Groupe de Recherche en
Économie Théorique et Appliquée

Multilateral versus sequential negotiations over climate change

Alejandro Caparrós

Consejo Superior de Investigaciones Científicas (CSIC)
Institute for Public Goods and Policies (IPP)
Albasanz 26, 28037 Madrid, Spain
E-mail: alejandro.caparros@csic.es

&

Jean-Christophe Péreau

GREThA, CNRS, UMR 5113
University of Bordeaux
Avenue Léon Duguit, 33608 Pessac cedex, France
E-mail: jean-christophe.pereau@u-bordeaux.fr

Cahiers du GREThA
n° 2015-34
Novembre

GRETHA UMR CNRS 5113
Université de Bordeaux
Avenue Léon Duguit - 33608 PESSAC - FRANCE
Tel : +33 (0)5.56.84.25.75 - Fax : +33 (0)5.56.84.86.47 - www.gretha.fr

Négociations séquentielles ou multilatérales dans la lutte contre le changement climatique

Résumé

Cet article propose un modèle de formation endogène de coalition avec externalités positives dans lequel un pays dit leader décide si les négociations doivent être menées de manière séquentielle ou multilatérale dans la lutte contre le changement climatique. Les résultats montrent que le choix d'une négociation séquentielle dépend de la propriété de convexité du jeu à utilité transférable et des gains de déviation des pays suiveurs. Excepté des cas clairement identifiés, la négociation conduira toujours à la formation de la grande coalition mais le processus pourra être graduel. Ce résultat est validé dans le modèle standard de recherche d'un accord environnemental international en présence d'agents hétérogènes même quand la grande coalition n'est pas stable dans un cadre multilatéral. L'article analyse aussi le rôle d'une agence qui pourrait inciter à accélérer les négociations intra-round et accroître le délai entre les rounds de négociations dans un processus séquentiel.

Mots-clés : négociation multilatérale, formation endogène de coalition, négociations internationales, médiateur, accords environnementaux internationaux.

Multilateral versus sequential negotiations over climate change

Abstract

We discuss a model of gradual coalition formation with positive externalities in which a leading country endogenously decides whether to negotiate multilaterally or sequentially over climate change. We show that the leader may choose a sequential path, and that the choice is determined by the convexity of the TU-game and the free-rider payoffs of the followers. Except in a few clearly defined cases, the outcome of the negotiation process is always the grand coalition, although the process may need some time. This holds for the standard IEA game with heterogeneous players even if the grand coalition is not stable in a multilateral context. We also analyze the role of a facilitating agency. The agency has an incentive to speed up intra-stage negotiations and to extend the period between negotiation stages in a sequential process.

Keywords: multilateral bargaining, endogenous coalition formation, international negotiations, mediator, international environmental agreements.

JEL: C78, Q54

Reference to this paper: CAPARROS Alejandro, PEREAU Jean-Christophe (2015) *Multilateral versus sequential negotiations over climate change*, *Cahiers du GREThA*, n°2015-34.
http://ideas.repec.org/p/grt/wpegrt/2015-34.html.

GRETHA UMR CNRS 5113
Université de Bordeaux
Avenue Léon Duguit - 33608 PESSAC - FRANCE
Tel : +33 (0)5.56.84.25.75 - Fax : +33 (0)5.56.84.86.47 - www.gretha.fr

1 Introduction

Climate change negotiations have favored a multilateral approach over the last two decades that has yielded only a limited success. Given these difficulties, Benedick (2007) questioned already some years ago whether this was a good strategy for climate change. In fact, as the author points out, the main success story in international environmental agreements, the Montreal Protocol, was adopted following a sequential path[1] (Benedick, 1998). This issue resembles the question of whether regionalism or multilateralism is the most effective strategy for achieving global free trade, which has been debated extensively in international economics (Bhagwati, 1993; Yi, 1996). As noted by Aghion *et al.* (2007), another way of posing this question is to ask whether multilateral or sequential bargaining[2] is more likely to lead to a global agreement. Aghion *et al.* (2007) focus on Free Trade Agreements (FTA) but state that their results may also be applied to International Environmental Agreements (IEA), as their model is based on a partition function approach[3]. The authors propose a model where one leader country endogenously chooses whether to negotiate sequentially or multilaterally, assuming that the leader makes only take-it-or-leave-it offers. The key message is that the grand coalition is always formed and that the leader prefers multilateral negotiations if the game has positive coalition externalities. For the FTA game, positive and negative externalities may arise, the latter mainly due to political economy reasons, but for the IEA game externalities are clearly positive (Barrett, 1994; Carraro and Siniscalco, 1993; Finus and Caparrós, 2015). Hence, their model predicts that IEA will always be negotiated multilaterally.

Assuming that a country plays a leader role in international negotiations is a reasonable assumption not only for trade negotiations but also for environmental negotiations. During the negotiations of the Montreal Protocol the US played a clear leader role (Benedick, 1998; Sandler, 2004), and in climate change negotiations both, the US and the EU, have played this role at different times[4]. Nevertheless, it is most likely too strong an assumption that the

[1] The first efforts to protect the ozono layer were an informal accord between six nations, the US, Australia, Canada, Norway and Sweden to ban CFC from aerosol spray cans. Then, under the leadership of the US, the Montreal Protocol was negotiated in about nine months between 30 countries. The coverage of the Protocol grew gradually and is now universal. For details, see Benedick (1998).

[2] Note that the Kyoto Protocol may also be seen as a first stage in a sequential bargaining process, as only some of the world's countries have targets under this Protocol (Carraro, 2005).

[3] By the same token, our analysis is relevant for FTA as well.

[4] The US played a prominent role during the negotiations of the Kyoto Protocol, until it lost this position by not ratifying the Protocol. After the withdrawal of the US, the EU took the leader role until the Protocol was ratified. Once the Kyoto Protocol entered into force, negotiations were resumed to reach a post-Kyoto agreement. In the negotiation round that took place in Copenhagen in 2009, the US recovered its leader role by reaching a partial agreement with a handful of emerging economies. This agreement was subsequently adopted by the remaining countries and eventually led to the modest Cancun Agreements. For Sandler (2004:

leader only makes take-it-or-leave-it offers, as in Aghion *et al.* (2007). The implication is that the entire surplus generated by the agreement is appropriated by the leader (the US in their discussion), while the other countries of the world are indifferent between facing unrestricted climate change and cooperating. We therefore propose a dynamic model that fully specifies the bargaining process and in which a leading country endogenously decides whether to negotiate multilaterally or sequentially (it is the agenda-setter), but in which distribution of the surplus is endogenously determined by the bargaining protocol followed. In each stage, we assume that bargaining takes place à la Rubinstein (1982) between the different players involved (building on Huang (2002), when three or more players bargain). Changing this assumption has profound implications on the results obtained, showing that sequential bargaining may be chosen with positive externalities and that the choice is determined by the convexity of the game.

The second difference between our model and Aghion's et al. is that we introduce discounting[5], a crucial feature to model climate change negotiations as delaying an agreement has costs. As sequential bargaining implies welfare losses once discounting is introduced, even in equilibrium, we postulate that there is an international agency (e.g. the UNFCCC Secretariat) that can intervene in the negotiations. This mediator or facilitating agency does not have the power to impose an agreement. Instead, we assume that, through lobbying, the organization of international conferences and meetings, or other means, the facilitating agency only has the power to influence the amount of time that elapses between negotiation stages or the amount of time that elapses between offers in the intra-stage negotiation. Our results show the circumstances under which such mediator has a role to play, and that it will generally have an incentive to separate the negotiation rounds (or stages in our model) to induce the leader to choose multilateral bargaining.

Our benchmark model assumes that all players terminate an agreement if one country breaks it. This assumption is responsible for the result that the grand coalition is achieved, regardless of the path followed. The same assumption was implicit in Aghion's et al. (2007) and has been applied to climate change negotiations by Caparrós et al. (2004) or Harstad (2012). However, this is not the standard assumption in IEA analyses, which generally look for "self-enforcing" agreements[6] (Barrett, 1994). IEA analyses generally use the concept of

224), this lack of a clear leader explains the difficulties encountered in reaching an effective agreement.

[5] As in any bargaining model, the discount factor summarizes all reasons why a deal is more valuable the sooner it is reached (Muthoo, 1999). The reasons can be impatience, haggling costs, or that a climate threshold may be reached during the negotiations if they continue.

[6] Barrett (1994) popularized in the IEA literature the term "self-enforcing" to refer to agreements reached under the assumption that countries will continue cooperating when one country defects. However, this may be confusing to some readers as the term "self-enforcing" was also popularized by Telser (1980) in the context of repeated interactions, and for Telser a self-enforcing agreement implies that "if one party takes an

internal and external stability (d'Aspremont, 1983). This solution concept assumes that all countries believe that if one country leaves a coalition the remaining countries will continue cooperating, which is most likely not so far fetched for international negotiations. Countries can certainly not assume that international agreements are completely binding because enforcement at the international level is rather weak (although the assumption that international agreements are not binding at all is also not satisfactory). In addition, Parties to a treaty tend to continue cooperating even if one country fails to do so, at least if the number of countries involved is large[7]. Real-life climate change negotiations provide examples for the IEA game (although Canada announced that it was not meeting its commitment under the Kyoto Protocol, the European Union continued its efforts to meet its targets). We show that if the free-riding incentives are smaller for the leader than for the followers, sequential bargaining will be chosen more often under this alternative assumption. Furthermore, we show that a stable grand-coalition may be achieved following a sequential path if coalitions *consolidate*[8] and superadditivity holds, and that if the grand coalition is not stable in a multilateral setting, the sequential approach may, in fact, be the only way to reach a stable global agreement.

As already mentioned, the model presented here is at the intersection of two strands of literature. On one side is the literature on FTA. Aghion et al. (2007) is the closest precedent, but there are a large number of papers on this issue. Sen and Biswas (2015) extend Aghion's et al. analysis to different protocols for the multilateral option, keeping the assumption of take-it-or-leave-it offers and the absence of discounting. Using a political economy approach, Levy (1997) and Krishna (1998) show that bilateral FTAs can undermine political support for further multilateral trade liberalization. On the contrary, Saggi and Yildiz (2010 and 2011) show that bilateral FTAs can be necessary to achieve global free trade in a three player model (they focus on the impact of banning bilateral agreements). Other analyses of the FTA game can be found in Yi (1996), who analyzes a game with identical players and particular functions, showing that the grand coalition is achieved under the open membership rule, or in Macho-Stadler and Xue (2007), who use a three player model to discuss the winners and losers in a sequential coalition formation game.

action... [that]... constitutes a violation of the agreement, then the theory assumes that the victim responds by terminating the agreement". Note that this resembles the assumption in our benchmark model.

[7] If there are only three countries it is hard to believe that if one country leaves an agreement the other two will continue cooperating under the terms of the old agreement.

[8] Consolidation of coalitions means an intermediate coalition cannot break up in subsequent negotiations. In other words, we assume that, if the EU and the US form a coalition first, they will develop an integrated emissions trading scheme that makes it impossible for one of them to break up a subsequent deal with China. As discussed in footnote 18, this is not the same than the assumption that one player commits in Carraro and Siniscalco (1993).

On the other side, there is also an extensive literature applying game theoretic concepts to IEA analysis (see Finus and Caparrós (2015) for a survey). Most of these papers assume identical players and model coalition formation as a simultaneous move game (although there is typically a second stage where emissions are determined). The main result is that only a reduced number of countries cooperate (Barrett, 1994). Closer related to our paper, Asheim et al. (2006), and previously Carraro and Siniscalco (1998), among others, have shown that social welfare may be higher with multiple agreements than with a single global agreement. Asheim et al. (2006) propose a dynamic framework with homogeneous players (they also provide an overview of the different papers addressing this issue). However, their focus is different to ours, as they do not allow the two regional coalitions to reach a new agreement to enlarge the scope of cooperation. Examples of sequential games can be found in Caparrós et al. (2004) and Finus et al. (2010). The former is a three player game focused on asymmetric information and the latter analyzes a sequential coalition formation game with heterogeneous players using simulations and an Integrated Assessment Model. Finus et al.'s model is based on the sequential move unanimity game (Bloch, 1995). In this game (based in turn on Chatterjee et al. (1993)), an initiator proposes a coalition and if all potential members agree the coalition is formed. Then, a new proposer is selected from the players not participating in the coalition. By contrast, in our model it is the coalition which negotiates with the remaining players to enlarge the coalition. This implies that only one coalition forms in equilibrium, which is more in line with the empirical finding that all IEA represent single agreements that are gradually extended to cover more countries. Examples where coverage, both in terms of effort and in terms of countries, evolves over time are the agreements on ozone depleting substances (the Montreal Protocol and related treaties), on sulphur reduction (the Helsinki Protocol and related treaties), or on nitrogen oxide (the Sophia Protocol and related treaties).

2 The model

2.1 Benchmark model

Building on Aghion *et al.* (2007), we start by considering a transferable-utility game in a world with three countries: a, b, and c. The game is described in partition form and we define a coalition structure as a partition Γ of $\{a, b, c\}$. A coalition is a group of countries that have agreed to form an IEA. For every partition Γ and every player j (singleton or coalition $C \in \Gamma$), the value function $v(j; \Gamma)$ assigns a payoff to j given the coalition structure Γ (this payoff is gross of lump-sum transfers).

One country is the leader, which means that it is the agenda-setter (without loss of generality, we assign this role to country c). In the first stage of the game, the leader chooses between multilateral and sequential bargaining. As we are looking for a subgame perfect equilibrium (SPE), country c chooses the path that maximizes its payoff.

If the leader goes for sequential bargaining, it has to decide whether to bargain first with a or with b (we discuss in detail only the case where it approaches a first, as the other subgame is symmetric). One of the differences with Aghion et al. (2007) is that we substitute take-it-or-leave-it offers for an alternate offers bargaining framework. If the leader negotiates first with a the offer consists of a coalition between a and c and lump-sum transfers from c to a a. Country a can accept the offer, or reject it and make a counter-offer to c, which may be accepted by c or rejected with a new counter-offer. The alternative-offers protocol à la Rubinstein follows until an agreement is reached, yielding a payoff of $P(i, \Gamma_{ac}, s_a)$ for $i = a, c$ if the coalition is not expanded further, where s_a indicates that the sequential process started with a, and $\Gamma_{ac} = \langle\{ac\}, \{b\}\rangle$ indicates that a and c form a coalition (where there is no risk of confusion, we write P_i instead of $P(i, \Gamma_{ac}, s_a)$). In case of perpetual disagreement[9] between a and c, the coalition structure is $\Gamma_\phi = \langle\{a\}, \{b\}, \{c\}\rangle$. In this intra-stage[10] negotiations, offers are made at discrete points in time, and the duration of the period between offers is $\tau > 0$. Thus, the discount factor in the intra-stage negotiations is $\delta = e^{-r\tau t} < 1$, where $r > 0$ is the discount rate and $t = \{0, 1, 2, ...\}$ are the periods when the offers are made.

Whenever a and c reach an agreement, the coalition $\{ac\}$ proceeds to bargain with b. However, after the first negotiation (between a and c) and before the second negotiation (between $\{ac\}$ and b) starts, a period of duration $\theta > 0$ elapses. Thus, the inter-stage discount factor is $\sigma = e^{-r\theta t} < 1$ (that is we consider two discount factors, although there is only one discount rate). The new negotiation then starts with an offer from $\{ac\}$ to b, offering an expansion of the coalition to include all three countries and lump-sum transfers from $\{ac\}$ to b. If b accepts the offer, the coalition structure is $\Gamma_{abc} = \langle\{abc\}\rangle$, and each player i receives a payoff equal to $P(i, \Gamma_{abc}, s_a)$. If b rejects the offer, it proposes a counter-offer to $\{ac\}$, which may be accepted or rejected with a new counter-offer, and so on. In case of perpetual disagreement between b and $\{ac\}$, the coalition structure is $\Gamma_{ac} = \langle\{ac\}, \{b\}\rangle$.

If c chooses multilateral bargaining, it bargains simultaneously with both follower countries. Thus, country c makes an offer that consists of a coalition including all countries and a system of lump-sum transfers. The transfers determine the payoffs $P(i, \Gamma, m)$ of countries

[9] That there are no additional negotiations if a and c reach no agreement may seem a strong assumption, but note that under perfect information disagreement is never the outcome in a negotiation à la Rubinstein.

[10] *Intra-stage* refers the bargaining process (and the discount factor) that takes place *within* each of the stages of the sequential negotiation process, or within the unique stage of the multilateral process. *Inter-stage* refers to the period *between* two negotiation stages.

$i \in c, a, b$ (where m stands for multilateral). Following Huang (2002), we allow for conditional and unconditional offers. A conditional offer only binds if both countries accept the offer, while an unconditional offer only binds between the proposer and the country accepting the offer[11]. In the latter case, the proposer buys out the right of the accepting country and continues negotiating with the other country (but within the same negotiation stage). Although we present a more general model in Section 3, for the time being we assume that countries talk following a cyclical protocol c, a, b, c, a, b .... Thus, country a talks second and decides to accept the offer or to make a counter-offer, then country b talks and so on. If an offer is finally accepted by both countries, $\Gamma = \langle\{abc\}\rangle$ is the resulting coalition structure and the game ends. In the event of perpetual disagreement, where no offer is accepted by all countries, the coalition structure is $\Gamma = \langle\{a\}, \{b\}, \{c\}\rangle$.

The bilateral bargaining process described in the sequential case above is unproblematic because uniqueness is ensured by Rubinstein's alternative-offers procedure (Rubinstein, 1982). However, in the multilateral bargaining case requiring unanimity may yield multiple equilibria, as shown by Shaked (reported by Sutton (1986)), if we do not wish to focus solely on the case where the history of the negotiations has no impact. There are different proposals to restoring uniqueness, such as the "exit" game proposed by Krishna and Serrano (1996) or the possibility of making unconditional offers financed by outside money (Huang, 2002), which we have decided to follow[12]. Huang's solution has the following welcome features: (i) it degenerates to the Rubinstein Bargaining Solution (RBS) (Rubinstein , 1982) for the case of two players; (ii) it tends to the Nash Bargaining Solution (NBS) if the discount factor tends to one; and (iii) if all players are allowed to talk once in the cycle, the equilibrium is the same as the unique Stationary SPE of the unanimity game (Osborne and Rubinstein, 1990). It is therefore an adequate generalization of the RBS to more than two players (focusing directly on stationary strategies yields similar results, but this is a strong assumption (Osborne and Rubinstein, 1990)). One can see Huang's proposal only as a refinement to select the most reasonable equilibria in multilateral bargaining, but allowing for conditional offers is also an adequate manner of modeling what actually takes place in international negotiations, as the number of players finally involved in the key negotiations in a meeting is generally small (with each negotiator representing the interests and positions of a large set of countries). However, one has to accept that the outside money needed to buy out

[11] Note that this distinction is not possible in the set of bilateral agreements that take place in the sequential path.

[12] In fact, Huang (2002) also shows that the multiplicity of equilibria in multilateral bargaining appears if only conditional offers are possible. Krishna and Serrano (1996) already showed that if only unconditional offers are possible (the exit game), unicity is restored. Huang's proposal has the advantage of allowing for both types of offers.

the right to negotiate for another country cannot be interpreted literally, and it should be taken as a compromise to accept the outcome that the other party will negotiate (maybe in exchange for a reciprocal deal in another negotiation). Finally, Huang's proposal is more adequate for our model than Gomes' (2005), as the latter selects the first proposer randomly and we want to assign this role to the leader.

Before moving on to the results, let us introduce additional notation to allow for a more compact presentation. The additional surplus generated by forming an intermediate coalition between countries i and j is denoted by $\Delta_{ij} = v(ij;\Gamma_{ij}) - v(i;\Gamma_\phi) - v(j;\Gamma_\phi)$. The surplus generated by forming the grand coalition is denoted $\Delta_{abc} = v(abc;\Gamma_{abc}) - \sum_{i=a,b,c} v(i;\Gamma_\phi)$ when it was formed starting from the all singletons situation and $\Delta_{abc-ij} = v(abc;\Gamma_{abc}) - v(ij;\Gamma_{ij}) - v(k;\Gamma_{ij})$ when it was formed enlarging a coalition formed by countries i and j (to allow for more compact results we write $\Delta_{-ij} = \Delta_{abc-ij}$).

For future reference, we provide the following definitions.[13]

Definition 1 *Positive Coalition externalities*[14]*. There are positive coalition externalities in country j when $E(j) = v(j;\Gamma_{kl}) - v(j;\Gamma_\phi) > 0$.*

Given our focus on climate change negotiations, we assume throughout the paper that the game has positive coalition externalities.

Definition 2 *Cohesiveness and superadditivity. The game is cohesive if $\Delta_{abc} > 0$ and $\Delta_{-jc} > 0$, for all $j \neq c$. The game is superadditive if, in addition, $\Delta_{jc} > 0$, for all $j \neq c$.*

That is, cohesiveness (called GC superadditivity in Aghion *et al.* (2007)) requires the joint payoffs of all three countries to be larger when they act together, compared with under no agreement or a partial agreement (for general definitions see, e.g., Osborne and Rubinstein (1994)). Superadditivity also requires that a partial agreement increases the payoffs for the members of the agreement. As discussed below, the models with particular functions usually used to analyze IEA exhibit cohesiveness, but not always superadditivity. In fact, there are no reasons to assume that cohesiveness fails in the IEA game, and we will therefore rule this possibility out[15].

Finally, we define convex[16]/concave games for our three player case as follows (for a general definition, see Shapley (1971)).

[13] For simplicity, Definitions 2 and 3 are only valid for games where c is involved in the coalitions.

[14] This definition covers cases where non-members only partially benefit from the formation of the coalition and cases where non-members obtain the same increase in benefits as members (pure public goods).

[15] The literature on FTA (Aghion et al, 2007; Saggi and Yildiz 2010 and 2011) has considered the possibility that cohesiveness does not hold. We consider this possibility in Appendix B, showing that, as in the papers just mentioned, both stumbling and building block equilibria are possible.

[16] Convex games are supermodular.

Definition 3 *Convex/concave game. A game is convex if $\Delta_{-kc} > \Delta_{jc}$ and concave if the opposite relation holds, for all $j \neq k$.*

This definition shows that a game is convex if the marginal contribution of country j is not decreasing when it joins a coalition of a greater size. It is easy to show that a convex game is superadditive, although the converse is not necessarily true.

Applying the equilibrium concept developed in Huang (2002), with some modifications and extensions discussed in the Appendix, we obtain the following lemma:

Lemma 1 *The payoff for the leader in the unique equilibrium outcome is:*
(i) if the leader follows a multilateral path:

$$P(c, \Gamma_{abc}, m) = v(c; \Gamma_{\phi}) + \frac{\Delta_{abc}}{\left(1 + \delta + \delta^2\right)}, \tag{1}$$

(ii) if the leader follows a sequential path:

$$P(c, \Gamma_{abc}, s) = v(c; \Gamma_{\phi}) + \max\left(\frac{\Delta_{ac}}{(1+\delta)} + \frac{\sigma\Delta_{-ac}}{(1+\delta)^2}, \frac{\Delta_{bc}}{(1+\delta)} + \frac{\sigma\Delta_{-bc}}{(1+\delta)^2}\right). \tag{2}$$

Proof: Appendix A.1.

Note that if the intra-stage discount factor δ tends to 1, the leader obtains in multilateral bargaining its disagreement payoff plus one third of the additional surplus created if a deal is struck. As the discount factor δ decreases, the share obtained by the leader increases, as the first mover advantage becomes more important.

If the leader follows the sequential path, it has to decide which player to approach first. As should be expected, the leader chooses the option that maximizes the additional surplus that it gets, in addition to its disagreement payoff (see equation 2). However, to ease the exposition we assume henceforth that the leader approaches country a first in the eventuality of choosing sequential bargaining (because we can rename the countries at will, this implies no loss of generality). When the leader negotiates first with a, it obtains a share $\frac{1}{(1+\delta)}$ from the surplus, Δ_{ac}, created by this negotiation. From the second negotiation the leader obtains $\frac{\sigma\Delta_{-ac}}{(1+\delta)^2}$. That is, as the agreement will be obtained in the second negotiation stage the payoff secured has to be discounted using the inter-stage discount factor σ. From the additional surplus create by this second agreement, Δ_{-ac}, a share equal to $\frac{1}{(1+\delta)}$ is obtained by the coalition $\{ac\}$, and out of this the leader obtains a share equal to $\frac{1}{(1+\delta)}$ (this share was decided during the negotiation in the first stage).

To facilitate the comparison with Aghion's et al (2007) results, let us first assume that there is no discounting, either within one negotiation stage or between the different negotiation stages in a sequential coalition formation process. Note that, without discounting, assuming that the leader approaches a first implies that

$$\Delta_{ac} + E(a) > \Delta_{bc} + E(b). \tag{3}$$

Hence, the leader is interested in approaching first the country to which it adds more in a partial coalition and which has larger positive externalities.

We obtain the following proposition:

Proposition 1 *Assume that the game is cohesive and that the discount factors tend to one. Then, the grand coalition is formed in all cases and the leader prefers multilateral bargaining iff*

$$[\Delta_{-ac} - 2\Delta_{ac}] + 4E(b) > 0 \tag{4}$$

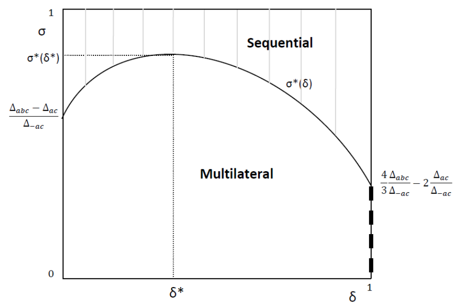
and sequential bargaining iff the opposite relation holds. Thus, convexity of the game and positive coalition externalities in the follower countries favor multilateral bargaining.

Proof: Appendix A.2.

From equation (4) it is clear that only if $\Delta_{-ac} > 2\Delta_{ac}$, when the surplus generated by moving from the intermediate coalition to the GC is twice as large as the one generated by forming the coalition $\{ac\}$, is it possible to determine directly what the leader prefers (as $E(b)$ is always non-negative with positive externalities). The intuition for this result comes from the fact that, when discount factors tend to one, in a multilateral negotiation the leader obtains one third of the total additional surplus, while in a sequential negotiation the leader obtains one half of the surplus created wit the first agreement, Δ_{ac}, and one fourth of the surplus created with the second agreement, Δ_{-ac}. Thus, if the surplus of a three player coalition is clearly larger than that of a two player coalition, the leader prefers multilateral bargaining (to obtain one third instead of one fourth of this additional surplus).

To discuss the role of convexity, let us focus on a case without externalities. Then, as we have assumed that c approaches a first, we know that $\Delta_{ac} > \Delta_{bc}$ and therefore convexity becomes a necessary condition for multilateral bargaining (because a concave game implies $\Delta_{bc} > \Delta_{-ac}$ and the first term in equation 4 becomes negative). Finally, positive coalition externalities in country b, $E(b)$ in the second term in equation 4, favor multilateral bargaining because the larger these externalities are the larger is the disagreement point of player b in the second negotiation of a sequential process, and the lower the payoff obtained by the leader in this second negotiation (inducing it to adopt a multilateral approach).

Aghion et al. (2007) found that positive coalition externalities imply that the leader prefers multilateral bargaining. We also find that positive coalition externalities favor multilateral bargaining. Nevertheless, the role of externalities is not as prominent in our framework, as the leader may prefer sequential bargaining even with positive externalities as long



as $\Delta_{-ac} < 2\Delta_{ac} - 4E(b)$. That is, when player b adds relatively little in the second negotiation round (this is perfectly possible as shown in section 4 below).

For the general case where the discount factors do not tend to one, we obtain the following proposition:

Proposition 2 *Assume that the game is cohesive. The grand coalitions is always formed and, for a given δ, the leader chooses multilateral bargaining iff $\sigma \leq \sigma^*(\delta)$ and sequential bargaining iff $\sigma > \sigma^*(\delta)$, with*

$$\sigma^*(\delta) = \frac{(1+\delta)^2}{1+\delta+\delta^2}\frac{\Delta_{abc}}{\Delta_{-ac}} - (1+\delta)\frac{\Delta_{ac}}{\Delta_{-ac}} \tag{5}$$

The function $\sigma^*(\delta)$ is concave and is maximized for δ^*, with $0 < \delta^* < 1$.

Proof: Appendix A.3.

Although the interpretation is more difficult once discounting has been introduced, Proposition 2 shows that the relation between Δ_{ac} and Δ_{-ac} remains key in defining the set of discount factors for which the leader chooses multilateral bargaining (although the relationship between Δ_{abc} and Δ_{-ac} also impacts the outcome). As Figure 1 shows, any pair of discount factors (δ, σ) below the solid line induce the leader to choose a multilateral bargaining process, while any pair above this line induces a sequential process. Hence, both negotiation processes can be chosen by the leader for a given valuation function, depending on the discount factors.

The intra-stage discount factor δ only affects the distribution of the surplus to be shared within the round, as in equilibrium the first offer made by the leader is immediately accepted. In other words, as in any complete information analysis based on the RBS (or Huang's extension), if the leader chooses multilateral bargaining there is no delay and no "wasted" surplus. Nevertheless, if the leader decides to follow the sequential path, there is indeed a welfare loss given by the fact that the global agreement will not be achieved immediately (i.e., once the assumption of no discounting between stages is relaxed, global welfare is reduced if the leader chooses sequential bargaining). Given this potential global welfare loss, we now postulate that there is an international agency (e.g. the UNFCCC Secretariat) that can intervene in the negotiations. As discussed in the introduction, this *mediator* or *facilitating agency* has only the power to influence the amount of time that elapses between negotiation stages and/or the amount of time that elapses between offers in the intra-stage negotiation. That is, the agency can influence both discount factors, δ and σ, by modifying the duration of the periods in the intra-stage negotiation and the duration of the period between rounds (i.e., the agency modifies τ and θ and thus the discount factors, although the discount rate remains obviously the same). In modeling terms, this implies having an additional stage before the game starts where the facilitating agency chooses the discount factors δ and σ.

Adapting the utility function of the mediator in Camiña and Porteiro (2009) to our game, we obtain the following definition.

Definition 4 *Given a generic vector of payoffs for players a, b and c, (P_a, P_b, P_c), the preferences of an α-type mediator, with $\alpha \in \mathbb{R}_{\geq 0}$, are represented by the following utility function*

$$U(P_a, P_b, P_c) = \sum_{i \in a,b,c} (P_i - v(i; \Gamma_\phi)) - \alpha \sum_{i,j \in a,b,c} |(P_i - v(i; \Gamma_\phi)) - (P_j - v(j; \Gamma_\phi))|$$

That is, the mediator positively values the additional surplus created compared with the initial situation and negatively values the unequal distribution of this additional surplus. The relevance of the latter component of the utility function is larger the larger α becomes. We can now write:

Proposition 3 *Assume that the game is cohesive.*
(i) If $\sigma^*(\delta) > 1 \ \ \forall \delta$, then multilateral bargaining is always the outcome, and if $\sigma^*(\delta) < 0$ $\forall \delta$ then sequential bargaining is always the outcome (i.e., the mediator cannot influence the outcome).
(ii) If $0 \leq \sigma^*(\delta) \leq 1$ holds for some values of δ:
- a purely efficiency seeking mediator, $\alpha = 0$, favors multilateral bargaining and sets any δ for which $\sigma^*(\delta) > 0$ and any σ such that $\sigma \leq \min\{1, \sigma^*(\delta)\}$.
- any mediator with $\alpha \in (0, +\infty]$ favors multilateral bargaining and sets $\delta \to 1$ and any σ for which $\sigma \leq \min\{1, \sigma^*(\delta)\}$.

Proof: Appendix A.4.

As the proposition shows, there are situations where the agency has no role because the negotiation will always, or never, be multilateral. In the intermediate case where the choice of the discount factors has an impact on the outcome, as soon as the mediator is only marginally interested in equity it will favor multilateral bargaining and set δ as large as possible, as this will imply that all three players obtain an equal share of the additional surplus $v(abc; \Gamma_{abc})$ (any sequential outcome is less egalitarian). That is, the mediator has an interest to speed up negotiations within each stage (i.e. increase δ). However, it is also interested in delaying the start of a new negotiation round (i.e. reduce σ). In fact, one of the optimal equilibria is to set $\sigma = 0$, which implies to block any possibility to negotiate in future rounds. In Figure 1, the set of optimal pairs of discount factors for the leader are those on the bold dashed line.

On the other hand, a purely efficiency seeking mediator, $\alpha = 0$, is also interested in multilateral bargaining, as this implies distributing $v(abc; \Gamma_{abc})$, which is the largest gross payoff if the game is cohesive, without delay. However, assuming perfect information and a subgame perfect equilibrium, the mediator has no interest in modifying δ and should simply try to keep σ equal or lower than the level $\sigma^*(\delta)$ defined in Lemma (2).

2.2 Internally stable agreements

Our benchmark model has assumed that all players terminate an agreement if one country breaks it (as in Aghion et al. (2007), or in Caparrós et al. (2004) and Harstad (2012) in the context of climate change negotiations). This "benchmark" assumption explains that a global agreement[17] is reached in all cases. However, most of the literature on the IEA game is based on the concept of internal and external stability (d'Aspremont *et al.*, 1983). As already noted, this solution concept assumes that if one country leaves a coalition the remaining countries will continue cooperating.

The transfers implicit in the surplus sharing agreement shown in the previous sections are sub-game perfect, and a country would therefore never break an equilibrium agreement if it assumes that the remaining countries would terminate it; nevertheless, if a country assumes that the remaining countries will continue cooperating even if it breaks the multilateral agreement, it may find it beneficial to leave the coalition. To prevent this possibility, researchers focusing on the IEA game have proposed a variety of transfer schemes that allow Potentially Internally Stable (PIS) coalitions to effectively become internally stable. A PIS coalition has sufficient surplus to grant each coalition member the payoff it would obtain

[17] In the literature focused on the IEA game, a similar assumption is made in the papers based on the core (Chander and Tulkens, 1997), which also reach a global agreement as a result.

as a member of the fringe if it would leave the coalition and the remaining countries would continue cooperating (Carraro *et al.*, 2006). Carraro *et al.* (2006) show that any distribution of the remaining surplus (after granting each member its fringe payoff) is equally effective in stabilizing the PIS coalition. In our model, we assume that countries bargain over this remaining surplus.

Before proceeding, let us define formally the concepts mentioned in the last paragraphs.

Definition 5 *A coalition $C \in \Gamma$ is internally stable if $v(i, \Gamma_C) \geq v(i, \Gamma_{C\setminus\{i\}})$ $\forall i \in C$, and externally stable if $v(j, \Gamma_C) \geq v(j, \Gamma_{C\cup\{j\}})$ $\forall j \notin C$. The coalition is stable if both conditions hold. The coalition is PIS if $\sum_{j\in C} \left[v(j, \Gamma_C) - v(j, \Gamma_{C\setminus\{j\}})\right] \geq 0$.*

The results in this section crucially depend on whether intermediate coalitions "consolidate" between negotiation rounds (see footnote 8 for a justification). We say that a coalition "consolidates" if it cannot break up in subsequent negotiations.[18] If c approaches a first, we check whether ac is internally stable, but in the negotiation between ac and b, we only check for internal stability against defections from ac and b (i.e., we rule out the possibility that a or c go back to their singleton behavior; they can only go back to the ac coalition). In fact, if coalitions do not consolidate, the whole idea of sequential bargaining is meaningless if one requires all agreements to be internally stable because, in each negotiation round, previous agreements would play no role and players would grant each member of the future coalition its free-rider payoff under the assumption that the remaining players would continue to cooperate.

The bargaining protocol considered is the same as before and we denote Δ^F the remaining surplus after each player receives free-riding payoff. If the leader chooses the multilateral negotiation, the remaining surplus is $\Delta^F_{abc} = v(abc; \Gamma_{abc}) - \sum_{i=a,b,c} v(i; \Gamma_{jk})$, with $\Delta^F_{abc} < \Delta_{abc}$ for positive externalities. However, as there are no "remaining" countries that could continue cooperating in bilateral agreements, we have that $\Delta^F_{ij} = \Delta_{ij}$, and $\Delta^F_{-ij} = \Delta_{-ij}$. Thus, to obtain the equivalent to Lemma 1 it suffices to substitute Δ_{abc} by Δ^F_{abc} in equation (1), as equation (2) remains unchanged. We can now write the following proposition, which compares the results obtained under both assumptions (assuming as in Proposition (1) that σ and δ tend to one):

[18] Carraro and Siniscalco (1993) analyze the role of commitment in a one-shot game with identical players, showing that it favors larger coalitions. They assume that a sub-set of the otherwise identical countries are committed to cooperation. As in our model time appears explicitly, we do not assume commitment but rather that after some time players forming a coalition essentially become one player. Nevertheless, this new player is not more committed to cooperation than the rest, and all equilibria are sub-game perfect.

Proposition 4 *Assume that internally stable agreements are required and that the discount factors tend to one. If coalitions consolidate, the GC is achieved and is stable if superadditivity holds. Furthermore, for the set of value functions for which*

$$E(a) + E(b) < 2E(c), \tag{6}$$

the subset of value functions for which multilateral bargaining is chosen is larger if internally stable agreements are required than under our benchmark assumption (the opposite is true if the relationship is reversed).

Proof: Appendix A.5.

The Proposition shows that cohesiveness is no longer sufficient to ensure the formation of the grand-coalition. The reason is that cohesiveness only implies that the GC generates more surplus than any intermediate coalition structure, but not that the additional surplus is enough to grant each country what it would obtain free-riding if all other countries continue to cooperate. However, if coalitions "consolidate", the grand coalition is still the outcome if superadditivity holds (in the Proposition, superadditivity is a sufficient but not a necessary condition).

Condition (6) tells us that sequential bargaining is more likely to occur if internally stable agreements are required than under our benchmark assumption if externalities are smaller for the leader than for the followers. The intuition for this result is that large fringe payoffs in the follower countries imply that they obtain a larger share of the surplus in a multilateral negotiation with internally stable agreements, reducing the share obtained by the leader, while the surplus secured by the leader in a sequence of bilateral negotiation is the same under both assumptions.

To obtain the equivalent to condition (5) in Proposition (2), it is sufficient to substitute Δ_{abc} by Δ_{abc}^F. However, the concavity of the function $\sigma^{*F}(\delta)$ now also depends on whether Δ_{abc}^F is positive or negative (note that superadditivity does not guarantee that $\Delta_{abc}^F > 0$). If $\Delta_{abc}^F > 0$, the GC is PIS and $\sigma^{*F}(\delta)$ is concave. For this case, Proposition (2) and (3) continue to hold. That is, the role played by cohesiveness corresponds now to PIS.

On the contrary, if $\Delta_{abc}^F < 0$ the GC is not PIS and Propositions (2) and (3) no longer hold. In this case, multilateral bargaining does not increase the payoff for the leader compared with the initial situation, and therefore, the leader will always choose sequential bargaining. More precisely, if superadditivity holds, the leader will choose a sequential path that will ultimately lead to the grand coalition (as, with superadditivity, $\Delta_{ij} > 0$ and $\Delta_{-ij} > 0$). In this case, the mediator discussed above is also not interested in promoting a multilateral negotiation, as it would be doomed. In fact, with the powers that our mediator has (influencing the discount factors), it could not induce the leader to choose a multilateral approach. If, in addition to $\Delta_{abc}^{F} < 0$, superadditivity does also not hold, i.e. $\Delta_{ac} < 0$, the leader will stay as a singleton.

3 A general model

We now generalize the model to a world with many countries that are allowed to bargain following any fixed bargaining protocol. Although the analysis becomes inevitably more complicated, the main message is that the results discussed in the previous section continue to hold, with some minor variations. Thus, readers that are not particularly interested in the technical details can skip this section and move directly to the illustration of the different configurations using particular functions discussed in section 4.

Let $c = C_0$ be the agenda setter and assume that there are $N \geq 2$ follower countries indexed by $a_1, ..., a_N$. The set of all countries, the grand coalition, is denoted by $C_G = \{c, a_1, ..., a_N\}$, and the set of all follower countries that are initially acting as singletons is denoted by $C_\phi = \{a_1, a_2, ..., a_N\}$. We only allow for the formation of one coalition at any moment in time, although coalition size and composition may vary.

In stage one country c chooses to bargain with any subset $S_1 \subset C_\phi$ of the follower countries. Bargaining in stage one follows a fixed protocol that specifies who has the right to propose at what time (this protocol is assumed to be cyclic). Let $p_1(t)$ be the proposer in period $t = 1, 2, 3, ...$ of stage 1 and assume that the periodicity of the protocol is $p_1 < \infty$. From period to period there is a common discount factor δ that applies. An offer from proposer $p_1(t)$ consists of payoffs $P(j, \Gamma_{C_1}, p_1(t))$ for all $j \in c \cup S_1 = C_1$, i.e., the members of the new coalition (to simplify notation, we denote $\Gamma_{C_i} = \langle \{C_i\}, \{h_1\}, \{h_2\}, ..., \{h_m\} \rangle$ any partition formed by a coalition C_i and m countries acting as singletons). As in section 2.1, offers can be conditional or unconditional, allowing us to use the solution concept proposed in Huang (2002). Conditional offers are binding if all the members of $C_1 = c \cup S_1$ accept the offer, while unconditional offers are binding between the proposer and the subset of countries accepting the offer (as before, the proposer buys out the right of the accepting countries and continues negotiating). Stage 1 only ends when an agreement is accepted by all members of C_1 (otherwise, negotiations go on forever following the protocol).

As stated above, Huang (2002) analyzes a multilateral negotiation over the distribution of a cake of size 1 in one stage. The salient feature of our model is that after the negotiation in stage 1 has finished (the case analyzed in Huang (2002)), the players can engage in a second negotiation with an additional set of players. Thus, in the second stage, coalition C_1 decides to bargain with a subset S_2 of the remaining follower countries, i.e., $S_2 \subset C_\phi \backslash S_1$. Stage 2 bargaining follows the same logic as in stage 1, again with a fixed protocol (not necessarily the same). From one stage to the next, the discount factor is σ. An offer from each proposer now consists of payoffs $P(j, \Gamma_{C_2}, p_2(t))$ for all the members of the new coalition ($j \in C_2 = C_1 \cup S_2$). The coalition C_1 formed in the previous stage makes the first offer and members of the coalition decide by unanimity.

The game only moves to the third stage if all countries in S_2 accept one of the offers received/proposed. More generally, if agreements were reached in the first $z-1$ rounds, then in stage z the coalition C_{z-1} bargains with a subset $S_z \subset C_\phi \backslash \cup_{i=1}^{z-1} S_i$ of the follower countries. The first offer is proposed by the coalition C_{z-1} and consists of a new coalition $C_z = C_{z-1} \cup S_z$ and payoffs $P(j, \Gamma_{C_z}, p_z(1))$ for all $j \in C_{z-1} \cup S_z$. Bargaining continues thereafter following the protocol p_z. The game ends at some stage Z if the leader (or the leading coalition) decides not to make any offer to the remaining set of countries, or if the grand coalition is formed. In addition, any of the stages $z \leq Z$ can go on forever following the cyclic protocol of offers. A particular sequence of negotiation stages is denoted as ω. When c decides at stage 1 to bargain with all the follower countries simultaneously, we say that c has chosen multilateral bargaining (i.e., $S_1 = C_\phi$). When c chooses at stage 1 to bargain with any subset $S_1 \neq C_\phi$, we say that c has chosen sequential bargaining (as a tie-breaking rule, we assume that c prefers less stages to more and that for the same number of stages it prefers to reach an agreement with the country or countries that talk first).

We now define our main building blocks. The additional surplus generated by moving from a coalition C_{z-1} to a coalition C_z is denoted by $\Delta_{C_z - C_{z-1}}$ and is defined as:

$$\Delta_{C_z - C_{z-1}} = v(C_z, \Gamma_{C_z}) - \sum_{j \in C_{z-1} \cup S_z} v(j, \Gamma_{C_{z-1}}),$$

where $v(j, \Gamma)$ is the value function already defined in section 2.1.

The share of the surplus obtained by j when the proposer number t is proposing in stage z out of the total Z stages of the sequential path is denoted by $\Phi(j, z/Z, t)$. Following Huang (2002), we define this share as:

$$\Phi(j, z/Z, t) = \frac{\sum_{s \in A_j^{z/Z}(t)} \delta^s}{\delta^t \sum_{s=1}^{p_{z/Z}} \delta^{s-1}}, \tag{7}$$

where $A_i^z(t)$ is the set[19] of periods where player i is the proposer in the first cycle of the proposing protocol starting from period t (with the superscript z/Z standing for stage z out of the total number of Z stages).

[19] More precisely, for stage z, $A_j^z(t) = \{s \mid t \leq s \leq t + p_z - 1 \text{ and } p_z(s) \in D_j(t)\}$, for all $j \in F(t)$, with $F(t)$ denoting the players still in the bargaining game of stage z and $D_j(t)$ the set of players whose right to propose is owned by j at the beginning of period t. See Huang (2002) for details.

3.1 Benchmark model

The following proposition shows our main result for the general model[20]:

Proposition 5 *If the game is cohesive, the grand coalition is always formed, i.e.,* $\Gamma(\omega^*) = \Gamma_{C_G}$*. The equilibrium net payoff for the leader in the unique SPE is given by:*

$$P^*(c, \Gamma(\omega^*), p_{1/Z^*}(1)) = v(c, \Gamma_{C_O}) + \sum_{k=1}^{Z(\omega^*)} \sigma^{k-1} \Delta^{\omega^*}_{C_k - C_{k-1}} \prod_{i=1}^{k} \Phi(C_{i-1}, i/Z(\omega^*), 1), \qquad (8)$$

where ω^* *is the sequence of negotiations that maximizes the payoff for the leader.*

Proof: Appendix A.6.

Proposition (5) shows that cohesiveness is sufficient to ensure formation of the grand coalition (as Proposition (1) for the 3 player case). The intuition is that any intermediate coalition may improve its situation by sharing the additional surplus obtained by moving to the grand coalition. The proposition also shows that the offer proposed by the leader in the first round of negotiations already takes into account all the rounds that will take place in equilibrium (this offer will be immediately accepted). To ease the interpretation of Proposition (5), let us write out the RHS in equation (8):

$$\begin{aligned} & v(c, \Gamma_{C_O}) + \Phi(C_0, 1/Z, 1)\Delta_{C_1 - C_0} + \sigma\Phi(C_0, 1/Z, 1)\Phi(C_1, 2/Z, 1)\Delta_{C_2 - C_1} + .. \\ = \; & v(c, \Gamma_{C_O}) + \frac{1-\delta}{1-\delta^{S_1+1}}\Delta_{C_1 - C_0} + \sigma\frac{1-\delta}{1-\delta^{S_1+1}}\frac{1-\delta}{1-\delta^{S_2+1}}\Delta_{C_2 - C_1} + .. \end{aligned}$$

The first term indicates that the leader country always obtains at least its impasse point value, $v(c, \Gamma_{C_O})$. The second term shows that it also obtains a share equal to $\Phi(C_0, 1/Z, 1)$ of the additional surplus generated in the first stage of the sequential process ($\Delta_{C_1 - C_0}$). The third term shows that the payoff obtained in the second round (if there is a second round in ω^*) has to be discounted using the intra-stage discount factor, σ. The share obtained by the leading coalition (that formed in stage 1) of the additional surplus generated by expansion of the coalition is given by $\Phi(C_1, 2/Z, 1)$. The part of this additional share relevant to the leader country is the portion that it will obtain, and this is a function of the share it secured in the first negotiation stage ($\Phi(C_0, 1/Z, 1)$). Thus, the relevant share is obtained by multiplying both shares. The remaining terms of the RHS of (8) follow the same logic for the remaining potential stages.

The next Proposition shows that the role of the sign of the externalities and the convexity of the game in the general model is also similar to the case analyzed in Proposition (1)

[20] The general definition of cohesiveness is $v(C_G, \Gamma_{C_G}) > \sum_{j \in \Gamma} v(j; \Gamma)$ for every $\Gamma \neq \Gamma_{C_G}$. Superadditivity requires that this hold for all intermediate coalitions as well.

focusing, to allow the comparison, on a situation where the leader can only choose between multilateral bargaining or sequential bargaining in two stages.

Proposition 6 *If sequential bargaining can only take two rounds, the leader prefers multilateral bargaining iff*

$$\left[\Delta_{-C_{C_1}} - \frac{e_3}{e_1}\Delta_{C_{C_1}}\right] + \frac{e_2}{e_1}E(S_0 \diagdown S_1) > 0 \tag{9}$$

where $\Delta_{C_G-C_{C_1}} = \Delta_{-C_{C_1}} = v(C_G, \Gamma_{C_G}) - v(C_1, \Gamma_{C_1}) - \sum_{j \in S_0 \diagdown S_1} v(j, \Gamma_{C_1})$; $\Delta_{C_{C_1}} = v(C_1, \Gamma_{C_1}) - \sum_{j \in S_1} v(j, \Gamma_{C_0}) - v(c, \Gamma_{C_0})$; $E(S_0 \diagdown S_1) = \sum_{j \in S_0 \diagdown S_1} v(j, \Gamma_{C_1}) - \sum_{j \in S_0 \diagdown S_1} v(j, \Gamma_{C_0})$; $e_1 = \Phi(c, 1/1, 1) - \sigma\Phi(c, 1/2, 1)\Phi(C_1, 2/2, 1)$; $e_2 = \sigma\Phi(c, 1/2, 1)\Phi(C_1, 2/2, 1)$ *and* $e_3 = \Phi(c, 1/2, 1) - \Phi(c, 1/1, 1)$. *Thus, convexity and positive coalition externalities in the follower countries generally favor multilateral negotiations while concavity favors sequential negotiations.*

Proof: Appendix A.7.

The Proposition shows that positive coalition externalities favor the multilateral approach as long as $\frac{e_2}{e_1} > 0$ because $E(S_0 \diagdown S_1)$ is positive when coalition externalities are positive. Additionally, $\frac{e_2}{e_1}$ is positive as long as $\frac{\Phi(c,1/1,1)}{\Phi(c,1/2,1)\Phi(C_1,2/2,1)} > \sigma$, which generally holds (unless the bargaining protocol is much more favorable for the leader in the two sequential negotiations than in the unique multilateral negotiation). To see this, assume for simplicity that the bargaining protocol implies at all stages that the leader (or the leading coalition) talks first and all the followers are then allowed to talk once before the leader starts talking again (we call this the "standard" protocol hereafter). Then, the last condition becomes:

$$\frac{\left(1-\delta^{S_1+1}\right)\left(1-\delta^{S_2+1}\right)}{\left(1-\delta^{N+1}\right)(1-\delta)} > \sigma$$

which always holds with $\sigma \leq 1$ and $S_1 + S_2 = N$ because the LHS is always larger than 1.

On the other hand, assuming that there are no coalition externalities to focus on the terms influenced by the convexity of the game, we can see that multilateral bargaining is favored by the leader if $\Delta_{C_G-C_{C_1}}$ is $\frac{e_3}{e_1}$ times $\Delta_{C_{C_1}}$. Finally, $\frac{e_3}{e_1}$ is larger than one as long as $\Phi(c, 1/2, 1)(1 + \sigma\Phi(C_1, 2/2, 1)) > 2\Phi(c, 1/1, 1)$, that is, if the share obtained by the leader country in the first negotiation times the discounted share that it obtains in the second negotiation is more than twice the share obtained in the multilateral negotiation. For the "standard" protocol, $\frac{e_3}{e_1}$ is larger than one as long as

$$\sigma > \frac{\left(1-\delta^{S_2+1}\right)\left(\delta^{N+1} - 2\delta^{S_1+1} + 1\right)}{\left(1-\delta^{N+1}\right)(1-\delta)}$$

and this may or may not hold depending on the value of σ because the RHS is smaller than one. Thus, for a given value function, the smaller σ, the more the leader will favor multilateral bargaining. This result also allows generalizing Propositions (2) and (3), after determining the precise bargaining protocol to be considered.

Before moving on, we show that the three player model discussed above and the game in Aghion *et al.* (2007) are particular cases of Proposition (6).

Example 1 *To obtain the game in Section 2.1, set $N = 2$, $a_1 = a$, $a_2 = b$, $\delta = \sigma \to 1$ and the cyclic protocol to c, a, b, c, a, b and so on. This implies $\Phi(c, 1/1, 1) = \frac{1}{3}$, $\Phi(c, 1/2, 1) = \frac{1}{2}$, $\Phi(C_1, 2/2, 1) = \frac{1}{2}$, $\sigma\Phi(c, 1/2, 1)\Phi(C_1, 2/2, 1) = \frac{1}{4}$ and, therefore, $\Delta_{C_G - C_{C_1}} = \Delta_{-ac}$, $\Delta_{C_{C_1}} = \Delta_{ac}$, $E(S_0 \diagdown S_1) = E(b)$, $\frac{e_2}{e_1} = 3$ and $\frac{e_3}{e_1} = 2$. Substituting these values, equation (9) simplifies to equation (4).*

Example 2 *To obtain the game in Aghion et al. (2007) set $N = 2$, $a_1 = a$, $a_2 = b$, and $\delta = \sigma \to 1$. The assumption that the leader makes only take-it-or-leave-it offers implies in our framework that only the leader talks in the protocol. This implies $\Phi(c, 1/1, 1) = 1$, $\Phi(c, 1/2, 1) = 1$, $\Phi(C_1, 2/2, 1) = 1$, $\sigma\Phi(c, 1/2, 1)\Phi(C_1, 2/2, 1) = 1$, $e_1 = 0$, $e_2 = 1$ and $e_3 = 0$. Hence, equation (9) simplifies to $E(S_0 \diagdown S_1) = E(b) > 0$, and whether multilateral or sequential bargaining is adopted only depends on the sign of the externalities.*

3.2 Internally stable agreements

As in section 2.2, the additional surplus to be shared by moving from a coalition C_{z-1} to a coalition C_z is, now, what is left over after granting each member of the new coalition C_z its free-riding gross payoff. We denote this surplus as $\Delta^F_{C_z - C_{z-1}}$ and define it as:

$$\Delta^F_{C_{z(\omega)} - C_{z(\omega)-1}} = v(C_{z(\omega)}, \Gamma_{C_{z(\omega)}}) - \sum_{j \in C_{z(\omega)-1} \cup S_{z(\omega)}} v(j, \Gamma_{C_{z(\omega)} - \{j\}})$$

Feasible negotiation sequences, $\bar{\omega}$, are those where in every stage $\Delta^F_{C_{z(\omega)} - C_{z(\omega)-1}} > 0$ (the accent above indicates values obtained requiring internally stable agreements). The following proposition extends the result in Proposition (4) that cohesiveness is no longer sufficient to ensure the formation of the grand-coalition, but that superadditivity is.

Proposition 7 *If coalitions consolidate in sequential bargaining and internally stable agreements are required, the GC is achieved and is stable if superadditivity holds, i.e., $\Gamma(\omega^*) = \Gamma_{C_G}$.*

The equilibrium net payoff for the leader in the unique SPE is given by:

$$\begin{aligned} \bar{P}^*(c,\Gamma(\bar{\omega}^*),p_{1/Z^*}(1)) &= v(c,\Gamma_{C_1-\{c\}})+V \quad \textit{if } V>0; \ \textit{otherwise } v(c,\Gamma_{C_O}) \\ \textit{with} \qquad V &= \sum_{k=1}^{Z(\omega^*)} \sigma^{k-1}\Delta^{F\omega^*}_{C_k-C_{k-1}} \prod_{i=1}^{k} \Phi(C_{i-1}, i/Z(\omega^*),1) \end{aligned} \tag{10}$$

where $\bar{\omega}^*$ is the sequence of negotiations that maximizes the payoff for the leader.

Proof: Appendix A.8.

The intuition for the proof is that any intermediate coalition can always follow a one by one enlargement process because in a sequence of bilateral agreements superadditivity ensures that each enlargement is profitable for both parts (as before, superadditivity is a sufficient condition). In fact, if only two parties are negotiating, requiring internally stable agreements and our benchmark assumption lead to the same outcome, because it is not possible to free-ride on the "remaining" countries once one of the two countries defects.

Generalizing the rest of Proposition (4) to more than three players would require taking into account the differences in the sequential bargaining process under both types of assumptions (although smaller fringe payoffs of the leader would still favor sequential bargaining).

Finally, as for the three players case, multilateral bargaining only yields for the leader a payoff larger than its singleton value if the GC is PIS, or if

$$\Delta^F_{C_G-C_0} = v(C_G,\Gamma_{C_G}) - \sum_{j\in c\cup S_0} v(j,\Gamma_{C_G-\{j\}}) > o$$

and, hence, if this does not hold the leader will always choose a sequential path.

4 Illustration with particular functions

This sub-section illustrates the implications of the model previously discussed using the quadratic functions for heterogeneous countries defined in McGinty (2007), which are in turn an adaptation of the functions used by Barrett (1994) for homogeneous countries. In agreement with Barrett (1994), Carraro and Siniscalco (1993) and McGinty (2007) we assume that only one coalition can be formed, which assigns abatement efforts to its members in order to maximize the gross payoff of the coalition, and that the remaining countries act as singletons. However, we do not assume, as did McGinty (2007) and previously Barrett (1994), that the cooperating coalition acts as a Stackelberg leader[21]. We assume instead

[21] Assumption of a Stackelberg leader implicitly assumes that the coalition is the largest player and has therefore more power to commit than the rest of the countries. Nevertheless, with asymmetric countries this

that the cooperating coalition plays a Nash equilibrium against the remaining singletons if no agreement is reached to enlarge the coalition. The other difference is that we model negotiations explicitly and recognize that delays are costly (a climate change agreement today is more effective than the same agreement two decades later).

As before, we consider a set of $C_G = \{c, a_1, ..., a_N\}$ heterogeneous countries. Global benefit is a concave function $B(Q) = \beta(\gamma Q - (Q^2/2))$ where $Q = \sum_{i \in C_G} q_i$ stands for global abatement and q_i for individual abatement. Parameters γ and β are strictly positive. We assume that each nation receives a share of the benefit equal to λ_i, where $\lambda_i > 0$ $\forall i \in C_G$ and $\sum_{i \in C_G} \lambda_i = 1$. All countries are assumed to have convex abatement cost functions depending only on their individual abatement, with $\varsigma_i > 0$. Therefore, the individual gross payoff (before any transfers) is given by:

$$\pi_i(\lambda_i, \varsigma_i, q_i, Q_{-i}) = B_i(Q, \lambda_i) - C_i(q_i, \varsigma_i) = \lambda_i \beta \left(\gamma Q - \frac{Q^2}{2}\right) - \frac{\varsigma_i}{2} q_i^2 \tag{11}$$

The following proposition is useful to analyze this game since it yields the value function:

Proposition 8 *Assume that a coalition C_K with K members has been formed. The individual abatements for signatories and for non-signatories resulting from the Nash equilibrium between the coalition and the singletons on the fringe are given by:*

$$q_i^s = \frac{\gamma\beta \sum_{i \in K} \frac{\lambda_i}{\varsigma_i}}{1 + \beta\Omega_K}; \quad q_j^{ns} = \frac{\gamma\beta \frac{\lambda_j}{\varsigma_j}}{1 + \beta\Omega_K} \tag{12}$$

The gross payoffs of the Nash equilibrium are:

$$v(C_K, \Gamma_{C_K}) = \gamma^2\beta^2 \frac{\lambda_j \left[2\Omega_K + \beta\Omega_K^2\right] - \frac{1}{\varsigma_j}\left(\sum_{i \in K} \lambda_i\right)^2}{2\left(1 + \beta\Omega_K\right)^2} \tag{13}$$

$$v(j, \Gamma_{C_K}) = \gamma^2\beta^2 \frac{\lambda_j \left[2\Omega_K + \beta\Omega_K^2\right] - \frac{1}{\varsigma_j}\lambda_j^2}{2\left(1 + \beta\Omega_K\right)^2} \tag{14}$$

with $\Omega_K = \sum_{i \in K} \lambda_i \sum_{i \in K} \frac{1}{\varsigma_i} + \sum\limits_{j \notin K} \frac{\lambda_j}{\varsigma_j}$.

Proof: Appendix A.9

To illustrate the different configurations that can appear we provide now some simulation results[22]. To simplify, we focus on a game with 5 players[23] (one leader, c, and 4

is not necessarily the case because a group of small countries may form a coalition which is too small to act as a Stackelberg leader.

[22] All simulations have been done using Scilab 5.3.3. Codes are available from the authors upon request.

[23] Although we focus only on the five player case to ease the exposition, one can also find cases where sequential bargaining is chosen with three players.

followers a_i, $i = 1..4$). Before moving to Monte-Carlo simulations, we discuss in more detail three examples, assuming in all cases $\delta \to 1$, $\gamma = 5$ and $\beta = 50$. On the costs side we denote $\boldsymbol{\varsigma} \equiv (\varsigma_c, \varsigma_{a_1}, \varsigma_{a_2}, \varsigma_{a_3}, \varsigma_{a_4})$ and consider two parameter sets: $\boldsymbol{\varsigma}^l \equiv (100, 80, 75, 70, 25)$, implying that the leader has larger abatement costs, and $\boldsymbol{\varsigma}^f \equiv (25, 80, 75, 70, 100)$. On the benefits side, we denote $\lambda \equiv (\lambda_c, \lambda_{a_1}, \lambda_{a_2}, \lambda_{a_3}, \lambda_{a_4})$ and consider three cases: (i) an equal distribution of benefits $\lambda^e \equiv (0.2, 0.2, 0.2, 0.2, 0.2, 0.2)$,(ii) a distribution with one large follower $\lambda^f \equiv (0.03, 0.08, 0.08, 0.08, 0.73)$, and (iii) a distribution with one large leader $\lambda^l \equiv (0.73, 0.08, 0.08, 0.08, 0.03)$.

To get a feeling of the likelihood of the different paths, we also perform Monte-Carlo simulations[24], with 1000 repetitions. The game is superadditive[25] in 584 of the 1000 cases considered and, as expected, always cohesive.

4.1 Benchmark model

Table 1 shows the results for our examples. As the game is cohesive, the grand coalition is always formed. However, the path chosen by the leader depends on the parameters of the model. For example, for $\sigma \to 1$, the leader chooses a multilateral approach with $(\lambda^e, \boldsymbol{\varsigma}^l)$ and a sequential approach with $(\lambda^f, \boldsymbol{\varsigma}^l)$. As our theoretical model has highlighted, lowering the discount factors influences these results, reducing the number of stages with $\sigma = 0.90$ for $(\lambda^f, \boldsymbol{\varsigma}^l)$, and ultimately inducing the leader to choose a multilateral approach for $(\lambda^f, \boldsymbol{\varsigma}^l)$ (for values of σ below 0.4). Thus, if this option is available, the mediator would like to enlarge the period between negotiations stages, θ, to make sure that $\sigma = e^{-r\theta t} < 0.40$ for $(\lambda^f, \boldsymbol{\varsigma}^l)$.

The Monte-Carlo simulations confirm that sequential bargaining is by no means an exception, as sequential bargaining is chosen by the leader in 857 out of 1000 cases. Although not in one step, in all these cases the GC is the final outcome. As the game is cohesive, "all singletons" or partial cooperation are never the final outcome, as predicted by Proposition (5), so that multilateral bargaining is the option selected in the remaining 143 cases (yielding, again, the GC).

[24] λ values are randomly selected, ensuring that $\sum \lambda_i = 1$. The remaining parameters are randomly selected from the following intervals: $\gamma \in [1, 10]$, $\beta \in [1, 100]$, $c \in [1, 100]$, $\delta \in [0.95, 1]$, $\sigma \in [0.8, \delta]$.

[25] Note that although convexity always holds superadditivity does not hold generally. Nevertheless, it holds for a sub-class of the payoff function shown in (11) that has been used frequently in the analysis of IEA, namely the case of identical players with linear benefits and quadratic (Breton et al., 2006). Thus, for the linear-quadratic payoff function, Propositions 5 and 7 imply that the GC is achieved under both assumptions considered, although not necessarily in one step.

Table 1. Examples of negotiation paths

parameter values	σ $\rightarrow 1$	0.9	max multi
Benchmark			
λ^e, ς^l	C_G	C_G	$\rightarrow 1$
λ^f, ς^l	$a_2a_4, a_2a_3a_4, C_G$	$a_2a_3a_4, C_G$	$\sigma = 0.40$
λ^l, ς^f	$a_3, a_2a_3, a_1a_2a_3, C_G$	$a_2a_3, a_1a_2a_3, C_G$	ϕ
Internally stable agreements			
λ^e, ς^l	$a_2, a_2a_4, a_2a_3a_4, C_G$	$a_2, a_2a_4, a_2a_3a_4, C_G$	ϕ
λ^f, ς^l	$a_4, a_3a_4, a_2a_3a_4, C_G$	$a_4, a_3a_4, a_2a_3a_4, C_G$	ϕ
λ^l, ς^f	C_G	C_G	$\rightarrow 1$

Note: The second column shows the path chosen for the leader when $\sigma \rightarrow 1$, the third the path chosen for $\sigma = 0.9$ and the fourth the maximum level of σ for which multilateral bargaining is the outcome. In each cell, commas separate the agreements formed in the sequence (excluding c).

4.2 Internally stable agreements

We also obtain multilateral and sequential bargaining under this assumption. For (λ^l, ς^f) the GC is PIS and the leader chooses multilateral bargaining even for $\sigma \rightarrow 1$. On the contrary, for (λ^e, ς^l) and (λ^f, ς^l) the GC is not PIS and the leader chooses always a sequential approach, as a multilateral agreement is not possible under this assumption if the GC is not PIS.

The Monte-Carlo simulations show that the leader selects sequential bargaining in 926 cases. Out of these cases, only in 7 cases is the GC not achieved. Multilateral bargaining is selected in 55 cases, yielding always the GC when it is selected. Finally, the leader decides to remain as a singleton in 12 cases. That is, only in 19 cases is the GC not achieved. This is a remarkable result considering that the GC is PIS in a multilateral setting only in 103 out of the 1000 cases considered.

5 Conclusion

This paper has discussed a model of gradual coalition formation with positive externalities in which a leading country endogenously decides whether to negotiate multilaterally or sequentially over climate change. Our analysis has shown that the two features driving the results are how much bargaining power the followers have, and the assumption made about the behavior of the remaining players if one player breaks an agreement.

Assuming followers with no bargaining power (take-it-or-leave-it offers from the leader) and that agreements are terminated if one player breaks them, Aghion *et al.* (2007) show that the grand coalition will be achieved and that the leader would always prefer a multilateral negotiation process. Relaxing the bargaining power assumption, we have highlighted the role played by the convexity of the game, showing that the leader may prefer a sequential path even with positive externalities. This extends considerably the cases where sequential negotiations may be chosen and helps explain the recent trend toward bilateral or regional climate change negotiations.

Requiring agreements to be internally stable, we have shown that reaching a stable GC is only guaranteed if superadditivity holds, and that the leader chooses sequential bargaining more often if its free-rider payoff is small. We have also shown that even if the grand coalition is not stable in a multilateral setting, the sequential approach may reach a stable global agreement. Furthermore, our simulations have shown that this is in fact the case for most parameter values. For this path to be feasible, coalitions need to consolidate. In real-life negotiations, this implies giving intermediate coalitions some time to build up common institutions, such an emissions trading market, to make it impossible for them to break up in subsequent negotiations to enlarge cooperation. This clearly shows that a multilateral approach is not necessarily the best strategy to negotiate a climate change agreement, and provides a rationale for focusing on linking the European emissions trading system to other systems in a sequential process.

References

[1] Aghion, P., Antràs, P. and Helpman, E., 2007. Negotiating free trade. *J. Int. Econ.* 73:1-30.

[2] Asheim, G.B., Froyn, C.B., Hovi, J. and Menz, F.C., 2006. Regional versus global cooperation for climate control. *J. Environ. Econ. Manage* 51:93-109.

[3] d'Aspremont, C., A. Jacquemin, J.J. Gabszewicz and J. Weymark, 1983. On the Stability of Collusive Price Leadership. *Can. J. Econ.* 16, 17-25.

[4] Barrett, S., 1994. Self Enforcing International Environmental Agreements, *Oxford Econ. Pap.* 46:878-894.

[5] Benedick, R.E., 1998. *Ozone Diplomacy: New Directions in Safeguarding the Planet - Enlarged Edition.* Harvard University Press, Cambridge.

[6] Benedick, R.E., 2007. Avoiding Gridlock on Climate Change. *Issues in Science and Technology* 23: 2.

[7] Bhagwati, J.N., 1993. Regionalism and multilateralism: an overview. In: De Melo, J., Panagariya, A. (eds.), *New Dimensions in Regional Integration*. Cambridge University Press, Cambridge, UK:22-51..

[8] Bloch, F., 1995. Endogenous structures of association in oligopolies. *RAND J. Econ.* 26(3): 537-556.

[9] Breton, M., Fredj, K. and Zaccour, G., 2006. International cooperation, coalitions stability and free riding in a game of pollution control. *Manchester Sch.* 74(1): 103-122.

[10] Camiña, E. and Porteiro, N., 2009. The role of mediation in peacemaking and peacekeeping negotiations. *Europ. Econ. Rev.* 53: 73–92.

[11] Caparrós A., Péreau, J.-C. and Tazdaït, T., 2004. North-South Climate Change Negotiations: a Sequential Game with Asymmetric Information, *Public Choice* 121 (3-4):455-480.

[12] Carraro, C. and Siniscalco, D., 1993. Strategies for the International Protection of the Environment. *J. Public Econ.* 52:309-328.

[13] Carraro, C. and Siniscalco, D., 1998. International environmental agreements: incentives and political economy. *Europ. Econ. Rev.* 42: 561-572.

[14] Carraro, C., 2005. Institution design for managing global commons: lessons from coalition theory. In G. Demange and M. Wooders (eds.), *Group Formation in Economics: Networks, Clubs and Coalitions*. Cambridge University Press, Cambridge, UK: 354-380.

[15] Carraro, C., Eyckmans, J. and Finus, M., 2006. Optimal transfers and participation decisions in international environmental agreements. *Rev. Int. Organ.* 1:379–396..

[16] Chander, P. and Tulkens, H., 1997. The Core of an Economy With Multilateral Environmental Externalities. *Int. J. Game Theory* 26: 379-401.

[17] Chatterjee, K., Dutta, B., Ray, D. and Sengupta, K., 1993. A noncooperative theory of coalitional bargaining. *Rev. Econ. Stud.* 60(2): 463-477.

[18] Finus, M. and Caparrós, A., 2015. *Game Theory and International Environmental Cooperation: Essential Readings*. Edward Elgar, Cheltenham.

[19] Finus, M., Rundshagen, B. and Eyckmans, J., 2010. Simulating a sequential coalition formation process for the climate change problem: first come, but second served? *Ann. Oper. Res.* (DOI: 10.1007/s10479-010-0815-x).

[20] Gomes, A. 2005. Multilateral Contracting with Externalities. *Econometrica* 73(4):1329-1350.

[21] Harstad, B., 2012. Climate Contracts: A Game of Emissions, Investments, Negotiations, and Renegotiations. *Review of Economic Studies* 79(4): 1527-57.

[22] Huang, C.-Y., 2002. Multilateral bargaining: conditional and unconditional offers. *Econ. Theory* 20:401–412.

[23] Krishna, P., 1998. Regionalism and Multilateralism: A Political Economy Approach. *Q. J. Econ.* 113 (1): 227-251.

[24] Krishna, V. and Serrano, R., 1996. Multilateral bargaining. *Rev. Econ. Stud.* 63:61–80.

[25] Levy, P.I., 1997. A Political-Economic Analysis of Free-Trade Agreements. *Am Econ Rev* 87(4): 506-51.

[26] Macho-Stadler, I, and Xue, L., 2007. Winners and Losers from the Gradual Formation of Trading Blocs. *Economica* 74, 664–681.

[27] McGinty, M., 2007. International environmental agreements among asymmetric nations. *Oxford Econ. Pap.* 59: 45-62.

[28] Muthoo, A., 1999. *Bargaining Theory with Applications.* Cambridge University Press, UK, 357p.

[29] Osborne, M.J. and Rubinstein, A., 1990. *Bargaining and Markets.* Academic Press. San Diego, U.S.

[30] Rubinstein, A., 1982. Perfect Equilibrium in a Bargaining Model. *Econometrica* 50:97-108.

[31] Saggi, K. and Yildiz, H.M., 2010. Bilateralism, Multilateralism, and the Quest for Global Free Trade. *J. Int. Econ.* 81:26–37.

[32] Saggi, K. and Yildiz, H.M., 2011. Bilateral Trade Agreements and the Feasibility of Multilateral Free Trade. *Rev. of Int. Econ.*, 19(2):356–373

[33] Sandler, T., 2004. *Global Collective Action.* Cambridge University Press. Cambridge, U.K.

[34] Sen, N. and Biswas, R., 2015. Trade negotiations under alternative bargaining structure. *Econ. & Politics* 27(3): 509-523.

[35] Shapley, L.S., 1971. Cores of Convex Games. *Int. J. Game Theory* 1:11–26.

[36] Sutton, J., 1986. Non-cooperative bargaining theory: An introduction. *Rev. Econ. Stud.* 53:709-724.

[37] Telser, L.G., 1980. A Theory of Self-Enforcing Agreements. J Bus 53 (1): 27-44

[38] Yi, S.S., 1996. Endogenous formation of customs unions under imperfect competition: open regionalism is good. *J. Int. Econ.* 41:153-177.

A Appendix (proofs)

A.1 Lemma 1

For multilateral bargaining there is a single stage and the bargaining game is similar to the one analyzed in Theorem 3 in Huang (2002), except that (i) the surplus shared is not 1, but the surplus generated by moving from the all singletons situation to the grand coalition and (ii) the reservation payoffs are not zero. As these changes do not modify the proof substantially, it is omitted (the full proof is available from the authors). Thus, from this analysis we know that the equilibrium is unique, that the first offer made by the leader will be immediately accepted and that the offer grants each player j its disagreement payoff $v(c;\Gamma_\phi)$ plus a share $\Phi(j,m)$ of the additional surplus created through the agreement. For the cyclical protocol assumed, the share obtained by the leader is given by:

$$\Phi(c,m) = \frac{\sum_{s\in A_j^m(t)} \delta^s}{\delta^t \sum_{s=1}^{3} \delta^{s-1}} = \frac{\delta^1}{\delta^1\left(\delta^0+\delta^1+\delta^2\right)} = \frac{1}{1+\delta+\delta^2},$$

where $A_i^m(t)$ is the set of periods where player c is the proposer in the first cycle of the negotiation protocol (which is 1 as the protocol is $c,a,b,c,a,b...$) starting from period $t=1$. For a more general presentation of this share see section 3. Hence, the payoff obtained by the leader in equilibrium, where the first offer is immediately accepted, is given by equation (1).

For bilateral bargaining, the share obtained in the first (and in the second) stage of the game in equilibrium is given by:

$$\Phi(c,s) = \frac{\sum_{s\in A_j^s(t)} \delta^s}{\delta^t \sum_{s=1}^{2} \delta^{s-1}} = \frac{\delta^1}{\delta^1 \left(\delta^0 + \delta^1\right)} = \frac{1}{1+\delta}.$$

If the leader country approaches first a, in equilibrium it obtains its disagreement payoff, $v(c;\Gamma_\phi)$ plus a share $\Phi(c,s)$ of the additional surplus generated by forming the first coalition (Δ_{ac}), as we assume the protocol $c,a,c,a...$ In addition, after θ periods the coalition $\{ac\}$ will start negotiating with player b. As the protocol in this second negotiation is $ac, b, ac, b, ...$ the coalition obtains a share $\Phi(ac,s)$ of the additional surplus created Δ_{-ac}, with $\Phi(ac,s) = \Phi(c,s) = 1/(1+\delta)$. The leader knows at the beginning of the game, that out of this surplus obtained by the coalition $\{ac\}$, it obtains a share $\Phi(c,s)$, which needs to be discounted using the discount factor σ.

The payoff that the leader obtains under sequential bargaining if it approaches b first is calculated using the same reasoning. The leader approaches first the player with which it obtains a higher payoff, yielding equation (2).

A.2 Proposition 1

The leader prefers multilateral if $P(c,\Gamma_{abc},m) - P(c,\Gamma_{abc},s) > 0$, where both expressions are obtained applying Lemma (1), and taking the limit as the discount factors tend to 1. Assuming that the leader approaches a first, or that $\Delta_{ac} + E(a) > \Delta_{bc} + E(b)$, this inequality simplifies to (4). Positive externalities in the follower countries favor multilateral bargaining because they increase the second term in (4). To see that convexity also favors multilateral bargaining note that it is a necessary condition when there are no externalities. A concave game implies $\Delta_{bc} > \Delta_{-ac}$ and, as c approaches a first, we know that $\Delta_{ac} > \Delta_{bc}$. Thus, we know that for a concave game $\Delta_{ac} > \Delta_{-ac}$, making the first term in (4) negative. The same argument holds if the conditions are met for the leader to approach b first.

A.3 Proposition 2

From Lemma (1) we know that, assuming that c approaches a first in sequential bargaining, c chooses multilateral if $P(c,\Gamma_{abc},m) - P(c,\Gamma_{abc},s) > 0$. Solving this equation for σ equation

(5) is found. As

$$\begin{aligned}
\frac{\partial \sigma^*(\delta)}{\partial \delta} &= \frac{\left(1-\delta^2\right)}{\left(1+\delta+\delta^2\right)^2}\frac{\Delta_{abc}}{\Delta_{-ac}} - \frac{\Delta_{ac}}{\Delta_{-ac}} \\
\frac{\partial^2 \sigma^*(\delta)}{\partial \delta^2} &= -\frac{2\left(3\delta+1-\delta^3\right)}{\left(\delta^2+\delta+1\right)^3}\frac{\Delta_{abc}}{\Delta_{-ac}},
\end{aligned}$$

the function is concave if $\Delta_{-ac} > 0$ and convex if $\Delta_{-ac} < 0$. Given that $\frac{\left(1-\delta^2\right)}{\left(1+\delta+\delta^2\right)^2}$ is a continuously decreasing function that starts at 1 for $\delta = 0$ and takes a value of zero for $\delta = 1$, and by cohesiveness $\Delta_{-ac} > 0$ and $\Delta_{abc} > \Delta_{ac}$, the first derivative is positive for low values of δ and negative for large values, implying that $\sigma^*(\delta)$ is maximized for δ^*, with $0 < \delta^* < 1$.

A.4 Proposition 3

If $\sigma^*(\delta) \leq 1$ never holds, then $\sigma \leq \sigma^*$ always holds; and if $0 \leq \sigma^*(\delta)$ never holds then $\sigma \leq \sigma^*$ never holds. An efficiency seeking mediator, $\alpha = 0$, prefers a multilateral negotiation because it yields an immediate agreement with $\sum_{i \in a,b,c} P(i, \Gamma_{abc}, m) = v(abc; \Gamma_{abc})$, while with sequential bargaining, at least one period elapses, implying that $\sum_{i \in a,b,c} P(i, \Gamma_{abc}, s) < v(abc; \Gamma_{abc})$, with $\sigma < 1$. If $0 < \sigma^*(\delta) < 1$ holds for some values of δ, by Proposition (2) the leader will select multilateral bargaining for any $\sigma \leq \sigma^*$ and the agreement will be reached immediately with the first offer. Thus, irrespective of the value of δ and as long as $\sigma \leq \sigma^*$, there will be no welfare loss. However, if $\alpha > 0$, the mediator will choose the most egalitarian outcome out of the set for which multilateral bargaining is chosen by the leader, and this is achieved by setting $\delta \to 1$. Irrespective of the value of α, the mediator has no incentive to induce sequential bargaining, as this reduces the efficiency and cannot improve the equity of the sharing compared with a multilateral deal with $\delta \to 1$.

A.5 Proposition 4

Using this assumption multilateral bargaining yields the following payoffs for the leader:

$$\bar{P}(c, \Gamma, s) = v(c; \Gamma_{ab}) + \Phi(c, m)\Delta_{abc}^F$$

with $\Phi(c, m) = 1/3$ if the discount factors tend to one. Assuming consolidation of partial coalitions, the results of sequential bargaining are the same as shown for the benchmark assumption. From the proof of Proposition (1) we know that under the benchmark assumption the leader prefers multilateral over sequential bargaining if $P(c, \Gamma, m) - P(c, \Gamma, s) > 0$

(and sequential bargaining if < 0). Because the payoff for the leader under sequential bargaining is the same under both assumptions (i.e., $\bar{P}(c,\Gamma,s) = P(c,\Gamma,s)$), the new condition for preferring multilateral bargaining is $\bar{P}(c,\Gamma,m) - P(c,\Gamma,s) > 0$ (the accent above indicates values obtained requiring internally stable agreements). Therefore, as long as $P(c,\Gamma,m) > \bar{P}(c,\Gamma,m)$, the leader prefers sequential bargaining in more cases requiring internally stable agreements. Substituting the values for $P(c,\Gamma,m)$ and $\bar{P}(c,\Gamma,m)$ shown above, we obtain: $v(c;\Gamma_\phi) + \frac{\Delta_{abc}}{3} > v(c;\Gamma_{ab}) + \frac{\Delta^F_{abc}}{3}$. Basic manipulations yield condition (6).

A.6 Proposition 5

The proof has three steps. First, we define in Lemma 2 the equilibrium outcome under multilateral bargaining, since this is a special case of the analysis in Huang (2002). Then we show the equilibrium transfer under sequential bargaining, in Lemma 3, and finally we show that the grand coalition is always achieved.

A.6.1 Step 1. Equilibrium outcome of the multilateral bargaining process

Lemma 2 *Assume that c has decided at stage 1 to bargain with all follower countries simultaneously (i.e. c has chosen multilateral bargaining and $C_1 = C_G$). The unique equilibrium outcome is reached in period 1 of the unique stage and is given by*

$$P^*(j, \Gamma_{C_G}, p_{1/1}(1)) = v(j; \Gamma_{C_O}) + \Phi(j, 1/1, 1)\Delta_{C_G - C_0} \textit{ for all } j \in c \cup S_1 = c \cup C_\phi$$

As already mentioned for the three players case, if the leader chooses multilateral bargaining the bargaining game is similar to the one analyzed in Theorem 3 in Huang (2002), with the modifications discussed above. As these changes do not modify the (rather long) proof substantially, it is omitted (the full proof is available from the authors).

A.6.2 Step 2. Equilibrium outcome of the sequential bargaining process

Lemma 3 *Assume that c has decided at stage 1 to follow a sequential bargaining process. The unique equilibrium in any stage z of the game is:*

$$\begin{aligned} P^*(j, \Gamma_{C_z}, p_{z/Z}(1)) &= P^*(j, \Gamma_{C_{z-1}}, p_{z-1/Z}(1)) + \sum_{k=z}^{Z} \sigma^{k-z} \Delta_{C_k - C_{k-1}} \prod_{i=z}^{k} \Phi(n, i/Z, 1) \\ \textit{for all } j &\in C_{z-1} \cup S_z, \textit{ with } n = j \textit{ if } i = z \textit{ and } n = C_{i-1} \textit{ if } i > z \end{aligned}$$

Suppose that in stage Z the coalition negotiates with all the remaining players. The reservation payoff of the coalition, which includes the leader, is the payoff obtained in the previous negotiation in phase $Z-1$: $P^*(C_{Z-1}, \Gamma_{C_{Z-1}}, p_{Z-1}(1))$. Similarly, the remaining players have a reservation payoff of $P^*(j, \Gamma_{C_{Z-1}}, p_{Z-1}(1))$, for all $j \in \cup S_Z$. Hence, by Lemma (2) the equilibrium offer proposed by the coalition in period 1 of stage Z and immediately accepted is $P^*(j, \Gamma_{C_Z}, p_{Z/Z}(1)) = P^*(j, \Gamma_{C_{Z-1}}, p_{Z-1}(1)) + \Phi(j, Z, 1)\Delta_{C_Z - C_{Z-1}}$, for all $j \in C_{Z-1} \cup S_Z$.

In the previous stage $(Z-1)$, the reservation payoff of the players is $P^*(j, \Gamma_{C_{Z-2}}, p_{Z-2}(1))$, for all $j \in C_{Z-2} \cup S_{Z-1}$, and the surplus shared is (i) the surplus produced by moving from C_{Z-2} to C_{Z-1} plus (ii) the discounted additional surplus to be obtained at stage Z (discounted using the between-stages factor σ). Thus, following a similar reasoning as in Lemma (2) the payoff for j in stage $Z-1$ is, for all $j \in C_{Z-2} \cup S_{Z-1}$,

$$
\begin{aligned}
&P^*(j, \Gamma_{C_{Z-1}}, p_{Z-1/Z}(1)) = P^*(j, \Gamma_{C_{Z-2}}, p_{Z-2/Z}(1)) + \\
&+\Phi(j, Z-1, 1)\left\{\Delta_{C_{Z-1} - C_{Z-2}} + \sum_{k=(Z-1)+1}^{Z} \sigma^{k-(Z-1)} \Delta_{C_k - C_{k-1}} \prod_{i=(Z-1)+1}^{k} \Phi(C_{i-1}, i, 1)\right\}
\end{aligned}
$$

More generally, for any stage z the equilibrium outcome is:

$$
\begin{aligned}
P^*(j, \Gamma_{C_z}, p_{z/Z}(1)) &= P^*(j, \Gamma_{C_{z-1}}, p_{z-1}(1)) + \\
&\quad +\Phi(j, z, 1)\left\{\Delta_{C_z - C_{z-1}} + \sum_{k=z+1}^{Z} \sigma^{k-z} \Delta_{C_k - C_{k-1}} \prod_{i=z+1}^{k} \Phi(C_{i-1}, i, 1)\right\}
\end{aligned}
$$

and this can be rewritten as in (8). Given the tie-breaking rule assumed this equilibrium is unique since multiple equilibria only appear if the leader obtains the same payoff with different negotiation sequences.

A.6.3 Step 3. Grand coalition

The generalized definition of cohesiveness implies that $v(C_G, \Gamma_{C_G}) > \sum_{j \in \Gamma} v(j; \Gamma)$ for every $\Gamma \neq \Gamma_{C_G}$. Any coalition smaller than the GC prefers moving to the GC in one step to staying as an intermediate coalition (it may also prefer moving in several stages, but to show that the GC forms it is sufficient to show that it prefers to move in one stage). Suppose that at stage T a coalition has been formed with $m < c \cup N$ members. This coalition may offer the remaining players a share $\Phi(j, 1/1, 1)$ of $\Delta_{C_G - C_T}$, which will be accepted and increase the payoff for all players since $\Delta_{C_G - C_T} > 0$ by cohesiveness.

A.7 Proposition 6

From Lemma (2) we know that the payoff for the leader from multilateral bargaining is $P(c, \Gamma_{C_G}, p_{1/1}(1)) = v(j; \Gamma_{C_0}) + \Phi(c, 1, 1)\Delta_{C_G - C_0}$. From Lemma (3), we know that the leader obtains from a sequential bargaining process with 2 stages:

$$P(c, \Gamma_{C_G}, p_{1/2}(1)) = v(j; \Gamma_{C_0}) + \Phi(c, 1/2, 1)\Delta_{C_1 - C_0} + \sigma\Phi(c, 1/2, 1)\Phi(C_1, 2/2, 1)\Delta_{C_G - C_1}.$$

The leader prefer sequential bargaining if $P(c, \Gamma_{C_G}, p_{1/2}(1)) > P(c, \Gamma_{C_G}, p_{1/1}(1))$, or if (9).

A.8 Proposition 7

Assume consolidation of coalitions. To show that the grand coalition forms if superadditivity holds, suppose that coalition C_T has been formed at stage T with $m < c \cup N$ members and that there are $h_1, h_2, ...h_M$ countries which are not part of the coalition. The coalition can propose h_1 a payoff $\Phi(C_T \cup h, 1/, 1)\Delta^F_{C_T \cup h_1 - C_T}$ equal to

$$\Phi(C_T \cup h, 1/, 1) \left\{ v(C_T, \Gamma_{\{C_T \cup h_1, h_2, ...h_M\}}) - v(C_T, \Gamma_{\{C_T, h_1, h_2, ...h_M\}}) - v(h_1, \Gamma_{\{C_T, h_1, h_2, ...h_M\}} \right\},$$

which is immediately accepted by h_1 and positive by superadditivity (because $\Delta^F_{C_T \cup h_1 - C_T} = \Delta_{C_T \cup h_1 - C_T}$). This holds for the initial coalition structure where no coalitions are formed and for all intermediate coalitions until the grand coalition is achieved. That is, as in bilateral negotiations the surplus to be shared between the countries is the same under both assumptions, the leader can at least choose a bilateral bargaining path to reach the grand coalition and no intermediate coalition will be interested in interrupting the bilateral sequence until the grand coalition.

Nevertheless, cohesiveness is no longer sufficient to ensure that the grand coalition is formed. The reason is that the payoff just defined is not necessarily positive if only cohesiveness holds. In addition, the leader cannot achieve the GC directly, because cohesiveness does not ensure that all the free-riding gross payoffs can be allocated to each country. I.e., $\Delta^F_{C_G - C_T}$ may be negative even if cohesiveness holds.

To obtain equation (10) follow the proof to obtain equation (8) substituting Δ^{ω^*} with $\Delta^{F\bar{\omega}^*}$ and $v(c, \Gamma_{C_O})$ with $v(c, \Gamma_{C_1 - \{c\}})$. In addition, note that if no sequence $\bar{\omega}$ yields a positive V the leader remains as a singleton.

A.9 Proposition 8

We consider a Nash equilibrium between the coalition of size K and the $(N+1)-K$ non signatories acting as singletons. The program of the signatories coalition is to maximize

$$\sum_{j\in K}\pi_i^s(Q,\lambda_i)=\sum_{j\in K}\lambda_i\beta\left(\gamma\left(Q_K+Q_{(N+1)-K}\right)-\frac{\left(Q_K+Q_{(N+c)-K}\right)^2}{2}\right)-\frac{1}{2}\sum_{j\in K}\varsigma_i q_j^2 \qquad (15)$$

while the $(N+1)-K$ singletons maximize

$$\pi_i^{ns}(Q,\lambda_i)=\lambda_i\beta\left(\gamma\left(Q_K+q_i+Q_{(N+1)-K-i}\right)-\frac{\left(Q_k+q_i+Q_{n-k-i}\right)^2}{2}\right)-\frac{1}{2}\varsigma_i q_i^2 \qquad (16)$$

The reaction function for a signatory is $q_i^s=\frac{\beta\sum_{i\in K}\lambda_i}{\varsigma_i\left(1+\beta\sum_{i\in K}\lambda_i\sum_{i\in K}\frac{1}{\varsigma_i}\right)}(\gamma-Q_{ns})$ and for a non-signatory $q_i^{ns}=\frac{\beta\frac{\lambda_i}{\varsigma_i}}{1+\beta\sum_{j\notin K}\frac{\lambda_i}{\varsigma_i}}(\gamma-Q_s)$. After substitution, individual abatements for signatories, q_i^s, and non-signatories, q_j^{ns}, are the ones shown in equation (12). The payoffs shown in equations (13) and (14) of the main text are obtained substituting these values into equation (11), taking into account that $v(C_K,\Gamma_{C_K})=\sum_{i\in K}\pi_i^s$ and $v(j,\Gamma_{C_K})=\pi_i^{ns}$.

B Cohesiveness fails

In the main text, we have assumed that cohesiveness holds. This is a rather mild requirement for most public goods and is generally accepted in IEA analyses. Nevertheless, the literature on FTA (Aghion et al, 2007; Saggi and Yildiz 2010 and 2011) has considered the possibility that cohesiveness (or GC superadditivity) does not hold, and we will therefore analyze it. Cohesiveness can fail because $\Delta_{abc}<0$ and/or $\Delta_{-ij}<0$. As in the papers just mentioned, the consequence is that both stumbling and building block equilibria are possible.

If $\Delta_{abc}<0$, a multilateral agreement is not possible. However, if $\Delta_{ac}>0$ and $\Delta_{-ac}>0$ sequential bargaining will yield the GC, and the intermediate coalition is a building block. In fact, if the discount rate σ is sufficiently low, sequential bargaining may lead to the GC even if $\Delta_{ac}<0$, as long as $\Delta_{-ac}>0$ and the gain in the second stage is enough to compensate the loss in the first stage. Hence, the mediator could have a role to play if it can decrease σ.

If $\Delta_{-ac}<0$ and $\Delta_{abc}>0$, multilateral bargaining would yield the grand coalition, while a partial agreement between a and c would be a stumbling block that would avoid achievement of the GC. A purely efficiency seeking mediator would have no incentive to favor multilateral bargaining, as it would reduce the aggregated payoff of the three countries. Nevertheless, if the mediator has an $\alpha > 0$, it may be interested in promoting multilateral bargaining, as the intermediate situation can be very unequal. Finally, if $\Delta_{-ac} < 0$ and $\Delta_{abc} < 0$, the GC would not be achieved following a sequential nor a multilateral approach.

Cahiers du GREThA
Working papers of GREThA

GREThA UMR CNRS 5113

Université de Bordeaux

Avenue Léon Duguit
33608 PESSAC - FRANCE
Tel : +33 (0)5.56.84.25.75
Fax : +33 (0)5.56.84.86.47

http://gretha.u-bordeaux.fr/

Cahiers du GREThA (derniers numéros – last issues)

2015-22 : BOKINO Régis, Transparence dans les deux banques centrales de la zone franc africaine : évaluation et comparaison internationale

2015-23 : DUCRAY François, *Les aides aux étudiants - Une question d'arbitrage entre les bourses sociales et la demi-part fiscale*

2015-24 : FLAMAND Marina, FRIGANT Vincent, *L'hétérogénéité des comportements des constructeurs automobiles vis-à-vis du Corporate Venture Capital*

2015-25 : BRESCHI Stefano, LISSONI Francesco , MIGUELEZ Ernest*, Foreign inventors in the US: Testing for Diaspora and Brain Gain Effects*

2015-26 : ASSOGBA Guillaume, KLEBANER Samuel*, Vers un cadre d'analyse institutionnaliste de la politique de filière : Quelle cohérence pour la politique de filière française ?*

2015-27 : LALANNE Aurélie, ZUMPE Martin*, Zipf's law, Gibrat's law and Cointegration*

2015-28 : CHARLES Léo, *Evolution of trade patterns and economic performance: the case of France and Switzerland during the nineteenth century*

2015-29 : LLOPIS Oscar, *Who do you care about? Scientists' personality traits and perceived beneficiary impact*

2015-30 : ZUMPE Martin, Le nouveau modèle keynésien : une interprétation d'«Interest and Prices» de Michael Woodford

2015-31 : BERGÉ Laurent, WANZENBOCK Iris, SCHERNGELL Thomas : *Centrality of regions in R&D networks: Conceptual clarifications and a new measure*

2015-32 : ODUOR Jacob, MABROUK Fatma, SHIMELES Abebe : Remittances and Youth Labor Market Participation in Africa

2015-33 : CHOUNET François, MAVEYRAUD Samuel : Corrélation des taux de change et régime d'étalon-or durant la Première Guerre mondiale

La coordination scientifique des Cahiers du GREThA est assurée par Emmanuel PETIT.
La mise en page est assurée par Julie VISSAGUET