Environment and Development Economics 18: 69–92 © Cambridge University Press 2012
doi:10.1017/S1355770X12000411

Forming coalitions to negotiate North–South climate agreements

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

JEAN-CHRISTOPHE PÉREAU
Université Montesquieu-Bordeaux IV, GREThA (UMR CNRS 5113), Avenue Léon Duguit, 33608 Pessac cedex, France.
E-mail: jean-christophe.pereau@u-bordeaux4.fr

Submitted November 8, 2011; revised June 21, 2012, August 29, 2012; accepted September 7, 2012; first published online October 26, 2012

ABSTRACT. This paper analyzes North-South negotiations over climate change abatement. We consider that northern countries have an incentive to negotiate over a transfer to the southern countries in exchange for their abatement efforts rather than reducing their emissions at home. We study the incentives for northern and southern countries to form negotiation-coalitions at each side of the bargaining table and the impact of these negotiation-coalitions on the final outcome. We show that the incentives can be separated into direct efficiency gains, as fixed costs savings, and indirect bargaining power gains. Depending on the relative values of these gains, we determine the equilibrium of the game. We also show that bargaining power gains encourage southern countries to negotiate separately while they encourage northern countries to unite, and that this hinders the formation of the grand coalition.

1. Introduction

In 1992 virtually all the countries in the world signed and ratified the United Nations Framework Convention on Climate Change. Five years later, the Kyoto Protocol was signed, engaging countries included in Annex I of this Protocol (OECD countries and economies in transition, hereinafter 'northern countries') to reduce their overall emission by five per cent in 2012 compared to 1990, but leaving non-Annex I countries ('southern

This work was carried out with the financial support of the CNRS-CSIC PICS program under the project SEIME.

countries') without abatement commitments.[1] Nevertheless, every year southern countries represent a more relevant part of the total emissions with the result that the current negotiation rounds are focusing on the best way to convince these countries to accept binding abatement efforts.[2] However, given their focus on development and the historical responsibility of northern countries, it is unlikely that southern countries will accept binding targets without some kind of transfer from northern countries. This 'transfer' could take the form of money (unlikely in large sums) or any other form of benefit (technology transfers or better conditions in other negotiations).

Coalitions play a significant role in these negotiations (Hampson and Hart, 1995). During all the negotiation rounds, countries have organized themselves into coalitions, such as the Umbrella Group (US and similarly minded countries during the 1990s), the European Union, the G77 and China (a large coalition of developing countries), the Least Developed Countries, the AOASIS (small island countries) and, more recently, the Rainforest Coalition (a coalition including almost all countries with rainforests, except Brazil). These coalitions have been relatively stable, although some variations have occurred over the already long history of negotiations. Nevertheless, a constant throughout the process has been that coalitions have been formed between northern or between southern countries.[3] This concept of 'coalition' differs from the one used in most of the literature on International Environmental Agreements (IEA), as in this literature a coalition is a group of countries that have decided to determine jointly their greenhouse gas (GHG) abatement effort (this is usually done by assuming that they have merged and now maximize their welfare jointly). In our framework, however, a coalition is a group of countries, either northern or southern, that have decided to join forces in international negotiations over a public good. In other words, for us, 'G77 and China' or the 'Rainforest Coalition' are coalitions, whereas they are not for most of the remaining literature on IEA. To avoid confusion, we will use hereafter the term 'negotiation-coalition' (NC) to refer to the coalitions set up for negotiation purposes that we are going to analyze (our concept is similar to the one used in Hampson and Hart (1995)).

[1] In other words, right from the beginning, the international community recognized that there were two types of fundamentally different countries: so-called Annex I countries (developed countries that accepted a 'historical responsibility' in the current levels of atmospheric CO_2) and non-Annex I countries (developing countries with a reduced 'historical responsibility' and an urgent need for development).

[2] Abatement made by one country, i.e its reduction of GHG emissions, is a public good that generates positive benefits for other countries, and this explains the difficulties encountered in reaching a global agreement for climate change (Carraro and Siniscalco, 1993; Barrett, 1994).

[3] In the official website of the United Nations Framework Convention on Climate Change, one can find the description of the different groupings of countries during the negotiations. See http://unfccc.int/parties_and_observers/parties/negotiating_groups/items/2714.php.

The question which we are looking to answer is whether or not there is a natural tendency for this type of negotiation to be carried out between NCs and whether the reason for this possible tendency is based on *efficiency gains* or *bargaining power gains* (or both). *Efficiency gains* refer to the direct gains which countries can expect when they negotiate together minus what they can expect negotiating separately. These gains can come from fixed-costs savings[4] but also from the different amounts of abatement to which the southern players rationally expect they are going to commit themselves while acting united instead of separated. If these gains are positive, countries will have an incentive to enter the negotiations as a NC. Nevertheless, there are reasons for negotiating together, or not, which are not defined by the comparison of the valuation of the two expected outcomes. In fact, even if efficiency gains did not exist, because the value of the expected outcomes under both scenarios is the same, countries would not be indifferent about negotiating together or separately. We refer to these additional reasons to form, or not, a NC as 'bargaining power' reasons. We show below that these bargaining power reasons are explained by the indirect gains obtained by countries not taking part in partial agreements (even though these partial agreements will never be the outcome of the negotiation).

Existing literature adds relatively little to the question of the formation of NCs. The bulk of the literature on IEA focuses on the formation of a single coalition, chiefly using the concept of internal and external stability from non-cooperative game theory (Finus, 2001). In the initial papers, all countries were assumed to be identical (Carraro and Siniscalco, 1993; Barrett, 1994), although this assumption has been relaxed in recent years (McGinty, 2007). Another branch of the literature initiated by Chander and Tulkens (1997) uses cooperative game theory to determine transfer schemes that ensure the non-emptiness of the core, but without modeling negotiations explicitly and without analyzing the role of bargaining power (see Breton *et al.* (2006) for a discussion of the differences between these two approaches). However, this literature does not analyze the negotiation process itself and the concept of 'coalition' used also differs, as already mentioned. In addition, contrary to our framework, these coalitions can be formed by northern and southern countries. Thus, our contribution to the literature on IEA is to focus on the bargaining *process* between northern and southern countries. More related to our approach, Rotillon *et al.* (1996) and Caparrós *et al.* (2004) model international negotiations on climate change abatement efforts using a bargaining procedure *à la* Rubinstein, in the latter case focusing on asymmetries of information. Both assume a unique North (see Carraro *et al.* (2005) or Harstad (2010) for a survey on bargaining theory and the use of this theory in the analysis of international negotiations), which reduces the relevance of the public good aspect of the problem because the northern coalition can never free ride. In addition, these papers do not distinguish, as we

[4] Acting together may also provide the opportunity to share abatement technologies (between the southern players) and this may reduce marginal abatement costs.

https://doi.org/10.1017/S1355770X12000411 Published online by Cambridge University Press

do below, between bargaining power and efficiency reasons to form NCs. This distinction has only been analyzed by Chipty and Snyder (1999) in the context of bilateral bargaining between one upstream firm and several downstream firms over a private good. Finally, our focus on NCs also distinguishes our paper from the now large literature that focuses on the formation of coalitions, such as Bloch (1996), Ray and Vohra (2001), Maskin (2003) or de Clippel and Serrano (2008). These papers take into account the interactions between coalitions when deciding the coalition structure to be formed, but they do not focus on the negotiations between coalitions and they are therefore not interested in NCs which are only set up for the negotiation process (bargaining power reasons are also not isolated in this literature).

This article is organized as follows. Section 2 presents a simple model for the negotiation between two northern countries and two southern countries. Both groups of countries can decide whether or not to form NCs and, once this decision has been taken, they engage in a negotiation *à la* Rubinstein to decide transfers from northern to southern players. Section 3 solves the model by backward induction, analyzing the incentives to form NCs and distinguishing between efficiency and bargaining power motivations. Section 4 discusses the implications of our model for the analysis of climate negotiations. Section 5 concludes.

2. The model

In the basic version of the model[5] we consider a world with two northern countries, denoted by 1 and 2, and two southern countries, denoted by 3 and 4. As described in the introduction, we reserve the term 'negotiation-coalition' for the strategic alliance between two northern (or two southern) countries before the bilateral negotiations take place. Since the empirical evidence briefly discussed in the introduction shows that northern and southern countries never join the same NC, we exclude the possibility of the formation of a NC between northern and southern players. When both northern countries decide to form a NC we use the notation (12), reserving the notation (34) for the case when both southern countries decide to form a NC. We use the term 'player' to describe the countries or NC that can eventually participate in the bilateral negotiations (i.e., the possible northern players are 1, 2 and (12) while the southern players are 3, 4 and (34)).

We denote by τ_N the northern NC coalition structure and by τ_S the southern NC structure, reserving the notation Ω_N for the set of all possible northern *NC structures* (with Ω_S for the southern NC). Since the northern and southern countries can only go separately or together in our simple model, Ω_N has only two possible elements, either $\tau_N = \{1, 2\}$ or $\tau_N = \{(12)\}$; and the set Ω_S has also only two possible elements, either southern countries negotiate separately, $\tau_S = \{3, 4\}$, or together, $\tau_S = \{(34)\}$.

[5] A version with n northern countries and m southern countries can be found in Caparrós and Péreau (2010).

https://doi.org/10.1017/S1355770X12000411 Published online by Cambridge University Press

For each pair (τ_N, τ_S) an *agreement structure* $\phi(\tau_N, \tau_S)$ is a set of bilateral agreements.[6] In our simplified framework, there are eight possible agreement structures: $\phi_1 = \{[(12), (34)]\}$, $\phi_2 = \{[(12), 3], [(12), 4]\}$, $\phi_3 = \{[1, (34)], 2\}$, $\phi_4 = \{[2, (34)], 1\}$, $\phi_5 = \{[1, 3], [1, 4], 2\}$, $\phi_6 = \{[2, 3], [2, 4], 1\}$, $\phi_7 = \{[1, 3], [2, 4]\}$, $\phi_8 = \{[1, 4], [2, 3]\}$. Square brackets are used to denote bilateral agreements and brackets are used, as before, to denote NC. For instance, $\phi_1 = \{[(12), (34)]\}$ means that the northern NC (12) reaches an agreement with the southern NC (34). In addition, there can be no agreement for each possible pair of NC structures. We call the set including all possible agreement structures $\Phi(\tau_N, \tau_S)$.

2.1. Game tree and assumptions

The game Γ is under perfect information and has four stages: (i) stage one, formation of the northern NC structure τ_N^*; (ii) stage two, formation of the southern NC structure τ_S^*; (iii) stage three, choice of the bilateral agreements to be negotiated between northern and southern players, i.e., the agreement structure ϕ^*; and (iv) stage four, the simultaneous negotiation over the transfer in the different bilateral agreements, following Rubinstein's alternating-offers protocol (Rubinstein, 1982).

The resolution of the game by backward induction, using the subgame perfect equilibrium (SPE) concept, is based on several assumptions:[7]

A1: The NC structure τ_N (respectively τ_S) is decided by the northern (southern) players maximizing the aggregated worth for the northern (southern) players
A2: For each pair (τ_N, τ_S), the agreement structure $\phi(\tau_N, \tau_S)$ with the highest aggregated worth is selected out of $\Phi(\tau_N, \tau_S)$.
A3: The tie-breaking rule in A1 and A2 is random selection.
A4: The distribution of the surplus of each bilateral agreement is negotiated following Rubinstein's alternating-offer procedure, with the northern

[6] Note that our 'agreements' are similar to the 'coalitions' in the standard IEA game, as they are formed by countries that have decided to determine jointly the abatement to be done. The main difference is that we allow for several agreements to co-exist and that we allow northern players to take part in more than one agreement, while there is a unique coalition in the basic IEA game (Carraro and Siniscalco, 1993) and countries can only participate in one agreement in the extensions based on coalition formation theory (Carraro, 2005). An additional difference, already mentioned, is that we model the whole process and determine the distribution of the worth within the agreement, although not within the NC.

[7] Most of the IEA literature is focused on the analysis of the stability of the coalitions. We do not analyze the stability of the agreements in our framework (the equivalent to the coalitions in the IEA literature). Nevertheless, it is important to highlight that the equilibrium agreements in our framework are SPE and, therefore, no player would have at any stage an incentive to deviate. In addition, once an agreement is reached there is no future period in our model where the agreement could be broken. As pointed out in our conclusion, analyzing this issue in more detail is an interesting future research topic.

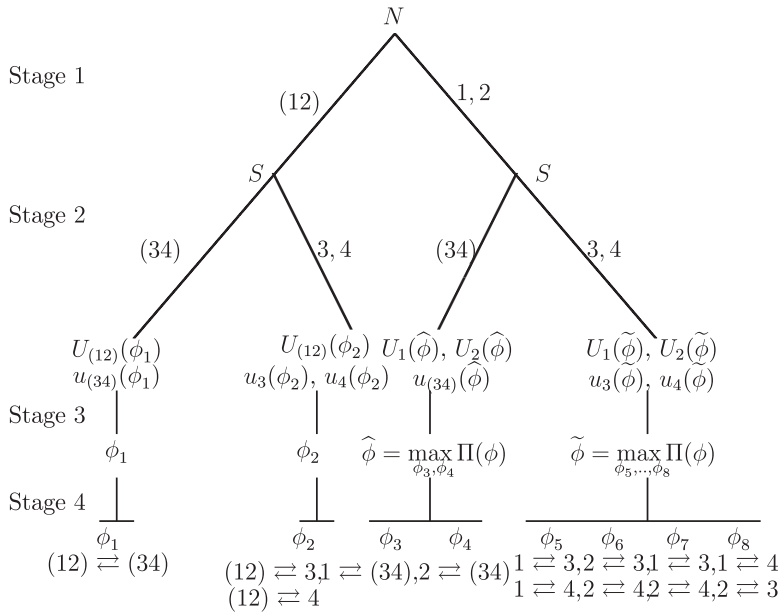

Figure 1. *Four-stage game tree with two northern and southern countries*

players proposing the first offer and under the belief that the optimal agreements are implemented in the remaining simultaneous negotiations.
A5: The southern players can participate in only one bilateral agreement.

We will now discuss the different assumptions while describing the decision tree of the game Γ shown in figure 1. In the first stage (analyzed in section 3.4), the northern players decide whether they prefer to go to the negotiations separately, as 1 and 2, or forming the NC (12). In the second stage (analyzed in section 3.3), the southern players decide whether to go to the negotiations separately, as 3 and 4, or jointly as the NC (34). In both cases, the southern players take into account the decision taken by the northern players in the previous stage. In these two stages, assumption A1 allows us to select the particular NC structures that the northern and the southern players will choose. More complex coalition formation procedures could also be relevant in a world with many countries (see, e.g., Ray and Vohra, 1999); however, in a world with only two northern countries and two southern countries Assumption A1 simply implies that when a NC is formed it compensates internally any potential losers as long as sufficient surplus is available (nevertheless, we do not analyze explicitly the distribution of the worth within the NC).

In the third stage (analyzed in section 3.2), the agreement structure to be negotiated at stage four is decided, out of all the possible agreement structures in each branch of the game tree. In the first branch starting from the LHS in figure 1, the northern NC structure is $\tau_N = \{(12)\}$ and the southern NC structure is $\tau_S = \{(34)\}$. In this case the only possible agreement is [(12), (34)] and the negotiation in the fourth stage will necessarily take

https://doi.org/10.1017/S1355770X12000411 Published online by Cambridge University Press

place between (12) and (34), represented by (12) $\rightleftarrows$ (34) in figure 1, i.e., the agreement structure selected at stage three is necessarily $\phi_1 = \{[(12), (34)]\}$.

In the second branch of the tree starting from the LHS there is also only one possibility, as by assumption A5 each southern player can only sign one agreement (although northern players can sign more than one agreement).[8] That is, in this case at stage three the NC structures are $\tau_N = \{(12)\}$ and $\tau_S = \{3, 4\}$, and the agreement structure selected is necessarily $\phi_2 = \{[(12), 3], [(12), 4]\}$. In the third branch from the LHS, where $\tau_N = \{1, 2\}$ and $\tau_S = \{(34)\}$, there are two possible agreement structures, either $\phi_3 = \{[1, (34)], 2\}$ or $\phi_4 = \{[2, (34)], 1\}$. We use assumption A2 to select the agreement structure generating the largest surplus to be negotiated in stage four. The reason for this assumption is that the higher the aggregated worth of the agreement structure, the higher will be the aggregated worth which the players will get from the negotiation in the next stage.[9] We call the agreement structure selected $\widehat{\phi}$. In the last branch from the LHS, where $\tau_N = \{1, 2\}$ and $\tau_S = \{3, 4\}$, we assume again that the agreement structure generating the largest surplus will be selected amongst $\phi_5, \ldots, \phi_8$ and we call this agreement structure $\widetilde{\phi}$. Assumption A3 ensures that only one structure is selected when several structures yield the maximum worth.

In the fourth stage (analyzed in section 3.1), the different bilateral agreements in the agreement structure are negotiated simultaneously. In the first branch of the game tree starting from the LHS, the only agreement negotiated is [(12), (34)]. The outcome of this negotiation is a payoff for the northern player involved, $U_{(12)}$, and a payoff for the southern player involved, $u_{(34)}$. In the second branch, the two agreements in $\phi_2 = \{[(12), 3], [(12), 4]\}$ are negotiated simultaneously. The outcome of these two negotiations determines the payoffs for the different players involved: $U_{(12)}$, u_3 and u_4. In the third branch the bilateral agreements in $\widehat{\phi}$ are negotiated simultaneously (and $\widetilde{\phi}$ in the fourth branch). In all cases the negotiations in this fourth stage take place following Rubinstein's offer

[8] Allowing southern countries to reach partial agreements with different northern countries would complicate the expressions derived below without any significant gain.

[9] To justify assumption A2 we will focus on the third branch, but the need for an assumption instead of a complete analysis of the different cases becomes clear in the fourth branch. Our assumption implies that ϕ_3 will be selected in stage three if the aggregate payoff of ϕ_3 is higher than that of ϕ_4. This relationship will hold either because the gains obtained by the members of the agreement are larger if 1 signs an agreement with (34) than if 2 signs it, in which case (34) will have an incentive to sign an agreement with 1, or because player 2 obtains a large benefit from staying outside of the agreement, which will see its interest in bargaining a deal by itself reduced, which will again favor an agreement between 1 and (34). Although under some conditions 1 may prefer that 2 sign the agreement and 2 that 1 sign it, as long as the aggregate payoff of ϕ_3 is higher than that of ϕ_4, player 2 could compensate player 1 for signing the agreement with (34). Analyzing these compensations, especially in the fourth branch, would significantly complicate the analysis in section 3.2, without any real gain for the main goal of the paper (to analyze NCs).

and counter-offer protocol (Rubinstein, 1982), with the northern players proposing the first offer (assumption A4). If there are two simultaneous negotiations (i.e., in ϕ_2, ϕ_5, ϕ_6, ϕ_7 and ϕ_8), both negotiations take place under the assumption that the optimal agreement is reached in the other negotiation. It means that in every negotiation the northern and the southern players believe that their failure to reach an agreement would not affect the other optimal agreement, which is negotiated simultaneously. It is unclear whether simultaneous or sequential bargaining better describes real-life negotiations and, as we point out at the end of the paper, comparing these two bargaining protocols is an interesting avenue for future research (see also Aghion *et al.*, 2007). Nevertheless, we believe that simultaneous bargaining better captures international negotiations on climate change, since currently several negotiations are taking place at the same time without one negotiation waiting until the previous one has finished (the US with China outside of the Kyoto Protocol, Europe with several developing countries within the Clean Development Mechanism of the Kyoto Protocol, Norway with Indonesia on Reducing Emissions from Deforestation and Degradation, etc.).

2.2. Payoff functions

We denote the abatement performed by the southern player j by q_j. The aggregated abatement performed by the southern players is given by the vector $Q = \{q_3, q_4\}$ if they remain independent and $Q = \{q_{(34)}\}$ if they form a NC (we also use Q to denote the sum of the components of the vector where appropriate). The abatement performed by the northern players i is denoted by $\bar{q}_i$ and the aggregated abatement by the northern players is $\overline{Q} = \{\bar{q}_1, \bar{q}_2\}$ when they act independently and $\overline{Q} = \{\overline{q}_{(12)}\}$ if they form a NC. The total abatement is then $\overline{Q} + Q$.

We further assume that the emission abatements made by the northern players are an optimal reaction to the abatement of the southern players, after the agreements discussed below have been signed. That is, we assume that northern players act as Stackelberg leaders and decide their abatement levels according to the reaction function $\overline{q}_i = f_i(Q)$ with $\frac{\partial \overline{q}_i}{\partial Q} < 0$.

The gross payoff of the northern players is given by the function $V_i(Q, \overline{Q}, \overline{q}_i)$ which depends on the total abatement done by the South, Q; the total abatement done by the North, $\overline{Q}$; and the abatement done at home by the northern player i, $\overline{q}_i$. That is, all costs and benefits, except the transfer T_{ij}, are included in V_i. As $\overline{q}_i$ and $\overline{Q}$ are functions of total southern abatement Q, we can omit them in the gross payoff expressions and use $V_i(Q)$. Slightly abusing notation whenever we are interested in detailing the components of the vector Q, we write $V_i(q_3, q_4)$ instead of $V_i(\{q_3, q_4\})$. Hence the net payoff function for the northern players $i = 1, 2, (12)$ is:

$$U_i = V_i(Q) - \sum_{j \in J_i} T_{ij}, \tag{1}$$

where J_i is the set of the bilateral agreement signed by i with the southern players j. The set J_i can consist in one agreement, {3}, {4} or {(34)}; in two, {3, 4}; or in none, {∅}.

https://doi.org/10.1017/S1355770X12000411 Published online by Cambridge University Press

Although this is not necessary to derive the results below, in order to facilitate the interpretation and to connect with the IEA literature, we assume that the value function takes the form $V_i(.) = B_i(\overline{Q}, Q) - C_i(\overline{q}_i)$, where i benefits from the total emission reductions (B_i) but only bears the cost (C_i) of the abatement that it performs at home ($\frac{\partial C_i}{\partial \overline{q}_i} > 0$). Nevertheless, when a northern player signs an agreement with some (at least one) southern player j, it also bears the cost of the transfer T_{ij} for each one of the agreements it signs. The abatement q_j performed by a southern player j in exchange for the transfer granted by the northern player i, benefits i directly ($\frac{\partial B_i}{\partial q_j} \geq 0$) and reduces the abatement to be done by i at home (i.e., $\frac{\partial C_i}{\partial \overline{q}_i} \frac{\partial \overline{q}_i}{\partial q_j} \leq 0$).

For southern players we also assume that all costs and benefits except the transfer are included in v_j and we assume that $v_j(Q, \overline{Q}, q_j) = b_j(\overline{Q}, Q) - c_j(q_j)$, where b_j stands for the benefits from the total abatement obtained by j and c_j for its abatement costs (we assume $\frac{\partial b_j}{\partial q_j} \geq 0$ and $\frac{\partial c_j}{\partial q_j} \geq 0$). We simplify the notation of v_j by using $v_j(Q)$, as $\overline{Q}$ is a function of Q and the latter vector incorporates the information of the total abatement done by the southern players and the amount done at home by southern player j, i.e., q_j (as for V, when detailing the different components of Q we write $v_i(q_3, q_4)$). As by assumption A5 southern players can only be involved in one bilateral agreement, the net payoff function for a southern player j is given by:

$$u_j = v_j(Q) + T_{ij}. \tag{2}$$

The total gross payoff for all the players in each agreement structure ϕ_k is denoted by $\Pi(\phi_k) = \sum_i V_i(Q(\phi_k)) + \sum_j v_j(Q(\phi_k))$.

We also assume that northern and southern players are fundamentally different and in particular that there is at least one level of abatement $q_j > q_j^d$ for which $V_i(q_j, q_l^*) - V_i(q_j^d, q_l^*) > v_j(q_j^d, q_l^*) - v_j(q_j, q_l^*)$ for any pair of a northern player i and a southern player j, where q_j^d is the abatement that player j would perform when maximizing its payoff independently (in both cases under the assumption that the other southern player, if there is one, does its optimal abatement). This can happen because an increase in southern abatement benefits more the northern than the southern player, through B_i and b_j, or because the reduction in costs for the North, $\frac{\partial C_i}{\partial \overline{q}_i} \frac{\partial \overline{q}_i}{\partial q_j} \leq 0$, is larger than the increase in the costs in the South (this latter case is probably the most relevant one for the climate change game). Of course, a combination of the two reasons just described can also explain why northern players are 'buyers' of the abatement made by southern players. This assumption implies that the northern player is interested in granting the southern player a transfer in order to increase the abatement of the southern player, and that all southern countries will sign an agreement in equilibrium.

We also assume that northern and southern players have different discount factors, δ_N and δ_S respectively, although all northern (southern)

https://doi.org/10.1017/S1355770X12000411 Published online by Cambridge University Press

players share the same discount factor, with $1 > \delta_N, \delta_S > 0$. That is, in addition to the differences in the valuation functions described in the previous paragraph, we are assuming that northern players (respectively southern players) are relatively similar one to another in their valuation of the time needed to reach an agreement, but that northern and southern players have different perceptions of the costs of delay.[10]

3. Results

3.1. Bilateral bargaining (stage four)

Since we solve the model by backward induction we analyze first the last stage of the game. As indicated in assumption A4, the transfer is determined by the Rubinstein Bargaining Solution (RBS). At this stage, the two parts of each bilateral agreement are already known and so is the abatement effort that the South must perform if the agreement is reached.

To illustrate our framework, we focus initially on a particular agreement structure $\phi_5 = \{[1, 3], [1, 4], 2\}$. Let us consider the bargaining process between 1 and 3, both assuming that 1 and 4 will reach an optimal agreement at the same time. In case of agreement,[11] we have

$$U_1^a = V_1(q_3^*, q_4^*) - T_{13} - T_{14}^*$$
$$u_3^a = v_3(q_3^*, q_4^*) + T_{13},$$

while in case of non-agreement, we have

$$U_1^{na} = V_1(q_3^d, q_4^*) - T_{14}^*$$
$$u_3^{na} = v_3(q_3^d, q_4^*),$$

with $q_3^d = \arg\max_x v_3(x, q_4^*)$, that is, in case of disagreement country 3 determines its deviation abatement level maximizing its own payoff function (under the assumption that player 4 performs the optimal abatement level coming out from its agreement with player 1).

We define the net payoff functions (R) from an agreement at period τ as:

$$R_1(T_{13}, \tau) = \delta_N^\tau(U_1^a - U_1^{na}) = \delta_N^\tau(V_1(q_3^*, q_4^*) - V_1(q_3^d, q_4^*) - T_{13})$$
$$R_3(T_{13}, \tau) = \delta_S^\tau(u_3^a - u_3^{na}) = \delta_S^\tau(v_3(q_3^*, q_4^*) - v_3(q_3^d, q_4^*) + T_{13}).$$

Let $T_{13}^{(1)}$ be the offer (or the counter-offer) made by player 1 over the transfer T_{13} to be received by 3, and $T_{13}^{(3)}$ the counter-offer made by 3 (although the counter-offer is made by 3 it is still a transfer from player 1

[10] Let us note that discount factors in bargaining models are not necessarily related to financial discount rates but capture all the reasons why one player is more or less impatient about reaching an agreement (Muthoo, 1999).

[11] The abatement performed by the southern players depends on the agreement structure, i.e $q_3(\phi_5)$ and $q_4(\phi_5)$ but for simplicity we omit these terms.

https://doi.org/10.1017/S1355770X12000411 Published online by Cambridge University Press

to 3). Rubinstein's bargaining procedure mentioned above is as follows: 1 makes an offer of a transfer $T_{13}^{(1)}$ to 3. If 3 accepts then the bargaining is over. If 3 declines the offer, 3 proposes a counteroffer $T_{13}^{(3)}$. If 1 accepts this offer the agreement is struck, and if 1 does not accept, 1 makes a new counter-offer. The alternating-offers procedure continues until an agreement is reached. The RBS that we are looking for is the unique SPE given by the following two conditions (Rubinstein, 1982; Muthoo, 1999): $R_1(T_{13}^{(3)}, 0) = R_1(T_{13}^{(1)}, 1)$ and $R_3(T_{13}^{(1)}, 0) = R_3(T_{13}^{(3)}, 1)$. This yields:

$$V_1(q_3^*, q_4^*) - V_1(q_3^d, q_4^*) - T_{13}^{(3)} = \delta_N[V_1(q_3^*, q_4^*) - V_1(q_3^d, q_4^*) - T_{13}^{(1)}]$$
$$v_3(q_3^*, q_4^*) - v_3(q_3^d, q_4^*) + T_{13}^{(1)} = \delta_S[v_3(q_3^*, q_4^*) - v_3(q_3^d, q_4^*) + T_{13}^{(3)}].$$

The first equation asserts that 1 is indifferent in terms of expected payoffs as regards accepting 3's offer $T_{13}^{(3)}$ in the current period or rejecting it and making in the following period the counteroffer $T_{13}^{(1)}$ that will be accepted by 3 (as the negotiation is stationary, comparing period 1 and 2 is the same as comparing periods 0 and 1). The second equation reflects the same indifference for 3. Since we assume (assumption A4) that 1 makes the first offer, the equilibrium transfer is

$$T_{13}^* = \varphi[V_1(q_3^*, q_4^*) - V_1(q_3^d, q_4^*)] + (1 - \varphi)[v_3(q_3^d, q_4^*) - v_3(q_3^*, q_4^*)],$$

with $\varphi = \frac{\delta_S(1-\delta_N)}{1-\delta_N\delta_S}$. As shown by Rubinstein (1982), under perfect information the equilibrium offer is proposed by player 1 in the first round and it is immediately accepted by player 3, i.e., $T_{13}^{(1)} = T_{13}^*$ (delays in reaching an agreement can only take place with incomplete information (Caparrós *et al.*, 2004)).

The transfer $T_{13}^* > 0$ to be paid by 1 to 3 depends positively on the marginal contribution of 3 to the surplus obtained by 1 (first term in brackets) and on the free-riding payoff of 3 (second term in brackets).

At the same time, a similar program between 1 and 4 yields the agreement payoff functions $U_1^a = V_1(q_3^*, q_4^*) - T_{13}^* - T_{14}$ and $u_4^a = v_4(q_3^*, q_4^*) + T_{14}$, and the impasse point in case of a failure of the negotiation $U_1^{na} = V_1(q_3^*, q_4^d) - T_{13}^*$ and $u_4^{na} = v_4(q_3^*, q_4^d)$, with $q_4^d = \arg\max_x v_4(x, q_3^*)$. Hence, simultaneously 1 reaches an agreement with 4 based on a similar transfer:

$$T_{14}^* = \varphi[V_1(q_3^*, q_4^*) - V_1(q_3^*, q_4^d)] + (1 - \varphi)[v_4(q_3^*, q_4^d) - v_4(q_3^*, q_4^*)].$$

The resulting payoffs for all the players are

$$U_1(\phi_5) = V_1(q_3^*, q_4^*) - T_{13}^* - T_{14}^*$$
$$U_2(\phi_5) = V_2(q_3^*, q_4^*)$$
$$u_3(\phi_5) = v_3(q_3^*, q_4^*) + T_{13}^*$$
$$u_4(\phi_5) = v_4(q_3^*, q_4^*) + T_{14}^*.$$

To deal with all the other agreement structures, let us assume, without loss of generality, that the northern player i is bargaining with the southern

https://doi.org/10.1017/S1355770X12000411 Published online by Cambridge University Press

player j. An agreement between them yields:

$$U_i^a = V_i(q_j^*, Q_{-j}^*) - T_{ij} - \sum_{l \in J_i/j} T_{il}^* \tag{3}$$

$$u_j^a = v_j(q_j^*, Q_{-j}^*) + T_{ij}, \tag{4}$$

where $Q_{-j} = Q - q_j$ denotes the abatement made by the southern player not involved in the bilateral agreement, if there is any (i.e., if player j is for example 3 then Q_{-j} is q_4, but if j is (34) then Q_{-j} is empty). In case of disagreement, southern players will implement their deviation abatement $q_j^d = \arg\max_x v_j(x, Q_{-j}^*)$, under the assumption that the other southern player will implement the abatement coming out from its optimal agreement. Hence, the impasse point in case of non-agreement is:

$$U_i^{na} = V_i(q_j^d, Q_{-j}^*) - \sum_{l \in J_i/j} T_{il}^* \tag{5}$$

$$u_j^{na} = v_j(q_j^d, Q_{-j}^*). \tag{6}$$

Since we assume (assumption A4) that i makes the first offer, we obtain the following lemma.

Lemma 1. *The equilibrium transfer is defined by:*

$$T_{ij}^* = \varphi[V_i(Q^*) - V_i(q_j^d, Q_{-j}^*)] + (1-\varphi)[v_j(q_j^d, Q_{-j}^*) - v_j(Q^*)]. \tag{7}$$

As for the particular case discussed above, the transfer T_{ij}^* to be paid by i to j depends positively on the marginal contribution of j to the surplus obtained by i (first term in brackets) and on the free-riding payoff of j (second term in brackets). That is, the more i gets out of the agreement and the more j benefits from deviating from the agreement, the more willing i is to pay. As before, this offer is proposed in the first period and immediately accepted. The resulting payoffs for i and j are, for every agreement structure ϕ_k:

$$U_i(\phi_k) = V_i(Q^*) - T_{ij}^* - \sum_{l \in J_i/j} T_{il}^* = V_i(Q^*) - \sum_{j \in J_i} T_{ij}^*$$

$$u_j(\phi_k) = v_j(Q^*) + T_{ij}^*.$$

3.2. Agreement structure and abatement (stage three)

In stage three, for each agreement in every agreement structure ϕ_k, both players involved know that they will distribute the worth on stage four according to the RBS. Thus, they are interested in maximizing the worth of the bilateral agreement by choosing the level of abatement q_j^* that maximizes the surplus. Hence for each possible bilateral agreement within every

https://doi.org/10.1017/S1355770X12000411 Published online by Cambridge University Press

agreement structure ϕ_k, we have

$$q_j^*(\phi_k) = \arg\max_x [V_i(x, Q_{-j}^*) + v_j(x, Q_{-j}^*)]. \tag{8}$$

This gives the abatement effort that would be performed for each possible bilateral agreement (recall that each southern player j can only participate in one agreement within ϕ_k). As stated during the presentation of the model, we assume that $V_i(q_j^*, Q_{-j}^*) - V_i(q_j^d, Q_{-j}^*) > v_j(q_j^d, Q_{-j}^*) - v_j(q_j^*, Q_{-j}^*)$ with $q_j^d = \arg\max_x v_j(x, Q_{-j}^*)$.

Substituting in $\Pi(\phi_k) = \sum_i V_i(Q(\phi_k)) + \sum_j v_j(Q(\phi_k))$, we obtain the total value of each agreement structure ϕ_k. As stated above, for the first two branches on the LHS of the game tree there is only one agreement structure so that no further selection is needed. Nevertheless, for the two branches on the RHS we use assumption A2 to select the agreement structure ϕ_k with the largest aggregated value for the players involved, i.e.:

$$\widehat{\phi} = \arg\max_{\phi_3, \phi_4} \Pi(\phi)$$

$$\widetilde{\phi} = \arg\max_{\phi_5, .., \phi_8} \Pi(\phi),$$

where $\widehat{\phi}$ is the agreement structure selected for $\tau_N = \{1, 2\}$ and $\tau_S = \{(34)\}$ and $\widetilde{\phi}$ for $\tau_N = \{1, 2\}$ and $\tau_S = \{3, 4\}$. This selection procedure ensures that, for each branch of the game tree in figure 1, only one agreement structure is selected and that an abatement effort in case of agreement is selected for each potential bilateral agreement within each agreement structure.

3.3. Formation of a southern negotiation-coalition (stage two)

We now analyze the implications of the formation of a southern NC τ_S at stage two. In particular we analyze the case in which the southern countries 3 and 4 face the question of whether or not to form a NC (34). This is the lower part in the tree shown in figure 1. Equations (1)–(8) have to be calculated for the two countries acting separately and for the NC (note that the impasse point for the NC (34) differs from that of the separated countries 3 and 4).

The formation of a southern NC at stage two is profitable under the condition that a southern NC (34) is able to get a higher payoff than two southern countries separately (this condition is equivalent to assumption A1 in a world with two countries):

$$u_{(34)} > u_3 + u_4. \tag{9}$$

Let us start by considering a particular example. Suppose that, if southern countries form a NC, (34), the outcome of stages three and four would be that (34) will sign an agreement with the northern player 1 (i.e., that agreement structure ϕ_3 will be reached) and that, if southern countries act separately, southern player 3 signs an agreement with northern player 1, while 4 signs an agreement with 2 (i.e., that agreement structure ϕ_7 will

https://doi.org/10.1017/S1355770X12000411 Published online by Cambridge University Press

be reached if southern countries act separately). In this case, condition (9) reads:

$$v_{(34)}(q^*_{(34)}) + T^*_{1(34)} > v_3(q^*_3, q^*_4) + T^*_{13} + v_4(q^*_3, q^*_4) + T^*_{24},$$

with T^*_{ij} given by the equivalent to equation (7). We can further rewrite this condition as

$$SE + NE + (NBP - SBP) > 0$$

with

$$\begin{aligned} SE = {} & [v_{(34)}(q^*_{(34)}) + Dv_{(34)}(q^d_{(34)})] \\ & - [v_3(q^*_3, q^*_4) + Dv_3(q^d_3, q^d_4)] - [v_4(q^*_3, q^*_4) + Dv_4(q^d_3, q^d_4)] \qquad (10) \\ NE = {} & [V_1(q^*_{(34)}) - V_1(q^d_{(34)})] \\ & - [V_1(q^*_3, q^*_4) - V_1(q^d_3, q^d_4)] - [V_2(q^*_3, q^*_4) - V_2(q^d_3, q^d_4)] \qquad (11) \\ SBP = {} & D[v_3(q^d_3, q^*_4) - v_3(q^d_3, q^d_4)] + D[v_4(q^*_3, q^d_4) - v_4(q^d_3, q^d_4)] \qquad (12) \\ NBP = {} & [V_1(q^d_3, q^*_4) - V_1(q^d_3, q^d_4)] + [V_2(q^*_3, q^d_4) - V_2(q^d_3, q^d_4)], \qquad (13) \end{aligned}$$

and $D = \frac{1-\varphi}{\varphi} = \frac{1-\delta_S}{\delta_S(1-\delta_N)}$.

From equation (10), SE represents the '*southern efficiency gain*', that is, the gain which the southern players can expect when they act together (first square bracket) minus what they can expect acting separately (second and third square brackets). SE focuses on direct gains[12] obtained by the different southern players. Note that to compare the two situations the southern players add what they would obtain in the event of an agreement and their net benefit in the event of a disagreement, since their expected payoff is a function of both eventualities. From equation (11), NE represents the '*northern efficiency gain*', the direct gain which the northern players would obtain from an agreement with the NC minus the gain they would obtain when dealing with two separated southern players. Unlike for SE, for NE only the net benefit for the northern players from the agreement is relevant (i.e., the three square brackets show the difference before and after the different agreements). The reason is that the southern players, when deciding whether or not to form a NC, are only interested in the part of the benefit for the northern players that is relevant to them, i.e., the part that explains the transfer they can expect to obtain.

We now turn to the most interesting terms for our analysis, NBP and SBP. The first square bracket in equation (12) shows the indirect gain that an agreement between the northern player 2 and the southern player 4 induces for the southern player 3 as long as 3 signs no agreement (the

[12] As mentioned in the introduction, direct net gains can come from fixed-costs savings, from the different amounts of abatement and from the opportunity to share abatement technologies (between the southern players). We call these benefits 'efficiency gains' as most of them fall into this category.

https://doi.org/10.1017/S1355770X12000411 Published online by Cambridge University Press

second square bracket shows the indirect gain for 4 of an agreement between 1 and 3). The first square bracket in equation (13) shows the indirect gain which the agreement between 2 and 4 induces for the northern player 1 if this player does not sign any agreement (the second square bracket shows the indirect gain for 2 of an agreement between 1 and 3). We refer to *SBP* and *NBP* as 'bargaining power' terms and, as all the gains involved are positive, we have $NBP \geq 0$ and $SBP \geq 0$. These gains are irrelevant in equilibrium (where all players sign their optimal agreements) but, as we show below, they determine whether or not forming a NC increases bargaining power.

Finally, D collects the discount terms and has the following properties: (i) for a given δ_S, the larger δ_N the larger D (ii) for a given δ_N, the larger δ_S the smaller D (iii) if $\delta_N = \delta_S$, the larger the discount factor the smaller D, and (iv) $D > 0$. In other words, a decrease in D implies more patient players or a more patient southern player compared to the northern player.[13]

To deal now with all possible agreement structures, let us assume, without loss of generality, that the outcome of stages three and four is that the northern player $i = 1, 2$, (12) signs an agreement with the southern NC (34) if the South acts united, while the northern player $k = 1, 2$, (12) (respectively p) signs an agreement with the southern country $j = 3, 4$ (respectively $l \neq j$) when they are separate. Thus, condition (9) yields:

$$v_{(34)}(q^*_{(34)}) + T^*_{i(34)} > v_j(q^*_3, q^*_4) + T^*_{kj} + v_l(q^*_3, q^*_4) + T^*_{pl}, \tag{14}$$

with $T^*_{i(34)}$, T^*_{kj} and T^*_{pl} given by the equivalent to equation (7). This expression can be manipulated[14] to obtain a form in which the motives for the formation of a NC can be distinguished. To this end, we define:

$$\underline{SE} = [v_{(34)}(q^*_{(34)}) + Dv_{(34)}(q^d_{(34)})] - [v_l(q^*_3, q^*_4) + Dv_l(q^d_3, q^d_4)] - [v_j(q^*_3, q^*_4) + Dv_j(q^d_3, q^d_4)] \tag{15}$$

$$\underline{NE} = [V_i(q^*_{(34)}) - V_i(q^d_{(34)})] - [V_p(q^*_3, q^*_4) - V_p(q^d_3, q^d_4)] - [V_k(q^*_3, q^*_4) - V_k(q^d_3, q^d_4)] \tag{16}$$

$$\underline{SBP} = D[v_l(q^d_l, q^*_j) - v_l(q^d_3, q^d_4)] + D[v_j(q^d_j, q^*_l) - v_j(q^d_3, q^d_4)] \tag{17}$$

$$\underline{NBP} = [V_p(q^d_l, q^*_j) - V_p(q^d_3, q^d_4)] + [V_k(q^d_j, q^*_l) - V_k(q^d_3, q^d_4)]. \tag{18}$$

[13] In most economic analysis, discount factors are assumed to be smaller for southern players than for northern players (or discount rates larger). Nevertheless, the terms included in D are not necessarily related to financial discount rates but incorporate all 'political' reasons that may explain why one player is more impatient than another to reach an agreement. For example, internal political pressure may be stronger in northern players, implying that they are more impatient to reach an agreement, and this would yield $\delta_N < \delta_S$.

[14] We have added and subtracted in (9) $Dv_j(q^d_3, q^d_4)$, $Dv_l(q^d_3, q^d_4)$, $V_k(q^d_3, q^d_4)$ and $V_p(q^d_3, q^d_4)$.

https://doi.org/10.1017/S1355770X12000411 Published online by Cambridge University Press

Based on these terms, condition (19) shows that southern countries have an incentive to form a NC if

$$\underline{SE} + \underline{NE} + (\underline{NBP} - \underline{SBP}) > 0. \tag{19}$$

An increase in the efficiency terms $\underline{SE}$ and $\underline{NE}$ will always encourage southern countries to form a NC. Nevertheless, even if the sum of the efficiency terms is positive, the bargaining power terms can invert the condition above. In order to focus only on the bargaining power terms, condition (19) shows even if forming a southern NC has no efficiency gains (i.e., $\underline{SE} = \underline{NE} = 0$), southern countries have a incentive to form a NC if $\underline{NBP} > \underline{SBP}$.

That is, even in the eventuality that $\underline{SE} = \underline{NE} = 0$ (i.e., that the countries involved in the agreements obtain the same direct gain by passing from the initial situation to the final situation with all the optimal agreements in place), the southern countries may have incentives to form a NC, depending on the relative values of the bargaining power terms $\underline{NBP} \geq 0$ and $\underline{SBP} \geq 0$. In other words, southern countries have an incentive to form a NC if, while acting separately, a partial agreement induces a larger indirect gain for northern countries that do not participate than for southern countries that are not part of the agreement (taking into account in the latter case the impact of D). The intuition for this result is that, by acting together, southern countries can expect to obtain part of the benefit obtained by a northern NC not participating in the agreement while they know that the additional rent that they would potentially need to share, the benefit obtained by a non-participating southern NC, is comparatively small.

3.4. Formation of a northern negotiation-coalition (stage one)

This section analyzes the implications of the formation of a northern NC at stage one. The formation of a northern NC will be profitable if:

$$U_{(12)} > U_1 + U_2. \tag{20}$$

The main difference from the previous stage is that now we have to consider that northern players can sign multiple agreements, unlike southern players. Without loss of generality, we assume that when the northern countries act separately, 1 signs an agreement with the set J_1 of southern players and 2 signs an agreement with the set J_2 of southern players. The sets J_1 or J_2 can be empty and are disjoint such that $J_1 \cap J_2 = \emptyset$. The set J_1 can cover one agreement $\{3\}$, $\{4\}$ or $\{(34)\}$, two agreements $\{3, 4\}$ or no agreement $\{\emptyset\}$. The same holds for J_2. For the northern NC, we can have one agreement $\{(34)\}$ or two agreements $\{3, 4\}$. We use the notation $|J_i|$ to denote the number of elements in the set. Thus, we can rewrite (20) as:

$$V_{(12)}(q_w^*) - \sum_{w \in J_{(12)}} T^*_{(12)w} > V_1(q_z^*, q_y^*) - \sum_{z \in J_1} T^*_{1z} + V_2(q_z^*, q_y^*) - \sum_{y \in J_2} T^*_{2y}, \tag{21}$$

https://doi.org/10.1017/S1355770X12000411 Published online by Cambridge University Press

with $T^*_{(12)w}$, T^*_{1z} and T^*_{2y} given by the equivalent to equation (7). To get the motives that drive the formation of a northern NC, we define:[15]

$$\overline{SE} = \sum_{w\in J_{(12)}} [v_w(q^*_w) - v_w(q^d_w)] - \sum_{z\in J_1}[v_z(q^*_z, q^*_y) - v_z(q^d_z, q^d_y)] - \sum_{y\in J_2}[v_y(q^*_z, q^*_y) - v_y(q^d_z, q^d_y)] \quad (22)$$

$$\overline{NE} = \frac{(1-\varphi|J_{(12)}|)V_{(12)}(q^*_w) - (1-\varphi|J_1|)V_1(q^*_z, q^*_y) - (1-\varphi|J_2|)V_2(q^*_z, q^*_y)}{1-\varphi} + \frac{\varphi}{1-\varphi}\left(\sum_{w\in J_{(12)}} V_{(12)}(q^d_w) - \sum_{z\in J_1} V_1(q^d_z, q^d_y) - \sum_{y\in J_2} V_2(q^d_z, q^d_y)\right) \quad (23)$$

$$\overline{SBP} = \sum_{z\in J_1}(v_z(q^d_z, q^*_y) - v_z(q^d_z, q^d_y)) + \sum_{y\in J_2}(v_y(q^d_y, q^*_z) - v_y(q^d_z, q^d_y)) \quad (24)$$

$$\overline{NBP} = \frac{\sum_{z\in J_1}(V_1(q^d_z, q^*_y) - V_1(q^d_z, q^d_y)) + \sum_{y\in J_2}(V_2(q^d_y, q^*_z) - V_2(q^d_z, q^d_y))}{D}. \quad (25)$$

The fact that each northern player can sign multiple agreements (unlike southern players) explains the differences from the definitions of $\underline{SE}$, $\underline{NE}$, $\underline{NBP}$ and $\underline{SBP}$ in the previous section. If each northern player signs an agreement with only one player ($|J_1| = |J_2| = |J_{(12)}| = 1$), equations (22)–(25) are similar to (15)–(18): (i) $\overline{SE}$ and $\overline{NE}$ are focused on direct gains while $\overline{NBP}$ and $\overline{SBP}$ are given by indirect gains (ii) D appears in $\overline{NE}$ and $\overline{NBP}$ and not in $\overline{SE}$ and $\overline{SBP}$ (iii) $\overline{SE}$ focuses on incremental benefits as does $\underline{NE}$ and (iv) $\overline{NE}$ adds up agreement and disagreement outcomes as does $\underline{SE}$. Interpretations are therefore symmetrical.

Based on these terms, condition (26) shows that northern countries have an incentive to form a NC if

$$\overline{NE} + \overline{SE} - (\overline{NBP} - \overline{SBP}) > 0. \quad (26)$$

As in condition (19), we obtain that an increase in the efficiency terms tends to favor a united North. However, the bargaining power terms ($\overline{NBP} - \overline{SBP}$) appear in condition (26) with the opposite sign. The interpretation of $\overline{NBP} \geq 0$ and $\overline{SBP} \geq 0$ is similar to that of $\underline{NBP}$ and $\underline{SBP}$, although now the expressions have to take into account the fact that northern players can sign agreements with different southern players. Focusing again on the eventuality that $\overline{SE} = \overline{NE} = 0$, the incentive for the northern countries to form a NC depends on the relative values of the bargaining terms $\overline{NBP}$

[15] By adding and subtracting $\sum_{z\in J_1} v_z(q^d_z, q^d_y)$, $\sum_{y\in J_2} v_y(q^d_z, q^d_y)$, $D^{-1}\sum_{z\in J_1} V_1(q^d_z, q^d_y)$ and $D^{-1}\sum_{y\in J_2} V_2(q^d_z, q^d_y)$ in (21).

https://doi.org/10.1017/S1355770X12000411 Published online by Cambridge University Press

and $\overline{SBP}$. Northern countries have an incentive to form a NC if $\overline{SBP} > \overline{NBP}$ (i.e., if a partial agreement induces a larger indirect gain in southern players that do not participate than in northern players that are not part of the partial agreement). That is, if bargaining power favors one of the sides of the negotiations acting in a united way, the general tendency will go in the opposite direction for the other side. The intuition is the same as for the formation of a southern NC, but since we are comparing the potential gains from non-participating countries from the northern perspective the balance that favors forming a NC is the opposite. In addition, as in our framework the grand coalition can only be the outcome of a negotiation between a northern and a southern NC, bargaining power terms make the formation of the grand coalition more difficult.

3.5. Equilibrium

Collecting previous results, we can write the following proposition that describes the equilibrium of the game and shows that the equilibrium is unique:

Proposition 1. *The game* Γ *has a unique SPE in which transfers are given by Lemma 1. The equilibrium NC structure and agreement structure are*

1. $(\tau_N^*; \tau_S^*) = \{(12); (34)\}$ *with* $\phi_1^* = \{[(12), (34)]\}$ *iif*

$$u_{(34)}(\phi_1) > u_3(\phi_2) + u_4(\phi_2) \quad \textit{and}$$

$$\begin{cases} U_{(12)}(\phi_1) > U_1(\widehat{\phi}) + U_2(\widehat{\phi}) & \textit{if } u_{(34)}(\widehat{\phi}) > u_3(\widetilde{\phi}) + u_4(\widetilde{\phi}) \\ U_{(12)}(\phi_1) > U_1(\widetilde{\phi}) + U_2(\widetilde{\phi}) & \textit{if } u_{(34)}(\widehat{\phi}) < u_3(\widetilde{\phi}) + u_4(\widetilde{\phi}) \end{cases}$$

with $\widehat{\phi} = \arg\max_{\phi_3,\phi_4} \Pi(\phi)$ *and* $\widetilde{\phi} = \arg\max_{\phi_5,\ldots,\phi_8} \Pi(\phi)$.

2. $(\tau_N^*; \tau_S^*) = \{(12); 3, 4\}$ *with* $\phi_2^* = \{[(12), 3], [(12), 4]\}$ *iif*

$$u_{(34)}(\phi_1) < u_3(\phi_2) + u_4(\phi_2) \quad \textit{and}$$

$$\begin{cases} U_{(12)}(\phi_2) > U_1(\widehat{\phi}) + U_2(\widehat{\phi}) & \textit{if } u_{(34)}(\widehat{\phi}) > u_3(\widetilde{\phi}) + u_4(\widetilde{\phi}) \\ U_{(12)}(\phi_2) > U_1(\widetilde{\phi}) + U_2(\widetilde{\phi}) & \textit{if } u_{(34)}(\widehat{\phi}) < u_3(\widetilde{\phi}) + u_4(\widetilde{\phi}) \end{cases}$$

3. $(\tau_N^*; \tau_S^*) = \{1, 2; (34)\}$ *with* $\widehat{\phi}^* = \arg\max_{\phi_3,\phi_4} \Pi(\phi)$ *iif*

$$u_{(34)}(\widehat{\phi}) > u_3(\widetilde{\phi}) + u_4(\widetilde{\phi}) \textit{ and}$$

$$\begin{cases} U_{(12)}(\phi_1) < U_1(\widehat{\phi}) + U_2(\widehat{\phi}) & \textit{if } u_{(34)}(\phi_1) > u_3(\phi_2) + u_4(\phi_2) \\ U_{(12)}(\phi_2) < U_1(\widehat{\phi}) + U_2(\widehat{\phi}) & \textit{if } u_{(34)}(\phi_1) < u_3(\phi_2) + u_4(\phi_2) \end{cases}$$

4. $(\tau_N^*; \tau_S^*) = \{1, 2; 3, 4\}$ *with* $\widetilde{\phi}^* = \arg\max_{\phi_5,\ldots,\phi_8} \Pi(\phi)$ *iif*

$$u_{(34)}(\widehat{\phi}) < u_3(\widetilde{\phi}) + u_4(\widetilde{\phi}) \textit{ and}$$

$$\begin{cases} U_{(12)}(\phi_1) < U_1(\widetilde{\phi}) + U_2(\widetilde{\phi}) & \textit{if } u_{(34)}(\phi_1) > u_3(\phi_2) + u_4(\phi_2) \\ U_{(12)}(\phi_2) < U_1(\widetilde{\phi}) + U_2(\widetilde{\phi}) & \textit{if } u_{(34)}(\phi_1) < u_3(\phi_2) + u_4(\phi_2). \end{cases}$$

https://doi.org/10.1017/S1355770X12000411 Published online by Cambridge University Press

The uniqueness of the SPE of the game Γ comes directly from the uniqueness of the SPE of the Rubinstein bargaining model. Since the equilibrium transfers are unique, it implies that both the NC structure and the agreement structure are also unique (taking our assumptions into account).

If, for example, $u_{(34)}(\phi_1) > u_3(\phi_2) + u_4(\phi_2)$, $u_{(34)}(\widehat{\phi}) > u_3(\widetilde{\phi}) + u_4(\widetilde{\phi})$ and $U_{(12)}(\phi_1) > U_1(\widehat{\phi}) + U_2(\widehat{\phi})$, the equilibrium is the one described in Proposition (1.1). In this case, both northern and southern countries will form a NC, and the outcome of the negotiations will be a global agreement (i.e., $\phi_1^* = \{[(12), (34)]\}$). The abatement in this case is $q^*_{(34)}(\phi_1) = \arg\max_x [V_{(12)}(x) + v_{(34)}(x)]$, which is the abatement that a central planner would decide. This global agreement corresponds to the grand coalition using the standard definition of coalition used in the IEA literature.

The resulting payoffs are given by the worth for each player of the abatement, plus or minus the transfer from the North to the South:

$$U_{(12)}(\phi_1) = V_{(12)}(q^*_{(34)}(\phi_1)) - T^*_{(12),(34)}(\phi_1)$$
$$u_{(34)}(\phi_1) = v_{(34)}(q^*_{(34)}(\phi_1)) + T^*_{(12),(34)}(\phi_1)$$

with

$$T^*_{(12),(34)}(\phi_1) = \varphi[V_{(12)}(q^*_{(34)}(\phi_1)) - V_{(12)}(q^d_{(34)}(\phi_1))] + (1-\varphi)[v_{(34)}(q^d_{(34)}(\phi_1)) - v_{(34)}(q^*_{(34)}(\phi_1))].$$

However, if $u_{(34)}(\widehat{\phi}) < u_3(\widetilde{\phi}) + u_4(\widetilde{\phi})$, $U_{(12)}(\phi_1) < U_1(\widetilde{\phi}) + U_2(\widetilde{\phi})$ and $u_{(34)}(\phi_1) > u_3(\phi_2) + u_4(\phi_2)$, the equilibrium will be the one described in Proposition (1.4), and the agreement structure will be the one with highest aggregated payoff out of $\phi_5, \ldots, \phi_8$. If this is, for example $\phi_5 = \{[1,3],[1,4],2\}$, the equilibrium abatement $q_3^*(\phi_5)$ and $q_4^*(\phi_5)$ will be $q_j^*(\phi_5) = \arg\max_x [V_1(x, q_k^*(\phi_5)) + v_k(x, q_k^*(\phi_5))]$ with $j \neq k = 3, 4$ and the equilibrium payoffs will be:

$$U_1(\phi_5) = V_1(q_3^*(\phi_5), q_4^*(\phi_5)) - T^*_{1,3}(\phi_5) - T^*_{1,4}(\phi_5)$$
$$U_2(\phi_5) = V_2(q_3^*(\phi_5), q_4^*(\phi_5))$$
$$u_3(\phi_5) = v_3(q_3^*(\phi_5), q_4^*(\phi_5)) + T^*_{1,3}(\phi_5)$$
$$u_4(\phi_5) = v_4(q_3^*(\phi_5), q_4^*(\phi_5)) + T^*_{1,4}(\phi_5)$$

with

$$T^*_{1,3}(\phi_5) = \varphi[V_1(q_3^*(\phi_5), q_4^*(\phi_5)) - V_1(q_3^d(\phi_5), q_4^*(\phi_5))] + (1-\varphi)[v_3(q_3^d(\phi_5), q_4^*(\phi_5)) - v_3(q_3^*(\phi_5), q_4^*(\phi_5))]$$
$$T^*_{1,4}(\phi_5) = \varphi[V_1(q_3^*(\phi_5) + q_4^*(\phi_5)) - V_1(q_3^*(\phi_5), q_4^d(\phi_5))] + (1-\varphi)[v_4(q_3^*(\phi_5), q_4^d(\phi_5)) - v_4(q_3^*(\phi_5), q_4^*(\phi_5))].$$

In this case, neither the North nor the South will be interested in forming any NC, and in equilibrium northern country 1 will sign one agreement

https://doi.org/10.1017/S1355770X12000411 Published online by Cambridge University Press

with southern country 3 and another agreement with southern country 4, $\phi_5 = \{[1,3],[1,4],2\}$. Using the standard concept of coalitions in the IEA literature, the coalitions will now be $\{1,3\}$ and $\{1,4\}$, showing that in our framework northern countries can take part in more than one 'coalition' (in the IEA literature countries can generally participate in only one coalition, whether there is only one possible coalition (Carraro and Siniscalco, 1993) or several coalitions (Carraro, 2005)).

4. Discussion

Let the northern players be the US (player 1) and the EU[16] (player 2) and the southern players be China (player 3) and the remaining developing countries forming the G77 (player 4). Are northern (southern) players going to act united or separately during the climate change negotiations? Before answering this question, let us see to what extent our assumptions are reasonable for describing these negotiations.

The main limitation is that we have assumed that only bilateral agreements are possible, meaning that the only way to reach a global agreement would be a deal between a northern NC and a southern NC. Fortunately, this is not as far from reality as it could appear at first glance since the North (US–EU) and the South (China–G77) would probably need to reach a common position before they negotiate between themselves if a final global agreement is to be reached. If the US proposes a different deal to that of the EU, then a final global agreement is unlikely. In other words, our bilateral agreement framework does not preclude a global agreement; it only imposes a precise structure for the negotiations in order to reach this agreement. This brings us to the second strong assumption of our framework, the sequential NC structure formation: first the North, then the South. As pointed out above, this assumption is not essential for our analysis but it is probably reasonable if the North acts as a Stackelberg leader, a plausible assumption for a negotiation in which the North 'grants' a transfer to the South in exchange for an abatement effort which ultimately benefits the South more than the North (as climate change is expected to harm southern countries more). Furthermore, the EU has approved its strategy until 2020 by committing to a 30 per cent reduction in its 1990 emissions if other industrialized countries join the effort, or 20 per cent if they do not.[17] Since this was done before negotiations for the post-Kyoto agreement really started, and the EU has good chances to meet its Kyoto targets, the commitment of the EU is *a priori* and credible (almost defining a Stackelberg strategy).

Then, are the countries going to act united or separately? First they will check whether the agreement they can expect by joining forces is better or worse than the agreement they can expect by acting separately. This is captured in the direct efficiency terms given by $\underline{SE}$, $\underline{NE}$, $\overline{SE}$ and $\overline{NE}$ and is

[16] The EU and G77 are already groups of countries, but we will treat them as individual players.

[17] Brussels European Council 8/9 March 2007.7224/1/07Rev1. Brussels, 2 May 2007.

https://doi.org/10.1017/S1355770X12000411 Published online by Cambridge University Press

hardly surprising. As conditions (19) and (26) show, the different efficiency terms that we have defined favor southern and northern players forming NCs and, as proposition (1) has shown, the only way to reach the grand coalition is to have a negotiation between a united North and a united South. Hence, the larger the efficiency terms, the larger the options to reach a grand coalition.

Let us now move to the bargaining power terms, which are probably the most interesting ones as they may offset the efficiency terms. Condition (19) tells us that if $\underline{NBP}$ is larger than $\underline{SBP}$, then China and the G77 will have an additional reason to act together, not related to the different agreements that would actually be signed while acting together or separately. Should we expect $\underline{NBP}$ to be larger than $\underline{SBP}$? Probably not. The major beneficiaries of GHG abatement efforts are the countries of the South since they are expected to be the most damaged by climate change (IPCC, 2007). Therefore, the G77 would benefit from a partial agreement involving only China and an agreement signed only by the G77 would probably benefit China (i.e., $\underline{SBP}$ will presumably be large). On the other hand, the EU or the US would benefit relatively little from a partial agreement involving only the other northern player. Furthermore, a partial agreement might even harm one of them in political terms, as the Bush administration in the US probably suffered when the EU and the remaining countries included in Annex I of the Protocol continued with the Kyoto Protocol (i.e., $\underline{NBP}$ is probably small).

And what about the EU and the US? Do they have an incentive based on bargaining power to unite in climate negotiations? Condition (26) gives us the terms which the EU and the US have to consider while deciding whether or not to form a NC. Assume, for the sake of the argument, that each northern NC signs only one agreement and that both discount factors tend to one (i.e., $|J_1| = |J_2| = |J_{(12)}| = 1$; $\delta_S = \delta_N \to 1$ and hence $D \to 1$). In this case, $\overline{NBP} \simeq \underline{NBP}$ and $\overline{SBP} \simeq \underline{SBP}$. Owing to the fact that these terms enter with the opposite sign in condition (19) in condition (26), the same arguments used to support the lack of incentives for the merging of China and the G77 can be used to show that the EU and the US do indeed have an incentive to create a northern NC for bargaining power reasons. Thus, in our framework bargaining power gains hinder the formation of the grand coalition while they encourage southern countries to negotiate separately and northern countries to act united (equilibrium number 2 in Proposition (1)). In real-life negotiations, having a united North and a separated South probably does not preclude the formation of the grand coalition, but it certainly renders a global agreement more difficult. In any case, the take-home message is that, as shown by Hampson and Hart (1995) in their political science based discussion, which NC are formed is key to predicting the outcome of the negotiations.

Have we seen this behavior in past climate change negotiations? On the road to Kyoto in 1997 and Marrakech in 2001, not really. G77 and China negotiated together for most of the negotiation process whereas the EU and the US held opposite positions on most issues. However, the Kyoto Protocol and the Marrakech Accords include no obligations for the southern

countries, meaning that we can hardly talk of a North–South agreement, in our sense. Nevertheless, current climate change negotiations are more in line with the predictions of the bargaining power part of our argument. Before the failure in Copenhagen in 2009, the G8 members agreed in July 2009 upon a common abatement target for 2050 (80 per cent reduction of GHG emissions compared to 1990 levels and a 2°C increase in temperature target) and proposed a 50 per cent global abatement target that would imply abatement efforts for southern countries[18] (G8 countries also accepted that they would need to contribute financially). The opposition of China and India prevented an agreement within the G17 on this proposal and in Copenhagen no agreement was reached either (i.e., the North held a common position and offered the South a deal, which was rejected). In Cancun a modest agreement was achieved in 2010 but it is more a deal to save the negotiation process than a deal to tackle climate change once and for all. Northern countries this time had a more or less common position (with the relative exception of Japan), while southern countries were more divided, with China bargaining hard before accepting the deal and with Bolivia refusing the deal altogether until the last minute.

5. Conclusion

This paper has studied the reasons why northern and southern countries decide to form NCs when negotiating over climate change mitigation. The main question that we have analyzed is whether or not these negotiations have a natural tendency to be carried out between NCs and whether the reason for this possible tendency is based on efficiency gains or bargaining power gains (or both). Whether or not bargaining power gains induce the formation of NCs depends on the gains obtained by those players not involved in the eventual partial agreements (which, in our framework, never take place). Furthermore, if bargaining power encourages the formation of a NC on one side of the table, on the other side the tendency will go in the opposite direction. We have also shown that, for the climate change negotiations, we should expect bargaining power terms to favor a united North and a separated South. In our framework this precludes any possibility of reaching the grand coalition. Although this result can probably not be transposed directly to real-life negotiations, what holds is that having the South acting separately will always render a global agreement more difficult. In other words, as in Hampson and Hart (1995), we have shown that the stages where the NCs are formed have a great impact on the final outcome that one should expect.

[18] 'There has been an important convergence in G8 positions, which has provided a strong impulse to the upcoming negotiations to reach an ambitious and effective global agreement in Copenhagen. [...] the G8 countries have committed to reduce their greenhouse gas emissions by 80 per cent or more by 2050 with reference to 1990 or more recent years. [...] The G8 countries confirmed their willingness to contribute their fair share of the financing [...] to ensure the necessary actions to combat climate change also in developing countries.' G8 Fact Sheet – Climate Change ⟨www.g8italia2009.it⟩.

We have left possible extensions of this paper for future research. One alternative would be to assume sequential bargaining of the different agreements instead of simultaneous bargaining. This would allow, for example, analysis of the type of bargaining protocol preferred by the North or the South, respectively. As already pointed out, the NC formation protocol which we have assumed in stages one and two could be replaced by a more complex procedure. This would probably not change the results obtained but it would allow us to study the incentives to deviate within the NCs which we have not analyzed. Finally, the model presented here could be extended to a context with asymmetric information, which would advance the model nearer to reality and would avoid one of the shortcomings of the analysis presented here, namely that agreements are reached without delay.

References

Aghion, P., P. Antràs, and E. Helpman (2007), 'Negotiating free trade', *Journal of International Economics* **73**(1): 1–30.

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

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

Breton, M., K. Fredj, and G. Zaccour (2006), 'International cooperation, coalitions stability and free riding in a game of pollution control', *Manchester School* **74**(1): 103–122.

Caparrós, A. and J.-C. Péreau (2010), 'Coalition formation and bargaining power: theory and application to international negotiations on public goods', Instituto de Políticas y Bienes Públicos (IPP) Working Paper No. 17, Madrid.

Caparrós, A., J.-C. Péreau, and T. Tazdaït (2004), 'North–south climate change negotiations: a sequential game with asymmetric information', *Public Choice* **121**(3–4): 455–480.

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, UK: Cambridge University Press, pp. 354–380.

Carraro, C. and D. Siniscalco (1993), 'Strategies for the international protection of the environment', *Journal of Public Economics* **2**: 309–328.

Carraro, C., C. Marchiori, and A. Sgobbia (2005), 'Advances in negotiation theory: bargaining, coalitions and fairness', FEEM Working Paper No. 66.05, Milan.

Chander, P. and H. Tulkens (1997), 'The core of an economy with multilateral environmental externalities', *International Journal of Game Theory* **26**: 379–401.

Chipty, T. and C.M. Snyder (1999), 'The role of firm size in bilateral bargaining: a study of the cable television industry', *Review of Economics and Statistics* **81**(2): 326–340.

de Clippel, G. and R. Serrano (2008), 'Bargaining, coalitions and externalities: a comment on Maskin', Working Paper No. 2008-16, Department of Economics, Brown University, Providence, RI.

Finus, M. (2001), *Game Theory and International Environmental Cooperation*, Cheltenham: Edward Elgar.

Hampson, O. and M. Hart (1995), *Multilateral Negotiations: Lessons from Arms Control, Trade, and the Environment*, Baltimore and London: Johns Hopkins University Press.

Harstad, B. (2010), 'How to negotiate and update climate agreements', in J.E. Aldy and R.N. Stavins (eds), *Post-Kyoto International Climate Policy: Implementing Architectures for Agreement*, Cambridge, UK: Cambridge University Press, pp. 273–299.

IPCC (2007), *Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, M.L. Parry, O.F. Canziani, J.P. Palutikof, P.J. van der Linden and C.E. Hanson (eds), Cambridge, UK: Cambridge University Press.

Maskin, E. (2003), 'Bargaining, coalitions and externalities', Presidential Address to the Econometric Society, Institute for Advanced Study, Princeton.

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

Muthoo, A. (1999), *Bargaining Theory with Applications*, Cambridge, UK: Cambridge University Press.

Ray, D. and R. Vohra (1999), 'A theory of endogenous coalition structures', *Games and Economic Behavior* **26**: 286–336.

Ray, D. and R. Vohra (2001), 'Coalitional power and public goods', *Journal of Political Economy* **109**(6): 1355–1384.

Rotillon, G., T. Tazdaït, and S. Zeghni (1996), 'Bilateral or multilateral bargaining in the face of global environmental change?', *Ecological Economics* **18**: 177–187.

Rubinstein, A. (1982), 'Perfect equilibrium in a bargaining model', *Econometrica* **50**: 97–109.