Strategic decentralization and the provision of global public goods*

Renaud Foucart[†] Cheng Wan[‡]

November 13, 2015

This paper studies strategic decentralization and strategic (non) activation of a potential coalition in the provision of global public goods. We show that a coalition, with the aim of maximizing the aggregate utility of its members, may find it strategically advantageous to decentralize its provision, so that members act autonomously to maximize their own utility. In particular, strategic decentralization can lead to the phenomenon that a multipolar world with two active coalitions provides less public good than a world with only one of the coalitions while the other countries act independently. Furthermore, we show that it can be in the joint interest of a group of countries susceptible to form a coalition to commit themselves to acting independently, even if they share a strong interest in the public good. Finally, we discuss how the choice of coalition decentralization and that of activating a potential coalition vary with the players' weight and taste for the public good.

JEL: C72, H41

Keywords: Global public goods, Coalitions, Decentralization, Commitment, Emission game.

*We thank Paola Conconi, Grégoire Garsous, Victor Ginsburgh, Georg Kirchsteiger, Patrick Legros, Debraj Ray, André Sapir and Annalisa Vinella for useful comments and discussion.

[†]Humboldt University, Berlin, renaud.foucart@hu-berlin.de

[‡]Nuffield College and Department of Economics, Oxford University, cheng.wan@economics.ox.ac.uk

1 Introduction

Contribution to global public goods is one of the domains where international cooperation matters the most. Actors, which can be states, countries or regions, decide whether to act together for their common interest on issues such as global warming, production of scientific knowledge, international security or preservation of natural resources. These issues have in common that the establishment of a cooperative relationship leads to an improvement in the total provision of a global public good. Typical cooperation takes the form of a coalition, where members delegate decision-making to a centralized level, which maximizes their total surplus.

This paper investigates the causes and consequences of strategic decentralization of decision-making in a coalition. First, we discuss the potential strategic incentive for a coalition to decentralize the decision-making about the provision of a global public good. Decentralization is a commitment to free riding on other actors: the members of a coalition may all benefit from acting independently if this incites other actors to provide more public good. Second, we take one step back and show that a group of actors may collectively benefit from not creating the institutions that make cooperation possible, if its creation induces the decentralization of other coalitions.

Our approach is in deliberate contrast with the traditional view of coalition formation according to which individual actors seek to cooperate on a single issue. In particular, the literature on self-enforcing International Environmental Agreements (IEAs) has developed a very pessimistic view on the capacity of individual countries to form coalitions providing a global public good, suggesting that "an IEA is unable to improve much on the noncooperative outcomes" (Barrett, 2005, p.1480). It is because members often have incentives to quit the coalition and become free riders. On the contrary, we start by assuming that countries belong to pre-existing groups - bounded by a common history of cooperation, common interest, or mutual trust - and are able to choose the institutional design to maximize their combined surplus.

As noted by Kosfeld *et al.* (2009), political and economic institutions with a coercive power on its members do exist in practice. Once an institution and the mutual trust between its members exist, nothing prevents them to vote unanimously to extend the scope of the institution's intervention, including potential compensations. For instance, agriculture was not part of the issues on which the institutions of the European Union were built. But there was no process of coalition formation on the common agricultural policy. It was an unanimous choice to extend the scope of the UE, while the British government managed to receive a rebate on its contribution to the EU budget as a compensation.[1]

However, while coalitions exist and are relatively stable over time, it is a striking fact that they often act in a decentralized way on specific global issues. Two well-known examples are the EU and the US.

[1] In 1984, the UK negotiated a mechanism (the "Fontainebleau agreement") wherein it automatically gets back about two-thirds of the difference between what it contributes to and what it receives from the EU budget at the end of each year (Lowe *et al.*, 2002).

Individual member countries of the European Union have long acted in a decentralized way on most military interventions (Howorth, 2001; Kirchner, 2006). This includes conflicts on the European continent such as the Balkan wars in the 1990s, where most of the leadership was left to the US (Gordon, 1997, p.74). Since the early 2000s, some smaller-scale military interventions have been made on behalf of the European Union's Common Security and Defense Policy (CSDP) (Kaldor and Salmon, 2006), but when the 2011 crisis in Libya escalated, "no one apparently seriously considered intervention under the framework of the CSDP" and "the European Union stood on the sidelines and watched as France and the United Kingdom, acting within a NATO framework, intervened militarily on the Union's doorstep" (Menon, 2011, p.75).

The United States is a federal country whose centralized level of government takes decisions on several global issues, such as national defense. However, regarding global climate change, the policy of the US has always been "bottom up," in that the central government delegates to states the choice of taking constraining decisions on abatement targets (Lutsey and Sperling, 2008).[2] A consequence has been the refusal of the US administration to commit on abatements, making the European Union complain about the "lack of American Leadership."[3]

A typical explanation of such phenomena is that centralized decision fails to be effective, either because coalition members have too different preferences to reach an agreement, or because proper compensations cannot be efficiently designed on all issues. Our first point provides an alternative explanation. We claim that decentralization is strategic if it is in the joint interest of the coalition members. Indeed, it may allow the member countries to commit to free riding in the presence of other active coalitions or countries. Decentralization as a means to free ride is derived from the fact that smaller players have less incentive to contribute in a noncooperative game of public goods (see for instance Olson and Zeckhauser, 1966). Hence a coalition prefers to decentralize if the benefits to its members from their free riding on the actors outside the coalition are higher than their gain from cooperating within the coalition.

A consequence of such a strategic behavior is that a unipolar world with a single coalition in addition to noncooperative countries can sometimes provide more public good than the multipolar world in which these noncooperative countries also form a coalition. As a matter of fact, a large coalition with small enough rivals usually has no incentive for decentralization. On the contrary, the formation of a smaller coalition by these initially noncooperative rivals may lead to a strategic decentralization of the larger coalition.

Our second point is deduced from this reasoning: a group of actors susceptible to cooperation but without pre-existing institutions can strategically choose whether or not, according to their common interest, to activate their coalition by building the necessary institutions. It can turn out to be detrimental for a group of countries to activate their

[2] The US government seems, however, able to deal in an efficient and centralized way on environmental issues with a national impact only, such as acid rains, as shown by the Bush senior administration in the early 1990s (Joskow *et al.*, 1998).

[3] Andreas Carlgren, Sweden's Environment Minister talking on behalf of the EU presidency (Copenhagen talks, 2009, cited by The Guardian, November 2, 2009).

coalition if it results in the strategic decentralization of an existing coalition. If these countries happen to be more sensitive to the level of public good than the others, the outcome can be worse for them, because their pro-public-good feature makes them exploited even more aggressively by the free riders. It is thus beneficial for each country in this group to avoid creating an active coalition in the first place, so that they are *committed to not being committed* to the provision of public good.

This scenario can be applied to the recent efforts of the so-called BRICS (Brazil, Russia, India, China and South Africa) to establish a political institution, holding annual meetings since 2009. While the five countries share similar concerns about "the international dominance of the United States, the threat of terrorism from religious fundamentalists and ethnic movements, and the need to prioritize economic development," they were "reluctant to share any burdens" (Pant, 2013). Indeed, till 2014, the only substantial cooperative institution launched by the BRICS had been the project of a joint development bank, making Rodrik (2013) write that "[what] the world needs from the BRICS is not another development bank, but greater leadership on today's great global issues. The BRICS countries are home to around half of the world's population and the bulk of unexploited economic potential. If the international community fails to confront its most serious challenges - from the need for a sound global economic architecture to addressing climate change - they are the ones that will pay the highest price."

Observers often take the behavior of these countries as a sign of lack of interest in the global public goods, while they actually care much more than they appear to do. We claim that their reluctance to further organize as an active coalition is not due to their seemingly carelessness but is indeed a strategic behavior supported by our second point. If building institutions that increase trust and cooperation among a certain group of countries leads the rest of the world to expect their further cooperation on other issues, these countries may do better by remaining noncooperative.

The difference between decentralization and non-activation is crucial: an active coalition with existing institutions may end up choosing decentralization, but only if it is profitable for all the members given the expectation on the behavior of other coalitions. In contrast, a group of countries may together benefit from deliberately avoiding creating institutions in the first place, even if its members would unanimously choose to centralize if the institutions existed.

This paper adopts a game theoretic approach. We study a game of pollution emission because of the familiarity of the reader with the subject as well as the existing extensive literature in this field. However, our results hold in a more general setting of public goods. Consider a game taking place over three stages. In the first stage, each inactive (potential) coalition decides whether or not to activate the coalition. Coalitions already activated do nothing at this stage. In the second stage, active coalitions decide to centralize or decentralize the choice of abatement levels. In the third stage, active centralized coalitions and individual countries (including the member countries of the inactive coalitions and those of the active but decentralized coalitions) play a noncooperative global public good game. In our setting, this underlying global public game is just the emission game discussed in Carraro and Siniscalco (1993). In such a game, the public good is the aggregate abatement, i.e. the aggregate pollution not emitted. Each player chooses her emission quantity in the limit of her capacity, characterized by her weight. A player benefits from her abuse of the environment as a reservoir for her own emissions, but also suffers from the aggregate pollution.

We look for the subgame perfect Nash equilibria of this three-stage game. In Section 2, we study the emission game in the last stage and characterize its unique equilibrium. In Section 3, we discuss strategic decentralization in the second stage. After deriving conditions for decentralization to be disadvantageous in the general case, we then focus on two-coalition decentralization games and give a full description of their equilibria. In Section 4, we turn to the first stage of the game by defining the game form in the general setting, then clarify under which conditions it is in the common interest of the members of a potential coalition to become active facing another active coalition. Section 5 provides four numerical examples. The first example shows how the choice between centralizing and decentralizing made by a coalition in the second stage depends on the existence of other active coalitions. The second example displays a situation where two coalitions simultaneously choose to be active and centralized. The third example exhibits the possibility for a world with one large active coalition and small individual countries to provide more public good than a world in which these independent countries have formed a second active coalition, even if its size is smaller than the first one. The last example shows that a group of individuals sharing a strong preference for the public good may be better off by not forming a coalition, when they confront some coalition with milder interest in the public good. In Section 6, we discuss how the variation in the characteristics of the players, such as their size and preference for the public good, change their decision in the three stages of the game. Section 7 concludes. All proofs and some auxiliary results are collected in the Appendix 2.

Related literature

To our knowledge, this paper constitutes the first attempt to study the strategic decentralization of coalitions in the context of global public good games.

The literature on fiscal federalism has generally taken as given that public goods (such as national defense) were precisely the cause of the formation of a federation (Oates, 2005, p.366), but not a cause of secession or decentralization. However, economists have long been interested in the formation and the maintenance of cartels producing public goods. The generally studied procedure of coalition formation is a two-stage game. In the first stage, countries decide whether or not to join a coalition. In the second stage, each coalition acts as a single player who maximizes the aggregate welfare of its members. This setting presumes that, conditional on being members of a coalition, countries are able to sign binding contracts and punish deviators. The three most widely considered procedures for the first stage are respectively proposed by Bloch (1996), Ray and Vohra (1997) and d'Aspremont *et al.* (1983).[4]

[4] Yi (1997) presents number of properties of those different rules for the creation of coalitions with either positive or negative externalities. Belleflamme (2000) allows for asymmetric countries in an open membership game with negative externalities and McGinty (2007) for one with positive externalities.

This approach has been at the basis of most of the economic literature on self-enforcing IEAs, the most studied case of coalition formation concerning global public goods (see for instance Barrett, 1994 and Barrett and Mattei, 1993). Since individual countries have a strong incentive to free ride, these models generally predict that IEAs can achieve little more than a noncooperative framework (Barrett, 2005).[5] Nevertheless, in a global public good game, the coalition maximizing the aggregate welfare of all the players (i.e. the grand coalition) should be feasible with transfers and/or credible punishment of deviators (see Carraro and Siniscalco, 1993 and Carraro *et al.*, 2006 for a discussion of this argument).[6] There must however be a history of sequential commitments, in the sense that a first group of countries commit to act together, and then jointly choose to expand the coalition by providing a transfer to an additional member. Ray and Vohra (2001) show that the existence of transfers is not sufficient to ensure the formation of the grand coalition when countries are free to sequentially offer a new partition, even after entering a coalition.

The novelty of our approach is that, instead of focusing on the motive of individual countries to join or quit coalitions, we consider pre-existing coalitions capable of implementing any kind of transfers and punishment schemes. We define strategic decentralization as the collective decision of a group of countries to act separately on a certain issue, in order to maximize their total surplus. This specification is similar to the idea of players "delegating" themselves in congestion games (Sorin and Wan, 2013).

This result of profitable decentralization is also reminiscent of at least two important results in Industrial Organization. First, Salant *et al.* (1983) show that in a Cournot environment, a horizontal merger may lower the profits of firms. Second, in a similar environment, Baye *et al.* (1996) show that large firms may benefit from divisionalisation and franchising. If a firm can ex-ante commit itself to delegating the production choice to several smaller franchises, this is equivalent to a commitment to a higher level of production. However, these papers consider comparative statics without studying the strategic impact of one firm's divisionalisation on the decision of other firms to centralize or not.

Our focus is further close to two recent papers on the link between country size, centralization and the provision of global public goods. First, Eckert (2003) shows that a federal government which delegates the negotiation and enforcement power to the local government of a pollution-producing region may benefit from a better position in the negotiation over a climate agreement with another country, because its disagreement point is more favourable. Second, Buchholz *et al.* (2014) shows that if a group of countries can choose a matching ratio for their contributions to a public good, the existence of a coalition may decrease the aggregate level of public good. However, although the first paper discusses a country's strategic choice of a certain form of constitution, it considers neither decen-

[5] More recent work (Barrett, 2013; Foucart and Garsous, 2013) shows that this result holds because of the hypothesis of the convexity of pollution costs, hence the concavity of the benefits from the abatements. A threshold leading to a catastrophe, as it is the case with climate change, could lead to a large abating coalition in equilibrium if identified with sufficiently high precision.

[6] Most of the environmental literature focuses on a single coalition. In an open membership game, multiple coalitions may coexist in equilibrium, but must be stand-alone stable, in the sense that no individual should be better off by deviating (Yi, 1997).

tralization nor coalition formation, where as the second paper does not consider strategic interaction among (potential) coalitions beyond the last stage of contribution to the global public good.

2 Last stage: Emission game

To solve the three-stage game by backward induction, we begin the analysis by the last stage. The players in this stage play an emission game where the public good is the aggregate abatement of all players, i.e. the aggregate pollution not emitted. This section gives the definition and some important properties of an emission game.[7]

There are N players, indexed by $i \in \mathcal{N} = \{1, ..., N\}$. A player i of a strictly positive weight m_i is either an autonomous country or a centralized coalition composed of a finite number of member countries. Centralized coalitions choose their aggregate quantity of emissions, in order to maximize their aggregate welfare. The weight of a country characterizes its capacity of pollution. For instance, the level of emissions from a country of weight 1 is between 0 and 1. The weight of a coalition is the total weight of its member countries. Let the total weight of all the players be denoted by $M = \sum_{i \in \mathcal{N}} m_i$.

Every player i chooses a level of emissions $q_i \in [0, m_i]$. The profile of choices is a vector $\mathbf{q} = (q_1, \ldots, q_N)$. The aggregate level of emissions is denoted by $Q = \sum_{i \in \mathcal{N}} q_i \in [0, M]$.

The total utility function of player i when the players' emission profile is $\mathbf{q}$ is

$$U_i(\mathbf{q}) = G_i(q_i) - s_i F_i(Q), \tag{1}$$

where $G_i(q_i)$ is the player's benefit from her own emissions (by using the environment for her production or consumption), $s_i > 0$ is the measure of the player's sensitivity to global pollution, and $F_i(Q)$ stands for the loss caused by the aggregate level of emissions. The fact that the loss depends on the aggregate level of pollution implies that each player's emissions generate a negative externality on all the others. Besides, the sensitivity of a coalition to global pollution is the common sensitivity of its member countries.

In this paper, we consider the following specific forms of functions G_i and F_i:

$$G_i(q_i) = m_i \cdot g\big(\frac{q_i}{m_i}\big), \tag{2}$$

$$F_i(Q) = m_i \cdot f(Q). \tag{3}$$

The interpretation is as follows. When comparing different players' welfare, it is necessary to discard the influence of their different weights, so that all other things held constant the utility of a group of players is equal to the sum of the utilities of the group members. The gain from one's own pollution and the loss from the aggregate pollution are thus measured for per-unit of weight, and these functions should be symmetric across the countries. Formula (2) signifies that player i's gain by each unit of her weight from her own emissions

[7]We show in Appendix 1 that our emission game is identical to a public good provision game.

into the environment, $g(\frac{q_i}{m_i})$, depends only on her per-unit pollution level. Owing to this assumption, two countries, both having weight 1 and emitting a quantity $\frac{1}{2}$, gain in sum $2g(\frac{1}{2})$. This is the same as a country of weight 2 gains from emitting a quantity 1. Similarly, formula (3) means that each unit of weight of each player bears the same loss $f(Q)$ from the aggregate emissions.

The *(per-unit) utility function* of player i is thus:

$$u_i(\mathbf{q}) = \frac{U_i(\mathbf{q})}{m_i} = g\big(\frac{q_i}{m_i}\big) - s_i f(Q). \tag{4}$$

Assumption 1. *f is defined from $\mathbb{R}_+$ to $\mathbb{R}_+$. It is twice differentiable, strictly increasing and strictly convex.*

g is defined from $[0,1]$ to $\mathbb{R}_+$. It is twice differentiable, strictly increasing and strictly concave. Besides, $\lim_{x\to 0+} g(x) = +\infty$, $g'(1) > 0$.

The assumptions on f signify marginally increasing damage resulting from the aggregate pollution, whereas those on g signify marginally decreasing returns on a player's use of the environment as a reservoir for her own emissions.

It follows immediately from Assumption 1 that, given any choice of the others, player i's utility u_i is always strictly convex in her own choice q_i. Indeed, $\frac{\partial^2 u_i(\mathbf{q})}{\partial q_i^2} = \frac{1}{m_i^2} g''(\frac{q_i}{m_i}) - s_i f''(Q) < 0$.

Denote this emission game by $\Gamma(\mathcal{N})$. A strategy profile $\mathbf{q}$ is a Nash equilibrium if for all player $i \in \mathcal{N}$,

$$u_i(q_i, \mathbf{q}_{-i}) = \max_{0 \le x \le m_i} u_i(x, \mathbf{q}_{-i}), \tag{5}$$

where $\mathbf{q}_{-i} = (q_1, \ldots, q_{i-1}, q_{i+1}, \ldots, q_N)$.

Lemma 1. *An emission game admits a unique Nash equilibrium.*

As the following lemma shows, a player's contribution and utility at the equilibrium depend on her weight and her pollution sensitivity.

Lemma 2. *At the equilibrium $\mathbf{q}$ of the emission game, consider two players i and j.*

(i) If they have the same weight and the same sensitivity, then they produce the same amount of public good and have the same per-unit utility.

(ii) If they have the same sensitivity but different weights, then either the bigger one produces strictly more public good than the smaller one w.r.t. their respective weight, while receiving a strictly lower per-unit utility, or neither of them produces any public good so that both receive the same per-unit utility.

(iii) If they have the same weight but different sensitivity, then either the more sensitive one produces strictly more public good than the less sensitive one w.r.t. their respective weight, or neither of them produces any public good. The more sensitive player has strictly lower utility.

Here "player i producing more public good than player j w.r.t. their respective weight" means $\frac{q_i}{m_i} \leq \frac{q_j}{m_j}$.

The following lemma further shows that when a player is "small" enough, it always produces as much pollution as its capacity permits, providing no public good at all.

Lemma 3. *Given the total weight of the players and the range of their sensitivities, there is a threshold $\epsilon > 0$ such that any player who has a weight inferior to ϵ never contributes to the production of public good.*

Lemma 2(ii) and Lemma 3 recover the well-known result (Olson and Zeckhauser, 1966) that being small is an advantage in the contribution to a global public good. Roughly speaking, smaller players free ride more by producing more pollution per unit of weight. Intuitively, the marginal return from a unit of investment in the public good is increasing in the weight of a player. Hence, being small is a commitment to investing less in the public good. A small player can thus more easily free ride on the others, and her equilibrium utility is higher. However, the pollution sensitivity should also be taken into account as Lemma 2(iii) reveals. A small country that is highly sensitive to pollution may abstain from polluting too much.

3 Second stage: Strategic decentralization

This section focuses on the second stage of the game where strategic choices of *decentralization* are taken. We first give the definition of a decentralization game, then provide some general results on the externalities that unilateral decentralization generates on the other players as well as on the aggregate provision of the public good, and finally investigate in detail a specific class of strategic decentralization games.

3.1 A decentralization game

First define a strategic decentralization game where active coalitions can *simultaneously choose to decentralize.*

There are a finite number of players. The player set $\mathcal{N}$ is made up of two disjoint subsets $\mathcal{N}_1$ and $\mathcal{N}_2$. A player $k \in \mathcal{N}_1$ is an autonomous country. A player $k \in \mathcal{N}_2$ is a coalition composed of a finite number n_k of countries. Denote the finite set of member countries of a coalition k by $\mathcal{J}_k$. For the sake of simplicity and without loss of generality, we assume that the member countries of a coalition k have the same weight $\frac{m_k}{n_k}$. The coalition *decentralizes* if each of its member countries is free to choose its quantity of emissions.

A decentralization game, denoted by $D(\mathcal{N})$, is defined as follows. Autonomous countries have no choice to make. Each coalition has two choices at this stage: *centralization* (C), i.e. to act as one player maximizing the aggregate utility of its member countries in the third stage, and *decentralization* (D), i.e. to let the member countries act independently in the third stage.

Once all the coalitions have made their choice, an emission game is induced by the pure strategy profile $\mathbf{s} = (s_k)_{k \in \mathcal{N}_2}$. It is played in the third stage by (i) the initially autonomous countries, (ii) the member countries of the coalitions that have chosen decentralization, and (iii) the coalitions that have chosen centralization. Denote the set of these players by $\mathcal{N}^{\mathbf{s}}$. Section 2 shows that a unique equilibrium exists in emission game $\Gamma(\mathcal{N}^{\mathbf{s}})$. Let us define the players' utilities in $\mathcal{N}$ associated to a pure strategy profile $\mathbf{s}$ in the following manner:

The utility of an initially autonomous country ($l \in \mathcal{N}_1$) or a coalition having chosen centralization ($l \in \mathcal{N}_2$ such that $s_l = C$) is its equilibrium utility in the induced emission game $\Gamma(\mathcal{N}^{\mathbf{s}})$:

$$v_l(\mathbf{s}) = u_l(\Gamma(\mathcal{N}^{\mathbf{s}})).$$

The utility of a coalition having chosen decentralization ($k \in \mathcal{N}_2$ such that $s_k = D$) is the sum of the equilibrium utilities of its member countries in $\Gamma(\mathcal{N}^{\mathbf{s}})$.

$$v_k(\mathbf{s}) = \frac{1}{n_k} \sum_{k_j \in \mathcal{J}_k} u_{k_j}(\Gamma(\mathcal{N}^{\mathbf{s}})).$$

In particular, since the member countries of k have the same weight and the same sensitivity to pollution, they have the same equilibrium utility in the induced emission game. Hence $v_k(\mathbf{s})$ is the common equilibrium per-unit utility of all its member countries in $\Gamma(\mathcal{N}^{\mathbf{s}})$.

To simplify the notation, we denote a strategy profile $\mathbf{s}$ by the list of coalitions that choose decentralization: $\{k : k \in \mathcal{N}_2, s_k = D\}$. For example, if only one coalition k decentralizes, the new set of players is simply denoted by $\mathcal{N}^{(k)} = \mathcal{N} \cup \mathcal{J}_k \setminus \{k\}$.

Two remarks are necessary here.

First, member countries make the choice between centralization and decentralization of a coalition in a collective manner. In the case of centralization, transfers between countries can be made so as to ensure that all the member countries achieve the same per-unit utility. In the case of decentralization, they still receive the same per-unit utility in the induced emission game because of their equal weight. Hence a collective decision of (de)centralization can effectively be obtained without controversy. Had countries different weights, the same results could be obtained using transfers at the decentralization stage. If the aggregate utility of a coalition is higher using decentralization, there exist transfers to be implemented in the second stage so that decentralization is Pareto-improving within the coalition. Indeed, transfers made in the decentralization stage have no impact on the incentives of independent countries in the third (emission) stage.

Second, different from the literature on endogenous coalition formation, we discard the possibility that certain member countries of a coalition form one or several smaller coalitions, for two reasons. First, when the coalition decentralizes the decision-making, it is natural to assume that this power is returned to the original components of the coalition instead of some newly formed sub-coalition(s). Second, even if we allow new sub-coalitions, they can either be created according to a joint decision of the decentralizing coalition's members or come into existence through an endogenous coalition formation process. On the one hand, if it is the member countries of the coalition who have jointly and deliberately chosen to divide their coalition into certain sub-coalitions, this choice needs to be justified by their anticipation of what will happen to these sub-coalitions. For example, they must anticipate whether the sub-coalitions will continue to decentralize so that they decompose themselves to even smaller sub-sub-coalitions. Also, they need to anticipate the behavior of the newly formed sub-coalitions of the rival coalitions, and so on. Not only is this approach of reasoning incoherent with our initial assumption that a coalition decentralizes if it is in the common interest of each of its member countries, but also it makes the analysis so complicated that we can draw no conclusion from the outcome of such recursive behaviors. On the other hand, though individual countries may form an endogenous sub-coalition spontaneously, it is irrelevant to our analysis, which focuses on the motive of an existing coalition's decentralization instead of coalition's formation.

3.2 Externality of decentralization

This subsection examines the consequences of the unilateral decentralization of one coalition. In the remaining of this section, by "initially" we mean at the outcome of the emission game where all the activated coalitions are centralized.

Lemma 4. *If a coalition initially contributes to the public good, then its unilateral decentralization has the following effects:*

- *(i) The aggregate contribution to the public good is strictly decreased.*
- *(ii) Each of the other players contributes more to the public good while receiving a strictly lower utility.*
- *(iii) The total contribution to the public good from the members of the decentralized coalition is strictly decreased.*

If a coalition initially contributes no public good, then its unilateral decentralization does not change the contribution and the utility of any other player; the total contribution from its own members does not change either.

Lemma 2 asserts that larger groups contribute more to the provision of public good. As a matter of fact, their unilateral decentralization also exerts greater negative externalities under some mild conditions.

Lemma 5. *Consider two coalitions with the same sensitivity to pollution but different weights. Suppose that the larger one initially contributes to the public good. If the respective unilateral decentralization of either of them leads to the full-scale pollution of its* *member countries, then the decentralization of the larger one exerts a strictly greater negative externality on the other countries' utilities as well as on the aggregate contribution of the public good.*

Lemma 3 ensures that the member countries of a coalition provide no public good after its decentralization as long as their weights are all inferior to ϵ. Under this assumption, the decentralization of the larger coalition always exerts a strictly greater negative externality.

3.3 Nash equilibria of the decentralization game

We now return to the decentralization game $D(\mathcal{N})$ where coalitions simultaneously choose whether or not to decentralize decision-making.

3.3.1 General case

The decentralization game $D(\mathcal{N})$ is finite (i.e. with a finite number of players each possessing a finite number of choices). Hence the game admits mixed-strategy equilibria.

Lemma 6. *In a decentralization game:*

(i) If a centralized coalition i initially produces no public good, then it has no strictly positive gain from decentralization, whatever the choices of the other coalitions.

(ii) If a centralized coalition i initially contributes to the provision of public good, but none of the other players makes any contribution both before and after the unilateral decentralization of coalition i, then coalition i strictly prefers not to decentralize, whatever the choices of the other coalitions.

The condition that none of the other players provides any public good after the unilateral decentralization of i in (ii) cannot be dropped. Indeed, a player who does not provide public good in one situation can well do so when other players' behavior change. Example 3 in Section 5 presents a case where a coalition does not contribute to the public good in the presence of another centralized coalition, but contributes when the latter decentralizes.

The intuition behind Lemma 6 is that, for decentralization to be profitable for a coalition, it should be neither too big nor too small compared with the other players. On the one hand, if it is too big, its decentralization does not induce a large increase in the provision of public good made by its opponents. Hence, the harm done by the free riding of its own members on each other outweighs the benefit from their free riding on the others. On the other hand, if it is too small, then it completely free rides even as a centralized coalition; hence it is indifferent between centralization and decentralization.

The following is an immediate corollary of Lemma 6(ii).

Corollary 1. *In a decentralization game, suppose that coalition i initially contributes to the provision of public good. Also suppose that all the autonomous countries and the member countries of all the other coalitions have a weight inferior to ϵ. Then, there is no equilibrium at which coalition i chooses decentralization while all the other coalitions also choose decentralization.*

3.3.2 Two-coalition case or a bipolar world

A characterization of all the Nash equilibria of a decentralization game is hardly tractable for an arbitrary number of players. Therefore, we focus the rest of our analysis on a particular class of decentralization games, in which there are only two coalitions and no initially autonomous countries, and all the member countries of the two coalitions have weights lower than ϵ. A thorough analysis of the Nash equilibria of games in this class is possible.

The matrices in Table 1 represent this 2×2 decentralization game. The row player is coalition 1 while the column player is coalition 2. Recall that choice C stands for "centralization" and D for "decentralization". The quantities of emissions from the two coalitions at the equilibrium of the induced emission game are listed in the matrix on the left hand side of Table 1 while their corresponding per-unit utilities are listed in the matrix on the right hand side.

Whenever a coalition i decentralizes, the aggregate emission quantity from its member countries is m_i. Indeed, since all the member countries of both coalitions have weight lower than ϵ, according to Lemma 3, they emit as much pollution as their capacities permit when acting autonomously.

We only consider the nontrivial case where at least one coalition, say coalition 1, initially provides some public good ($q_1 < m_1$). Besides, we also concentrate on the nontrivial case where coalition 2 is not too small with respect to coalition 1 so that it produces some public good when coalition 1 decentralizes unilaterally ($q_2^{(1)} < m_2$). Indeed, if it is not the case, then according to Lemma 4, it also provides no public good when coalition 1 is centralized. In other words, coalition 2 is so small that it never contributes any public good, in which case Lemma 6(ii) implies that coalition 1 strictly prefers centralization while coalition 2 is indifferent between centralization and decentralization. Finally, we consider only (locally) stable equilibria.[8]

Table 1: A 2×2 delegation game.

(a) Emissions

	C	D
C	q_1, q_2	$q_1^{(2)}, m_2$
D	$m_1, q_2^{(1)}$	m_1, m_2

(b) Per-unit utilities

	C	D
C	$u_1,\ u_2$	$v_1^{(2)},\ v_2^{(2)}$
D	$v_1^{(1)},\ v_2^{(1)}$	$v_1^{(1,2)},\ v_2^{(1,2)}$

Proposition 1. *The set of stable Nash equilibria of the decentralization game, denoted by NE, depends only on whether the coalitions gain from their respective unilateral decentralization. More precisely,*

(i) if unilateral decentralization is weakly deleterious for both coalitions ($u_1 \geq v_1^{(1)}$, $u_2 \geq v_2^{(2)}$), then at the unique equilibrium both coalitions centralize ($NE = \{(C, C)\}$);

[8] By "locally stable" we mean that a sequence of alternative best replies triggered by any small perturbation to a player's equilibrium strategy brings the game back to the equilibrium.

(ii) if unilateral decentralization is weakly deleterious for coalition 1 but strictly profitable for coalition 2 ($u_1 \geq v_1^{(1)}$, $u_2 < v_2^{(2)}$), then at the unique equilibrium coalition 1 centralizes while coalition 2 decentralizes ($NE = \{(C, D)\}$);

(iii) if unilateral decentralization is strictly profitable for coalition 1 but weakly deleterious for coalition 2 ($u_1 < v_1^{(1)}$, $u_2 \geq v_2^{(2)}$), then at the unique equilibrium coalition 1 decentralizes while coalition 2 centralizes ($NE = \{(D, C)\}$);

(iv) if unilateral decentralization is strictly profitable for both coalitions (i.e. $u_1 < v_1^{(1)}$ and $u_2 < v_2^{(2)}$), then there are two equilibria: one coalition centralizes while the other decentralizes ($NE = \{(C, D), (D, C)\}$.

As from Lemma 6 and Corollary 1, at least one coalition remains centralized at a pure equilibrium. While it is fairly plain that a higher sensitivity to pollution results in a stronger incentive to remain centralized, the role of the weight of a country is much less straightforward. We deduce from Lemma 6 that very small and very large coalitions are less likely to decentralize. We also know that a large centralized coalition is always worse off than a smaller one in the emission game (Lemma 2). This does not mean however that a large coalition always has more incentive to decentralize than a smaller one, because the negative consequence of her decentralization on the aggregate provision of public good is also more severe (Lemma 5).

4 First stage: Activation of the coalition

This section turns to the first stage of the game, where the member countries of a potential coalition make a collective decision on whether or not to activate the coalition.

4.1 Setting of the game

There are a finite number of players who are divided into G disjoint *groups*. Let the set of G groups be denoted by $\mathcal{G} = \mathcal{G}_1 \cup \mathcal{G}_2 \cup \mathcal{G}_3$. Each group in $\mathcal{G}_1$ is an *autonomous country*. Each group in $\mathcal{G}_2$ is an *active coalition* containing a finite number of member countries. Each group in $\mathcal{G}_3$ is a finite number of *autonomous countries that are susceptible to form a coalition.*

In the first stage, only the groups in $\mathcal{G}_3$, i.e. the *potential* coalitions, have a choice. The autonomous countries in the group make the choice jointly. They choose whether to activate their coalition (A) or not to activate it (NA). Denote the profile of choices by $\mathbf{a} = (a_g)_{g \in \mathcal{G}_3}$, where $a_g \in \{A, NA\}$ for all $g \in \mathcal{G}_3$. It induces a decentralization game $D(\mathcal{N}^{\mathbf{a}})$ which is played at the second stage. The players in the induced decentralization game, $\mathcal{N}^{\mathbf{a}}$, are composed of autonomous countries and coalitions. Among the former, there are initial autonomous countries in $\mathcal{G}_1$ and the member countries of all the potential coalitions in $\mathcal{G}_3$ having chosen not to activate the coalition. Among the latter, there are initial coalitions in $\mathcal{G}_2$ and the potential coalitions in $\mathcal{G}_3$ having chosen to activate the coalition. If a potential coalition's members choose not to activate coalition, then in the induced decentralization subgame $D(\mathcal{N}^{\mathbf{a}})$, each of them plays as an autonomous country and receives its equilibrium utility in $D(\mathcal{N}^{\mathbf{a}})$; on the other hand, if they choose to activate the coalition, then the whole group acts as an activated coalition in the induced decentralization game $D(\mathcal{N}^{\mathbf{a}})$, and it receives the per-unit utility at the equilibrium of $D(\mathcal{N}^{\mathbf{a}})$.

A problem arises when it comes to decide the players' utilities associated to a choice profile. As Proposition 1 shows, there may be multiple equilibria in the induced decentralization subgame $D(\mathcal{N}^{\mathbf{a}})$. While it is possible to define a selection or rationalization criterion of the equilibria, or to keep on working with all the equilibria, we choose the simplest case immune to controversy. Firstly, we assume that there is only one potential coalition g in $\mathcal{G}_3$, the unique group having a choice in the first stage. Thus, only g's anticipation of what will happen thereafter counts. Secondly, when member countries of g compare their equilibrium utilities in the two decentralization subgames induced by the two choices (A and NA), they may be comparing two sets which are not necessarily singletons because of the multiple equilibria in theses subgames. To define a criterion for the comparison between two sets of values is not evident because it depends on the anticipation of g's members on what will happen in each of the subgames as well as on other elements in the game. A universal rule is not at hand. This difficulty can be bypassed if we further suppose that there is only one other player in the game, who is an active coalition, as shown in next subsection.

4.2 A potential coalition versus a coalition

Let us look at the particular case where there are an active coalition (called coalition 1) and a group of autonomous countries susceptible to form a coalition (called group 2). All member countries' weights are assumed to be lower than ϵ. In this setting, the three-stage game can be well defined and a full analysis of its equilibria is possible.

In the first stage, members of group 2 have a choice to make: A or NA.

Not activating the coalition (NA) induces a decentralization game where they act as autonomous countries, so that they have no choice in the induced decentralization game in the second stage. The matrix on the left hand side of Table 2 presents this decentralization subgame. There, the common per-unit utility to the members of group 2 depends on coalition 1's choice of (de)centralization. According to the proof of Proposition 1, $v_1^{(2)} > v_1^{(1,2)}$. Hence the unique and pure equilibrium of the subgame is C, so that the per-unit utility to the member countries of group 2 is $v_2^{(2)}$.

Activating the coalition (A) induces a decentralization game in the second stage where group 2 acts as an active coalition. This is the setting in Section 3.3.2. The decentralization subgame is presented in the matrix on the right hand side of Table 2. Its equilibria are fully characterized in Proposition 1. There are three candidates for stable Nash equilibrium in this game, (C, C), (D, C) and (C, D). Let the set of coalition 2's per-unit utility at all the stable equilibria in this decentralization game be denoted by $\mathcal{E}_2$. Then $\mathcal{E}_2$ has at least one and at most two elements.

The member countries of group 2 compare their per-unit utilities in the two decentralization games. More precisely, they compare $v_2^{(2)}$ with the elements in $\mathcal{E}_2$. We show that the comparison is feasible because there are only three possibilities: (i) $v_2^{(2)} \geq e$ for all $e \in \mathcal{E}_2$, and there is $\hat{e} \in \mathcal{E}_2$ such that $v_2^{(2)} > \hat{e}$, (ii) $v_2^{(2)} \leq e$ for all $e \in \mathcal{E}_2$, and there is $\hat{e} \in \mathcal{E}_2$ such that $v_2^{(2)} < \hat{e}$, and (iii) $v_2^{(2)} = e$ for all $e \in \mathcal{E}_2$.

Table 2: First stage.

(a) Coalition 2 is not active

C	$v_1^{(2)}*,\ v_2^{(2)}*$
D	$v_1^{(1,2)},\ v_2^{(1,2)}$

(b) Coalition 2 is active

	C	D
C	$u_1,\ u_2$	$v_1^{(2)},\ v_2^{(2)}$
D	$v_1^{(1)},\ v_2^{(1)}$	$v_1^{(1,2)},\ v_2^{(1,2)}$

Proposition 2. *In the nontrivial case discussed in Section 3.3.2, group 2's choice in the first stage of the game is as follows.*

(i) If $u_1 \geq v_1^{(1)}$ and $u_2 \geq v_2^{(2)}$ (the two equalities do not hold simultaneously), group 2 chooses A.

(ii) If $u_1 \geq v_1^{(1)}$ and $u_2 < v_2^{(2)}$, group 2 is indifferent between A and NA.

(iii) If $u_1 < v_1^{(1)}$, group 2

(a) chooses A if $v_2^{(1)} > v_2^{(2)}$;

(b) chooses NA if $v_2^{(1)} < v_2^{(2)}$;

(c) is indifferent between A and NA if $v_2^{(1)} = v_2^{(2)}$.

Recall that $u_1 \geq v_1^{(1)}$ means coalition 1 does (weakly) better by centralizing than by decentralizing, facing a centralized coalition 2; $u_2 < v_2^{(2)}$ means that group 2 does strictly better in a decentralized way than in a centralized way, facing a centralized coalition 1; and $v_2^{(1)} > v_2^{(2)}$ means that the per-unit utility of coalition 2 as a decentralized group facing a centralized coalition 1 is strictly higher than its per-unit utility as a centralized coalition facing a decentralized coalition 1.

Considering Proposition 2 together with the previous results allows us to characterize the motives of group 2 for choosing to activate its coalition in the first stage. A first possibility is that group 2 cares about the public good and expects the active coalition 1 to also care enough to stay centralized after the activation of coalition 2. This case corresponds to Proposition 2(i) and is illustrated in Example 2 in the next section. It is the most optimistic case since it leads to two centralized coalitions in the third stage. A second possibility is that group 2 does not care that much about the public good, so that it can expect to decentralize decision-making in the second stage and free ride in the third stage. In that case group 2 is indifferent, as the activation of the coalition has no impact on the decision-making. This case corresponds to Proposition 2(ii). The third possibility is that, even if coalition 1 chooses to decentralize as a consequence of the activation of coalition 2, the members of group 2 are still better off by jointly contributing than by jointly free riding in the third stage. This case corresponds to Proposition 2 (iii.a).

The reason for the members of group 2 to choose not to activate their coalition in the first stage is to commit to decentralization in the second stage. This corresponds to Proposition 2(iii.b). There are two scenarios where this can happen. In the first one, the presence of two coalitions leads to multiple equilibria in the decentralization subgame, and only the equilibrium where coalition 1 centralizes and coalition 2 decentralizes is advantageous for group 2. It is the case in Example 1 in the next section, where group 2 chooses not to activate the coalition in the first stage so as to ensure that at the unique equilibrium in the second stage, coalition 1 centralizes. This intuition is reminiscent of the first mover advantage in a Stackelberg game. In the second scenario, centralization is a weakly or strictly dominant strategy for an active coalition 2 in the second stage, and at the equilibrium coalition 1 decentralizes; however, it is a disadvantageous outcome for group 2. Therefore, it chooses not to activate the coalition in the first stage in order to "force" coalition 1 to centralize in the second stage. This case is illustrated in Example 3 in the next session. Such a scenario can even take place if group 2 is smaller. It gets further exacerbated if the preference for the public good is stronger for the members of this smaller group. This last case is illustrated in Example 4 in the next session.

5 Examples

This section provides four numerical examples to illustrate the results of the previous sections. In each example, there are two groups of countries. The first group is an active coalition of weight m_1, called coalition 1. The second group is a potential coalition of total weight $m_2 \leq m_1$, called group 2. When the member countries of group 2 activate their coalition, the group is also called coalition 2. In the first stage, the member countries of group 2 choose whether to activate their coalition. In the second stage, the active coalitions decide whether to centralize or decentralize decision-making of emissions. The third and last stage is the emission game induced by the decisions made in the second stage. We focus on subgame perfect Nash equilibria (SPNE). The four examples respectively show that:

(i) whether an active coalition has an incentive for strategic decentralization depends on the nature of its opponent(s);

(ii) two coalitions may both choose centralizing in equilibrium;

(iii) a unipolar world may provide more public good than a multipolar world; and

(iv) a group of countries, whose dominant strategy when acting as a coalition is to centralize decision-making, may find it better not to activate their coalition in the first place.

In each example, we first study the decentralization subgame beginning at the second stage. There are thus two possible scenarios: either group 2 has chosen not to activate the coalition (NA) or to activate it (A) in the first stage. In the first scenario, the world is composed of one coalition and a group of small autonomous countries, and is called a *unipolar world*. In the second one, the world is composed of two coalitions, and is called a *multipolar world*. After obtaining the SPNE in these two decentralization subgames, we return to the first stage to find the SPNE for the full game. The notations are the same as in the previous sections. We provide the detailed computation in Appendix 3.

5.1 Example 1: Strategic decentralization according to the nature of opponents

Assume $m_1 = m_2 = 1$, and the per-unit utility function of players in the induced emission game is

$$u_i(q) = 3q_i^{\frac{1}{2}} - Q^{\frac{3}{2}}. \tag{6}$$

Second stage, scenario 1 (unipolar world): Group 2 has chosen not to activate the coalition. Thus, in the second stage, only coalition 1 has a choice between centralization (C) and decentralization (D). We first solve the two emission games in the last stage induced respectively by the two choices of coalition 1. Since individual countries are sufficiently small, $q_2 = m_2 = 1$ in both emission games. If coalition 1 chooses C, the value of q_1 at the equilibrium of the induced emission game is $q_1 = \frac{\sqrt{5}-1}{2} = 0.62$, and its per-unit utility is $u_1 = 0.30$. If coalition 1 chooses D, then none of its member countries contribute any public good in the induced emission game, thus $q_1^{(1)} = 1$, and their common per-unit utility is $v_1^{(1)} = 0.17$, lower than 0.30.

The matrix on the left hand side of Table 3 describes the decentralization game. Coalition 1 is the row player. Group 2 is the column player but it has no choice (because it is inactive). As Lemma 6 shows, coalition 1 never benefits from decentralizing and free riding on the small countries that are already committed to produce no public good at all. Hence, at the unique equilibrium of the subgame, coalition 1 chooses C, and this is true for all the four examples in this section. Therefore, the analysis of scenario 1 is omitted in the remaining examples.

Second stage, scenario 2 (multipolar world): Group 2 has chosen to activate the coalition. In the second stage, two identical coalitions, 1 and 2, choose between centralization (C) and decentralization (D).

Four possible emission games can be induced by the choices of the two coalitions. First, if they both choose C, then at the equilibrium of the induced emission game, both of them discharge pollution of quantity $q_1 = q_2 = \frac{\sqrt{2}}{2} = 0.71$, and their per-unit utilities are $u_1 = u_2 = 0.84$. If one of the coalitions chooses D, say coalition 1, while the other one chooses C, the induced emission game is similar to the one in the first case of scenario 1 (just by swapping the roles of the two coalitions). Hence the quantity of pollution from coalition 1 is $q_1^{(1)} = m_1 = 1$ while that from coalition 2 is $q_2^{(1)} = 0.62$ at the equilibrium. The per-unit utility to coalition 1 is $v_1^{(1)} = 0.94 > u_1 = 0.84$. The benefits gained by coalition 1 by free riding on a centralized coalition 2 outweigh the loss from cooperation between its own member countries. This result is in contrast with scenario 1, where coalition 1 has no benefits from free riding on a group 2 of small autonomous countries. The two other emission games are similar to those in scenario 1.

The matrix on the right hand side of Table 3 describes the decentralization game. Coalition 1 is the row player and coalition 2 the column player. There are two stable Nash equilibria which are pure: (C, D) and (D, C), highlighted by stars. Indeed, a coalition has an incentive to decentralize if the other one is centralized, but not if the other one is decentralized.

This scenario could apply to the examples of the EU and the US in the introduction, where two active coalitions, the EU and the US, are present, but decision-making over each global public good (climate change and world security) is centralized in one coalition and decentralized in the other.

Table 3: Strategic decentralization depends on the nature of the other player.

(a) Coalition 2 is not active.

C	0.30∗, 0.94∗
D	0.17, 0.17

(b) Coalition 2 is active.

	C	D
C	0.84, 0.84	0.30∗, 0.94∗
D	0.94∗, 0.30∗	0.17, 0.17

First stage: The previous scenarios are induced by the choice of group 2 in the first stage. If it chooses to activate (resp. not to activate) the coalition, then the induced decentralization subgame is scenario 2 (resp. scenario 1). Comparing the equilibria in the two scenarios, we see that, unless group 2 is certain that the equilibrium attained in Scenario 2 is (C, D) (for this to be true, they must let coalition 1 be convinced that an active coalition 2 will choose decentralization), they strictly prefer not to activate the coalition. Instead of trying to convince coalition 1 that an active coalition 2 decentralizes, group 2 can simply choose not to activate the coalition in the first place. This choice of group 2 forces coalition 1 to commit itself to centralization in the second stage, which allows the member countries in group 2 to free ride.

5.2 Example 2: Two centralized coalitions in equilibrium

Assume $m_1 = m_2 = 1$, and the per-unit utility function of each player in the emission game in the last stage is

$$u_i(q) = q_i^{\frac{1}{2}} - Q^{\frac{3}{2}}. \tag{7}$$

Compared with the utility function (6) in Example 1, all the countries care more about the public good.

Second stage, scenario 2 (multipolar world): Group 2 has chosen to activate the coalition so that there are two identical coalitions who choose between C and D in the second stage. If both coalitions choose C, the equilibrium of the induced emission game is attained at $q_1 = q_2 = \sqrt{1/18} \approx 0.24$, with per-unit utilities $u_1 = u_2 = 0.16$. If coalition 1 chooses D, the equilibrium of the induced emission game becomes $q_1^{(1)} = 1, q_2^{(1)} = 0.1$. Coalition 2, the centralized one, decreases its emissions to compensate the free riding of coalition 1, the decentralized one. But the compensation is not very effective, since its emissions were already low. In consequence, the utility of the decentralized coalition is $v_1^{(1)} = -0.16 < u_1 = 0.16$. Therefore, when one coalition stays centralized, the other has no incentive for decentralization, which is different from the situation in Example 1. The unique equilibrium of this decentralization subgame represented by the matrix on the right hand side of Table 4 is (C, C).

Table 4: Two coalitions stay united at the equilibrium.

(a) Coalition 2 is not active.

C	$-0.84*$, $-0.16*$
D	-1.83, -1.83

(b) Coalition 2 is active.

	C	D
C	$0.16*$, $0.16*$	-0.84, -0.16
D	-0.16, -0.84	-1.83, -1.83

First stage: The previous scenarios are induced by the choice of group 2 in the first stage. Comparing the per-unit utility in equilibrium to the member countries of group 2 in the two scenarios, we can derive that group 2 chooses to activate the coalition in the first stage. Therefore, in the only SPNE of this three-stage game, both coalitions are active and centralized, and they provide a high level of public good. Besides, the multipolar world is strictly more efficient in providing public goods than the unipolar one. The reason for this result is that the countries care so much about the public good that the gain from free riding on the other coalition is offset by the loss of cooperation within a coalition.

5.3 Example 3: A unipolar world providing more public good than a multipolar world

In this example, the two groups have different weights: $m_1 = 1.18, m_2 = 0.82$. All players in the induced emission game have the same utility functions as defined by (6). In particular, the per-unit utility functions for the two groups when they are centralized coalitions are respectively

$$u_1(q) = 3\left(\frac{q_1}{1.18}\right)^{\frac{1}{2}} - Q^{\frac{3}{2}}, \tag{8}$$

$$u_2(q) = 3\left(\frac{q_2}{0.82}\right)^{\frac{1}{2}} - Q^{\frac{3}{2}}. \tag{9}$$

Second stage, scenario 2 (multipolar world): Group 2 has chosen to activate the coalition in the first stage. First, consider the case where coalition 1 chooses C. If coalition 2 makes the same choice, the equilibrium of the induced emission game is a corner solution: $q_2 = m_2 = 0.82$. This is due to the fact that coalition 2 is much smaller than coalition 1 so that it free rides on the latter even when it behaves as a centralized coalition. The emissions from both groups are thus the same as in the case where coalition 2 chooses D (with $q_1 = 0.60$, $u_1 = 0.45$). Given coalition 1's choice of C, coalition 2 is indifferent between C and D.

Next, consider the case that coalition 2 chooses C. The subcase that coalition 1 also chooses C has been discussed. If coalition 1 chooses D, the equilibrium of the induced emission game is attained at $q_1^{(1)} = m_1 = 1.18$ and $q_2^{(1)} = 0.66$. The latter is strictly lower than m_2, because coalition 2 prefers to provide a strictly positive amount of public good if no one else is willing to do so. The per-unit utility of coalition 1 is $v_1^{(1)} = 0.50 > u_1 = 0.45$. Coalition 1 has thus a strictly positive gain from decentralization by free riding on the centralized opponent.

The matrix on the right hand side of Table 5 describes the decentralization subgame. There are two equilibria which are pure, (D, C) and (C, D). However, the second one (distinguished by double stars) is not stable. This contrast results from the different weights of the two groups.

Table 5: A unipolar world provides more public good than a multipolar world.

(a) Coalition 2 is not active.

C	$0.45*$, $1.31*$
D	0.17, 0.17

(b) Coalition 2 is active.

	C	D
C	0.45, 1.31	$0.45**$, $1.31**$
D	$0.50*$, $0.20*$	0.17, 0.17

First stage: Comparing the per-unit utility of the member countries of group 2 at the equilibria in the two scenarios, we derive that they choose not to activate the coalition in the first stage. By doing so, they ensure a per-unit utility of 1.31, while if they activate the coalition, they obtain only 0.20 at the stable equilibrium. This is a good news for those caring about the aggregate provision of public good. Indeed, if group 2 activates the coalition, and coalition 1 decentralizes while a centralized coalition 2 has to provide all of the public good, it will provide less than what coalition 1 does in the case that group 2 free rides on a centralized coalition 1. The aggregate pollution in the first case is $Q^{(1)} = m_1 + q_2^{(1)} = 1.18 + 0.66 = 1.84$, while that in the second case is $Q = q_1 + m_2 = 0.60 + 0.82 = 1.42$.

Therefore, in this example (and in contrast to the previous example), a multipolar world provides less public good than a unipolar world.

5.4 Example 4: A group of countries whose dominant strategy as a coalition is centralization preferring not to activate the coalition

Consider two groups of countries with different weights and different preference for the public good. Group 1 has total weight $m_1 = 1.1$. Group 2 is smaller with total weight $m_2 = 0.9$. However, group 2 is much more affected by pollution than group 1: $s_1 = 1$, $s_2 = \frac{5}{2}$. In the context of climate change, group 2 can be composed of coastal countries, or can reflect the fact that developed countries are more affected by the impact of climate change. The per-unit utility functions for the two groups when they are active centralized coalitions in an emission game are thus respectively

$$u_1(q) = 3\left(\frac{q_1}{1.1}\right)^{\frac{1}{2}} - Q^{\frac{3}{2}}, \tag{10}$$

$$u_2(q) = 3\left(\frac{q_2}{0.9}\right)^{\frac{1}{2}} - \frac{5}{2}Q^{\frac{3}{2}}. \tag{11}$$

Second stage, scenario 2 (multipolar world): Group 2 has chosen to activate the coalition in the first stage. Consider the emission game induced by (C, C), the centralization of both coalitions. At the equilibrium, $q_1 = 0.87$, $q_2 = 0.17$, and the per-unit utilities are $u_1 = 1.60$ and $u_2 = -1.36$. Thus, even if coalition 1 is larger, coalition 2 makes a higher contribution because of its strong preference for the public good. Now suppose that coalition 1 chooses D while coalition 2 chooses C. At the equilibrium of the induced emission game, a decentralized coalition 1 emits $q_1^{(1)} = m_1 = 1.1$ of pollution while a centralized coalition 2 emits $q_2^{(1)} = 0.14$, which is lower than the level of its emissions 0.17 in the case where both coalitions are centralized. The utility of coalition 1 is $v_1^{(1)} = 1.61 > u_1 = 1.60$. Therefore, given the centralization of coalition 2, which is smaller but more sensitive to pollution, coalition 1, which is larger but less sensitive to pollution, prefers decentralizing.

We turn to the case where coalition 2 chooses D while coalition 1 chooses C. At the equilibrium of the induced emission game, compared with the case of (C, C), the pollution of coalition 2 is increased to $q_2^{(2)} = m_2 = 0.90$, while that of coalition 1 is reduced to $q_1^{(2)} = 0.60$. However, the unilateral decentralization is not advantageous for coalition 2: $v_2^{(2)} = -1.61 < u_2 = -1.36$. Indeed, according to matrix (b) in Table 6, C is a dominant strategy of coalition 2 in this multipolar subgame. The unique pure equilibrium is (D, C), so that the larger but less sensitive to pollution coalition 1 chooses strategic decentralization, while the smaller but more sensitive to pollution coalition 2 stays centralized and provides all of the public good.

First stage: By comparing the per-unit utility of the member countries of group 2 in the two scenarios, we derive that group 2 chooses not to activate the coalition in the first place. Indeed, otherwise they would be committed to centralization in the second stage, which allows coalition 1 to decentralize and free ride. Note that choice pair (C, D), though leading to the same utilities as in a unipolar world, is not an equilibrium of the decentralization game in a multipolar world. For coalition 2, strategy D is strictly dominated by C in

Table 6: Centralization as a dominant strategy.

(a) Coalition 2 is not active.

C	$0.38*,\ -1.61*$
D	$0.17,\ -4.07$

(b) Coalition 2 is active.

	C	D
C	$1.60,\ -1.36$	$0.38,\ -1.61$
D	$1.61*,\ -2.27*$	$0.17,\ -4.07$

scenario 2. On the contrary, by not activating the coalition in the first place, the member countries of group 2 "commit" themselves to playing D so as to "force" coalition 1 to play C, i.e. contributing to the public good. It is a typical example of situations where more choices induce a worse outcome.

Finally, note that the outcome of the decentralization game of two active coalitions (D, C) provides more public good ($Q^{(1)} = 1.24$) than the outcome when coalition 2 is not activated ($Q^{(2)} = 1.5$).

6 Comparative statics

6.1 Comparative statics in the emission game

This subsection discusses the impact of the variation of a player's weight or sensitivity to pollution on the equilibrium of an emission game.

The following proposition signifies that if the sensitivity to pollution of a player is increased, then its contribution to the public good will be increased while that of any other players will be reduced, and the aggregate level of the provision of the public good will be increased. Besides, the utility of that player will be decreased while that of the others will be increased.

Proposition 3. *In emission game $\Gamma(\mathcal{N})$, suppose that the sensitivity of player i to pollution is increased and becomes $\tilde{s}_i$ ($\tilde{s}_i > s_i$). Let $\tilde{\mathbf{q}} = (\tilde{q}_j)_{j\in\mathcal{N}}$ denote the new equilibrium and $\tilde{Q}$ the aggregate level of emission at the new equilibrium. Then*

(i) $\tilde{Q} \le Q$, $\tilde{q}_i \le q_i$, and each of the equalities holds if and only if $\tilde{q}_i = q_i = m_i$.

(ii) For each $j \in \mathcal{N} \setminus \{i\}$, $\tilde{q}_j \ge q_j$, and equality holds if and only if either $\tilde{q}_i = q_i = m_i$ or $\tilde{q}_j = q_j = m_j$.

(iii) If $\tilde{q}_i = q_i = m_i$, then $u_j(\tilde{\mathbf{q}}) = u_j(\mathbf{q})$ for all $j \in \mathcal{N}$. Otherwise, $u_i(\tilde{\mathbf{q}}) < u_i(\mathbf{q})$, and $u_j(\tilde{\mathbf{q}}) > u_j(\mathbf{q})$ for all $j \neq i$.

For the impact of the variation of the sensitivity s_i, we need the following assumption on function g. One can easily check the g's used in the numerical examples in Section 5 satisfy this assumption.

Assumption 2. *For all $q > 0$, function $x \mapsto \frac{g'(\frac{q}{x})}{x}$ is strictly decreasing for all $x \ge q$.*

The following proposition implies that, when a player i's weight is increased, there are two possibilities for its impact on the equilibrium. Roughly speaking, in the case that player i has not been polluting with full capacity, his emission level and utility will both be decreased, and so will the aggregate emission level; while for the other players, their emission levels and utilities will be increased. In the case that player i has been polluting with full capacity, then his emission level and the aggregate emission level will first be increased then be decreased, but his utility will always be decreased; while for the other players, their emission levels and utilities will all first be decreased then be increased.

Proposition 4. *In emission game $\Gamma(\mathcal{N})$, suppose that the weight of player i is increased and becomes $\tilde{m}_i$ ($\tilde{m}_i > m_i$). Let $\tilde{\mathbf{q}} = (\tilde{q}_j)_{j\in\mathcal{N}}$ denote the new equilibrium and $\tilde{Q}$ the aggregate level of emission at the new equilibrium. Then,*

1. $\frac{\tilde{q}_i}{\tilde{m}_i} \leq \frac{q_i}{m_i}$.
2. *If* $\frac{g'(\frac{q_i}{m_i})}{m_i} = s_i f(Q)$*, then, under Assumption 2,*

(i) $\tilde{Q} < Q$*,* $\tilde{q}_i < q_i$*, and for all* $j \neq i$*,* $\tilde{q}_j \geq q_j$ *where equality holds if and only if* $q_j = m_j$*.*

(ii) $u_i(\tilde{\mathbf{q}}) < u_i(\mathbf{q})$ *and for all* $j \neq i$*,* $u_j(\tilde{\mathbf{q}}) > u_j(\mathbf{q})$*.*

3. *If* $\frac{g'(\frac{q_i}{m_i})}{m_i} > s_i f(Q)$ *(so that* $q_i = m_i$*), then*

(1) for all $\tilde{m}_i$ *that is close enough to* m_i *for* $\frac{g'(\frac{q_i}{\tilde{m}_i})}{\tilde{m}_i} > s_i f(Q)$*, one has*

(i) $\tilde{Q} > Q$*,* $\tilde{q}_i > q_i$*, and for all* $j \neq i$*,* $\tilde{q}_j \leq q_j$ *where equality holds if and only if* $\tilde{q}_j = m_j$*.*

(ii) $u_i(\tilde{\mathbf{q}}) < u_i(\mathbf{q})$ *and for all* $j \neq i$*,* $u_j(\tilde{\mathbf{q}}) < u_j(\mathbf{q})$

(2) for all $\tilde{m}_i$ *such that* $\frac{g'(\frac{q_i}{\tilde{m}_i})}{\tilde{m}_i} \leq s_i f(Q)$*, under Assumption 2, one has*

(i) $\tilde{Q} \leq Q$ *and* $\tilde{q}_i \leq q_i$*, where each of the equalities holds if and only if* $\frac{g'(\frac{q_i}{\tilde{m}_i})}{\tilde{m}_i} = s_i f(Q)$*; for all* $j \neq i$*,* $\tilde{q}_j \geq q_j$ *where equality holds if* $\frac{g'(\frac{q_i}{\tilde{m}_i})}{\tilde{m}_i} = s_i f(Q)$ *or* $q_j = m_j$*.*

(ii) $u_i(\tilde{\mathbf{q}}) < u_i(\mathbf{q})$*; for all* $j \neq i$*,* $u_j(\tilde{\mathbf{q}}) \geq u_j(\mathbf{q})$ *where equality holds if and only if* $\frac{g'(\frac{q_i}{\tilde{m}_i})}{\tilde{m}_i} = s_i f(Q)$*.*

6.2 Comparative statics for the first two stages in the bipolar case

In this subsection, we first study the impact of the variation of a coalition's weight or sensitivity to pollution on the equilibrium of a two-coalition decentralization game. Then we further discuss how the increase of a potential coalition's weight or sensitivity to pollution affects its choice between activating and not activating the coalition, in the setting of a coalition versus a potential coalition analysed in Section 4.2.

First let us focus on the setting of Section 3.3.2 where there are two coalitions playing a decentralization game. Fix the characteristics of coalition 1, i.e. its weight m_1 and sensibility s_1, and one of the two characteristics of coalition 2. Let the other characteristic of coalition 2 increase, so that we can study the change of the equilibria of the game presented in Table 1(b). Let the lower bound of the weight of coalition 2 be $\underline{s}_2$ and that of its weight be $\underline{m}_2$. When coalition 2 has weight $\underline{m}_2$ and sensitivity $\underline{s}_2$, denote the equilibrium of the emission game played by two centralized coalitions 1 and 2 by $\underline{q}$. Assume that $\underline{q}_1 < m_1$ and $\underline{q}_2 < m_2$, i.e. both coalitions contribute to the public good in this case. Besides, for the sake of simplicity of the analysis, we assume that the weights of individual countries are negligible (mathematically, of measure zero) so that they will always decentralize whenever they determine the emission level independently.

In the following proposition, we fix m_2 and let s_2 be increased. One can take the equilibrium of the emission game corresponding to each entry of Table 1(b) and the resulting aggregate pollution as well as the utilities as functions of s_2. The first part of the proposition states that when s_2 is large enough, strategy "decentralization" D is strictly dominated by strategy "centralization" C for coalition 2 and, in particular, (C, C) is a better outcome than (C, D) for it. The second part of the proposition states that, for most of the values of (m_1, s_1), when s_2 is large enough, strategy "decentralization" D is either strictly or weakly dominated by strategy "centralization" C for coalition 1 and, in particular, (C, C) is the unique (locally) stable equilibrium of the game; in the case that (D, C) can also be an equilibrium, it is better for coalition 2 than (C, D).

Proposition 5. *In a two-coalition decentralization game, let m_1, s_1 and m_2 be fixed.*

1.For s_2 large enough, $u_2(s_2) > v_2^{(2)}(s_2)$ and hence strategy D is strictly dominated by strategy C for coalition 2. In particular, (C, D) is worse than (C, C) for coalition 2, and it is not an equilibrium of the game.

2. (1) If $g'(1) < s_1 f'(m_1)$, then $u_1(s_2) > v_1^{(1)}(s_2)$ for s_2 large enough and hence strategy D is strictly dominated by strategy C for coalition 1. In particular, (C, C) is the unique equilibrium of the game.

(2) If $g'(1) > s_1 f'(m_1)$, then $u_1(s_2) = v_1^{(1)}(s_2)$ for s_2 large enough and hence strategy D is weakly dominated by strategy C for coalition 1. In particular, (C, C) is the unique (locally) stable equilibrium of the game.

(3) If $g'(1) = s_1 f'(m_1)$, then $v_2^{(2)}(s_2) < v_2^{(1)}(s_2)$ for all s_2 large enough, i.e. (C, D) is worse than (D, C) for coalition 2. The equilibrium in this case can be (C, C) or (C, D).

Assumption 3. *For all $q > 0$, function $x \mapsto \frac{g'(\frac{q}{x})}{x}$ is strictly decreasing for all $x \geq q$ and $\frac{g'(\frac{q}{x})}{x}$ tends to 0 when x tends to $+\infty$.*

In the following proposition, we fix s_2 and let m_2 be increased. One can take the equilibrium of the emission game corresponding to each entry of Table 1(b) and the resulting aggregate pollution as well as the utilities as functions of m_2. The first part of the proposition states that when m_2 is large enough, strategy "decentralization" D is strictly dominated by strategy "centralization" C for coalition 2 and, in addition, (C, D) is worse than both (C, C) and (D, C). The second part of the proposition states similar results as Proposition 5: roughly speaking, in most of the cases strategy D is strictly or weakly dominated by strategy C for coalition 1 and the only (locally) stable equilibrium of the game is (C, C).

Proposition 6. *In a two-coalition decentralization game, let m_1, s_1 and s_2 be fixed.*

1. Under Assumption 2, there exists a $\bar{m}_2 \geq \underline{m}_2$ such that, for all $m_2 > \bar{s}_2$, $u_2(m_2) > v_2^{(2)}(m_2)$ and $v_2^{(1)}(m_2) > v_2^{(2)}(m_2)$. In particular, (C, C) and (D, C) are both better for coalition 2 than (C, D), and (C, D) is not an equilibrium of the game.

2. Under Assumption 3,

(1) If $g'(1) < s_1 f'(m_1)$, then $u_1(s_2) > v_1^{(1)}(s_2)$ for s_2 large enough and hence strategy D is strictly dominated by strategy C for coalition 1. In particular, (C, C) is the unique equilibrium of the game.

(2) If $g'(1) > s_1 f'(m_1)$, then $u_1(s_2) = v_1^{(1)}(s_2)$ for s_2 large enough and hence strategy D is weakly dominated by strategy C for coalition 1. In particular, (C, C) is the unique (locally) stable equilibrium of the game.

Finally consider the three-stage games with one active coalition 1 and a group 2 composed of countries which can potentially form a coalition.

Corollary 2. *In the setting of Proposition 5, for s_2 large enough, group 2 chooses to activate the coalition in the first stage of the game at the equilibrium.*

Corollary 3. *In the setting of Proposition 6, under Assumption 2, for m_2 large enough, group 2 chooses to activate the coalition in the first stage of the game at the equilibrium.*

The results of this section may serve to explain the change in the attitude of BRICS countries towards environmental cooperation. On April 22, 2015, the Environment Ministers of the five countries gathered in Moscow for their first official meeting to discuss green economy development and cooperation in tackling climate change.[9] Either their economic growth or their increasing concern for a sustainable development could have acted as a trigger for this summit.

Finally, we present two groups of simulations to illustrate the evolution of equilibrium behaviour respectively in stage 2 and stage 3, in the bipolar case, as described in Propositions 5–6 and Corollaries 2–3. In both groups, the utility functions are defined by $u_i(\mathbf{q}) = 3q_i^{\frac{1}{2}} - Q^{\frac{3}{2}}$.

The first group corresponds to case that $g'(1) > s_1 f'(m_1)$. The first simulation focuses on the evolution of the equilibria of a two-coalition decentralization game, i.e. the game from stage 2, when the weight and sensitivity of one of the coalitions vary while those of the other one are fixed. Fix the weight of an active coalition 1 to be $m_1 = 0.9$, and its sensitivity to be $s_1 = 1$. Let the weight of an active coalition 2, m_2, vary from 0.5 to 3, and its sensitivity, s_2, vary from 0.9 to 2.9.

In Figure 1, the areas in deep blue correspond to the range of (m_2, s_2) where the unique equilibrium is (C, C); the area in orange corresponds to the range of (m_2, s_2) where the

[9] http://www.brics.utoronto.ca/docs/150422-environment.html

Figure 1: Eq. in stage 2

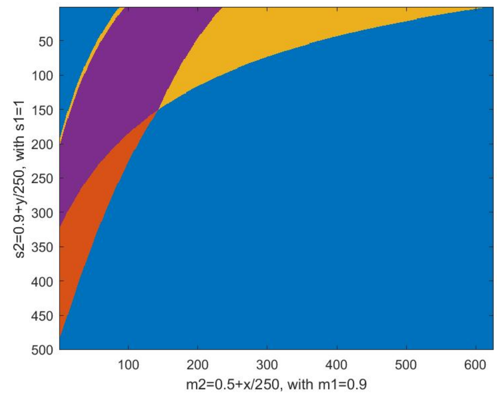
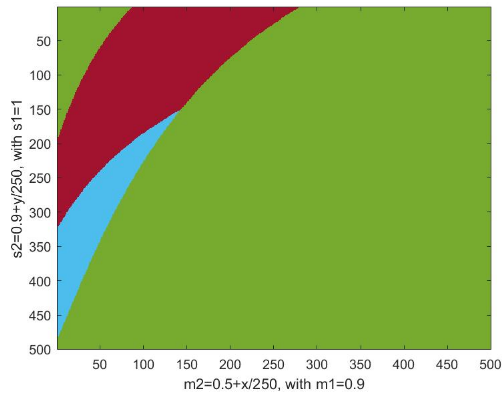


Figure 2: Eq. in stage 1



unique equilibrium is (C, D); the areas in yellow corresponds to the range of (m_2, s_2) where the unique equilibrium is (D, C); and the area in purple corresponds to the range of (m_2, s_2) where there are mulitple equilibria: (C, D) and (D, C). This figure is consistent with the results in Propositions 5 and 6 which state that when m_2 or s_2 is large enough, (C, C) is the unique equilibrium.

The second simulation focuses on the evolution of the choice of a potential coalition, group 2, in stage 1, facing coalition 1. The values of m_1 and s_1 and the ranges of m_2 and s_2 are the same as in the previous simulation.

In Figure 2, the areas in green correspond to the range of (m_2, s_2) where group 1 chooses to activate the coalition; the area in red correspond to the range of (m_2, s_2) where group 1 chooses not to activate the coalition; and the area in pale blue corresponds to the area where group 1 is indifferent between activating or not the coalition. This figure is consistent with the results in Corollaries 2 and 3, which state that when m_2 or s_2 is large enough, group 2 will activate the coalition.

Figure 3: Eq. in stage 2

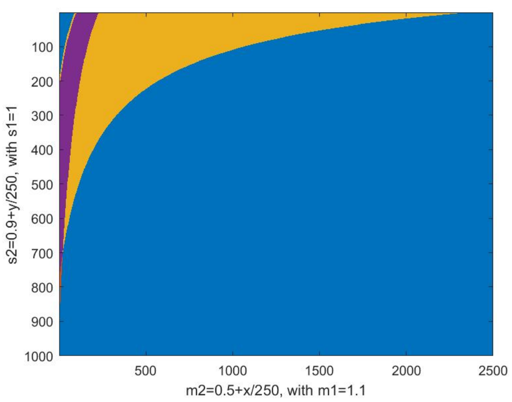
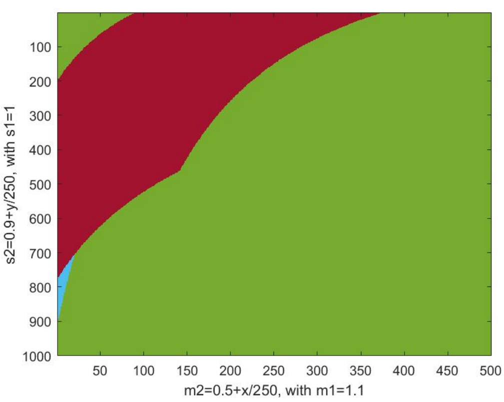


Figure 4: Eq. in stage 1



The second group corresponds to case that $g'(1) < s_1 f'(m_1)$. Fix the weight of an active coalition 1 to be $m_1 = 1.1$, and its sensitivity to be $s_1 = 1$. The ranges of the weight and the sensitivity of coalition/group 2 are $0.5 \leq m_2 \leq 10.5$, and $0.9 \leq s_2 \leq 4.9$. The colours have the same meaning as in the first group of simulations. It is easy to see that the simulations (Figures 3–4) are consistent with the results in Propositions 5–6 and Corollaries 2–3.

7 Conclusion

Coalitions of countries build institutions that make them able to act in a centralized way on many global issues. But some countries would never be able to be united into the same coalition, because binding agreements including compensations and punishments involve a high level of trust. This is why we observe in practice competing coalitions of countries, states and regions. These coalitions act in concert on some global issues, but also let their members act separately on other ones. In this paper, we have shown that decentralization can be strategic and benefit all the members of a coalition. Such behavior acts as a commitment to free riding. Decentralization therefore benefits a coalition when the gain from free riding on other actors exceeds the loss caused by not internalizing the externalities exerted by the coalition members on each other. Therefore, if the existence of a second, smaller, coalition gives the first, larger, coalition an incentive to decentralize, a multipolar world with these two coalitions can provide less public good than the unipolar world where only the first coalition exists.

A striking consequence of this analysis is that a relatively small group of countries with a strong interest in the public good is doomed if they are organized as a coalition. Indeed, for such a coalition, the gain from cooperation is high and investing in the public good is a dominant strategy, regardless of the behavior of the other players. Therefore, other coalitions less interested in the public good may decide to decentralize their provision in order to free ride on this highly motivated coalition. This potential risk invites these pro-public good countries to think twice before activating their coalition.

Our static analysis does not embody the fact that when institutions are built, actors do not necessarily anticipate every single future global issue. For instance, it is very likely that the founding fathers of the United States did not have climate change in mind at the time of drafting the constitution. What we claim is that, even in the presence of uncertainty, actors anticipate that building cooperative institutions may encourage the rest of the world to see them as a motivated coalition on all global issues, and thus generate more free riders.

Appendix 1: The emission game expressed in terms of public good

Let us show that the emission game studied in this paper is exactly the mirror of a public good provision game, with identical properties. For the sake of simplicity, suppose that all the actors have the same taste for the public good.

First assume an individual j living in autarchy. She provides a public good, "abatements" in quantities a_j. The marginal benefits from the public good are decreasing, and the marginal cost of production is increasing (either because of decreasing marginal returns or the fact that taxation is distortionary). Let the benefits function and the cost function be denoted by $\tilde{f}$ and $\tilde{g}$ respectively. Then her utility function is

$$u = \tilde{f}(a_j) - \tilde{g}(a_j),$$

with $\tilde{f}' > 0$, $\tilde{f}'' < 0$, $\tilde{g}' > 0$, $\tilde{g}'' > 0$.

Consider now that individual j is part of a group i composed of m_i individuals identical to j. Hence, group i produces a quantity $a_i = m_i a_j$ of the public good. There are no other agents in this economy. The utility function of individual j becomes

$$u = \tilde{f}(a_i) - \tilde{g}(a_j) = \tilde{f}(a_i) - \tilde{g}\left(\frac{a_i}{m_i}\right)$$

Next, assume this group i exists in a larger world which is composed of M individuals identical to j (possibly divided into several groups) who produce an aggregate amount A of public good. Hence, individual j now consumes a quantity A of public good. The utility function of individual j becomes

$$u = \tilde{f}(A) - \tilde{g}\left(\frac{a_i}{m_i}\right) \tag{12}$$

Finally, in the context of pollution emission, equation (4) is just a reformulation of (12). To see this, notice that abatements correspond to pollution not emitted. Identify individual j to a unit of weight and group i to a country of weight m_i. Hence, $a_i = m_i - q_i$ and $A = M - Q$. Also, define $f(x) = -\tilde{f}(M - x)$ and $g(x) = -\tilde{g}(1 - x)$. Then, (12) rewrites

$$u_j = g\left(\frac{q_i}{m_i}\right) - f(Q)$$

with $g' > 0$, $g'' < 0$, so that g is strictly increasing and strictly concave, and $f' > 0$, $f'' > 0$ so that f is strictly increasing and strictly convex.

Appendix 2: Formal proofs

The following auxiliary result is used in several proofs.

Lemma 7. *A Nash equilibrium* $\mathbf{q} = (q_1, \ldots, q_N)$ *of the emission game satisfies that, for every player* i*, either* $0 < q_i < m_i$ *and* $s_i\, f'(Q) = g'(\frac{q_i}{m_i})/m_i$*, or* $q_i = m_i$ *and* $s_i\, f'(Q) \leq g'(1)/m_i$*.*

Proof. It is simply a reformulation of the first order condition of the optimization problem (5). □

Proof of Lemma 1. The existence of an equilibrium can be proved via a standard approach of variational inequalities or fixed point theorem.

The uniqueness of the equilibrium can be proved by contradiction. Suppose that $\mathbf{q}$ and $\mathbf{q}^*$ are two equilibria. Let us first show that $Q = Q'$, where $Q = \sum_i q_i$ and $Q' = \sum_i q_i'$. If it is not true, say, for example, $Q > Q'$, then $f(Q) > f(Q')$. According to Lemma 7, there are only two possibilities:

(i) $s_i\, f'(Q) = \frac{g'(\frac{q_i}{m_i})}{m_i}$ and $0 < q_i < m_i$, then $q_i < q_i' \leq m_i$;

(ii) $s_i\, f'(Q) \leq \frac{g'(\frac{q_i}{m_i})}{m_i}$ and $q_i = m_i$, then $q_i = q_i' = m_i$.

In both cases, one has $\sum_i q_i \leq \sum_i q_i'$, which is contradictory to the hypothesis that $Q > Q'$. Therefore, $Q = Q'$. □

Proof of Lemma 2. (i) If $q_i < q_j$, then $s_i\, f'(Q) = \frac{g'(\frac{q_j}{m_i})}{m_i} > \frac{g'(\frac{q_i}{m_j})}{m_j}$. This is impossible according to Lemma 7 considering that $s_i = s_j$. Hence, $q_i \geq q_j$. Symmetrically, $q_j \geq q_i$, thus $q_i = q_j$ and, consequently, $u_i = u_j$.

(ii) Suppose that $\frac{q_i}{m_i} < 1$. Then $s_i\, f'(Q) = \frac{g'(\frac{q_i}{m_i})}{m_i} > \frac{g'(\frac{q_i}{m_i})}{m_j}$. Since $s_j\, f'(Q) \leq \frac{g'(\frac{q_j}{m_j})}{m_j}$, one has $\frac{g'(\frac{q_i}{m_i})}{m_j} < \frac{g'(\frac{q_j}{m_j})}{m_j}$ because $s_i = s_j$. This implies that $\frac{q_j}{m_j} < \frac{q_i}{m_i}$.

Suppose that $\frac{q_i}{m_i} = 1$, then it is clear that $\frac{q_j}{m_j} \leq 1 = \frac{q_i}{m_i}$.

In the case that $\frac{q_j}{m_j} < \frac{q_i}{m_i} \leq 1$, since $g(\frac{q_j}{m_j}) > g(\frac{q_i}{m_i})$ and $s_i = s_j$, $u_i = g(\frac{q_i}{m_i}) - s_i\, f(Q) > g(\frac{q_j}{m_j}) - s_j\, f(Q) = u_j$. In the case that $\frac{q_j}{m_j} = \frac{q_i}{m_i} = 1$, one has $u_i = u_j = g(1) - s_i\, f(Q)$.

(iii) Suppose that $\frac{q_i}{m_i} < 1$. If $\frac{q_j}{m_j} = 1$, then $\frac{g'(1)}{m_j} = s_j\, f'(Q) > s_i\, f'(Q)$, which implies that $\frac{q_i}{m_i} = 1$ since $m_i = m_j$. Hence $\frac{q_j}{m_j} < 1$. Consequently, $\frac{g'(\frac{q_j}{m_j})}{m_j} = s_j\, f'(Q) > s_i\, f'(Q) = \frac{g'(\frac{q_i}{m_i})}{m_i}$. Hence $g'(\frac{q_j}{m_j}) > g'(\frac{q_i}{m_i})$ considering that $m_i = m_j$. Finally $\frac{q_j}{m_j} < \frac{q_i}{m_i}$.

Suppose that $\frac{q_i}{m_i} = 1$, then it is clear that $\frac{q_j}{m_j} \leq 1 = \frac{q_i}{m_i}$.

In the case that $\frac{q_j}{m_j} < \frac{q_i}{m_i} \leq 1$, since $g(\frac{q_j}{m_j}) > g(\frac{q_i}{m_i})$ and $s_i < s_j$, $u_i = g(\frac{q_i}{m_i}) - s_i\, f(Q) > g(\frac{q_j}{m_j}) - s_j\, f(Q) = u_j$. In the case that $\frac{q_j}{m_j} = \frac{q_i}{m_i} = 1$, one has $u_i = g(1) - s_i\, f(Q) < g(1) - s_j\, f(Q) = u_j$. □

Proof of Lemma 3. By assumption, $g'(1) > 0$. Hence, there exists ϵ such that $s_i f'(M) \leq \frac{g'(1)}{\epsilon}$ for all i. Consequently, for all strategy profile $\mathbf{q}$, $s_i f'(Q) \leq s_i f'(M) \leq \frac{g'(1)}{m_i}$ as long as $m_i \leq \epsilon$. □

Proof of Lemma 4. Let $\mathbf{q}$ be the equilibrium in emission game $\Gamma(\mathcal{N})$, and $\mathbf{q}^{(i)}$ the equilibrium in emission game $\Gamma(\mathcal{N}^{(i)})$ induced by the decentralization of coalition i. Let $Q = \sum_{k\in\mathcal{N}} q_k$ and $Q^{(i)} = \sum_{k\in\mathcal{N}^{(i)}} q_k^{(i)}$. The Lemma can be reformulated as follows.

If $q_i < m_i$, then, (i) $Q^{(i)} > Q$; (ii) for each $j \in \mathcal{N}$ and $j \neq i$, either $q_j^{(i)} < q_j$ or $q_j^{(i)} = q_j = m_j$; $\sum_{j\in\mathcal{J}_i} q_{i_j}^{(i)} > q_i$; (iii) for each $j \in \mathcal{N}$ and $j \neq i$, $u_j(\mathbf{q}^{(i)}) < u_j(\mathbf{q})$. If $q_i = m_i$, then $Q^{(i)} = Q$, $q_{i_j}^{(i)} = m_{i_j}$ for all $j \in \mathcal{J}_i$, $q_j^{(i)} = q_j$ and $u_j(\mathbf{q}^{(i)}) = u_j(\mathbf{q})$ for all $j \in \mathcal{N}$ and $j \neq i$.

It is sufficient to prove for the particular case where coalition i is composed of two countries i_1 and i_2 respectively of weight μ_1 and μ_2. The general result can be obtained by induction on the number of the coalition's members. Suppose that $\mu_1 \leq \mu_2$. For the simplicity of notation, denote the emissions of μ_1 by η_1 that of μ_2 by η_2 at the equilibrium $\mathbf{q}^{(i)}$ after the decentralization of i.

First consider the case $q_i < m_i$.

(i) Since $\mu_1 \leq \mu_2$, according to Lemma 2, $\frac{\eta_1}{\mu_1} \geq \frac{\eta_2}{\mu_2}$. There are two possibilities.

a) $\frac{\eta_2}{\mu_2} > \frac{q_i}{m_i}$. Then $\frac{\eta_1+\eta_2}{\mu_1+\mu_2} > \frac{q_i}{m_i}$, i.e. $\eta_1 + \eta_2 > q_i$. Suppose that $Q^{(i)} \leq Q$. Then there must be some player $j \neq i$ who discharges fewer emissions now, i.e. $q_j^{(i)} < q_j \leq m_j$. Thus $s_i f'(Q^{(i)}) = \frac{g'(\frac{q_j^{(i)}}{m_j})}{m_j} > s_i \frac{g'(\frac{q_j}{m_j})}{m_j} \geq f'(Q)$, which implies $Q^{(i)} > Q$. This contradicts the assumption that $Q^{(i)} \leq Q$, hence $Q^{(i)} > Q$.

b) $\frac{\eta_2}{\mu_2} \leq \frac{q_i}{m_i}$ and hence lower than 1. Suppose that $Q^{(i)} \leq Q$. Then $g'(\frac{q_i}{m_i}) \geq m_i s_i f'(Q) > \mu_2 s_i f'(Q^{(i)}) = g'(\frac{\eta_2}{\mu_2})$. In particular, $g'(\frac{q_i}{m_i}) > g'(\frac{\eta_2}{\mu_2})$ which implies that $\frac{q_i}{m_i} < \frac{\eta_2}{\mu_2}$. This contradicts the assumption that $\frac{\eta_2}{\mu_2} \leq \frac{q_i}{m_i}$, hence $Q^{(i)} > Q$.

(ii) Suppose that there exists a player $j \neq i$ such that $q_j < m_j$ and $q_j^{(i)} \geq q_j$. Then $g'(\frac{q_j}{m_j}) = m_j s_j f'(Q) < m_j s_j f'(Q^{(i)}) \leq g'(\frac{q_j^{(i)}}{m_j})$, which implies $q_j > q^{(i)}$, a contradiction. If $q_j = m_j$, then $q_j^{(i)} \leq m_j = q_j$. Since $q_j^{(i)} \leq q_j$ for all $j \in \mathcal{N}$ and $j \neq i$, while $Q^{(i)} > Q$, one deduces that $\eta_1 + \eta_2 > q_i$.

(iii) Clear from (i), (ii) and the definition of utilities.

Now consider the case $q_i = m_i$. According to Lemma 7, $s_i f'(Q) \leq \frac{g'(1)}{m_i} < \frac{g'(1)}{m_{i_j}}$ for all $j \in \mathcal{J}_i$, where the second inequality is due to the fact that the weight of any member of coalition i is lower than the total weight of coalition. Therefore, the Nash equilibrium conditions in Lemma 7 are satisfied for all player $l \neq i$ and all member $j \in \mathcal{J}_i$ of coalition i for $Q^{(i)} = Q$. The other players' behavior does not change after the decentralization of i while its own members still provide no public good. □

Proof of Lemma 5. Let $\mathbf{q}$ be the equilibrium in emission game $\Gamma(\mathcal{N})$, and $\mathbf{q}^{(i)}$ the equilibrium in emission game $\Gamma(\mathcal{N}^{(i)})$ induced by the decentralization of coalition i, for $i = 1, 2$. Denote $Q = \sum_{k\in\mathcal{N}} q_k$ and $Q^{(i)} = \sum_{k\in\mathcal{N}^{(i)}} q_k^{(i)}$, for $i = 1, 2$. Suppose that $s_1 = s_2$, $m_1 > m_2$, $q_1 < m_1$, and $q_{i_l}^{(1)} = m_{1_l}$ for each $l \in \mathcal{J}_1$ and $q_{2_k}^{(2)} = m_{2_k}$ for each $k \in \mathcal{J}_2$. The Lemma can be reformulated as follows: (i) $Q^{(1)} > Q^{(2)}$; (ii) for each $j \in \mathcal{N}$ and $j \neq 1, 2$, $q_j^{(1)} < q_j^{(2)}$ or $q_j^{(1)} = q_j^{(2)} = m_j$; (iii) for each $j \in \mathcal{N}$ and $j \neq 1, 2$, $u_j(\mathbf{q}^{(1)}) < u_j(\mathbf{q}^{(2)})$.

(i) If at $\mathbf{q}$, $q_2 = m_2$, then the decentralization of coalition 2 does not change the emissions of any player. Thus $Q^{(2)} = Q$. Besides, $Q^{(1)} > Q$ according to Lemma 4. Hence $Q^{(1)} > Q^{(2)}$.

If $q_2 < m_2$, then by Lemma 2, $\frac{q_1}{m_1} < \frac{q_2}{m_2} < 1$. Suppose that $Q^{(1)} \leq Q^{(2)}$. Let us first show that $m_1 + q_2^{(1)} \geq m_2 + q_1^{(2)}$.

Case 1: There are only two players, coalitions 1 and 2, in the game. Then $m_1 + q_2^{(1)} \leq m_2 + q_1^{(2)}$ is equivalent to $Q^{(1)} \leq Q^{(2)}$.

Case 2: There are more than two players in the game. Consider player $j \neq 1, 2$. If $q_j^{(2)} = m_j$, then $m_j s_j f'(Q^{(1)}) \leq m_j s_j f'(Q^{(2)}) \leq g'(1)$, which implies that $q_j^{(1)} = m_j = q_j^{(2)}$. If $q_j^{(2)} < m_j$, then $m_j s_j f'(Q^{(1)}) \leq m_j s_j f'(Q^{(2)}) = g'(\frac{q_i^{(2)}}{m_j})$. Thus, either $q_j^{(1)} \geq q_j^{(2)}$ when $q_j^{(1)} < m_j$, or $q_j^{(1)} = m_j > q_j^{(2)}$ when $q_j^{(1)} = m_j$. In any case, $q_j^{(1)} \geq q_j^{(2)}$. In consequence, $\sum_{j \neq 1,2} q_j^{(1)} \geq \sum_{j \neq 1,2} q_j^{(2)}$. Considering the hypothesis $Q^{(1)} \leq Q^{(2)}$, one deduces that $m_1 + q_2^{(1)} \leq m_2 + q_1^{(2)}$.

Having shown that $m_1 + q_2^{(1)} \leq m_2 + q_1^{(2)}$, one obtains $m_1 - q_2^{(1)} \leq m_2 - q_2^{(1)}$, which implies that $\frac{m_1 - q_2^{(1)}}{m_1} < \frac{m_2 - q_2^{(1)}}{m_2}$. Consequently, $\frac{q_2^{(1)}}{m_1} > \frac{q_1^{(2)}}{m_2}$, which yields

$$\frac{g'(\frac{q_1^{(2)}}{m_1})}{m_1} < \frac{g'(\frac{q_2^{(1)}}{m_2})}{m_2}. \tag{13}$$

On the other hand, since $q_1 < m_1$ and $q_2 < m_2$, Lemma 4 implies that $q_1^{(2)} < q_1$ and $q_2^{(1)} < q_2$. Therefore, $\frac{g'(\frac{q_1^{(2)}}{m_1})}{m_1} = s_1 f'(Q^{(2)}) \geq s_2 f'(Q^{(1)}) = \frac{g'(\frac{q_2^{(1)}}{m_2})}{m_2}$. This is contradictory to (13).

(ii) First notice that, according to (i), $f'(Q^{(2)}) < f'(Q^{(1)})$. Consider player $j \neq 1, 2$.

If $q_j^{(1)} = m_j$, then $m_j s_j f'(Q^{(2)}) < m_j s_j f'(Q^{(1)}) \leq g(1)$. Thus, $q_j^{(2)} = m_j$.

If $q_j^{(1)} < m_j$, then one has only to prove the result in the case that $q_j^{(2)} < m_j$. One has $g'(\frac{q_j^{(2)}}{m_j}) = m_j s_j f'(Q^{(2)}) < m_j s_j f'(Q^{(1)}) = g'(\frac{q_j^{(1)}}{m_j})$, which implies that $q_j^{(2)} > q_j^{(1)}$.

(iii) Clear from (i), (ii) and the definition of utilities. □

Proof of Lemma 6. First consider the unilateral decentralization of coalition i.

(i) According to Lemma 4, the emission level of the members of coalition i does not change before and after its unilateral decentralization: they never provide any public good, while the other players' behavior does not change either.

(ii) First notice that $Q = q_i + \sum_{j \neq i} q_j$, and $Q^{(i)} = \sum_{l \in \mathcal{J}_i} q_{i_l}^{(i)} + \sum_{j \neq i} q_j$. For the simplicity of notation, let $\omega_i = \sum_{l \in \mathcal{J}_i} q_{i_l}^{(i)}$ and $Q^{-i} = \sum_{j \neq i} q_j$. According to Lemma 4, $Q^{(i)} > Q$, i.e. $w_i > q_i$. Besides, $Q^{(i)} - Q^i = \omega_i - q_i$.

Since $q_i < m_i$, one knows that $s_i f'(Q) = \frac{g'(\frac{q_i}{m_i})}{m_i}$ by Lemma 2. Let this strictly positive real number be denoted by A. Then, because of the strict convexity of f and the strict concavity of g, one deduces that

$$f'(x) > A, \quad \forall\, x \in (Q, M),$$
$$\frac{g'(\frac{x}{m_i})}{m_i} = \frac{\mathrm{d}}{\mathrm{d}x} g(\frac{x}{m_i}) < A, \forall\, x \in (q_i, m_i),$$

Therefore,

$$u_i(\mathbf{q}) - \sum_{i_l \in \mathcal{J}_i} \frac{m_{i_l}}{m_i} u_{i_l}(\mathbf{q}^{(i)}) = \big[g\big(\frac{q_i}{m_i}\big) - s_i f(Q)\big] - \big[g\big(\frac{\omega_i}{m_i}\big) - s_i f(Q^{(i)})\big]$$
$$= s_i[f(Q^{(i)} - f(Q)] - \big[g\big(\frac{\omega_i}{m_i}\big) - g\big(\frac{q_i}{m_i}\big)\big]$$
$$= \int_Q^{Q^{(i)}} s_i f'(x)\mathrm{d}x - \int_{q_i}^{\omega_i} \frac{\mathrm{d}}{\mathrm{d}x} g\big(\frac{x}{m_i}\big)\mathrm{d}x > (Q^{(i)} - Q)A - (\omega_i - q_i)A = 0.$$

Finally, notice that coalitions i's utility can only be further reduced by the simultaneous decentralization of any of the other coalitions, according to Lemma 4. □

Proof of Proposition 1. According to Lemma 6, $v_1^{(1,2)} < v_1^{(2)}$ and $v_2^{(1,2)} < v_2^{(1)}$ (since $q_2^{(1)} < m_2$). Hence the strategy pair (D, D) is never a Nash equilibrium.

The following results can be obtained by Lemma 4:

$$\begin{aligned} q_1^{(2)} \leq q_1 < m_1 \Rightarrow &\quad v_1^{(2)} < u_1, \\ q_2^{(1)} \leq q_2 \leq m_2 \Rightarrow &\quad v_2^{(1)} < u_2, \\ q_2^{(1)} < m_2 \Rightarrow &\quad v_1^{(1,2)} < v_1^{(1)}, \\ q_1^{(2)} < m_1 \Rightarrow &\quad v_2^{(1,2)} \leq v_2^{(2)}. \end{aligned}$$

First notice that it is impossible to have $u_1 = v_1^{(1)}$ and $u_2 = v_2^{(2)}$. Otherwise, $q_1 = m_1$, absurd according to the hypothesis.

(i): If $u_1 > v_1^{(1)}$ and $u_2 > v_2^{(2)}$, it is clear that (C, C) is the unique, strict and thus stable equilibrium. If $u_1 > v_1^{(1)}$ and $u_2 = v_2^{(2)}$, then (C, C), (C, D) and all the strategy pairs where coalition 1 plays C and coalition 2 plays a mixed strategy are equilibria. However, only (C, C) is stable. Indeed, consider an equilibrium where coalition 2 plays D with a strictly positive probability $\lambda \in]0, 1]$. If coalition 1 perturbs its strategy by playing D with a strictly positive probability, then the unique best reply of coalition 2 is to play C with probability 1, since $v_2^{(1,2)} < v_2^{(1)}$ and $u_2 = v_2^{(2)}$. The symmetric analysis applies to the case where $u_1 = v_1^{(1)}$ and $u_2 > v_2^{(2)}$.

(ii): If $u_1 > v_1^{(1)}$ and $u_2 < v_2^{(2)}$, it is clear that (C, D) is the unique, strict and thus stable equilibrium. If $u_1 = v_1^{(1)}$ and $u_2 < v_2^{(2)}$, then there is another equilibrium (D, C). However this equilibrium is not stable because if coalition 2 perturbs its strategy by playing D with a strictly positive probability, then the unique best reply of coalition 1 is to play C with probability 1, since $u_1 = v_1^{(1)}$ and $v_1^{(1,2)} < v_1^{(2)}$.

(iii): Symmetric to (ii).

(iv): It is clear that (C, D) and (D, C) are two strict and thus stable equilibria. There is another equilibrium which is completely mixed: $\big(\frac{v_2^{(1)}-v_2^{(1,2)}}{v_2^{(1)}-v_2^{(1,2)}+v_2^{(2)}-u_2}C+\frac{v_2^{(2)}-u_2}{v_2^{(1)}-v_2^{(1,2)}+v_2^{(2)}-u_2}D,$ $\frac{v_1^{(2)}-v_1^{(1,2)}}{v_1^{(2)}-v_1^{(1,2)}+v_1^{(1)}-u_1}C+\frac{v_1^{(1)}-u_1}{v_1^{(2)}-v_1^{(1,2)}+v_1^{(1)}-u_1}D\big)$. However it is not stable. For example, for all the strategy of coalition 2 which plays C with probability greater than $\frac{v_1^{(2)}-v_1^{(1,2)}}{v_1^{(2)}-v_1^{(1,2)}+v_1^{(1)}-u_1}$, the unique best replay of coalition 1 is to play D with probability 1. In turn, the unique best reply of coalition 2 is to play C with probability 1.

□

Proof of Proposition 2. In each case, group 1 compare $v_2^{(2)}$ and all its equilibrium utilities in the decentralization game on the right hand side of Table 2. It activates the coalition if it is better off at at least one of the latter; it does not activate the coalition if it is worse off at at least one of the latter; it is indifferent if it has the same utilities at the latter equilibria. The remaining is easy to deduce from Proposition 1. □

Proof of Proposition 3. (i)-(ii): Consider the case $\tilde{Q} \geq Q$. On the on hand, by the same arguments as in the proof of Lemma 1, one can show that for each $j \in \mathcal{N}$ and $j \neq i$, $\tilde{q}_j \leq q_j \leq m_j$. On the other hand, for i, note that $\tilde{s}_i f'(\tilde{Q}_i) > s_i f(Q)$. If $q_i = m_i$, then $\tilde{q}_i \leq q_i$. If $q_i < m_i$, then $\frac{g'(1)}{m_i} < \frac{g'(\frac{q_i}{m_i})}{m_i} = s_i f'(Q) < \tilde{s}_i f'(\tilde{Q})$. Hence $\tilde{q}_i < m_i$ and $\frac{g'(\frac{\tilde{q}_i}{m_i})}{m_i} = \tilde{s}_i f'(\tilde{Q}) > s_i f'(Q) = \frac{g'(\frac{q_i}{m_i})}{m_i}$, which implies that $\tilde{q}_i < q_i$. Therefore $\sum_{j\in\mathcal{N}} \tilde{q}_j \leq \sum_{j\in\mathcal{N}} q_j$. Since $\tilde{Q} \geq Q$, this is possible only if $\tilde{q}_i = q_i = m_i$ and $\tilde{q}_j = q_j$ for all $j \neq i$.

Consider the case $\tilde{Q} < Q$. By the same arguments as in the proof of Lemma 1, for each $j \neq i$, either $\tilde{q}_j > q_j$, or $\tilde{q}_j = q_j = m_j$. Hence $\tilde{q}_i < q_i$.

(iii) The result for the case $\tilde{q}_i = q_i = m_i$ is immediate.

In the case that $\tilde{q}_i < q_i \leq m_i$, recall that $\tilde{Q} < Q$. For each $j \neq i$, since $\tilde{q}_j \geq q_j$, one has $u_j(\mathbf{q}) < u_j(\tilde{\mathbf{q}})$. Let constant $B = s_i f'(Q)$. It is not greater than $\frac{g(\frac{q_i}{m_i})}{m_i}$ (condition for equilibrium $\mathbf{q}$). Then, for i,

$$
\begin{aligned}
u_i(\mathbf{q}) - u_i(\tilde{\mathbf{q}}) &= \big[g\big(\frac{q_i}{m_i}\big) - s_i f(Q)\big] - \big[g\big(\frac{\tilde{q}_i}{m_i}\big) - \tilde{s}_i f(\tilde{Q})\big] \\
&= \big[g\big(\frac{q_i}{m_i}\big) - g\big(\frac{\tilde{q}_i}{m_i}\big)\big] + [s_i f(\tilde{Q}) - s_i f(Q)] + [\tilde{s}_i f(\tilde{Q}) - s_i f(\tilde{Q})] \\
&= \int_{\tilde{q}_i}^{q_i} \frac{\mathrm{d}}{\mathrm{d}x} g\big(\frac{x}{m_i}\big)\,\mathrm{d}x - \int_{\tilde{Q}}^{Q} s_i f'(x)\,\mathrm{d}x + (\tilde{s}_i - s_i) f(\tilde{Q}) \\
&> (q_i - \tilde{q}_i)B - (Q - \tilde{Q})B = [(q_i - \tilde{q}_i) - (Q - \tilde{Q})]B \\
&\geq 0
\end{aligned}
$$

The first inequality is due to the concavity of g and the convexity of f. The second inequality is because $q_j \geq \tilde{q}_j$ for all $j \neq i$. □

Proof of Proposition 4. 1. Suppose that $\frac{\tilde{q}_i}{\tilde{m}_i} > \frac{q_i}{m_i}$. Then $\tilde{q}_i > q_i$ since $\tilde{m}_i > m_i$. Also,

$$\frac{g'(\frac{\tilde{q}_i}{\tilde{m}_i})}{\tilde{m}_i} < \frac{g'(\frac{q_i}{m_i})}{\tilde{m}_i} < \frac{g'(\frac{q_i}{m_i})}{m_i}.$$

If $\tilde{Q} \leq Q$, then $\tilde{q}_j \geq q_j$ for all j. This is impossible because $\tilde{q}_i > q_i$. Hence $\tilde{Q} > Q$, and thus $s_i f'(Q) < s_i f'(\tilde{Q}) \leq \frac{g'(\frac{\tilde{q}_i}{\tilde{m}_i})}{\tilde{m}_i} < \frac{g'(\frac{q_i}{m_i})}{m_i}$, which implies that $q_i = m_i$. This is absurd because $\frac{\tilde{q}_i}{\tilde{m}_i} > \frac{q_i}{m_i}$.

2. Under Assumption 2, consider the case where $\frac{g'(\frac{q_i}{m_i})}{m_i} = s_i f(Q)$. Suppose $\tilde{Q} \geq Q$. Then for each $j \neq i$, $\tilde{q}_j \leq q_j$, hence $\tilde{q}_i \geq q_i$ and in consequence,

$$\frac{g'(\frac{\tilde{q}_i}{\tilde{m}_i})}{\tilde{m}_i} \leq \frac{g'(\frac{q_i}{\tilde{m}_i})}{\tilde{m}_i} < \frac{g'(\frac{q_i}{m_i})}{m_i},$$

where the second inequality is because $\frac{g'(\frac{q_i}{x})}{x}$ is strictly decreasing in x. One has $s_i f(\tilde{Q}) \leq \frac{g'(\frac{\tilde{q}_i}{\tilde{m}_i})}{\tilde{m}_i} < \frac{g'(\frac{q_i}{m_i})}{m_i} = s_i f(Q)$ which implies that $\tilde{Q} < Q$. This is absurd. We have thus proved $\tilde{Q} < Q$. Consequently, $\tilde{q}_j \geq q_j$ for all $j \neq i$ where equality holds only if $q_j = m_j$. Therefore $\tilde{q}_i < q_i$. It is obvious that $u_j(\tilde{\mathbf{q}}) > u_j(\tilde{q})$ for all $j \neq i$. Let constant $B = s_i f'(Q) = \frac{g(\frac{q_i}{m_i})}{m_i}$. Then, for i,

$$\begin{aligned}
u_i(\mathbf{q}) - u_i(\tilde{\mathbf{q}}) &= \big[g\big(\frac{q_i}{m_i}\big) - s_i f(Q)\big] - \big[g\big(\frac{\tilde{q}_i}{\tilde{m}_i}\big) - s_i f(\tilde{Q})\big] \\
&= \big[g\big(\frac{q_i}{m_i}\big) - g\big(\frac{\tilde{q}_i}{m_i}\big)\big] - [s_i f(Q) - s_i f(\tilde{Q})] + \big[g\big(\frac{\tilde{q}_i}{m_i}\big) - g\big(\frac{\tilde{q}_i}{\tilde{m}_i}\big)\big] \\
&= \int_{\tilde{q}_i}^{q_i} \frac{\mathrm{d}}{\mathrm{d}x} g\big(\frac{x}{m_i}\big)\,\mathrm{d}x - \int_{\tilde{Q}}^{Q} s_i f'(x)\,\mathrm{d}x + \big[g\big(\frac{\tilde{q}_i}{m_i}\big) - g\big(\frac{\tilde{q}_i}{\tilde{m}_i}\big)\big] \\
&> (q_i - \tilde{q}_i)B - (Q - \tilde{Q})B = [(q_i - \tilde{q}_i) - (Q - \tilde{Q})]B \\
&\geq 0
\end{aligned}$$

The first inequality is due to the concavity of g, the convexity of f and the fact that $m_i < \tilde{m}_i$. The second inequality is because $q_j \geq \tilde{q}_j$ for all $j \neq i$.

3.(1) Consider the case where $\frac{g'(\frac{q_i}{m_i})}{m_i} > s_i f(Q)$ (so that $q_i = m_i$) and suppose that $\tilde{m}$ is close enough to m_i for $\frac{g'(\frac{q_i}{\tilde{m}_i})}{\tilde{m}_i} > s_i f(Q)$.

Suppose that $\tilde{Q} \leq Q$, then $\tilde{q}_j \geq q_j$ for all $j \neq i$. Hence $\tilde{q}_i \leq q_i = m_i < \tilde{m}_i$. As a consequence, $s_i f(\tilde{Q}) = \frac{g'(\frac{\tilde{q}_i}{\tilde{m}_i})}{\tilde{m}_i} \geq \frac{g'(\frac{q_i}{\tilde{m}_i})}{\tilde{m}_i} > s_i f(Q)$. It implies that $\tilde{Q} > Q$, which is absurd. We have thus proved $\tilde{Q} > Q$. It follows that $\tilde{q}_j \leq q_j$ for all $j \neq i$, and hence $\tilde{q}_i > q_i$. It is obvious that $u_j(\tilde{\mathbf{q}}) < u_j(\tilde{q})$ for all $j \neq i$. For i,

$$\begin{aligned} u_i(\mathbf{q}) - u_i(\tilde{\mathbf{q}}) &= \big[g\big(\frac{q_i}{m_i}\big) - s_i f(Q)\big] - \big[g\big(\frac{\tilde{q}_i}{\tilde{m}_i}\big) - s_i f(\tilde{Q})\big] \\ &= \big[g\big(\frac{q_i}{m_i}\big) - g\big(\frac{\tilde{q}_i}{\tilde{m}_i}\big)\big] + [s_i f(\tilde{Q}) - s_i f(Q)] > 0, \end{aligned}$$

because $\frac{q_i}{m_i} \geq \frac{\tilde{q}_i}{\tilde{m}_i}$ and $\tilde{Q} > Q$.

3.(2) Under Assumption 2, consider the case where $\frac{g'(\frac{q_i}{m_i})}{m_i} > s_i f(Q)$. Let $\hat{m}_i > m_i$ be such that $\frac{g'(\frac{q_i}{\hat{m}_i})}{\hat{m}_i} = s_i f(Q)$. Then it is not difficult to see that $\hat{\mathbf{q}} = \mathbf{q}$, and consequently $\hat{Q} = Q$. It follows immediately the $u_j(\bar{\mathbf{q}}) = u_j(\mathbf{q})$ for all $j \leq i$ and $u_i(\hat{\mathbf{q}}) = g(\frac{q_i}{\hat{m}_i}) - s_i f(Q) < g(\frac{q_i}{m_i}) - s_i f(Q) = u_i(\mathbf{q})$. For the result for all $\tilde{m}_i > \hat{m}_i$, one has only to apply the results in (2) by replacing m_i there by $\hat{m}_i$. □

Proof of Proposition 5. 1. First note that $q_1^{(2)}(s_2)$ is a constant. Indeed, when coalition 1 centralizes and coalition 2 decentralizes, the behaviour of the members of coalition 2 does not change when their sensitivity varies, and the remaining parameters – m_1, s_1 and m_2 – keep the same. Hence the equilibrium behaviour of coalition 1 does not change either. One has

$$\begin{aligned} u_2(s_2) &= g\big(\frac{q_2(s_2)}{m_2}\big) - s_2 f(Q(s_2)), \\ v_2^{(2)}(s_2) &= g(1) - s_2 f(q_1^{(2)}(\underline{s}_2) + m_2), \\ v_2^{(1)}(s_2) &= g\big(\frac{q_2^{(1)}(s_2)}{m_2}\big) - s_2 f(m_1 + q_2^{(1)}(s_2)). \end{aligned}$$

By Lemma 4, $q_1^{(2)}(\underline{s}_2) + m_2 > Q(\underline{s}_2)$, hence constant $D := f(q_1^{(2)}(\underline{s}_2) + m_2) - f(Q(\underline{s}_2))$ is strictly positive. According to Proposition 3, for all $s_2 \geq \underline{s}_2$, $Q(s_2) \leq Q(\underline{s}_2)$. Therefore $u_2(s_2) \geq g(0) - s_2 f(Q(\underline{s}_2))$. Hence

$$u_2(s_2) - v_2^{(2)}(s_2) \geq g(0) - s_2 f(Q(\underline{s}_2)) - g(1) + s_2 f(q_1^{(2)}(\underline{s}_2) + m_2) = g(0) - g(1) + s_2 D.$$

This implies that $u_2(s_2) - v_2^{(2)}(s_2)$ tends to $+\infty$ when s_2 goes to $+\infty$. In particular, there exists $\bar{s}_2$ such that $u_2(s_2) > v_2^{(2)}(s_2)$ for all $s_2 > \bar{s}_2$. Considering Lemma 6(ii), this shows that D is a strictly dominated strategy for coalition 2 for all $s_2 > \bar{s}_2$. Then (C, D) is not an equilibrium of the game.

2. According to Proposition 3, $q_2(s_2)$ and $q_2^{(1)}(s_2)$ are both strictly decreasing in s_2 for $s_2 \geq \frac{g'(1)}{f'(m_1+m_2)}$ (in particular $q_2(s_2) < m_2$ and $q_2^{(1)}(s_2) < m_2$). There are two consequences. Firstly, this means there exists a limit of $q_2(s_2)$ and a limit of $q_2^{(1)}(s_2)$ as s_2 goes to $+\infty$,

which are respectively denoted by $q_{2,l}$ and $q_{2,l}^{(1)}$. Secondly,

$$s_2 f'(q_1(s_2) + q_2(s_2)) = \frac{g'(\frac{q_2(s_2)}{m_2})}{m_2}, \tag{14}$$

and $s_2 f'(m_1 + q_2^{(1)}(s_2)) = g'(\frac{q_2^{(1)}(s_2)}{m_2})/m_2$. Let us show that $q_{2,l} = q_{2,l}^{(1)} = 0$. Otherwise if $q_{2,l} > 0$ then the left hand side of (14) tends to $+\infty$ when s_2 goes to $+\infty$, while the right hand side of (14) has a strictly positive limit $\frac{g'(\frac{q_{2,l}}{m_2})}{m_2}$. This is absurd. Similar arguments apply to $q_{2,l}^{(1)}$. Hence $q_{2,l} = q_{2,l}^{(1)} = 0$ and, consequently, $q_1(s_2)$ tends to $q_{1,l}$, where $q_{1,l}$ is also the optimal level of emissions of coalition 1 if it was the only player.

If $q_{1,l} < m_1$ (hence $g'(1) < s_1 f'(m_1)$), then

$$\lim_{s_2 \to +\infty} u_1(s_2) = g(\frac{q_{1,l}}{m_1}) - s_1 f(q_{1,l}) > g(1) - s_1 f(m_1) = \lim_{s_2 \to +\infty} v_1^{(1)}(s_2).$$

Hence C strictly dominates D for coalition 1 for s_2 large enough. The unique equilibrium of the decentralization game is (C, C).

If $q_{1,l} = m_1$ and $g'(1) > s_1 f'(m_1)$, then for s_2 large enough, $g'(1) > s_1 f'(m_1 + q_2(s_2))$. Therefore $q_1(s_2) = m_1$ and, consequently, coalition 1 is indifferent between playing C and D. Since we consider only the (locally) stable equilibrium, the weakly dominated strategy D is not played so that (C, C) is the unique equilibrium.

If $q_{1,l} = m_1$ and $g'(1) = s_1 f'(m_1)$, then $\lim_{s_2 \to +\infty} u_2(s_2) = \lim_{s_2 \to +\infty} v_2^{(1)}(s_2)$. According to the result of the first part of this proposition, this implies that for s_2 large enough, $v_2^{(1)}(s_2) > v_2^{(2)}(s_2)$. □

Proof of Proposition 6. 1. One has $v_2^{(2)}(m_2) = g(1) - s_2 f(q_1^{(2)}(m_2) + m_2) \leq g(1) - s_2 f(m_2)$, which implies that $v_2^{(2)}(m_2)$ tends to $-\infty$ as m_2 goes to $+\infty$. On the other hand, $u_2(m_2) = g(\frac{q_2(m_2)}{m_2}) - s_2 f(Q(m_2)) > g(0) - s_2 f(Q(\underline{m}_2))$, and $v_2^{(1)}(m_2) = g(\frac{q_2^{(1)}(m_2)}{m_2}) - s_2 f(m_1 + q_2^{(1)}(m_2)) > g(0) - s_2 f(m_1 + q^{(1)}(\underline{m}_2))$, because $Q(m_2) < Q(\underline{m}_2)$ and $q_2^{(1)}(m_2) < q_2^{(1)}(\underline{m}_2)$ according to Proposition 4. These imply that $u_2(m_2)$ and $v_2^{(1)}(m_2)$ have a finite lower bound as m_2 goes to $+\infty$.

2. The proof is similar to that of Proposition 5.2. □

Proof of Corollary 2. According to Proposition 5, if group 2 activates the coalition, it will have the utility at the equilibrium (C, C) for s_2 large enough. If it does not activate the coalition, it will have the utility of an active coalition 2 at (C, D). However, the same proposition also says that the average utility of coalition 2 at (C, D) is strictly lower than at (C, C) for s_2 large enough. □

Proof of Corollary 3. The same is similar to that of Corollary 2. □

Appendix 3: Detailed resolution of the examples

The four examples all follow a generic form for the per-unit utilities

$$u_1 = 3l\big(\frac{q_1}{m_1}\big)^{\frac{1}{2}} - (q_1 + q_2)^{\frac{3}{2}}, \quad u_2 = 3l\big(\frac{q_2}{m_2}\big)^{\frac{1}{2}} - s(q_1 + q_2)^{\frac{3}{2}}$$

When both groups act as centralized coalition, maximizing with respect to q_1 and q_2 respectively yields

$$q_1 + q_2 = \frac{l^2}{m_1 q_1} = \lambda, \quad q_1 + q_2 = \frac{l^2}{s^2 m_2 q_2} = \lambda, \tag{15}$$

from which we obtain

$$q_1 = \frac{l^2}{m_1 \lambda}, \quad q_2 = \frac{l^2}{s^2 m_2 \lambda}, \quad q_1 + q_2 = \frac{l^2}{\lambda}\big(\frac{1}{m_1} + \frac{1}{s^2 m_2}\big). \tag{16}$$

Using $q_1 + q_2 = \lambda$, we deduce $\lambda^2 = l^2(\frac{1}{m_1} + \frac{1}{s^2 m_2})$, i.e.

$$\lambda = l\sqrt{\frac{1}{m_1} + \frac{1}{s^2 m_2}}.$$

If both q_1 and q_2 have interior solutions $\in (0, m_i)$, their equilibrium values are found by plugging the value of λ into (16):

$$q_1^* = \frac{l}{m_1\sqrt{\frac{1}{m_1} + \frac{1}{s^2 m_2}}}, \quad q_2^* = \frac{l}{s^2 m_2\sqrt{\frac{1}{m_1} + \frac{1}{s^2 m_2}}}$$

If for some $j \in \{1, 2\}$, $q_j^* > m_j$, then there is no interior solution and the equilibrium value of q_j is $q_j = m_j$. In this case, the equilibrium emissions of the other player i are computed by taking the maximum emission of j as given. Let $\hat{q}_i$ be a potential interior optimal value of q_i given $q_j = m_j$. The first equation in (15) implies that $\hat{q}_i + m_j = \frac{l^2}{m_1 \hat{q}_1}$ or, equivalently, $m_1 \hat{q}_1{}^2 + m_1 m_2 \hat{q}_1 - l^2 = 0$, which yields the equilibrium value:

$$q_1 = \min\{m_1, \hat{q}_1\}, \text{ where } \hat{q}_1 = \frac{\sqrt{m_2^2 + \frac{4l^2}{m_1}} - m_2}{2}. \tag{17}$$

Similarly,

$$q_2 = \min\{m_2, \hat{q}_2\}, \text{ where } \hat{q}_2 = \frac{\sqrt{m_1^2 + \frac{4l^2}{s^2 m_2}} - m_1}{2}. \tag{18}$$

Now if only group i acts as a centralized coalition, while group j is either a decentralized coalition or a group of countries that has not activated the coalition, then countries in group j free ride so that $q^j = m^j$. In consequence, q^i is given by (17)-(18).

Finally, if both groups are either a decentralized coalition or a group of countries who have not activated the coalition between them, then $q^i = m^i$ for both $i = 1, 2$.

The parameters used in the four examples are as follows:

- **In example 1**, $l = s = m_1 = m_2 = 1$.
- **In example 2**, $l = \frac{1}{3}$ and $s = m_1 = m_2 = 1$.
- **In example 3**, $l = s = 1$, $m_1 = 1.18$ and $m_2 = 0.82$.
- **In example 4**, $l = 1$, $s = 10$; $m_1 = 1.4$ and $m_2 = 0.6$.

References

Barrett, S. (1994). Self-enforcing international environmental agreements. *Oxford Economic Papers*, **46**, 878–894.

— (2005). The theory of international environmental agreements. *Handbook of Environmental Economics*, **3**, 1457–1516.

— (2013). Climate treaties and approaching catastrophes. *Journal of Environmental Economics and Management*, **66** (2), 235–250.

— and Mattei, F. E. E. (1993). *Heterogeneous international environmental agreements.* Centre for Social and Economic Research on the Global Environment.

Baye, M. R., Crocker, K. J. and Ju, J. (1996). Divisionalization, franchising, and divestiture incentives in oligopoly. *The American Economic Review*, **86** (1), 223–236.

Belleflamme, P. (2000). Stable coalition structures with open membership and asymmetric firms. *Games and Economic Behavior*, **30** (1), 1–21.

Bloch, F. (1996). Sequential formation of coalitions in games with externalities and fixed payoff division. *Games and Economic Behavior*, **14** (1), 90–123.

Buchholz, W., Cornes, R. and Rübbelke, D. (2014). Potentially harmful international cooperation on global public good provision. *Economica*, **81** (322), 205–223.

Carraro, C., Eyckmans, J. and Finus, M. (2006). Optimal transfers and participation decisions in international environmental agreements. *The Review of International Organizations*, **1** (4), 379–396.

— and Siniscalco, D. (1993). Strategies for the international protection of the environment. *Journal of Public Economics*, **52** (3), 309–328.

d'Aspremont, C., Jacquemin, A., Gabszewicz, J. J. and Weymark, J. A. (1983). On the stability of collusive price leadership. *Canadian Journal of Economics*, **16** (1), 17–25.

Eckert, H. (2003). Negotiating environmental agreements: Regional or federal authority? *Journal of Environmental Economics and Management*, **46** (1), 1–24.

Foucart, R. and Garsous, G. (2013). Threshold effects in self-enforcing international environmental agreements. *ECARES Working Papers*.

Gordon, P. H. (1997). Europe's uncommon foreign policy. *International Security*, **22** (3), 74–100.

Howorth, J. (2001). European defence and the changing politics of the European Union: hanging together or hanging separately? *JCMS: Journal of Common Market Studies*, **39** (4), 765–789.

Joskow, P. L., Schmalensee, R. and Bailey, E. M. (1998). The market for sulfur dioxide emissions. *The American Economic Review*, **88** (4), 669–685.

Kaldor, M. and Salmon, A. (2006). Military force and European strategy. *Survival*, **48** (1), 19–34.

Kirchner, E. J. (2006). The challenge of European Union security governance. *JCMS: Journal of Common Market Studies*, **44** (5), 947–968.

Kosfeld, M., Okada, A. and Riedl, A. (2009). Institution formation in public goods games. *The American Economic Review*, **99** (4), 1335–1355.

Lowe, P., Buller, H. and Ward, N. (2002). Setting the next agenda? British and French approaches to the second pillar of the Common Agricultural Policy. *Journal of Rural Studies*, **18** (1), 1–17.

Lutsey, N. and Sperling, D. (2008). America's bottom-up climate change mitigation policy. *Energy Policy*, **36** (2), 673–685.

McGinty, M. (2007). International environmental agreements among asymmetric nations. *Oxford Economic Papers*, **59** (1), 45–62.

Menon, A. (2011). European defence policy from lisbon to libya. *Survival*, **53** (3), 75–90.

Oates, W. E. (2005). Toward a second-generation theory of fiscal federalism. *International tax and public finance*, **12** (4), 349–373.

Olson, M. and Zeckhauser, R. (1966). An economic theory of alliances. *The Review of Economics and Statistics*, **48** (3), 266–279.

Pant, H. V. (2013). The BRICS fallacy. *The Washington Quarterly*, **36** (3), 91–105.

Ray, D. and Vohra, R. (1997). Equilibrium binding agreements. *Journal of Economic Theory*, **73** (1), 30–78.

— and — (2001). Coalitional power and public goods. *Journal of Political Economy*, **109** (6), 1355–1385.

Rodrik, D. (2013). What the world needs from the BRICS. *Project Syndicate*, **April, 10**.

Salant, S. W., Switzer, S. and Reynolds, R. J. (1983). Losses from horizontal merger: the effects of an exogenous change in industry structure on cournot-nash equilibrium. *The Quarterly Journal of Economics*, **98** (2), 185–199.

Sorin, S. and Wan, C. (2013). Delegation equilibrium payoffs in integer-splitting games. *RAIRO-Operations Research*, **47** (1), 47–58.

Yi, S.-S. (1997). Stable coalition structures with externalities. *Games and Economic Behavior*, **20** (2), 201–237.