CESifo
AREA CONFERENCES 2020

Energy & Climate Economics

Munich, 1–2 October 2020

A Dynamic Analysis of International Agreements: Loose Versus Tight Coalitions

Luca Colombo, Paola Labrecciosay, and Ngo Van Long

CESifo

A dynamic analysis of international agreements: Loose versus tight coalitions

Luca Colombo[*] Paola Labrecciosa[†] Ngo Van Long[‡]

September 23, 2020

Abstract

We study equilibrium membership of an international agreement aimed to increase the stock of a global public good in a continuous-time game where, at each point in time, countries make two sequential decisions: participation in the agreement and contribution to the public good. We depart from the existing literature by assuming partial rather than full cooperation among participating countries. We prove existence and uniqueness of equilibrium membership and contribution to the public good and study, among other things, how a decrease in the coefficient of cooperation impacts on the equilibrium. We show that the coalition size can be large, thus challenging the conventional wisdom according to which, under full cooperation, only small coalitions can be stable. In contrast with previous studies, we also show that the coalition size can be increasing over time, even when coalitions are small. We highlight a novel trade-off between agreements that are *narrow-but-deep-and-long-lived* vs *broad-but-shallow-and-short-lived*. In the case of a time horizon that is sufficiently long, loose cooperative agreements, which are broad-but-shallow-and-short-lived, are both welfare- and Pareto-superior to tight cooperative agreements, which are narrow-but-deep-and-long-lived. Conditions exist under which the equilibrium coalition size is efficient.

JEL Classification: H41; C72; D60.

Keywords: differential games; voluntary provision of public goods; stable coalition; coefficient of cooperation; social welfare.

[*]Deakin Business School, Department of Economics, Burwood Campus, 221 Burwood Hwy, Burwood, 3125 VIC, Australia. Email: luca.colombo@deakin.edu.au.

[†]Corresponding author. Monash Business School, Department of Economics, Clayton Campus, Wellington Road, Clayton, 3800 VIC, Australia. Email: paola.labrecciosa@monash.edu.

[‡]McGill University, Department of Economics, 855 Sherbrook St. W., Montreal, Quebec, H3A 2T7. Email: ngo.long@mcgill.ca.

1 Introduction

Many public goods are funded predominantly through voluntary contributions. This is especially the case for global public goods such as climate change mitigation, widespread peace, financial stability, and global public health.[1] In the case of global public goods, international cooperation is needed for the attainment of an efficient outcome. However, experience from climate change policy indicates that full cooperation among all countries is hard to achieve.[2] Partial cooperation seems to be a more realistic prospect due to conflicting national interests, disagreements on what constitutes a fair burden, and a general distrust among countries.

Most of the literature on international agreements (e.g. Hoel, 1992; Carraro and Siniscalco, 1993; Barrett, 1994; Rubio and Ulph, 2007; Eichner and Pethig, 2013; Battaglini and Harstad, 2016; Karp and Sakamoto, 2019), assumes that signatory countries agree to "give up their own self interest" and act in a way to maximize the sum of all members' utilities.[3] A robust theoretical result in this literature is that the equilibrium coalition size is very small (except possibly when the potential benefits of cooperation are also small). This theoretical result is only partly supported by the fact that, in reality, proposed agreements that prescribe full cooperation (e.g. the Kyoto Protocol) often fail to attract a sufficient number of truly committed signatories: the pessimistic conclusion reached in the literature under the assumption of full cooperation that only agreements involving a very limited number of countries are stable (in the sense of self-enforcing) seems too pessimistic. Many real-world coalitions can be quite large.[4]

[1] Problems concerning voluntary provision of public goods have received a lot of attention in the literature. Classical references on voluntary provision of public goods include Chamberlin (1974), Morrison (1978), Bergstrom at al. (1986), Bernheim and Douglas (1986), Cornes and Sandler (1986), and Andreoni (1988). On the voluntary provision of public goods in dynamic settings see Fershtman and Nitzan (1991), Wirl (1996), Marx and Matthews (2000), Itaya and Shimomura (2001), Yanase (2006), Long and Shimomura (2007), Benchekroun and Long (2008), Fujiwara and Matsueda (2009), Battaglini et al. (2014), Georgiadis (2015, 2017), and Bowen et al. (2019), *inter alia*.

[2] For example, the Kyoto Protocol, which was adopted in 1997 and entered into force in February 2005, in its attempt to achieve a global agreement on the actions to take to reduce greenhouse gas emissions was a failure. Although it was signed by 178 countries, only a small number of countries were required to reduce emissions in the first phase. The second phase specified targets for 37 countries, but so far only seven have ratified.

[3] A variation to the standard setup is considered in Buchholtz et al. (2014). Specifically, this paper analyzes a model in which there are two different groups of countries. One is a coalition of like-minded cooperating countries whose members are mutually matching their public good provision, and the other consists of outsiders which - without any matching - act non-cooperatively playing Nash against the coalition.

[4] The inability of the theory of international cooperation to explain observed large coalition sizes is referred to

In the realm of climate change policy, it has recently been recognized that full cooperation could be less efficient than partial cooperation. A distinctly different approach from the "top-down" approach characterizing the Kyoto Protocol was taken at the Paris International COP21 Conference on Climate Change in 2015: countries agreed on an overall objective of limiting global warming to 2 degrees C relative to the pre-industrial temperature, but no country was required to set a specific target by a specific date. Indeed, unlike the Kyoto Protocol, the Paris Accord does not insist on the coordination of climate actions among signatories: instead of setting commitments through centralized bargaining, the "bottom-up" approach of the Paris Accord allows countries to make their own commitments. There were some indications that this form of loose agreement, by attracting more participants, could turn out to be more effective in reducing emissions than the Kyoto Protocol.[5]

Previous studies on international agreements seem to have been overoptimistic in assuming away all obstacles that are typically observed in international negotiations, reaching an overpessimistic conclusion. In this paper, we propose and analyze a full-fledged continuous-time game of voluntary provision of a global public good with endogenous number of contributors that departs from the existing literature in a fundamental way. Instead of assuming full cooperation, we assume partial cooperation among signatory countries: each signatory country agrees to maximize a *weighted* sum of utilities of all members rather than the sum of utilities of all members. In this respect, our approach shares some features with that taken by Harstad (2020a), who shows that each contribution to a public good maximizes an *asymmetric* Nash product where the weight on others' payoffs is smaller than in the Nash Bargaining Solution. We believe that, with the exception of those situations characterized by negligible conflicts among nations, partial cooperation is a better description of reality than full cooperation. We model partial cooperation by using a coefficient of

as a "paradox" by Kolstad and Toman (2005) and Nordhaus (2015). A solution to this paradox can be found in the presence of policy makers'reelection concerns (Battaglini and Harstad, 2020) or when parties are troubled by time inconsistency (Gerlagh and Liski, 2018; Harstad, 2020c). Other solutions to the above paradox are discussed in Rubio and Ulph (2006), allowing for Stackelberg leadership in the contribution stage, and in Finus and Maus (2008), considering a modesty parameter.

[5] The reactions of the stock markets after the Paris Agreement was reached could be one such indication. Renewable energy share prices rose after the Paris Agreement. The iShare Global Clean Energy Exchange Trade Fund rose by 1.4% and the MAC Global Solar Energy index rose by 1.9%. Stock prices of coal companies fell sharply (11.3% for Peabody Energy, 4.9% for Consol Energy Inc.). The U.S. oil and gas index dropped by 0.5%. See Mukanjari and Sterner (2018), van der Ploeg and Rezai (2019).

cooperation ranging from zero to one, zero corresponding to Nash behavior, and one corresponding to full cooperation (see Cyert and deGroot, 1973).[6] For every interior values, countries place a lower weight on the interest of others. As pointed out in Harstad (2020b) with reference to the Paris Accord: "Scholars and observers naturally expect a party's contribution to reflect that party's own interests to a larger extent, and the interests of other parties to a lower extent". Our modelling approach makes it possible to compare a Kyoto-style agreement, with a coefficient of cooperation close to one (meaning that countries are not allowed to place a lower weight on the interest of others) and a Paris-sytle agreement, with a coefficient of cooperation significantly lower than one (meaning that countries are allowed to place a lower weight on the interest of others).

The dynamic game at hand consists of a sequence of two-stage games. At each point in time there are two stages: (i) a participation stage; (ii) a contribution stage. In (i), each country chooses independently and noncooperatively whether or not to participate in an international agreement, for instance, an international environmental agreement aimed to tackle global warming, or an international agreement for global health security. We use the stability concept that is widely used in the literature on international environmental agreements (e.g. Rubio and Ulph, 2007; de Zeeuw, 2008).[7] In (ii), each signatory country contributes to the public good the amount that maximizes its own utility plus a fraction (the coefficient of cooperation) of all the other members' utilities; each non-signatory country behaves as a Nash player by contributing to the public good the amount that maximizes its own utility. Countries are free to join or leave the agreement at any point in time.

Our equilibrium concept is Markov Perfect Equilibrium. Countries condition their participation to time and the current stock of public good, and their contribution to the vector of participation decisions, time, and the current stock of public good. We assume symmetry among countries within the same group, either the group of signatories or the group of non-signatories. We prove equilibrium existence and uniqueness. If the time horizon is finite (resp. infinite) then the equilibrium size of the

[6] The assumption of full cooperation is relaxed also in Hoel and de Zeeuw (2014). They condider a three-stage game with a participation stage (Stage 1), an R&D stage (Stage 2), and an emission stage (Stage 3). They assume cooperation in Stage 2 but no cooperation in Stage 3.

[7] Rubio and Ulph (2007) and de Zeeuw (2008) build on earlier works on cartel stability in oligopolistic markets by d'Aspremont et al. (1983), and on earlier works on international environmental agreements by Carraro and Siniscalco (1993) and Barrett (1994). On the stability of oligopolistic cartels and international environmental agreements see also Diamantoudi (2005) and Diamantoudi and Sartzetakis (2015), respectively.

agreement increases (resp. is constant) over time. In equilibrium, signatory countries continuously contribute to the public good until a certain (endogenously determined) point in time is reached, beyond which an agreement among contributing countries cease to exist; non-signatory countries free-ride on the contributions made by the coalition. In a comparison with social optimum, we show that the equilibrium duration of the agreement is shorter than socially desirable, unless countries fully cooperate. Since participation in an international agreement takes place on a voluntary basis, when the agreement is "too demanding", in the sense that it specifies joint utility maximization or a behavior close to it, participation is weak. When, instead, the agreement specifies a low coefficient of cooperation, participation is strong. In this case, the equilibrium coalition size can be large. This is one of the key findings of our analysis.[8] Intuitively, when the agreement is such that a country can place a lower weight on the interest of others, it is not that costly for a country to participate, and this explains why the equilibrium coalition size is larger than in the case in which the agreement specifies joint-utility maximization. In the context of international environmental agreements, the trade-off between agreements that are narrow-but-deep vs. broad-but-shallow has been extensively studied (e.g. Schmalensee, 1998; Barrett, 2002; Aldy et al., 2003; Finus and Maus, 2008; Harstad, 2020b). We contribute to this stream of literature by showing how a change in the coefficient of cooperation impacts on equilibrium membership. To our knowledge, the coefficient of cooperation has not been considered in the literature prior to our paper, neither in static nor in dynamic models. Even more importantly, we add a third dimension, the time dimension, to the classical trade-off between agreements that are narrow-but-deep vs. broad-but-shallow, thus enriching the comparative analysis of international agreements. We highlight a novel trade-off between agreements that are *narrow-but-deep-and-long-lived* vs *broad-but-shallow-and-short-lived*. We show that loose cooperative agreements are broad-but-shallow-and-short-lived in that they imply strong participation but low cooperation and short duration, whereas tight cooperative agreements are narrow-but-deep-and-long-lived in that they imply weak participation but high cooperation and long duration. Our finding that small coalitions can be long-lasting is in contrast with previous studies on dynamic voluntary provision of public goods, which predict that agreements will be long-lasting if and only if the coalition is large (e.g. Battaglini and Harstad, 2016; Karp and Sakamoto, 2019; Kovác and Schmidt, 2019). In the case in which the time horizon is

[8] A similar result is obtained in Karp and Simon (2013). They develop a non-parametric approach and, keeping the assumption of full cooperation, they show that the conventional wisdom according to which the equilibrium coalition size is small is not robust. We reach a similar conclusion, but our approach is different.

sufficiently long, we demonstrate that loose cooperative agreements are superior to tight cooperative agreements in terms of discounted welfare. To our knowledge, this is a novel result in the literature.[9] The policy implications of this result are clear: international agreements aimed at maximizing the welfare of its members as well as global welfare should be designed in a way to give more weight to individual rather than collective rationality. Relying exclusively on collective rationality usually leads to very small coalitions and inefficient outcomes.

The remainder of this paper is organized as follows. The game theoretical model is laid down in Section 2. Sections 3 derives the social optimum. Section 4 characterizes the equilibrium of the game. Section 5 provides a welfare analysis. Section 6 concludes.

2 The Game

The game is specified in continuos time. Time is denoted by $t \in [0, T\rangle$. There are $n \geq 2$ *a priori* identical countries. At each $t \in [0, T\rangle$, countries make two sequential decisions: whether or not to participate in an international agreement aimed to increase the stock of a global public good (such as climate change mitigation), and how much to contribute to the public good. For all countries, participation decisions are made independently and noncooperatively; for each participating country, contribution levels are jointly decided by all participants, with the aim of maximizing the discounted value of a weighted sum of utilities of all participants, whereas non-participating countries decide how much to contribute independently and noncooperatively, each with the aim of maximizing its own discounted utility. The agreement exists until time $\widehat{T} \leq T$, with $\widehat{T}$ being *endogenously* determined. A country's decision to join the agreement at any $t_0 \in [0, \widehat{T}\rangle$ has to be rational at that time (in terms of discounted sum of utilities), but it may become irrational for a country to remain in the agreement at $t_1 \in [0, \widehat{T}\rangle$, with $t_1 > t_0$. Similarly, a country's decision to stay out of the agreement at any $t_0 \in [0, \widehat{T}\rangle$ has to be rational at that time; joining the agreement may become rational at $t_1 \in [0, \widehat{T}\rangle$, with $t_1 > t_0$. For the sake of simplicity, we assume that coun-

[9] Keeping the assumption of full cooperation, the possibility that a change that makes an international environmental agreement more efficient can lower welfare by lowering equilibrium membership is also discussed in Battaglini and Harstad (2016), who show that the ability of coalition members to contract on both investment and emissions lowers welfare relative to the case where members can contract only on emissions. In our paper, a change that makes an international agreement more efficient (for a given size) is represented by greater cooperation among coalition members.

tries are free to join or leave the agreement at any $t \in [0, \widehat{T}\rangle$ without having to pay any additional entry/exit cost. Let $m(t) \leq n$ denote the number of signatories at t ($n - m(t)$ denotes the number of non-signatories); signatories are denoted by the index $i = 1, ..., m$, and non-signatories by the index j. For each country, the utility is linear in the stock of public good, $K(t)$.[10] Without any loss in generality, we normalize the benefit of each unit of additional K to 1. The public good is entirely financed with countries' contributions, with $x_k(t)$ denoting country k's contribution at t, $k = i, j$. Each country's total cost of contributing is $cx_k(t)$, with $c > 0$. Further restrictions on c will be specified later on (see Assumption A2). The restrictions imposed on $x_k(t)$ are given in Assumption A1.

Assumption A1: $x_k(t) \in [0, 1]$ for all $t \in [0, T\rangle$.

A1 implies that each county's contribution cannot be negative and cannot exceed its maximum capacity, normalized to 1 throughout the game.

When country i participates in an international agreement, we distinguish its material payoff from its objective function. Its instantaneous material payoff is

$$u_i(t) = K(t) - cx_i(t), \tag{1}$$

and its instantaneous objective function is given by $v_i(t)$ as defined below

$$v_i(t) = u_i(t) + \phi_i \sum_{k \neq i}^{m(t)} u_k(t), \tag{2}$$

where $u_k(t)$ is defined analogously to (1) and $\phi_i \in [0, 1]$ denotes the coefficient of cooperation (see Cyert and deGroot, 1973), the weight that country i places on country k's utility when making decisions, with $i, k = 1, ..., m(t)$, $k \neq i$ (i.e. country k is a signatory country other than i).[11] Note that $\phi_i = 1$ means that country i fully internalizes the benefit that its contribution x_i confers on all other signatories. Therefore $\phi_i = 1$ *corresponds to the case of full internalization (joint utility maximization)*, whereas any $\phi_i \in (0, 1)$ describes partial internalization. When $\phi_i = 0$, countries act noncooperatively. We assume that $\phi_i = \phi$ for all $i = 1, ..., m(t)$, and that ϕ is specified by the

[10] This assumption is common in the literature on voluntary provision of public goods (or, equivalently, abatament of public bads). See, for instance, Barrett (1999), Hong and Karp (2012), and Battaglini *et al.* (2014).

[11] The coefficient of cooperation was called the coefficient of "effective sympathy" by Edgeworth (1881).

agreement.[12] Those countries entering an agreement contribute to the public good so as to maximize (2). For agreements specifying a high ϕ we use the expression **tight cooperative agreements**; for agreements specifying a low ϕ we use the expression **loose cooperative agreements**.

The instantaneous objective function of country j (a non-signatory) is given by

$$u_j(t) = K(t) - cx_j(t). \tag{3}$$

The main difference between (2) and (3) is that while signatories agree to cooperate (for any $\phi \in (0,1]$), non-signatories choose to act in isolation.

The evolution of the stock of public good is governed by the following differential equation:

$$\frac{dK(t)}{dt} = X(t), \; K(0) = K_0 \geq 0,$$

where $X(t)$ denotes the sum of all contributions at t.

Country i's contribution is specified by the agreement and solves the following problem:

$$\begin{cases} \max\limits_{x_i(t)} J_i = \int\limits_0^T e^{-rt} v_i(t)\, dt \\ s.t.\ x_i(t) \in [0,1] \text{ and } \frac{dK(t)}{dt} = x_i(t) + X_{-i}(t), \end{cases} \tag{4}$$

where $r > 0$ is the discount rate and X_{-i} the sum of all contributions except the contribution made by country i.

Country j's contribution solves the following problem:

$$\begin{cases} \max\limits_{x_j(t)} J_j = \int\limits_0^T e^{-rt} u_j(t)\, dt \\ s.t.\ x_j(t) \in [0,1] \text{ and } \frac{dK(t)}{dt} = x_j(t) + X_{-j}(t), \end{cases} \tag{5}$$

where X_{-j} denotes the sum of all contributions except the contribution made by country j.

We are interested in characterizing a Markov Perfect Equilibrium of the dynamic game at hand. We assume that strategies are of the Markovian type: countries condition their participation to time and the current stock of public good, and their contribution to the vector of participation decisions $\mathbf{p}(t) = \{p_1(t), ..., p_n(t)\}$, where $p_z \in \{in, out\}$, with $z = 1, ..., n$, time, and the current stock of public good. We consider strategies of the form: $p_z(t) = \psi_z(K,t)$, where ψ_z denotes the

[12] The usual assumption in the coalition literature is that $\phi = 1$, i.e. each member fully takes into account the welfare of all the other members. For recent applications of the coefficient of cooperation to environmental economics and industrial organization see Colombo and Labrecciosa (2018) and Lopez and Vives (2019), respectively.

decision rule, with $z = 1, ..., n$, and $x_k(t) = \sigma_k(K, t, \mathbf{p}(t))$, where σ_k denotes the decision rule, with $k = i, j$, (i.e. we assume symmetry among countries within the same group, either the group of signatories or the group of non-signatories).

In order to sharpen our results, we make the following assumption:

Assumption A2: $c \in (\underline{c}, \overline{c})$, with $\underline{c} = 1/r$ and $\overline{c} = n(1 - e^{-rT})/r$.

Our specification that $c > \underline{c}$ implies that for a country acting in isolation (i.e., not participating in the agreement) providing the public good will not be rational in equilibrium. The second specification, that $c < \overline{c}$, implies that it is socially desirable (in terms of sum of utilities for all countries) that all countries contribute, at least in the short run.

For future reference, we define global welfare at t,

$$w(t) = nK(t) - c\left[mx_i^*(t) + (n - m)\, x_j^*(t)\right],$$

which can be used to compute discounted global welfare,

$$W = \int_0^T e^{-rt} w(t)\, dt. \tag{6}$$

3 The Social Optimum Benchmark

In this section, we derive the socially optimal level of contributions to the public good. The socially optimal solution will be used as a benchmark against which comparing the equilibrium of the game.

The social planner's problem can be written as

$$\begin{cases} \max\limits_X W = \displaystyle\int_0^T e^{-rt}\left[nK(t) - cX(t)\right] dt \\ s.t.\ X \in [0, n] \text{ and } \frac{dK(t)}{dt} = X(t), \end{cases} \tag{7}$$

and the corresponding HJB equation is given by

$$rV(K, t) = \max_{X \in [0,n]} \left\{ nK - cX + \frac{\partial V(K, t)}{\partial K} X + \frac{\partial V(K, t)}{\partial t} \right\}. \tag{8}$$

Maximization of (8) implies that[13]

$$X^* = \begin{cases} n \text{ if } \frac{\partial V(K,t)}{\partial K} \geq c \\ 0 \text{ otherwise.} \end{cases} \tag{9}$$

[13] We assume that if $\frac{\partial V(K,t)}{\partial K} = c$ then the social planner will use the tie-breaking rule that $X^* = \sup[0, n]$.

Assume that $V(K,t) = A(t)K + B(t)$ and that $A(t) \geq c$. From (8), it follows that

$$r\left[A(t)K + B(t)\right] = n\left[K - c + A(t)\right] + \frac{\partial A(t)}{\partial t}K + \frac{\partial B(t)}{\partial t},$$

which implies that

$$A(t) = \frac{n\left[1 - e^{r(t-T)}\right]}{r},$$

and that

$$B(t) = \frac{n\left\{e^{r(t-T)}\left\{cr + n\left[r(t-T) - 1\right]\right\} - cr + n\right\}}{r^2},$$

where we have used the boundary condition $A(T) = B(T) = 0$. Since $A(t)$ is decreasing in t, then $A(t) \geq c$ for $t \in [0, \widetilde{T}\rangle$, with $\widetilde{T}$ solving $A(t) = c$. Note that the highest value that $A(t)$ can assume is $n(1 - e^{-rT})/r$. By Assumption A2, we have $A(0) \geq c$. In the long run $(t > \widetilde{T})$, $X^*(K) = 0$. From (8), it follows that

$$r\left[A(t)K + B(t)\right] = nK + \frac{\partial A(t)}{\partial t}K + \frac{\partial B(t)}{\partial t},$$

which implies that $A(t)$ is as in the short run $(t < \widetilde{T})$, while $B(t) = 0$.

The above discussion yields the following proposition.

Proposition 1 *The socially optimal level of contributions to the public good is given by*

$$X^{so}(t) = \begin{cases} n & \text{for } t \in [0, \widetilde{T}\rangle \\ 0 & \text{otherwise,} \end{cases}$$

with

$$\widetilde{T} = T + \frac{1}{r}\log\left(1 - \frac{cr}{n}\right) \leq T.$$

Proposition 1 establishes that it is socially desirable to stop contributing after time $\widetilde{T}$: when the terminal date approaches, the shadow price of the public good becomes smaller than the cost of contributing.

4 The Equilibrium of the Game

The game at hand consists of a sequence of two-stage games. At any $t \in [0, T\rangle$, there are two stages: (i) a participation stage; (ii) a contribution stage. We proceed as follows: first, we determine the contribution levels at t for a given number of participants; then, we endogenize the number of participants at t; finally, we check whether the equilibrium conditions for the two-stage game at t are satisfied for all two-stage games, i.e. for all $t \in [0, T\rangle$, or they are satisfied only for a time interval within $[0, T\rangle$.

4.1 The Contribution Stage at t

By standard arguments, MPE strategies have to satisfy the following HJB equations:

$$\begin{aligned} rV_i(K,t) &= \max_{x_i\in[0,1]}\left\{K - cx_i + \phi\sum_{k\neq i,k=1}^{m(t)}[K - c\sigma_k(K,t,\mathbf{p}(t))] + \frac{\partial V_i(K,t)}{\partial K}\right. \\ &\quad \left.\times\left[x_i + \sum_{k\neq i,k=1}^{m(t)}\sigma_k(K,t,\mathbf{p}(t)) + \sum_{j=m(t)+1}^{n}\sigma_j(K,t,\mathbf{p}(t))\right] + \frac{\partial V_i(K,t)}{\partial t}\right\}, \end{aligned} \tag{10}$$

and

$$\begin{aligned} rV_j(K,t) &= \max_{x_j\in[0,1]}\left\{K - cx_j + \frac{\partial V_j(K,t)}{\partial K}[x_j + \sum_{i=1}^{m(t)}\sigma_i(K,t,\mathbf{p}(t))\right. \\ &\quad \left. + \sum_{k\neq j,k=m(t)+1}^{n-1}\sigma_k(K,t,\mathbf{p}(t))] + \frac{\partial V_j(K,t)}{\partial t}\right\}, \end{aligned} \tag{11}$$

with $i = 1,...,m(t)$, $j = m(t)+1,...,n$. Maximization of (10) implies that[14]

$$x_i^* = \begin{cases} 1 \text{ if } \frac{\partial V_i(K,t)}{\partial K} \geq c \\ 0 \text{ otherwise,} \end{cases} \tag{12}$$

whereas maximization of (11) implies that[15]

$$x_j^* = \begin{cases} 1 \text{ if } \frac{\partial V_j(K,t)}{\partial K} \geq c \\ 0 \text{ otherwise.} \end{cases} \tag{13}$$

The (Stage 2) equilibrium payoff of country i is given by

$$\Pi_i(K,t;m) = \begin{cases} \int_t^T e^{-r(s-t)}\left[K_{ij}^*(t) - c\right] ds \text{ if } \frac{\partial V_i(K,t)}{\partial K}, \frac{\partial V_j(K,t)}{\partial K} \geq c \\ \int_t^T e^{-r(s-t)}\left[K_i^*(t) - c\right] ds \text{ if } \frac{\partial V_i(K,t)}{\partial K} \geq c,\ \frac{\partial V_j(K,t)}{\partial K} < c \\ \int_t^T e^{-r(s-t)}\left[K_j^*(t)\right] ds \text{ if } \frac{\partial V_i(K,t)}{\partial K} < c,\ \frac{\partial V_j(K,t)}{\partial K} \geq c \\ \int_t^T e^{-r(s-t)}\left[K^*(t)\right] ds \text{ if } \frac{\partial V_i(K,t)}{\partial K},\ \frac{\partial V_j(K,t)}{\partial K} < c \end{cases} \tag{14}$$

[14] We assume that if $\frac{\partial V_i(K,t)}{\partial K} = c$ then signatories will use the tie-breaking rule that $x_i^* = \sup[0,1]$.

[15] We assume that if $\frac{\partial V_j(K,t)}{\partial K} = c$ then non-signatories will use the tie-breaking rule that $x_j^* = \sup[0,1]$.

and that of country j by

$$\Pi_j(K,t;m) = \begin{cases} \int_t^T e^{-r(s-t)} \left[K_{ij}^*(t) - c\right] ds \text{ if } \frac{\partial V_i(K,t)}{\partial K}, \frac{\partial V_j(K,t)}{\partial K} \geq c \\ \int_t^T e^{-r(s-t)} \left[K_i^*(t)\right] ds \text{ if } \frac{\partial V_i(K,t)}{\partial K} \geq c,\ \frac{\partial V_j(K,t)}{\partial K} < c \\ \int_t^T e^{-r(s-t)} \left[K_j^*(t) - c\right] ds \text{ if } \frac{\partial V_i(K,t)}{\partial K} < c,\ \frac{\partial V_j(K,t)}{\partial K} \geq c \\ \int_t^T e^{-r(s-t)} \left[K^*(t)\right] ds \text{ if } \frac{\partial V_i(K,t)}{\partial K}, \frac{\partial V_j(K,t)}{\partial K} < c \end{cases} \tag{15}$$

where $K_{ij}^*(t)$, $K_i^*(t)$, $K_j^*(t)$ and $K^*(t)$ denote the trajectories of K when all countries contribute, only signatories contribute, only non-signatories contribute, and no countries contribute, respectively: $K_{ij}^*(t)$ solves $dK/dt = n$, $K_i^*(t)$ solves $dK/dt = m(t)$, $K_j^*(t)$ solves $dK/dt = n - m(t)$, and $K^*(t) = K_0$.

4.2 The Participation Stage at t

In equilibrium, two groups of players can be identified: a group $G_1(t)$ consisting of $m(t)$ signatories (where $m(t)$ is endogenously determined), and a group $G_2(t)$ consisting of $n-m(t)$ non-signatories. For a (Stage 1) equilibrium with $m(t)$ signatories to exist it must be that each $i \in G_1(t)$ has no unilateral incentive to deviate and join $G_2(t)$, and each $j \in G_2(t)$ has no unilateral incentive to deviate by joining $G_1(t)$. Formally, for each $i \in G_1(t)$, it must hold that

$$\Pi_i(K,t;m) \geq \Pi_j(K,t;m-1), \tag{16}$$

and for each $j \in G_2(t)$, it must hold that

$$\Pi_j(K,t;m) \geq \Pi_i(K,t;m+1), \tag{17}$$

for all $K \geq 0$ and $t \in [0,T\rangle$. Inequality (16) states that, for a participating country, it must be individually rational to participate in the agreement (in which case the total number of participating countries is m) rather than acting in isolation.[16] We will refer to condition (16) as the

[16] The agent assumes that if it leaves group G_1 to join group G_2, then (i) the number of signatories will become $m-1$, i.e., the other $m-1$ signatories stay in G_1; and (ii) the remaining signatories will adjust their contribution level to the Nash equilibrium level with $m-1$ signatories.

contributor-rationality condition. (16) corresponds to the internal stability condition in the coalition literature.

Inequality (17) states that, for a non-signatory, it must be better to stay out of the agreement rather than joining group $G_1(t)$ (in which case it assumes that the number of signatories becomes $m+1$, and that all of them contribute at their new symmetric Nash equilibrium level with $m+1$ signatories). We will refer to condition (17) as the **free-rider-rationality** condition. (17) corresponds to the external stability condition in the coalition literature. In line with the bulk of the literature, we assume that (17) needs to hold with strict inequality sign, i.e. a country which is indifferent between joining and not joining will join.

4.3 Equilibrium Characterization

Proposition 2 *If* $c \leq \widehat{c} \in [\underline{c}, \overline{c}]$,

$$\widehat{c} = \frac{[\phi(n-1)+1]\left(1-e^{-rT}\right)}{r},$$

then, for all $t \in [0, \widehat{T}\rangle$, *there exists an agreement among* $m^*(t)$ *countries contributing to the public good which is unique, with*

$$m^*(t) = f\left(1+\frac{1}{\phi}\left(\frac{cr}{1-e^{r(t-T)}}-1\right)\right) \leq n,$$

where $f(x)$ *is the smallest integer larger than or equal to* x, *and*

$$\widehat{T} = T + \frac{1}{r}\log\left(1-\frac{cr}{1+\phi(n-1)}\right) \leq T.$$

The equilibrium level of contributions to the public good is given by

$$X^*(t) = \begin{cases} m^*(t) & \text{for } t \in [0, \widehat{T}\rangle \\ 0 & \text{otherwise.} \end{cases}$$

Proof. The proof proceeds in three steps. In Step 1, we derive the (Stage 2) value functions and equilibrium payoffs. In Step 2, we prove that $m^*(t)$ given in Lemma 1 is an equilibrium. Then, in Step 3, we show that it is unique.

Step 1. Assume $V_k(K,t) = A_k(t)K + B_k(t)$, with $k = i, j$, and $A_k(T) = B_k(T) = 0$. There are four sub-cases to consider:

1a. $\frac{\partial V_i(K,t)}{\partial K}, \frac{\partial V_j(K,t)}{\partial K} \geq c$. In this sub-case, $x_i^* = x_j^* = 1$. Value functions for the signatories are given by

$$V_i = A_i(t) K + B_i(t),$$

where

$$A_i(t) = \frac{[\phi(m-1)+1]\left(1-e^{r(t-T)}\right)}{r},$$

and

$$B_i(t) = \frac{[\phi(m-1)+1]\left\{e^{r(t-T)}\{cr+n[r(t-T)-1]\}+(n-cr)\right\}}{r^2},$$

and value functions for the non-signatories are given by

$$V_j = A_j(t) K + B_j(t),$$

where

$$A_j(t) = \frac{1-e^{r(t-T)}}{r},$$

and

$$B_j(t) = \frac{e^{r(t-T)}\{cr+n[r(t-T)-1]\}+(n-cr)}{r^2}.$$

Since by assumption $\frac{1}{r} < c$ then $A_j(t) < c$ implying that sub-case 1a does not exist.

1b. $\frac{\partial V_i(K,t)}{\partial K} \geq c$, $\frac{\partial V_j(K,t)}{\partial K} < c$. In this sub-case, $x_i^* = 1$ and $x_j^* = 0$. Value functions for the signatories are given by

$$V_i = A_i(t) K + B_i(t),$$

where $A_i(t)$ is as in sub-case 1a, and

$$B_i(t) = \frac{[\phi(m-1)+1]\left\{e^{r(t-T)}\{cr+m[r(t-T)-1]\}+(m-cr)\right\}}{r^2},$$

and value functions for the non-signatories are given by

$$V_j = A_j(t) K + B_j(t),$$

where $A_j(t)$ is as in sub-case 1a, and

$$B_j(t) = \frac{m\left\{e^{r(t-T)}[r(t-T)-1]+1\right\}}{r^2}.$$

This sub-case exists when $\frac{1}{r} < c$ and $\frac{[\phi(m-1)+1]\left(1-e^{r(t-T)}\right)}{r} \geq c$. The first inequality is satisfied by Assumption A2; the second inequality holds true if $m \geq m^*(t)$, with $m^*(t)$ given in Lemma 1.

The (Stage 2) equilibrium payoffs are given by

$$\Pi_i = V_i|_{\phi=0},$$

and

$$\Pi_j = V_j.$$

1c. $\frac{\partial V_i(K,t)}{\partial K} < c$, $\frac{\partial V_j(K,t)}{\partial K} \geq c$. In this sub-case, $x_i^* = 0$ and $x_j^* = 1$. Value functions for the signatories are given by

$$V_i = A_i(t) K + B_i(t),$$

where $A_i(t)$ is as in sub-case 1a, and

$$B_i(t) = \frac{(n-m)[\phi(m-1)+1]\left\{e^{r(t-T)}[r(t-T)-1]+1\right\}}{r^2},$$

and value functions for the non-signatories are given by

$$V_j = A_j(t) K + B_j(t),$$

where $A_j(t)$ is as in sub-case 1a, and

$$B_j(t) = \frac{e^{r(t-T)}\{cr + (n-m)[r(t-T)-1]\} + n - m - cr}{r^2}.$$

This sub-case exists when $\frac{[\phi(m-1)+1]\left(1-e^{r(t-T)}\right)}{r} < c$ and $\frac{1-e^{r(t-T)}}{r} > c$ which is impossible.

1d. $\frac{\partial V_i(K,t)}{\partial K}, \frac{\partial V_j(K,t)}{\partial K} < c$. In this sub-case, $x_i^* = x_j^* = 0$. Value functions for the signatories are given by

$$V_i = A_i(t) K + B_i(t),$$

where $A_i(t)$ is as in sub-case 1a, and

$$B_i(t) = 0,$$

and value functions for the non-signatories are given by

$$V_j = A_j(t) K + B_j(t),$$

where $A_j(t)$ is as in sub-case 1a, and

$$B_j(t) = 0.$$

This sub-case exists when $\frac{[\phi(m-1)+1]\left(1-e^{r(t-T)}\right)}{r} < c$, which is the case if $m < m^*(t)$, with $m^*(t)$ given in Lemma 1. The (Stage 2) equilibrium payoffs are given by

$$\Pi_i = V_i|_{\phi=0},$$

and

$$\Pi_j = V_j.$$

To sum-up, only sub-cases 1b and 1d are possible, and the (Stage 2) equilibrium payoffs turn out to be

$$\Pi_i(K,t;m) = \begin{cases} \frac{m+r(K-c)-e^{r(t-T)}\{r(K-c)+m[r(T-t)+1]\}}{r^2} & \text{for } t \in [0,\widehat{T}\rangle \\ \frac{\left(1-e^{r(t-T)}\right)K}{r} & \text{for } t \in (\widehat{T},T\rangle \end{cases}$$

and

$$\Pi_j(K,t;m) = \begin{cases} \frac{m+rK-e^{r(t-T)}\{rK+m[r(T-t)+1]\}}{r^2} & \text{for } t \in [0,\widehat{T}\rangle \\ \frac{\left(1-e^{r(t-T)}\right)K}{r} & \text{for } t \in (\widehat{T},T\rangle. \end{cases}$$

Step 2. Consider first the case where

$$f\left(1+\frac{1}{\phi}\left(\frac{cr}{1-e^{r(t-T)}}-1\right)\right) < n,$$

so that there are $m^* < n$ members. From (16) and (17), we have that m^* is an equilibrium if

$$\begin{aligned} \Pi_i(K,t;m^*) &= \frac{m^*+r(K-c)-e^{r(t-T)}\{r(K-c)+m^*[r(T-t)+1]\}}{r^2} \\ &\geq \Pi_j(K,t;m^*-1) = \frac{\left(1-e^{r(t-T)}\right)K}{r}, \end{aligned} \tag{18}$$

and

$$\begin{aligned} \Pi_j(K,t;m^*) &= \frac{m^*+rK-e^{r(t-T)}\{rK+m^*[r(T-t)+1]\}}{r^2} \\ &> \Pi_i(K,t;m^*+1) = \frac{m^*+1+r(K-c)-e^{r(t-T)}}{r^2} \\ &\quad \times \{r(K-c)+(m^*+1)[r(T-t)+1]\}. \end{aligned} \tag{19}$$

Inequality (18) can be rewritten as

$$\Pi_j(K,t;m^*) - \Pi_j(K,t;m^*-1) = \frac{e^{r(t-T)}\{cr-m^*[r(T-t)+1]\}+m^*-cr}{r^2} \geq 0,$$

which is decreasing in t for $m^* > c/(T-t)$ and nil at $t = T$. $m^* - c/(T-t)$ is increasing in c, and nil at $c = c_1$, with

$$c_1 = \frac{(1-\phi)(T-t)\left(e^{r(t-T)}-1\right)}{\phi\left(1-e^{r(t-T)}\right)-r(T-t)} < \frac{1}{r}.$$

Recall that $c > 1/r$ by Assumption A2. Therefore, $m^* > c/(T-t)$. This proves that $\Pi_i(K,t;m^*) > \Pi_j(K,t;m^*-1)$.

Inequality (19) can be rewritten as

$$\Pi_j(K,t;m^*) - \Pi_i(K,t;m^*+1) = \frac{e^{r(t-T)}[1-r(c+t-T)]+cr-1}{r^2},$$

which is increasing in c and nil at $c = c_2$, with

$$c_2 = (T-t)\left(1+\frac{1}{e^{r(t-T)}-1}\right)+\frac{1}{r} < \frac{1}{r}.$$

Therefore, $\Pi_j(K,t;m^*) > \Pi_i(K,t;m^*+1)$.

Consider next the case where

$$f\left(1+\frac{1}{\phi}\left(\frac{cr}{1-e^{r(t-T)}}-1\right)\right) = n,$$

so that there are $m^* = n$ members. From (16) we have that $m^* = n$ is an equilibrium if

$$\begin{aligned}\Pi_i(K,t;n) &= \frac{n+r(K-c)-e^{r(t-T)}\{r(K-c)+n[r(T-t)+1]\}}{r^2} \\ &\geq \Pi_j(K,t;n-1) = \frac{\left(1-e^{r(t-T)}\right)K}{r}. \qquad (20)\end{aligned}$$

By the same logic as that used to prove that $m^* > c/(T-t)$ we have that $\Pi_i(K,t;n) > \Pi_j(K,t;n-1)$. Since $m^* = n$, the free-rider-rationality condition does not apply. Note that when $m^* = n$ the equilibrium level of contributions is nil.

Step 3. Consider m^*+1 members, with $m^* \leq n-1$. We have that m^*+1 is an equilibrium if

$$\begin{aligned}\Pi_i(K,t;m^*+1) &= \frac{m^*+1+r(K-c)-e^{r(t-T)}\{r(K-c)+(m^*+1)[r(T-t)+1]\}}{r^2} \\ &\geq \Pi_j(K,t;m^*) = \frac{m^*+rK-e^{r(t-T)}\{rK+m^*[r(T-t)+1]\}}{r^2},\end{aligned}$$

which can be rewritten as

$$\Pi_i(K,t;m^*+1) - \Pi_j(K,t;m^*) = \frac{e^{r(t-T)}[r(c+t-T)-1]-cr+1}{r^2} \geq 0.$$

Since $c > 1/r$ by Assumption A2 then $\Pi_i(K,t;m^*+1) < \Pi_j(K,t;m^*)$ implying that m^*+1 does not satisfy the contributor-rationality condition: the extra member would find it rational to leave. Analogously, consider m^*+k, with $k \geq 1$ and $m^* \leq n-k$. We have

$$\Pi_i(K,t;m^*+k) < \Pi_j(m^*+k-1).$$

Hence, any outcome with more than $m^* < n$ members does not satisfy the contributor-rationality condition. Consider now $m^* - 1$ members, with $m^* < n$. We have

$$\begin{aligned}\Pi_j (K, t; m^* - 1) &= \frac{\left(1 - e^{r(t-T)}\right) K}{r} < \Pi_i (K, t; m^*) = \frac{m^* + r (K - c) - e^{r(t-T)}}{r^2} \\ &\times \{r (K - c) + m^* [r (T - t) + 1]\} .\end{aligned}$$

Therefore, $m^* - 1$ does not satisfy the free-rider-rationality condition: the first excluded member would find it rational to join. Finally, consider $m^* - k$, with $k \geq 1$ and $m^* < n$. We have

$$\Pi_i (K, t; m^* - k) = \frac{\left(1 - e^{r(t-T)}\right) K}{r} = \Pi_j^* (K, t; m^* - k - 1) = \frac{\left(1 - e^{r(t-T)}\right) K}{r}.$$

Under the tie-breaking assumption that a country which is indifferent between joining and not joining will join any outcome with less than $m^* < n$ members does not satisfy the free-rider-rationality condition. We can then conclude that there exists a unique m^*, with m^* given in Lemma 1.

It is immediate to verify that $\widehat{T}$ is the value of t that solves

$$1 + \frac{1}{\phi} \left(\frac{cr}{1 - e^{r(t-T)}} - 1 \right) = n,$$

and that $\widehat{T} \in (0, T)$ for $c < \widehat{c}$. For $c \geq \widehat{c}$ we have either $\widehat{T} < 0$ or $\widehat{T} \notin \mathbb{R}$. As a consequence, since

$$\frac{[\phi (n - 1) + 1] \left(1 - e^{r(t-T)}\right)}{r} < c,$$

then $m^* (t) = n$ and $X^* (t) = 0$ for all $t \in [0, T\rangle$. ■

Proposition 1 derives the threshold of c below which there exists an agreement among m^* countries contributing to the public good. Such a threshold is increasing in ϕ, n, T, and decreasing in r, implying that the likelihood of cooperative agreements to exist is higher the tighter the agreement (i.e., the higher ϕ), the higher the number of countries which could potentially participate in the agreement, the longer the time horizon of each country, or the less each country discounts future payoffs. When it is too costly for countries to contribute, an agreement among contributing countries does not arise in equilibrium, and the stock of public good remains at its initial stock, K_0, for all $t \in [0, T\rangle$.

Corollary 1 *An agreement among $m^*(t) \leq n$ contributing countries specifying a coefficient of cooperation $\phi < \underline{\phi}$ does not exist, with*

$$\underline{\phi} = \frac{e^{-Tr} + cr - 1}{(n - 1) (1 - e^{-Tr})}.$$

If $T \geq \underline{T}$ then $\underline{\phi} \in (0,1]$, with

$$\underline{T} = \frac{1}{r} \log \left(\frac{n}{n - cr} \right).$$

If $T < \underline{T}$ then $\underline{\phi} > 1$ and an agreement among $m^*(t) \leq n$ contributing countries does not exist.

Proof. The inequality $c \leq \widehat{c}$ in Proposition 1 can be solved for ϕ to get $\underline{\phi}$, which reduces to $(cr-1)/(n-1)$ as T tends to infinity. If $\phi > \underline{\phi}$ then $m^*(t) < n$; if $\phi = \underline{\phi}$ then $m^*(t) = n$; if $\phi < \underline{\phi}$ then $c > \frac{\partial V_i(K,t)}{\partial K}$ implying that $X^*(t) = 0$. $\underline{\phi}$ is decreasing in T and $\underline{\phi} = 1$ at $\underline{T}$. Therefore, $\underline{\phi} \leq (>)1$ for $T \geq (<)\underline{T}$. ■

The above corollary implies that if the time horizon is very short ($T < \underline{T}$) then voluntary public good provision is not an equilibrium outcome of the game, even if signatories agree to fully cooperate. In the social optimum, instead, even when the time horizon is very short, it is always optimal for the social planner to provide the public good.

Corollary 2 *A loose cooperative agreement exists for a shorter time than a tight cooperative agreement.*

Proof. Immediate since $\widehat{T}$ is increasing in ϕ. ■

An agreement among $m^*(t) \leq n$ contributing countries exists until $\widehat{T}$. Only when T tends to infinity T and $\widehat{T}$ are equal to each other and independent of ϕ, otherwise $\widehat{T} < T$, i.e., there exists a time interval where countries make no contributions to the public good, with $\widehat{T}$ being lower for loose cooperative agreements. T can be interpreted as the expected duration of a government. For instance, a shorter T can be the result of a more democratic regime where elections are periodic and occur more frequently than in a less democratic regime. In the extreme case of dictatorship, T is very long, eventually going to infinity. Under this interpretation, at any given time, countries with long-lasting welfare-maximizing governments are more likely to contribute to the public good than countries where governments change over time and have different objectives than welfare maximization.

Corollary 3 *If the time horizon is finite (resp. infinite) then the equilibrium size of the agreement increases (resp. is constant) over time.*

Proof. $m^*(t)$ given in Proposition 2 is clearly increasing in t for finite T, and independent of t for $T \rightarrow \infty$. ■

The result that participation to an international agreement increases over time is in line with what has been observed in several real-world situations (e.g. increasing participation to the European Union, which started in the 1950s with six members and nowadays has 27 members; increasing participation to NATO, which started in 1949 with 12 founding members and, as of today, has 30 members). Indeed, for the case of a finite planning horizon, our model is capable of accounting for an important well-documented fact about international cooperation.

Corollary 4 *The equilibrium size of the agreement is larger for a loose cooperative agreement than for a tight cooperative agreement.*

Proof. Immediate since m^* is decreasing in ϕ. ■

Intuitively, an increase in the coefficient of cooperation increases the likelihood of a positive contribution (from (12)), thus leading to an increase in the incentive to free ride. The result that participation to an international agreement is larger for loose cooperative agreements than to tight cooperative agreements can explain the larger participation to a loose cooperative agreement such as the Paris Accord than to a tight cooperative agreements such as the Kyoto Protocol.

Loose cooperative agreements are more successful in attracting participation by being less ambitious. The logic resembles that behind partial collusion in repeated games: when the critical discount factor sustaining full collusion is too high compared with the discount factor used by firms to discount future profits, it is still possible for firms to sustain some degree of cooperation rather than behaving as Nash players provided that firms' discount factor is sufficiently high.

In the light of Corollaries 2 and 4, we are in a position to establish a novel trade-off between agreements that are, using the jargon of the literature on international environmental agreements (see Harstad 2020b and references therein), **narrow-but-deep-and-long-lived** vs **broad-but-shallow-and-short-lived**.

Proposition 3 *A loose cooperative agreement is broad-but-shallow-and-short-lived (i.e. it implies strong participation but low cooperation and short duration) whereas a tight cooperative agreement is narrow-but-deep-and-long-lived (i.e. it implies weak participation but high cooperation and long duration).*

Proposition 3 adds a third dimension, the time dimension, to the classical trade-off between agreements that are narrow-but-deep vs broad-but-shallow. In our dynamic game, strong participation has to be weighted against not only low cooperation but also short duration of the agreement.

We then expect tight cooperative agreements to be welfare-superior to loose cooperative agreements in a larger subset of parameters than in the case in which both agreements have the same duration, which is the case in static models.

In the static counterpart of our model, there exists a unique trade-off, that between participation in the agreement and cooperation in the agreement. In our dynamic model, instead, there are three trade-offs: (i) a trade-off between participation in the agreement and cooperation in the agreement; (ii) a trade-off between participation in the agreement and duration of the agreement; (iii) a trade-off between cooperation in the agreement and duration of the agreement. Compared with the static literature on narrow-but-deep vs broad-but-shallow agreements, we provide a richer comparative analysis of international agreements, both positive and normative. Our framework yields results that are coherent with well-established facts about international cooperation that cannot be accounted for in static models, such as increasing participation over time, and shed new light on the social desirability of international cooperation. Which of the two cooperative agreements, loose or tight, should be favored on welfare grounds? Is efficiency achievable, and, if so, under what circumstances? We will tackle these questions in the next section.

We now turn our attention to a comparison between the equilibrium of the game and the social optimum. It can be verified that $\widehat{T} \leq \widetilde{T}$ (with strict inequality if $\phi < 1$), since the shadow price of the public good is higher for the social planner, who fully internalizes the impact that a country's contribution has on other countries, than for individual countries, which give more weight to their own payoff rather than to the payoff of other countries in the agreement. This implies that, compared with social optimum, there is under-contribution to the public good for all $t \in [0, \widetilde{T}\rangle$. Indeed, there exists a dynamic-free riding problem. Interestingly, for $t \in [0, \widehat{T}\rangle$, $X^*(t)$ increases over time, therefore free-riding diminishes over time; for $t \in (\widehat{T}, \widetilde{T}\rangle$, $X^*(t) = 0$ while $X^{so}(t) = n$, thus free-riding reaches its peak. We can then argue that, provided that $c < \widehat{c}$, free riding is non-monotone w.r.t. time. If, instead, $c \geq \widehat{c}$ then free riding will be constant over time (at its peak): $X^*(t) = 0$ while $X^{so}(t) = n$ for all $t \in [0, \widetilde{T}\rangle$. When $\phi = 1$, we have $\widehat{T} = \widetilde{T}$, therefore the duration of the agreement in equilibrium turns out to be socially optimal. Despite this, compared with social optimum, there is under-contribution. The intuitive explanation is that while the social planner makes it compulsory for countries to participate in the agreement, in the unregulated scenario, countries' participation is voluntary, and there are private incentives for countries to free ride on the contributions to the public good made by signatories, thus leading to a lower participation than that which is socially optimal. As in the case in which $\phi < 1$, for $t \in [0, \widehat{T}\rangle$, free riding decreases over time until reaching zero. For $t \in (\widehat{T}, T\rangle$ free-riding is nil. Hence, if $\phi < 1$ free-riding will be always non-increasing over time.

Remark 1 *Partial cooperation is responsible for free-riding to be non-monotone with respect to time.*

We conclude this section with the impact of c and r on the equilibrium size of the agreement, m^*. The following comparative statics properties hold: (i) m^* is nondecreasing in c, either constant, if dc is such that $f(x)$ remains unchanged, or increasing. Intuitively, an increase in c increases the incentive to "share the burden", thus leading to bigger participation. (ii) m^* is nondecreasing in r, either constant, if dr is such that $f(x)$ remains unchanged, or increasing. Intuitively, an increase in the discount rate decreases the likelihood of a positive contribution (from (12)), thus leading to a decrease in the incentive to free ride. Note that m^* is independent of n if $m^* < n$, otherwise it is increasing in n.

5 Welfare Analysis

In this section, we compare and contrast loose and tight cooperative agreements in terms of discounted global welfare, and we show that conditions exist under which efficiency can be achieved. For expository convenience, we focus on the case in which T is infinite. By continuity, the results for $T \to \infty$ extend to the case in which T is very large.

From (6), discounted global welfare turns out to be

$$W = \frac{rnK_0 + (n - cr)\, m^*}{r^2}.$$

It follows that

$$\frac{\partial W}{\partial \phi} = \frac{(cr-1)\,(cr-n)}{r^2 \phi^2} \frac{\partial m^*}{\partial \phi}.$$

Proposition 4 *Let $T \to \infty$. Discounted global welfare is higher under a loose than under a tight cooperative agreement.*

Proof. Immediate since m^* is decreasing in ϕ (for $d\phi$ sufficiently large) and $c \in [\underline{c}, \overline{c}]$. ■

Since both Π_i and Π_j are increasing in m, loose cooperative agreements are not only welfare-superior but also Pareto-superior. The fact that W is decreasing in ϕ (unless $d\phi$ is such that m^*

remains unchanged) and the fact that $m^* = n$ for $\phi = \underline{\phi}$, with $\underline{\phi}$ given in Corollary 1, enable us to state the following important corollary to Proposition 4.

Corollary 5 *Let $T \to \infty$. If $\phi = \underline{\phi}$ then $m^* = n$ for all $t \in [0, \infty)$. In this case, the equilibrium size of the agreement and the contributions to the public good are both socially optimal.*

When $T \to \infty$ (therefore the difference between the durations of the two agreements converges to zero) discounted global welfare is maximized by setting $\phi = \underline{\phi}$, with $\underline{\phi}$ denoting the lower bound of ϕ for an agreement to exist (see Corollary 1). Any increase in the coefficient of cooperation above $\underline{\phi}$ leads to a reduction in discounted global welfare by reducing the equilibrium number of contributors and, in turn, the stock of public good. Interestingly, efficiency is achieved despite the fact that signatory countries do not fully coordinate. $\underline{\phi}$ is increasing in c and r and decreasing in n. Not surprisingly, discounted global welfare is increasing in n. As to the impact of c, we have

$$\frac{\partial W}{\partial c} = \frac{(n - cr)}{r\phi} \frac{\partial m^*}{\partial c} - \frac{m^*}{r}.$$

Since $\partial m^* / \partial c \geq 0$ and $c < n/r$ (by Assumption A2) *a priori*, $\partial W / \partial c$ can be either positive or negative. Counterintuitively, a small increase in c can be welfare-improving because it can lead to a discrete increase in m^* which outweighs the negative direct effect. When instead $dc > 0$ is sufficiently large, despite a discrete increase in m^*, discounted global welfare decreases.

6 Concluding Remarks

We have proposed and analyzed a n-country continuous-time game of voluntary provision of a global public good. Our dynamic game consists of a sequence of two-stage games. At each point in time, each country has two sequential decisions to make: participation in an international agreement aimed to increase the stock of public good, and contribution to the public good, which depends on whether a country is in or out of the agreement. In the case in which a country participates in the agreement, the contribution level is determined by the maximization of a weighted sum of utilities of all participants, with the weight being specified in the agreement, whereas a non-signatory country decides how much to contribute to the public good, independently and noncooperatively, with the aim of maximizing its own utility. The assumption of partial rather than full cooperation among signatory countries represents one of the main departures from the existing literature on coalition formation and international agreements. This seemingly small departure has important consequences for equilibrium membership and welfare.

Our analysis has shown that a loose cooperative agreement, specifying a low coefficient of cooperation, is associated with larger participation than a tight cooperative agreement, specifying a high degree of cooperation. In contrast with the conventional wisdom according to which the equilibrium coalition size is small and typically inefficient, we have shown that loose cooperative agreements can lead to an equilibrium coalition size that is large and efficient. In the case of a time horizon that is sufficiently long, we have shown that discounted global welfare is higher under loose than under tight cooperative agreements; efficiency is achieved when the coefficient of cooperation is set at its lower bound. A policy implication of this finding is that insisting on coordination among voluntary contributors to a public good is generally welfare reducing. In the realm of climate change policy, for instance, a loose agreement in the style of the Paris Accord is likely to be more successful than its predecessor, the Kyoto Protocol, which proved to be "too demanding" for countries to join.

A particularly important contribution of our paper is the addition of a third dimension, the time dimension, to the classical trade-off between agreements that are narrow-but-deep vs broad-but-shallow. In our dynamic game, strong participation has to be weighted against not only low cooperation but also short duration of the agreement. This leads to a novel trade off between agreements that are narrow-but-deep-and-long-lived vs broad-but-shallow-and-short-lived.

In contrast with previous studies on dynamic voluntary provision of public goods (e.g. Battaglini and Harstad, 2016; Karp and Sakamoto, 2019; Kovác and Schmidt, 2019), we have shown that relatively small coalitions can become bigger over time, which is in line with what has been observed in relation to several real-world international agreements such as the EU and NATO.

References

[1] Aldy, J. E., Barrett, S., Stavins, R.N., 2003. Thirteen plus one: a comparison of global climate policy architectures. *Climate Policy* 3, 373-397.

[2] Andreoni, J., 1988. Privately provided public goods in a large economy: The limits of altruism. *Journal of Public Economics* 35, 57-73.

[3] d'Aspremont, C., Jacquemin, J., Gabszewicz, J. Weymark, J., 1983. On the stability of collusive price leadership. *Canadian Journal of Economics* 16, 17-25.

[4] Barrett, S., 2002. Consensus Treaties. *Journal of Institutional and Theoretical Economics* 158, 529-547.

[5] Barrett, S., 1994. Self-enforcing international environmental agreements. *Oxford Economics Papers* 46, 878-894.

[6] Barrett, S., 1999. A theory of full international cooperation. *Journal of Theoretical Politics* 11, 519-541.

[7] Barrett, S., 2002. Consensus treaties. *Journal of Institutional and Theoretical Economics* 158, 529-547.

[8] Battaglini, M., Harstad, B., 2020. The political economy of weak treaties. *Journal of Political Economy*, forthcoming.

[9] Battaglini, M., Harstad, B., 2016. Participation and duration of environmental agreements. *Journal of Political Economy* 124, 160-204.

[10] Battaglini, M., Nunnari, S., Palfrey, T.R., 2014. Dynamic free riding with irreversible investments. *American Economic Review* 104, 2858-2871.

[11] Benchekroun, H., Long, N.V., 2008. The build-up of cooperative behavior among non-cooperative selfish agents. *Journal of Economic Behavior and Organization* 67, 239-252.

[12] Bergstrom, T.C., Blume, L.E., Varian, H.R., 1986. On the private provision of public goods. *Journal of Public Economics* 29, 25-49.

[13] Bernheim, B Douglas, 1986. On the voluntary and involuntary provision of public goods. *American Economic Review* 76, 789-793.

[14] Bowen, R., Georgiadis, G., Lambert, N.S., 2019. Collective choice in dynamic public good provision. *American Economic Journal: Microeconomics* 11, 243-298.

[15] Buchholz, W., Cornes, R., Rubbelke, D., 2014. Potentially harmful international cooperation on global public good provision. *Economica* 81, 205-223.

[16] Carraro, C., Siniscalco, D., 1993. Strategies for the international protection of the environment. *Journal of Public Economics* 52, 309-28.

[17] Chamberlin, J., 1974. Provision of collective goods as a function of group size. *American Political Science Review* 65, 707-716.

[18] Colombo, L., Labrecciosa, P., 2018. Consumer surplus-enhancing cooperation in a natural resource oligopoly. *Journal of Environmental Economics and Management* 92, 185-193.

[19] Cyert, R.M., deGroot, M.H., 1973. An analysis of cooperation and learning in a duopoly context. *American Economic Review* 63, 24–37.

[20] Cornes, R., Sandler, T., 1986. The theory of externalities, public goods and club goods, Second Edition. Cambridge, U. K.: Cambridge University Press.

[21] de Zeeuw, A., 2008. Dynamic effects on the stability of international environmental agreements. *Journal of Environmental Economics and Management* 55, 163-174.

[22] Diamantoudi, E., 2005. Stable cartels revisited. *Economic Theory* 26, 907-921.

[23] Diamantoudi, E., Sartzetakis, E.S., 2015. International environmental agreements: coordinated action under foresight. *Economic Theory* 59, 527-546.

[24] Dutta, P.K., Radner, R., 2004. Self-enforcing climate-change treaties. *Proceedings of the National Academy of Sciences* 101, 4746-4751.

[25] Edgeworth, F.Y., 1881. Mathematical psychics: an essay on the application of mathematics to the moral sciences. Kegan Paul, London.

[26] Eichner, T., Pethig, R., 2013. Self-enforcing environmental agreements and international trade. *Journal of Public Economics* 102, 37-50.

[27] Fershtman, C., Nitzan, S., 1991. Dynamic voluntary provision of public goods. *European Economic Review* 35, 1057-1067.

[28] Finus, M., Maus, S., 2008. Modesty may pay!. *Journal of Public Economic Theory* 10, 801-826.

[29] Fujiwara, K., Matsueda, N., 2009. Dynamic voluntary provision of public goods: An extension. *Journal of Public Economic Theory* 11, 27-36.

[30] Georgiadis, G., 2015. Project and Team Dynamics. *Review of Economic Studies* 82, 187-218.

[31] Georgiadis, G., 2017. Deadlines and infrequent monitoring in the dynamic provision of public goods. *Journal of Public Economics* 152, 1-12.

[32] Gerlagh, R., Liski, M., 2018. Consistent climate policies. *Journal of the European Economic Association* 16, 1-44.

[33] Harstad, B., 2020a. Pledge-and-review bargaining: A theory. Mimeo.

[34] Harstad, B., 2020b. Pledge-and-review bargaining: From Kyoto to Paris. Mimeo.

[35] Harstad, B., 2020c. Technology and Time Inconsistency. *Journal of Political Economy*, forthcoming.

[36] Hoel, M., 1992. International environmental conventions: the case of uniform reductions of emissions. *Environmental and Resource Economics* 2, 141-159.

[37] Hoel, M., de Zeeuw, A., 2014. Technology agreements with heterogeneous countries. CESifo Working Paper No. 4635.

[38] Hong, F.,Karp, L., 2012. International environmental agreements with mixed strategies and investment. *Journal of Public Economics* 96, 685-697.

[39] Itaya, J., Shimomura, K., 2001. A dynamic conjectural variations model in the private provision of public goods: a differential game approach. *Journal of Public Economics* 81, 153-172.

[40] Karp, L., Sakamoto, H., 2019. Sober optimism and the formation of international environmental agreements. Mimeo.

[41] Karp, L., Simon, L.K., 2013. Participation games and international environmental agreements: A non-parametric model. *Journal of Environmental Economics and Management* 65, 326-344.

[42] Kolstad, C. D., Toman, M., 2005. The economics of climate policy. Handbook of Environmental Economics 3, 1562-93.

[43] Kovác, E., Schmidt, R.C., 2019. A simple dynamic climate cooperation model. Mimeo, University of Hagen.

[44] Long, N.V., Shimomura, K., 2007. Voluntary contributions to a public good: Non-neutrality results. *Pacific Economic Review* 12, 153-170.

[45] Lopez, A., Vives, X., 2019. Overlapping ownership, R&D spillovers and antitrust policy. *Journal of Political Economy* 127, 2394-2437.

[46] Yanase, A., 2006. Dynamic voluntary provision of public goods and optimal steady-state subsidies. *Journal of Public Economic Theory* 8, 171-179.

[47] Mukanjari, S., Sterner, T., 2018. Do markets trump politics? Evidence from fossil market reactions to the Paris Agreement and the U.S. election. Working Paper in Economics, No. 728, University of Gothenburg, Gothernburg.

[48] Marx, L.M., Matthews, S.A., 2000. Dynamic voluntary contribution to a public project. *Review of Economic Studies* 67, 327-358.

[49] Morrison, C.C., 1978. A note on providing public goods through voluntary contributions. *Public Choice* 33, 119-123.

[50] Nordhaus, W.D., 2015. Climate clubs: Overcoming free-riding in international climate policy. *American Economic Review* 105, 1339-1370.

[51] Rubio, S.J., Ulph, A., 2006. Self-enforcing international environmental agreements revisited. *Oxford Economic Papers* 58, 233-263.

[52] Rubio, S.J., Ulph, A., 2007. An infinite-horizon model of dynamic membership of international environmental agreements. *Journal of Environmental Economics and Management* 54, 296-310.

[53] Schmalensee, R., 1998. Greenhouse policy architectures and institutions. Economics and Policy Issues in Climate Change, ed. by W. D. Nordhaus, Resources for the Future Press, Washington, D. C.

[54] van der Ploeg, R., Rezai, A., 2019. The risk of policy tipping and stranded carbon assets. CESifo Working Paper No. 7769.

[55] Varian, H. R. 1994. Sequential contributions to public goods. *Journal of Public Economics* 53, 165-186.

[56] Wirl, F., 1996. Dynamic voluntary provision of public goods: Extension to nonlinear strategies. *European Journal of Political Economy* 12, 555-560.