HIAS-E-92

A Dynamic Analysis of Climate Change Mitigation with Endogenous Number of Contributors: Loose vs Tight Cooperation

Luca Colombo[(a)], Paola Labrecciosa[(b)] and Ngo Van Long[(c),(d)]

(a) Deakin University
(b) Monash University
(c) McGill University
(d) Hitotsubashi Institute for Advanced Study, Hitotsubashi University

December, 2019

Hitotsubashi Institute for Advanced Study, Hitotsubashi University
2-1, Naka, Kunitachi, Tokyo 186-8601, Japan
tel:+81 42 580 8604 http://hias.hit-u.ac.jp/

A dynamic analysis of climate change mitigation with endogenous number of contributors: loose vs tight cooperation

Luca Colombo (Deakin University)

Paola Labrecciosa (Monash University)

Ngo Van Long (McGill University, and Hitotsubashi Institute for Advanced Study, Hitotsubashi University)

Abstract

We propose a dynamic model of climate change abatement in which the number of contributors is endogenous and thus may differ between two modes of cooperation, namely, loose vs tight. In the tight mode of cooperation, each member is prescribed a specific target, whereas in the loose one, members choose their own abatement levels as Nash players. Conditions exist such that the incentive to free ride is lower and the number of contributors is higher in the loose cooperation framework, and this can lead to higher welfare, both in the steady state and along the transition path. Our theoretical results suggest that the loose coalition mode, such reflected in the spirit of the Paris International COP21 Conference on Climate Change, by attracting more participants, could turn out to be more effective in reducing emissions than the Kyoto Protocol.

JEL Classification: Q2, Q52, C73.

Keywords: differential games; pollution abatement; climate change; mode of cooperation.

1 Introduction

The literature on international environmental agreements (IEAs) assumes that a target is set for each country as part of the agreement - tight cooperation (see, e.g., Carraro and Siniscalco, 1993; Barrett, 1994; Rubio and Ulph, 2006; Eichner and Pethig, 2013). A robust theoretical result under that assumption is that the equilibrium number of signatories to a tight IEA is very small.[1] In practice, proposed agreements that prescribe tight cooperation often fail to attract a sufficient number of truly committed signatories. For example, the Kyoto Protocol, which was adopted in 1997 and entered into force in February 2005, 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. Anticipation that no country will meet their prescribed targets has led to the Paris Agreement in 2015, which used a distinctly different approach from the Kyoto Protocol. This new approach may be termed a loose mode of cooperation. At the Paris International COP21 Conference on Climate Change, countries have agreed on an overall objective of limiting global warming to 2 degrees C relative to the pre-industrial temperature, but no country is required to set a specific target by a specific date. Signatories are free to determine their own target, and there is no penalty if a target is not met. There were some indications that the this form of loose agreement, by attracting more participants, could turn out to be more effective in reducing emissions than the Kyoto Protocol.[2]

A reason why the IEAs with a tight mode of cooperation do not attract many truly committed signatories is that the incentive to free ride is very strong. In this paper, we develop a model of loose cooperation, where countries/members agree in principle to aggregate targets but do not commit to country-specific targets. One of the advantages of loose targets is that more potential contributors are willing to join. To the best of our knowledge, this paper represents the first theoretical attempt to show that, compared to tight IEAs, a looser form of cooperation can work better in terms of improving welfare.

We propose and analyze a dynamic (continuous time) model of voluntary abatement of a public bad that departs from the existing literature in two ways: (i) the number of contributors is endogenous and dependent on the mode of cooperation; (ii) in addition to the standard cost which is increasing in the amount of contribution, abating agents must also incur a fixed participation cost.[3]

[1] Barrett performed numerical simulations. The results obtained were proven analytically in Rubio and Ulph (2006).

[2] Indeed, 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).

[3] Classical references on the voluntary provision of public goods include Chamberlin (1974), Bergstrom, Blume and Varian (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),

As to (i), we take a similar approach to that taken by Rubio and Ulph (2007) and de Zeeuw (2008), who study the dynamic incentives for countries to sign international environmental agreements, with one important difference.[4] Rubio and Ulph (2007) and de Zeeuw (2008) assume that members of an IEA fully coordinate on their contributions. We, instead, do not restrict our attention to that case, but consider also the scenario in which active contributors decide non-cooperatively on how much to contribute. As to (ii), the existence of a fixed participation cost, besides its theoretical interest, makes the model applicable to a wide range of real-world problems. Consider, for example, the Paris Accord mentioned above. There are fixed costs of administration for each country, e.g., they have to set up a bureaucratic machinery for the operation of the emission-permits market and/or the collection of carbon taxes, assessment of firms' emission levels, auditing firms' report on emissions of carbon, etc. Besides administrative costs, the fixed cost may include the political economy cost, for instance, the cost of overcoming opposition from pressure groups to pass the legislation and the cost of redistributing from winners to losers within the country.[5]

We compare the welfare outcomes across two scenarios of private contributions to abatement of a public bad (such as the stock of pollution) with fixed participation cost, such that in equilibrium some agents choose to be free riders, thus avoiding the fixed cost. In the non-coordination scenario, after paying the fixed cost, contributors choose their contribution levels as Nash players. In the coordination scenario, those who have chosen to be contributors must coordinate on their contribution levels. The main result of our analysis is that there exist parameter regions in which non-coordination results in an increase in both the number of contributors and the aggregate contribution, and consequently, higher social welfare. Loose cooperation can be welfare-improving not only in the steady state equilibrium but also in terms of present value of net utility.

Another advantage of loose cooperation is that it allows individual countries the flexibility to respond to idiosyncratic shocks. Many of these shocks are related to political economy considerations, such as discontents from a powerful segment of the electorate. Shocks of this nature are quite often private information (i.e., not verifiable by third parties) and therefore state-contingent transfer payments are not feasible (Bagwell and Staiger, 2005; Amador and Bagwell, 2013). These are important considerations. However, for simplicity, as a first step, we do not model idiosyncratic shocks and asymmetry among countries.

There are many examples of successful loose cooperation in world history. Of note is the

Fujiwara and Matsueda (2009), Battaglini *et al.* (2014), Georgiadis (2015, 2017), and Bowen *et al.* (2019), *inter alia.* All of these papers suppose that there are no fixed participation costs.

[4] 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 IEAs by Carraro and Siniscalco (1993) and Barrett (1994). On the stability of oligopolistic cartels and IEAs see also Diamantoudi (2005) and Diamantoudi and Sartzetakis (2015), respectively.

[5] Another example of contributions to a pure public good with fixed cost of participation is R&D with complete spillovers. The fixed participation cost arises because a firm that does R&D must incur the cost of setting up a research lab. Similarly, those who want to contribute articles to Wikipedia must at first incur the fixed costs of learning the technical side of navigating that website.

Hanseatic League, a Central European loose confederation of merchant guilds and market cities that came to dominate Baltic maritime trade for over 300 years, reaching its peak in the 15th century, with over 100 member cities (Atatüre, 2008). A more recent example is the Association of South East Asian Nations (ASEAN) which promote free trade among member countries without requiring them to abide by strict rules. In modern democracies, political parties are notable instances of loose cooperation. In the USA, for example, Republicans and Democrats can move in and out of their parties without penalties, and while political donations are encouraged, party members are not asked to commit to specific donations. Similarly, religious organizations typically do not insist on tight cooperation. It seems that our theory of loose cooperation can be useful to understand phenomena outside the domain of environmental economics.

The remainder of this paper is organized as follows. The model is specified in Section 2. Sections 3 and 4 analyze the non-cooperative and the cooperative case, respectively. A welfare comparison is provided in Section 5. Section 6 concludes.

2 The Set-up

Consider an economy populated by $n \geq 2$ infinitely-lived agents. Time is continuous and denoted by $t \in [0, \infty)$. At $t = 0$, each agent decides whether or not to contribute to the abatement of a stock of a public bad (henceforth, the stock of pollution). Contributors are denoted by the index $i = 1, ..., m$, where $m \leq n$. The remaining $n - m$ individuals are called free riders, and we use the index j to refer to free riders. We assume that to be a contributor, an agent must incur a fixed participation cost $F > 0$.

All agents are *ex-ante* identical. Each agent's maximum productive capacity of the final consumption good is a positive number a, which we refer to as their "business-as-usual" level of output. Each unit of output generates a unit of emission of a pollutant. The stock of pollution is a public bad. Individuals realize that if they produce the consumption good at their maximum productive capacity, they will each add a units of pollutant to the accumulation of the public bad. Cutting output below the maximum capacity is referred to as "abatement." If an individual i chooses the abatement level $x_i(t)$, where $0 \leq x_i(t) \leq a$, he will have only $a - x_i(t)$ units of output to consume. We assume that the direct utility derived from consumption is quadratic:

$$U(a - x_i(t)) \equiv [a - x_i(t)] - [a - x_i(t)]^2,$$

where we assume that $a < 1/2$, so that any $x_i > 0$ constitutes a sacrifice of direct utility. The stock of pollution is denoted by $k(t) \geq 0$. At any time $t \geq 0$, the stock $k(t)$ inflicts a damage flow $bk(t)^2$ to each agent, where $b > 0$. The instantaneous (net) utility function of a contributor is defined as

$$u_i(k(t), x_i(t)) \equiv U(a - x_i(t)) - b[k(t)]^2 \equiv [a - x_i(t)] - [a - x_i(t)]^2 - b[k(t)]^2,$$

where $x_i(t) \in [0, a]$.[6] The instantaneous utility function of a non-contributor is given by

$$u_j(k(t)) \equiv u_i(k\,(t)\,, 0) \equiv a - a^2 - b[k(t)]^2.$$

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

$$\frac{dk\,(t)}{dt} = na - \sum_{i=1}^{m} x_i\,(t) - \delta k\,(t)\,, \tag{1}$$

where δ is the natural decay rate of the stock of pollution, with $\delta > 0$. By (1), the addition to the stock of pollution is increasing in na and it decreases as the sum of abatement levels $(\sum_{i=1}^{m} x_i\,(t))$ increases.

Let r be the positive rate at which future payoffs are discounted. Once an agent has decided to be a contributor, his objective functional is

$$J_i^C \equiv \int_0^\infty e^{-rt} \left\{ -b\,[k\,(t)]^2 + a - x_i\,(t) - [a - x_i\,(t)]^2 \right\} dt.$$

The payoff of a non-contributor (i.e., a free rider) is

$$J_j^N \equiv \int_0^\infty e^{-rt} \left\{ -b\,[k\,(t)]^2 + a - a^2 \right\} dt. \tag{2}$$

In our model, all agents are identical *ex-ante*, but they behave differently *ex-post.* We focus on the case where in equilibrium there are exactly two groups: a group G^C consisting of m contributors (where m is endogenously determined) who, in equilibrium, contribute the same amount, $x_i^* > 0$ for all $i \in G^C$, and a group G^N consisting of $n - m$ free riders, whose contribution is nil, $x_j = 0$ for all $j \in G^N$.[7]

At the beginning of the game ($t = 0$), individuals decide on whether they are members of group G^C or of group G^N. We will refer to $t = 0$ as Stage 1 and to the time period $(0, \infty)$ as Stage 2. Accordingly, we will refer to the decision at $t = 0$ as Stage 1 decision, and to the decisions at $t \in (0, \infty)$ as Stage 2 decisions. In Stage 1, each member of group G^C must incur a participation cost, $F > 0$. In Stage 2, members of group G^C, fully informed about the size of G^C, choose their contribution levels, while members of group G^N (who by choice did not incur the participation cost F in Stage 1) are free riders: their contribution is nil.

Let V_i^C and V_j^N denote the value functions of a contributor and a free rider, respectively. (These functions are obtained by solving Stage 2 game). Given V_i^C and V_j^N, for a (Stage 1) equilibrium with m contributors to exist it must be that each $i \in G^C$ has no unilateral incentive to deviate and

[6] We will impose parameter restrictions which guarantee that u_i is decreasing in k and increasing in a, and that consumption net of the abatament, $a - x_i$, is nonnegative.

[7] We do not consider asymmetric equilibria where abatement levels differ across the m contributors.

join G^N, and each $j \in G^N$ has no unilateral incentive to deviate by joining G^C. Formally, for each $i \in G^C$, it must hold that[8]

$$V_i^C(k_0; m) - F \geq V_j^N(k_0; m-1), \tag{3}$$

and, for each $j \in G^N$, it must hold that

$$V_j^N(k_0; m) \geq V_i^C(k_0; m+1) - F. \tag{4}$$

Inequality (3) states that no contributor would gain by deviating, i.e., it must be individually rational to contribute (in which case the total number of contributors is m) rather than free riding (in which case, the total number of contributors becomes $m-1$).[9] We will refer to condition (3) as the *contributor-rationality* condition.

Inequality (4) states that, for a free rider, it must be better to contribute nil rather than joining group G^C (in which case it assumes that the number of contributors becomes $m+1$, and that all of them contribute at their new symmetric equilibrium level with $m+1$ contributors). We will refer to condition (4) as the *free-rider-rationality* condition.

As is well-known in the literature, if a plan is optimal at time t=0 it does not necessarily mean that it will remain optimal at time $\tau > 0$: time consistency at $t = 0$ and $k(0) = k_0$ requires that, at any future date $\tau \in (0, \infty)$, each contributor (resp. free rider) will find it individually rational to continue to contribute (resp. not contribute). Formally, for each $i \in G^C$, it must hold that

$$V_i^C(k(\tau; m); m) \geq V_j^N(k(\tau; m); m-1), \tag{5}$$

and, for each $j \in G^N$, it must hold that

$$V_j^N(k(\tau; m); m) \geq V_i^C(k(\tau; m); m+1) - F, \tag{6}$$

where $k(\tau; m)$ denotes the equilibrium state trajectory with $m < n$ contributors. If (5) and (6) are both satisfied then the decisions made at $t = 0$ by the contributors and the free riders will be time-consistent, i.e., they will remain optimal (payoff-wise) at any future date.[10]

In what follows, we will consider two scenarios of contributions to the abatement of the stock of pollution with fixed cost of participation, namely, a non-coordination scenario and a coordination scenario. In the non-coordination scenario, in Stage 2, contributors choose their contribution levels taking the contributions of the others as given (i.e., as Nash players). We assume that contributors

[8] These inequalities are standard in the tight coalitions literature (see D'Aspremont *et al.*, 1983; Barret, 1994) but they have never been applied to loose coalitions.

[9] The agent assumes that if he leaves group G^C to join group G^N, then (i) the number of contributors will become $m-1$, i.e., the other $m-1$ contributors stay in G^C; and (ii) the remaining contributors will adjust their contribution level to the new equilibrium level with $m-1$ contributors.

[10] Note that in (5) and (6) the value functions are evaluated under the requirement that the number of contributors is equal to $m < n$ and constant from time t $= 0$ until the comparison time τ. Note also that the fixed cost (paid at the comparison time τ) appears only in (6). This is so because at time τ the fixed cost has to be paid only by a free rider who decides to join the group of contributors; a contributor who at time $\tau > t$ decides to join the group of free riders cannot recover the fixed cost that he paid at time $t = 0$.

use (stationary) feedback strategies, i.e., they condition their contributions at time t on the current level of the stock of pollution, exclusively. In the coordination scenario, those who have chosen to be contributors collectively decide on their common contribution level with the aim of maximizing the discounted sum of their utilities.

3 Non-coordination Behavior

We consider Stage 2 decisions first. Under non-coordination, each contributor i takes the strategies of other contributors as given. We must then solve for a non-cooperative feedback equilibrium of the dynamic game. Given a number of contributors $m \geq 1$, non-cooperative feedback equilibrium strategies of a generic contributor have to satisfy the following Hamiltonian-Jacobian-Bellman (HJB) equation, where $V_i^C(k;m)$ denotes the value function of a contributor, given that there are m contributors:

$$\begin{aligned} rV_i^C(k;m) &= \max_{x_i\in[0,a]} \Big\{ -bk^2 + a - x_i - (a-x_i)^2 \\ &\quad + \frac{dV_i^C(k;m)}{dk}\left(na - x_i - X_{-i}(k) - \delta k\right) \Big\}, \end{aligned} \tag{7}$$

where $X_{-i}(k) = \sum_{h\neq i}^{m} x_h(k)$, together with the usual transversality condition. Maximization of the right-hand side of (7) gives (for an interior solution):[11]

$$x_i^*(k) = a - \frac{1}{2}\left[1 + \frac{dV_i^C(k;m)}{dk}\right]. \tag{8}$$

Proposition 1 *The feedback equilibrium strategy of a contributor under non-cooperative behavior is given by ($i = 1, 2, ..., m$)*

$$x_i^*(k) = a - \frac{1 + A^C k + B^C}{2}$$

for k such that $x_i^*(k) \in (0,a)$, where $A^C, B^C < 0$ are constants (given in the Appendix) that depend on the parameters of the model.

Proof. See Appendix A. ■

Corollary 1 *The vector of strategies $(x_1^*(k), x_2^*(k), ..., x_m^*(k))$ induces a trajectory of k given by*

$$k^*(t) = k_{ss}^* + e^{-\phi t}(k_0 - k_{ss}^*)$$

where $\phi > 0$ is the speed of convergence (given in the Appendix) and k_{ss}^* is the steady state of k under non cooperative behavior, given by

$$k_{ss}^* = \frac{m\left(1 + B^C\right) + 2a\,(n-m)}{2\delta - A^C m}$$

which is stable.

[11] For an interior solution $x_i^*(k) \in (0,a)$ it must be that $dV_i^C(k)/dk \in (-1, 2a-1)$. We will verify later that conditions on the parameters of the model exist such that $dV_i^C(k)/dk \in (-1, 2a-1)$ at any point in time, implying that the equilibrium trajectory of x_i^* remains between 0 and a throughout the entire planning horizon.

Proof. See Appendix B. ■

Two remarks are in order. First, the equilibrium strategy given in Proposition 1 is for an interior solution, $x_i^*(k) \in (0, a)$. Given the purpose of our analysis, corner solutions are clearly not interesting: when $x_i^*(k) = 0$ the distinction between contributors and free riders vanishes; when $x_i^*(k) = a$ abatement levels become state-independent and private consumption turns out to be nil. Second, the equilibrium strategy given in Proposition 1 is linear in k, with a positive slope equal to $-A^C/2 > 0$. Thus $x_i^*(k)$ is increasing in k: the higher the stock of pollution, the higher the contribution to the abatement by agent $i = 1, 2, ..., m$, for any given $m \geq 1$.[12]

Figure 1 below depicts the abatement strategy $x_i^*(k)$ for the case in which $x_i^*(0) > 0$.[13]

[Insert Fig. 1 about here]

Figure 2 below illustrates the steady-state equilibrium. (The aggregate emission is measured along the vertical axis.)

[Insert Fig. 2 about here]

Let us now turn to the following question: does an increase in the number of non-cooperative contributors lead to a lower long-run stock of pollution? Interestingly, the impact of m on k_{ss}^* is ambiguous, depending on the parameters of the model. For instance, if we evaluate the derivative of k_{ss}^* w.r.t. m at $m = 1$ we find that the sign depends on a. Specifically, if a is below a certain threshold, $\widehat{a}$, then the derivative is positive (therefore k_{ss}^* increases in m), otherwise it is negative. Surprisingly, having more contributors to the abatement of the public bad can make things worse (i.e., the long-run stock of the public bad is bigger when there is an additional contributor). The intuition is as follows. We note that the smaller a is, the greater is the marginal loss of direct utility caused by a given abatement level x_i (because, for any given x_i, the marginal utility of consuming $a - x_i$ is higher when a is lower). Therefore, when a is small, a contributor's best reply to an increase in the sum of contributions by others tends to be a big reduction in her own contribution. This can lead to the result that an increase in m leads to a decrease in total abatement, and hence an increase in k_{ss}^*.

The impact of an increase in m (starting from $m = 1$) on the steady-state stock of pollution is illustrated in Figure 3. As shown in Figure 3, when the number of contributors increases from 1 to a higher number, the steady-state stock k_{ss}^* moves to the right (respectively, left) if $a < \widehat{a}$ (respectively, $a > \widehat{a}$).

[Insert Fig. 3 about here]

[12] As is well-known in the differential game literature, the linear feedback strategy is only one of the infinitely many feedback strategies that satisfy the differential equation resulting from differentiating the maximized HJB equation w.r.t. the state variable. However, value functions associated with nonlinear feedback strategies can be obtained only implicitly, whereas value functions associated with linear feedback strategies are polynomials of degree two, and can be easily used for the derivation of the equilibrium number of contributors.

[13] All the pictures in the paper are drawn for parameter values such that $x_i^*(0) > 0$. The definition of $\overline{k}$ is given in the Appendix (see Appendix A).

Remark 1 *The above surprising result (i.e., an increase in m leads to a decrease in total abatement) cannot be obtained in a corresponding static version of the model. In the static version, with m contributors and n agents, contributor i chooses x_i to maximize*

$$(a - x_i) - (a - x_i)^2 - b(k_0 + na - X_{-i} - x_i)^2$$

The FOC is

$$-1 + 2(a - x_i) + 2b(k_0 + na - X_{-i} - x_i) = 0$$

Let $x^*_{i,m}$ denote the symmetric Nash equilibrium, then

$$-1 + 2(a - x^*_{i,m}) + 2b(k_0 + na - mx^*_{i,m}) = 0$$

It follows that

$$x^*_{i,m} = \frac{2b(k_0 + na) + 2a - 1}{(2bm + 2)} \text{ and } x^*_{i,(m+1)} = \frac{2b(k_0 + na) + 2a - 1}{(2b(m + 1) + 2)}$$

Clearly, an increase in m will always increase total abatement and therefore decreases the pollution:

$$mx^*_{i,m} = \frac{2b(k_0 + na) + 2a - 1}{2b + \frac{2}{m}} < \frac{2b(k_0 + na) + 2a - 1}{2b + \frac{2}{m+1}} = (m + 1)x^*_{i,(m+1)}.$$

In view of the above Remark, one may ask why there is no static counterpart to the result (obtained in our dynamic game model) that a bigger m can lead to increased pollution. The key to the answer is that in a dynamic model, each agent expects that if he increases his emission today, the pollution stock will be bigger tomorrow, which would in turn induce other agents to emit somewhat less tomorrow than otherwise; this dynamic strategic consideration may give him an incentive to undertake less abatement today when he learns that the number of contributors has increased. In a static model, by definition, such dynamic strategic considerations do not exist.

The impact of m on x^*_{ss} is also ambiguous: the derivative of x^*_{ss} w.r.t. m evaluated at $m = 1$ is positive if a is below a certain threshold, $\widetilde{a}$, otherwise it is negative. One can verify that $\widehat{a} > \widetilde{a}$. Starting from $m = 1$, if $a < \widetilde{a}$ then both k^*_{ss} and x^*_{ss} increase in m; if $\widetilde{a} < a < \widehat{a}$ then k^*_{ss} increases and x^*_{ss} decreases in m; if $a > \widehat{a}$ then both k^*_{ss} and x^*_{ss} decrease in m (see Appendix C).

How does m impact on the speed of convergence, ϕ? We find that ϕ is increasing in m for every $m \geq 1$: more contributors imply faster convergence to the steady state.

Using x^*_i given in Proposition 1 we can compute the value function of a free rider, which has to satisfy the following HJB equation:

$$rV_j^N(k;m) = -bk^2 + a(1 - a) + \frac{dV_j^N(k;m)}{dk}\left(na - mx^*_i(k) - \delta k\right).$$

It can be checked that $V_j^N(k;m) = A^N k^2/2 + B^N k + E^N$, with $A^N < 0$, $B^N < 0$, and E^N given in the Appendix (see Appendix D).

Let us now turn to the Stage 1 decision and the investigation of the existence and uniqueness (or otherwise) of the equilibrium number of contributors under non-coordination, m^*. Clearly, the contributor-rationality condition (3) is satisfied if and only if $F \leq \overline{F}^{(m)}$, where

$$\begin{aligned} F &\leq \overline{F}^{(m)} \equiv V_i^C(k_0; m) - V_j^N(k_0; m-1) \\ &= \left[A^C(m) - A^N(m-1)\right] k_0^2/2 + \left[B^C(m) - B^N(m-1)\right] k_0 + E^C(m) - E^N(m-1), \quad (9) \end{aligned}$$

and the free-rider-rationality condition (4) is satisfied if and only if $F \geq \underline{F}^{(m)}$, where

$$\begin{aligned} F &\geq \underline{F}^{(m)} \equiv V_i^C(k_0; m+1) - V_j^N(k_0; m) \\ &= \left[A^C(m+1) - A^N(m)\right] k_0^2/2 + \left[B^C(m+1) - B^N(m)\right] k_0 + E^C(m+1) - E^N(m) \quad (10) \end{aligned}$$

Interestingly, the curve $\underline{F}^{(m)}$ is graphically a horizontal left-ward displacement of the curve $\overline{F}^{(m)}$, because of the following property:

Lemma 1 $\overline{F}^{(m)} = \underline{F}^{(m-1)}$.

Proof. By definition. ■

If both $\overline{F}^{(m)}$ and $\underline{F}^{(m)}$ are monotone decreasing in m (in the first quadrant) then the graph of the curve $\underline{F}^{(m)}$ lies below that of the curve $\overline{F}^{(m)}$, with $\overline{F}^{(m)} > \underline{F}^{(m)} > 0$ for all $m \in [0, \widetilde{m} - 1]$ and $\overline{F}^{(m)} > \underline{F}^{(m)} < 0$ for all $m \in (\widetilde{m} - 1, \widetilde{m})$, with $\widetilde{m}$ being the unique real number such that $\overline{F}^{(\widetilde{m})} = 0$ (assuming that $\overline{F}^{(m)}$ intersects the horizontal axis). Let $\lfloor \widetilde{m} \rfloor$ denote the largest integer that is less than or equal to $\widetilde{m}$. Assume that $n > \lfloor \widetilde{m} \rfloor$. Denote by $\overline{S}$ the set of $\lfloor \widetilde{m} \rfloor$ critical values of fixed cost at which the constraint (3) is binding:

$$\overline{S} \equiv \left\{ \overline{F}^{(1)}, \overline{F}^{(2)}, ..., \overline{F}^{(\lfloor \widetilde{m} \rfloor)} \right\}.$$

Clearly, if the fixed participation cost F is such that $\overline{F}^{(s)} > F > \overline{F}^{(s+1)}$ for some integer $s \in \{1, 2, ..., \lfloor \widetilde{m} \rfloor\}$, then the *unique* equilibrium configuration of contributors and free riders $(m^*, n-m^*)$ is $(s, n-s)$. To see this, note that, by Lemma 1, $\overline{F}^{(s)} > F > \overline{F}^{(s+1)}$ is equivalent to $\overline{F}^{(s)} > F > \underline{F}^{(s)}$, which means that both rationality constraints (3) and (4) are satisfied, and therefore $(s, n-s)$ is an equilibrium configuration. Uniqueness follows from the fact that if F satisfies the condition $\overline{F}^{(s)} > F > \underline{F}^{(s)}$, then F cannot satisfy $\overline{F}^{(h)} \geq F \geq \underline{F}^{(h)}$ for any other integer $h \neq s$.

It follows that for *almost every* F in the real interval $\left[\overline{F}^{(1)}, \overline{F}^{(\lfloor \widetilde{m} \rfloor)}\right]$, there is a unique equilibrium configuration of contributors and free riders. The only cases in which non-uniqueness of equilibrium could arise are when F happens to coincide with one member of the set $\overline{S}$ defined above. In such cases, there are exactly two equilibria. More precisely, for $s = 1, 2, ..., \lfloor \widetilde{m} \rfloor$, if $F = \overline{F}^{(s)}$, then there are exactly two equilibrium configurations, $(s, n-s)$ and $(s-1, n-(s-1))$. At the former equilibrium, the rationality constraint of contributors is binding. At the latter one, the rationality constraint of free riders is binding.

The above discussion can be summarized by the following proposition:

Proposition 2 *In the non-cooperative scenario, if* $\underline{F}^{(m^*)} < F < \overline{F}^{(m^*)}$ *then there exists a unique equilibrium in which* m^* *agents contribute and* $n - m^*$ *agents free ride, with* $n > \lfloor \widetilde{m} \rfloor$, $\lfloor \widetilde{m} \rfloor$ *being the largest integer that is less than or equal to* $\widetilde{m}$.

Corollary 2 *If* $\overline{F}^{(m)} > 0$ *for all* $m \geq 0$ *or* $n < \lfloor \widetilde{m} \rfloor$, *conditions on the parameters of the model exist under which* $m^* = n$, *i.e. there are no free riders.*

The following numerical example illustrates Proposition 1: Let the parameter values be $a = 0.4825$, $b = 0.012$, $n = 10$, $\delta = 0.5$, $k_0 = 0$, $r = 0.1$ and $F = 0.045$. Can $m = 10$ be the equilibrium number of contributors? As $\underline{F}^{(10)} = 0.0423$ and $\overline{F}^{(10)} = 0.0499$, clearly if $F \in (0.0423, 0.0499)$ then $m^* = 10$.[14]

One can also check that $m^* = 10$ satisfies the time-consistency condition (5), since $V_i^C(k(\tau); 10) - V_j^N(k(\tau); 9) = 0.0616 + 0.0029e^{-1.1877\tau} - 0.0146e^{-0.5939\tau} > 0$ for all $\tau \in [0, \infty)$. Therefore, the decisions made at $t = 0$ by all agents to join group G^C are time-consistent.

4 Cooperative Behavior

We now turn our attention to the cooperative scenario. We need to derive $V_i^C(k_0; m|coop)$ and $V_j^N(k_0; m|coop)$ under the assumption that those who contribute must act cooperatively, i.e., they must coordinate their contributions. Under cooperation among contributors, the objective functional of contributor i is given by

$$J_i^C(k_0; m|coop) \equiv \int_0^\infty e^{-rt} \left\{ -b [k(t)]^2 + \left[a - \frac{X(t)}{m} \right] - \left[a - \frac{X(t)}{m} \right]^2 \right\} dt,$$

and that of non-contributor j is given by (2), where $X(t)/m$ denotes the coordinated contribution of the representative contributor i.

As in the previous section, we consider Stage 2 decisions first. Given a number of contributors $m \geq 1$, equilibrium strategies of a generic contributor under cooperative behavior have to satisfy the following HJB equation:

$$rV_i^C(k; m|coop) = \max_{X \in [0, na]} \left\{ -bk^2 + a - \frac{X}{m} - \left(a - \frac{X}{m} \right)^2 + \frac{dV_i^C(k; m|coop)}{dk} (na - X - \delta k) \right\}, \tag{11}$$

together with the usual transversality condition. Maximization of the right-hand side of (11) gives (for an interior solution):[15]

$$X = m \left[a - \frac{1}{2} \left(1 + m \frac{dV_i^C(k; m|coop)}{dk} \right) \right]. \tag{12}$$

[14] One can check that $x^*(0) = 0.0428$, and $\lim_{t \to \infty} x^*(t) = 0.1123$, thus $a - x^*(t) > 0$ for all $t \in [0, \infty)$.

[15] For an interior solution $x_i^{**}(k) \in (0, a)$ it must be that $dV_i^C(k; m|coop)/dk \in (-1/m, (2a-1)/m)$. We will verify later that conditions on the parameters of the model exist such that $dV_i^C(k; m|coop)/dk \in (-1/m, (2a-1)/m)$ at any point in time implying that the equilibrium trajectory of x_i^{**} remains between 0 and a troughout the entire planning horizon.

Proposition 3 *The feedback equilibrium strategy of a contributor under cooperation is given by $(i = 1, 2, ..., m)$*

$$x_i^{**}(k) = a - \frac{1 + m\left(\alpha^C k + \beta^C\right)}{2}$$

for k such that $x_i^{**}(k) \in (0, a)$, where $\alpha^C, \beta^C < 0$ are constants (given in the Appendix) that depend on the parameters of the model.

Proof. See Appendix E. ■

Corollary 3 *The vector of strategies $(x_1^{**}(k), x_2^{**}(k), ..., x_m^{**}(k))$ induces a trajectory of k given by*

$$k^{**}(t) = k_{ss}^{**} + e^{-\sigma t}(k_0 - k_{ss}^{**})$$

where σ is the speed of convergence (given in the Appendix) and k_{ss}^{**} is the steady state of k under cooperative behavior, given by

$$k_{ss}^{**} = \frac{m\left(1 + \beta^C\right) + 2a\,(n - m)}{2\delta - \alpha^C m}$$

which is stable.

Proof. See Appendix F. ■

Since $\alpha^C < 0$, clearly $x_i^{**}(k)$ is increasing in k: the higher the stock of public bad, the higher the contribution to the abatement. This is the same as under non cooperation. It can be verified that $m\alpha^C < A^C$ for all $m > 1$, implying that $x_i^{**}(k)$ is steeper than $x_i^{*}(k)$. As the public bad increases, the response of contributors in the cooperative scenario is stronger than that of their counterparts in the non-cooperative scenario.

Figure 4 below shows that, if the number of contributors is greater than 1 and is the same under cooperation and under non-cooperation, then the long-run pollution stock under non-cooperation is higher than under cooperation, i.e., $k_{ss}^{*} > k_{ss}^{**}$.[16]

[Insert Fig. 4 about here]

Would an increase in the number of cooperative contributors result in a lower long-run stock of pollution? Surprisingly, the answer is: "It depends on parameter values." Indeed, the derivative of k_{ss}^{**} w.r.t. m evaluated at $m = 1$ is positive or negative, depending on whether the following expression is positive or negative:

$$\Phi(a) \equiv -\left\{2ab\,(2n - 1) + \delta\,(2a - 1)\,(\delta + r) + b\right\}.$$

Notice that $\Phi(a)$ is decreasing in a and is equal to zero at $a = \overline{a}$, where

$$\overline{a} = \frac{1}{2} - \frac{n}{2n + \left[(\delta/b)\,(\delta + r) - 1\right]} > 0 \text{ if } (\delta/b)\,(\delta + r) > 1.$$

[16] The definition of $\overline{\overline{k}}$ is given in the Appendix (see Appendix E).

Thus, if $a < \overline{a}$ (resp. $> \overline{a}$) then k_{ss}^{**} increases (resp. decreases) in m (starting from $m = 1$). Similar to the non-cooperative case, we find that in the cooperative case, a surprising result can arise: if a and b are small, having more contributors can lead to a greater stock of public bad in the long run. It is instructive to compare the threshold $\overline{a}$ of the cooperative case with the threshold $\widehat{a}$ of the non-cooperaive case. One can verify that $\widehat{a} > \overline{a}$. This implies that there exists an interval of a, (i.e. $\overline{a} < a < \widehat{a}$) where an increase in m (starting from $= 1$) leads to an increase in the steady-state stock of public bad under non-cooperation, but a decrease in the steady-state stock of public bad under cooperation.

The comparative steady-state analysis w.r.t. m is summarized in Figure 5.

[Insert Fig. 5 about here]

As in the case of non-cooperation, the speed of convergence, σ, increases in m for every $m \geq 1$. By comparing σ with ϕ, we see that, in general, σ can be either higher or lower than ϕ. However, if $r \to 0$, we have that $\phi > \sigma$. By continuity, $\phi > \sigma$ for r small enough. If agents are sufficiently patient then convergence is faster under non-cooperation than under cooperation.

Using x_i^{**} given in Proposition 3 we can compute the value function of a free rider, which has to satisfy the following HJB equation:

$$rV_j^N(k; m|coop) = \left\{ -bk^2 + a(1-a) + \frac{dV_j^N(k; m|coop)}{dk}\left(na - mx_i^{**}(k) - \delta k\right) \right\}.$$

We can show that $V_j^N(k; m|coop) = \alpha^N k^2/2 + \beta^N k + \epsilon^N$, with α^N, β^N, and ϵ^N given in the Appendix (see Appendix G).

Let us now turn to the investigation of the existence and uniqueness (or otherwise) of the equilibrium number of contributors under cooperation.

The following inequalities are the counterparts of inequalities (9) and (10) for the case of cooperation among contributors:

$$\begin{aligned} F \quad &\leq \quad \overline{\overline{F}}^{(m)} \equiv V_i^C(k_0; m|coop) - V_j^N(k_0; m-1|coop) \\ &= \quad \left[\alpha^C(m) - \alpha^N(m-1)\right] k_0^2/2 + \left[\beta^C(m) - \beta^N(m-1)\right] k_0 + \epsilon^C(m) - \epsilon^N(m-1), \end{aligned} \tag{13}$$

$$\begin{aligned} F \quad &\geq \quad \underline{\underline{F}}^{(m)} \equiv V_i^C(k_0; m+1|coop) - V_j^N(k_0; m|coop) \\ &= \quad \left[\alpha^C(m+1) - \alpha^N(m)\right] k_0^2/2 + \left[\beta^C(m+1) - \beta^N(m)\right] k_0 + \epsilon^C(m+1) - \epsilon^N(m). \end{aligned} \tag{14}$$

The following Lemma is the counterpart of Lemma 1 for the case of cooperation among contributors.

Lemma 2 $\overline{\overline{F}}^{(m)} = \underline{\underline{F}}^{(m-1)}$.

Proof. By definition. ■

If both $\overline{\overline{F}}^{(m)}$ and $\underline{\underline{F}}^{(m-1)}$ are monotone decreasing in m (in the first quadrant), and if there exists an $\widehat{m} > 0$ such that $\overline{\overline{F}}^{(\widehat{m})} = 0$, then the graph of the curve $\underline{\underline{F}}^{(m)}$ lies below that of the curve $\overline{\overline{F}}^{(m)}$, with $\overline{\overline{F}}^{(m)} > \underline{\underline{F}}^{(m-1)} > 0$ for all $m \in [0, \widehat{m}-1]$ and $\overline{\overline{F}}^{(m)} > \underline{\underline{F}}^{(m-1)} < 0$ for all $m \in (\widehat{m}-1, \widehat{m})$.

Proposition 4 *In the cooperative scenario, if* $\underline{\underline{F}}^{(m^{**})} < F < \overline{\overline{F}}^{(m^{**})}$ *then there exists a unique equilibrium in which* m^{**} *agents contribute and* $n - m^{**}$ *agents free ride, with* $n > \lfloor \widehat{m} \rfloor$, $\lfloor \widehat{m} \rfloor$ *being the largest integer that is less than or equal to* $\widehat{m}$.

Corollary 4 *If* $\overline{\overline{F}}^{(m)} > 0$ *for all* $m \geq 0$ *or* $n < \lfloor \widehat{m} \rfloor$, *conditions on the parameters of the model exist under which* $m^{**} = n$, *i.e. there are no free riders.*

As in the non-cooperative scenario, there are two cases in which the equilibrium number of contributors is non-unique: (i) $F = \underline{\underline{F}}^{(m^{**})}$, with $\underline{\underline{F}}^{(m^{**})} > 0$, in which case there exist two equilibria, m^{**} and $m^{**}+1$; (ii) $F = \overline{\overline{F}}^{(m^{**})}$, with $\overline{\overline{F}}^{(m^{**})} > 0$, in which case there exist two equilibria, m^{**} and $m^{**}-1$.

The following numerical example illustrates Proposition 2: $a = 0.4825$, $b = 0.012$, $n = 10$, $\delta = 0.5$, $k_0 = 0$, $r = 0.1$, and $F = 0.045$. Can $m = 2$ be the equilibrium number of contributors in the cooperative case? We get $\underline{\underline{F}}^{(2)} = 0.0289$ and $\overline{\overline{F}}^{(2)} = 0.4518$, implying that if $F \in (0.0289, 0.4518)$ then $m^{**} = 2$.[17] One can also check that $m^{**} = 2$ satisfies the time-consistency condition (5), since $V_i^C(k(\tau); 2) - V_j^N(k(\tau); 1) = 0.4678 + 0.0008e^{-2.3515\tau} - 0.0167e^{-1.1758\tau} > 0$, and (6), since $F = 0.045 > 0.0313 + 0.0002e^{-2.3515\tau} - 0.0026e^{-1.1758\tau}$ for $\tau \in [0, \infty)$. Therefore, the decisions made at $t = 0$ by two agents to join group G^C and by eight agents to join group G^N are time-consistent.

Figure 6 below depicts the four thresholds of F, namely $\underline{F}$, $\overline{F}$, $\underline{\underline{F}}$ and $\overline{\overline{F}}$ (for the parameter values used in the previous numerical examples).

[Insert Fig. 6 about here]

As can be seen in Fig. 6, suppose we draw a horizontal line representing a give fixed cost level F and suppose this line intersects the vertical line (representing a given integer m, say $m = 2$) at a point in the region $CDGF$ then we can deduce that the number of contributors in the cooperative and the non-cooperative scenarios is the same m (see for instance point H in Fig. 6).[18] Outside this region, the numbers of contributors in the two scenarios are different from each other: if the horizontal line representing F is above C then $m^{**} > m^{*}$; if F is below E then $m^{**} < m^{*}$. For $F = 0.045$, the number of contributors is 10 under non-cooperation (see point B in Fig. 6) and 2 under cooperation (see point A in Fig. 6).

[17] One can check that $x^{**}(0) = 0.1415$, and $\lim_{t\to\infty} x^{**}(t) = 0.3176$, so that $a - x^{**}(t) > 0$ for all $t \in [0, \infty)$. Moreover, $\lim_{t\to\infty} k^{**}(t) = 8.3796$. Note that, compared with the non-cooperative case, where $m^{*} = 10$, we have $\lim_{t\to\infty} x^{**}(t) > \lim_{t\to\infty} x^{*}(t)$ but, because $m^{**} = 2 < 10$, $\lim_{t\to\infty} k^{**}(t) > \lim_{t\to\infty} k^{*}(t)$.

[18] For $a = 0.4825$, $b = 0.012$, $n = 10$, $\delta = 0.5$, $k_0 = 0$ and $r = 0.1$ we get $\underline{F}^{(2)} = 0.167$ and $\overline{F}^{(2)} = 0.2023$. Hence, if $F = 0.2$ then $m^{*} = m^{**} = 2$ (recall that $\underline{\underline{F}}^{(2)} = 0.0289$ and $\overline{\overline{F}}^{(2)} = 0.4518$).

The trajectories of $k^*(t)$ and $k^{**}(t)$ (for the parameter values used in the previous numerical examples) are depicted in Figure 7.

[Insert Fig. 7 about here]

As shown in Figure 7, under the parameter values stated above, the stock of public bad turns out to be higher under cooperation than under non-cooperation not only at the steady state but also at any point in time during the transition phase. This is so because the free riding incentive is stronger when potential free-riding agents know that the active contributors will coordinate their contributions rather than behaving as Nash players.

Finally, Figures 8a and 8b below plot the trajectories of individual and total contributions in the cooperative and non-cooperative scenarios (for the parameter values used in the previous numerical examples).

[Insert Fig. 8 about here]

Individual contributions are higher under cooperation than under non cooperation at any point in time (see Fig. 8a). However, the number of contributors in the non cooperative scenario ($m^* = 10$) exceeds that in the cooperative scenario ($m^{**} = 2$), and total contribution turns out to be higher in the former than in the latter (see Fig. 8b).

5 Welfare Comparison between Cooperation and Non-cooperation

Welfare with m contributors and $n - m$ free riders is given by

$$W^* (k_0, m, n) = mV_i^C (k_0; m) + (n - m)V_j^N (k_0; m) - mF, \tag{15}$$

if the contributors do not coordinate their contributions, and by

$$W^{**} (k_0, m, n) = mV_i^C (k_0; m|coop) + (n - m)V_j^N (k_0; m|coop) - mF, \tag{16}$$

if the contributors act cooperatively. Equations (15) and (16) make use of the value functions and therefore measure the discounted sum of (net) utilities. From (15) and (16), using Corollaries 1 and 2, we can derive the steady-state welfare levels in the cooperative and non-cooperative scenarios, W_∞^{**} and W_∞^*, respectively, which we will use later in this section.

While the equilibrium number of contributors under non-coordination is in general not the same as that under coordination, it turns out to be useful to define, as an intermediate step in our analysis, for the same common m, the difference $\Delta W \equiv W^{**} - W^*$, with

$$\begin{aligned} \Delta W &= m \left[V_i^C (k_0; m|coop) - V_i^C (k_0; m)\right] \\ &\quad +(n - m) \left[V_j^N (k_0; m|coop) - V_j^N (k_0; m)\right]. \end{aligned}$$

It can be checked that the term between the first set of square brackets is positive. The term between the second set of square brackets can be rewritten as

$$b \int_0^\infty e^{-rt} \left\{[k^* (t)]^2 - [k^{**} (t)]^2\right\} dt,$$

since the consumption level of the non contributors is the same irrespective of whether the contributors cooperate or not. Under the hypothesis that the number of contributors is the same under cooperation and under non-cooperation, it is easy to verify that $k^*(t) > k^{**}(t)$ for any t, implying that the above integral is positive. We can then state the following (unsurprising) proposition:

Proposition 5 *For any given common $m \geq 1$, it holds that welfare is higher under cooperation than under non cooperation.*

The above proposition establishes that if the equilibrium m is the same, then the welfare outcome when the contributors coordinate their contributions is superior to the outcome under non-cooperation. However, when the endogenous number of contributors under non-cooperation is larger than that under cooperation, it is possible that non-cooperation may lead to higher welfare than cooperation. The following example illustrates this interesting result.

Numerical Example: $a = 0.4825$, $b = 0.012$, $n = 10$, $\delta = 0.5$, $k_0 = 0$, $r = 0.1$. Under the above parameter values, we determined before that $m^* = 10$ and $m^{**} = 2$. If $F = 0.045$ we find that welfare when the contributors cooperate is smaller than under non-cooperation:

$$W^{**}(2, 10|F = 0.045) = -42.1706 < W^*(10, 10|F = 0.045) = -28.842.$$

Going beyond numerical examples, let us formalize a key result of our paper:

Proposition 6 ***(Cooperation can reduce welfare when the number of free riders is endogenous)*** *Take parameter values such that $m^* > m^{**}$. If $m^* - m^{**}$ is large enough then there exists an interval of F such that $W^{**}(m^{**}, n) < W^*(m^*, n)$.*

Proof. See Appendix H. ■

For the parameter values in the numerical example above, the set of admissible fixed costs such that $m^* = 10$ and $m^{**} = 2$ is given by $[0.0289, 0.0499]$. Any $F \in [0.0289, 0.0499]$ would lead to $W^{**}(2, 10) < W^*(10, 10)$.

In the remainder of this section, we will compare the steady-state welfare level under cooperation with that under non-cooperation. Define

$$\begin{aligned} \Delta W_\infty \quad &\equiv \quad W_\infty^{**} - W_\infty^* = m\left[V_i^C(k_\infty^{**}; m|coop) - V_i^C(k_\infty^*; m)\right] \\ &\quad +(n-m)\left[V_j^N(k_\infty^{**}; m|coop) - V_j^N(k_\infty^*; m)\right]. \end{aligned}$$

It can be checked that $\Delta W_\infty > 0$. Hence, for the same number of contributors, we obtain the same qualitative result as that previously obtained with the discounted welfare measures. In general, the equilibrium number of contributors under non-cooperation differs from that under cooperation. Using the parameter values of the numerical example above we find that

$$W_\infty^{**}(2, 10|F = 0.045) = -61.6167 < W^*(10, 10|F = 0.045) = -42.9171,$$

which is in line with the discounted welfare comparison.

In conclusion, we have shown that there exists parameter values such that cooperation is welfare-reducing, both at the steady state and along the transition path.

6 Concluding Remarks

We have analyzed a dynamic game of voluntary abatement of a public bad with fixed cost of participation, in which the number of contributors is endogenously determined. One of our striking results is that cooperation among contributors can worsen social welfare and lead to a higher steady-state level of public bad compared to Nash behavior.

When the number of contributors is the same in the cooperative and the non-cooperative settings, social welfare turns out to be higher if contributors coordinate on their contribution levels rather than choosing their contribution levels as Nash players. However, our analysis has shown that the number of contributors in the cooperative and non-cooperative settings may differ significantly. Conditions on the parameters of the model exist such that the incentive to free ride is higher and, consequently, the number of contributors is lower in the cooperative setting, compared to the non-cooperative one, leading to a lower social welfare in the former than in the latter. A policy implication of this finding is that insisting on a tight coalition is not necessarily welfare improving. In the context of international environmental agreements, our analysis suggests that an accord in the spirit of COP 21, i.e., the Paris Accord, might be able to bring more countries on board, leading to higher total contribution to the abatement of the public bad compared to an agreement imposing specific targets for each country, which is what the Kyoto Protocol was trying to achieve.

Possible extensions of our framework include: (i) relaxing the assumption of full cooperation by considering a coefficient of cooperation ranging from zero to one (see, for instance, Vives, 2008; Colombo and Labrecciosa, 2018); (ii) allowing for agents heterogeneity (see McGinty, 2007; Pavlova and de Zeeuw, 2013); (iii) allowing for asymmetric information about agents' characteristics (see Bagwell and Staiger, 2005; Amador and Bagwell; 2013); (iv) accounting for uncertainty and ambiguity about the evolution of the stock of public bad (see, for instance, Lemoine and Traeger, 2014, 2016).

Acknowledgment: Part of this research was carried out at Hitotsubashi Institute for Advanced Study, Hitotsubashi University.

Figures

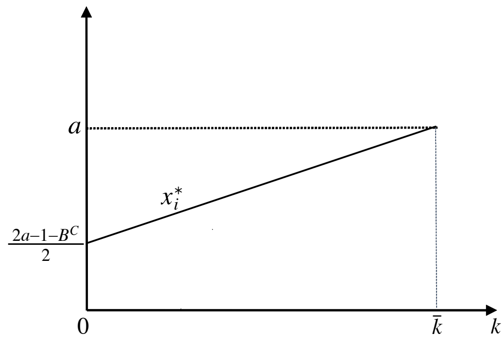

Figure 1: Equilibrium strategy of a contributor

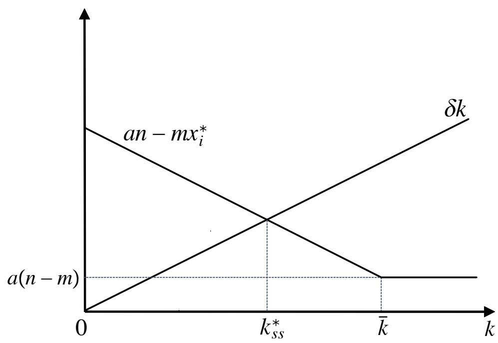

Figure 2: Steady-state equilibrium

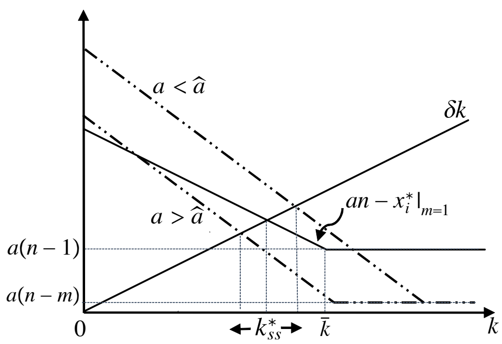

Figure 3: The impact of an increase in m on k_{ss}^*

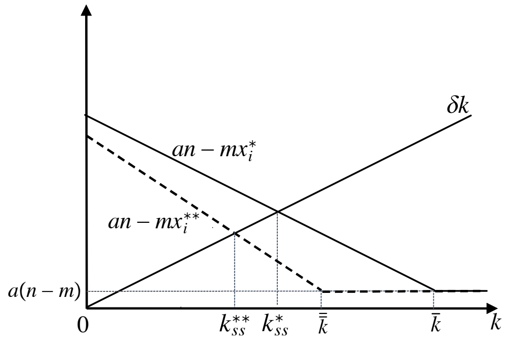

Figure 4: Steady-state comparison for a given m

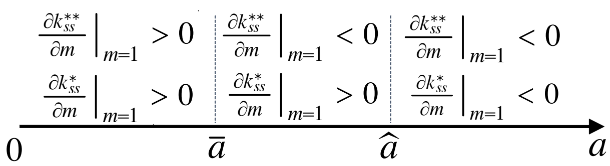

Figure 5: Comparative steady-state analysis w.r.t. m

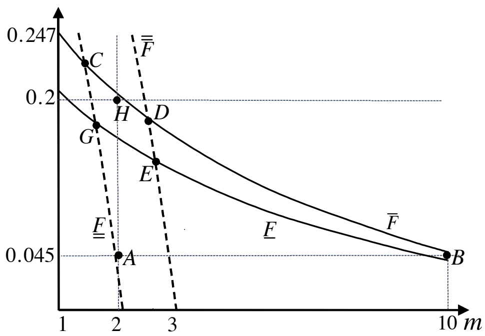

Figure 6: determination of m^* and m^{**}

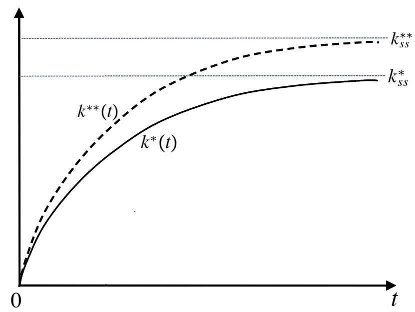

Figure 7: state trajectories comparison ($m^* = 10$, $m^{**} = 2$)

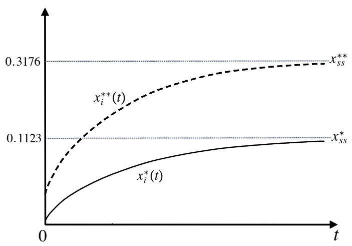

Figure 8a: individual contributions

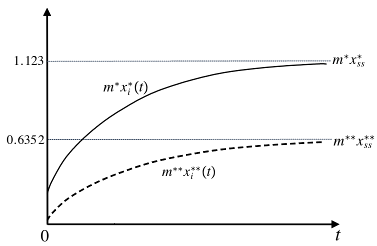

Figure 8b: total contributions

Appendix

Appendix A: Proof of Proposition 1

Inserting (8) into (7) yields (under symmetry)

$$\begin{aligned} rV_i^C(k;m) &= \frac{1}{4}\left\{1-4bk^2-\left[\frac{dV_i^C(k;m)}{dk}\right]^2\right\} \\ &\quad +\frac{dV_i^C(k;m)}{dk}\left\{an+\frac{m}{2}\left[1-2a+\frac{dV_i^C(k;m)}{dk}\right]-\delta k\right\}. \end{aligned} \tag{A.1}$$

Given the linear quadratic structure of the game at hand we guess a value function of the form

$$V_i^C(k;m) = A^C\frac{k^2}{2} + B^C k + E^C, \tag{A.2}$$

where A^C, B^C, and E^C are constants. Let $\Delta = \sqrt{4b\,(2m-1) + (2\delta + r)^2}$. It can be checked that (A.2) with

$$A^C = \frac{2\delta + r - \Delta}{2m-1} < 0,$$

$$B^C = \frac{A^C\,(2a\,(n-m)+m)}{A^C(1-2m) + 2\,(\delta + r)} < 0,$$

and

$$E^C = \frac{B^C\left[2m(1-2a) + 4an + B^C(2m-1)\right] + 1}{4r},$$

satisfies (A.1) for any $k \in (\underline{k}, \overline{k})$, with $\underline{k} = (2a-1-B^C)/A^C$ and $\overline{k} = -(1+B^C)/A^C$. For $k \notin (\underline{k}, \overline{k})$ we have a corner solution, either $x_i^* = 0$ or $x_i^* = a$. Specifically, $x_i^* = 0$ for $k < \underline{k}$ and $x_i^* = a$ for $k > \overline{k}$. Note that $\underline{k} \leq 0$ if a is sufficiently close to 1/2; in that case $x_i^*(k) > 0$ for all $k > 0$.

Appendix B: Proof of Corollary 1

The equilibrium trajectory of the stock of pollution under no cooperation, $k^*(t)$, is the solution of the following first-order differential equation:

$$\frac{dk(t)}{dt} = an - mx_i^*(k(t)) - \delta k(t),$$

with $x_i^*(k)$ given in Proposition 1. It can be checked that

$$k^*(t) = k_{ss}^* + e^{-\phi t}(k_0 - k_{ss}^*),$$

where

$$\phi = \frac{2\delta\,(1-m) - m\,(\Delta - r)}{2\,(1-2m)} > 0$$

is the speed of convergence to k_{ss}^*, with k_{ss}^* given in Corollary 1 (the expression of Δ is given in Appendix A). It is immediate to verify that $\lim_{t\to\infty} k^*(t) = k_{ss}^*$.

Appendix C: Thresholds of a

The derivative of k^*_{ss} w.r.t. m evaluated at $m = 1$ is decreasing in a and nil at $a = \widehat{a}$, with

$$\widehat{a} = \frac{-2b^2 + b(\delta + r)\left(\Delta|_{m=1} - 2\delta - r\right) - \delta(\delta + r)^2\left(\Delta|_{m=1} - r\right)}{2\left\{2b^2(n-1) + b(\delta + r)\left[(1-2n)\left(\Delta|_{m=1} - r\right) + 2\delta(n-1)\right] - \delta(\delta + r)^2\left(\Delta|_{m=1} - r\right)\right\}}.$$

The derivative of x^*_{ss} w.r.t. m evaluated at $m = 1$ is decreasing in a and nil at $a = \widetilde{a}$, with

$$\widetilde{a} = \frac{\delta\left(4\delta + 3r - \Delta|_{m=1}\right)}{2\left\{2bn - \delta\left[2\delta(n-2) + r(n-3) - (n-1)\Delta|_{m=1}\right]\right\}}.$$

Appendix D: Coefficients of $V_j^N(k; m)$

The coefficients of the value function of a non contributor under non cooperation are given by:

$$A^N = \frac{2b(1-2m)}{m\Delta + (2\delta + r)(m-1)}$$

$$\begin{aligned} B^N &= \left[\frac{m(\Delta - (2\delta + r))}{2(2m-1)} + \delta + r\right]^{-1} \\ &\quad \times A^N\left[\frac{m(2a(m-n) - m)(2b(2m-1) - (\delta + r)(\Delta - (2\delta + r)))}{4(2m-1)(b(2m-1) + \delta(\delta + r))} + a(n-m) + \frac{m}{2}\right] \end{aligned}$$

$$\begin{aligned} E^N &= r^{-1}\Big[a(1-a) + aB^N(n-m) + B^N\frac{m}{2} \\ &\quad + \frac{mB^N(2a(m-n) - m)(2b(m-1) - (\delta + r)\Delta - (2\delta + r))}{4(2m-1)(b(2m-1) + \delta(\delta + r))}\Big] \end{aligned}$$

with Δ given in Appendix A.

Appendix E: Proof of Proposition 2

Inserting (12) into (11) yields

$$\begin{aligned} rV_i^C(k; m|coop) &= \frac{1}{4}\Bigg\{1 - 4bk^2 + \frac{dV_i^C(k; m|coop)}{dk} \\ &\quad \times\left[2m(1-2a) - 4k\delta + 4an + m^2\frac{dV_i^C(k; m|coop)}{dk}\right]\Bigg\}. \end{aligned} \tag{E.1}$$

As in the non cooperative case, we guess a value function of the form

$$V_i^C(k; m|coop) = \alpha^C\frac{k^2}{2} + \beta^C k + \epsilon^C, \tag{E.2}$$

where α^C, β^C, and ϵ^C are constants. Let $\Gamma = \sqrt{4bm^2 + (2\delta + r)^2}$. It can be checked that (E.2) with

$$\alpha^C = \frac{2\delta + r - \Gamma}{m^2} < 0,$$

$$\beta^C = \frac{\alpha^C [2a(m-n)-m]}{\alpha^C m^2 - 2(\delta + r)} < 0,$$

and

$$\begin{aligned}\epsilon^C &= r^{-1}\Big[a(1-a) + aB^N(n-m) + B^N\frac{m}{2} \\ &+ \frac{mB^N(2a(m-n)-m)(2b(m-1)-(\delta+r)\Delta-(2\delta+r))}{4(2m-1)(b(2m-1)+\delta(\delta+r))}\Big],\end{aligned}$$

satisfies (E.1) for any $k \in (\underline{\underline{k}}, \overline{\overline{k}})$, with $\underline{\underline{k}} = ((2a-1)/m - \beta^C)/\alpha^C$ and $\overline{\overline{k}} = -(1/m + \beta^C)/\alpha^C$. For $k \notin (\underline{\underline{k}}, \overline{\overline{k}})$ we have a corner solution, either $x_i^{**} = 0$ or $x_i^{**} = a$. Specifically, $x_i^{**} = 0$ for $k < \underline{\underline{k}}$ and $x_i^{**} = a$ for $k > \overline{\overline{k}}$. Note that $\underline{\underline{k}} \leq 0$ if a is sufficiently close to 1/2; in that case $x_i^{**}(k) > 0$ for all $k > 0$.

Appendix F: Proof of Corollary 2

The equilibrium trajectory of the stock of pollution under cooperation, $k^{**}(t)$, is the solution of the following first-order differential equation:

$$\frac{dk(t)}{dt} = an - mx_i^{**}(k(t)) - \delta k(t),$$

with $x_i^{**}(k)$ given in Proposition 2. It can be checked that

$$k^{**}(t) = k_{ss}^{**} + e^{-\sigma t}(k_0 - k_{ss}^{**}),$$

where

$$\sigma = \frac{\Gamma - r}{2} > 0$$

is the speed of convergence to k_{ss}^{**}, with k_{ss}^{**} given in Corollary 2 (the expression of Γ is given in Appendix E). It is immediate to verify that $\lim_{t\to\infty} k^{**}(t) = k_{ss}^{**}$.

Appendix G: Coefficients of $V_j^N(k; m|coop)$

The coefficients of the value function of a non contributor under cooperation are given by:

$$\alpha^N = -\frac{2b}{\Gamma} < 0,$$

$$\beta^N = -\frac{\alpha^N(\delta + r)[m(2a-1) - 2an](\Gamma - r)}{2[bm^2 + \delta(\delta + r)](\Gamma + r)} < 0,$$

and

$$\begin{aligned}\epsilon^N &= r^{-1}\Big[a(1-a) + a\beta^N(n-m) + \beta^N\frac{m}{2} \\ &+ \frac{B^N[2a(m-n)-m]\left\{2bm^2 - (\delta+r)[\Gamma - (2\delta+r)]\right\}}{4[bm^2 + \delta(\delta+r)]}\Big],\end{aligned}$$

with Γ given in Appendix E.

Appendix H. Proof of Proposition 6

The difference $W^{**}(m^{**}, n) - W^{*}(m^{*}, n)$ is increasing in F (for $m^{*} > m^{**}$) and equal to zero at $\widetilde{F}$, where

$$\begin{aligned}\widetilde{F} &= \frac{1}{m^{*} - m^{**}} \left\{ m^{*} V_i^C(k_0; m^{*}) + (n - m^{*}) V_j^N(k_0; m^{*}) \right. \\ &\quad \left. - \left[m^{**} V_i^C(k_0; m^{**}|coop) + (n - m^{**}) V_j^N(k_0; m^{**}|coop) \right] \right\},\end{aligned}$$

which is positive if $m^{*} - m^{**}$ is large enough. If $F \in \mathcal{F} \cap \mathcal{F}^{*} \cap \mathcal{F}^{**}$, with $\mathcal{F} \equiv \{F : 0 \leq F \leq \widetilde{F}\}$, $\mathcal{F}^{*} \equiv \{F : \max\{0, \underline{F}^{(m^{*})}\} < F < \overline{F}^{(m^{*})}\}$, and $\mathcal{F}^{**} \equiv \{F : \max\{0, \underline{\underline{F}}^{(m^{**})}\} < F < \overline{\overline{F}}^{(m^{**})}\}$, then the equilibrium number of contributors is m^{*} in the non-cooperative scenario and m^{**} in the cooperative scenario, and $W^{**}(m^{**}, n) < W^{*}(m^{*}, n)$. For the parameter values in the numerical example in Section 5, we obtain $\widetilde{F} = 1.7110$. Therefore, any $F \in [0.0289, 0.0499]$ would lead to $W^{**}(2, 10) < W^{*}(10, 10)$.

References

[1] Amador, M., Bagwell, K., 2013. The theory of optimal delegation with an application to tariff caps. *Econometrica* 81, 1541-1599.

[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] Atatüre, S., 2008. The historical roots of European Union: integration, characteristics, and responsibilities for the 21st century. *European Journal of Social Sciences. Eurojournal* 7(2).

[5] Bagwell, K., Staiger, R.W., 2005. Enforcement, private political pressure and the general agreement on tariffs and trade. *Journal of Legal Studies* 3, 471-513.

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

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

[8] 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.

[9] Benchekroun, H., Long, N.V., 2012. Collaborative environmental management: A review of the literature. *International Game Theory Review* 14, 1-22.

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

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

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

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

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

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

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

[17] Cornes, R., Hartley, R., Sandler, T., 1999. Equilibrium existence and uniqueness in public good models: An elementary proof via contraction. *Journal of Public Economic Theory* 14, 499-509.

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

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

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

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

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

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

[24] Gaudet, G., Salant, S.W., 1991. Uniqueness of Cournot equilibrium: New results from old methods. *Review of Economic Studies* 58, 399-404.

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

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

[27] Ihori, T., 1996. International public goods and contribution productivity differentials. *Journal of Public Economics* 61, 139-154.

[28] 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.

[29] Kahana, N., Klunover, D., 2016. Private provision of a public good with time-allocation choice. *Social Choice and Welfare* 47, 379-386.

[30] Lemoine, D., Traeger, C., 2014. Watch your step: optimal policy in a tipping climate. *American Economic Journal: Economic Policy* 6, 137-166.

[31] Lemoine, D., Traeger, C., 2016. Ambiguous tipping points. *Journal of Economic Behavior and Organization* 132, 5-18.

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

[33] Long, N.V., Soubeyran, A., 2000. Existence and uniqueness of Cournot equilibrium: A contraction mapping approach. *Economics Letters* 67, 345-348.

[34] McGinty, M., 2007. International environmental agreements among asymmetric nations. *Oxford Economic Papers* 59, 45–62.

[35] 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.

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

[37] Pavlova, Y., de Zeeuw, A., 2013. Asymmetries in international environmental agreements. *Environmental and Development Economics* 18, 51-68.

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

[39] 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.

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

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

[42] Vives, X., 2008. Innovation and competitive pressure. *Journal of Industrial Economics* 56, 419-496.

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

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