

NBER WORKING PAPER SERIES

PARTICIPATION AND DURATION OF ENVIRONMENTAL AGREEMENTS

Marco Battaglini  
Bård Harstad

Working Paper 18585  
<http://www.nber.org/papers/w18585>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
December 2012

We have benefitted from participants at the Climate and the Economy conference, IIES - Stockholm University, the conference Instruments to curb global warming: Recent developments, Paris, and a seminar at the University of Oslo. We are particularly grateful to the comments from Tore Ellingsen and Kjetil Storesletten. We thank Nemanja Antic for outstanding research assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2012 by Marco Battaglini and Bård Harstad. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Participation and Duration of Environmental Agreements  
Marco Battaglini and Bård Harstad  
NBER Working Paper No. 18585  
December 2012  
JEL No. D86,F53,H87,Q54

**ABSTRACT**

We analyze participation in international environmental agreements (IEAs) in a dynamic game where countries pollute and invest in green technologies. If complete contracts are feasible, participants eliminate the hold-up problem associated with their investments; however, most countries prefer to free-ride rather than participate. If investments are non-contractible, countries face a hold-up 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 hold-up problem. Under well-specified conditions even the first-best outcome is possible when the contract is incomplete. Since real-world IEAs fit in the incomplete contracting environment, our theory may help explaining the rising importance of IEAs and how they should be designed.

Marco Battaglini  
Department of Economics  
Fisher Hall  
Princeton University  
Princeton, NJ 08544  
and NBER  
mbattagl@princeton.edu

Bård Harstad  
Kellogg School of Management  
Northwestern University  
2001 Sheridan Road  
Evanston, IL 60208  
and NBER  
harstad@kellogg.northwestern.edu

# 1. Introduction

One striking feature of the post-World War II period is the rise of international environmental agreements (henceforth, IEAs). More than 350 treaties are currently in force; a number that has grown exponentially over the last sixty years.<sup>1</sup> Three features seem to characterize these agreements. First, they are voluntary: no international organization can force sovereign countries to cooperate; even side payments to motivate countries to participate are rare.<sup>2</sup> 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.<sup>3</sup> 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 76% of them include more than 5 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% compared to 1990-levels.

It is not surprising that countries may desire cooperation in order to limit free riding, since the 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. As such, it should be plagued by free rider problems. Indeed, a recent influential literature has shown 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.<sup>4</sup>

---

<sup>1</sup>This information and the information presented below is derived from the rich data set on IEAs presented in Barrett [2003].

<sup>2</sup>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 2012-2014 triennium is US \$450 million, this for an agreement that includes 175 countries. The resources available to the Kyoto agreement for the 2010-2012 period amounted to less than \$434 million.

<sup>3</sup>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 §114 of the Cancun Agreement (UNFCCC, 2011) and confirmed by the Durban Platform (UNFCCC, 2012).

<sup>4</sup>A detailed survey of the literature is presented at the end of this section.

Summarizing this literature in their survey of environmental economics, Kolstad and Toman [2003] describe its findings as the “Paradox of International Agreements:” while IEAs seem to be ubiquitous, economic theory suggest that they should not exist, or at least they should not be effective in the form in which they are observed. How, then, can we explain that a large number of countries do indeed participate in IEAs? How detrimental is the fact that such agreements are typically “incomplete contracts” which do not specify 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.<sup>5</sup> Countries also decide whether to participate in an IEA or stay out. The length and the depth of the cooperative agreement are endogenous: members of the IEA negotiate the number of years for which the agreement holds and the abatement level for each participant. We show that the effectiveness of IEAs crucially depends on the contractual environment in a surprising way. Two cases are possible: a complete contracting environment, in which the IEA can also specify the green technology investments; and an incomplete contracting environment, in which technological investments are not contractible. We show that, in every Markov equilibrium, only very few countries find it optimal to cooperate in an environment of contractual completeness, regardless of the discount factor and other parameters of the model. In an environment of contractual incompleteness, however, the coalition may be much larger. Indeed under some conditions, even the first-best outcome may be feasible. Thus, our analysis shows that contract incompleteness can be beneficial and explain why environmental coalitions are often quite large.

At the core of the paper is the recognition of a particular type of hold-up problem that arises when countries can invest in green technologies and the impact this can have at the coalition formation stage. Effective cooperation implies that the more efficient a country is in reducing pollution, the larger its share of pollution reduction should be. This implies that, prior to pollution agreements, countries may have incentives to delay

---

<sup>5</sup>Using a terminology standard in the literature, we refer to technologies that reduce the cost of cutting pollution as “green technologies.”

investments. While this observation is not new, the contribution of this paper is to integrate it in an endogenous theory of coalition formation to show that when the length and the size of participation are both endogenous, an IEA may be successful precisely because it is plagued by a potential hold-up problem.

In a nutshell, the intuition behind our results is that although the hold-up problem could decrease the benefit of an agreement with many participants, the value of an agreement with a smaller number of countries is reduced more, so that the temptation to stay out declines. Suppose a country that is expected to participate in equilibrium chooses instead to deviate by not participating in a particular period. This generates two effects: First, it makes the agreement less effective 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, it may induce a shorter 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. Although this second effect is present both in complete and incomplete contracting environments, it is more significant in the latter case. When contracting is complete, the duration of the contract is not very important: countries in the IEA will use the contract to make sure that green investments are efficient. This is not possible when contracts are incomplete, and short-term agreements will then discourage investments thanks to the hold-up problem. The hold-up problem generated by a short-term agreement is thus a credible “threat” which reduces the incentives to free-ride, and it is materialized only off the equilibrium path.

Our positive analysis has important normative implications. 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).<sup>6</sup> 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 hold-up problems associated with noncontractible investments (going back to Grossman and Hart, 1986, and surveyed by Segal and Whinston, 2010), but this contractual incompleteness is always harmful. At best, the fact that investments are non-contractible is irrelevant if the externalities are small and the contract is sufficiently long-lasting (Guriev and Kvasov, 2005). Our paper shows that contract incompleteness is actually *strictly beneficial* when the coalition size is endogenous. This finding may have extensive applications beyond the design of environmental treaties, such as for cartels and other cooperative arrangements between firms.

In environmental economics, there is an emerging literature that uses insights from the hold-up problem to study the relationship between investments in green technologies and international cooperation (see Buchholz and Konrad [1994], Harstad [2012a, 2012b], Becherle and Tirole [2011] and Schmidt [2012]). 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 hold-up problem. We integrate the hold-up problem with an endogenous model of coalition formation and agreement length to show how the hold-up problem can be

---

<sup>6</sup>We explore the robustness of the results to some of these extensions in Section 5.

beneficial and lead to a larger equilibrium coalition.<sup>7</sup>

A second strand of related literature in environmental economics focuses on the size of coalitions or IEAs;<sup>8</sup> this is related to the “Paradox of International Agreements” mentioned above. Building on the work by Palfrey and Rosenthal [1984] and D’Aspermont et al. [1983], 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. 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 [1983], Carraro et al. [2006], Barrett [1994], Dixit and Olson [2000]). The timing in these models, as in our’s, 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.<sup>9</sup> The 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 et al. [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 [2012]). Although this literature is primarily static, dynamic extensions have been presented by Barrett [1994], Rubio and Casino [2005], 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 con-

---

<sup>7</sup>While relatively few papers focus on the hold-up problem, several permit both technological investments and emissions: see Dutta and Radner (2004) or van der Ploeg and de Zeeuw (1992). Hoel and de Zeeuw (2010) and Barrett (2006) even include a coalition formation stage.

<sup>8</sup>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 Toma (2005).

<sup>9</sup>With the two stages, Coasian bargaining is prevented since a party can commit to not negotiate later (Dixit and Olsson [2000], Ellingsen and Paltseva [2011]). Alternative coalition-formation models are presented by, among others, Chwe (1994), Rey and Vohra (2001) and, applied to a dynamic model of climate treaties, de Zeeuw (2008).

tracts 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. 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).

The paper is organized as follows. The next section presents the model, the equilibrium concept, and two benchmark cases: the first-best solution and the noncooperative “business as usual” outcome. Section 3 derives the equilibrium if both emissions and investments are contractible. Section 4 derives the equilibrium in the incomplete contracting setting and shows why this environment is more efficient. Various extensions are discussed in Section 5, while Section 6 concludes. Proofs omitted from the main text are in the Appendix.

## 2. Model and Preliminaries

### A. Consumption, Pollution, and Technology

While the model may have several applications, it is particularly fit to analyze climate change. We consider an economy with many countries and an infinite number of periods. In every period  $t \geq 1$ , each country  $i \in N = \{1, \dots, n\}$  benefits from consuming  $y_{i,t}$ , perhaps best interpreted as country  $i$ 's level of energy. As in much of the literature, we assume the benefit of consumption,  $B_i(y_{i,t})$ , is represented by a quadratic and concave function:

$$B_i(y_{i,t}) = -\frac{b}{2}(\bar{y}_i - y_{i,t})^2. \quad (2.1)$$

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:

$$g_{i,t} = y_{i,t} - R_{i,t}, \quad (2.2)$$



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 (2.2). 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. Since pollution is costly, other countries benefit if  $i$  reduces its emissions. According to (2.2),  $i$  can do this in two ways: by decreasing the consumption level  $y_{i,t}$  or by increasing the technology stock  $R_{i,t}$ . The depreciation rate of  $R_{i,t}$  is  $1 - q_R \in [0, 1]$  but if  $i$  invests  $r_{i,t}$  units today, it can build technology available for tomorrow:<sup>10</sup>

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

Since  $R_{i,t}$  is given by the investment stage in the previous period, deciding on  $r_{i,t}$  is equivalent to directly choosing  $R_{i,t+1}$ . In general, the investment cost,  $\kappa_t(\cdot)$ , may depend on both the investment level and the level of existing technology. We find it natural and helpful to assume that the marginal investment cost of raising  $R_{i,t+1}$  is increasing and linear in the stock itself,  $R_{i,t+1}$ . Intuitively, the cheapest technology options are developed and installed first; the further down the line one goes, the more expensive (relative to its effectiveness) the next unit of technology is. Technically, we assume  $\partial \kappa(\cdot) / \partial R_{i,t+1} = k R_{i,t+1}$ , where  $k > 0$  is a constant. We also require that there is no cost when investing nothing. Combined, it follows that  $\kappa(\cdot)$  must take the following form:<sup>11</sup>

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

If there is full depreciation ( $q_R = 0$ ), the investment cost is simply  $kr_{i,t}^2/2$ .<sup>12</sup>

<sup>10</sup>We do not need to assume  $r_{i,t} \geq 0$ .

<sup>11</sup>To see this, just solve the differential equation  $\partial \kappa(\cdot) / \partial R_{i,t+1} = k R_{i,t+1}$  to get  $\kappa(\cdot) = k R_{i,t+1}^2 / 2$  plus a constant or variable which 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  $\kappa(\cdot)$ .

<sup>12</sup>If one relaxes the typical assumption on quadratic functional forms, then larger coalitions may be

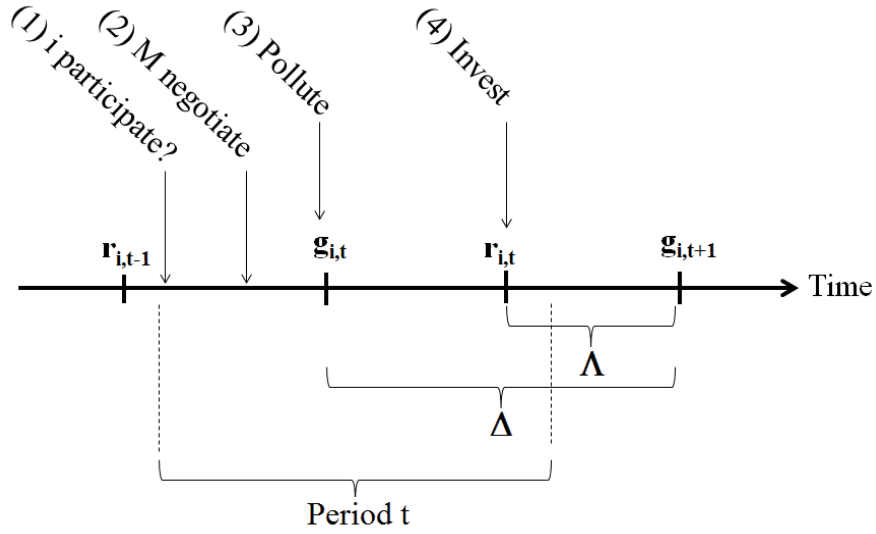


Figure 2.1: *The timing of the game*

### 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 investment stages alternate. In a continuous-time setting, let  $\rho$  be the discount rate,  $\Delta$  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  $\Lambda$ , 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 at the rate  $\delta \equiv e^{-\rho\Delta}$ .<sup>13</sup>

feasible (Karp and Simon, 2010). By sticking to the quadratic formulae, we show that our mechanism works even in this traditionally dismal case.

<sup>13</sup>If  $\Lambda = \Delta$ , emissions and investments are decided simultaneously. If  $\rho \rightarrow \infty$  or, equivalently,  $\delta \rightarrow 0$ , then there will be no investment and the next period becomes irrelevant. The model is then as in Barrett (2005, Section 6.4).

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 3 analyzes the complete contracting environment. In this case, the stage game is the following (see Figure 2.1 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  $t \in \{1, \dots, T\}$ .<sup>14</sup> (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. If a coalition existed at the start of the period, steps (1)-(2) are skipped.

Section 4 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 time (4).

As will become clear in our robustness section, we do not need to impose strong assumptions on the stage (2) bargaining process, where the coalition members negotiate the terms of the contract. 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

---

<sup>14</sup>Whether the choices of  $T$  and the policies 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.

as a Stackelberg leader since the environmental harm is linear in the stock.

For simplicity, we will not consider any international trade in the pollution permits  $g_{i,t}$ s. Section 5-D explains, however, that allowing for such trade among the participating countries will not affect any of the results.

### 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 in pure strategies since these are simple, robust, and the strategies depend on the payoff relevant variables only. These equilibria are also empirically plausible.<sup>15</sup>

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 preference of every  $i \in N$  is independent of any stock or past action, and it can be represented by the continuation value function  $v_{i,t} = \sum_{\tau=t}^{\infty} \delta^{\tau-t} \widehat{u}_{i,\tau}$ , where*

$$\begin{aligned} \widehat{u}_{i,t} &\equiv -\frac{b}{2}d_{i,t}^2 - C \sum_{j \in N} (\bar{y}_i - d_{j,t}) - \delta \frac{K}{2} R_{i,t+1}^2 + \delta C \sum_{j \in N} R_{j,t+1}, & (2.3) \\ d_{i,t} &\equiv \bar{y}_i - (g_{i,t} + R_{i,t}), \\ K &\equiv k(1 - e^{-\rho\Delta} q_R^2) e^{\rho\Lambda}, \text{ and} \\ C &\equiv \frac{c}{1 - \delta q_G}. \end{aligned}$$

**Proof.** Note that  $\sum_{\tau=t}^{\infty} \delta^{\tau-t} \widehat{u}_{i,\tau} = \sum_{\tau=t}^{\infty} \delta^{\tau-t} u_{i,\tau} + e^{-\rho(\Delta-\Lambda)} q_R^2 R_t^2 k/2$ , where the latter term is a constant which therefore does not affect the ranking of any vectors of future actions.

■

The representation in (2.3) 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

<sup>15</sup>There is an emerging experimental literature showing that Markov perfect equilibria provide a good description of behavior in dynamic free rider problems, see Battaglini et al. [2012a], [2012b] and Vespa [2012] for recent contributions. Dixit and Olson (2000) and Hong and Karp (2010) analyze equilibria in mixed strategies.

players' reaction functions. This fact is key for a simple characterization of a Markov equilibrium and its associated strategies. Since the stocks are “payoff-irrelevant”, the Markov strategies are not conditioned on  $G_t$  and the  $R_{i,t}$ s. The only relevant state variables are whether an IEA is in force or not and, if an IEA is in force, the prescription of the IEA. In particular, every node at which there is no IEA in force is equivalent.<sup>16</sup>

#### D. The First-Best

Consider a welfare function  $W(v_{1,t}, v_{2,t}, \dots, v_{n,t})$  which is symmetric, concave, and increasing in each of its arguments. A special case is the utilitarian 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 (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 invest 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:

$$\frac{d_{i,t}}{R_{i,t}} = x \equiv \frac{K}{b}, \quad \forall t > 1. \quad (2.4)$$

---

<sup>16</sup>A detailed description of the players' strategies will be presented in Sections 3 and 4 when we discuss in details the games with contractual completeness and incompleteness.

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 (as shown in the second equality of (2.4)). 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 a Markov equilibrium,  $i$  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  $j$ . Thus, when each country is simply maximizing  $v_{i,t}$ , or equivalently  $\hat{u}_{i,t}$  in (2.3), we get the following outcome.

**Proposition 2.** *There is a unique Markov-perfect equilibrium.*

(i) *The noncooperative investments ensure that:*

$$R_{i,t+1} = \frac{C}{K} \Leftrightarrow r_{i,t} = \frac{C}{K} - q_R R_{i,t} \quad \forall t \geq 1. \quad (2.5)$$

(ii) *The noncooperative emission levels are given by:*

$$d_{i,t} = \frac{C}{b} \Leftrightarrow g_{i,t} = \bar{y}_i - \frac{C}{b} - R_{i,t} \quad \forall t \geq 1. \quad (2.6)$$

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:

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

## 3. Contractible Investment

This section analyzes the model in Section 2 assuming that the coalition can contract on investment as well as emission levels. In this environment a pure strategy equilibrium can be described as a fixed 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 set  $M^*$  describes the countries that choose to join an

IEA when there is an opportunity to do it (that is at period one, or at any period following the expiration of a previously formed IEA).<sup>17</sup> The function  $T^*(M)$  specifies, for any coalition of countries that has chosen to join the IEA, the length of the agreement.<sup>18</sup> The functions  $(g_{i,t}(T, M), R_{i,t+1}(T, M))_{t=1}^T$  specify the levels of pollutions and investments for all periods following the formation of the IEA.<sup>19</sup> The participants collectively choose  $g_{i,t}$  and  $r_{i,t}$  for every  $i \in M$  and  $t \in \{1, 2, \dots, T\}$  at the start of period 1.<sup>20</sup> 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 is symmetric, the identity of the countries in  $M$  is irrelevant, i.e., 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 we will simply focus on the characterization of the number of countries  $m^* = |M^*|$  that join the IEA.

#### A. Emissions and Investments

For the same reasons as in the business-as-usual case discussed above, every nonparticipant acts according to (2.5)-(2.6). For the coalition, negotiating the  $r_{i,t}$ s is equivalent to negotiating the  $R_{i,t+1}$  at time  $t$  when  $R_{i,t}$  must be taken as given. Furthermore, agreeing on  $g_{i,t}$  is equivalent to agreeing on  $d_{i,t} = \bar{y}_i - R_{i,t} - g_{i,t}$ . The coalition ensures that the externality of the  $m$  coalition members are taken into account, but it does not internalize the environmental harm on the nonparticipants.

**Proposition 3.** (i) *For every coalition member, equilibrium investment levels ensure that:*

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

<sup>17</sup>Because we study pure strategy Markov Perfect Equilibria, if a country's strategy prescribes to join with probability one at  $t = 0$ , then the same country will choose to join an IEA with probability one at any period following the expiration of a previously formed IEA.

<sup>18</sup>Naturally, in equilibrium we will observe only  $T^*(M^*)$ , since only countries in  $M^*$  joining 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.

<sup>19</sup>Because we focus on a Markov equilibrium, the period  $\tau$  in which the IEA is formed is irrelevant, so these function are independent of  $\tau$ . If the coalition is formed in period  $\tau$ , 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  $t = 1, \dots, T$ .

<sup>20</sup>In this and the following sections we normalize the period when the coalition is formed to "period one." Given this, a  $T$ -period agreement expires at the end of period  $T$ .

(ii) *Equilibrium consumption and emission are given by:*

$$\begin{aligned} d_{i,t} &= m \frac{C}{b}, t \in \{1, \dots, T\} \Rightarrow \\ g_{i,1} &= \bar{y}_i - R_{i,1} - m \frac{C}{b} \text{ and } g_{i,t} = \bar{y}_i - m \frac{C}{K} - m \frac{C}{b}, t \in \{2, \dots, T\}. \end{aligned}$$

**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; the second-order conditions are trivially fulfilled. ■

Every coalition member invests more and consumes less if the coalition size,  $m$ , is large. The investment and abatement levels are at first-best if  $m = n$ , but they are otherwise too low. It is interesting to note that independently 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.** *The ratio  $d_{i,t}/R_{i,t} = x$  is first-best for every  $t \in \{2, \dots, T\}$ .*

(ii) *If we had  $m = n$  in every agreement, the outcome would be first-best for any  $T \in \{1, \dots, \infty\}$ .*

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

### *B. Duration of the Agreement*

While Proposition 3 holds for any  $T$ , no matter where it comes from, we can also ask for the equilibrium  $T$  when the countries can freely negotiate this length. 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 we said, in a Markov-perfect equilibrium, the equilibrium coalition,  $M^*$ , will be independent of any stock, history, or time. 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 at a stage when the actual coalition is  $M$  (potentially



$M \neq M^*$  out of equilibrium).<sup>21</sup>

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

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

From Proposition 4 we know 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 will form and negotiate the identical terms in period  $T + 1$ , generating the same payoffs to everyone, irrespective of the choice of  $T$ .

### C. Participation

We can now analyze the first stage of the game. When a country contemplates on whether or not to join the coalition, it anticipates the reaction function described in Proposition 4. In particular, if a country which is supposed to participate in equilibrium considers to deviate, 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 if  $i$  does not participate). This trade-off determines whether a country would like to join the coalition.

**Proposition 5.**  *$M^*$  is an equilibrium coalition if and only if  $m^* = |M^*| \in \{2, 3\}$ .*

**Proof.** 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:

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

---

<sup>21</sup>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  $M^* \subseteq M$ .

Substituting expression (7.1) for  $v(m^*, T^*)$ , this condition can be written as:

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

Simplified, this becomes

$$\left( \frac{m^{*2}}{2} - m^* \right) \left( \frac{1}{b} + \frac{\delta}{K} \right) \geq \left( m^{*2} - 3m^* + \frac{3}{2} \right) \left( \frac{1}{b} + \frac{\delta}{K} \right), \quad (3.1)$$

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; a contradiction. ■

The result is dismal. Even with extremely patient players and high pollution externalities, equilibrium coalition size will be very small. The gain from signing an IEA is the fact that the other coalition members will take the entrant's environmental 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 the coalition.<sup>22</sup>

Recall that a special case of our model is the workhorse model with one-period and without 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, investments, as well as emissions and if countries can contract on all these choices for any length of time.

## 4. Incomplete Contracts

As discussed in the Introduction, real climate negotiations have mainly focused on emission levels, leaving the investment decisions to individual countries. We therefore now relax the assumption that the environmental policy is fully contractible, and we assume that, while countries can commit to pollution levels, they cannot commit to specified

---

<sup>22</sup>The reason that the discount factor doesn't help in obtaining a larger IEA is intuitive: as  $\delta$  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 of the IEA remains small even as  $\delta \rightarrow 1$ .

levels of green investments. We study how investments are influenced by the negotiated emission quotas, how the emission quotas are decided taking into account the effect on investments, and how the contract incompleteness influences the equilibrium length of the contracts as well as coalition size.

As in the previous section, a pure strategy equilibrium can be described as a fixed 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 participants collectively choose the length  $T$  of the IEA and commit to  $g_{i,t}$  for every  $i \in M$  and  $t \in \{1, 2, \dots, T\}$  at the start of period 1; the level of investments, however, are 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 (2.5)-(2.6). 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:

$$R_{i,t} = \frac{b(\bar{y}_i - g_{i,t})}{b + K}, t \in \{2, \dots, T\}, \text{ but } R_{i,T+1} = \frac{C}{K}.$$

(ii) Taking these functions into account, the equilibrium emission levels satisfy:

$$\begin{aligned} g_{i,1} &= \bar{y}_i - R_{i,1} - m\frac{C}{b} \text{ and } g_{i,t} = \bar{y}_i - m\frac{C}{K} - m\frac{C}{b}, t \in \{2, \dots, T\} \\ \Rightarrow R_{i,t} &= m\frac{C}{K} > R_{i,T+1} = \frac{C}{K}, t \in \{2, \dots, T\} \\ \Rightarrow d_{i,t} &= m\frac{C}{b}, t \in \{1, \dots, T\}. \end{aligned}$$

Part (i) first states that  $i$  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

$d_{i,T+1}$ 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 hold-up 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  $t \in \{1, \dots, T\}$ . 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, when we take  $T$  and  $M$  as given.

**Corollary to Proposition 6.** (i) *The ratio  $d_{i,t}/R_{i,t} = x$  is first-best for every  $t \in \{2, \dots, T\}$ .*

(ii) *If we had  $m = n$  in every agreement, the outcome would be first-best 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 cannot be negotiated. If we had  $m = n$  and  $T$  finite, complete contracts would lead to the first-best while incomplete contracts would not.

### *B. Duration of the Agreement*

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

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

periods, where:

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

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

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 hold-up 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 none-the-less preferred. The threshold making the coalition indifferent,  $\widehat{m}(x, m^*)$ , is thus strictly smaller than  $m^*$ .

Proposition 7 allows to predict what happens in equilibrium if a country that is expected to join the IEA chooses to opt out instead. A particularly important case is when such a unilateral deviation induces a short agreement with  $T = 1$ . By Proposition 7, we are in this case when  $m^* - 1 \leq \widehat{m}(x, m^*)$ . The following Corollary characterizes this condition.

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

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

It is easy to see why the threshold  $m_M(x)$  plays a special role in the characterization of Markov equilibria. As we have seen from Proposition 4, if a country deviates from an equilibrium with complete contracts by staying out of an IEA, then the length of the agreement is  $T = 1$ . The temptation in that case consists of the difference between the utility of being in coalition of size  $m^*$  for one period and the utility of free-riding by opting out of the IEA and letting the others form a coalition of size  $m^* - 1$  for one period; after that period the agent receives the same expected continuation payoff regardless of the deviation. If  $m^* > m_M(x)$  the temptation with incomplete contracts consists of the

difference between the utility of being in a coalition of size  $m^*$  forever and the utility of free-riding by staying out of the IEA and letting the others form a coalition of size  $m^* - 1$  forever. When  $m^* > m_M(x)$ , therefore, the temptation with incomplete and incomplete contracts have the same sign and differ only by a factor of proportionality  $1/(1 - \delta)$ . Since we are not able to sustain an IEA larger than 3 with complete contracts, the same would be true with incomplete contracts when  $m^* > m_M(x)$ . In order to obtain a large coalition we will need a harsher punishment which can only be achieved if  $m^* \leq m_M(x)$ . We will refer to this as the *discipline constraint*.

### C. Participation

To induce a stable IEA in equilibrium we need to satisfy two incentive constraints: a *participation constraint*, so that any country in the IEA is not tempted to stay out of it; and a *non-participation constraint*, so that countries outside the IEA are not tempted to join. The latter constraint is satisfied whenever  $m^* > 1$ , so we will focus on the first.

The larger is a coalition, the more tempting it is for its members to opt out, because larger coalitions require larger reductions in pollution from their members. Participation in an IEA of size  $m$ , therefore, is incentive compatible only if  $m$  is not too large. The upper bound on  $m$ , below which the participation constraint holds, depends on whether the discipline constraint is satisfied or not: when the discipline constraint is satisfied, the punishment for not joining the IEA is higher and so the participation constraint can be satisfied for a higher  $m^*$ . 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  $m^* = |M^*| \in \{2, 3\}$  or  $3 < m^* \leq \min(n, m(x))$ , where*

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

with

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

and

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

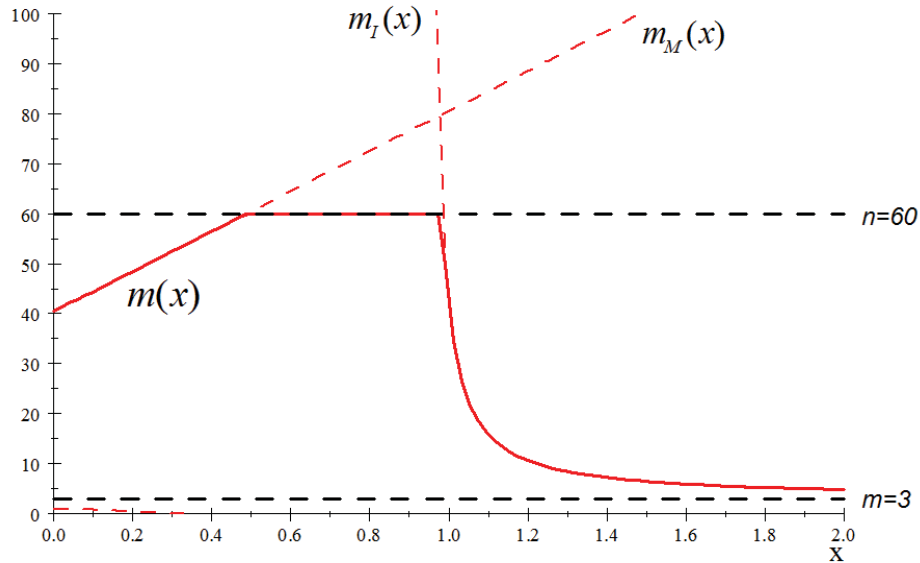


Figure 4.1: *The coalition size  $m^*$  must be below all three curves*

Just as before, we do have equilibria where 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 4.1 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 3 countries, which is the upper bound with complete contracts. In the example in Figure 4.1 there is an interval for  $x$  in which all countries choose to join the IEA and thus the outcome is efficient.

As the propositions presented above make clear a key determinant of the size of the IEA is the relative cost of technology,  $x$ . This variable has interesting but ambiguous effects on the 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 under-investment 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: its incentive 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 hold-up problem is, in any case, less important: the individual incentive 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

Proposition 8 makes clear that in an incomplete contracting environment we can always sustain at least an IEA with 3 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  $m^* \leq m_M(x)$  requires:

$$x \geq \underline{x}(\delta, m^*) \equiv \frac{\left(\frac{m^*-2}{m^*-1}\right)^2 - \delta}{1 - \left(\frac{m^*-2}{m^*-1}\right)^2}. \quad (4.1)$$

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

$$x \leq \bar{x}(\delta, m) \equiv \delta + \frac{2\delta}{m-3}. \quad (4.2)$$

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:

$$\underline{x}(\delta, m^*) \leq x \leq \bar{x}(\delta, m^*). \quad (4.3)$$

Since utilitarian welfare is increasing in  $m^*$ , (4.3) allows us to characterize when a coalition of size  $m > 3$  is feasible and, therefore, when the best Markov equilibrium with incomplete contracts is strictly superior to the best equilibrium with complete contracts. Expression (4.3) also allows us to characterize when a coalition of size  $m = n$  is feasible and, thus, when the best Markov equilibrium with incomplete contracts achieves the first-best outcome.

**Proposition 9.** *(i) The size of the coalition and the associated utilitarian welfare are always weakly larger in the best equilibrium with incomplete contracts than in the best equilibrium with complete contracts.*



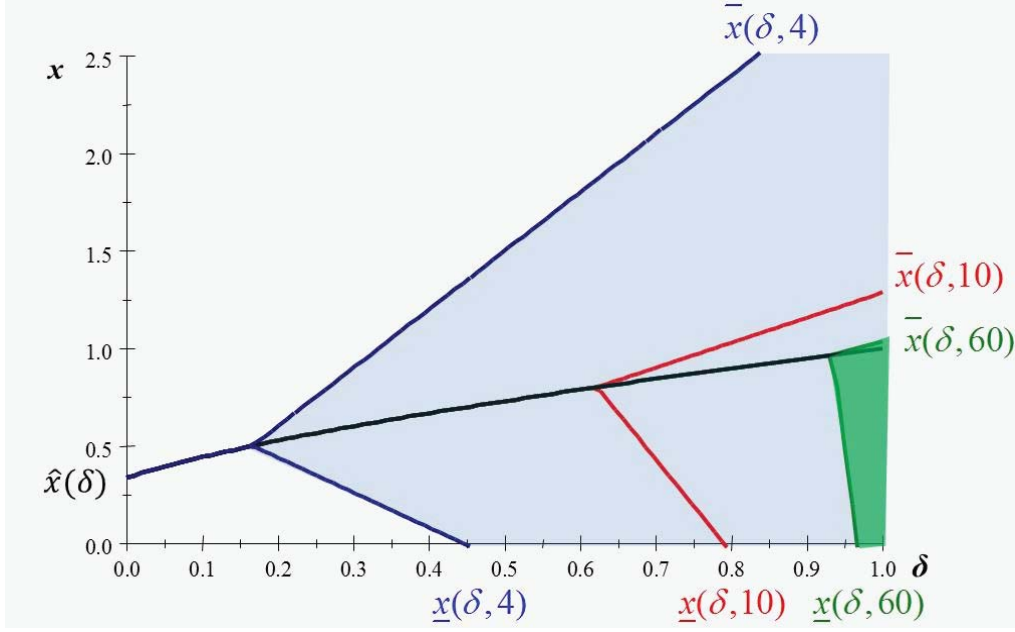


Figure 4.2: *Feasible IEAs.*

(ii) They are strictly larger in the best equilibrium with incomplete contracts if and only if:

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

a set that is non-empty if  $\delta \geq 1/6$ .

(iii) Moreover, for any  $n$ , the best equilibrium with incomplete contracts is first-best if and only if:

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

a set that is non-empty if  $\delta \geq \frac{(n-2)(n-3)}{n(n-1)}$ .

The conditions in parts (ii) and (iii) are illustrated in Figure 4.2. The figure plots  $\underline{x}(\delta, n)$ ,  $\bar{x}(\delta, n)$  and  $\hat{x}(\delta)$ , where  $\hat{x}(\delta)$  is the locus of the intersection of the first two curves.<sup>23</sup> The lightly shaded area in the figure describes the region of the parameter space in which an equilibrium coalition size can be larger than 3, so that the IEA is strictly superior with incomplete than with complete contracts.<sup>24</sup> The darkly shaded area corresponds to the

<sup>23</sup>We have  $\bar{x}(\delta, n) \geq \underline{x}(\delta, n)$  (respectively,  $\bar{x}(\delta, n) \leq \underline{x}(\delta, n)$ ) for  $x \geq \hat{x}(\delta)$  (respectively,  $x \leq \hat{x}(\delta)$ ).

<sup>24</sup>The sum of utilities is larger with incomplete contracts than with complete contracts. However, it is

region in which there is efficient equilibrium when  $n = 60$ . This region is non-empty 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  $\delta$ , then it remains feasible for any  $\delta' > \delta$ : the more patient the agents are, the larger the set of parameters that support a given coalition size is. However, an efficient outcome is not always possible, even if  $\delta$  is arbitrarily large. From (iii) we can see that if  $x > \bar{x}(1, n) = \frac{n-1}{n-3}$ , then there is no  $\delta \leq 1$  such that all countries find it optimal to join the IEA.<sup>25</sup>

A key assumption in the above analysis is that the contract length can be endogenously negotiated among participating countries. If the duration  $T$  were instead exogenous, it is easy to show that the equilibrium number of participants would be  $m^* \in \{2, 3\}$  regardless of the contractual environment. Since the incomplete contracts generate under-investments in period  $T$ , the complete contracting environment would strictly Pareto dominate the incomplete contracting environment for any fixed  $T < \infty$ . Thus, Proposition 9 would be over-turned.

## 5. 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.

### A. Commitment and Time Horizon

In the preceding analysis we have assumed that countries in the IEA can commit to a policy for an arbitrary amount of time. Indeed, we have seen that, in equilibrium, they find it optimal to commit for an infinite number of periods. How does the analysis change when we consider environments in which countries can only commit for a finite number of

---

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 where it would not be a member, and this might be the expected outcome if contracts were complete.

<sup>25</sup>We have chosen to emphasize the effects of  $x$  rather than the impact of  $\delta$  since the discount factor has multiple interpretations (as patience or period-length, for example). The alternative interpretations would have conflicting implications for how one should change the model's other parameters when  $\delta$  changes.

periods? An interesting benchmark is the case in which countries cannot commit for more than  $\bar{T}$  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 3. In an incomplete contracting environment we can show that an equilibrium exists if and only if:

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

where  $\underline{x}(\delta, m^*)$  is defined as in Corollary 1 to Proposition 8 and:

$$\bar{x}(\delta, m^*, \bar{T}) = \delta \left( 1 - \frac{(1 - \delta) \delta^{\bar{T}-1}}{1 - \delta^{\bar{T}}} \right) \frac{m^* - 1}{m^* - 3}. \quad (5.1)$$

The analysis is therefore as in the previous section, 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 explanation is that if  $i \in M^*$  deviates by not participating, the hold-up problem is moved forward from  $\bar{T}$  to the current period. If  $\bar{T}$  is small, this "penalty" is small so the incentive constraint strengthens and, to satisfy it,  $x$  must be smaller. However, as it can be easily verified from (5.1), in an incomplete contracting environment quite large coalitions are easily feasible even when the expected length of the agreement is short. Naturally, the upper bound converges to  $\bar{x}(\delta, m^*)$  as  $\bar{T} \rightarrow \infty$ .

### *B. Noncooperative Bargaining in an IEA*

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 ( $\sum_N z_i = 0$ ).<sup>26</sup> Each country has the same

---

<sup>26</sup>In Baron and Ferejohn's bargaining model countries are allowed to make monetary transfers among

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 chosen with the non-cooperative bargaining protocol described above. The intuition behind this result is as follows.<sup>27</sup> Take the problem faced by a country selected to propose an IEA. For simplicity, consider only the case with incomplete 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.<sup>28</sup> 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:

$$\max_{g_{j,t}, t_j, T} \left\{ \begin{array}{l} \sum_{l=t}^{t+T} \delta^{l-t} \left( u_l(g_{j,l}) - c \sum_{j \in M} g_{j,l} \right) - \sum z_j \\ \text{s.t. } z_j + \sum_{l=t}^{t+T} \delta^{l-t} \left( u_l(g_{j,l}) - c \sum_{j \in M} g_{j,l} \right) + \delta^T v_j \geq V_j(M) \end{array} \right\}, \quad (5.2)$$

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 (5.2) 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). Without loss of generality, assuming that the inequality holds as an equality, we have:

$$z_j = V_j(M) - \sum_{l=t}^{t+T} \delta^{l-t} \left( u_l(g_{j,l}) - c \sum_{j \in M} g_{j,l} \right) - \delta^T v_j. \quad (5.3)$$

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.

<sup>27</sup>Proofs for this result, and the other results in this section, are available from the authors.

<sup>28</sup>Formally  $u_l(g_{j,l})$  is equal to  $-\frac{b}{2} (Y_{i,l} - g_{i,l} - R_{i,l})^2$  for  $l = t$ , where  $R_{i,t}$  is taken as given from the previous period; to  $-\frac{b}{2} (Y_{i,l} - g_{i,l} - R_{i,t}(g_{i,l}))^2 - \frac{K}{2} R_{i,l} (g_{i,l})^2$  for  $l = t+1, \dots, T-1$  where  $R_{i,l}(g_{i,l})$  is given by Proposition 4.1; 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  $l = T$ .

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 (5.2) as:

$$\max_{g_{j,l}, T} \left\{ \sum_{j \in M} \sum_{l=t}^{t+T} \delta^{l-t} \left( u_l(g_{j,l}) - c \sum_{j \in M} g_{j,l} \right) + \delta^T v_j \right\}, \quad (5.4)$$

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 are expecting a utilitarian solution with no transfer, their expected continuation is  $V_j(M) = \frac{1-\delta^{T^*}}{1-\delta} \left( u_l(g_{j,l}^*) - c \sum_{j \in M} g_{j,l}^* \right) + \delta^{T^*} v_j$ , where  $g_{j,l}^*, T^*$  is the solution of (5.4). Condition (5.3) then implies that  $z_j = 0$ . Therefore, the cooperative solution assumed in Sections 3 and 4 is an outcome of this non-cooperative bargaining.

### C. Heterogenous Investment Cost

Symmetry is obviously a strong assumption when discussing international agreements of any sort. We have permitted only a few types of heterogeneity (i.e., the energy satiation point and the initial technology stock). In reality, countries are also likely to face different environmental costs as well as investment costs. Since the latter type of heterogeneity is analyzed in the related model of Harstad (2012a), here we describe how our results are likely to generalize in that setting.

When investments are private (without spillovers) and the countries have committed to a long-term agreement, then each country will invest optimally irrespective of any heterogeneity regarding the investment cost. At the negotiation stage, however, it is efficient that the countries with low investment costs receive a smaller emission quota, but this may be acceptable only with some accompanying side transfers. So, with this type of heterogeneity, side transfers between the coalition members may be necessary to achieve efficiency.

In the above analysis the exact composition of the equilibrium coalition,  $M^*$ , was not unique. With some types of heterogeneity, it may be natural to order the countries according to their benefit from an IEA, and this ordering of countries may serve as a

natural focal point for the composition of the equilibrium coalition, suggesting that the coordination problem (of who should participate and who should not) is a problem that is likely to be overcome in reality.<sup>29</sup>

#### *D. Technological Spillovers*

Another strong assumption in the model is that technological investments are private (i.e. they do not generate positive externalities on the other countries). This was important for the claim that, when the agreement lasts forever, participants invest optimally, given the allocated emission quota. With technological spillovers, however, investments are likely to be too small since an investing country will not take into account the direct benefit of neighboring countries, once the neighbors' emission quotas are fixed. This implies that incomplete contracts will always result in insufficient investment. With complete contracts, on the other hand, the coalition would be able to specify that each member invests more, by accounting for the positive externality on other coalition members. This could lead to a new trade-off when comparing the two contractual environments: while incomplete contracts may lead to larger coalitions, they also result in less investment compared to the coalition's optimum. We conjecture that the larger the technological spillover, the more likely that complete contracts dominate incomplete ones.

While technological spillovers are discussed by Harstad (2012a), Harstad (2012b) discusses renegotiation (none of these papers endogenizes the coalition size). With renegotiation, the coalition can easily force each member to take into account the technological spillover, even if investments are noncontractible: the coalition can simply specify very low emission quotas for each member, forcing each to invest quite a lot and, after the investments have been made, these quotas will be relaxed and renegotiated to a lower, optimal, level. With such renegotiation, therefore, our conjecture is that our main result still holds, just as in our basic workhorse model.

---

<sup>29</sup>Regarding the most efficient equilibrium, it is not clear whether the equilibrium coalition should include the countries that have high or low cost of investing in green technology. Instead, what seems to be of importance is how much a country is likely to invest as a member of the coalition compared to the same country's investment level as a nonparticipant. We conjecture that the larger this difference is (i.e., the larger the elasticity of the investment level with respect to the allocated emission quota is), the more likely that the country is a member of the welfare-maximizing equilibrium coalition.

### *E. Trade in Permits*

We have simplified the discussion above by not considering trade in pollution permits within the coalition. However, allowing for such trade would leave the results unchanged. For the complete contracting environment, this should be obvious: the coalition is then specifying identical  $R_{i,t+1}$ s and  $g_{i,t}$ s for every country, implying that the  $d_{i,t}$ s are identical, and so are the marginal benefits ( $bd_{i,t}$ ) of being allowed to consume more fossil fuel. Trade in permits has thus no value and allowing for such trade has no impact on the contract or on participation.

If investments are non-contractible, then the benefit of investing is different if permits can be traded. Increasing  $r_{i,t}$  is then reducing the need to pollute and country  $i$  can then sell permits. The willingness to pay is only  $bd_{j,t}$  for country  $j$ , say, and  $i$ 's optimal investment level implies  $KR_{i,t} = bd_{j,t}$ , which is satisfied also in Section 4. Allowing for trade will thus not change the equilibrium investments as a function of  $g_{i,t}$ s, and the optimal choice of the  $g_{i,t}$  is thus unchanged. Investments in period  $T$  will, as before, be sunk in the following period and the only effect of  $r_{i,T}$  is a reduction in  $g_{i,T+1}$ , leaving every  $d_{j,T+1}$  unchanged.

## **6. Conclusion**

This paper detects and analyzes novel and intriguing relationships between coalition size, contract length, and contract incompleteness. The results are important for designing environmental agreements. While critics have suggested that the UN's 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 — 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 on emission levels and not as well on investments. However, since our results are consequential and perhaps provocative, we suggest that future research should extend the model in a number of directions to further investigate the results' robustness in detail. Our model should fit well as a workhorse for such extensions as it has proven to be both simple and tractable.



## References

- [1] Aldy, Joseph and Stavins, Robert (Ed.) (2007): *Architectures for Agreement*, Cambridge U. Press.
- [2] Aldy, Joseph and Stavins, Robert (Ed.) (2009): *Post-Kyoto International Climate Policy*, Cambridge U. Press.
- [3] Barrett, Scott (1994): "Self-enforcing international environmental agreements," *Oxford Economic Papers*, 46, p. 878-94.
- [4] Barrett, Scott (1997): "The strategy of trade sanctions in international environmental agreements," *Resource and Energy Economics* 19, p. 345-61
- [5] Barrett, Scott (2002): "Consensus Treaties," *Journal of Institutional and Theoretical Politics*, 158: 519-41."
- [6] Barrett, Scott (2003): *Environment and Statecraft*, Oxford University Press.
- [7] Barrett, Scott (2005): "The Theory of International Environmental Agreements," *Handbook of Environmental Economics* 3, edited by K.-G. Mäler and J.R. Vincent.
- [8] Barrett, Scott (2006): "Climate Treaties and "Breakthrough" Technologies," *American Economic Review, papers and proceedings* 96 (2): 22-5.
- [9] Battaglini, Marco, Nunnari, Salvatore and Palfrey, Thomas (2012a), "Legislative Bargaining and the Dynamics of Public Investment," *American Political Science Review*, in press.
- [10] Battaglini, Marco, Nunnari, Salvatore and Palfrey, Thomas (2012b), "The Dynamic Free Rider Problem: A Laboratory Study," in preparation.
- [11] Beccherle, Julien and Tirole, Jean (2011): "Regional Initiatives and the Cost of Delaying Binding Climate Change Agreements," *Journal of Public Economics* 95(11-12): 1339-48.
- [12] Bosello, Francesco, Buchner, Barbara and Carraro, Carlo (2003): "Equity, Development, and Climate Change Control," *Journal of the European Economic Association* 1: 601-11.
- [13] Buchholz, Wolfgang and Konrad, Kai (1994): "Global Environmental Problems and the Strategic Choice of Technology," *Journal of Economics* 60 (3): 299-321.
- [14] Calvo, Emilio and Rubio, Santiago (2012): "Dynamic Models of International Environmental Agreements: A Differential Game Approach", mimeo.
- [15] Carraro, Carlo, Eyckmans, Johan, and Finus, Michael (2006): "Optimal Transfers and Participation Decisions in International Environmental Agreements," *The Review of International Organizations* 1(4): 379-96.

- [16] Carraro, Carlo and Siniscalco, Domenico (1993): "Strategies for the international protection of the environment," *Journal of Public Economics* 52 (3): 309-28.
- [17] Chwe, Michael (1994): "Farsighted Coalitional Stability," *Journal of Economic Theory* 63: 299-325.
- [18] d'Aspremont, Claude., Jacquemin, Alexis, Gabszewicz, Jean Jaskold and Weymark, John A. (1983): "On the stability of collusive price leadership," *The Canadian Journal of Economics* 16(1): 17-25.
- [19] de Zeeuw, Aart (2008): "Dynamic effects on the stability of international environmental agreements," *Journal of Environmental Economics and Management* 55(2): 163-74.
- [20] Dixit, Avinash and Olson, Mancur (2000): "Does voluntary participation undermine the Coase Theorem?" *Journal of Public Economics* 76 (3): 309-35.
- [21] Dutta, Prajit K. and Radner, Roy (2004): "Self-enforcing climate-change treaties," *Proc. Nat. Acad. Sci. U.S.*, 101: 4746-51.
- [22] Ellingsen, Tore and Paltseva, Elena (2011): "The Private Provision of Excludable Public Goods: An Inefficiency Result," forthcoming, *Journal of Public Economics*.
- [23] Finus, M., and S. Maus (2008): "Modesty may pay!" *Journal of Public Economic Theory* 10: 801-26.
- [24] Guriev, Sergei and Kvasov, Dmitriy (2005): "Contracting on Time," *The American Economic Review* 95 (5): 1269-1385.
- [25] Harstad, Bård (2012a): "Climate Contracts: A Game of Emissions, Investments, Negotiations, and Renegotiations," *Review of Economic Studies* 79(4): 1527-57.
- [26] Harstad, Bård (2012b): "The Dynamics of Climate Agreements," mimeo.
- [27] Heal, Geoffrey, Kunreuther, H. (2011): "Tipping Climate Negotiations," NBER Working Paper No 16954.
- [28] Hoel, Michael (1992): "International environmental conventions: the case of uniform reductions of emissions," *Environmental and Resource Economics* 2(2): 141-159.
- [29] Hoel, Michael and Schneider, Kerstin (1997): "Incentives to Participate in an International Environmental Agreement," *Environmental and Resource Economics*, 9: 153-70.
- [30] Hoel, Michael and de Zeeuw, Aart (2010): "Can a Focus on Breakthrough Technologies Improve the Performance of International Environmental Agreements?" *Environmental and Resource Economics* 47 (3): 395-406.
- [31] Hong, Fuhai and Karp, Larry (2010): "International Environmental Agreements with Mixed Strategies and Investment," mimeo, UC Berkeley.

- [32] Karp, Larry and Simon, Leo (2012): "Participation games and international environmental agreements: a nonparametric model," forthcoming *Journal of Environmental Economics and Management*.
- [33] Kolstad, Charles D. and Toman, Michael (2005): "The Economics of Climate Policy," *Handbook of Environmental Economics* 3: 1562-93.
- [34] Nordhaus, William D. (2006): "After Kyoto: Alternative Mechanisms to Control Global Warming," *American Economic Review* 96(2): 31-4.
- [35] Palfrey, Thomas R. and Rosenthal, Howard (1984): "Participation and the provision of discrete public goods: a strategic analysis," *Journal of Public Economics* 24(2): 171-93.
- [36] Ploeg, Frederick van der, and de Zeeuw, Aart (1992): "International aspects of pollution control," *Environmental and Resource Economics* 2 (2): 117-39.
- [37] Rey, Debraj and Vohra, Rajiv (2001): "Coalitional Power and Public Goods," *Journal of Political Economy* 109 (6): 1355-84.
- [38] Rubio, Santiago J., and Casino, Begona (2005): "Self-enforcing international environmental agreements with a stock pollutant," *Spanish Economic Review* 7 (2): 89-109.
- [39] Rubio, Santiago J., and Ulph, Alistar (2007): "An infinite-horizon model of dynamic membership of international environmental agreements," *Journal of Environmental Economics and Management* 54 (3): 296-310.
- [40] Schmidt, Robert C. (2012): "Early investments in low-carbon technologies and the success of climate cooperation," mimeo.
- [41] Segal, Ilya and Whinston, Michael D. (2010): "Property Rights," prepared for *Handbook of Organizational Economics* (Gibbons and Roberts, Eds), Elsevier.
- [42] UNFCCC (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," United Nations FCCC/CP/2010/7/Add.1.
- [43] UNFCCC (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," United Nations FCCC/CP/2011/9/Add.1.
- [44] Vespa, E. (2012), "Cooperation in Dynamic Games: An Experimental Investigation," mimeo, UCSB.

## 7. Proofs

### 7.1. 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):

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

This implies:

$$v(m^*, T^*) = -\frac{1}{1 - \delta} C \left[ \bar{y}_i - C \left( \frac{m^{*2}}{2} + n - m^* \right) \left( \frac{1}{b} + \frac{\delta}{K} \right) \right], \quad (7.1)$$

and therefore:

$$\begin{aligned} v(m, T) &= -\frac{1 - \delta^T}{1 - \delta} C \left[ \bar{y}_i - C \left( \frac{m^2}{2} + n - m \right) \left( \frac{1}{b} + \frac{\delta}{K} \right) \right] \\ &\quad - \frac{\delta^T}{1 - \delta} C \left[ \bar{y}_i - C \left( \frac{m^{*2}}{2} + n - m^* \right) \left( \frac{1}{b} + \frac{\delta}{K} \right) \right] \end{aligned} \quad (7.2)$$

Note that the derivate of  $v(m, T)$  w.r.t.  $T$ , or equivalently w.r.t.  $-\delta^T$ , 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  $m \leq m^*$ . ■

### 7.2. Proof of Proposition 6

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

$$\begin{aligned} v_i &= \sum_{t=1}^T \delta^{t-1} \left[ -\frac{b}{2} (\bar{y}_i - g_{i,t} - R_{i,t})^2 - C \left( \sum_{j \in N} g_{j,t} \right) - \delta \frac{K}{2} R_{i,t+1}^2 \right] \\ &\quad + \delta^T v_i + \delta^T C \sum_{j \in N} R_{j,T+1}. \end{aligned} \quad (7.3)$$

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).<sup>30</sup>

It follows that the first-order conditions for the  $R_{i,t}$ s are:

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

This implies:

$$R_{i,t} = \frac{\bar{y}_i - g_{i,t}}{K/b + 1} \Rightarrow \bar{y}_i - g_{i,t} - R_{i,t} = \frac{K}{b} \frac{\bar{y}_i - g_{i,t}}{K/b + 1}, \quad t \in \{2, \dots, T\}. \quad (7.4)$$

(ii) Substituting (7.4) into (7.3) 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 w.r.t.  $a_{i,t} = a_t$ ,  $t \in \{2, \dots, T\}$  gives:

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

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

### 7.3. 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  $T^* < \infty$ . First note that in Markov equilibrium the decision to join a coalition is stationary, so the continuation value for a participant can be written recursively as:

$$\begin{aligned} v(m^*, T^*) &= -\frac{1 - \delta^{T^*}}{1 - \delta} C \left[ \bar{y}_i - C \left( \frac{m^{*2}}{2} + n - m^* \right) \left( \frac{1}{b} + \frac{\delta}{K} \right) \right] \\ &\quad - \delta^{T^*} \frac{C^2}{2K} (m^* - 1)^2 + \delta^{T^*} v(m^*, T^*) \end{aligned} \quad (7.5)$$

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

<sup>30</sup>Also, note  $v_i$  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 (7.3).

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 (7.5) implies that:

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

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. ■

We can now prove Proposition 7. Given Lemma A.1, the value of a  $T$ -period agreement for each member of a coalition of size  $m$  is:

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

Note that the derivate of  $v(m, T)$  w.r.t.  $T$ , or equivalently w.r.t.  $-\delta^T$ , is always negative if and only if:

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

That is, after some algebra, if and only if  $m \leq \hat{m}(x)$ , as defined in Proposition 7. ■

#### 7.4. Proof of Proposition 8

Suppose  $m^* \leq \bar{m}_M$ . If a country which 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:

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

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

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

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

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

After some algebra, this inequality reduces to:

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

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:

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

which is implied by  $m^*(m^* - 2) \geq 0$ , or  $m^* \geq 2$ , which is always satisfied. From Proposition 7 we can conclude that an equilibrium of size  $m^* \in [2, n]$  exists if  $2\frac{\delta}{x} \geq (m^* - 3) \left(1 - \frac{\delta}{x}\right)$  and  $m^* \leq 1 + \frac{1}{1 - \sqrt{\frac{x+\delta}{x+1}}}$  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 2 steps.

**Step 1.** Assume  $m^* = 2$ . In this case  $m^* \leq 1 + \frac{1}{1 - \sqrt{\frac{x+\delta}{x+1}}}$  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  $m^* \geq 3 + \frac{2\delta}{x-\delta}$ , that is if  $x \geq \delta + \frac{2\delta}{m^*-3} = -\delta$ , which is always true. If  $m^* = 3$ , condition  $2\frac{\delta}{x} \geq (m^* - 3) \left(1 - \frac{\delta}{x}\right)$  is always true. Condition  $m^* \leq 1 + \frac{1}{1 - \sqrt{\frac{x+\delta}{x+1}}}$ , is true if  $x > \frac{\left(\frac{m^*-2}{m^*-1}\right)^2 - \delta}{1 - \left(\frac{m^*-2}{m^*-1}\right)} = \frac{1/4 - \delta}{3/4} = \frac{1}{3} - \frac{4}{3}\delta$ . Assume  $x < \frac{1}{3} - \frac{4}{3}\delta$ . 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:

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

Note that this inequality is the same as (3.1) 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 be 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 (4.1) and (4.2). We therefore need to consider only 3 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 1 to infinity. The deviation of agent  $i$  is profitable if:

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

Note that:

$$\begin{aligned} v'(m^* - 1, T') &= -\frac{1}{1-\delta} C \left[ -\frac{C}{2b} + \bar{y}_i - C \left( \begin{array}{c} (m^* - 1)^2 + n \\ -(m^* - 1) \end{array} \right) \frac{1}{b} \right. \\ &\quad \left. + \frac{C\delta}{2K} - n \frac{C}{K} \right] \\ &\quad + \frac{\delta}{1-\delta} \frac{C^2}{K} ((m^* - 1)^2 - (m^* - 1)) \frac{C^2 \delta}{K} \\ &\quad - \frac{\delta^{T'}}{1-\delta^{T'}} \frac{C^2}{K} (m-2)(m-1) \\ &= -\frac{1}{1-\delta} C \left[ -C \left( (m^* - 1)^2 + n - (m^* - 1) \right) \frac{1}{b} + \frac{C\delta}{2K} - n \frac{C}{K} \right] \\ &\quad + \left( \frac{\delta}{1-\delta} - \frac{\delta^{T'}}{1-\delta^{T'}} \right) \frac{C^2}{K} (m-2)(m-1). \end{aligned} \tag{7.6}$$

The right-hand side of (7.6) 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 (7.6) 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. ■