Third, there are multiple equilibria regarding the coalition size. If one could coordinate on the equilibrium with the largest coalition size, then the coalition members would benefit and welfare would increase. Perhaps likely participants can influence the equilibrium selection by announcing an appropriate target for the coalition size.

Given the complexity of the problems we study, it is not surprising that our model has many limitations. We abstract from norms or ethical arguments that may compel countries to participate in IEAs. We also abstract from private information and many types of heterogeneity. There are no technological spillovers and technology cannot be traded. Firms are absent, and each country acts as a single player (so we abstract from important domestic political economy forces).[5] The relationship between the contractual environment and IEAs, which is the primary focus of our paper, however, seems to be an important factor that has not been sufficiently explored by the preceding literature.

From a theoretical point of view, we are not aware of other studies that link contractual incompleteness with the possibility of cooperation in public good problems. There is a huge literature on the holdup problems associated with noncontractible investments (going back to Grossman and Hart [1986] and surveyed by Segal and Whinston [2010]), but contractual incompleteness is generally either harmful or, at best, irrel-

[5] We explore several extensions in Sec. VI; others are explored in our working paper (Battaglini and Harstad 2012).

evant, if the externalities are small and the contract is sufficiently long lasting (Guriev and Kvasov 2005). In an important exception, Bernheim and Whinston (1998) construct simple two-player, two-stage games in which, if some aspects of performance are not contractible, the optimal contract may also leave other aspects of performance unspecified when the players' actions are strategic complements. In our paper we do not require preexisting contractual incompleteness, and because we focus on coalition formation, we study games with many players and an infinite horizon.

In environmental economics, there is an emerging literature that uses insights from the holdup problem to study the relationship between investments in green technologies and international cooperation (see Buchholz and Konrad 1994; Becherle and Tirole 2011; Harstad 2012b, 2015; Helm and Schmidt 2015). These papers develop the idea that individual countries fear that investments in green technology today will weaken their bargaining position in the future, when new commitments are to be negotiated. However, these papers take participation as exogenously given and focus on the harmful effects of the holdup problem. We integrate the holdup problem with an endogenous model of coalition formation and agreement length to show how the holdup problem can be beneficial and lead to a larger equilibrium coalition.[6]

A second strand of related literature in environmental economics focuses on the size of coalitions or IEAs. Building on the work by d'Aspremont et al. (1983) and Palfrey and Rosenthal (1984), this research has highlighted the fact that cooperative agreements are a form of public good, so countries should be expected to free ride on any form of negotiation.[7] The main result of this literature is that international agreements are incentive compatible only if they involve a very small number of countries (Hoel 1992; Carraro and Siniscalco 1993; Barrett 1994; Dixit and Olson 2000; Carraro, Eyckmans, and Finus 2006). This is related to the paradox of international agreements mentioned above. The timing in these models, as in ours, is that countries first decide whether or not to participate in a coalition, and second, the coalition members negotiate an agreement that maximizes the sum of the members' payoffs.[8] The

[6] While relatively few papers focus on the holdup problem, several permit both technological investments and emissions (van der Ploeg and de Zeeuw 1992; Dutta and Radner 2004). Barrett (2006) and Hoel and de Zeeuw (2010) even include a coalition formation stage. More recently, Harstad, Lancia, and Russo (2015) derive the best subgame-perfect equilibria in a model with both investment and emission stages, similar to the one of this paper.

[7] See the surveys by Barrett (2005) and Aldy and Stavins (2007, 2009), among others. A more general survey of the field of climate change economics can be found in Kolstad and Toman (2005).

[8] With the two stages, Coasian bargaining is prevented since a party can commit to not negotiating later (Dixit and Olson 2000; Ellingsen and Paltseva 2012). Alternative coalition formation models are presented by, among others, Chwe (1994), Ray and Vohra (2001), and,

prediction of small coalitions has been found to be robust by a large subsequent literature, which concludes that significant international cooperation is possible only if monetary transfers between countries are feasible (Carraro and Siniscalco 1993; Hoel and Schneider 1997; Bosello, Buchner, and Carraro 2003) or if the environmental technology is characterized by increasing returns or similar technical conditions (Barrett 2005, 2006; Heal and Kunreuther 2011; Karp and Simon 2013). Although this literature is primarily static, dynamic extensions have been presented by Barrett (1994), Rubio and Casino (2005), and Rubio and Ulph (2007) with similar conclusions (see Calvo and Rubio [2012] for a survey).

We build on this literature and extend it in two directions. First, in the preceding literature, negotiations of IEAs are confined to pollution limits lasting for an exogenous length, typically one period. In our dynamic model, the duration of the agreement is endogenously negotiated, so the length becomes a function of the coalition size. Second, we allow for investments in technology and consider environments in which complete contracts are admissible and environments in which only emission levels are contractible. We find that the small-coalition prediction is robust to each of these realistic extensions in isolation, but not when they are combined.[9]

Finally, our work is related to the literature on international trade agreements.[10] Particularly related is the paper by Bagwell and Staiger (2001), who study an economy in which countries choose tariffs and other domestic policies to manipulate their terms of trade. As in our model, with no international agreement, countries achieve an inefficient equilibrium (inefficiently low market access to foreign competitors). A trade agreement can be signed to achieve an efficient outcome if it allows countries to commit to a given level of market access to foreign competitors. Incomplete trade agreements that only set limits on tariffs, however, are inefficient. In contrast to us, Bagwell and Staiger study a static model in which participation is given and not voluntary, and they explicitly rule out nonpecuniary externalities. As we explain in Section IV.D, our model can achieve an efficient equilibrium because it is dynamic, the length of the agreement is endogenous, and participation is voluntary. Because of the differences in the economic environment, the two papers present complementary analyses of international agreements.

The next section presents the model, the equilibrium concept, and two benchmark cases: the first-best solution and the noncooperative

applied to a dynamic model of climate treaties, de Zeeuw (2008). If extraction rights are tradable, then Harstad (2012a) finds that efficiency is possible even with small coalitions.

[9] We also find that there is a positive relationship between the coalition size and depth, which contrasts with the typical observations in the literature (Barrett 2002; Finus and Maus 2008).

[10] See Bagwell and Staiger (2010) for an extensive review of this literature.

"business as usual." Section III solves the game in an environment in which contracts can be complete and confirms the classic result that few countries are willing to participate. Incomplete contracts are considered and proven to be more efficient in Section IV. Section V endogenizes the contractual environment and derives the optimal degree of incompleteness. Various extensions are presented in Section VI, while Section VII presents conclusions. The Appendix contains nontrivial proofs omitted from the main text.

II. Model and Preliminaries

A. *Consumption, Pollution, and Technology*

We consider an economy with many countries and an infinite number of periods. In every period $t \geq 1$, each country $i \in N = \{1, \ldots, n\}$ benefits from consuming $y_{i,t}$, perhaps best interpreted as country i's level of energy. As in much of the literature, we assume that the benefit of consumption, $B_i(y_{i,t})$, is represented by a quadratic and concave function:

$$B_i(y_{i,t}) = -\frac{b}{2}(\bar{y}_i - y_{i,t})^2. \tag{1}$$

The variable $\bar{y}_i$ is an exogenous satiation point that should be assumed to be large: it represents the consumption or energy level country i would choose if there were no concern for climate change. The parameter $b > 0$ measures the disutility of reducing consumption relative to the satiation point.

While consumption is privately beneficial, it contributes to a public bad. We will say that the emission level of country i at time t is

$$g_{i,t} = y_{i,t} - R_{i,t}, \tag{2}$$

where $R_{i,t}$ represents the level of green technology. The stock $R_{i,t}$ may therefore measure the quantity of potential emissions $(y_{i,t})$ that country i can clean thanks to the accumulated abatement technology. Or, as in our favored interpretation, $R_{i,t}$ can measure the quantity of energy generated by country i's renewable energy sources. When $g_{i,t}$ is the quantity of fossil fuel consumption, i's total energy consumption is $y_{i,t} = g_{i,t} + R_{i,t}$, implying (2). We allow $\bar{y}_i$ and the initial stock $R_{i,1}$ to vary across the i's, but countries are otherwise assumed to be identical.

The stock of pollution is $G_t = q_G G_{t-1} + \sum_{i \in N} g_{i,t}$, where $1 - q_G \in [0, 1]$ measures the natural depreciation rate of greenhouse gases. At each point in time, country i's environmental harm is cG_t, where $c > 0$ is assumed to be a constant.

The technology stock depreciates at the rate $(1 - q_R) \in [0, 1]$, and if country i invests $r_{i,t}$ units today, the technology available tomorrow is[11]

$$R_{i,t+1} = q_R R_{i,t} + r_{i,t}. \tag{3}$$

In general, the investment cost, $\kappa_t(\cdot)$, may depend on both the investment level and the level of existing technology. Because of this, we assume that the cost is convex and the marginal cost increases proportionally with the stock of capital. This reflects the fact that existing technological solutions can be ranked according to costs and that the cheapest technology options are developed and installed first. Specifically, we assume that the marginal cost of a unit of technology is

$$\partial\kappa(\cdot)/\partial R_{i,t+1} = kR_{i,t+1}. \tag{4}$$

It follows that $\kappa(\cdot)$ takes the form

$$\kappa(R_{i,t+1},\ R_{i,t}) = \frac{k}{2}(R_{i,t+1}^2 - q_R^2 R_{i,t}^2)$$

when the investment is durable $(q_R > 0)$ and $kr_{i,t}^2/2$ with full depreciation $(q_R = 0)$.[12] Although there may be uncertainty, learning by doing, and increasing returns to scale in reality, cost functions are normally assumed to be both increasing and convex in the literature in order to ensure interior solutions. Assumption (4) is also standard in the literature and so makes our work more comparable with existing findings.[13]

In Section VI.A we extend the analysis to other functional forms, showing that the quadratic forms of $\kappa(\cdot)$ and $B_i(\cdot)$ are not driving the results. In our working paper (Battaglini and Harstad 2012), moreover, we discuss how to allow for heterogeneous investment costs, technological spillovers, tradable permits, and renegotiation.

B. *Timing*

Time can be continuous or discrete. However, we assume that countries invest simultaneously at discrete points in time, they consume simultaneously at discrete points in time, and the consumption stages and the

[11] We do not assume that $r_{i,t}$ is necessarily positive.

[12] To see this, just solve the differential equation $\partial\kappa(\cdot)/\partial R_{i,t+1} = kR_{i,t+1}$ to get $\kappa(\cdot) = kR_{i,t+1}^2/2$ plus a constant or variable that must be independent of $R_{i,t+1}$. Requiring $\kappa = 0$ when $r_{i,t} = 0 \Rightarrow R_{i,t+1} = q_R R_{i,t}$ pins down this constant and thus $\kappa(\cdot)$.

[13] For example, the same assumption of a convex cost of investments in abatement technology with marginal costs that increase linearly in the stock of capital is made by Dutta and Radner (2004), who empirically calibrate their theoretical model to study the dynamic effect of environmental agreements. Dutta and Radner also assume as we do that the marginal benefit of investments is linear (see [3]).

investment stages alternate. In a continuous-time setting, let $\rho > 0$ be the discount rate, $\Delta > 0$ be the time from one emission/consumption decision to the next, and $\Lambda \in (0, \Delta]$ be the time required to develop new technology. The optimal and equilibrium time between the investment stage and the next emission stage is then Λ; thus the time between the emission stage and the next investment stage is $\Delta - \Lambda$. We define a period to start with the emission stage and end with the investment stage. Given this, the utility of country i in period t is

$$u_{i,t} = -\frac{b}{2}(\bar{y}_i - g_{i,t} - R_{i,t})^2 - cG_t - \frac{k}{2}(R_{i,t+1}^2 - q_R^2 R_{i,t}^2)e^{-\rho(\Delta-\Lambda)}$$

for every $i \in N$. Country i at time t seeks to maximize $\sum_{\tau \geq t} \delta^{\tau - t} u_{i,\tau}$, where next period's utility is discounted by the factor $\delta \equiv e^{-\rho\Delta} \in (0, 1)$.[14]

We do not take a stand on what the contractual environment actually is. Instead, we analyze and compare all scenarios we believe are of interest. At the end of this section we derive two benchmark cases: the first-best outcome and the noncooperative, business-as-usual environment, in which nothing is contractible.

Section III analyzes the complete contracting environment. In this case, the stage game is as follows (see fig. 1 for an illustration). (1) Coalition formation stage: if there exists no coalition, every $i \in N$ independently and simultaneously decides whether to become a member of a new coalition, M. The remaining countries, $L \equiv N \setminus M$, remain independent. (2) Negotiation stage: the coalition members first negotiate the duration of the agreement T and then every $g_{i,t}$ and $r_{i,t}$ for $i \in M$ and $t \in \{1, \ldots, T\}$.[15] (3) Emission stage: every nonparticipant $i \in L$ simultaneously and independently chooses $g_{i,t}$, while the coalition members pollute as agreed. (4) Investment stage: every nonparticipant $i \in L$ simultaneously and independently chooses $r_{i,t}$, while the coalition members invest as agreed. Since $R_{i,t}$ is given by the investment stage in the previous period, deciding on $r_{i,t}$ is equivalent to directly choosing $R_{i,t+1}$. If an agreement already existed at the start of the period, the first two stages are skipped.

Section IV considers an incomplete contracting environment in which emissions, but not investments, are contractible. In this case, the coalition members negotiate the $g_{i,t}$'s while the $r_{i,t}$'s are chosen noncooperatively at stage 4.

[14] If $\Lambda = \Delta$, emissions and investments are decided simultaneously. If $\rho \to \infty$ or, equivalently, $\delta \to 0$, then there will be no investment and the next period becomes irrelevant. The model is then as in Barrett (2005, sec. 6.4).

[15] Whether the choices of the policies and the duration are simultaneous or sequential is irrelevant for the results. In the following it will prove convenient for expositional reasons to separate these decisions as if they were sequential.

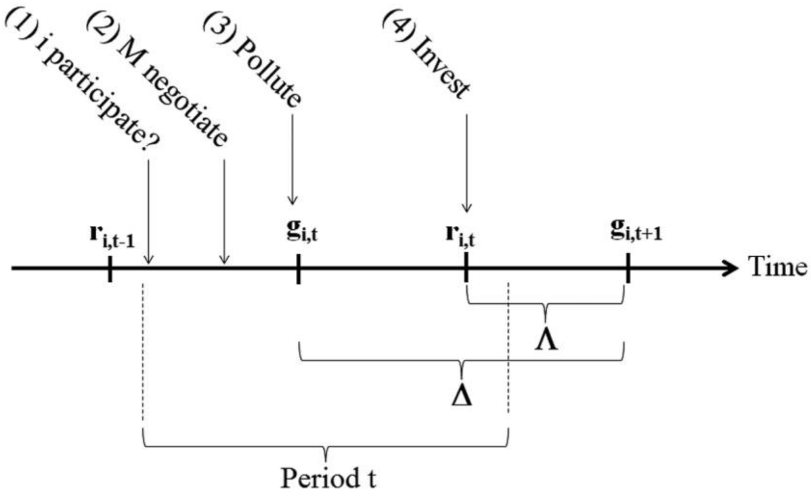

FIG. 1.—The timing of the game

As will be clarified in Section VI.C, we do not need to impose strong assumptions on the outcome of the bargaining stage 2. As a start, however, it is convenient to simply assume that any coalition M cooperatively chooses a policy vector (T, $g_{i,t}$, and, if contracts are complete, $r_{i,t}$) that maximizes the utilitarian welfare of the coalition without any accompanying side transfers. This is the standard assumption in the literature (see the survey by Barrett [2005]).

Our results are also quite robust with respect to timing. For example, stage 3 and stage 4 can occur simultaneously or their timing can be reversed (requiring $\Delta < \Lambda$) without affecting any of the conclusions. Stage 2 and stage 3 may also occur simultaneously or in reversed order: to us it is irrelevant whether or not the coalition acts as a Stackelberg leader since the environmental harm is linear in the stock.

C. *The Equilibrium Concept and Preliminaries*

There is typically a large number of subgame-perfect equilibria in dynamic games. We focus on Markov-perfect equilibria (MPEs) in pure strategies since these are simple and robust and the strategies depend on the payoff-relevant variables only. These equilibria are also empirically plausible.[16]

[16] There is an emerging experimental literature showing that MPEs provide a good description of behavior in dynamic free-rider problems; see Vespa (2011) and Battaglini, Nunnari, and Palfrey (2012, 2015) for recent contributions. Dixit and Olson (2000) and Hong and Karp (2012) analyze equilibria in mixed strategies.

Because of the linearity of the payoffs and technology, the game has a simple structure that allows a practical characterization of all equilibria. To see this, note that the players' preferences can be restated as follows.

LEMMA 1. At any time t, the utility of $i \in N$ is independent of all past stocks and can be represented by the continuation value function $v_{i,t} = \sum_{\tau=t}^{\infty} \delta^{\tau-t} \hat{u}_{i,\tau}$, where

$$\hat{u}_{i,t} \equiv -\frac{b}{2} d_{i,t}^2 - C \sum_{j \in N} (\bar{y}_i - d_{j,t}) - \delta \frac{K}{2} R_{i,t+1}^2 + \delta C \sum_{j \in N} R_{j,t+1}, \tag{5}$$

$$d_{i,t} \equiv \bar{y}_i - (g_{i,t} + R_{i,t}),$$
$$K \equiv k(1 - e^{-\rho\Delta} q_R^2) e^{\rho\Lambda},$$
$$C \equiv \frac{c}{1 - \delta q_G}.$$

Proof. Note that

$$\sum_{\tau=t}^{\infty} \delta^{\tau-t} \hat{u}_{i,\tau} = \sum_{\tau=t}^{\infty} \delta^{\tau-t} u_{i,\tau} + e^{-\rho(\Delta-\Lambda)} q_R^2 R_t^2 k/2,$$

where the latter term is a constant not affecting the ranking of any vectors of future actions. QED

The present-discounted cost of emission is represented by C, while K is the net cost of technology given that some of it survives to later periods. The variable $d_{i,t}$ measures how much i decreases consumption relative to the bliss level. Since $g_{i,t} = \bar{y}_i - d_{i,t} - R_{i,t}$, country i can reduce $g_{i,t}$ either by decreasing consumption or by investing in technology.

The representation in (5) makes clear that the accumulated stocks of greenhouse gases and green technologies enter linearly in the players' objective functions. Because of this, these stocks do not affect the marginal cost or benefit of the policies, nor the players' reaction functions. This fact is key for a simple characterization of the MPEs and their associated strategies. Since the stocks are "payoff irrelevant," the Markov-perfect strategies are conditioned on neither G_t nor the $R_{i,t}$'s. The only relevant state variables are whether an IEA is in force or not and, if so, the prescription of that contract. In particular, all nodes at which there is no contract in place are equivalent.[17]

D. *The First-Best Outcome*

Consider a welfare function $W(v_{1,t}, v_{2,t}, \ldots, v_{n,t})$ that is symmetric, concave, and increasing in each of its arguments. A special case is the util-

[17] A detailed description of the players' strategies will be presented in Secs. III and IV before we analyze the games with contractual completeness and incompleteness.

itarian welfare function $W(\cdot) = \sum_{i \in N} v_{i,t}$. Since $W(\cdot)$ is symmetric and every function $v_{i,t}$ is symmetric and concave in the vectors of $d_{i,t}$'s and $R_{i,t}$'s, the first-best requires that the $d_{i,t}$'s and the $R_{i,t}$'s are identical across the countries. So, even if countries have different ideal points $\bar{y}_i$, it is efficient that they all decrease their consumption level, relative to their ideal point, by the same amount $d_{i,t}$. Furthermore, these uniform policies must be such that each $v_{i,t}$ is maximized. The first-order conditions are then straightforward to derive from (5) (the second-order conditions hold trivially).

PROPOSITION 1.

i. The first-best investments ensure that

$$R_{i,t+1} = n\frac{C}{K} \Leftrightarrow r_{i,t} = n\frac{C}{K} - q_R R_{i,t} \quad \forall t \geq 1.$$

ii. The first-best emission levels are given by

$$d_{i,t} = n\frac{C}{b} \Leftrightarrow g_{i,t} = \bar{y}_i - R_{i,t} - n\frac{C}{b} \quad \forall t \geq 1.$$

Intuitively, if the cost of emission and the number of countries are both large, then it is optimal that each country consumes less as well as invests more in green technology. The two means of reducing emissions should be combined in a sensible way: the technological solution ought to dominate the total abatement effort if K is small, while consumption reduction is cheaper if b is small. The first-best ratio between the two instruments is as follows:

$$\frac{d_{i,t}}{R_{i,t}} = x \equiv \frac{K}{b} \quad \forall t > 1. \tag{6}$$

By definition, x measures how the present value of the marginal cost of investing (taking future cost savings into account) increases in $R_{i,t}$ relative to how the marginal cost of reducing consumption from the bliss level increases in the level of this reduction. At the first-best, this ratio dictates by how much it is optimal to reduce consumption relative to the optimal green technology stock. Since both $d_{i,t}$ and $R_{i,t}$ are proportional to C, the ratio x is independent of C.

E. No Cooperation (Business as Usual)

Suppose instead that each country decides $g_{i,t}$ and $r_{i,t}$ noncooperatively. In an MPE, i anticipates that its choices of $d_{i,t}$ and $R_{i,t}$ do not affect the future choice of $d_{j,\tau}$ and $R_{j,\tau}$ for any player j or time τ. Thus, when each country is simply maximizing $v_{i,t}$ or, equivalently, $\hat{u}_{i,t}$ in (5), we get the following outcome.

PROPOSITION 2. There is a unique Markov-perfect equilibrium.

i. The noncooperative investments ensure that

$$R_{i,t+1} = \frac{C}{K} \Leftrightarrow r_{i,t} = \frac{C}{K} - q_R R_{i,t} \quad \forall t \geq 1. \tag{7}$$

ii. The noncooperative emission levels are given by

$$d_{i,t} = \frac{C}{b} \Leftrightarrow g_{i,t} = \bar{y}_i - \frac{C}{b} - R_{i,t} \quad \forall t \geq 1. \tag{8}$$

The noncooperative equilibrium coincides with the first-best only if $n = 1$. With multiple countries, each country invests too little while it pollutes and consumes too much. Note, however, that the ratio of consumption reduction to technology is exactly as in the first-best:

$$\frac{d_{i,t}}{R_{i,t}} = x \equiv \frac{K}{b} \quad \forall t > 1.$$

III. Contractible Investments

This section analyzes the model in Section II assuming that the coalition can contract on investment as well as emission levels. A pure-strategy equilibrium will specify a coalition M^*, a duration strategy $T^*(M)$, and a policy $(g_{i,t}(T, M), R_{i,t+1}(T, M))_{t=1}^{T}$. Here M^* is the set of countries whose strategy is to join the coalition when there is an opportunity to do so (i.e., in period 1 and in a period following the expiration of an agreement).[18] The function $T^*(M)$ specifies, for any coalition of countries that has chosen to join the IEA, the length of the agreement.[19] The functions $(g_{i,t}(T, M), R_{i,t+1}(T, M))_{t=1}^{T}$ specify the levels of emissions and investments for all periods following the formation of the IEA. The participants collectively choose $g_{i,t}$ and $r_{i,t}$ for every $i \in M$ and $t \in \{1, 2, \ldots, T\}$ at the start of period 1.[20] The nonparticipating countries choose $g_{i,t}$ and $r_{i,t}$ independently in every period.

We first present the equilibrium $g_{i,t}$'s and $r_{i,t}$'s, assuming a duration T and coalition M, before we derive T and, finally, M. Because the model

[18] Because we study pure-strategy MPEs, if a country's strategy prescribes to join with probability one at $t = 0$, then the same country will choose to join with probability one at any period following the expiration of an agreement.

[19] Naturally, in equilibrium we will observe only $T^*(M^*)$, since only countries in M^* join the IEA in equilibrium. However, we still need to specify the reaction function for all the possible coalitions M that can be reached by a unilateral deviation.

[20] Because we focus on MPEs, the period τ in which the IEA is formed is irrelevant, so these functions are independent of τ. If the coalition is formed in period τ, then pollution and investments in the following T periods will be $g_{i,\tau+t} = g_{i,t}(T, M)$ and $R_{i,\tau+t+1} = R_{i,t+1}(T, M)$ for $t = 1, \ldots, T$. In this and the following sections we normalize the period when the coalition is formed to "period 1." Thus, a T-period agreement expires at the end of period T.

is symmetric, the identity of the countries in M is irrelevant; that is, if we have an equilibrium with coalition M, then we have an equilibrium with any other coalition $M' \neq M$ with $|M'| = |M|$. In the remainder we will ignore the identity of countries in the equilibrium coalition and simply focus on the characterization of the number of countries $m^* = |M^*|$ that join the IEA.

A. *Emissions and Investments*

For the reasons described in the business-as-usual case above, every nonparticipant acts according to (7)–(8). The coalition ensures that the externalities of the m coalition members are taken into account, but it does not internalize the environmental harm on the nonparticipants. Negotiating the $r_{i,t}$'s is equivalent to negotiating the $R_{i,t+1}$. Furthermore, agreeing on $g_{i,t}$ is equivalent to agreeing on $d_{i,t} = \bar{y}_i - R_{i,t} - g_{i,t}$.

PROPOSITION 3.

i. For every coalition member, equilibrium investment levels ensure that

$$R_{i,t+1} = m\frac{C}{K} \Leftrightarrow r_{i,t} = m\frac{C}{K} - q_R R_{i,t} \quad \forall i \in M,\ t \in \{1, \ldots, T\}.$$

ii. Equilibrium consumption and emission are given by

$$d_{i,t} = m\frac{C}{b}, \quad t \in \{1, \ldots, T\} \Rightarrow$$
$$g_{i,1} = \bar{y}_i - R_{i,1} - m\frac{C}{b}$$

and

$$g_{i,t} = \bar{y}_i - m\frac{C}{K} - m\frac{C}{b}, \quad t \in \{2, \ldots, T\}.$$

Proof. Since every country has the identical preference $\hat{u}_{i,t}$, the negotiated $d_{i,t}$'s and the $R_{i,t}$'s will be identical for every $i \in M$, and these maximize $\sum_{j \in M} \hat{u}_{j,t}$. The first-order conditions in proposition 3 follow and the second-order conditions are trivially fulfilled. QED

Every coalition member invests more and consumes less if the coalition size is large. The investment and abatement levels are first-best if $m = n$, but they are otherwise too low. It is interesting to note that independent of m, and even if $m < n$, the ratio of consumption reduction to technology stock is efficient: the coalition chooses the right mixture of investments relative to general abatement.

Corollary to proposition 3.

i. We have $d_{i,t}/R_{i,t} = x$ for every $t \in \{2, \ldots, T\}$.

ii. If $m = n$, the outcome would be first-best for every $t \in \{1, \ldots, T\}$ regardless of T.

Finally, note that the coalition's optimal $d_{i,t}$ and $R_{i,t+1}$ are independent of any past stocks, the duration of the agreement, and what the countries expect will replace it.

B. *Duration of the Agreement*

While proposition 3 holds for any contract length, no matter where it comes from, we can also ask for the equilibrium T when the countries can freely negotiate it. The choice of T will depend on the composition of the current coalition, M, as well as on what the countries believe will replace the agreement. As noted already, the equilibrium coalition, M^*, will be independent of any stock, history, or time in an MPE. Thus, no matter the actual composition of the current coalition, M, everyone expects that, once the current agreement expires, the next coalition will be M^*. The next proposition characterizes the equilibrium duration as a function of the coalition's actual size.[21]

Proposition 4. Let M^* denote an equilibrium coalition of size $m^* \equiv |M^*|$. Then, a coalition of size $m = |M|$, satisfying $M \subseteq M^*$ or $M^* \subseteq M$, finds it optimal to contract for $T(m)$ periods, where

$$T(m) = \begin{cases} 1 & \text{if } m < m^* \\ \{1, \ldots, \infty\} & \text{if } m = m^* \\ \infty & \text{if } m > m^*. \end{cases}$$

From proposition 4 we learn that if the coalition happens to be smaller than the equilibrium coalition, the coalition strictly prefers a one-period agreement, since a larger coalition is to be expected next period. If the current coalition equals the equilibrium coalition, then any length is a best choice. If the length is $T < \infty$, for example, the identical coalition (comprising the same m countries) will form and negotiate the identical terms in period $T + 1$, generating the same payoffs to everyone, irrespective of the choice of T.

[21] Proposition 4 does not specify the players' reaction function when neither $M \subseteq M^*$ nor $M^* \subseteq M$. The reaction function after these out-of-equilibrium histories is irrelevant for the equilibrium conditions since a coalition reached after a unilateral deviation must be such that either $M \subseteq M^*$ or $M^* \subseteq M$.

C. *Participation*

We can now analyze the first stage of the game. For M^* to be an equilibrium coalition, it must be both externally stable and internally stable. External stability requires that every $j \in N \setminus M^*$ should be unwilling to join. It can be shown that this condition is satisfied whenever $|M^*| > 1$. Internal stability requires that every $i \in M^*$ does not strictly prefer to free ride.

When a country contemplates whether or not to join the coalition, it anticipates the reaction function described in proposition 4. In particular, if a country that is supposed to participate in equilibrium considers a deviation, then it understands that the consequence will be a one-period contract and that the country will be expected, and find optimal, to join the coalition next period. The country must then balance the gains from its own lower investment cost and higher consumption today, with the fact that the coalition members will not take the externality on i into account (i.e., they will consume more and invest less when the coalition is smaller). This trade-off determines whether a country would like to join the coalition.

PROPOSITION 5. M^* is an equilibrium coalition if and only if $m^* = |M^*| \leq 3$.

The result is dismal. Even with extremely patient players and large externalities, the equilibrium coalition size will be very small. The gain from participating is the fact that the other coalition members will take the entrant's externality into account and thus further reduce consumption and raise investment. Proposition 5 shows that these gains cannot motivate more than three countries to join.[22]

Recall that a special case of our model is the workhorse model with one period and no investments (achieved by letting $\delta = 0$ and x and K approach infinity). A well-known result from that literature is that at most three countries will join the coalition (Barrett 2005). This result is quite robust in that it is independent of any parameters of the model. Proposition 5 shows that this discouraging result continues to hold even if we have multiple periods, if we have investment in green technologies, and if countries can contract on all these choices for any length of time.

IV. Incomplete Contracts

As discussed in the introduction, real climate negotiations have mainly focused on emission levels, leaving the investment decisions to individual countries. To also capture this situation, we now relax the assumption that the policy is fully contractible and assume that countries can

[22] The reason that the discount factor does not help in obtaining a larger coalition is intuitive: as δ increases, the benefit of joining a coalition increases, but so does the benefit of staying out and free riding. The result is that the size remains small even as $\delta \to 1$.

commit to emission levels but not to specified levels of investments. We investigate how investments are influenced by the negotiated emission quotas, how the quotas are decided taking into account the effect on investments, and how the contractual incompleteness influences the equilibrium duration as well as coalition size.

As in the previous section, a pure-strategy equilibrium specifies a coalition M^*, a duration strategy $T^*(M)$, and a policy $(g_{i,t}(T, M), R_{i,t+1}(T, M))_{t=1}^{T}$. The coalition chooses the duration T and commits to $g_{i,t}$ for every $i \in M$ and $t \in \{1, 2, \ldots, T\}$ at the start of period 1. The level of investments, however, is independently chosen by the individual members in every period. Nonparticipating countries choose both $g_{i,t}$ and $r_{i,t}$ independently in every period.

A. *Emissions and Investments*

Just as in the previous sections, nonparticipants find it optimal to consume and invest according to (7)–(8). For coalition members, however, the optimal investment levels will depend on the negotiated quotas. If $g_{i,t}$ is small, then the marginal utility of energy consumption is very large unless $R_{i,t}$ is large. Thus, the smaller the quota, the larger the incentives to invest.

PROPOSITION 6.

i. For every $i \in M$, equilibrium investment ensures that the technology stock decreases in the emission quota:

$$R_{i,t} = \frac{b(\bar{y}_i - g_{i,t})}{b + K}, \quad t \in \{2, \ldots, T\},$$

but $R_{i,T+1} = C/K$.

ii. The equilibrium emission quotas satisfy

$$g_{i,1} = \bar{y}_i - R_{i,1} - m\frac{C}{b},$$

$$g_{i,t} = \bar{y}_i - m\frac{C}{K} - m\frac{C}{b}, \quad t \in \{2, \ldots, T\}$$

$$\Rightarrow R_{i,t} = m\frac{C}{K}, \quad \text{but} \quad R_{i,T+1} = \frac{C}{K}, \quad t \in \{2, \ldots, T\}$$

$$\Rightarrow d_{i,t} = m\frac{C}{b}, \quad t \in \{1, \ldots, T\}.$$

Part i states that country i, in general, invests more if $g_{i,t}$ is small, as is intuitive. In the last period of the agreement, however, the countries realize that the impact of a higher $R_{i,T+1}$ is simply to reduce total emissions (and i's quota) one by one: their investment choices are "sunk" and not payoff relevant in the following period when the countries will choose the $d_{i,T+1}$'s and $R_{i,T+2}$'s. Thus, the marginal benefit to country i of increasing the technological stock is just C: this explains why the equilibrium level of $R_{i,T+1}$ is only C/K. This underinvestment can be interpreted as a consequence of the traditional holdup problem, where parties invest too little when they fear being "held up" in future negotiations.

Part ii describes the equilibrium negotiated quotas. For every period and country, quotas ensure that the marginal benefit of another unit of consumption equals the coalition's cost of more emissions. Since the latter is constant over time, the implication is that $d_{i,t}$ is the same for every $i \in M$ and $t \in \{1, \ldots, T\}$. The countries will then invest the ideal amount for the coalition as a whole, except for the last period, in which every country invests too little. So, except for the last period, emission and investment levels are identical to the complete contracting outcome, if we take T and M as given.

Corollary to proposition 6.

i. We have $d_{i,t}/R_{i,t} = x$ for every $t \in \{2, \ldots, T\}$.

ii. If $m = n$, the outcome would be first-best for every $t \in \{1, \ldots, T\}$ if and only if $T = \infty$.

Note that the above corollary is similar to the corollary to proposition 3. The only difference is that if we had $m = n$ for every agreement, then complete contracts would implement the first-best for any T, while incomplete contracts would implement the first-best only if $T = \infty$. When T is finite, every country invests too little in the last period if investments are noncontractible. If we had $m = n$ and T finite, complete contracts would lead to the first-best while incomplete contracts would not.[23]

B. *Duration of the Agreement*

Proceeding as in the previous section, we next determine equilibrium contract length T, given an arbitrary coalition, M.

Proposition 7. Let M^* denote the equilibrium coalition of size $m^* \equiv |M^*|$. Then, a coalition of size $m = |M|$, satisfying $M \subseteq M^*$ or $M^* \subseteq M$, finds it optimal to contract for $T(m)$ periods, where

$$T(m) = \begin{cases} 1 & \text{if } m < \hat{m}(x, m^*) \\ \{1, \ldots, \infty\} & \text{if } m = \hat{m}(x, m^*) \\ \infty & \text{if } m > \hat{m}(x, m^*), \end{cases}$$

[23] A similar result is derived in the literature on international trade (see Bagwell and Staiger [2001], where $T = 1$ and $n = 2$).

with

$$\hat{m}(x,\ m^*) \equiv m^* - (m^* - 1)\left(1 - \sqrt{\frac{x+\delta}{x+1}}\right) < m^*.$$

In proposition 4, assuming complete contracts, the coalition was indifferent to T if $M = M^*$, and any smaller coalition made them strictly prefer a one-period contract. This is no longer the case. With incomplete contracts, the small investments generated by the holdup problem create a cost of signing short-term agreements. This cost must be weighed against the benefit of waiting for a larger coalition in the future. If the current coalition size, m, is smaller but close to the equilibrium size, m^*, then a long-term agreement with a smaller coalition is nonetheless preferred. The threshold making the coalition indifferent, $\hat{m}(x,\ m^*)$, is thus strictly smaller than m^*.

Proposition 7 allows us to predict the duration if a country deviates from the equilibrium by not participating. In particular, for a unilateral deviation to trigger $T = 1$, it must be the case that $m^* - 1 \leq \hat{m}(x,\ m^*)$. This inequality implies that m^* cannot be too large.

Corollary to proposition 7. If a single country deviates by not participating, the remaining coalition sets $T = 1$ only if $m^* \leq m_M(x)$, where

$$m_M(x) \equiv 1 + \frac{1}{1 - \sqrt{(x+\delta)/(x+1)}}.$$

We will refer to the inequality $m^* \leq m_M(x)$ as the *discipline constraint*. If it is violated, then even if a country $i \in M^*$ deviates by not participating, the remaining participants will proceed by signing a long-term agreement ($T = \infty$). If the discipline constraint is instead satisfied, then whenever some $i \in M^*$ deviates by not participating, the remaining coalition signs a one-period agreement only while it waits for i to return to the equilibrium strategy in the next period.

C. *Participation*

Just as before, any equilibrium coalition M^* must ensure that every $i \in M^*$ prefers to participate. Larger coalitions require larger reductions in pollution from their members (in line with proposition 6), and this makes it more tempting to free ride. The individual participation constraint thus requires that $m^* = |M^*|$ cannot be too large.

If $m^* > m_M(x)$, such that the discipline constraint is violated, then the coalition signs a long-lasting agreement ($T = \infty$) whether i participates or deviates. Proposition 6 then fully characterizes the impact of the smaller m, and investments are exactly as with complete contracts. Compared to the situation with complete contracts, the only differences are that now free riding gives a benefit in every period rather than just for one period; but the cost (i.e., the coalition pollutes more) is also suffered in every period rather than just one. By comparison, the participation constraint still requires that the (one-period) cost is larger than the (one-period) benefit.[24] As shown in the previous section, this participation constraint is satisfied if and only if $m^* \leq 3$.

If the discipline constraint holds, so that $m^* \leq m_M(x)$, then $i \in M^*$ anticipates that free riding would lead to a one-period agreement, triggering the holdup problem. That is, free riding implies that even participants will reduce technology stocks from m^*C/K to simply C/K rather than to $(m^* - 1)C/K$ as with complete contracts. Thus, the punishment for free riding is now higher, and so the individual participation constraint can be satisfied for a larger m^*. The next result determines this threshold, $m_I(x)$, and allows us to characterize all the Markov equilibria.

PROPOSITION 8. M^* is an equilibrium coalition if and only if either $m^* = |M^*| \leq 3$ or $3 < m^* \leq \min\{n, m(x)\}$, where

$$m(x) = \min\{m_I(x),\ m_M(x)\} = \begin{cases} m_M(x) & \text{if } x < \hat{x} \\ m_I(x) & \text{if } x \geq \hat{x}, \end{cases} \tag{9}$$

with

$$m_I(x) \equiv 3 + \frac{2\delta}{x - \delta}$$

and

$$\hat{x} = \frac{1}{6}\left[(1 + \delta) + \sqrt{(1 + \delta)^2 + 12\delta}\right] \in \left(\frac{1}{3},\ 1\right).$$

Just as before, we do have equilibria in which the coalition size is just two or three. In addition, the equilibrium coalition size m^* can now be much larger, as long as it satisfies $m^* \leq m(x)$. In fact, if $n \leq m(x)$, the grand coalition is an equilibrium outcome and the first-best outcome would be implemented. Figure 2 illustrates $m(x)$ as a function of x. The figure shows that, even for very small discount factors, equilibrium participation can be significantly larger than three countries, which is the upper bound with complete contracts. In the example in figure 2 there is an interval for x in which all countries choose to join the IEA and thus the outcome is efficient.

As the above formulas make clear, a key variable is the relative cost of technology, x. This variable has interesting but ambiguous effects on the

[24] When a country free rides in every period rather than in just one period, the benefit as well as the cost must be multiplied by $1/(1 - \delta)$.

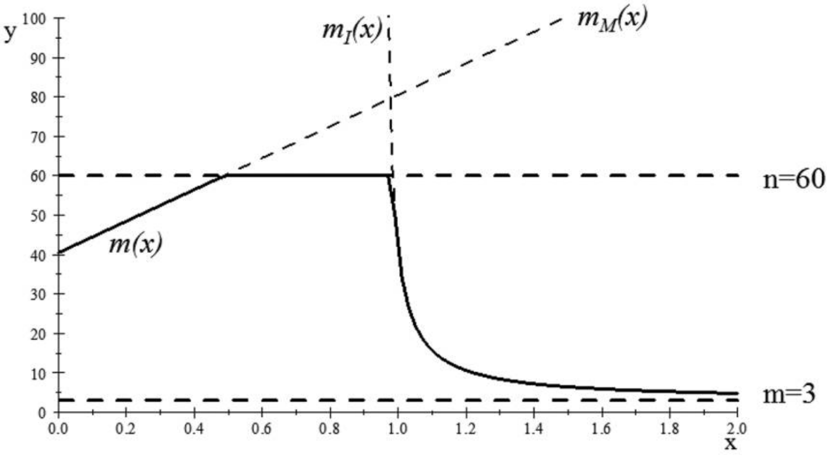

FIG. 2.—The coalition size m^* must be below all three curves

coalition size. Intuitively, a larger x means that technological investment becomes both more expensive and less important as a policy relative to simply reducing consumption. Thus, when x is large, the underinvestment problem following a short-term agreement is less important. This has two consequences. On the one hand, this makes the coalition more willing to sign a short-term agreement and wait for a larger coalition in the future: the discipline constraint is thus relaxed and $m_M(x)$ increases. On the other hand, it becomes more tempting for $i \in M^*$ to deviate, since the subsequent holdup problem is, in any case, less important: the participation constraint is thus strengthened and $m_I(x)$ declines. When $x < \hat{x}$, the binding constraint is $m_M(x)$. If $x > \hat{x}$, the binding constraint is $m_I(x)$. To satisfy both constraints, x must be moderate.

D. Comparing Contractual Environments

In both contractual environments, the equilibrium coalition will be formed and an everlasting agreement will be signed. Every $g_{i,t}$ and $r_{i,t}$ would thus be exactly the same in the two cases if the coalitions were the same. The comparison between the two environments thus boils down to the coalition sizes that can be sustained. The coalition size is important, since utilitarian welfare increases monotonically in the equilibrium coalition size, m^*.

Proposition 8 makes clear that in an incomplete contracting environment we can always sustain a coalition with three countries. For a precise comparison of the equilibrium outcomes in a complete and an incomplete contractual environment, it is useful to recast the result of proposition 8 to characterize the conditions under which a given coalition size can be supported in equilibrium. To this end, note that for every potential equilibrium coalition size m^*, the discipline constraint $m^* \le m_M(x)$ requires

$$x \ge \underline{x}(\delta,\ m^*) \equiv \frac{(m^* - 2)^2 - \delta(m^* - 1)^2}{(m^* - 1)^2 - (m^* - 2)^2}. \tag{10}$$

Similarly, the participation constraint $m^* \le m_I(x)$ requires

$$x \le \overline{x}(\delta,\ m) \equiv \delta + \frac{2\delta}{m - 3}. \tag{11}$$

It follows that a coalition size $m^* \in (3,\ n]$ is feasible in equilibrium if and only if x is moderate in the following sense:

$$\underline{x}(\delta,\ m^*) \le x \le \overline{x}(\delta,\ m^*). \tag{12}$$

Since utilitarian welfare is increasing in m^*, (12) allows us to characterize when a coalition of size $m > 3$ is feasible and, therefore, when the best MPE with incomplete contracts is strictly superior to the best MPE with complete contracts. Expression (12) also allows us to characterize when a coalition of size $m = n$ is feasible and, thus, when the best MPE with incomplete contracts achieves the first-best outcome.

Proposition 9.

i. The maximal coalition size is always weakly larger with incomplete contracts than with complete contracts.

ii. It is strictly larger if and only if

$$x \in [\underline{x}(\delta,\ 4),\ \overline{x}(\delta,\ 4)] = \left[\frac{1}{5}(4 - 9\delta),\ 3\delta\right],$$

a set that is nonempty if $\delta \ge 1/6$.

iii. Moreover, for any n, the best equilibrium with incomplete contracts implements the first-best outcome if and only if

$$x \in [\underline{x}(\delta,\ n),\ \overline{x}(\delta,\ n)] = \left[\frac{(n-2)^2 - \delta(n-1)^2}{(n-1)^2 - (n-2)^2},\ \delta + \frac{2\delta}{n-3}\right],$$

a set that is nonempty if $\delta \ge (n-2)(n-3)/n(n-1) < 1$.

The conditions in parts ii and iii are illustrated in figure 3. The figure plots $\underline{x}(\delta,\ n)$, $\overline{x}(\delta,\ n)$, and $\hat{x}(\delta)$, where $\hat{x}(\delta)$ is the locus of the intersection of the first two curves.[25] The lightly shaded area in the figure describes the region of the parameter space in which an equilibrium coalition size can be larger than three, so that the IEA is strictly superior with incom-

[25] We have $\overline{x}(\delta,\ n) \ge \underline{x}(\delta,\ n)$ (respectively, $\overline{x}(\delta,\ n) \le \underline{x}(\delta,\ n)$) for $x \ge \hat{x}(\delta)$ (respectively, $x \le \hat{x}(\delta)$).

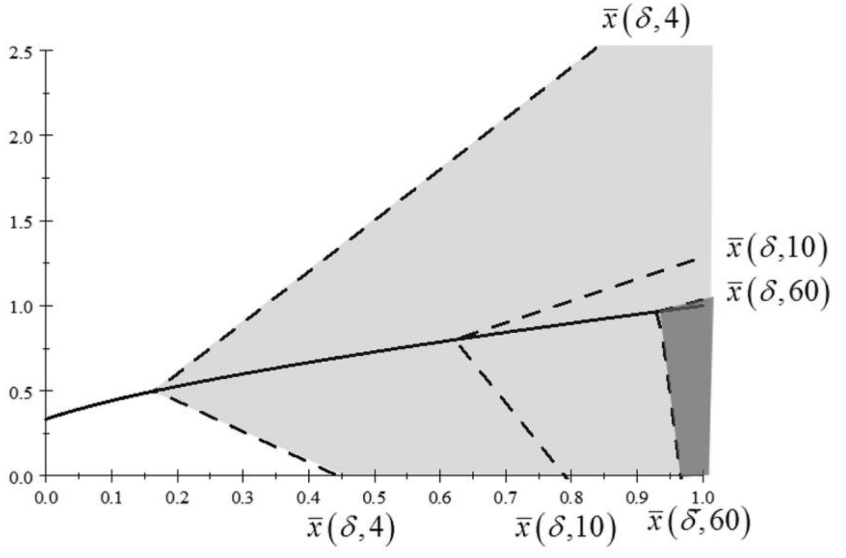

FIG. 3.—Feasible IEAs

plete than with complete contracts.[26] The darkly shaded area corresponds to the region in which there is a fully efficient equilibrium when $n = 60$. This region is nonempty for every finite n.

It is interesting to note how these constraints and regions depend on the discount factor. As expected, if a coalition of size m is feasible at x with some δ, then it remains feasible for any $\delta' > \delta$: the more patient the agents are, the larger is the set of parameters that support a given coalition size. However, an efficient outcome is not always possible, even if δ is arbitrarily large. From part iii we can see that if $x > \bar{x}(1, n) = (n-1)/(n-3)$, then there is no $\delta \leq 1$ such that all countries find it optimal to join.[27]

A crucial assumption in the above analysis is that the game is dynamic and that the contract length can be endogenously negotiated among participating countries. It is easy to show that if the duration T were exogenous, the equilibrium coalition size would be $m^* \leq 3$ regardless of the contractual environment. Since incomplete contracts generate underinvestments in period T, the complete contracting environment would strictly Pareto-dominate the incomplete contracting environment for

[26] The sum of utilities is larger with incomplete contracts than with complete contracts. However, it is not clear that contract incompleteness Pareto-dominates complete contracts. For example, if $m^* = 4$ is possible with incomplete contracts, then the fourth country might prefer an equilibrium with a smaller coalition and one in which it would not be a member, and this might be the expected outcome if contracts were complete.

[27] We have chosen to emphasize the effects of x rather than the impact of δ since the discount factor has multiple interpretations (as patience or period length, e.g.). The alternative interpretations would have conflicting implications for how one should change the model's other parameters when δ changes.

any fixed $T < \infty$.[28] In the robustness section, we discuss how our results survive if the duration is endogenous but limited by a finite upper threshold, $T \leq \overline{T}$.

V. Endogenous Incomplete Contracts

In the preceding sections we have analyzed several alternative situations: complete contracts, no contracts, and incomplete contracts.[29] This way, we did not need to take a definitive stand on what the appropriate contractual environmental is likely to be. Traditionally, the literature on incomplete contracts assumes that the nature of the contractual environment is exogenous: we should expect incomplete contracts when investments are ex post observable by the negotiating partners but not verifiable by a third party, such as an international court.[30] Contractual incompleteness, therefore, seems appropriate when it is hard or costly to describe the exact nature of the investment and its expected payoffs in all conceivable contingencies. For our specific application, there are at least two reasons for assuming contractual incompleteness. First, part of the investment in green technology is in basic research, and this may be difficult to describe ex ante. Second, establishing unanimous criteria to evaluate the effectiveness of new technologies is often more controversial in environmental matters because of its political nature: there are well-funded lobbyists that can produce countervailing evidence on the feasibility of new green technologies and make the issue controversial (even if the science is not).[31]

[28] The fact that, if T is exogenous, we obtain an inefficient outcome highlights the differences of our theory with Bagwell and Staiger's (2001) theory of trade agreement mentioned in the introduction. First, in our model, efficiency requires only control of pollution limits; in Bagwell and Staiger's model, if only tariffs are controlled, the equilibrium is inefficient. As the analysis in this section, however, highlights, our efficiency result is true only if the length and the participation in the agreement are endogenous, features that characterize our model. Second, in Bagwell and Staiger's model, an efficient allocation is possible even without explicit regulation of domestic standards (and so with an "incomplete contract" in our terminology) if countries can commit to a given level of market access. Efficiency, however, is possible because there are no nonpecuniary externalities and because participation in the agreement is exogenous. In our model, efficiency is possible because the model is dynamic and both participation and the length of the agreement are endogenous. Finally, in our model complete contracts make efficiency impossible; in Bagwell and Staiger's model, complete contracts are always good. This occurs because participation is exogenous in their paper; one of the main results of our paper is to show that contractual completeness is harmful to participation.

[29] For the sake of brevity, we have withheld the analysis of the (counterfactual) case in which investments are contractible but emissions are not.

[30] See Hart (1995) and Tirole (1999) for authoritative discussions of the conditions under which it is plausible to assume incomplete contracts.

[31] For example, in September 2013, after the administrator of the Environmental Protection Agency Gina McCarthy proposed to limit new coal power plants to 1,100 pounds of carbon dioxide per megawatt hour, opponents such as the Electric Reliability Coordinating Council responded that the technology to reduce emissions is not yet available (see *New York Times*, September 19, 2013).

There are situations in which the contracting environment is best viewed as endogenously determined. To consider this case, we first introduce the possibility of acquiring a "contracting technology" such that investments can be contracted on (e.g., by establishing standards of measurements and monitoring facilities). We prove that contracts will always remain incomplete in equilibrium, but the possibility to switch to complete contracts might nevertheless influence the outcome. In Section V.B we allow a variety of technologies in which some are contractible while others are not. Measuring the degree of contractual incompleteness, we can extend the results from Section IV as well as derive the optimal (and equilibrium) degree of incompleteness.

A. *Endogenizing the Contractual Environment*

We now let countries decide on whether to make the contractual environment complete: for instance, they can write detailed rules regarding how investments should be measured and establish regulatory agencies that verify and measure each country's investment.[32] Establishing such a monitoring technology on a country may potentially require a cost $h \geq 0$ and some time $\Delta_m \geq 0$. It may be reasonable to assume that (*a*) the monitoring technology is durable and so the cost h is paid only the first time investments are measured; but we will also consider (*b*) the nondurable case in which the cost must be paid in every period in which the monitoring technology is used.

As in the preceding analysis, let $m(x) \geq 3$ refer to the largest possible coalition size under incomplete contracts. If $m(x) = 3$, the equilibrium outcomes under incomplete and complete contracts coincide; so consider the case with $m(x) \geq 4$ and an equilibrium $m^* \in [4, m(x)]$.

Suppose first that $\Delta_m \geq \Delta$, so the decision to measure investments must be made before the coalition formation stage. In this case it is clear that, regardless of h and whether monitoring is (*a*) durable or (*b*) reversible, it is optimal for the coalition to leave contracts incomplete. The coalition is larger under incomplete contracts, and the duration will, in both cases, be infinite in equilibrium.[33]

Suppose next that $\Delta_m < \Delta$, so the coalition can decide whether to contract on investments even after the coalition formation stage. A complete contract is unnecessary if the actual number of coalition members

[32] On the importance of establishing technological standards, note that the Environmental Protection Agency has established a pilot program of cooperation with international authorities to establish mutual recognition of environmental technology verification programs. See, e.g., http://ec.europa.eu/environment/etv/international.htm.

[33] As explained in fn. 26, it is possible that a fourth country would strictly prefer a complete contracting environment, but Coasian bargaining would predict that the surplus-maximizing incomplete contracting environment would prevail.

turned out to be $m \geq m^*$ since then a long-term agreement will be chosen.[34] Thus, suppose $m < m^*$. This out-of-equilibrium possibility is (only) of interest to a country that is contemplating to free ride, so it is sufficient to consider the case $m = m^* - 1 \geq 3$. Two cases are relevant.

a. Assume first that the technology is durable so that the decision to move to complete contracts is essentially irreversible.[35] In this situation, the countries anticipate that after installing the measurement technology the equilibrium coalition size is forever three (at best). The sum of payoffs (for the coalition members) is then smaller than if the current coalition of size $m \geq 3$ commits to a long-term agreement, an option that is available without the measuring technology. Consequently, the coalition will never want to make an irreversible switch to a complete contracting environment no matter the levels of $h \geq 0$ or $m \geq 3$.

b. Assume next that the technology is not durable and measurement cost h must be paid in every period. Switching to a complete contract in this period will then not affect any future MPE. As before, if only $m = m^* - 1$ countries participate at time t, the coalition will prefer to negotiate a short-term agreement (since $m^* \leq m_M(x)$). If investments are not part of the contract, proposition 6 states that $R_{i,t+1} = C/K$. If investments are part of the contract, proposition 3 states that instead $R_{i,t+1} = mC/K$. The one-period gain from contracting on investments is $\delta(m-1)^2 C^2/2K$, and this is less than the cost if

$$\delta(m-1)^2 \frac{C^2}{2K} \leq h \Rightarrow m + 1 \leq m_h \equiv 2 + \frac{\sqrt{2Kh/\delta}}{C}. \tag{13}$$

Consequently, any $m^* \leq m_h$ can be an equilibrium coalition size without violating the constraint that the coalition will stick to the incomplete contracts even when $m = m^* - 1$. Combined with proposition 8, we can conclude that any $m^* > 3$ is an equilibrium coalition size if just $m^* \leq \min\{m_I(x), m_M(x), m_h, n\}$. While both m_M and m_h increase in K, m_I decreases in K. Thus, a simple sufficient condition for the threshold m_h to be nonbinding is $K > \delta C^2/2h \Leftrightarrow m_h > m_I$.

PROPOSITION 10. Suppose that the countries can choose to sign complete contracts.

i. In equilibrium, contracts are always incomplete.
ii. If the decision is irreversible or $\Delta_m \geq \Delta$, the equilibrium is as described by propositions 6–8.

[34] Of course a coalition can simultaneously make the contractual environment complete and choose $T < \infty$. This, however, would not be optimal since after the end of the agreement no coalition larger than three is formed.

[35] This assumption is reasonable if setting up a measurement technology requires a fixed cost, while the cost of subsequently applying and maintaining the technology is negligible (for the result, it is sufficient to assume that the subsequent cost is strictly lower than the initial setup cost).

iii. If the decision is reversible and $\Delta_m < \Delta$, propositions 6–8 hold if (9) is replaced by $m(x) = \min\{m_I(x), m_M(x), m_h\}$.

It is also possible to endogenize the measurement cost. Suppose that the recurring cost $h \in [\underline{h}, \overline{h}]$ can be reduced if the countries take appropriate action in advance (before the coalition formation stage). For example, countries might be able to exert some effort (or up-front payments) in order to reduce the future cost h. What, then, is the equilibrium effort and h? The simple answer is that the countries will never exert any effort in reducing the future h, so $h = \overline{h}$. The (only) consequence of exerting effort would be that m_h and thus m^* may be reduced according to (13). In fact, countries may instead prefer to raise h (and thus m_h). This way, our model can explain why contracting on investments is costly (and perhaps artificially costly).

B. *The Optimal Degree of Incompleteness*

So far, there has been a stark distinction between complete and incomplete contracting environments. Since the reality may be somewhere in between, consider now a situation in which there is a large set of green technology investments (a continuum of mass one). The technologies are identical and characterized by the same investment cost, depreciation rate, perfect substitutability, and effectiveness. The only difference is that a fraction $\alpha \in [0, 1]$ of these technologies (and the associated investments) are contractible while the others are not. In an agreement with duration T, this implies that in the last period, investments ensure that $R_{i,T+1} = mC/K$ for the mass α of contractible investments while $R_{i,T+1} = C/K$ for the noncontractible ones. The results from Section IV continue to hold as long as $\alpha < 1$, but now we obtain more nuanced results in which the feasibility set depends on α.

PROPOSITION 11. Suppose that a fraction $\alpha \in [0, 1]$ of investments are contractible. Then M^* is an equilibrium coalition if and only if either $m^* = |M^*| \leq 3$ or $3 < m^* \leq \min\{n, m(x; \alpha)\}$, where

$$m(x;\alpha) = \min\{m_I(x;\alpha),\ m_M(x;\alpha)\},$$

with

$$m_I(x;\alpha) \equiv 3 + \frac{2\delta(1-\alpha)}{x + 2\delta\alpha - \delta},$$
$$m_M(x;\alpha) \equiv 1 + \mu(\alpha) + \sqrt{\mu(\alpha)[\mu(\alpha) - 1]},$$
$$\mu(\alpha) \equiv \frac{1 + x - \alpha(1-\delta)}{(1-\alpha)(1-\delta)} > 1.$$

The proof is available in the online appendix. The threshold for the participation constraint, $m_I(x; \alpha)$, is decreasing in α for the same reason that $m_I(\cdot)$ was and still is decreasing in x: when the holdup problem becomes less important (because either x or α increases), then a country fears the consequences of a one-period agreement less and free riding becomes tempting unless the coalition is sufficiently small. At the same time, it becomes less costly for the coalition to sign a one-period agreement if a deviator free rides. Thus, the threshold for the discipline constraint, $m_M(x; \alpha)$, is increasing in α as well as in x. Combined, $m(x; \alpha)$ increases in α when $m_M(x; \alpha) < m_I(x; \alpha)$ but decreases in α otherwise.

To complement the previous subsection, we can endogenize the contractual environment by deriving the preferred level of incompleteness—if the countries could decide on α. For the sake of brevity, we consider only the case in which the countries cooperatively decide on α before the coalition formation stage (i.e., $\Delta_m \geq \Delta$, using the notation in the previous subsection). They would then prefer to set α such that the coalition size would be as large as possible. In the online appendix we prove the following proposition.

PROPOSITION 12. Let $\alpha^*(x) = \arg\max_\alpha m(x; \alpha)$: (i) if $x \geq \hat{x}$, then $\alpha^*(x) = 0$; (ii) if $x < \hat{x}$, then $\alpha^*(x) \in (0, 1)$, $\alpha^*(x)$ decreases in x, and it is such that $m_I(x; \alpha^*(x)) = m_M(x; \alpha^*(x))$.

Note that it is always efficient to have some degree of contractual incompleteness: $\alpha^* < 1$. The reason is that in the limit when $\alpha \uparrow 1$, then $m_I(x; \alpha) \downarrow 3$ at the same as $m_M(x; \alpha) \uparrow \infty$. So, for (almost) complete contracts, the binding constraint is always $m_I(x; \alpha)$, which is decreasing in α.

On the other hand, it is possible that $\alpha^* = 0$. If $x \geq \hat{x}$, defined in proposition 8, then $m_I(x; \alpha) \leq m_M(x; \alpha)$ even when $\alpha^* = 0$, and thus the binding constraint is $m_I(x; \alpha) < m_M(x; \alpha)$ for every $\alpha \in (0, 1]$. In this case, $m(x; \alpha) = m_I(x; \alpha)$ is always decreasing in α, and thus we have the corner solution $\alpha^* = 0$.

The importance of the threshold $\hat{x}$ is therefore intuitive: if $x < \hat{x}$, then $m_I(x; \alpha) > m_M(x; \alpha)$ when $\alpha^* = 0$. Since we also know that $m_I(x; 1) < m_M(x; 1)$ and because both thresholds are continuous in α, there exists an $\alpha^* \in (0, 1)$ such that the two thresholds cross, $m_I(x; \alpha^*) = m_M(x; \alpha^*)$. The best degree of contractual incompleteness is then ensuring that both constraints are binding and equalized. Since $\partial m_I(x; \alpha)/\partial x < 0$ while $\partial m_M(x; \alpha)/\partial x > 0$, we have that α^* must decrease in x to ensure $m_I(x; \alpha^*) = m_M(x; \alpha^*)$. Plainly, if the green investments are relatively expensive, then a larger fraction of them should remain noncontractible.[36]

[36] If $n < m(x; \alpha)$ for some $\alpha \in [0, 1]$, neither constraint is binding. There is then an interval $[\underline{\alpha}(x), \bar{\alpha}(x)] \subset \Re$ such that for every $\alpha \in [\underline{\alpha}(x), \bar{\alpha}(x)]$, $m(x; \alpha) = n$ and the first-best is possible. The lower threshold is defined by $m_M(x; \underline{\alpha}(x)) = n$ while the upper threshold is defined by $m_I(x; \bar{\alpha}(x)) = n$. If $\bar{\alpha}(x) < \underline{\alpha}(x)$, the interval is empty and the coalition size is maximized by $\alpha^* = \arg\max_\alpha m(x; \alpha) < n$, as described by proposition 12.

VI. Robustness

In this section we discuss a few extensions of the basic model to show that the results are robust with respect to a number of modeling choices we made for convenience. In particular, we (A) generalize the quadratic formulas above, (B) permit limits on the possibility to commit, (C) show that the bargaining outcome we have assumed can be derived in a noncooperative bargaining game, and (D) discuss how to relax our equilibrium refinement. All the extensions build on the model above (rather than building on each other), so they can be read isolated and in any order. Our working paper (Battaglini and Harstad 2012) discusses technological spillovers, tradable pollution permits, and heterogeneous investment costs.

A. Relaxing the Functional Forms

The adoption of a model with quadratic preferences and cost functions is a convenient choice in the preceding analysis: first, it allows us to directly compare our result with the previous literature (which has made the same assumption); second, it permits simple, closed-form solutions and thus keeps the analysis clean. The intuition for why incomplete contracts are helpful, however, does not hinge on the quadratic formulation. To see how this result generalizes, suppose that the disutility of consumption reduction, $B(d_{i,t})$, is a general increasing concave function while the investment cost is $\delta K(R_{i,t})$, a general increasing and convex function. Suppose $q_R = 0$ for simplicity. If, in addition, we continue to let the marginal disutility of pollution be the constant C, then a complete contract implies $d_{i,t} = B'^{-1}(Cm)$ and $R_{i,t} = K'^{-1}(Cm)$ for the members and $d_{i,t} = B'^{-1}(C)$ and $R_{i,t} = K'^{-1}(C)$ for nonparticipants. The same is true for the case in which an incomplete contract lasts forever.

Just as before, we can show that the largest possible coalition size under incomplete contracts is larger than the largest possible coalition size under complete contracts.

To see this result, note that each member of an m-sized coalition receives the following payoff from this period's choices (analogous to $\hat{u}_{i,t}$ in lemma 1):

$$\begin{aligned}\hat{u}_m^M = & -B(B'^{-1}(Cm)) - \delta K(K'^{-1}(Cm)) - C\sum_{i\in N}\bar{y}_i \\ & + mC[B'^{-1}(Cm) + \delta K'^{-1}(Cm)] \\ & + (n-m)C[B'^{-1}(C) + \delta K'^{-1}(C)].\end{aligned}$$

If a country deviates from this equilibrium, this nonparticipant's one-period payoff is given by the following if contracts are complete:

$$\begin{aligned}\hat{u}_{m-1}^{CC} = & -B(B'^{-1}(C)) - \delta K(K'^{-1}(C)) - C\sum_{i\in N}\bar{y}_i \\ & + (m-1)C[B'^{-1}(C(m-1)) + \delta K'^{-1}(C(m-1))] \\ & + (n-m+1)C[B'^{-1}(C) + \delta K'^{-1}(C)].\end{aligned}$$

If the contract is instead incomplete, investments will be lower in this period, so the deviator's payoff becomes

$$\begin{aligned}\hat{u}_{m-1}^{IC} = & -B(B'^{-1}(C)) - \delta K(K'^{-1}(C)) - C\sum_{i\in N}\bar{y}_i \\ & + (m-1)CB'^{-1}(C(m-1)) + (n-m+1)CB'^{-1}(C) \\ & + \delta nCK'^{-1}(C).\end{aligned}$$

Clearly, we must have $\hat{u}_{m-1}^{CC} > \hat{u}_{m-1}^{IC}$, since the participants invest more under complete contracts and this is beneficial for a nonparticipant.

The participation constraint for a coalition of size m requires that each member must find participation better than free riding one period. For complete contracts, this implies $\hat{u}_m^M \geq \hat{u}_{m-1}^{CC}$; but for incomplete contracts, the condition is $\hat{u}_m^M \geq \hat{u}_{m-1}^{IC}$, which is weaker since $\hat{u}_{m-1}^{CC} > \hat{u}_{m-1}^{IC}$. It follows that every potential member finds free riding less attractive if the contract is incomplete than if it is complete. Thus, the upper boundary m_I (i.e., the largest m satisfying $\hat{u}_m^M \geq \hat{u}_{m-1}^{IC}$) must be larger for incomplete contracts.

The necessary condition $\hat{u}_m^M \geq \hat{u}_{m-1}^{IC}$ is sufficient if it is indeed optimal for the coalition to sign a one-period agreement when one of the countries deviates by not participating. This requires

$$\frac{\hat{u}_{m-1}^M}{1-\delta} \leq \hat{u}_{m-1}^{M,1} + \frac{\delta \hat{u}_{m-1}^M}{1-\delta}, \tag{14}$$

where $\hat{u}_{m-1}^{M,1}$ is the first-period payoff for one of the $m-1$ coalition members. Clearly, this condition is nonbinding if δ is sufficiently close to one since $\hat{u}_m^M > \hat{u}_{m-1}^M$. If condition (14) fails, then the deviator's payoff is $\hat{u}_{m-1}^{CC}$ as with complete contracts. In other words, the equilibrium coalition size is larger for incomplete than for complete contracts even if $B(\cdot)$ and $K(\cdot)$ are nonquadratic.[37]

[37] Of course, even when contracts are complete, the coalition size might be much larger than three (this point has been made by Karp and Simon [2013]). Furthermore, we cannot conclude that the coalition is always strictly larger with incomplete contracts since m must be an integer and the largest integer satisfying $\hat{u}_m^M \geq \hat{u}_{m-1}^{IC}$ may equal the largest integer satisfying $\hat{u}_m^M \geq \hat{u}_{m-1}^{CC}$ even if $\hat{u}_{m-1}^{CC} > \hat{u}_{m-1}^{IC}$. It is for such reasons that it is helpful with specific functional forms, such as those we have above.

B. Commitment and Time Horizon

In the preceding analysis we have assumed that countries in the IEA can commit to a policy for the entire length of an agreement, and we have therefore focused on the issue of participation. In doing this we are following a typical approach in the literature on environmental international agreements (see Barrett 2005). The question of whether such commitment is actually possible is ultimately empirical, and it has been closely scrutinized in the literature. In one of the most comprehensive empirical studies, Breitmeier, Young, and Zurn (2006) conclude that although compliance problems are frequently encountered, "the majority of member states comply with the majority of international environmental rules most of the time" (chap. 3, 66). These significant levels of compliance are explained as the result of explicit enforcement mechanisms in the agreements, but more often by other factors that are often ignored in game theoretic models:[38] for example, establishing mechanisms of compliance monitoring or performance assessment that increase media scrutiny and peer pressure (see Peterson 1997; Young 2011).[39]

Still, it is clear that problems with incentive compatibility and compliance in international agreements may limit their effectiveness. To explore this issue we study how the analysis changes when the countries can commit only for the entire length of the agreement. An interesting benchmark is the case in which countries cannot commit for more than $\overline{T}$ periods. It can be shown that the presence of an upper bound does not change the equilibrium characterization when there is contractual completeness as in Section III. In an incomplete contracting environment, we can show following the exact same steps as in Section IV that an equilibrium exists if and only if

$$\underline{x}(\delta,\ m^*) \leq x \leq \overline{x}(\delta,\ m^*,\ \overline{T}),$$

where $\underline{x}(\delta,\ m^*)$ is defined as in corollary 1 to proposition 8 while

$$\overline{x}(\delta,\ m^*,\ \overline{T}) \equiv \delta\left[1 - \frac{(1-\delta)\delta^{\overline{T}-1}}{1-\delta^{\overline{T}}}\right]\frac{m^*-1}{m^*-3}. \tag{15}$$

[38] Explicit enforcement procedures are contemplated, e.g., in the Montreal Protocol, the protocols of the Geneva Convention, the Basel Protocol, the Aarhus Convention on the Access to Justice in Environmental Matters, the Carthagena Protocol on Biosafety, the International Treaty on Plant Genetic Resources for Food and Agriculture, and the Stockholm Convention on Persistent Organic Pollutant (see Breitmeier et al. 2006).

[39] Examples are mandatory reporting systems of routine information (e.g., in the International Maritime Organization), mechanisms to publicly report deviant behavior to the central organization (e.g., in the Montreal Protocol and the Madrid Protocol on Antarctic Environmental Protection), and mechanisms of performance assessment (e.g., in the Convention on North Pacific Anadromous Stock [Annex II]); see Peterson [1997] for details.

The analysis is therefore as in the previous sections, except that the upper bound of the feasibility set, $\bar{x}(\delta, m^*, \bar{T})$, is now an increasing function of $\bar{T}$: the smaller $\bar{T}$ is, the smaller the region of parameters that sustains an IEA of size m^* is. The intuition is that if $i \in M^*$ deviates by not participating, the holdup problem is moved forward from $\bar{T}$ to the current period. If $\bar{T}$ is small, this "penalty" is small, so the participation constraint strengthens and, to satisfy it, x must be smaller. However, as can be easily verified from (15), quite large coalitions are feasible in an incomplete contracting environment even when the expected length of the agreement is short. Naturally, the upper bound converges to $\bar{x}(\delta, m^*)$ as $\bar{T} \to \infty$.

C. *Noncooperative Bargaining*

In the analysis presented above we have assumed that the policies in the IEA are chosen cooperatively. In this section we present a simple microfoundation of the cooperative decision rule used in the previous sections. To achieve this, we adopt a bargaining model introduced by Baron and Ferejohn (1989), now a standard workhorse model in the political economy literature. Bargaining, in this model, follows a simple dynamic protocol. First, one of the signatory countries is randomly selected to make a proposal. The proposal consists of a time horizon, pollution limits $g_{i,t}$, and (if possible) investments $r_{i,t}$ for each country and each period of the agreement, and a vector of monetary transfers z_i for each country that satisfy budget balance ($\sum_{i \in N} z_i = 0$).[40] Each country has the same probability of being selected to make a proposal. Countries observe the proposal and unanimity is required. If the proposal is accepted, then it is implemented and bargaining ends; if the proposal is rejected, then another country is selected to be the proposer and the process is repeated. The process stops when a policy is chosen. The time between subsequent offers is close to zero, so we ignore discounting between offers.

It is relatively straightforward to prove that if an IEA is an equilibrium of the games studied in the previous sections, then it is an equilibrium of the corresponding game in which policies in the IEA are chosen with the noncooperative bargaining protocol described above. The intuition behind this result is as follows.[41] Take the problem faced by a country selected to propose an IEA. For simplicity, consider only the case with incomplete

[40] In Baron and Ferejohn's bargaining model, countries are allowed to make monetary transfers among each other. As we have said in the introduction, monetary transfers are not typically observed in IEAs. Since in the equilibrium described below transfers are zero, however, this evidence is not necessarily in contrast with the bargaining model with transfers of this section.

[41] Proofs for this result, and the other results in this section, are available from the authors.

contracts (the case with complete contracts is almost identical). Let $u_l(g_{j,l})$ be the indirect utility of country j at time l given the equilibrium investment in green technology $R_{j,l}(g_{j,l})$ from proposition 6.[42] The proposing country desires to maximize its expected utility but will be forced to make a proposal sufficiently appealing to be approved by all other participants. Formally, the proposer's problem at time t can be stated as

$$\begin{gathered}\max_{g_{j,l},z_j,T}\left\{\sum_{l=t}^{t+T}\delta^{l-t}\left[u_l(g_{j,l})-c\sum_{j\in M}g_{j,l}\right]-\sum z_j\right\}\\ \text{subject to } z_j+\sum_{l=t}^{t+T}\delta^{l-t}\left[u_l(g_{j,l})-c\sum_{j\in M}g_{j,l}\right]+\delta^T v_j\geq V_j(M),\end{gathered}\tag{16}$$

where $V_j(M)$ is the outside option for a country that refuses the proposal: that is, the expected utility of entering a new round of bargaining before knowing who the proposer will be. The inequality in (16) is the individual rationality constraint: each agent j must be better off accepting the proposer's offer (the left-hand side of the inequality) than by rejecting it (the right-hand side). It can be shown that the inequality holds as an equality, so we have

$$z_j=V_j(M)-\sum_{l=t}^{t+T}\delta^{l-t}\left[u_l(g_{j,l})-c\sum_{j\in M}g_{j,l}\right]-\delta^T v_j. \tag{17}$$

It is important to note that although endogenous in the model, from the point of view of the proposer, $V_j(M)$ is a constant independent of his or her proposal. Given this, it is easy to see that, modulo a constant that is irrelevant for the solution, we can rewrite (16) as

$$\max_{g_{j,l},T}\left\{\sum_{j\in M}\sum_{l=t}^{t+T}\delta^{l-t}\left[u_l(g_{j,l})-c\sum_{j\in M}g_{j,l}\right]+\delta^T v_j\right\}, \tag{18}$$

which is the utilitarian problem we have been assuming. Note, moreover, that the proposer does not need to make a transfer to have the policy accepted (and will not be able to extract any surplus). If the other countries

[42] Formally, $u_l(g_{j,l})$ is equal to $-(b/2)(Y_{i,l}-g_{i,l}-R_{i,l})^2$ for $l=t$, where $R_{i,t}$ is taken as given from the previous period; to

$$-\frac{b}{2}[Y_{i,l}-g_{i,l}-R_{i,l}(g_{i,l})]^2-\frac{K}{2}R_{i,l}(g_{i,l})^2$$

for $l=t+1,\ldots,T-1$, where $R_{i,l}(g_{i,l})$ is given by proposition 6; and to

$$-\frac{b}{2}(Y_{i,l}-g_{i,l}-R_{i,l})^2-\frac{K}{2}R_{i,l}^2-\delta\frac{K}{2}\left(\frac{C}{K}\right)^2+\delta^T C\sum_{j\in N}\frac{C}{K}$$

for $l=T$.

are expecting a utilitarian solution with no transfer, their expected continuation is

$$V_j(M) = \frac{1-\delta^{T^*-1}}{1-\delta}\left[u_l(g^*_{j,l}) - c\sum_{j\in M} g^*_{j,l}\right] + \delta^{T^*} v_j,$$

where $g^*_{j,l}$, T^* is the solution of (18). Condition (17) then implies that $z_j = 0$. Therefore, the cooperative solution assumed in Sections III and IV is an equilibrium of this noncooperative bargaining.

D. *The Equilibrium Concept*

Up to this point we have focused the analysis on the study of MPEs. These equilibria are appealing because they are simple and they do not rely on complex punishment strategies that may seem unrealistic in many environments, including the coalition formation problem studied above. Although it is hard to test empirically what type of equilibrium is actually played in real-world strategic interactions, recent experimental work has provided evidence in support of MPEs as the appropriate equilibrium concept in dynamic environment with state variables (see, e.g., Vespa 2011; Battaglini et al. 2012, 2015). Because of this, MPEs are widely adopted to study dynamic strategic interactions.[43]

It is, however, important to recognize that more efficient equilibria are possible using history-dependent strategies if the discount factor is sufficiently high. In an MPE, every time that the countries can choose an agreement, the coalition that is formed is history independent. A natural extension is to consider equilibria in which the coalitions that are formed after a deviation may depend on the history of coalitions before the deviation h^t (even though this history is payoff irrelevant). In this case, we can show not only that large agreements are possible using history-dependent strategies but that they can also be constructed with relatively simple strategies. We say that a subgame-perfect equilibrium (SPE) is *simple* if after any history h^t, a coalition $M(h^t)$ is formed for all remaining periods. Differently from MPEs, the coalition may be history dependent; but the SPE is "simple" since it is unnecessary to construct a complex sequence of changing coalitions to discourage free riding. In the online appendix we formally prove that, if δ is sufficiently large, there exists a simple equilibrium in which any number $m \leq n$ of countries join an agreement, even in environments with complete contracts.

[43] See, among others, Levhari and Mirman (1980), Dutta and Radner (2004), Battaglini and Coate (2007, 2008), Besley and Persson (2011), Harstad (2012b). In contrast, Harstad et al. (2015) has recently analyzed the best subgame-perfect equilibria in a game similar to the one of this paper.

In these equilibria, a deviation is punished by the formation of a particular coalition designed to penalize the deviator for the remaining periods: this is done by forming a smaller and less efficient coalition in which the deviating country has to participate. Of course, these punishing coalitions must be equilibrium coalitions in the subgame following the deviation. This implies that the equilibria in these subgames must be supported by even worse threats and even smaller coalitions. Although these equilibria are substantially more complicated than our MPEs because they require a nested chain of punishment phases, they may appear plausible in some environments. In these cases the differences between environments with and without complete contracts that we have highlighted in the previous pages may be less marked, since efficiency can be achieved in both cases, at least for high discount factors.

In an influential contribution, however, Barrett (1994, 2005) has argued that equilibria in coalition formation games should be at least weakly renegotiation proof as defined in Farrell and Maskin (1989).[44] The MPEs derived in the previous sections are all robust to this refinement, since MPEs are weakly renegotiation proof by construction.[45] In the online appendix, however, we formally prove that if weakly renegotiation-proof equilibria with $m > 4$ exist, then they cannot be simple as defined above. This result, therefore, suggests that if larger coalitions can be sustained as renegotiation-proof equilibria with complete contracts, then these equilibria must rely on quite complex punishment strategies. So complex strategies may be unrealistic in the context of international environmental agreements.

VII. Conclusion

This paper provides a theoretical framework for studying coalition formation in dynamic games. When complete contracts are feasible, countries avoid holdup problems associated with investments; few countries, however, choose to participate in the agreement. When contracts are incomplete, on the contrary, the holdup problem induces countries to invest little in green technologies unless the contract duration is sufficiently long. Since only large coalitions sign long-lasting agreements, this effect mitigates the free-rider problem and significant participation is feasible in equilibrium. Our theory therefore explains why coalitions may be larger when contracts are incomplete and even why countries

[44] In our game an equilibrium is weakly renegotiation proof if there are no two histories h^t and $\tilde{h}^t$ in which an agreement is formed in which the continuation equilibrium strategies σ_{h^t} and $\sigma_{\tilde{h}^t}$ are such that σ_{h^t} strictly Pareto-dominates $\sigma_{\tilde{h}^t}$.

[45] In each equilibrium, after any history, the continuation strategies and value function are uniquely defined, so the renegotiation-proofness condition is automatically satisfied.

may prefer to underinvest in "contractual technologies" that could have reduced the contractual incompleteness.

The results have a number of implications for the design of environmental agreements. While critics have suggested that the United Nations' approach is flawed because it focuses only on emissions and not on investments, we have found this to be a possible strength since this may allow for more participants. While some authors advocate a short duration for agreements and others a long duration, we show the importance of letting the duration be endogenously negotiated by the set of committed countries. Although many scholars have suggested that there is a trade-off between size, depth, and length, the Kyoto Protocol arguably fails on all these accounts; this is consistent with our theory, which suggests a positive relationship between depth, breadth, and length. To take advantage of these relationships, it is important that countries coordinate on an equilibrium with a large coalition, that the contract duration is endogenously negotiated, and that future agreements focus only on emission levels and not on investments as well.

Our model has proven to be simple and tractable, and therefore, it both can and should be extended in a number of directions. In particular, we have abstracted from compliance issues, private information, and more complicated equilibria, and we have ignored investments in "brown" technology as well as technological spillovers and trade. All these aspects are important, and they should be included when analyzing environmental agreements in future research.

Appendix

Proof of Proposition 4

Let $m^* \equiv |M^*|$ while T^* is the equilibrium agreement length. If m countries participate in a T-period contract, every i's continuation value can be written as (when substituting from proposition 3)

$$\begin{aligned} v(m,\ T) = {} & \sum_{t=1}^{T} \delta^{t-1}\left\{-\frac{b}{2}\left(\frac{mC}{b}\right)^2 - C\left[\bar{y}_i - (m^2 + n - m)\left(\frac{C}{b} + \frac{\delta C}{K}\right)\right]\right. \\ & \left. - \delta\frac{K}{2}\left(m\frac{C}{K}\right)^2\right\} + \delta^T v(m^*,\ T^*) \\ = {} & -\frac{1-\delta^T}{1-\delta}C\left[\bar{y}_i - C\left(\frac{m^2}{2} + n - m\right)\left(\frac{1}{b} + \frac{\delta}{K}\right)\right] + \delta^T v(m^*,\ T^*). \end{aligned}$$

This implies

$$v(m^*,\ T^*) = -\frac{1}{1-\delta}C\left[\bar{y}_i - C\left(\frac{m^{*2}}{2} + n - m^*\right)\left(\frac{1}{b} + \frac{\delta}{K}\right)\right], \qquad \text{(A1)}$$

and therefore,

$$v(m,\ T) = -\frac{1-\delta^T}{1-\delta} C\left[\bar{y}_i - C\left(\frac{m^2}{2} + n - m\right)\left(\frac{1}{b} + \frac{\delta}{K}\right)\right] - \frac{\delta^T}{1-\delta} C\left[\bar{y}_i - C\left(\frac{m^{*2}}{2} + n - m^*\right)\left(\frac{1}{b} + \frac{\delta}{K}\right)\right]. \quad \text{(A2)}$$

Note that the derivate of $v(m, T)$ with respect to T or, equivalently, with respect to $-\delta^T$ is always negative if and only if

$$\frac{C^2}{1-\delta}\left(\frac{m^2}{2} + n - m\right)\left(\frac{1}{b} + \frac{\delta}{K}\right) \leq \frac{C^2}{1-\delta}\left(\frac{m^{*2}}{2} + n - m^*\right)\left(\frac{1}{b} + \frac{\delta}{K}\right),$$

requiring $m \leq m^*$. QED

Proof of Proposition 5

First, note that a trivial equilibrium is one in which no country joins the coalition (if no one does, it is irrelevant if i does). In the formulas, this would correspond to the situation in which $m^* = 1$.

Following proposition 4, if a participant deviates, then $m = m^* - 1 < m^*$; so $T = 1$ and the participant is expected to join the coalition next period. Such a one-period deviation is not beneficial to i if

$$v(m^*,\ T^*) \geq -\frac{b}{2}\left(\frac{C}{b}\right)^2 - \left\{C\left[\bar{y}_i - (m^2 + n - m)\frac{C}{b}\right] + \delta\frac{K}{2}\left(\frac{C}{K}\right)^2 - \delta C(m^2 + n - m)\frac{C}{K} - \delta v(m^*,\ T^*)\right\}.$$

Substituting expression (A1) for $v(m^*, T^*)$, this condition can be written as

$$-C\left[\bar{y}_i - C\left(\frac{m^{*2}}{2} + n - m^*\right)\left(\frac{1}{b} + \frac{\delta}{K}\right)\right] \geq -\left\{\frac{b}{2}\left(\frac{C}{b}\right)^2 + C\left[\bar{y}_i - (m^2 + n - m)\frac{C}{b}\right] + \delta\frac{K}{2}\left(\frac{C}{K}\right)^2 - \delta C(m^2 + n - m)\frac{C}{K}\right\}.$$

Simplified, this becomes

$$\left(\frac{m^{*2}}{2} - m^*\right)\left(\frac{1}{b} + \frac{\delta}{K}\right) \geq \left(m^{*2} - 3m^* + \frac{3}{2}\right)\left(\frac{1}{b} + \frac{\delta}{K}\right), \quad \text{(A3)}$$

requiring $(m^* - 1)(m^* - 3) \leq 0$. It follows that if $m^* = 3$, each participant is indifferent whether to join; if $m^* = 2$, each participant strictly prefers to join. If $m^* > 3$, no participant would be willing to join. QED

Proof of Proposition 6

Part i: Once the quotas $g_{i,t}$ for $i \in M$ and $t \in \{1, \ldots, T\}$ are negotiated in period 1, country i's continuation payoff can be written recursively as follows (where we drop the subscripts for period t):

$$\begin{aligned} v_i = \sum_{t=1}^{T} \delta^{t-1}\left[-\frac{b}{2}(\bar{y}_i - g_{i,t} - R_{i,t})^2 - C\left(\sum_{j\in N} g_{j,t}\right) - \delta\frac{K}{2}R_{i,t+1}^2\right] \\ + \delta^T v_i + \delta^T C\sum_{j\in N} R_{j,T+1}. \end{aligned} \tag{A4}$$

This recursive formulation recognizes that the game starting at time $T + 1$ is identical to the game starting in period 1 (as before, the stocks are payoff irrelevant at the start of period $T + 1$ as well as period 1, since the stocks do not change the ranking of any vector of future actions).[46]

It follows that the first-order conditions for the $R_{i,t}$'s are

$$\begin{aligned} R_{i,t} &= \frac{b}{K}(\bar{y}_i - g_{i,t} - R_{i,t}) \quad \text{for } t \in \{2, \ldots, T\}, \\ R_{i,T+1} &= \frac{C}{K}. \end{aligned}$$

This implies

$$\begin{aligned} R_{i,t} &= \frac{\bar{y}_i - g_{i,t}}{K/b + 1} \Rightarrow \bar{y}_i - g_{i,t} - R_{i,t} \\ &= \frac{K}{b}\frac{\bar{y}_i - g_{i,t}}{K/b + 1}, \quad t \in \{2, \ldots, T\}. \end{aligned} \tag{A5}$$

Part ii: Substituting (A5) into (A4) and defining $a_{i,t} \equiv \bar{y}_i - g_{i,t}$, we see that every i is identical with respect to the $a_{i,t}$'s. Negotiating the $g_{i,t}$'s is equivalent to negotiating the $a_{i,t}$'s, so, in equilibrium, the $a_{i,t}$'s will be identical and such as to maximize a participant's continuation value. The first-order condition with respect to $a_{i,t} = a_t$, $t \in \{2, \ldots, T\}$, gives

$$-b\left(\frac{K/b}{K/b + 1}\right)^2 a_t + mC - K\left(\frac{1}{K/b + 1}\right)^2 a_t = 0 \Rightarrow \bar{y}_i - m\frac{C}{K} - m\frac{C}{b} = g_{i,t}.$$

For $t = 1$, the countries are, in effect, negotiating the $d_{i,1}$'s directly (since $R_{i,1}$ is given), and all countries have symmetric preferences over the $d_{i,1}$'s and the preferred $d_{i,1} = d_1$ is

$$d_1 = mC/b \Rightarrow g_{i,1} = \bar{y}_i - R_{i,1} - mC/b.$$

QED

[46] Also, note that v_i does not account for the fact that a larger technology stock at the outset reduces emission in the first period (this benefit has already been accounted for); this is why the term $\delta^T C\sum_{j\in N} R_{j,T+1}$ must be added at the end of (A4).

Proof of Proposition 7

It is first useful to prove the following lemma.

LEMMA A1. On the equilibrium path of a Markov equilibrium, $T^* = \infty$.

Proof. Assume not, so that $T^* < \infty$. First note that in a Markov equilibrium the decision to join a coalition is stationary, so the continuation value for a participant can be written recursively as

$$\begin{aligned} v(m^*, T^*) = & -\frac{1-\delta^{T^*}}{1-\delta} C\left[\bar{y}_i - C\left(\frac{m^{*2}}{2} + n - m^*\right)\left(\frac{1}{b} + \frac{\delta}{K}\right)\right] \\ & - \delta^T \frac{C^2}{2K}(m^* - 1)^2 + \delta^T v(m^*, T^*), \end{aligned} \tag{A6}$$

where the second term follows from the fact that, in an incomplete contracting environment, each coalition member receives the additional "benefit" that in the last period, it can invest less, although that, in turn, generates more pollution in period $T + 1$. Compared to the complete contracting situation, the net additional benefit is

$$\begin{aligned} \delta^{T-1}\left[\delta \frac{K}{2}\left(m\frac{C}{K}\right)^2 - \delta\frac{K}{2}\left(\frac{C}{K}\right)^2\right] - \delta^T C\left[m\left(m\frac{C}{K}\right) - m\left(\frac{C}{K}\right)\right] \\ = -\delta^T \frac{C^2}{2K}(m-1)^2 < 0. \end{aligned}$$

Equation (A6) implies that

$$\begin{aligned} v(m^*, T^*) = & -\frac{1-\delta^{T^*}}{1-\delta} C\left[\bar{y}_i - C\left(\frac{m^{*2}}{2} + n - m^*\right)\left(\frac{1}{b} + \frac{\delta}{K}\right)\right] \\ & - \frac{\delta^{T^*}}{1-\delta^{T^*}} \frac{C^2}{2K}(m^* - 1)^2 \\ < & -\frac{1-\delta^{T^*}}{1-\delta} C\left[\bar{y}_i - C\left(\frac{m^{*2}}{2} + n - m^*\right)\left(\frac{1}{b} + \frac{\delta}{K}\right)\right] \\ = & \, v(m^*, \infty), \end{aligned}$$

where the last term is the utility that the coalition would achieve if it committed to an infinite agreement. It follows that $T^* < \infty$ cannot be optimal. QED

We can now prove proposition 7. Given lemma A1, the value of a T-period agreement for each member of a coalition of size m is

$$\begin{aligned} v(m, T) = & -\frac{1-\delta^T}{1-\delta} C\left[\bar{y}_i - C\left(\frac{m^2}{2} + n - m\right)\left(\frac{1}{b} + \frac{\delta}{K}\right)\right] \\ & - \frac{\delta^T}{1-\delta} C\left[\bar{y}_i - C\left(\frac{m^{*2}}{2} + n - m^*\right)\left(\frac{1}{b} + \frac{\delta}{K}\right)\right] - \delta^T \frac{C^2}{2K}(m-1)^2. \end{aligned}$$

Note that the derivate of $v(m, T)$ with respect to T or, equivalently, with respect to $-\delta^T$ is always negative if and only if

$$\left(\frac{m^2}{2}-m\right)+\frac{1-\delta}{2K}\left(\frac{bK}{K+\delta b}\right)(m-1)^2 \le \frac{m^{*2}}{2}-m^*,$$

that is, after some algebraic manipulations, if and only if $m \le \hat{m}(x)$, as defined in proposition 7. QED

Proof of Proposition 8

Suppose $m^* \le \overline{m}_M$. If a country that joins the coalition in equilibrium deviates, then the coalition size will be $m = m^* - 1$ and the coalition will form a one-period contract rather than a long-term contract. The participant is expected to join the coalition next period. Such a one-period deviation is not strictly beneficial to i if

$$v(m^*, T^*) \ge -\left\{\frac{b}{2}\left(\frac{C}{b}\right)^2 + C\left[\bar{y}_i - (m^2+n-m)\frac{C}{b}\right] + \delta\frac{K}{2}\left(\frac{C}{K}\right)^2 - \delta Cn\frac{C}{K}\right\} + \delta v(m^*, T^*),$$

where $m = m^* - 1$. Simplifying, we obtain

$$\left(\frac{m^{*2}}{2}-m^*\right)\left(\frac{1}{b}+\frac{\delta}{K}\right) \ge \left(m^{*2}-3m^*+\frac{3}{2}\right)\frac{1}{b}-\frac{\delta}{2K}.$$

Summing and subtracting $[m^{*2} - 3m^* + (3/2)](\delta/K)$, we obtain

$$\left(\frac{m^{*2}}{2}-m^*\right)\left(\frac{1}{b}+\frac{\delta}{K}\right) \ge \left(m^{*2}-3m^*+\frac{3}{2}\right)\left(\frac{1}{b}+\frac{\delta}{K}\right)-(m^{*2}-3m^*+2)\frac{\delta}{K}.$$

After some algebra, this inequality reduces to

$$2\frac{\delta}{x} \ge (m^*-3)\left(1-\frac{\delta}{x}\right).$$

To prevent a deviation from a nonparticipating country, we also need to satisfy the condition that a nonparticipant does not find it profitable to join the coalition:

$$-\frac{C}{1-\delta}\left\{\bar{y}_i - C\left[\frac{(m^*+1)^2}{2}+n-m^*-1\right]\left(\frac{1}{b}+\frac{\delta}{K}\right)\right\}$$
$$\le -\frac{C}{1-\delta}\left[\bar{y}_i - C\left(m^{*2}+n-m^*-\frac{1}{2}\right)\left(\frac{1}{b}+\frac{\delta}{K}\right)\right],$$

which is implied by $m^*(m^*-2) \ge 0$, or $m^* \ge 2$, which is always satisfied. From proposition 7 we can conclude that an equilibrium of size $m^* \in [2, n]$ exists if

$$2\frac{\delta}{x} \ge (m^*-3)\left(1-\frac{\delta}{x}\right)$$

and

$$m^* \leq 1 + \frac{1}{1 - \sqrt{(x+\delta)/(x+1)}}$$

or, rewriting these two conditions, if $m^* \leq \min\{m_I(x),\ m_M(x)\}$. It is easy to verify that $m_I(x) \geq m_M(x)$ if and only if

$$x \leq \hat{x} = \frac{1}{6}\left[(1+\delta) + \sqrt{(1+\delta)^2 + 12\delta}\right],$$

which proves the sufficiency of $m^* \leq m(x)$.

The fact that $T^* = \infty$ follows from proposition 7. For the remaining results, we proceed in two steps.

Step 1: Assume $m^* = 2$. In this case,

$$m^* \leq 1 + \frac{1}{1 - \sqrt{(x+\delta)/(x+1)}}$$

is always satisfied. Condition

$$2\frac{\delta}{x} \geq (m^* - 3)\left(1 - \frac{\delta}{x}\right)$$

is satisfied if $x \geq \delta$ or, in case $x < \delta$, if $m^* \geq 3 + [2\delta/(x-\delta)]$, that is, if $x \geq \delta + [2\delta/(m^*-3)] = -\delta$, which is always true. If $m^* = 3$, condition $2(\delta/x) \geq (m^* - 3)[1-(\delta/x)]$ is always true. Condition $m^* \leq 1 + \{1/[1 - \sqrt{(x+\delta)/(x+1)}]\}$ is true if

$$x > \frac{\left(\dfrac{m^*-2}{m^*-1}\right)^2 - \delta}{1 - \left(\dfrac{m^*-2}{m^*-1}\right)^2} = \frac{1/4 - \delta}{3/4} = \frac{1}{3} - \frac{4}{3}\delta.$$

Assume $x < (1/3) - (4/3)\delta$. In this case a unilateral deviation is not optimal if $m^* \leq 3$. To see this, note that if a country does not join the coalition, the other countries in the coalition will still find it optimal to commit to an agreement that lasts for an infinite number of periods. In this case, staying out of the coalition is not profitable if

$$\begin{aligned} v(m^*,\ T^*) = &-\frac{C}{1-\delta}\left[\bar{y}_i - C\left(\frac{m^{*2}}{2} + n - m^*\right)\left(\frac{1}{b} + \frac{\delta}{K}\right)\right] \\ \geq &-\frac{1}{1-\delta}\left(\frac{b}{2}\left(\frac{C}{b}\right)^2 + C\left\{\bar{y}_i - [(m^*-1)^2 + n - m^* + 1]\frac{C}{b}\right\}\right. \\ &\left. + \delta\frac{K}{2}\left(\frac{C}{K}\right)^2 - \delta[(m^*-1)^2 + n - m^* + 1]\frac{C^2}{K}\right) \\ = &-\frac{C}{1-\delta}\left[\bar{y}_i - C\left(m^{*2} + n - 3m^* + \frac{3}{2}\right)\left(\frac{1}{b} + \frac{\delta}{K}\right)\right]. \end{aligned}$$

Note that this inequality is the same as (A3) studied in proposition 5: it is satisfied if $m^* \leq 3$.

Step 2: We now prove that the conditions of proposition 8 are necessary. To this end, it will suffice to show that $m^* > 3$ cannot be an equilibrium if it is not the case that $m^* < m_I(x)$ and $m < m_M(x)$. These two inequalities can be written as (10) and (11). We therefore need to consider only three cases:

a. $x > \bar{x}(m^*, \delta)$, $x > \underline{x}(m^*, \delta)$: By the definition of $\bar{x}(m^*, \delta)$, we have that at least one agent has an incentive to free ride by not participating.

b. $x < \underline{x}(m^*, \delta)$: In this case, if a country deviates and does not participate, the remaining coalition members commit to a contract that lasts for an infinite number of periods. In this case, the argument presented in step 1 above shows that it is optimal to deviate if $m^* > 3$.

c. $x > \bar{x}(m^*, \delta)$, $x = \underline{x}(m^*, \delta)$: In this case, if there are $m^* - 1$ countries in the coalition, then the coalition members are indifferent between choosing any T'. Assume that if there are $m^* - 1$ participants, then they choose to commit to an agreement for T' periods, where T' can be anything from one to infinity. The deviation of agent i is profitable if

$$v(m^*, T^*) < -\frac{1-\delta^{T'}}{1-\delta} C\left\{\bar{y}_i - C\left[(m^*-1)^2 + n - (m^*-1) - \frac{1}{2}\right]\left(\frac{1}{b} + \frac{\delta}{K}\right)\right\} - \delta^{T'}\frac{C^2}{K}(m-2)(m-1) + \delta^{T'} v(m^*, T^*) = v'(m^*-1, T').$$

Note that

$$\begin{aligned} v'(m^*-1, T') = & -\frac{1}{1-\delta} C\left\{-\frac{C}{2b} + \bar{y}_i - C[(m^*-1)^2 \right. \\ & \left. + n - (m^*-1)]\frac{1}{b} + \frac{C\delta}{2K} - n\frac{C}{K}\right\} \\ & + \frac{\delta}{1-\delta}\frac{C^2}{K}[(m^*-1)^2 - (m^*-1)]\frac{C^2\delta}{K} \\ & - \frac{\delta^{T'}}{1-\delta^{T'}}\frac{C^2}{K}(m-2)(m-1) \\ = & -\frac{1}{1-\delta} C\left\{-\frac{C}{2b} + \bar{y}_i - C[(m^*-1)^2 + n \right. \\ & \left. - (m^*-1)]\frac{1}{b} + \frac{C\delta}{2K} - n\frac{C}{K}\right\} \\ & + \left(\frac{\delta}{1-\delta} - \frac{\delta^{T'}}{1-\delta^{T'}}\right)\frac{C^2}{K}(m-2)(m-1). \end{aligned} \tag{A7}$$

The right-hand side of (A7) is increasing in T', so the condition is satisfied if it is satisfied for $T' = 1$. By the definition of $\bar{x}(m^*, \delta)$, we have that (A7) is satisfied for $T' = 1$ if $x > \bar{x}(m^*, \delta)$. So when $x > \bar{x}(m^*, \delta)$ and $x = \underline{x}(m^*, \delta)$, agent i has a profitable deviation. QED

References

Aldy, Joseph, and Robert Stavins. 2007. *Architectures for Agreement.* Cambridge: Cambridge Univ. Press.

———. 2009. *Post-Kyoto International Climate Policy.* Cambridge: Cambridge Univ. Press.

Bagwell, Kyle, and R. Staiger. 2001. "Domestic Policies, National Sovereignty and International Economic Institutions." *Q.J.E.* 116 (2): 519–62.

———. 2010. "The World Trade Organization: Theory and Practice." *Ann. Rev. Econ.* 2:223–56.

Baron, David P., and John A. Ferejohn. 1989. "Bargaining in Legislatures." *American Polit. Sci. Rev.* 83 (December): 1181–1206.

Barrett, Scott. 1994. "Self-Enforcing International Environmental Agreements." *Oxford Econ. Papers* 46:878–94.

———. 2002. "Consensus Treaties." *J. Inst. and Theoretical Politics* 158:519–41.

———. 2003. *Environment and Statecraft.* Oxford: Oxford Univ. Press.

———. 2005. "The Theory of International Environmental Agreements." In *Handbook of Environmental Economics*, vol. 3, edited by K.-G. Mäler and J. R. Vincent. Amsterdam: Elsevier.

———. 2006. "Climate Treaties and 'Breakthrough' Technologies." *A.E.R. Papers and Proc.* 96 (2): 22–25.

Battaglini, Marco, and Stephen Coate. 2007. "Inefficiency in Legislative Policy-Making: A Dynamic Analysis." *A.E.R.* 97 (1): 118–49.

———. 2008. "A Dynamic Theory of Public Spending, Taxation, and Debt." *A.E.R.* 98:201–36.

Battaglini, Marco, and Bård Harstad. 2012. "Participation and Duration of Environmental Agreements." Working Paper no. 18585, NBER, Cambridge, MA.

Battaglini, Marco, Salvatore Nunnari, and Thomas Palfrey. 2012. "Legislative Bargaining and the Dynamics of Public Investment." *American Polit. Sci. Rev.* 106 (May): 407–29.

———. 2015. "The Dynamic Free-Rider Problem: A Laboratory Study." Working Paper no. 48-2012, Econ. Theory Center, Princeton Univ.

Beccherle, Julien, and Jean Tirole. 2011. "Regional Initiatives and the Cost of Delaying Binding Climate Change Agreements." *J. Public Econ.* 95 (11–12): 1339–48.

Bernheim, Douglas, and Michael Whinston. 1998. "Incomplete Contracts and Strategic Ambiguity." *A.E.R.* 88 (4): 902–32.

Besley, Timothy, and Torsten Persson. 2011. *Pillars of Prosperity: The Political Economics of Development Clusters.* Princeton, NJ: Princeton Univ. Press.

Bosello, Francesco, Barbara Buchner, and Carlo Carraro. 2003. "Equity, Development, and Climate Change Control." *J. European Econ. Assoc.* 1:601–11.

Breitmeier, Helmut, Oran Young, and Michael Zurn. 2006. *Analyzing International Environmental Regimes: From Case Studies to Database.* Cambridge, MA: MIT Press.

Buchholz, Wolfgang, and Kai Konrad. 1994. "Global Environmental Problems and the Strategic Choice of Technology." *J. Econ.* 60 (3): 299–321.

Calvo, Emilio, and Santiago Rubio. 2012. "Dynamic Models of International Environmental Agreements: A Differential Game Approach." *Internat. Rev. Environmental and Resource Econ.* 6 (4): 289–339.

Carraro, Carlo, Johan Eyckmans, and Michael Finus. 2006. "Optimal Transfers and Participation Decisions in International Environmental Agreements." *Rev. Internat. Orgs.* 1 (4): 379–96.

Carraro, Carlo, and Domenico Siniscalco. 1993. “Strategies for the International Protection of the Environment.” *J. Public Econ.* 52 (3): 309–28.

Chwe, Michael. 1994. “Farsighted Coalitional Stability.” *J. Econ. Theory* 63:299–325.

d'Aspremont, Claude, Alexis Jacquemin, Jean Jaskold Gabszewicz, and John A. Weymark. 1983. “On the Stability of Collusive Price Leadership.” *Canadian J. Econ.* 16 (1): 17–25.

de Zeeuw, Aart. 2008. “Dynamic Effects on the Stability of International Environmental Agreements.” *J. Environmental Econ. and Management* 55 (2): 163–74.

Dixit, Avinash, and Mancur Olson. 2000. “Does Voluntary Participation Undermine the Coase Theorem?” *J. Public Econ.* 76 (3): 309–35.

Dutta, Prajit K., and Roy Radner. 2004. “Self-Enforcing Climate-Change Treaties.” *Proc. Nat. Acad. Sci.* 101:4746–51.

Ellingsen, Tore, and Elena Paltseva. 2012. “The Private Provision of Excludable Public Goods: An Inefficiency Result.” *J. Public Econ.* 96 (9–10): 658–69.

Farrell, Joseph, and Eric Maskin. 1989. “Renegotiation in Repeated Games.” *Games and Econ. Behavior* 1:327–60.

Finus, Michael, and Stefan Maus. 2008. “Modesty May Pay!” *J. Public Econ. Theory* 10:801–26.

Grossman, Sanford J., and Oliver D. Hart. 1986. “The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration.” *J.P.E.* 94 (4): 691–719.

Guriev, Sergei, and Dmitriy Kvasov. 2005. “Contracting on Time.” *A.E.R.* 95 (5): 1269–1385.

Harstad, Bård. 2012a. “Buy Coal! A Case for Supply-Side Environmental Policy.” *J.P.E.* 120 (1): 77–115.

———. 2012b. “Climate Contracts: A Game of Emissions, Investments, Negotiations, and Renegotiations.” *Rev. Econ. Studies* 79 (4): 1527–57.

———. 2015. “The Dynamics of Climate Agreements.” *J. European Econ. Assoc.* Electronically published June 4.

Harstad, Bård, Francesco Lancia, and Alessia Russo. 2015. “Compliance Technology and Self-Enforcing Agreements.” Manuscript, Univ. Oslo.

Hart, Oliver. 1995. *Firms, Contracts, and Financial Structure.* Oxford: Oxford Univ. Press.

Heal, Geoffrey, and Howard Kunreuther. 2011. “Tipping Climate Negotiations.” Working Paper no. 16954, NBER, Cambridge, MA.

Helm, Carsten, and Robert C. Schmidt. 2015. “Climate Cooperation with Technology Investments and Border Carbon Adjustment.” *European Econ. Rev.* 75: 117–30.

Hoel, Michael. 1992. “International Environmental Conventions: The Case of Uniform Reductions of Emissions.” *Environmental and Resource Econ.* 2 (2): 141–59.

Hoel, Michael, and Aart de Zeeuw. 2010. “Can a Focus on Breakthrough Technologies Improve the Performance of International Environmental Agreements?” *Environmental and Resource Econ.* 47 (3): 395–406.

Hoel, Michael, and Kerstin Schneider. 1997. “Incentives to Participate in an International Environmental Agreement.” *Environmental and Resource Econ.* 9: 153–70.

Hong, Fuhai, and Larry Karp. 2012. “International Environmental Agreements with Mixed Strategies and Investment.” *J. Public Econ.* 96 (9–10): 685–97.

Karp, Larry, and Leo Simon. 2013. “Participation Games and International Environmental Agreements: A Nonparametric Model.” *J. Environmental Econ. and Management* 65 (2): 326–44.

Kolstad, Charles D., and Michael Toman. 2005. "The Economics of Climate Policy." In *Handbook of Environmental Economics*, vol. 3, edited by K.-G. Mäler and J. R. Vincent, 1562–93. Amsterdam: Elsevier.

Levhari, David, and Leonard J. Mirman. 1980. "The Great Fish War: An Example Using a Dynamic Nash-Cournot Solution." *Bell J. Econ.* 11:322–34.

Palfrey, Thomas R., and Howard Rosenthal. 1984. "Participation and the Provision of Discrete Public Goods: A Strategic Analysis." *J. Public Econ.* 24 (2): 171–93.

Peterson, M. J. 1997. "International Organizations and the Implementation of Environmental Regimes." In *Global Governance: Drawing Insights from the Environmental Experience*, edited by Oran Young. Cambridge, MA: MIT Press.

Ray, Debraj, and Rajiv Vohra. 2001. "Coalitional Power and Public Goods." *J.P.E.* 109 (6): 1355–84.

Rubio, Santiago J., and Begona Casino. 2005. "Self-Enforcing International Environmental Agreements with a Stock Pollutant." *Spanish Econ. Rev.* 7 (2): 89–109.

Rubio, Santiago J., and Alistair Ulph. 2007. "An Infinite-Horizon Model of Dynamic Membership of International Environmental Agreements." *J. Environmental Econ. and Management* 54 (3): 296–310.

Segal, Ilya, and Michael D. Whinston. 2010. "Property Rights." In *Handbook of Organizational Economics*, edited by Roberts Gibbons. Amsterdam: Elsevier.

Tirole Jean. 1999. "Incomplete Contracts: Where Do We Stand?" *Econometrica* 67 (4): 741–81.

UNFCCC (UN Framework Convention on Climate Change). 2011. "Report of the Conference of the Parties on Its Sixteenth Session, Held in Cancun from 29 November to 10 December 2010, Addendum, Part Two: Action Taken by the Conference of the Parties at Its Sixteenth Session." http://unfccc.int/resource/docs/2010/cop16/eng/07a01.pdf.

———. 2012. "Report of the Conference of the Parties on Its Seventeenth Session, Held in Durban from 28 November to 11 December 2011, Addendum, Part Two: Action Taken by the Conference of the Parties at Its Seventeenth Session." http://unfccc.int/resource/docs/2011/cop17/eng/09a01.pdf.

van der Ploeg, Frederick, and Aart de Zeeuw. 1992. "International Aspects of Pollution Control." *Environmental and Resource Econ.* 2 (2): 117–39.

Vespa, E. 2011. "Cooperation in Dynamic Games: An Experimental Investigation." Manuscript, Univ. California, Santa Barbara.

Young, Oran. 2011. "Effectiveness of International Environmental Regimes: Existing Knowledge, Cutting-Edge Themes, and Research Strategies." *Proc. Nat. Acad. Sci.* 108 (50): 19854–60.