



HAL
open science

Inefficient coasean negotiations over emissions and transfers

B.A. Caparros, Jean-Christophe Pereau

► **To cite this version:**

B.A. Caparros, Jean-Christophe Pereau. Inefficient coasean negotiations over emissions and transfers. Journal of Economic Behavior and Organization, 2021, 189, pp.359-378. <10.1016/j.jebo.2021.06.044>. <hal-03432922>

HAL Id: hal-03432922

<https://hal.science/hal-03432922v1>

Submitted on 2 Aug 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC 4.0 - Attribution - Non-commercial use - International License

Inefficient Coasean Negotiations over Emissions and Transfers

By Alejandro Caparrós^{a,b} and Jean-Christophe Perea^c

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

^bDurham University Business School, Durham University, Durham DH1 3LB, UK

^cUniversity of Bordeaux, GREThA (UMR CNRS 5113), Avenue Léon Duguit, 33608 Pessac cedex, France; e-mail: jean-christophe.pereau@u-bordeaux.fr

Abstract: This article analyses the outcomes of multilateral and sequential negotiation procedures in a Rubinstein alternating-offers model where two polluters and a victim bargain over both, transfers and pollution levels. We show that the Coase Theorem does not hold in a multilateral framework if sequential negotiations are possible (not imposed), although there are no frictions and no delays between stages. Sequential negotiations lead to emission levels which are not socially optimal, but players involved in the first agreement in the sequential path may prefer this path and hence launch it. We also show that when negotiations focus only on transfers, as commonly assumed, the inefficiency vanishes. Finally, we show that the inefficiency can be explained by the player's inside options, which are given by their potential temporary disagreement payoffs, despite the fact that agreements are reached immediately in equilibrium. Results are generalized to a large number of polluters.

JEL Classification: C78

Keywords: polluter, victim, alternating-offers model, sequential bargaining, inside option, impasse point, Coase theorem.

Acknowledgments: We acknowledge financial support from the CNRS-CSIC International Research Program 'ALLIES' 2020-2023 GREThA-IPP. Alejandro Caparrós acknowledges financial support from the Spanish Ministry of Science and Innovation (project Tren-Good, ECO2017-84461-R). We thank Noemi Navarro, the three referees and the editor for their stimulating comments and suggestions.

1 Introduction

In one of the most influential articles in economics, Coase (1960) argued that bargaining with side payments can solve pollution problems without the need of government intervention, and that only transaction (or bargaining) costs can prevent voluntary bargaining from attaining Pareto-efficient outcomes. This basic idea is supported, in a more formal setting, by the well known result that, under complete information, Rubinstein's (1982) alternating-offer bargaining model yields an efficient outcome (Farrell, 1987; Kenan and Wilson, 1993; Muthoo, 1999; Lee and Sabourian, 2007). The negotiation ends immediately on an agreement located on the Pareto frontier which is also the social optimum. Although Coase claimed that his result holds generally, and not only for two players, most of the analyses based on Rubinstein's bargaining protocol have focused on the two players case. When moving to a multilateral framework, several questions arise. Are the outcomes in a multilateral framework always socially efficient? If not, when is the outcome efficient and what explains inefficient outcomes? Or, alternatively, how sensitive is the outcome to minor variations in the structure of the game?

Aivazian and Callen (1981) and Aivazian et al. (1987) addressed the problems that arise when the Coase theorem is generalized to more than two players, in particular two pollutants and a victim. They use numerical examples and cooperative game theory, focusing either on the core (Aivazian and Callen, 1981) or on the bargaining set (Aivazian et al. 1987). Their basic result is that if agreements can be renegotiated at any time, the three players involved in the game will be involved in "endless recontracting". As Coase points out in his comment to the first paper just mentioned, if the agreements considered by the authors are seen as contracts they are "peculiar in that their terms can be broken at will" (Coase, 1981). The goal of our paper is, therefore, to analyze this issue in a fully non-cooperative setting, where players can sign binding contracts.¹

Okada (1996) shows in a coalition formation framework that if intermediate coalitions are possible and players are sufficiently patient, the grand coalition is formed if and only if it has the largest value per capita among all coalitions (other authors have found similar results, see the survey in Ray and Vohra 2015). Basically, intermediate coalitions are found in equilibrium if at least one player has no interest in moving from the intermediate coalition to the grand coalition, as her payoff in the intermediate coalition structure is higher. As the grand coalition is not reached, not all possible contracts are signed and implemented. The inefficiency highlighted in the model outline below has a different source, and it holds

¹Hence, we essentially rule out endless recontracting by assumption. Although, this can be seen as a strong assumption, it provides a framework that complements the analysis in Aivazian and Callen (1981) and Aivazian et al. (1987).

despite the fact that all players are interested in reaching the grand coalition, where all possible contracts are in place. The inefficiency arises because in our model players bargain over emission and transfers and, depending on the path followed to reach the final situation where all contracts are signed, emissions, and hence payoffs, may differ from the socially optimal level (if players bargain only over transfers, optimality is restored).

Cai (2000) also finds an inefficient result in multilateral bargaining under perfect information, in his case due to the delay involved in reaching an agreement. Applying Cai's model to Aivazian and Callen's two-polluters-one-victim example, he imposes that the victim negotiates binding agreements, necessarily sequentially with the two polluters, one at a time and in a pre-established order. Although we share the assumption that the victim cannot go back and renegotiate an agreement, as we also consider binding contracts, we do not impose a sequential negotiation and our victim is free to choose a multilateral protocol. Furthermore, even if the sequential negotiation is chosen, the victim can decide the order in which she approaches the polluters. This implies that the series of interdependent hold-up problems that explain Cai's result do not appear in our model.

In particular, we compare the outcome of multilateral and sequential bilateral negotiation protocols with respect to the social optimum in a setting where two agents are polluting and one victim is suffering the pollution. As Aivazian and Callen (1981) we use a model with two polluters and one victim for definitiveness, although we do not rely on numerical examples and we subsequently generalize our key results to any number of polluters. In any case, to fix ideas, consider the two factories in Aivazian and Callen's example, labeled 1 and 2, that pollute with their emissions a neighborhood laundry, labeled firm 0, whose operating costs are raised due to the smoke. As already noted, in terms of institutional setting for the bargaining protocols, we do neither impose sequential bargaining nor multilateral bargaining, allowing the players to choose their preferred option.

The basic tenet from the Coase Theorem² is that, as players will agree on an optimal pollution level if left alone, the government has only to enforce property rights and private agreements in order to reach the social optimum. Nevertheless, we show that as soon as there are two polluters and one victim, the outcome may very well be inefficient from the point of view of society (although property rights are well defined in our model). We show that the agreed level of pollution is efficient if all three players bargain multilaterally (simultaneously), but that if players bargain sequentially the pollution level is inefficient, even if the time period between stages in the sequential procedure becomes infinitely small. This holds despite the fact that agreements are reached immediately in equilibrium (as our model assumes perfect information) and all players sign and implement an agreement in the final situation (unlike

²Coase never called his result a Theorem (Coase, 1960).

in Okada 1996). In a Coasean world one can argue that, knowing this result, the players involved in the game would never choose a sequential negotiation path. Nevertheless, our results show that the players involved in the first agreement in a sequential path may in fact prefer³ the path initiated with this partial agreement, under certain conditions. If there is no government intervention, nothing would prevent them from negotiating and implementing this partial agreement, and hence produce a socially inefficient outcome.

In modeling terms, the contribution of this paper is to emphasize the role of inside options in a Rubinstein framework with potential sequential bargaining procedures (inside options refer to the payoffs that players obtain when they temporarily disagree). By assuming in each case that delay between negotiation rounds vanishes, the goal is to explain the source of possible inefficiencies. As already mentioned, the intuition for the inefficiency that we observe in the sequential case comes from the fact that in our game agents bargain over transfers *and* emission levels. Thus, not only the distribution of the net surplus is different, the pollution level and, hence, the net surplus itself, is also different. We also show that the inefficiency in the sequential procedure comes from the inside options, despite the fact that in equilibrium the temporary disagreement payoffs are never implemented as players agree immediately. Finally, we show that one can remove the inefficiency by not implementing immediately partial agreements (note, however, that this would also imply some form of intervention, as parties are otherwise free to implement their partial agreement).

Previous polluter-victim literature on the role of bargaining protocols on the Coase Theorem in a multilateral setting is rather scarce. However, inside options have been analyzed in labour economics. Inside options have been studied by Haller and Holden (1990), Fernandez and Glazer (1991) and Muthoo (1999) in simultaneous two-player procedures, focusing on negotiations between a union and a firm. In a three-player bargaining framework, between one union and two competing firms on the product market, Banerji (2002) analyses under which conditions the union prefers to bargain in sequence with the firms. Our framework shares some features with that model, but adds how the choice of partial agreements can be related to the achievement of the social optimum⁴. In terms of results, employment levels are first best with both protocols in Banerji (2002), but the author only shows that sequential

³This amounts to assuming a pre-negotiation phase where one of the players chooses the bargaining protocol. In a recent paper, Hu and Rocheteau (2020) propose a similar pre-bargaining stage, showing in their case that the buyer would choose one protocol and the seller another one. In our case, under certain conditions discussed below, *both* the polluter and the victim involved in the first negotiation in the sequential path would choose the sequential path. Hence, in the absence of intervention they will launch it, irrespectively of who has the option to choose the path in the pre-negotiation phase.

⁴In Banerji (2002) the assumption is that the firm stops producing while negotiating, while in our case emissions continue. Another difference is that our polluters are not affected by the emissions of the other polluter.

bargaining is preferred by one side (the trade union). On the contrary, we will show that sequential bargaining can be inefficient in our model and preferred by both sides involved in the first negotiation.

We focus on the case where agreements are implemented as they are reached (sequential implementation), although we also consider the impact of waiting until all agreements are struck before implementing them (simultaneous implementation). Both assumptions have been analyzed in the literature on multi-issues bargaining, where typically two players bargaining over two pies (Flamini, 2007). In this context, players are either allowed to eat their piece of the first pie immediately or only when an agreement over both pies has been reached (e.g. Fershtman 1990, Busch and Horstmann 1997). As in these papers, we also find that alternative assumptions on the implementation have a clear impact on the results obtained. However, note that in our case the two issues over which bargaining takes place are negotiated simultaneously (i.e. emissions and transfers), and that what we are considering is whether or not an agreement reached with a subset of the players involved in the game can be implemented immediately, or not.

There is a very large body of literature, both theoretical and empirical, that finds that the Coase Theorem fails if there are frictions. As pointed out by Dixit and Olson (2000), the literature on the Coase Theorem is too vast to attempt to survey it. Examples of this literature can be found in Lee and Sabourian (2007) who show, introducing complexity considerations, that the Theorem fails if negotiations are costly (this holds even if negotiation costs are infinitely small, as long as players have a marginal preference for simple strategies). Dixit and Olson (2000), focusing on free-riding incentives, also find that the efficient equilibrium is not robust to even very small transaction costs. However, our results show that the Theorem may fail in a situation without transaction costs or negotiation costs (as the discount rate tends to zero, negotiations are not costly, agreements are binding and there is perfect information).

The outline of the paper is as follows. The model is presented in Section 2. Multilateral and sequential bargaining are analyzed in Section 3. Section 4 compares the results obtained under the two bargaining procedures considered, and presents several examples to illustrate the main results of the paper. The last section concludes.

2 The model

Consider a negotiation game where a victim (the neighborhood laundry firm) suffers from the emissions of two polluters (the factories) and bargains over emission levels and transfers with both polluters. Denote the victim by 0, the two polluters by $i = 1, 2$ and the potential

transfers by τ_i , $i = 1, 2$. We assume in the remainder of the paper that the initial situation refers to a world with pollution, meaning that the polluters have the right to pollute. This implies that the victim has to compensate the polluters for their emission abatements with respect to the initial situation. A generic negotiation between the victim and a polluter $i = 1, 2$ will consist in a pair of variables, the level of emissions and the transfer, denoted by $o_i = (e_i, \tau_i)$. The negotiation game is analyzed under perfect information according to the alternating-offers procedure (Rubinstein, 1982; Osborne and Rubinstein, 1994). For simplicity, we add the assumption that the victim makes the first offer and chooses with whom she wants to bargain first (in the sequential case). The results described below would not change under the assumption that the polluter makes the first offer, although the distributions of welfare would vary. We also assume that the game ends when a multilateral agreement has been reached, or when both sequential negotiations have concluded with an agreement. The payoff function of the victim is defined as the sum of the flow of damage suffered and the transfer paid to the polluters, while the payoffs for both polluters are given by the flow of revenue obtained through their polluting activities plus the transfer received. Although time is continuous in this model, offers and counteroffers are made at periods of time separated by Δ units of time, as detailed below, and potential transfers from the victim to a polluter are paid only once (see Muthoo 1999 for a similar framework in a bilateral setting, although focused only on shares of a fixed pie). The flow of transfer paid to polluter i is denoted by τ_i , with $\tau_i \geq 0$, while $B_i(e_i) \geq 0$, for $i = 1, 2$, is the benefit function for the polluters and $D(e_1, e_2) \geq 0$ the damage function for the victim⁵. These functions have the following properties: $\frac{\partial B(e_i)}{\partial e_i} > 0$ for $e_i < \bar{e}_i$, $\frac{\partial B(e_i)}{\partial e_i} = 0$ for $e_i = \bar{e}_i$; $\frac{\partial B(e_i)}{\partial e_i} < 0$ for $e_i > \bar{e}_i$; where \bar{e}_i stands for the level which maximizes the individual benefit function of the polluters (It also corresponds to the emission level in case of complete disagreement and is called the "threat" in the sequel of the paper.); $\frac{\partial^2 B(e_i)}{\partial e_i^2} < 0$; $\frac{\partial D(e_1, e_2)}{\partial e_i} > 0$ and $\frac{\partial^2 D(e_1, e_2)}{\partial e_i^2} > 0$, with respect to e_i , for $i = 1, 2$. Hence, the strategy space for polluters is given by $e_i \in [0, \bar{e}_i]$. To ensure an interior solution, we also assume that $\frac{\partial B_i(0)}{\partial e_i} - \frac{\partial D(0, e_j)}{\partial e_i} > 0$ for $e_i \leq \bar{e}_i$, $i = 1, 2$ and $i \neq j$ (i.e., for any value of e_j the first unit of pollution of player i provides her more benefits than damage to player 0).

Taking this into account, the instantaneous payoff functions can be written as follows,

⁵As these values only vary after an agreement is reached, we do not use any time index to simplify the notation.

using U for the victim and Π for the polluters:

$$U_0(o_1, o_2) = -D(e_1, e_2) - \tau_1 - \tau_2 \quad (1)$$

$$\Pi_1(o_1) = B_1(e_1) + \tau_1 \quad (2)$$

$$\Pi_2(o_2) = B_2(e_2) + \tau_2 \quad (3)$$

We denote discount rates by r_i and denote the discounted payoff functions by V_i , for $i = 0, 1, 2$. Finally, to allow for a more compact presentation of the results, we define the global net surplus created when an agreement implies that players 1 and 2 move from one generic situation, denoted by a and consisting in a pair of efforts (e_1^a, e_2^a) , to another situation denoted by b with a new pair of efforts (e_1^b, e_2^b) , as:

$$\Psi(e_1^b, e_2^b; e_1^a, e_2^a) = B_1(e_1^b) + B_2(e_2^b) - D(e_1^b, e_2^b) - [B_1(e_1^a) + B_2(e_2^a) - D(e_1^a, e_2^a)]. \quad (4)$$

In order to have a meaningful bargaining problem, we assume that $\Psi(e_1^b, e_2^b; e_1^a, e_2^a) > 0$ when $e_i^b \leq e_i^a$ for $i = 1, 2$, with at least one strict inequality. Hence if at least one polluter reduces its emission in situation b compared to situation a , the created surplus is positive.

For future reference, we denote the emission levels at the social optimum as (e_1^o, e_2^o) , noting that they imply⁶

$$\begin{aligned} \frac{\partial B_1(e_1^o)}{\partial e_1} &= \frac{\partial D(e_1^o, e_2^o)}{\partial e_1} \\ \frac{\partial B_2(e_2^o)}{\partial e_2} &= \frac{\partial D(e_1^o, e_2^o)}{\partial e_2} \end{aligned} \quad (5)$$

The social optimum states that emissions are such that marginal benefits and damages are equated.

3 Bargaining procedures

We consider and compare the outcome of two bargaining procedures, a multilateral and a sequential negotiation.

3.1 Multilateral procedure

As a benchmark, we consider a direct extension of Rubinstein's procedure to a multilateral setting, which we call game \mathcal{G}^m . For notational convenience, we define offers made by player

⁶To simplify the notation of the derivatives we include, where appropriate, in the parenthesis the point where the derivative is evaluated, i.e. $\frac{\partial B_1(e_1^o)}{\partial e_1} = \frac{\partial B_1(e_1)}{\partial e_1} \Big|_{e_1^o}$.

i ($i = 0, 1, 2$) by $o_{12}^{(i)} = (o_1^{(i)}; o_2^{(i)}) = (e_1^{(i)}, \tau_1^{(i)}; e_2^{(i)}, \tau_2^{(i)})$, showing in the superscript the player making the offer. Player 0 makes the first offer $o_{12}^{(0)} = (e_1^{(0)}, \tau_1^{(0)}; e_2^{(0)}, \tau_2^{(0)})$, at time 0. If the offer is accepted by players 1 and 2, the agreement is struck and immediately implemented. If players 1 or 2 reject the initial offer, then Δ time units later, at time 1Δ , player 1 makes the next offer $o_{12}^{(1)} = (e_1^{(1)}, \tau_1^{(1)}; e_2^{(1)}, \tau_2^{(1)})$ to players 0 and 2, who can accept it or reject it. If player 0 or player 2 reject the new offer, then Δ time units later, at time 2Δ , player 2 makes a new offer. If accepted, the agreement is struck and implemented. Otherwise, player 0 talks again and the process is repeated indefinitely, until a unanimous agreement is reached. If they reach an agreement at time $t\Delta$ ($t = 0, 1, \dots$ and $\Delta > 0$) on a pair $o_{12}^* = (e_1^*, \tau_1^*; e_2^*, \tau_2^*)$, the negotiations conclude and from that moment onwards the discounted instantaneous payoffs are $D(o_{12}^*)e^{-r_0 t\Delta}$ for player 0 and $B_i(o_i^*)e^{-r_i t\Delta}$ for player $i = 1, 2$. That is, if an agreement is reached at time $t\Delta$ the discounted payoffs (denoted V_i for $i = 0, 1, 2$) for the players are:

$$V_0(t, \Delta, o_{12}) = \int_0^{t\Delta} U(\bar{o}_{12})e^{-r_0 x} dx + \int_{t\Delta}^{\infty} U(o_{12}^*)e^{-r_0 x} dx \quad (6)$$

$$V_1(t, \Delta, o_1) = \int_0^{t\Delta} \Pi_1(\bar{o}_1) e^{-r_1 x} dx + \int_{t\Delta}^{\infty} \Pi_1(o_1^*) e^{-r_1 x} dx \quad (7)$$

$$V_2(t, \Delta, o_2) = \int_0^{t\Delta} \Pi_2(\bar{o}_2) e^{-r_2 x} dx + \int_{t\Delta}^{\infty} \Pi_2(o_2^*) e^{-r_2 x} dx \quad (8)$$

Note that using the notation introduced in equations (1)-(3) we have: $U(\bar{o}_{12}) = -D(\bar{e}_1, \bar{e}_2)$; $U(o_{12}^*) = -D(e_{12}^*) - (\tau_1^* + \tau_2^*)$; $\Pi_i(\bar{o}_i) = B_i(\bar{e}_i)$; $\Pi_i(o_i^*) = B_i(e_i^*) + \tau_i^*$.

The first terms on the RHS of the payoffs (6)-(8) show the inside options for the three players (Haller and Holden, 1990; Muthoo, 1999). The initial (status-quo) pollution level corresponds to a Nash equilibrium and continues as long as no agreement is reached. It gives the payoff that the agents obtain when they temporarily disagree. In our model, as we assume that the polluters have the right to pollute, the inside option is given by the threat emission levels. The impasse point is given by an infinite stream of pollution at this level, i.e. for player 0 the impasse point yields $(1/r_0)U_0(\bar{o}_{12})$ and for players $i = 1, 2$, $(1/r_i)\Pi_i(\bar{o}_i)$.

As we are interested in negotiations without frictions, where the Coase theorem is less likely to fail, we focus on the case where $\Delta \rightarrow 0$. To simplify⁷, we assume stationary strategies

⁷As shown by Shaked and reported by Sutton (1986), the generalization of the Rubinstein bilateral bargaining process to a multilateral case is problematic since it may yield multiple equilibria when unanimity is required and strategies are not stationary. There are different alternatives to restore uniqueness even with non stationary strategies, such as the possibility of making conditional and unconditional offers financed by outside money (Huang, 2002), which has been extended to games with externalities in Caparrós and Péraireu (2017). For our analysis what is relevant is that Huang (2002) shows that, as long as all players are allowed to talk once following a cyclical protocol, the solution of the game allowing for non-stationary strategies converges to the SSPE of the unanimity game described above when $\Delta \rightarrow 0$ (the same result is obtained with the 'exit-game' proposed by Krishna and Serrano (1996)).

and restrict our analysis to Stationary Subgame Perfect Equilibria (SSPE) (see Osborne and Rubinstein, 1994, p131). Taking all this into account, we obtain the following Proposition:

Proposition 1 *Assume the multilateral bargaining procedure (game \mathcal{G}^m) and $\Delta \rightarrow 0$. The SSPE strategies are: (i) for player 0 to offer, at period $t = 0$, $o_1^{(0)} = (e_1^*, \tau_1^*)$ to player 1 and $o_2^{(0)} = (e_2^*, \tau_2^*)$ to player 2 and (ii) for players 1 and 2 to accept the offers immediately. The emission levels at the equilibrium are equal to the socially optimal emission levels, i.e. $(e_1^*, e_2^*) = (e_1^o, e_2^o)$, as they are given by (5). After implementing the transfers (τ_1^*, τ_2^*) , shown in equations (52) and (53) in Appendix A.1, the net equilibrium instantaneous payoffs are:*

$$U_0^*(o_1, o_2) - U_0(\bar{o}_1, \bar{o}_2) = \frac{r_1 r_2}{r_0 r_1 + r_0 r_2 + r_1 r_2} \Psi(e_1^o, e_2^o; \bar{e}_1, \bar{e}_2) \quad (9)$$

$$\Pi_1^*(o_1) - \Pi_1(\bar{o}_1) = \frac{r_0 r_2}{r_0 r_1 + r_0 r_2 + r_1 r_2} \Psi(e_1^o, e_2^o; \bar{e}_1, \bar{e}_2) \quad (10)$$

$$\Pi_2^*(o_2) - \Pi_2(\bar{o}_2) = \frac{r_0 r_1}{r_0 r_1 + r_0 r_2 + r_1 r_2} \Psi(e_1^o, e_2^o; \bar{e}_1, \bar{e}_2) \quad (11)$$

Proof. See Appendix A.1. ■

The positive global net surplus created is $\Psi(e_1^o, e_2^o; \bar{e}_1, \bar{e}_2)$, using the notation introduced in (4), and the shares obtained by each player depend on the relative patience of that player compared to the remaining players. Finally, note that assuming identical values for the discount rates of all players yields an equal share of 1/3 for all players.

We also show the Nash bargaining problem that would yield the results shown in Proposition 1, omitting the proof as it is straightforward (we just note that we can write the objective function using instantaneous payoffs instead of discounted payoffs as all values are constant).

Corollary 1 *SSPE's emission levels, transfers and payoff pairs of \mathcal{G}^m with $\Delta \rightarrow 0$, shown in Proposition 1, converge to the asymmetric Nash Bargaining Solution (NBS) of the following problem:*

$$\max_{e_1, \tau_1, e_2, \tau_2} \Omega_{012} = (U(o_1, o_2) - U(\bar{o}_1, \bar{o}_2))^{\alpha_0} (\Pi_1(o_1) - \Pi_1(\bar{o}_1))^{\alpha_1} (\Pi_2(o_2) - \Pi_2(\bar{o}_2))^{\alpha_2}$$

with $\alpha_0 = \frac{r_1 r_2}{r_0 r_1 + r_0 r_2 + r_1 r_2}$, $\alpha_1 = \frac{r_0 r_2}{r_0 r_1 + r_0 r_2 + r_1 r_2}$ and $\alpha_2 = 1 - \alpha_0 - \alpha_1$.

Extending this corollary to N polluters is also straightforward, showing that our key results would hold for any number of polluters.

3.2 Sequential procedure

Consider now the sequential game, called \mathcal{G}^s , where the negotiation proceeds in two stages. In stage 1, players 0 and i ($i = 1, 2$) bargain over $o_i = (e_i, \tau_i)$ and in stage 2 players 0

and j ($j \neq i$) bargain over $o_j = (e_j, \tau_j)$ knowing that an agreement has been reached at stage 1. We assume that the victim bargains twice, makes the first offer in both stages and chooses the polluter with whom she wants to bargain in stage 1. The other polluter is assumed to wait until the next stage. In each stage, the negotiation takes place following the offer-counteroffer procedure. We assume, without loss of generality as we can rename players at will, that negotiations start with player 1. Thus, let us consider first a bilateral negotiation between players 0 and 1. As before, we define offers made by player i ($i = 0, 1$) by $o_1^{(i)} = (e_1^{(i)}, \tau_1^{(i)})$. Player 0 makes the first offer $o_1^{(0)} = (e_1^{(0)}, \tau_1^{(0)})$, at time 0, which can be accepted by player 1, or rejected. If player 1 rejects the initial offer, then $\Delta > 0$ time units later, at time 1Δ , player 1 makes a counteroffer $o_1^{(1)} = (e_1^{(1)}, \tau_1^{(1)})$ to player 0, who can accept it or reject it. If player 0 rejects the counter-offer, then Δ time units later, at time 2Δ , he makes a new offer, and so on. If they reach an agreement at time Δt_1 , ($t_1 = 0, 1, \dots$ and $\Delta > 0$) on a pair $o_1^s = (e_1^s, \tau_1^s)$, from that moment onwards the discounted flow of payoffs are $D(e_1^s, \bar{e}_2)e^{-r_0 t_1 \Delta}$ for player 0 and $B_1(e_1^s)e^{-r_1 t_1 \Delta}$ for player 1 (to avoid confusion, in this sub-section we denote equilibrium values by an s in superscript). As long as no agreement is reached, the negotiation at this stage continues (potentially for ever) and players obtain their per-period disagreement payoff. Once an agreement between players 0 and 1 is struck, stage 2 of the game starts immediately. For future reference, we call the (endogenous) time at which the second stage starts $\Gamma = t_1 \Delta$. Player 0 starts bargaining with player 2 over $o_2 = (e_2, \tau_2)$, again following the bilateral offer-counteroffer procedure described above, with player 0 opening the negotiation. As long as no agreement is reached at stage 2, all players receive their partial-agreement (first-stage agreement) per-period payoff, and the negotiation continues (potentially for ever). Once an agreement between players 0 and 2 is reached over $o_2^s = (e_2^s, \tau_2^s)$ at time⁸ $(t_1 + t_2) \Delta = \Gamma + \Delta t_2$ (with $t_2 = 0, 1, \dots$), the negotiations conclude and all player receive their full-agreement payoffs from that period onwards. That is, if an agreement with player 1 is reached at $t_1 \Delta$ and an agreement with player 2 at $(t_1 + t_2) \Delta$, the

⁸The number of offers and counter-offers needed before an agreement is reached in the second stage are denoted by t_2 (with $t_2 = 0, 1, \dots$). As each period between offers has still a duration of Δ , the time at which the second-stage agreement is reached is $\Gamma + \Delta t_2$, when counted from the beginning of the game.

payoffs for the players are:

$$V_0(t_1, t_2, \Delta, o_{12}) = \int_0^{t_1\Delta} U(\bar{o}_{12})e^{-r_0x}dx + \int_{t_1\Delta}^{(t_1+t_2)\Delta} U(o_1^s, \bar{o}_2)e^{-r_0x}dx + \int_{(t_1+t_2)\Delta}^{\infty} U(o_{12}^s)e^{-r_0x}dx \quad (12)$$

$$V_1(t_1, t_2, \Delta, o_1) = \int_0^{t_1\Delta} \Pi_1(\bar{o}_1)e^{-r_1x}dx + \int_{t_1\Delta}^{\infty} \Pi_1(o_1^s)e^{-r_1x}dx \quad (13)$$

$$V_2(t_1, t_2, \Delta, o_2) = \int_0^{(t_1+t_2)\Delta} \Pi_2(\bar{o}_2)e^{-r_2x}dx + \int_{(t_1+t_2)\Delta}^{\infty} \Pi_2(o_2^s)e^{-r_2x}dx \quad (14)$$

As noted after equations (6)-(8) these expressions (and those shown below), can be write as a function of D and B , and hence as a function of emissions and transfers.

Solving game \mathcal{G}^s by backward induction we have that, in stage two, players 0 and 2 bargain over $o_2 = (e_2, \tau_2)$, taking into account the fact that an agreement over $o_1 = (e_1, \tau_1)$ has been reached in stage one and that this agreement has already been implemented. This gives the following pair of equilibrium conditions for stage 2 (starting at Γ , recall our stationarity assumption):

$$\begin{cases} \int_{\Gamma}^{\infty} U(o_1, o_2^{(2)})e^{-r_0x}dx &= \int_{\Gamma}^{\Gamma+\Delta} U(o_1, \bar{o}_2)e^{-r_0x}dx + \int_{\Gamma+\Delta}^{\infty} (U o_1, o_2^{(0)})e^{-r_0x}dx \\ \int_{\Gamma}^{\infty} \Pi_2(o_2^{(0)})e^{-r_2x}dx &= \int_{\Gamma}^{\Gamma+\Delta} \Pi_2(\bar{o}_2)e^{-r_2x}dx + \int_{\Gamma+\Delta}^{\infty} \Pi_2(o_2^{(2)})e^{-r_2x}dx \end{cases} \quad (15)$$

The first equation of the system (15) is the indifference condition for player 0, implying that player 0 is indifferent between an agreement based on the offer of his opponent (LHS) and the sum of the payoff during the negotiation and what he expects to obtain in the future with a payoff based on his counter-offer (RHS) if he rejects the offer initially. More precisely, the LHS term is the payoff of player 0 from Γ to ∞ based on an agreement with the offer of 2, assuming that 0 has already reached an agreement with 1. Concerning the expression on the RHS, the first term, $\int_{\Gamma}^{\Gamma+\Delta} U(o_1, \bar{o}_2)e^{-r_0x}dx$ with $\bar{o}_2 = (\bar{e}_2, 0)$, refers to the payoff that player 0 obtains when he rejects the offer from 2 but benefits from the agreement with 1. This is his inside option (Haller and Holden, 1990; Muthoo, 1999; Banerji, 2002). The second term is the payoff of player 0 when the agreement is based on his counter-offer from $\Gamma + \Delta$ to ∞ . The second equation of the system (15) is the indifference condition for player 2, which follows the same logic. The inside option of player 2 is given by the term $\int_{\Gamma}^{\Gamma+\Delta} \Pi_2(\bar{o}_2)e^{-r_2x}dx$.

In stage one, players 0 and 1 bargain over $o_1 = (e_1, \tau_1)$. At this stage, the indifference

conditions are, for player 0:

$$\begin{aligned}
& \int_0^\Gamma U(o_1^{(1)}, \bar{o}_2) e^{-r_0 x} dx + \int_\Gamma^\infty U(o_1^{(1)}, o_2(o_1^{(1)})) e^{-r_0 x} dx \\
= & \int_0^\Delta U(\bar{o}_1, \bar{o}_2) e^{-r_0 x} dx + \int_\Delta^{\Gamma+\Delta} U(o_1^{(0)}, \bar{o}_2) e^{-r_0 x} dx + \int_{\Gamma+\Delta}^\infty U(o_1^{(0)}, o_2(o_1^{(0)})) e^{-r_0 x} dx
\end{aligned} \tag{16}$$

and for player 1:

$$\int_0^\infty \Pi_1(o_1^{(0)}) e^{-r_1 x} dx = \int_0^\Delta \Pi_1(\bar{o}_1) e^{-r_1 x} dx + \int_\Delta^\infty \Pi_1(o_1^{(1)}) e^{-r_1 x} dx \tag{17}$$

The indifference condition (16) for player 0 at stage 1 is more complex and deserves further attention. Concerning the LHS terms, both terms represent the total payoff of player 0 if she accepts the offer from 1, distinguishing two different periods of time. The first one is the payoff of the victim from 0 to Γ . During that period, no agreement exists with 2. The second term measures her payoff from time Γ onwards, when an agreement with 2 will be signed given the offer accepted in stage 1 (the term with $o_2(o_1^{(1)})$). Concerning the RHS of (16), it reflects the payoff obtained by player 0 if she rejects the initial offer and waits for a future deal based on her own offer. The first term is the payoff of player 0 when no agreement exists with any player, between time 0 and Δ . This is her inside option during the first-stage negotiation. The next terms on the RHS are her payoff when the agreement is signed based on her counter-offer over two sub-periods: (i) from Δ to $\Gamma + \Delta$ before an agreement is reached in round 2 with player 2, i.e. the term with \bar{o}_2 (this is his inside option during the second stage negotiation) and (ii) from $\Gamma + \Delta$ to ∞ , after the agreement with player 2 is reached given the offer accepted in stage 1 (the term with $o_2(o_1^{(0)})$). Equation (17) gives the indifference condition for 1 and follows a similar logic.

As before, the impasse point is given by the payoffs received by the players when they disagree perpetually. In stage 2, the impasse point for player 2 is $(1/r_2)\Pi_2(\bar{o}_2)$ and for player 0 it is $(1/r_0)U_0(o_1, \bar{o}_2)$. In stage 1, the impasse point for player 1 is $(1/r_1)\Pi_1(\bar{o}_1)$, and for 0 it is $(1/r_0)U_0(\bar{o}_1, \bar{o}_2)$. Note that the latter differs from the impasse point of player 0 at the first stage. This is a consequence from the fact that the status-quo varies from one stage to another, as we assume that at stage 2 the status quo is the agreement obtained at stage 1 (see Bowen *et al.* (2014) for a discussion on the role of alternative assumptions about the status quo).

The resolution of this version of the model yields the following proposition, focusing again on the case where $\Delta \rightarrow 0$, as we are interested in negotiations without frictions (existence and uniqueness of (e_1^s, e_2^s) and (τ_1^s, τ_2^s) are proven in Appendix A.2).

Proposition 2 *Assume the sequential procedure (game \mathcal{G}^s), with $\Delta \rightarrow 0$. The SSPE strategies are: (i) for player 0 to offer (e_1^s, τ_1^s) to player 1 at stage 1 and period $t = 0$, and to offer (e_2^s, τ_2^s) immediately to player 2 at stage 2 and period $t = \Gamma \rightarrow 0$, and (ii) for players 1 and 2 to accept the offers immediately. The equilibrium instantaneous emissions (e_1^s, e_2^s) are not socially optimal and imply:*

$$\frac{\partial B_1(e_1^s)}{\partial e_1} = \frac{\partial D(e_1^s, e_2^s)}{\partial e_1} + \frac{r_0}{r_0 + r_2} \left(\frac{\partial D(e_1^s, \bar{e}_2)}{\partial e_1} - \frac{\partial D(e_1^s, e_2^s)}{\partial e_1} \right) \quad (18)$$

$$\frac{\partial B_2(e_2^s)}{\partial e_2} = \frac{\partial D(e_1^s, e_2^s)}{\partial e_2} \quad (19)$$

After implementing the transfers (τ_1^s, τ_2^s) , shown in equations (63) and (64) in Appendix A.2, the instantaneous net equilibrium payoffs are:

$$U(o_1, o_2) - U(\bar{o}_1, \bar{o}_2) = \frac{r_1}{r_0 + r_1} \Psi(e_1^s, e_2^s; \bar{e}_1, \bar{e}_2) - \frac{r_0 r_1}{(r_0 + r_1)(r_0 + r_2)} \Psi(e_1^s, e_2^s; e_1^s, \bar{e}_2) \quad (20)$$

and

$$\Pi_1(o_1) - \Pi_1(\bar{o}_1) = \frac{r_0}{r_0 + r_1} \Psi(e_1^s, e_2^s; \bar{e}_1, \bar{e}_2) - \frac{r_0^2}{(r_0 + r_1)(r_0 + r_2)} \Psi(e_1^s, e_2^s; e_1^s, \bar{e}_2) \quad (21)$$

$$\Pi_2(o_2) - \Pi_2(\bar{o}_2) = \frac{r_0}{r_0 + r_2} \Psi(e_1^s, e_2^s; e_1^s, \bar{e}_2) \quad (22)$$

Proof. See Appendix A.2. ■

Equations (18) and (19) show that emissions do not maximize social welfare⁹, i.e. $e_i^s \neq e_i^o$ for $i = 1, 2$. Since the term $\left(\frac{\partial D(e_1^s, \bar{e}_2)}{\partial e_1} - \frac{\partial D(e_1^s, e_2^s)}{\partial e_1} \right)$ is positive due to the convexity of the damage function, it turns out that 1 will reduce more his emissions than the optimal situation for society, $e_1^s < e_1^o$, while 2 will pollute more, $e_2^s > e_2^o$. The net effect is that the net surplus in the sequential case is lower than in the simultaneous case, with $\Psi(e_1^s, e_2^s; \bar{e}_1, \bar{e}_2) < \Psi(e_1^o, e_2^o; \bar{e}_1, \bar{e}_2)$. As already advanced in the introduction, the intuition for this inefficiency comes from the fact that changing the bargaining protocol does not only change the distribution of the net surplus, but also the net surplus itself, as agents are bargaining over transfers and emission levels. Note also that a small value of r_2 relatively to r_0 ($r_2 \ll r_0$) yields more reduction from player 1, and more pollution from player 2.

In this configuration, players 0 and 1 get a share of the total net surplus minus a term that is given to player 2. The positivity of the net surplus $\Psi(e_1^s, e_2^s; e_1^s, \bar{e}_2)$ ensures that the net payoff of player 2 in (22) is also positive. The payoffs of player 0 and 1 given by (20)

⁹The social planner prefers the multilateral procedure over the sequential one, as it allows to take into account all the synergies between the two agreements. This is related to the result discussed in the multi-issues literature, where simultaneous bargaining is preferred because it allows the exploitation of all the synergies among issues. See Bac and Raff (1996), or Thompson (2008) for a more informal discussion.

and (21) are also positive under the condition $\Psi(e_1^s, e_2^s; \bar{e}_1, \bar{e}_2) > r_0/(r_0 + r_2)\Psi(e_1^s, e_2^s; e_1^s, \bar{e}_2)$. This condition states that the surplus when polluters 1 and 2 move from the threat to the equilibrium solution has to be greater than a fraction of the surplus when only one player moves from the threat to the equilibrium while the other plays his equilibrium offer. The least favorable case occurs when the discount rate r_2 is small relatively to r_0 ($r_2 \ll r_0$) such that the fraction $r_0/(r_0 + r_2)$ tends to 1.

The payoff obtained by polluter 2 depends on his own contribution to the global surplus obtained by moving to the optimal level of emissions, i.e. $\Psi(e_1^s, e_2^s; e_1^s, \bar{e}_2)$, and on his relative patience compared to the victim. On the contrary, the payoff obtained by polluter 1 depends on the relative patience of all three players.

The inside option of 0 in stage two, originated by the agreement reached in stage one, increases his impasse point and his bargaining power in stage 2. This gives 0 advantage when bargaining with 2, implying that the surplus shared in stage 2 is smaller. Thus, both 0 and 1 are better off at the expense of 2, which is only involved in the second stage of the negotiation. Since the problem is symmetric, a first stage between 0 and 2 (instead of 1) would imply $\frac{r_2}{r_0+r_2}\Psi(e_1^s, e_2^s; \bar{e}_1, \bar{e}_2) - \frac{r_0r_2}{(r_0+r_1)(r_0+r_2)}\Psi(e_1^s, e_2^s; \bar{e}_1, e_2^s)$ as the RHS of (20). The relative importance of these terms can be used by the victim, player 0, to decide whether to bargain first with polluter 1 or with polluter 2, showing that the order of the partners in the sequential process becomes a strategic variable. Simple manipulations show that the victim prefers to negotiate first with polluter 1 if:

$$r_0r_1(\Psi(e_1^s, e_2^s; \bar{e}_1, \bar{e}_2) - \Psi(e_1^s, e_2^s; e_1^s, \bar{e}_2)) > r_0r_2(\Psi(e_1^s, e_2^s; \bar{e}_1, \bar{e}_2) - \Psi(e_1^s, e_2^s; \bar{e}_1, e_2^s)) \quad (23)$$

Note that, for $r_1 = r_2$, this condition holds if $\Psi(e_1^s, e_2^s; \bar{e}_1, e_2^s) > \Psi(e_1^s, e_2^s; e_1^s, \bar{e}_2)$, implying that the polluter negotiates first with the player with a larger marginal contribution.

We present the Nash bargaining problem that would yield the results shown in Proposition 2 (again, this corollary can be easily generalized to a sequential negotiation with N polluters).

Corollary 2 *SSPE's emission levels, transfers and payoff pairs of \mathcal{G}^s with $\Delta \rightarrow 0$, shown in Proposition 2, converge to the asymmetric Nash Bargaining Solution (NBS) of the following problem: In stage 2, the Nash program between 0 and 2 is*

$$\max_{e_2, \tau_2} \Omega_{02} = (U(o_1, o_2) - U(o_1, \bar{o}_2))^{\frac{r_2}{r_0+r_2}} (\Pi_2(o_2) - \Pi_2(\bar{o}_2))^{\frac{r_0}{r_0+r_2}} \quad (24)$$

and in stage 1

$$\max_{e_1, \tau_1} \Omega_{01} = (U(o_1, o_2(o_1)) - U(\bar{o}_1, \bar{o}_2))^{\frac{r_1}{r_0+r_1}} (\Pi_1(o_1) - \Pi_1(\bar{o}_1))^{\frac{r_0}{r_0+r_1}} \quad (25)$$

Proof. See Appendix A.3. ■

A slight modification of this Corollary also shows that the inefficiency in terms of emissions can be removed by assuming that partial agreements are implemented only after all agreements have been reached, instead of immediately as assumed in Proposition 2 and Corollary 2. Under this alternative assumption the corresponding Nash Bargaining problem in stage 2 would be given by

$$\max_{e_2, \tau_2} \Omega_{02} = (U(o_1, o_2) - U(\bar{o}_1, \bar{o}_2))^{\frac{r_2}{r_0+r_2}} (\Pi_2(o_2) - \Pi_2(\bar{o}_2))^{\frac{r_0}{r_0+r_2}}$$

It turns out that SSPE emissions and transfers obtained with this program are the same as under multilateral bargaining, the ones shown in Proposition 1. This also shows that the inefficiency of the sequential procedure comes from the inside option of player 0, and, consequently, from its impasse point in the second stage. However, note that imposing that partial agreements are implemented only after all agreements have been reached would also imply some form of intervention by the government or an outside agent, as parties are otherwise free to implement their partial agreement.

4 Comparison of results

From Proposition 2 we know that with sequential bargaining, under the assumption that partial agreements are implemented immediately, emissions are not optimal and do not maximize the social surplus. This would be especially problematic if players involved in the first negotiation in the sequential procedure would both prefer the outcome in the sequential procedure, as they could initiate the sequential procedure without the intervention of the other polluter. As we have assumed, without loss of generality, that negotiations start with player 1 (i.e. that condition 23 holds), we analyze the conditions under which players 0 and 1 prefer this procedure.

Comparing Propositions 1 and 2 we can see that players 0 and 1 prefer the sequential procedure, with immediate implementation of partial agreements, to the multilateral case if:

$$\Psi(e_1^s, e_2^s; \bar{e}_1, \bar{e}_2) > \frac{r_2(r_0+r_1)}{r_2(r_0+r_1)+r_0r_1} \Psi(e_1^o, e_2^o; \bar{e}_1, \bar{e}_2) + \frac{r_0}{r_0+r_2} \Psi(e_1^s, e_2^s; e_1^s, \bar{e}_2) \quad (26)$$

In words, the total surplus gain in the sequential case, $\Psi(e_1^s, e_2^s; \bar{e}_1, \bar{e}_2)$, has to be larger than a fraction of the total surplus in the multilateral case, $\Psi(e_1^o, e_2^o; \bar{e}_1, \bar{e}_2)$, note that $\frac{r_2(r_0+r_1)}{r_2(r_0+r_1)+r_0r_1} < 1$, plus a term influenced by the marginal contribution of 2 in the sequential case, $\Psi(e_1^s, e_2^s; e_1^s, \bar{e}_2)$. This expression is more likely to hold if the total surplus in the sequential case is large. However, the expression also calls for a small contribution of 2 in the

sequential case (recall that condition 23 goes in the same direction). That is, players 0 and 1 are more likely to follow the sequential path if it has a large surplus to be shared, and there is a small contribution of 2 in the sequential case (which implies that 0 and 1 are going to appropriate most of the surplus generated in the sequential path).

Focusing on the role of the discount rates, condition (26) tends to hold when player 1 is very impatient (higher r_1) as this reduces the value of $\frac{r_2(r_0+r_1)}{r_2(r_0+r_1)+r_0r_1}$. The impact of an increase in r_0 or r_2 is not clear-cut, as it has a different impact on the two terms collecting discount rates.

The following proposition summarizes these results:

Proposition 3 *With sequential bargaining under the assumption that partial agreements are implemented immediately and $\Delta \rightarrow 0$, emissions are not optimal and do not maximize the social surplus. Moreover, both players involved in the first partial agreement prefer the sequential path instead of the multilateral path if condition (26) holds.*

We now show, for particular functions, cases and examples where condition (26) is satisfied, and hence both players involved in the first partial agreement prefer the sequential path instead of the multilateral path (see footnote 3). We consider in all cases that, in the sequential case, player 0 can bargain first with 1 and second with 2, or vice-versa (as we can no longer rename players at will once we have specified the values of the parameters defining them).

Consider the following specifications for the damage and benefit functions (1)-(2)-(3):

$$D(e_1, e_2) = \frac{d}{2}(e_1 + e_2)^2 \quad (27)$$

$$B_1(e_1) = c_1 e_1 - \frac{w_1}{2} e_1^2 \quad (28)$$

$$B_2(e_2) = c_2 e_2 - \frac{w_2}{2} e_2^2 \quad (29)$$

with d, c_1, c_2, w_1 and $w_2 > 0$. The threat emissions are obtained by taking the derivative of the benefit function with respect to the level of emissions, $\frac{\partial B_i(e_i)}{\partial e_i} = 0$, yielding $\bar{e}_i = \frac{c_i}{w_i}$ for $i = 1, 2$. The equilibrium emissions obtained in the multilateral negotiation, which are equal to the socially optimal emissions as shown above, for $i = 1, 2$ and $j \neq i$, are given by

$$e_i^o = \frac{d(w_i \bar{e}_i - w_j \bar{e}_j) + w_i w_j \bar{e}_i}{d(w_i + w_j) + w_i w_j}.$$

Using eqs. (18) and (19), the equilibrium emissions obtained in the sequential bargaining

are given by

$$\begin{aligned} e_1^s &= \frac{(w_1(d+w_2)\bar{e}_1 - dw_2\bar{e}_2)(r_0+r_2) - d^2r_0\bar{e}_2}{d^2r_0 + (d(w_1+w_2) + w_1w_2)(r_0+r_2)}, \\ e_2^s &= \frac{(w_2(d+w_1)\bar{e}_2 - dw_1\bar{e}_1)(r_0+r_2) + d^2r_0\bar{e}_2}{d^2r_0 + (d(w_1+w_2) + w_1w_2)(r_0+r_2)}. \end{aligned}$$

We can now write the following Proposition.

Proposition 4 *Assuming the functional forms shown in (27)-(29), and that both polluters have identical benefit functions, i.e. $c_i = c$, $w_i = w$ for $i = 1, 2$, players involved in the first partial agreement always prefer the sequential path to the multilateral negotiation if $r_0 \leq \max\{r_1, r_2\}$.*

Proof. Consider first the case where $r_1 \geq r_0 \geq r_2$, with at least one strict inequality. In this case, the victim prefers to negotiate first with player 1 in the sequential case as, for $c_i = c$, $w_i = w$ for $i = 1, 2$, we have that $\Psi(e_1^s, e_2^s; e_1^s, \bar{e}_2) = \Psi(e_1^s, e_2^s; \bar{e}_1, e_2^s)$ and condition (23) simplifies to $r_1 > r_2$. Furthermore, condition (26) can be rewritten as

$$\frac{r_1}{r_2} (1 + 2dw^{-1}) + \frac{r_1}{r_0} \left(1 + 2\frac{dw^{-1}}{1 + dw^{-1}} \right) > 1 \quad (30)$$

and this holds for $r_1 \geq r_0 \geq r_2$.

When $r_2 \geq r_0 \geq r_1$, with at least one strict inequality, the victim prefers to negotiate first with player 2 in the sequential case and condition (26) also holds as it simplifies to

$$\frac{r_2}{r_1} (1 + 2dw^{-1}) + \frac{r_2}{r_0} \left(1 + 2\frac{dw^{-1}}{1 + dw^{-1}} \right) > 1, \quad (31)$$

When $r_2 \geq r_0$ and $r_1 \geq r_0$ condition (26) also holds. To see this note that it simplifies to (30) if $r_1 > r_2$ as player 0 prefers to negotiate first with player 1; to (31) if $r_2 > r_1$ as player 0 prefers to negotiate first with player 2; and to either (30) or (31) if $r_1 = r_2$ or $r_1 = r_2 = r_0$, as the victim is indifferent between negotiating first with player 1 or 2. ■

Hence, similar polluters (in terms of benefits) and a relatively patient victim (in comparison to at least one of the polluters), or equally patient players, suffice to launch the sequential path. To illustrate numerically these results consider the following two examples, which are based on the benefit and damage functions (27)-(29) and for which condition (26) holds.

Example 1 *Let $r_i = 1\%$, $c_i = 10$, and $w_i = d = 1$ for $i = 0, 1, 2$. Equilibrium emissions are shown in Table 1 and instantaneous payoffs in Table 2. Players 0 and 1 prefer the sequential*

path that starts with their bilateral negotiation over the multilateral procedure (they obtain a payoff equal to 57.4, which is higher than what they obtain in the multilateral case). However the sequential negotiation yields a lower total surplus (130.62 compared to 133.32) and lower total emissions (5.70 compared to 6.66). The same holds with player 2, making player 0 indifferent to choose with whom to negotiate first.

Example 2 Let players differ only in their discount rates, with $r_0 = 1\%$, $r_1 = 10\%$ and $r_2 = 0.4\%$; and $c_1 = c_2 = 10$, $w_1 = w_2 = d = 1$. As shown in Table 2, both players 0 and 1 are interested in choosing the sequential path that starts with the negotiation between these two players (player 0 obtains a payoff equal to 97.9 and player 1 a payoff of 9.79). As also shown in Table 2, the overall surplus created with this particular sequential path is the smallest one from the alternatives considered. Emissions are also lower in the sequential path initiated with player 1 (see Table 1).

Finally, note that Proposition 4 only contains a sufficient condition and that condition (26) can also hold when the polluters differ in their benefit functions, as the next example, also based on the benefit and damage functions introduced in (27)-(29), shows.

Example 3 Let players differ in their discount rates and in their benefit functions, with $r_0 = 0.5\%$, $r_1 = 17\%$ and $r_2 = 1\%$; and $c_1 = c_2 = 10$, $w_1 = 1, w_2 = 0.6$ and $d = 1$. Emissions are shown in Table 1 and instantaneous payoffs can be found in Table 2. Both players 0 and 1 are interested in choosing the sequential path that starts with their bilateral negotiation. Again, the overall surplus created with this sequential path is smaller than in the multilateral case, while emissions are also lower.

5 Discussion

In this section we show first that optimality is restored when negotiations deal only over transfers. Then, we discuss informally the likely impact of modifying other key assumptions of our analysis.

5.1 Negotiations only over transfers

Using our main model, we have shown above that negotiations over emission levels and transfers can lead to an inefficient outcome, as sequential negotiations may be preferred by both players involved in the first negotiation, and this path would lead to socially inefficient emissions. To disentangle the role of our assumption that players negotiate simultaneously

over multiple issues, emissions and transfers, we simplify the model above and assume that players negotiate exclusively over transfers. In the multilateral bargaining protocol we assume that players negotiate about the implementation of the emission levels that maximize the overall surplus, as they have no reason to leave unexhausted surplus. In the sequential procedure, we adopt a sequential version of passive beliefs (McAfee and Schwartz, 1994), and assume that players in the first sequential negotiation bargain over the level that maximizes their bilateral surplus, assuming that in the second stage the emission level agreed upon will also be the one that maximizes the bilateral surplus of the players involved in that negotiation. This yields the following proposition (we added a τ in superscript to indicate values obtained in this alternative framework):

Proposition 5 *Assume $\Delta \rightarrow 0$ and that players negotiate exclusively over transfers. The form of the equilibrium strategies is analogous to those described in Propositions 1 and 2:*

(i) *For the transfers only - multilateral bargaining procedure (game $\mathcal{G}^{m\tau}$), the emission levels at the equilibrium are equal to the socially optimal emission levels, i.e. $(e_1^{*\tau}, e_2^{*\tau}) = (e_1^o, e_2^o)$ and net equilibrium instantaneous payoffs are equal to those shown in (9) to (11).*

(ii) *For the transfers only - sequential procedure (game $\mathcal{G}^{s\tau}$) the emission levels at the equilibrium are equal to the socially optimal emission levels, i.e. $(e_1^{s\tau}, e_2^{s\tau}) = (e_1^o, e_2^o)$, and net equilibrium instantaneous payoffs are:*

$$\begin{aligned} U(o_1, o_2) - U(\bar{o}_1, \bar{o}_2) &= \frac{r_1}{r_0 + r_1} \Psi(e_1^o, e_2^o; \bar{e}_1, \bar{e}_2) - \frac{r_0 r_1}{(r_0 + r_1)(r_0 + r_2)} \Psi(e_1^o, e_2^o; e_1^o, \bar{e}_2) \\ \Pi_1(o_1) - \Pi_1(\bar{o}_1) &= \frac{r_0}{r_0 + r_1} \Psi(e_1^o, e_2^o; \bar{e}_1, \bar{e}_2) - \frac{r_0^2}{(r_0 + r_1)(r_0 + r_2)} \Psi(e_1^o, e_2^o; e_1^o, \bar{e}_2) \\ \Pi_2(o_2) - \Pi_2(\bar{o}_2) &= \frac{r_0}{r_0 + r_2} \Psi(e_1^o, e_2^o; e_1^o, \bar{e}_2) \end{aligned}$$

Both players involved in the first partial agreement prefer the sequential to the multilateral path if:

$$\left(1 - \frac{r_2(r_0 + r_1)}{r_2(r_0 + r_1) + r_0 r_1}\right) \Psi(e_1^o, e_2^o; \bar{e}_1, \bar{e}_2) > \frac{r_0}{r_0 + r_2} \Psi(e_1^o, e_2^o; e_1^o, \bar{e}_2)$$

Proof. Follow the proofs for Propositions 1 to 3 but considering that offers and counter-offers only deal over transfers, assuming that emissions in case of agreement maximize (i) the global surplus under the multilateral protocol or (ii) the bilateral surplus of the players involved in each one of the stages of the sequential protocol. For the $\mathcal{G}^{m\tau}$ game, the SSPE strategies are now: for player 0 to offer, at period $t = 0$, $\tau_1^{*\tau}$ to player 1 and $\tau_2^{*\tau}$ to player 2, and for players 1 and 2 to accept the offers immediately. The emission levels at the equilibrium are given by (5). For the $\mathcal{G}^{s\tau}$ game, the SSPE strategies are: for player 0 to offer $\tau_1^{s\tau}$ to player 1 at stage 1 and period $t = 0$, and to offer $\tau_2^{s\tau}$ immediately afterwards to player

2 at stage 2 and period $t = \Gamma \rightarrow 0$, and for players 1 and 2 to accept the offers immediately. The emission levels at the equilibrium are given by (5). ■

That is, emissions are now optimal under both procedures, but the sequential path can still be preferred by both players involved in the first partial agreement. However, from the point of view of the social planner this is not problematic, as the global surplus is maximized and the optimal level of emissions is attained in both cases. Only the distribution of the surplus is different under the two procedures considered.

5.2 Alternative assumptions

A key assumption of our model is that player 0 can choose with which player to negotiate first. One could imagine a situation in which the player that would intervene in the first negotiation in the sequential case would be chosen randomly. Our result that emissions are not optimal if the sequential path is followed would continue to hold, as they do not depend on the identity of the first player in the sequential path. The results that the sequential path may be preferred by both players involved in the first negotiation in the sequential case would also hold, although with some modifications. In our model, the polluter negotiates first with the player with a larger marginal contribution. If the random process selects this player, our results would remain unchanged. If the other player is selected, the sequential path will be chosen less often, as the second term on the RHS of condition (26) would be larger (because now the player contributing more would negotiate at the second stage). Thus, in expectation the sequential path may still be selected, but less often.

One could also imagine that players 1 and 2 compete to become the first player in the sequential path. This would modify the payoffs in the sequential path in all cases, as player 0 would probably obtain a larger share of the surplus. The results that emissions are not optimal if the sequential path is followed would continue to hold, for the reason discussed above. However, the sequential path would now be more attractive for player 0, as players 1 and 2 would only compete when it is attractive for them to be the first player.

Another alternative that one could consider is that players 1 and 2 are allowed to form a coalition before the negotiation starts. This strategy was investigated, in a framework where multilateral negotiations are compared to a set of bilateral simultaneous negotiations, in Chipty and Snyder (1999) and in Caparrós and Perea (2013). The results of these papers suggest that, indeed, under some circumstances this coalition would be formed, although not always. Determining the impact of interacting this possibility with our sequential procedure would require a new model, which we leave for future research.

Considering that pollution is a stock variable would also modify the analysis. For our

laundry, or Aivazian and Callen's laundry, flow pollution is the adequate assumption, but there are many relevant issues, such as climate change, where assuming a stock pollutant is more adequate. In this case one should probably focus on stationary Markov SPE, where the state variable is given by the stock of emissions, potentially giving the polluters the option to collude if there is no agreement. As this implies a substantial modification of our game, we have also left this for future research.

One could even consider that initiating the sequential case would require the consensus of all. This would substantially modify the game and potentially lead to multiple equilibria. However, note that imposing this consensus rule to start the sequential process would already require some form of intervention by the government or an external agent. On the contrary, we are interested in the case where two players can reach an independent agreement if they so wish. That is, nothing prevents our neighborhood laundry from negotiating first with one of the polluting factories, and the manager of the other polluting factory has no say in this negotiation.

Also the rules in the multilateral case could be modified, for example assuming that instead of unanimity there is a form of majority voting, as in Baron and Ferejohn (1989). Note first that we would need a modified version of the majority rule, as the victim would need to be part of any agreement. Thus, the fact that in the multilateral case the agreement would need only one of the polluters, resembling the first stage of our sequential case. Note also that, for our game with a potentially infinite number of negotiation rounds, this procedure may yield multiple equilibria. In any case, the results that negotiations are not optimal if the sequential path is chosen would continue to hold; in fact, emissions under the multilateral case may also not be optimal. The conditions under which the sequential procedure is preferred would certainly vary, although sequential negotiations may still be preferred under some circumstance. We also leave the details of this alternative model for future research.

6 Conclusion

This article has analyzed a bargaining game between two polluters and a victim over two variables, polluting emissions and monetary compensations (transfers). Results show that multilateral (simultaneous) bargaining leads to a global agreement which is optimal, as social welfare is maximized. On the contrary, the sequential procedure considered leads to agreements that are not efficient for society. More importantly, we have shown that if players are free to choose the bargaining protocol (multilateral or sequential), the solution may be inefficient. The reason is that when players bargain over emissions and transfers, and not only over transfers as commonly assumed, not only the distribution of the net surplus varies

with the protocol, also the pollution level, and hence the net surplus itself, varies with the protocol. This holds despite the fact that all players are interested in reaching the grand coalition, where all possible contracts are in place, with both protocols; and even when time between offers and between negotiation rounds tends to zero. We have also highlighted the key role of inside options (temporary disagreement payoffs) in explaining this inefficiency, even though all possible agreements are reached immediately in equilibrium and, therefore, inside options are never implemented. Finally, we have shown that if partial agreements in the sequential negotiation are only implemented after the whole sequence of agreements are struck, the solution is optimal from the point of view of society.

Our results show that the Coase Theorem only holds if sequential negotiations are ruled out, or if the possibility to implemented partial agreements is prohibited by a central authority. Note that this result is not restricted to the case of two polluters and holds also for a larger number of polluters. As for other difficulties associated with the Coase Theorem, one could argue that in a Coasean world the players should be able to find the efficient bargaining protocol. However, we have shown that, under certain conditions, both players involved in the first negotiation of the sequential path prefer the outcome of this negotiation to the one that they would obtain in a multilateral negotiation. Thus, without government intervention, nothing would prevent them from signing and implementing this partial agreement, which would launch a path leading to a socially inefficient situation.

References

- [1] Aivazian, V. A., and Callen, J. L. (1981), The Coase Theorem and the Empty Core, *Journal of Law and Economics* 24(1):175–81.
- [2] Aivazian, V. A., Callen, J. L. and Lipnowski, I. (1987), The Coase Theorem and Coalitional Stability, *Economica* 54(216):517-520.
- [3] Bac, M. and Raff, H. (1996), Issue-by-issue negotiations: the role of information and time preference, *Games and Economic Behavior*, 13(1):125–134.
- [4] Banerji, A. (2002), Sequencing strategically: wage negotiations under oligopoly, *International Journal of Industrial Organization*, 20:1037-1058.
- [5] Baron, D.P., and Ferejohn, J.A. (1989), Bargaining in legislatures, *American Political Science Review*, 83,1181–1206
- [6] Busch, L.A and I. Horstmann (1997), Bargaining Frictions, Bargaining Procedures and Implied Costs in Multiple-Issue Bargaining, *Economica*, 64(256):669-680

- [7] Cai, H (2000), Delay in Multilateral Bargaining under Complete Information, *Journal of Economic Theory*, 93:260-276
- [8] Caparrós, A. and J.C. Péreau (2013), Forming coalitions to negotiate North–South climate agreements, *Environment and Development Economics*, 18:69-92
- [9] Caparrós, A. and J.C. Péreau (2017), Multilateral versus sequential negotiations over climate change, *Oxford Economic Papers*, 69(2):365–387.
- [10] Chipty, T and C.M Snyder (1999) The role of firm size in bilateral bargaining: A study of the cable television industry, *The review of Economics and Statistics*, 81(2):326-340
- [11] Coase, R.H. (1960), The Problem of Social Cost, *Journal of Law and Economics*, 3:1-44.
- [12] Coase, R.H. (1981), The Coase Theorem and the Empty Core: A Comment, *Journal of Law and Economics* 24(1):183-187.
- [13] Dixit, A. and M. Olson (2000), Does voluntary participation undermine the Coase Theorem?, *Journal of Public Economics*, 76: 309-335.
- [14] Farrell, J. (1987), Information and the Coase Theorem, *Journal of Economic Perspectives*, 1(2):113-129.
- [15] Fernandez, R. and J. Glazer (1991), Striking for a bargain between two completely informed agents, *American Economic Review*, 81:240-252.
- [16] Fershtman, C. (1990), The Importance of the Agenda in Bargaining, *Game and Economic Behavior*, 2:224-238.
- [17] Flamini, F. (2007) First things first? The agenda formation problem for multi-issue committees, *Journal of Economic Behavior and Organization*, 63:138–157
- [18] Haller, H and S. Holden (1990), A letter to the editor on wage bargaining, *Journal of Economic Theory*, 52:232-3
- [19] Hu, T.W. and Rocheteau, G. (2020), Bargaining under liquidity constraints: Unified strategic foundations of the Nash and Kalai solutions, *Journal of Economic Theory*, 189, 105098.
- [20] Huang, C.-Y. (2002), Multilateral bargaining: conditional and unconditional offers, *Economic Theory*, 20:401–12.

- [21] Kenan J. and R. Wilson (1993), Bargaining with Private Information, *Journal of Economic Literature*, 31:45-104.
- [22] Krishna, V. and R. Serrano (1996), Multilateral bargaining, *Review of Economic Studies*, 63:61–80.
- [23] Lee, J and H. Sabourian (2007), Coase theorem, complexity and transaction costs, *Journal of Economic Theory*, 135:214-235
- [24] McAfee, P. and M. Schwartz (1994), Opportunism in multilateral vertical contracting: nondiscrimination, exclusivity, and uniformity, *American Economic Review*, 84:210-230.
- [25] Muthoo, A (1999), *Bargaining Theory with Applications*, Cambridge University Press.
- [26] Okada, A. (1996), A Noncooperative Coalitional Bargaining Game with Random Proposers, *Games and Economic Behavior* 16, 97–108.
- [27] Osborne, M. and Rubinstein, A. (1994) *A course in Game Theory*, MIT Press, Cambridge, Massachusetts.
- [28] Ray, D. and Vohra, R., (2015). Coalition Formation. In H. Peyton Young, Shmuel Zamir (eds.) *Handbook of Game Theory with Economic Applications*, volume 4: 239-326.
- [29] Rubinstein A. (1982), Perfect equilibrium in a bargaining model, *Econometrica*, 50:97-109.
- [30] Sutton, J. (1986), Non-cooperative bargaining theory: an introduction, *Review of Economic Studies*, 53:709–24.
- [31] Thompson, L. (2008). *The truth about negotiations*. FT Press, Pearson Education, New Jersey.

A Proofs

A.1 Proposition 1

A.1.1 Indifference, existence, uniqueness and delay

Assume first, to be shown below, that there is no delay, i.e. that whenever a player has to make an offer, her equilibrium offer is accepted immediately by all other players. To simplify the presentation, we add the notation $\delta_i = e^{-r_i \Delta}$, $i = 0, 1, 2$. Consider a point where 0 has

to make an offer to 1 and 2. Player 1's equilibrium payoff from rejecting an offer from 0 is $\delta_1(1/(1-\delta_1))\Pi_1^{(1)*} + \Delta\Pi_1(\bar{o}_1)$, with $\Delta > 0$ and small. The reason is that after rejecting an offer from 0, player 1 offers an instantaneous payoff equal to $\Pi_1^{(1)*}$ in the next period, and this is accepted by our no delay assumption, yielding a continuous stream with a discounted value equal to $(1/(1-\delta_1))\Pi_1^{(1)*}$ from the next period onwards, plus an interim payoff (inside option) equal to $\Delta\Pi_1(\bar{o}_1)$.

Perfection requires that 1 rejects the offer from 0 when $(1/(1-\delta_1))\Pi_1^{(0)*} < (1/(1-\delta_1))\delta_1\Pi_1^{(1)*} + \Delta\Pi_1(\bar{o}_1)$ and accepts it when $(1/(1-\delta_1))\Pi_1^{(0)*} \geq (1/(1-\delta_1))\delta_1\Pi_1^{(1)*} + \Delta\Pi_1(\bar{o}_1)$. In addition, note that $\Pi_1^{(0)*} \geq \Pi_1(\bar{o}_1)$ and $\Pi_1^{(1)*} \geq \Pi_1(\bar{o}_1)$ have to hold, as both offers need to provide instantaneous payoffs that exceed the instantaneous threat payoff. To take this into account, we re-write the condition under which player 1 accepts the offer from 0 in net terms, that is:

$$(1/(1-\delta_1))W_1^{(0)*} \geq (1/(1-\delta_1))\delta_1W_1^{(1)*} + \Delta\Pi_1(\bar{o}_1), \quad (32)$$

where $W_1^{(i)} = \Pi_1^{(i)*} - \Pi_1(\bar{o}_1)$ and the superscript (i) continues to stand for the player proposing the offer. This ensure that any positive solutions of (32), in terms of $W_1^{(0)*}$ and $W_1^{(1)*}$, implies that $\Pi_1^{(0)*} \geq \Pi_1(\bar{o}_1)$ and $\Pi_1^{(1)*} \geq \Pi_1(\bar{o}_1)$. (For future reference, we also introduce the notation $W_0^{(i)} = U_0^{(i)*} - U_0(\bar{o}_1; \bar{o}_2)$ and $W_2^{(i)} = \Pi_2^{(i)*} - \Pi_2(\bar{o}_2)$.)

It follows from our no delay assumption, because 1 has to accept the equilibrium offer made by 0, that (32) holds. However, (32) cannot hold as an inequality, because otherwise 0 could increase her net payoff by offering $W_0^{(0)'} > W_0^{(0)*}$, implying a $W_1^{(0)'}$ such that

$$(1/(1-\delta_1))W_1^{(0)*} > (1/(1-\delta_1))W_1^{(0)'} > (1/(1-\delta_1))\delta_1W_1^{(1)*} + \Delta\Pi_1(\bar{o}_1).$$

Hence, (32) has to hold as an equality (or an indifference condition):

$$\frac{1}{1-\delta_1}W_1^{(0)*} = \frac{\delta_1}{1-\delta_1}W_1^{(1)*} + \Delta\Pi_1(\bar{o}_1) \quad (33)$$

By the same token, perfection requires

$$\frac{1}{1-\delta_2}W_2^{(0)*} = \frac{\delta_2}{1-\delta_2}W_2^{(1)*} + \Delta\Pi_2(\bar{o}_2) \quad (34)$$

$$\frac{1}{1-\delta_0}W_0^{(1)*} = \frac{\delta_0}{1-\delta_0}W_0^{(2)*} + \Delta U_0(\bar{o}_1; \bar{o}_2) \quad (35)$$

$$\frac{1}{1-\delta_2}W_2^{(1)*} = \frac{\delta_2}{1-\delta_2}W_2^{(2)*} + \Delta\Pi_2(\bar{o}_2) \quad (36)$$

$$\frac{1}{1-\delta_0}W_0^{(2)*} = \frac{\delta_0}{1-\delta_0}W_0^{(0)*} + \Delta U_0(\bar{o}_1; \bar{o}_2) \quad (37)$$

$$\frac{1}{1-\delta_1}W_1^{(2)*} = \frac{\delta_1}{1-\delta_1}W_1^{(0)*} + \Delta\Pi_1(\bar{o}_1) \quad (38)$$

and, for $i = 0, 1, 2$,

$$W_0^{(i)*} + W_1^{(i)*} + W_2^{(i)*} = \Psi(e_1^o, e_2^o; \bar{e}_1, \bar{e}_2) \quad (39)$$

The unique solution to this system of equations is:

$$W_0^{(0)*} = \frac{(1 - \delta_1)(1 - \delta_2)(\delta_2 + \delta_1\delta_2 + 1)\Psi(e_1^o, e_2^o; \bar{e}_1, \bar{e}_2)}{1 - \delta_0^2\delta_1 - \delta_0\delta_2^2 - \delta_1^2\delta_2 + \delta_0\delta_1\delta_2 + \delta_0^2\delta_1^2\delta_2^2} + k_0\Delta, \quad (40)$$

$$W_0^{(1)*} = \frac{\delta_0^2(\delta_1 - 1)(\delta_2 - 1)(\delta_2 + \delta_1\delta_2 + 1)\Psi(e_1^o, e_2^o; \bar{e}_1, \bar{e}_2)}{1 - \delta_0^2\delta_1 - \delta_0\delta_2^2 - \delta_1^2\delta_2 + \delta_0\delta_1\delta_2 + \delta_0^2\delta_1^2\delta_2^2} + k_1\Delta \quad (41)$$

$$W_0^{(2)*} = \frac{\delta_0(1 - \delta_1)(1 - \delta_2)(\delta_2 + \delta_1\delta_2 + 1)\Psi(e_1^o, e_2^o; \bar{e}_1, \bar{e}_2)}{1 - \delta_0^2\delta_1 - \delta_0\delta_2^2 - \delta_1^2\delta_2 + \delta_0\delta_1\delta_2 + \delta_0^2\delta_1^2\delta_2^2} + k_2\Delta \quad (42)$$

$$W_1^{(0)*} = \frac{\delta_1(1 - \delta_0)(1 - \delta_2)(\delta_0 + \delta_0\delta_2 + 1)\Psi(e_1^o, e_2^o; \bar{e}_1, \bar{e}_2)}{1 - \delta_0^2\delta_1 - \delta_0\delta_2^2 - \delta_1^2\delta_2 + \delta_0\delta_1\delta_2 + \delta_0^2\delta_1^2\delta_2^2} + k_3\Delta \quad (43)$$

$$W_1^{(1)*} = \frac{(1 - \delta_0)(1 - \delta_2)(\delta_0 + \delta_0\delta_2 + 1)\Psi(e_1^o, e_2^o; \bar{e}_1, \bar{e}_2)}{1 - \delta_0^2\delta_1 - \delta_0\delta_2^2 - \delta_1^2\delta_2 + \delta_0\delta_1\delta_2 + \delta_0^2\delta_1^2\delta_2^2} + k_4\Delta \quad (44)$$

$$W_1^{(2)*} = \frac{\delta_1^2(\delta_0 - 1)(\delta_2 - 1)(\delta_0 + \delta_0\delta_2 + 1)\Psi(e_1^o, e_2^o; \bar{e}_1, \bar{e}_2)}{1 - \delta_0^2\delta_1 - \delta_0\delta_2^2 - \delta_1^2\delta_2 + \delta_0\delta_1\delta_2 + \delta_0^2\delta_1^2\delta_2^2} + k_5\Delta \quad (45)$$

where k_i are constants defined by $\delta_0, \delta_1, \delta_2, U_0(\bar{o}_1; \bar{o}_2), \Pi_1(\bar{o}_1)$ and $\Pi_2(\bar{o}_2)$; and $\Psi(e_1^o, e_2^o; \bar{e}_1, \bar{e}_2)$ uses the notation introduced in (4). These equations, together with (39), define the equilibrium offers at any period of the game, with player i making an offer at time Δt .

Focusing on the offers made by player 0, the relevant indifference conditions are (40) and (43), plus a third one that is obtained specifying equation (39) for $i = 0$ and yields:

$$W_2^{(0)*} = \frac{\delta_2^2(1 - \delta_0)(1 - \delta_1)(\delta_1 + \delta_0\delta_1 + 1)\Psi(e_1^o, e_2^o; \bar{e}_1, \bar{e}_2)}{1 - \delta_0^2\delta_1 - \delta_0\delta_2^2 - \delta_1^2\delta_2 + \delta_0\delta_1\delta_2 + \delta_0^2\delta_1^2\delta_2^2} + (1 - k_0 - k_3)\Delta$$

Using $\delta_i = (1 - r_i\Delta)$ and for $\Delta \rightarrow 0$, these three equations simplify to (9), to (11), with $W_i^{(0)*} > 0$ for $i = 1, 2, 3$.

It is trivial to show that the offers from player 0 just described, and the equivalents for the remaining players, form indeed an SSPE for $\Delta \rightarrow 0$, which has no delay by assumption. Deviating by offering more to players 1 and 2 is not profitable for 0, as it will be accepted, and offering them less is also not profitable, as the offer will be rejected, and the equilibrium counteroffer yields a perpetual net instantaneous payoff for 0 that is less than or equal to $\max\{\delta_0 W_0^{(1)*} + \Delta \Pi_1(\bar{o}_1; \bar{o}_2), \delta_0^2 W_0^{(2)*} + \Delta(1 + \delta_1)\Pi_1(\bar{o}_1), \delta_1^3 W_0^{(0)*} + \Delta(1 + \delta_1 + \delta_1^2)\Pi_1(\bar{o}_1)\}$, or, for $\Delta \rightarrow 0$, $\max\{\delta_0 W_0^{(1)*}, \delta_0^2 W_0^{(2)*}, \delta_1^3 W_0^{(0)*}\}$.

We show now, by contradiction, that there is no SSPE with delay when $\Delta \rightarrow 0$. Suppose that there is an SSPE in which player 0's equilibrium offer, $W_1^{(0)*}$, is reject (recall that player 0 opens the game). Suppose further that it is player $i = 1, 2$ who rejects the offer from 0. That implies that either her net payoff with the offer from 0 is smaller than what she expects

in the next round from the equilibrium offer of player 1 (who might be herself):

$$\frac{1}{1 - \delta_i} W_i^{(0)*} < \frac{\delta_i}{1 - \delta_i} W_i^{*(1)} + \Delta \Pi_i(\bar{o}_i), \quad (46)$$

or it is smaller than what she expects to get in two periods from the equilibrium offer of player 2:

$$\frac{1}{1 - \delta_i} W_i^{(0)*} < \frac{\delta_i^2}{1 - \delta_i} W_i^{*(2)} + \Delta \Pi_i(\bar{o}_i), \quad (47)$$

From (36), the equivalent for player 1, and the equivalent indifference condition obtained applying the logic described above to the offers between periods 2 and 3 (where player 0 talks again), we obtain, for $i = 1, 2$

$$\begin{aligned} & \frac{1}{1 - \delta_i} W_i^{(1)*} \\ = & \frac{\delta_i}{1 - \delta_i} W_i^{(2)*} + \Delta \Pi_i(\bar{o}_i) \\ = & \delta_1 \left(\frac{\delta_i}{1 - \delta_i} W_i^{*(0)} \right) + (1 + \delta_i) \Delta \Pi_i(\bar{o}_i), \end{aligned} \quad (48)$$

For $\Delta \rightarrow 0$ and using (48), condition (46) simplifies to $W_i^{*(0)} < \delta_i^2 W_i^{*(0)}$ and equation (47) simplifies to $W_i^{*(0)} < \delta_i^3 W_i^{*(0)}$, for $i = 1, 2$. Thus, both yield a contradiction. Note that longer delays are also not possible because due to our stationarity assumption players cannot expect higher payoffs in future rounds. This completes the proof that there is a unique no delay SSPE for $\Delta \rightarrow 0$.

A.1.2 Determination of e_i^* and τ_i^*

Now, to characterize the SSPE in terms of e_i^* and τ_i^* , for $i = 1, 2$, we substitute the payoff functions (1) to (3) in the set of indifference conditions (33) to (38). We first determine the equilibrium emissions, as a function of equilibrium transfers, and we then obtain the equilibrium transfers.

When deciding the amount of emissions to offer, player 0 sets $(e_1^{(0)}, e_2^{(0)})$ at the level that maximize her payoff, taking into account that the transfers have to satisfy the indifference conditions derived in (33) and (34), the two affecting her offers. As we have shown that the indifference conditions defined in (33) and (34) hold as equalities, noting that all periods after a multilateral agreement are equal and taking into account $U_0(e_1^{(0)}, e_2^{(0)}) \geq U_0(\bar{e}_1, \bar{e}_2)$, we can rewrite the optimization problem solved by player 0 as

$$\max_{e_1^{(0)}, e_2^{(0)}} U_0 = -\tau_1^{(0)} - \tau_2^{(0)} - D(e_1^{(0)}, e_2^{(0)}) = -g_1(e_1^{(0)}, \tau_1^{(1)}, \cdot) - g_2(e_2^{(0)}, \tau_2^{(1)}, \cdot) - D(e_1^{(0)}, e_2^{(0)})$$

with

$$\begin{aligned}\tau_1^{(0)} &= -B_1(e_1^{(0)}) + \delta_1 \left(B_1(e_1^{(1)}) + \tau_1^{(1)} \right) + (1 - \delta_1) (\Delta B_1(\bar{e}_1) + B_1(\bar{e}_1)) = g_1(e_1^{(0)}, \tau_1^{(1)}, \cdot) \\ \tau_2^{(0)} &= -B_2(e_2^{(0)}) + \delta_2 \left(B_2(e_2^{(1)}) + \tau_2^{(1)} \right) + (1 - \delta_2) (\Delta B_2(\bar{e}_2) + B_2(\bar{e}_2)) = g_2(e_2^{(0)}, \tau_2^{(1)}, \cdot),\end{aligned}$$

where g_i , with $i = 1, 2$, are obtained substituting (1) to (3) into (33) and (34). Note that to simplify the notation, we write function $g_1(e_1^{(0)}, e_1^{(1)}, \tau_1^{(1)}, \bar{e}_1, \delta_1)$ as $g_1(e_1^{(0)}, \tau_1^{(1)}, \cdot)$, specifying as arguments only those variable that are directly relevant for our proof, denoting all the other variables by (\cdot) . We use the same notation for the remaining functions g_i ($i = 1, \dots, 8$) that we define. The maximization program can be rewritten as

$$\begin{aligned}\max_{e_1^{(0)}, e_2^{(0)}} W_0 &= -\tau_1^{(0)} - \tau_2^{(0)} - D(e_1^{(0)}, e_2^{(0)}) + D(\bar{e}_1, \bar{e}_2) \\ &= -g_1(e_1^{(0)}, \tau_1^{(1)}, \cdot) - g_2(e_2^{(0)}, \tau_2^{(1)}, \cdot) - D(e_1^{(0)}, e_2^{(0)}) + D(\bar{e}_1, \bar{e}_2).\end{aligned}$$

The corresponding FOC are

$$\begin{aligned}\frac{\partial W_0}{\partial e_1} &= 0 \Leftrightarrow \frac{\partial B_1(e_1^{(0)})}{\partial e_1} = \frac{\partial D(e_1^{(0)}, e_2^{(0)})}{\partial e_1} \\ \frac{\partial W_0}{\partial e_2} &= 0 \Leftrightarrow \frac{\partial B_2(e_2^{(0)})}{\partial e_2} = \frac{\partial D(e_1^{(0)}, e_2^{(0)})}{\partial e_2}\end{aligned}\tag{49}$$

Using a similar argument as for player 0, player 1 solves his net payoff (integrating $\Pi_1(e_1^{(1)}) \geq \Pi_1(\bar{e}_1)$)

$$\max_{e_1^{(1)}} W_1 = g_3(e_1^{(1)}, \cdot) + B_1(e_1^{(1)}) - B_1(\bar{e}_1)$$

with

$$\begin{aligned}g_3(e_1^{(1)}, \tau_2^{(1)}, \tau_2^{(2)}, \cdot) &= -\tau_2^{(1)} - D(e_1^{(1)}, e_2^{(1)}) - \delta_0 \left(-\tau_1^{(2)} - \tau_2^{(2)} - D(e_1^{(2)}, e_2^{(2)}) \right) \\ &\quad + (1 - \delta_0) (\Delta D(\bar{e}_1; \bar{e}_2) + D(\bar{e}_1; \bar{e}_2))\end{aligned}$$

and the FOC is:

$$\frac{\partial W_1}{\partial e_1} = 0 \Leftrightarrow \frac{\partial B_1(e_1^{(1)})}{\partial e_1} = \frac{\partial D(e_1^{(1)}, e_2^{(1)})}{\partial e_1}\tag{50}$$

Finally, player 2 solves his net payoff (integrating $\Pi_2(e_2^{(2)}) \geq \Pi_2(\bar{e}_2)$)

$$\max_{e_2^{(2)}} W_2 = g_4(e_2^{(2)}, \cdot) + B_2(e_2^{(2)}) - B_2(\bar{e}_2)$$

with

$$g_4(e_2^{(2)}, \tau_1^{(2)}, \tau_1^{(0)}, \tau_2^{(0)}, \cdot) = -\tau_1^{(2)} - D(e_1^{(2)}, e_2^{(2)}) - \delta_0 \left(-\tau_1^{(0)} - \tau_2^{(0)} - D(e_1^{(0)}, e_2^{(0)}) \right) \\ + (1 - \delta_0) (\Delta D(\bar{e}_1; \bar{e}_2) + D(\bar{e}_1; \bar{e}_2)),$$

yielding

$$\frac{\partial W_2}{\partial e_2} = 0 \Leftrightarrow \frac{\partial B_1(e_2^{(2)})}{\partial e_2} = \frac{\partial D(e_1^{(2)}, e_2^{(2)})}{\partial e_2}. \quad (51)$$

Note that (49) to (51) imply that $e_1^{(0)} = e_1^{(1)}$; $e_2^{(0)} = e_2^{(1)}$; $e_2^{(0)} = e_2^{(2)}$; $e_1^{(0)} = e_1^{(2)}$. Hence, in equilibrium, $e_1^{(0)} = e_1^{(1)} = e_1^{(2)} = e_1^*$ and $e_2^{(0)} = e_2^{(1)} = e_2^{(2)} = e_2^*$. Thus, (e_1^*, e_2^*) is given by

$$\frac{\partial B_1(e_1^*)}{\partial e_1} = \frac{\partial D(e_1^*, e_2^*)}{\partial e_1} \quad \text{and} \quad \frac{\partial B_2(e_2^*)}{\partial e_2} = \frac{\partial D(e_1^*, e_2^*)}{\partial e_2}$$

For transfers proposed by player 0, and using the system given by (9) to (11) and (1) to (4), it is straightforward to obtain

$$\tau_1^* = \frac{r_0 r_2}{r_0 r_1 + r_0 r_2 + r_1 r_2} (D(\bar{e}_1, \bar{e}_2) - D(e_1^o, e_2^o) - (B_2(\bar{e}_2) - B_2(e_2^o))) \\ + \frac{r_0 r_1 + r_1 r_2}{r_0 r_1 + r_0 r_2 + r_1 r_2} (B_1(\bar{e}_1) - B_1(e_1^o)) \quad (52)$$

$$\tau_2^* = \frac{r_0 r_1}{r_0 r_1 + r_0 r_2 + r_1 r_2} (D(\bar{e}_1, \bar{e}_2) - D(e_1^o, e_2^o) - (B_1(\bar{e}_1) - B_1(e_1^o))) \\ + \frac{r_0 r_2 + r_1 r_2}{r_0 r_1 + r_0 r_2 + r_1 r_2} (B_2(\bar{e}_2) - B_2(e_2^o)) \quad (53)$$

A.2 Proposition 2

A.2.1 Stage 2

Using backward induction, we start in stage 2, when players 0 and 2 bargain, assuming that a previous agreement $o_1 = (\tau_1, e_1)$ has been agreed upon by players 0 and 1. Recall that we use the notation $\Gamma = t_1 \Delta$ to refer to the (endogenous) period at which stage 2 starts, with player 0 making the first offer. As player 0 proposes her preferred offer among those acceptable by player 2, we start by identifying offers that are acceptable for 2. As discussed in the main text, at time Γ player 2 will accept any offer from player 0, denoted $o_2^{(0)} = (\tau_2^{(0)}, e_2^{(0)})$, which yields at least the same stream of payoffs that she expect from waiting a period equal to Δ and proposing her equilibrium offer $o_2^{(2)}$, which will be accepted immediately by player 0. This yields the second equation in (15). Solving the integrals and

recalling that $\Pi_2(o_2^{(0)}) = \tau_2^{(0)} + B_2(e_2^{(0)})$, the acceptance condition for player 2 becomes

$$\tau_2^{(0)} \geq -B_2(e_2^{(0)}) + (1 - e^{-r_2\Delta}) B_2(\bar{e}_2) + e^{-r_2\Delta} \left(B_2(e_2^{(2)}) + \tau_2^{(2)} \right) = g_5(e_2^{(0)}, e_2^{(2)}, \tau_2^{(2)}, \bar{e}_2, r_2) \quad (54)$$

As before, for the functions g_i , and to simplify the exposition, we specify as arguments only the variables that are relevant to follow the proof, denoting all the other variables by (\cdot) , e.g. writing $g_5(e_2^{(0)}, \tau_2^{(2)}, \cdot)$ instead of $g_5(e_2^{(0)}, e_2^{(2)}, \tau_2^{(2)}, \bar{e}_2, r_2)$. When choosing her preferred offer among those acceptable by player 2, player 0 solves $\max_{o_2^{(0)}} \int_{\Gamma}^{\infty} U_0(o_1, o_2^{(0)}) e^{-r_2x} dx$ or, equivalently, $\max_{\tau_2^{(0)}, e_2^{(0)}} -\tau_1 - \tau_2^{(0)} - D(e_1, e_2^{(0)})$, subject to (54), where $o_1 = (\tau_1, e_1)$ is the offer agreed upon at stage 1 during the bargaining between 0 and 1.

As $U_0(o_1, o_2^{(0)})$ is strictly decreasing in $\tau_2^{(0)}$ the constraint (54) is satisfied with equality at the solution. In addition, by assumption D is convex, hence U_0 is concave, and B is concave. Hence, the solution to the maximization problem (denoted with a hat) $(\hat{\tau}_2^{(0)}, \hat{e}_2^{(0)})$ satisfies $\frac{\partial B_2(\hat{e}_2^{(0)})}{\partial e_2} = \frac{\partial D(e_1, \hat{e}_2^{(0)})}{\partial e_2}$ and $\hat{\tau}_2^{(0)} = g_5(e_2^{(0)}, \tau_2^{(2)}, \cdot)$.

Now we calculate the optimal offer o_2 for player 2 when the offer by player 0 is rejected. Using a similar argument as before, when allowed to make an offer player 2 proposes her preferred offer $o_2^{(2)}$ among those acceptable by player 0. We start then by identifying the acceptable offer by player 0. At time $\Gamma + \Delta$, player 0 will accept any offer from player 2, denoted $o_2^{(2)} = (\tau_2^{(2)}, e_2^{(2)})$, which yields at least the same stream of payoffs that she expect from waiting a period equal to Δ and proposing her equilibrium offer $o_2^{(0)} = (\tau_2^{(0)}, e_2^{(0)})$, which will be accepted immediately by player 2. This yields the first equation in (15). Again solving the integrals and recalling that $U_0(o_1, o_2^{(2)}) = -\tau_1 - \tau_2^{(2)} - D(e_1, e_2^{(2)})$, the acceptance condition for player 0 can then be rewritten as

$$\tau_2^{(2)} \leq D(e_1, \bar{e}_2) - D(e_1, e_2^{(2)}) - e^{-r_0\Delta} \left(D(e_1, \bar{e}_2) - D(e_1, e_2^{(0)}) \right) + e^{-r_0\Delta} \tau_2^{(0)} = g_6(e_2^{(2)}, \tau_2^{(0)}, \cdot) \quad (55)$$

Thus, player 2 solves $\max_{o_2^{(2)}} \Pi_2(o_2^{(2)})$ subject to (55), where $o_1 = (\tau_1, e_1)$ is the offer agreed upon at stage 1. Again, as $\Pi_2(o_2^{(2)})$ is strictly increasing in $\tau_2^{(2)}$ the constraint (55) is satisfied with equality at the solution. Given our assumptions, the solution to the maximization problem $(\hat{\tau}_2^{(2)}, \hat{e}_2^{(2)})$ satisfies $\frac{\partial B_2(\hat{e}_2^{(2)})}{\partial e_2} = \frac{\partial D(e_1, \hat{e}_2^{(2)})}{\partial e_2}$ and $\hat{\tau}_2^{(2)} = g_6(e_2^{(2)}, \tau_2^{(0)}, \cdot)$. Note that the first condition defining $\hat{e}_2^{(2)}$ is the same than the one that we found for $\hat{e}_2^{(0)}$ before. This means that, in equilibrium, $\hat{e}_2^{(0)} = \hat{e}_2^{(2)} = e_2^s$, for a given e_1 . Defining $f(e_2) = \frac{\partial B_2(e_2)}{\partial e_2} - \frac{\partial D(e_1, e_2)}{\partial e_2}$, this discussion implies that as long as $f(e_2) = 0$ exists and defines a unique value of e_2 , we have that $\hat{e}_2^{(0)} = \hat{e}_2^{(2)} = e_2^s$ exists and is unique. Using Bolzano's Theorem, we know that $f(e_2) = 0$ indeed exists and defines a unique value of e_2 because:

- (i) For e_1 a real constant and $e_2 \in [0, \bar{e}_2)$, we know that $f(e_2)$ is continuous in e_2 (because

it is differentiable) and monotone decreasing (recall that $\frac{\partial^2 B_2(e_2)}{\partial e_2^2} < 0$ and $\frac{\partial^2 D(e_1, e_2)}{\partial e_2^2} > 0$ imply that $\frac{\partial f(e_2)}{\partial e_2} = \frac{\partial^2 B_2(e_2)}{\partial e_2^2} - \frac{\partial^2 D(e_1, e_2)}{\partial e_2^2} < 0$)

(ii) Given the assumption $\frac{\partial B_2(0)}{\partial e_2} - \frac{\partial D(e_1, 0)}{\partial e_2} > 0$ for all e_1 , we have $f(0) > 0$ and $f(z) = \frac{\partial B_2(z)}{\partial e_2} - \frac{\partial D(e_1, z)}{\partial e_2} < 0$ for some $0 < z < \bar{e}_2$ (recall that $\frac{\partial B_2(\bar{e}_2)}{\partial e_2} = 0$)

To prove that $\tau_2^s = \tau_2^{(0)}$ exists and is unique we show that it is a fixed point of a non-negative valued affine function with a slope smaller than 1. Substitute $\hat{\tau}_2^{(2)} = g_6(e_2^{(2)}, \tau_2^{(0)}, \cdot) = g_6(e_2^s, \tau_2^s, \cdot)$ into $\tau_2^s = \tau_2^{(0)} = g_5(e_2^{(0)}, \tau_2^{(2)}, \cdot) = g_5(e_2^s, \hat{\tau}_2^{(2)}, \cdot)$ to obtain

$$\begin{aligned} \tau_2^s &= \tau_2^{(0)} = g_5(e_2^s, g_6(e_2^s, \tau_2^s, \cdot), \cdot) = e^{-(r_2+r_0)\Delta} \tau_2^s + e^{-r_2\Delta} (1 - e^{-r_0\Delta}) (D(e_1, \bar{e}_2) - D(e_1, e_2^s)) \\ &\quad + (1 - e^{-r_2\Delta}) (B_2(\bar{e}_2) - B_2(e_2^s)) \end{aligned}$$

Define $F(x)$ as the RHS of this expression, with $x = \tau_2^s$, i.e.:

$$\begin{aligned} F(x) &= e^{-(r_2+r_0)\Delta} x + e^{-r_2\Delta} (1 - e^{-r_0\Delta}) (D(e_1, \bar{e}_2) - D(e_1, e_2^s)) \\ &\quad + (1 - e^{-r_2\Delta}) (B_2(\bar{e}_2) - B_2(e_2^s)) \end{aligned}$$

Note that, given real constants e_1 and e_2^s , the function $F(x)$ is an affine function of x , and hence continuous. Furthermore, as $F(x)$ is an affine increasing function with $0 < \frac{\partial F(x)}{\partial x} = e^{-(r_2+r_0)\Delta} < 1$ and $F(0) > 0$, it can only intersect the 45 degree line once, at $x = \tau_2^s$. Hence τ_2^s is unique, given e_1 and e_2^s , and is equal to

$$\tau_2^s = \frac{e^{-r_2\Delta} (1 - e^{-r_0\Delta})}{1 - e^{-(r_2+r_0)\Delta}} (D(e_1, \bar{e}_2) - D(e_1, e_2^s)) + \frac{1 - e^{-r_2\Delta}}{1 - e^{-(r_2+r_0)\Delta}} (B_2(\bar{e}_2) - B_2(e_2^s)) \quad (56)$$

A.2.2 Stage 1

At Stage 1, players 0 and 1 bargain over $o_1 = (\tau_1, e_1)$ and anticipate that if e_1 is agreed upon then players 0 and 2 will agree upon (τ_2^s, e_2^s) as shown above. We write $\tau_2^s(e_1)$ and $e_2^s(e_1)$ as they depend on e_1 .

As before, player 0 starts making an offer at period 0 and proposes her preferred offer among those acceptable by player 1. Player 1 will accept any offer from player 0, denoted $o_1^{(0)} = (\tau_1^{(0)}, e_1^{(0)})$, which yields at least the same stream of payoffs that she expects from waiting a period equal to Δ and proposing her equilibrium offer $o_1^{(1)}$, which will be accepted immediately by player 0. This yields the condition in (17). Solving the integral and recalling $\Pi_1(o_1^{(0)}) = \tau_1^{(0)} + B_1(e_1^{(0)})$, the acceptance condition for player 1 becomes

$$\tau_1^{(0)} \geq -B_1(e_1^{(0)}) + (1 - e^{-r_1\Delta}) B_1(\bar{e}_1) + e^{-r_1\Delta} (B_1(e_1^{(1)}) + \tau_1^{(1)}) = g_7(e_1^{(0)}, \tau_1^{(1)}, \cdot) \quad (57)$$

Recalling that Γ denotes the (endogenous) time at which stage 2 starts, the optimization

problem for player 0 at time 0, is given by the expression

$$\max_{o_1^{(0)}} \int_0^\Gamma U_0(o_1^{(0)}, \bar{o}_2) e^{-r_0 x} dx + \int_\Gamma^\infty U_0(o_1^{(0)}, o_2^s(e_1^{(0)})) e^{-r_0 x} dx$$

subject to (57). Again, solving the integral, recalling $U_0(o_1, o_2^s(e_1)) = -\tau_1 - \tau_2^s(e_1) - D(e_1, e_2^s(e_1))$, and noting that the objective function is strictly decreasing in $\tau_1^{(0)}$, and hence that the constraint is satisfied with equality at the solution, we can find the optimal $e_1^{(0)}$ by solving

$$\max_{e_1^{(0)}} -g_7(e_1^{(0)}, \tau_1^{(1)}, \cdot) - (1 - e^{-r_0 \Gamma}) D(e_1^{(0)}, \bar{e}_2) - e^{-r_0 \Gamma} D(e_1^{(0)}, e_2^s(e_1^{(0)})) - e^{-r_0 \Gamma} \tau_2^s(e_1^{(0)})$$

which gives

$$\frac{\partial B_1(\hat{e}_1^{(0)})}{\partial e_1} = (1 - e^{-r_0 \Gamma}) \frac{\partial D(\hat{e}_1^{(0)}, \bar{e}_2)}{\partial e_1} + e^{-r_0 \Gamma} \left(1 + \frac{\partial e_2^s(\hat{e}_1^{(0)})}{\partial e_1} \right) \frac{\partial D(\hat{e}_1^{(0)}, e_2^s(\hat{e}_1^{(0)}))}{\partial e_1} + e^{-r_0 \Gamma} \frac{\partial \tau_2^s(\hat{e}_1^{(0)})}{\partial e_1} \quad (58)$$

and $\hat{\tau}_1^{(0)} = g_7(e_1^{(0)}, \tau_1^{(1)}, \cdot)$.

We find now the optimal offer o_1 for player 1 when the offer by player 0 is rejected. Using a similar argument as before, the offer made by player 1, denoted $o_1^{(1)} = (\tau_1^{(1)}, e_1^{(1)})$, has to provide player 0 at least the same stream of payoffs that she expects from waiting a period equal to Δ and proposing her equilibrium offer $o_1^{(1)}$, which will be accepted immediately by player 0, denoted $o_1^{(0)} = (\tau_1^{(0)}, e_1^{(0)})$. This yields the indifference condition (16).

Solving again the integrals and recalling that $U_0(o_1, o_2^s(e_1)) = -\tau_1 - \tau_2^s(e_1) - D(e_1, e_2^s(e_1))$, the acceptance condition for player 0 can then be rewritten as

$$\begin{aligned} \tau_1^{(1)} \leq & -(1 - e^{-r_0 \Gamma}) D(e_1^{(1)}, \bar{e}_2) - e^{-r_0 \Gamma} D(e_1^{(1)}, e_2^s(e_1^{(1)})) + (1 - e^{-r_0 \Delta}) D(\bar{e}_1, \bar{e}_2) \quad (59) \\ & + e^{-r_0 \Delta} (1 - e^{-r_0 \Gamma}) D(e_1^{(0)}, \bar{e}_2) + e^{-r_0(\Delta + \Gamma)} D(e_1^{(0)}, e_2^s(e_1^{(0)})) \\ & + e^{-r_0(\Delta + \Gamma)} \tau_2^s(e_1^{(0)}) - e^{-r_0 \Gamma} \tau_2^s(e_1^{(1)}) + e^{-r_0 \Delta} \tau_1^{(0)} = g_8(e_1^{(1)}, \tau_1^{(0)}, \cdot) \end{aligned}$$

When choosing her most preferred offer among those acceptable by player 0, player 1 solves $\max_{o_1^{(1)} = (\tau_1^{(1)}, e_1^{(1)})} \Pi_1(o_1^{(1)})$ subject to (59). As the objective function is strictly decreasing in $\tau_1^{(1)}$ the constraint is satisfied with equality at the solution and we can find the optimal $e_1^{(1)}$ by solving $\max_{e_1^{(1)}} B_1(e_1^{(1)}) + g_8(e_1^{(1)}, \tau_1^{(0)}, \cdot)$. Taking into account our assumptions, the solution to the maximization problem (e_1^s, τ_1^s) yields:

$$\frac{\partial B_1(\hat{e}_1^{(1)})}{\partial e_1} = (1 - e^{-r_0 \Gamma}) \frac{\partial D(\hat{e}_1^{(1)}, \bar{e}_2)}{\partial e_1} + e^{-r_0 \Gamma} \left(1 + \frac{\partial e_2^s(\hat{e}_1^{(1)})}{\partial e_1} \right) \frac{\partial D(\hat{e}_1^{(1)}, e_2^s(\hat{e}_1^{(1)}))}{\partial e_1} + e^{-r_0 \Gamma} \frac{\partial \tau_2^s(\hat{e}_1^{(1)})}{\partial e_1}$$

and $\widehat{\tau}_1^{(1)} = g_8(e_1^{(1)}, \tau_1^{(0)}, \cdot)$.

Note that the first condition defining $\widehat{e}_1^{(1)}$ is the same than the one that we found for $\widehat{e}_1^{(0)}$ in (58). This means that, in equilibrium, $\widehat{e}_1^{(0)} = \widehat{e}_1^{(1)} = e_1^s$. Using the expression of τ_2^s obtained in stage 2, i.e. (56), we can calculate $\frac{\partial \tau_2^s}{\partial e_1}(e_1^s)$. Substituting the expression obtained into (58), for $e_1^{(0)} = e_1^s$, and recalling that $\frac{\partial D(e_1^s, e_2^s)}{\partial e_2} = \frac{\partial B_2(e_2^s)}{\partial e_2}$ from stage 2, gives

$$\begin{aligned} & \frac{\partial B_1(e_1^s)}{\partial e_1} - \left((1 - e^{-r_0\Gamma}) + \frac{e^{-r_0\Gamma} e^{-r_2\Delta} (1 - e^{-r_0\Delta})}{1 - e^{-(r_2+r_0)\Delta}} \right) \frac{\partial D(e_1^s, \bar{e}_2)}{\partial e_1} \\ & - e^{-r_0\Gamma} \left(\frac{1 - e^{-r_2\Delta}}{1 - e^{-(r_2+r_0)\Delta}} \right) \frac{\partial D(e_1^s, e_2^s)}{\partial e_1} \\ & = 0 \end{aligned} \quad (60)$$

Let us call the LHS of (60) $h(e_1)$. We know that e_1^s exists and is unique (again using Bolzano's Theorem) because:

(i) The function $h(e_1)$ for $e_1 \in [0, \bar{e}_1)$ is continuous and monotone decreasing, given that $\frac{\partial^2 B_1(e_1)}{\partial e_1^2} < 0$ and $\frac{\partial^2 D(e_1, e_2)}{\partial e_1^2} > 0$ on their domains, with $-1 < \frac{\partial e_2^s}{\partial e_1} = \frac{\frac{\partial^2 D(e_1, e_2^s(e_1))}{\partial e_1^2}}{\frac{\partial^2 B_1(e_1^s(e_1))}{\partial e_1^2} - \frac{\partial^2 D(e_1, e_2^s(e_1))}{\partial e_1^2}} < 0$.

All this implies that $\frac{\partial h(e_1)}{\partial e_1} < 0$.

(ii) Our assumptions ensure that $h(0) > 0$ and

$$h(z) = \frac{\partial B_1(z)}{\partial e_1} - \left(1 - e^{-r_0\Gamma} + \frac{e^{-r_0\Gamma} e^{-r_2\Delta} (1 - e^{-r_0\Delta})}{1 - e^{-(r_2+r_0)\Delta}} \right) \frac{\partial D(z, \bar{e}_2)}{\partial e_1} - \frac{e^{-r_0\Gamma} (1 - e^{-r_2\Delta})}{1 - e^{-(r_2+r_0)\Delta}} \frac{\partial D(z, e_2^s(z))}{\partial e_1} < 0$$

for some $0 < z < \bar{e}_1$.

To see this, note that, as $\frac{\partial D(0, e_2^s(0))}{\partial e_1} = 0$, we have

$$h(0) = \frac{\partial B_1(0)}{\partial e_1} - \left((1 - e^{-r_0\Gamma}) + \frac{e^{-r_0\Gamma} e^{-r_2\Delta} (1 - e^{-r_0\Delta})}{1 - e^{-(r_2+r_0)\Delta}} \right) \frac{\partial D(0, \bar{e}_2)}{\partial e_1}.$$

As the term in brackets is smaller than 1, the assumption $\frac{\partial B_1(0)}{\partial e_1} - \frac{\partial D(0, \bar{e}_2)}{\partial e_1} > 0$ discussed in the main text suffices to ensure that $h(0) > 0$. Moreover, since $\frac{\partial B_1(\bar{e}_1)}{\partial e_1} = 0$ we always have $h(z) < 0$ for the given interval on e_1 .

Finally, as before, we know that τ_1^s exists and is unique because it is a fixed point of a non-negative valued affine function with a slope smaller than 1. Replacing the expression for $\widehat{\tau}_1^{(1)} = g_8(e_1^{(1)}, \tau_1^{(0)}, \cdot)$ with $\tau_1^{(0)} = \tau_1^s$ into $\tau_1^s = \widehat{\tau}_1^{(0)} = g_7(e_1^{(0)}, \cdot)$ with $\tau_1^{(1)} = \widehat{\tau}_1^{(1)}$, writing it

for $e_1^{(0)} = e_1^{(1)} = e_1^s$ and using the expression of τ_2^s in (56) we obtain

$$\begin{aligned}\tau_1^s &= e^{-(r_0+r_1)\Delta}\tau_1^s + (1 - e^{-r_1\Delta}) (B_1(\bar{e}_1) - B_1(e_1^s)) \\ &\quad + (1 - e^{-r_0\Delta}) e_1^{-r_1\Delta} (D(\bar{e}_1, \bar{e}_2) - D(e_1^s, \bar{e}_2)) \\ &\quad + \frac{e^{-r_0\Gamma} e_1^{-r_1\Delta} (1 - e^{-r_0\Delta}) (1 - e^{-r_2\Delta})}{1 - e^{-(r_2+r_0)\Delta}} \Psi(e_1^s, e_2^s; e_1^s, \bar{e}_2)\end{aligned}\tag{61}$$

Proceeding as for stage 2, define $H(x)$ as the RHS of (61). Note that, given real constants e_1^s and $e_2^s(e_1^s)$, the function $H(x)$ is an affine function of $\tau_1 \in \mathfrak{R}_+$, with $H(0) > 0$ (since $\Psi(e_1^s, e_2^s; e_1^s, \bar{e}_2) > 0$) and $0 < \frac{\partial H(x)}{\partial x} < 1$. Hence it will intersect the 45 degree line in exactly one value of τ_1 , namely $x = \tau_1^s$. Thus, τ_1^s is given by

$$\begin{aligned}\tau_1^s &= \frac{(1 - e^{-r_1\Delta})}{1 - e^{-(r_0+r_1)\Delta}} (B_1(\bar{e}_1) - B_1(e_1^s)) + \frac{(1 - e^{-r_0\Delta}) e_1^{-r_1\Delta}}{1 - e^{-(r_0+r_1)\Delta}} (D(\bar{e}_1, \bar{e}_2) - D(e_1^s, \bar{e}_2)) \\ &\quad + \frac{e^{-r_0\Gamma} (1 - e^{-r_0\Delta}) e_1^{-r_1\Delta}}{1 - e^{-(r_0+r_1)\Delta}} (D(e_1^s, \bar{e}_2) - D(e_1^s, e_2^s)) - \frac{e^{r_0\Gamma} (1 - e^{-r_0\Delta}) e_1^{-r_1\Delta}}{1 - e^{-(r_0+r_1)\Delta}} \tau_2^s\end{aligned}\tag{62}$$

Finally, recalling $\Gamma = t_1^s \Delta$, for any small $\Delta > 0$ we have that $e^{-r_i\Delta} = (1 - r_i\Delta)$ and $e^{-r_0\Gamma} = (1 - r_0\Gamma)$. Thus, taking the limit $\Delta \rightarrow 0$, the expression in (60) simplifies to (18). Also, the expression in (56) becomes

$$\tau_2^s = \frac{r_2}{r_0 + r_2} (B_2(\bar{e}_2) - B_2(e_2^s)) + \frac{r_0}{r_0 + r_2} (D(e_1^s, \bar{e}_2) - D(e_1^s, e_2^s))\tag{63}$$

and the expression in (62) becomes

$$\begin{aligned}\tau_1^s &= \frac{r_1}{r_0 + r_1} (B_1(\bar{e}_1) - B_1(e_1^s)) + \frac{r_0}{r_0 + r_1} (D(\bar{e}_1, \bar{e}_2) - D(e_1^s, \bar{e}_2)) \\ &\quad + \frac{r_0}{r_0 + r_1} (D(e_1^s, \bar{e}_2) - D(e_1^s, e_2^s)) - \frac{r_0}{r_0 + r_1} \tau_2^s\end{aligned}\tag{64}$$

Simple manipulations yield the net instantaneous equilibrium payoffs in Proposition 2.

A.3 Corollary 2

In stage 2, the Nash program between 0 and 2 can be rewritten as

$$\begin{aligned}\max_{e_2, \tau_2} \Omega_{02} &= \left(\int_0^\infty [-\tau_2 + D(e_1, \bar{e}_2) - D(e_1, e_2)] e^{-r_0 x} dx \right)^{\frac{r_2}{r_0+r_2}} \\ &\quad \left(\int_0^\infty [B_2(e_2) + \tau_2 - B_2(\bar{e}_2)] e^{-r_2 x} dx \right)^{\frac{r_0}{r_0+r_2}}\end{aligned}$$

or equivalently

$$\max_{e_2, \tau_2} \Omega_{02} = (-\tau_2 + D(e_1, \bar{e}_2) - D(e_1, e_2))^{\frac{r_2}{r_0+r_2}} (B_2(e_2) + \tau_2 - B_2(\bar{e}_2))^{\frac{r_0}{r_0+r_2}} \quad (65)$$

The derivative of Ω_{02} with respect to τ_2 gives

$$\tau_2 = \frac{r_2}{r_0+r_2} (B_2(\bar{e}_2) - B_2(e_2)) \frac{r_0}{r_0+r_2} (D(e_1, \bar{e}_2) - D(e_1, e_2)) \quad (66)$$

Substituting τ_2 in (65) and taking the derivative of Ω_{02} with respect to e_2 gives

$$\frac{\partial B_2(e_2)}{\partial e_2} = \frac{\partial D(e_1, e_2)}{\partial e_2} \quad (67)$$

which is the same equation as (19) implying that e_2 is a function of e_1 as $e_2(e_1)$.

The program for 0 in stage 1 is

$$\max_{e_1, \tau_1} \Omega_{01} = (-\tau_1 - \tau_2 + D(\bar{e}_1, \bar{e}_2) - D(e_1, e_2))^{\frac{r_1}{r_0+r_1}} (B_1(e_1) + \tau_1 - B_1(\bar{e}_1))^{\frac{r_0}{r_0+r_1}} \quad (68)$$

The derivative of Ω_{01} with respect to τ_1 gives

$$\tau_1 = \frac{r_1}{r_0+r_1} (B_1(\bar{e}_1) - B_1(e_1)) + \frac{r_0}{r_0+r_1} (-\tau_2 + D(\bar{e}_1, \bar{e}_2) - D(e_1, e_2))$$

Substituting in (68) and taking the derivative with respect to e_1 yields

$$-\frac{\partial \tau_2}{\partial e_1} + \frac{\partial B_1(e_1)}{\partial e_1} = \left(1 + \frac{\partial e_2}{\partial e_1}\right) \frac{\partial D(e_1, e_2)}{\partial e_1} \quad (69)$$

From (66), we have

$$\frac{\partial \tau_2}{\partial e_1} = -\frac{r_2}{r_0+r_2} \frac{\partial e_2}{\partial e_1} \frac{\partial B_2(e_2)}{\partial e_2} + \frac{r_0}{r_0+r_2} \frac{\partial D(e_2, \bar{e}_2)}{\partial e_1} - \frac{r_0}{r_0+r_2} \left(1 + \frac{\partial e_2}{\partial e_1}\right) \frac{\partial D(e_1, e_2)}{\partial e_1}$$

Substituting in (69) and using the optimality condition (67) gives the same equation as (18)

$$\frac{\partial B_1(e_1)}{\partial e_1} = \frac{\partial D(e_1, e_2)}{\partial e_1} + \frac{r_0}{r_0+r_2} \left(\frac{\partial D(e_1, \bar{e}_2)}{\partial e_1} - \frac{\partial D(e_1, e_2)}{\partial e_1} \right) \quad (70)$$

Combining (67) and (70) gives the equilibrium emissions (e_1^s, e_2^s) .

Transfers are

$$\tau_2^s = \frac{r_2}{r_0+r_2} (B_2(\bar{e}_2) - B_2(e_2^s)) + \frac{r_0}{r_0+r_2} (D(e_1^s, \bar{e}_2) - D(e_1^s, e_2^s))$$

and

$$\begin{aligned}\tau_1^s &= \frac{r_1}{r_0 + r_1} (B_1(\bar{e}_1) - B_1(e_1^s)) + \frac{r_0}{r_0 + r_1} (D(\bar{e}_1, \bar{e}_2) - D(e_1^s, e_2^s)) \\ &\quad - \frac{r_0}{r_0 + r_1} \frac{r_2}{r_0 + r_2} (B_2(\bar{e}_2) - B_2(e_2^s)) - \frac{r_0}{r_0 + r_1} \frac{r_0}{r_0 + r_2} (D(e_1^s, \bar{e}_2) - D(e_1^s, e_2^s)).\end{aligned}$$

This completes the proof.

Emissions	Disagreement (threat)	Multilateral	Sequential, first 1 (0↔1;0↔2)	Sequential, first 2 (0↔2;0↔1)
<i>Example 1</i> ¹				
e_1	10.00	3.33	1.42	4.28
e_2	10.00	3.33	4.28	1.42
$e_1 + e_2$	20.00	6.66	5.70	5.70
<i>Example 2</i>				
e_1	10.00	3.33	0.77	3.53
e_2	10.00	3.33	4.60	2.94
$e_1 + e_2$	20.00	6.66	5.37	6.47
<i>Example 3</i>				
e_1	10.00	2.72	0.17	2.82
e_2	16.66	4.54	6.14	4.35
$e_1 + e_2$	26.60	7.26	6.31	7.17

¹Parameter values are: $r_i = 1\%$, $c_i = 10$, and $w_i = d = 1$ for $i = 0, 1, 2$ for example 1; $r_0 = 1\%$, $r_1 = 10\%$, $r_2 = 0.4\%$, $c_1 = c_2 = 10$ and $w_1 = w_2 = d = 1$ for example 2; and $r_0 = 0.5\%$, $r_1 = 17\%$, $r_2 = 1\%$, $c_1 = c_2 = 10$, $w_1 = 1$, $w_2 = 0.6$ and $d = 1$ for example 3.

Table 1. Equilibrium emissions for the Examples

Instantaneous payoffs	Multilateral	Sequential, first 1 ($0 \leftrightarrow 1; 0 \leftrightarrow 2$)	Sequential, first 2 ($0 \leftrightarrow 2; 0 \leftrightarrow 1$)
<i>Example 1</i> ¹			
U_0	44.44	57.14	57.14
Π_1	44.44	57.14	16.34
Π_2	44.44	16.34	57.14
$U_0 + \Pi_1 + \Pi_2$	133.32	130.62	130.62
<i>Example 2</i>			
U_0	37.03	97.9	36.97
Π_1	3.70	9.79	3.80
Π_2	92.59	20.71	92.43
$U_0 + \Pi_1 + \Pi_2$	133.32	128.40	133.20
<i>Example 3</i>			
U_0	169.07	218.14	171.39
Π_1	4.97	6.41	1.472
Π_2	84.53	29.54	85.69
$U_0 + \Pi_1 + \Pi_2$	258.57	254.09	258.55

¹Parameter values are: $r_i = 1\%$, $c_i = 10$, and $w_i = d = 1$ for $i = 0, 1, 2$ for example 1; $r_0 = 1\%$, $r_1 = 10\%$, $r_2 = 0.4\%$, $c_1 = c_2 = 10$ and $w_1 = w_2 = d = 1$ for example 2; and $r_0 = 0.5\%$, $r_1 = 17\%$, $r_2 = 1\%$, $c_1 = c_2 = 10$, $w_1 = 1$, $w_2 = 0.6$ and $d = 1$ for example 3.

Table 2. Equilibrium instantaneous payoffs for the Examples