Benevolent Arbitration in the Shadow of Conflict

Andrea Canidio
Joan Esteban

This version: April 2020
(March 2018)

Barcelona GSE Working Paper Series

Working Paper nº 1027

Benevolent arbitration in the shadow of conflict*

Andrea Canidio[†] *and Joan Esteban*[‡]

Abstract

We propose a novel arbitration procedure. This procedure allows a completely uninformed arbitrator to efficiently settle a dispute between two fully informed players. The novelty is that we allow the players to make wasteful, pre-negotiation investments aimed at manipulating the outcome of the arbitration. Our arbitration procedure is such that these investments are minimized. The core of the arbitration procedure is a "concession game" in which the third party shares the available peace dividend as a function of the concessions made by each player. The concession game can be used to eliminate offensive investments, that is, investments that make the conflict occurring in case no agreement is reached more destructive.

JEL classification: F51; J52.

Keywords: Bargaining; arbitration; concessions; conflict.

*We are grateful to Attila Ambrus, Sandeep Baliga, Mohamed Belhaj, Yann Bramoullé, Georgy Egorov, Thomas Gall, Itzhak Gilboa, Sidartha Gordon, Patrick Legros, Dilip Mookherjee, Massimo Morelli, Andy Newman, David Pearce, Debraj Ray, and Ariel Rubinstein, seminar and conference participants at AMSE, Paris Dauphine, IMT Lucca, ThReD 2017, VIII IBEO workshop for their comments and suggestions. Joan Esteban gratefully acknowledges financial support from the AXA Research Fund, the Generalitat de Catalunya AGAUR grant 2017SGR1359, the Spanish Ministry of Economy and Competitiveness grant ECO2015-66883-P, and the National Science Foundation grant SES-1629370. This version supersedes an earlier version titled "Benevolent Mediation in the Shadow of Conflict".

[†]Corresponding Author. IMT School for Advanced Studies, Piazza S. Francesco, Lucca, Italy; INSEAD, Boulevard de Constance, Fontainebleau, France; andrea.canidio@imt.it.

[‡]Institut d'Anàlisi Econòmica and Barcelona GSE, Campus UAB, 08193 Bellaterra, Barcelona, Spain; joan.esteban@iae.csic.es.

1 Introduction

Third parties often play a key role in resolving disputes and preventing conflicts. For this reason, a large literature has studied how mediators and arbitrators can influence the players' behavior within the negotiation and increase the probability of reaching an agreement. An often overlooked implication is that, anticipating the actions of this third-party, the players may make costly, pre-negotiation investments aimed at influencing the negotiating outcome. These investments are often a form of rent seeking, because they do not increase the total payoff to be shared during the negotiation, but only how this surplus is split. They are, however, quite common.

For instance, negotiations are often conducted under the shadow of conflict: in case an agreement is not reached, the negotiating parties will fight in a non-cooperative game. Because the outcome of the conflict defines the disagreement point of the negotiation, the bargaining parties may spend resources to prepare for conflict even if they expect to achieve a negotiated agreement. This is why, for example, a government and a rebel group may engage in military actions just before negotiating a peace agreement.[1] Similarly, before negotiating a settlement, two firms may ask for additional legal opinions or hire very expensive lawyers as a way to manipulate the outcome of the lawsuit that may follow the breakdown of the negotiation.[2] Crucially, the way in which the dispute is expected to be resolved will determine how the negotiated outcome changes with the conflict payoffs, and hence the total resources wasted in pre-negotiation investments.

In this paper we study the problem of a benevolent arbitrator who wishes to efficiently share the "peace dividend" (i.e, the aggregate benefit of finding an agreement rather than triggering a conflict) in a way that minimizes wasteful pre-negotiation investments. Our main assumption is that, whereas the negotiating parties are fully informed, the arbitrator is completely uninformed with respect to the details of the dispute, in the sense that he does not observe neither the size of the payoff to be shared by the two parties, nor the players' outside options, nor the players' pre-negotiation investments. The arbitrator is therefore an outsider knowing less about

[1] There is ample evidence that conflicts are reactivated prior to the beginning of peace negotiations. For example, the mass killing of civilians (thus permanently weakening the opponents) is significantly more probable during the process of democratization of a country. See **?**.

[2] For a review of the academic literature arguing that the bargaining parties may spend resources to prepare for conflict before the start of the negotiation see **?** and the literature review in **?** (forthcoming).

the dispute than the players. For this reason, the arbitrator will resolve the dispute via an arbitration procedure (or arbitration game).[3]

We consider two types of pre-negotiation wasteful investments: *offensive* and *defensive*.[4] Offensive investments by a player decrease the opponent's payoff in case of conflict. As examples of offensive investment, a player may purchase ballistic missiles or collect evidence against the opponent to be used in a court case. Instead, defensive investments by a player increase this player's payoff in case of conflict. As examples of defensive investments, a player may purchase antimissile system and bunkers, or move assets to jurisdictions where they are harder to seize in case the outcome of a lawsuit is negative. In the case of industrial conflict, a firm may invest resources to make the relocation of the factory a credible threat, which should be considered as offensive.

Our main result is that, despite his informational disadvantage, the arbitrator can construct an arbitration procedure that fully eliminates offensive investments, minimizes defensive investments, and achieves an efficient split of the peace dividend while always satisfying the players' outside options. The first piece of the mechanism is what we call a *random permanent proposer* bargaining protocol, which allows the arbitrator to implement a generalized Nash bargaining solution. In a random permanent proposer bargaining protocol, the arbitrator randomly chooses a player who, in every subsequent period, proposes a settlement to the other player. The player receiving the proposal can accept it—in which case the game ends with an agreement—or reject it and continue the negotiation—in which case the game continues to the next period in which, again, the permanent proposer makes an offer—or reject it and trigger a conflict. Unsurprisingly, in equilibrium the permanent proposer will propose to keep the opponent at his outside option, and the opponent will accept immediately.[5] Hence, the random permanent proposer allows the arbitrator to efficiently split the peace dividend among the two players—where the

[3] The most famous example of an arbitration game is final offer arbitration, in which the parties submit a proposed settlement and the arbitrator chooses a solution among the submitted proposal. Other arbitration procedures have been proposed in the literature—see the next section.

[4] This is similar to **?**, who distinguish between investment in predatory activities and investment in defense.

[5] This extreme outcome is a consequence of the fact that the arbitrator is completely uninformed—which is usually not the case in practice. Nonetheless, it is interesting to note that "extreme" arbitration outcomes are quite common. See the empirical evidence discussed in **?**, in particular pages 724-725.

split is given by the probability of being the permanent proposer—without knowing neither the size of the peace dividend, nor the players' outside options.[6]

The second (and most important) piece of the game is what we call a *concession game*. The arbitrator announces that the probability that each player is the permanent proposer is a function of the players' concessions: costly actions that are visible to the arbitrator and can benefit the other player. Each player, therefore, may want to make a concession as a way to obtain a better treatment from the arbitrator. The concession game is build so that, in equilibrium, no concessions are made and the probability that a player is the permanent proposer is set at some default level. However, even a tiny increase in the peace dividend will trigger a costly "concession contest" between the players. By doing so, the arbitrator achieves two goals: he can elicit truthful revelation (because the arbitrator can now punish the players by triggering a costly a "concession contest"); and he can prevent offensive investments (which, because they make the conflict more costly, also increase the peace dividend).[7] Finally, the default probability that a player is the permanent proposer is chosen so that defensive investments are minimized. We show that achieving this objective may be incompatible with fairness: efficiency requires splitting the peace dividend unevenly. More precisely, in our preferred functional form specification, the waste minimizing sharing rule gives a larger share to the strongest player, where "weak" and "strong" are defined by the outcome of the potential conflict in absence of investments.

Literature Our paper belongs to the literature studying hold-up problems, that is, how to achieve efficiency (or reduce inefficiencies) when investments are not contractible. The vast majority of papers in this literature, however, assume that the

[6] Note that other mechanisms that achieve this goal exist in the literature (see, for example, **?** and **?**). These mechanisms could replace the random permanent proposer bargaining game, while leaving our results unchanged. They are, however, more convoluted than the random permanent proposer, which is why we choose such mechanism.

[7] There are similarities between our concession contest and money burning as in **?**. Using a forward induction argument, **?** show that the possibility of burning money allows a player to signal her intention to take an action and hence select her preferred outcome. Interestingly, this works also if in equilibrium no money is burned. Note, however, that whereas money burning occurs before the players take their actions, the concession contest occurs after the investment stage. The concession contest therefore affects the players' actions via a more standard backward induction logic.

outside options of the ex-post negotiation can be decided contractually ex-ante in order to induce the efficient level of investments.[8] Here instead we are interested in situations in which the outside options are determined by the players' investments and therefore cannot be specified contractually ex-ante. These situations have received considerably less attention.[9]

Whereas many papers have noted that the negotiating parties may want to invest before the start of the negotiation, existing economic models of third party intervention assume that the mediator/arbitrator sole role is to maximize surplus *within the negotiation*. A particularly relevant example is **?**, who compare mediated and unmediated negotiation and argue that mediated negotiation generates lower pre-negotiation wasteful investment in arms. In their framework, due to informational asymmetries, inefficient negotiation breakdowns may occur. The mediator's role is to regulate the flow of information among parties. His goal is to maximize the probability that an efficient settlement is reached. The authors show that, by doing so, he also unintentionally reduces the players' pre-negotiation investments. Hence, our model differs from **?** mainly because the arbitrator's goal is to achieve efficiency, which explicitly includes reducing wasteful pre-negotiation investments.

A number of other authors also noted that the way the negotiation is conducted can affect pre-negotiation actions by the players. Both **?** and **?** show that the surplus share accruing to each player can have an effect on decisions made prior to the beginning of the negotiation. In **?** the surplus share obtained by each party in a negotiation may affect the intensity of the pre-negotiation conflict. They show that an equal surplus-split rule may be welfare decreasing relative to an asymmetric surplus-split rule. **?** notice that by investing in arms players influence the probability of winning in a conflict—and hence the disagreement point—and the share of surplus in the case of a peaceful agreement. Their main result is that when fighting is not

[8] For example, in the seminal work of **?** the allocation of ownership indirectly determines the players' outside options in the ex-post negotiation. In **?** the players can directly specify ex-ante the outside options of the ex-post negotiation (together with the surplus share accruing to each player in the ex-post negotiation).

[9] There are papers in which the outside options of the ex-post negotiation depend both on the ex-ante allocation of ownership and on subsequent, non contractible investments (see **?**, **?**, and **?**). A few papers study a network of buyers and sellers, in which each player can spend resources to link with an additional buyer/seller and therefore increase his bargaining power (see **?**, **?** and **?**). Also relevant is **?**, who study bargaining protocols leading to efficient non-contractible investments prior to matching.

sufficiently destructive, arming will be unavoidable within the class of distribution rules they consider. Also related is the model in **?**, where each party starts by making wasteful investments in armaments. The paper compares the waste produced by three cooperative bargaining solutions: equal sacrifice, equal benefit, and Kalai-Smorodinski. The main result is that if players are symmetric equal sacrifice is the solution generating the lowest waste. Interesting, all these papers consider only what we call "offensive investments," which, under our assumptions, can be completely eliminated via the concession game.

Finally, our paper belongs to the literature studying cooperation in strategic games, that is, the fact that competing players may resolve their dispute using cooperative mechanisms typically interpreted as arbitration procedures (for a recent contribution, see **?**). Within this literature, we are most closely related to **?**, who show that an uninformed arbitrator can construct an arbitration procedure that induces truth-telling from the players (who are assumed fully informed). Relative to our model, they consider a more general bargaining game but, crucially, they do not allow the players to make pre-negotiation investments. Hence, whereas they also characterize the efficient arbitration procedure, this has a very different meaning in our paper since it also implies minimizing wasteful pre-negotiation investments.

The organization of the remainder of the paper is standard. The next section presents the model, the following section solves it, and the last section concludes. All mathematical derivations missing from the text are in Appendix.

2 The model

Two players, 1 and 2 can either cooperatively share a total payoff S or trigger a conflict. At the beginning of the game each player can make investments aimed at shifting the payoffs in case of conflict. Next, each player can unilaterally trigger a conflict. If no player triggers a conflict, then the parties settle their dispute with the help of an arbitrator, who implements an arbitration procedure. The arbitration procedure is common knowledge and is fully incorporated into the players' investment choices.

Conflict

The players are initially characterized by their ex-ante power levels ϕ_1 and ϕ_2, which define their payoffs in the conflict game in case no investments are made.[10] These payoffs may depend on natural elements (e.g. the presence of mountains may make one country harder to attack) or by the merit of the legal dispute. They may also depend on prior investments. The conflict payoffs can be manipulated by the players' offensive o_i and defensive investments d_i. By investing in o_i player i decreases player $-i$'s payoff in the conflict game, while by investing d_i player i can increase his own payoff in the conflict game.[11]

Let us denote by $\underline{u}_i$ the payoff of player i in the conflict game, taking into account her ex-ante power ϕ_i, own defensive investment d_i and the opponent's offensive investment o_{-i}, that is, $\underline{u}_i = u(\phi_i, d_i, o_{-i})$. We specify this payoff function to[12]

$$u(\phi_i, d_i, o_{-i}) \equiv \phi_i e^{-o_j}(2 - e^{-d_i}), \; i, j = 1, 2.$$

The marginal cost of investing in defensive and offensive technology are c_d and c_o, assumed constant.

Note that, if the players expect no negotiation to occur–and hence invest solely in view of influencing the payoff of the conflict game–their offensive investment will be zero and hence the joint payoff is

$$\max_{d_1} \{u(\phi_1, d_1, 0) - c_d \cdot d_1\} + \max_{d_2} \{u(\phi_2, d_2, 0) - c_d \cdot d_2\}.$$

Our main assumption is that sharing S dominates the conflict payoff, that is:

$$S \geq \max_{d_1} \{u(\phi_1, d_1, 0) - c_d \cdot d_1\} + \max_{d_2} \{u(\phi_2, d_2, 0) - c_d \cdot d_2\} \tag{A1}$$

[10] We call it "conflict game" because, in case conflict is triggered, players may take additional costly actions, such as effort, giving rise to a game.

[11] Note that some investments may be simultaneously offensive and defensive, in the sense that they simultaneously increase a player's utility and decrease the opponents utility in case of conflict (for example, hiring a very competent but expensive lawyer or purchasing tanks). Some other investments decrease both players utilities in case of conflicts, and are therefore mutually offensive (for example, nuclear weapons). As a first approximation here we only consider purely offensive or purely defensive investments, however our main result is robust to considering different types of investments (see Remark **??**).

[12] This functional form is chosen for ease of derivations as it allows to compute simple, closed form expressions. However, again, our main result generalizes to any functional form (see Remark **??**).

Information structure

We assume that the two players are fully informed: they observe the power levels ϕ_1 and ϕ_2, the total payoff to be shared S, both players' offensive and defensive investments. The arbitrator instead does not observe neither ϕ_1 or ϕ_2, nor S, nor the players' investments. He only observes the cost of investing c_d and c_o.

An important observation is that the arbitrator's lack of information prevents him from simply imposing a binding settlement—because this settlement may not be feasible, or may be worse than conflict for one of the players (who therefore would prefer triggering a conflict). The arbitration procedure maximizing welfare is therefore a game, in the sense that the resulting settlement is a function of the players' actions.

The arbitration

The arbitration procedure has three steps:

- first, the arbitrator asks the players to report S, ϕ_1, ϕ_2 and their respective investments. We call these reports m_1 and m_2.

- then, both players make simultaneous concessions. If player i makes concessions $b_i \geq 0$, player i bears a cost equal to b_i while player $-i$ enjoys a benefit equal to $\alpha \cdot b_i$, where $\alpha < 1$.[13] Concessions can therefore be interpreted as in-kind gifts, that generate a welfare loss measured by the parameter α. Together with the players' reports, the concessions are used by the arbitrator to set the sharing rule γ (see the next point), which is:

$$\gamma = f(b_1, b_2, m_1, m_2),$$

  where $f(.,.,.,.)$ is bounded between 0 and 1, increasing and concave in the first argument, decreasing and convex in the second argument, continuous and differentiable. We also assume that $\frac{\partial^2 f(b_1,b_2)}{\partial b_1 \partial b_2} \geq 0$.[14]

[13] Our results can be easily extended to more general expressions for the cost and benefit of concessions, including asymmetries between the two players. However, for ease of notation, here we assume a simple, common linear function.

[14] A particular functional form that fits all these requirements is a standard additive contest function of the form

$$f(b_1, b_2, m_1, m_2) = \tau + (1-\tau)\mu_1(b_1) - \tau\mu_2(b_2)$$

- finally, the two parties play an infinitely-repeated bargaining game in which one of the two players is the permanent proposer. In every period, the permanent proposer makes an offer to the other player, who can accept (in which case the game ends), reject and continue the negotiation (in which case the following period repeats identical), or reject and trigger a conflict. The players have discount factors $\beta_1, \beta_2 < 1$. At the start of the game, the permanent proposer is determined: with probability γ player 1 is the permanent proposer, otherwise player 2 is the permanent proposer.

The above defines a family of arbitration procedures, one for every specific $f(b_1, b_2, m_1, m_2)$ for which the concession game has a unique solution. Within this class of arbitration procedures, we will characterize the second-best efficient one, that is, no other $f(b_1, b_2, m_1, m_2)$ achieves higher social welfare given the constraints faced by arbitrator. As we will see, the random permanent proposer bargaining game achieves an efficient split of the peace dividend. Second-best efficiency therefore here amounts to minimizing the investment in arms $c_o(o_1 + o_2) + c_d(d_1 + d_2)$.

Timeline

To summarize, the timeline of the game is the following:

1. Each player chooses his offensive and defensive investments and then decides whether to trigger a conflict.
2. If no player triggers a conflict, then the arbitration procedure starts:
   (a) The players report S, ϕ_1, ϕ_2 and their respective investments to the arbitrator.
   (b) Each player chooses his level of concessions.
   (c) The arbitrator draws the permanent proposer.
   (d) The players play a permanent proposer bargaining game.

where the functions $\mu_1()$ and $\mu_2()$ are strictly increasing, between 0 and 1. Implicitly, both the parameter τ (here a measure of bias) and the shape of $\mu_1()$ and $\mu_2()$ depend on the messages m_1 and m_2.

3 Solution

We solve the game backward. First, for given b_1, b_2, m_1, m_2 (and hence for given γ), we solve for random permanent proposer bargaining game. We then derive the welfare maximizing $f(b_1, b_2, m_1, m_2)$.

3.1 Random permanent proposer.

Without loss of generality, assume that player 1 was selected as random permanent proposer. Call v_1 and v_2 the players' utilities in the subgame perfect equilibrium of the permanent proposer game. Player 1 offer will be such that player 2 is indifferent between accepting the offer, and either going to the following period and earning v_2 or triggering a conflict immediately. The players' utilities are

$$v_1 = S - \max\{\beta_2 v_2, \underline{u}_2\} \qquad v_2 = \max\{\beta_2 v_2, \underline{u}_2\}$$

or

$$v_1 = S - \underline{u}_2 \qquad v_2 = \underline{u}_2$$

Similarly, if player 2 is the permanent proposer, the players' utilities must satisfy

$$v_1 = \max\{\beta_1 v_1, \underline{u}_1\} \quad v_2 = S - \max\{\beta_1 v_1, \underline{u}_1\}$$

or

$$v_1 = \underline{u}_1 \qquad v_2 = S - \underline{u}_1$$

Hence, in the subgame perfect equilibrium of the random permanent proposer game, the players' payoffs are

$$U_1 = \underline{u}_1 + \gamma \left(S - \underline{u}_1 - \underline{u}_2\right), \text{ and } U_2 = \underline{u}_2 + (1-\gamma)\left(S - \underline{u}_1 - \underline{u}_2\right).$$

Note that γ is therefore the share of the peace dividend accruing to each player (see Figure **??**). It follows that the random permanent-proposer bargaining game implements the generalized Nash bargaining solution, in which player 1 receives a share γ of the peace dividend, while player 2 receives the rest, even if the arbitrator does not observe S, $\underline{u}_1$ or $\underline{u}_2$.

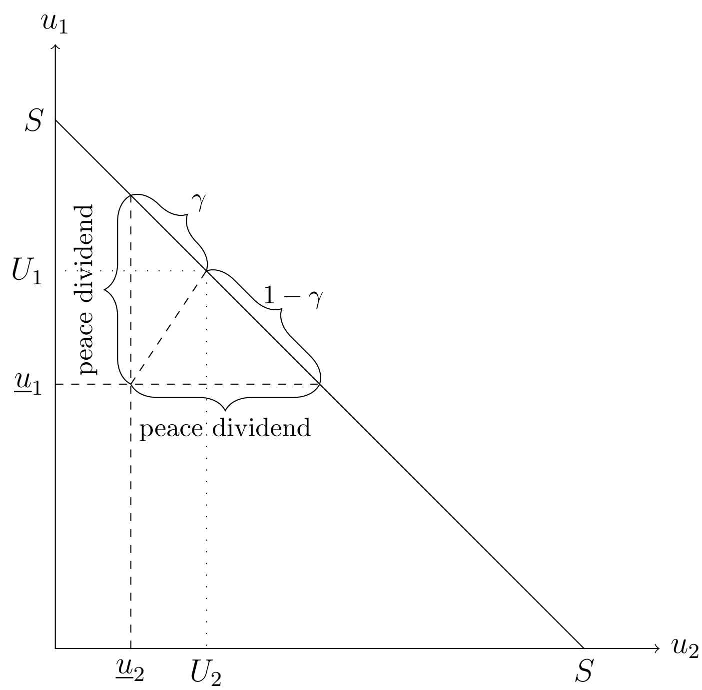

Fig. 1: Solution to the negotiation for given γ

3.2 The concession game

We first present our argument under the assumption that S, ϕ_1, ϕ_2 and the investments have been truthfully revealed by the players. For this reason, here for simplicity we omit the fact that $f(.,.)$ also depends on the previous messages m_1 and m_2, and only write $f(b_1, b_2)$. We then argue that, indeed, the concession game can be constructed so to induce truthful revelation.

Define the peace dividend as $P \equiv S - \underline{u}_1 - \underline{u}_2$, and the equilibrium level of concessions as $b_1(P);\ b_2(P)$. To start, we want to show that the equilibrium level of concessions are zero for P sufficiently low, and are strictly increasing in P for P sufficiently large.

For given b_2, player 1 chooses his level of concessions by solving:

$$\max_{b_1 \geq 0} \left\{ f(b_1, b_2)P - b_1 + \alpha b_2 \right\},$$

The first order condition is:

$$\frac{\partial f(b_1, b_2)}{\partial b_1} = P^{-1}. \tag{1}$$

Hence, if there exists a $b_1 \geq 0$ such that (**??**) has a solution, this b_1 is player 1 optimal concession level. In this case, we say that player 1's optimization problem has an interior solution. Otherwise, the optimal concession level is $b_1 = 0$, in which case we say that player 1's optimization problem has a corner solution. Note also that, for given b_2, as the peace dividend P increases player 1's problem is more likely to have an interior solution. In an interior solution, for given b_2 the level of concession of player 1 is strictly increasing with P.

Similarly, for given b_1 player 2 solves:

$$\max_{b_2 \geq 0} \left\{ (1 - f(b_1, b_2))P - b_2 + \alpha b_1 \right\},$$

with FOC:

$$-\frac{\partial f(b_1, b_2)}{\partial b_2} = P^{-1}. \tag{2}$$

Again, player 2's problem has an interior solution whenever there exists a $b_2 \geq 0$ that solves the above expression, and has a corner solution at $b_2 = 0$ otherwise. Also here, for given b_1, for low P player 2's problem is more likely to have a corner solution at 0. For sufficiently large P, player 2's problem will have an interior solution which is strictly increasing in P.

We now turn to the equilibrium level of concessions. By the implicit function theorem, whenever both (**??**) and (**??**) hold we have:

$$b_1'(P)\frac{\partial^2 f(b_1, b_2)}{\partial b_1^2} + b_2'(P)\frac{\partial^2 f(b_1, b_2)}{\partial b_1 \partial b_2} = -\frac{1}{P^2}$$

$$b_2'(P)\frac{\partial^2 f(b_1, b_2)}{\partial b_2^2} + b_1'(P)\frac{\partial^2 f(b_1, b_2)}{\partial b_1 \partial b_2} = \frac{1}{P^2}.$$

Similarly to a standard contest, also here each player's equilibrium concession $b_i(P)$ is increasing in P, strictly so whenever P is sufficiently large so that player i's problem has an interior solution. Note that this implies that both $b_1'(P)$ and $b_2'(P)$ have a discontinuity at the value of P such that zero concessions solve (**??**) and (**??**).

The important observation is that if the function $f(.,.)$ is such that

$$\frac{\partial f(b_1, b_2)}{\partial b_1}|_{b_1 = b_2 = 0} = P^{-1}.$$

$$-\frac{\partial f(b_1, b_2)}{\partial b_2}|_{b_1 = b_2 = 0} = P^{-1}.$$

then at P there are no concessions in equilibrium. However, if P increases, then the equilibrium concessions can be made arbitrarily large by setting $\frac{\partial^2 f(b_1,b_2)}{\partial b_1^2}$ and $\frac{\partial^2 f(b_1,b_2)}{\partial b_2^2}$ sufficiently low. That is, the equilibrium concessions can be set to zero at a specific peace dividend, but can be made arbitrarily large at any larger peace dividend. Such arbitration procedure is our candidate efficient arbitration procedure. The rest of the paper is devoted to showing that this is indeed the case.

Truthful revelation

Truthful revelation follows from the fact that the arbitrator can punish the players for lying by triggering some arbitrarily large equilibrium concessions.

The argument is familiar. Suppose both players expect truthful revelation. In this case, $m_1 = m_2$ and the arbitration procedure will be such that equilibrium concessions are zero. Suppose now that $m_1 \neq m_2$, and hence one of the players is lying. In this case, the mediator can construct a concession game such that the equilibrium γ is the same as if there were no concessions, while simultaneously generating two arbitrarily large concessions in equilibrium. As a result, both players receive an arbitrarily large punishment in case one of them misreports. Hence, truthful revelation is an equilibrium.

3.3 Offensive and defensive investment

Given this, we can analyze the players' choice of offensive and defensive investments. Player 1 solves

$$\max_{o_1,d_1} \left\{ \{u(\phi_1,d_1,o_2)+\gamma P\} - c_o \cdot o_1 - c_d \cdot d_1 - b_1(P) + \alpha_2 b_2(P) \right\},$$
$$\text{s.t.} \begin{cases} \gamma = f(b_1(P), b_2(P)) \\ P = S - u(\phi_1,d_1,o_2) - u(\phi_2,d_2,o_1) \end{cases}$$

By the envelope theorem, we can ignore the effect of o_1 and d_1 on b_1. The FOC with respect to o_1 is:

$$-\frac{\partial f(.,.)}{\partial b_2}\frac{\partial u(\phi_2,d_2,o_1)}{\partial o_1} b_2'(P)^+ P - \gamma\frac{\partial u(\phi_2,d_2,o_1)}{\partial o_1} - \alpha\frac{\partial u(\phi_2,d_2,o_1)}{\partial o_1} b_2'(P)^+ = c_o,$$

where $b_2'(P)^+$ is $b_2(P)$ right derivative (remember that $b_2'(P)$ may be discontinuous and that offensive investments increase P). If player 2's optimal concession problem has interior solution, then (**??**) holds and the above FOC becomes:

$$-\gamma\frac{\partial u(\phi_2,d_2,o_1)}{\partial o_1}+(1-\alpha)\frac{\partial u(\phi_2,d_2,o_1)}{\partial o_1}b_2'(P)^+=c_o. \tag{3}$$

If player 2's optimal concession problem has corner solution, then $b_2'(P)^+ = 0$. It follows that the above expression defines player 1 optimal offensive investment, both when the subsequent concession game has an interior solution and when the subsequent concession game has an interior solution.

The important observation is that, if the subsequent concession game has an interior solution, player 1 anticipates that by investing in o_1 he will increase the peace dividend and therefore the concessions made by player 2. This has two effects. It directly benefits player 1 because concessions are something valuable to the player receiving them. It, however, indirectly hurts player 1 because concessions by player 2 increase the share of ex-post surplus accruing to player 2. Because $\alpha < 1$ the negative effect dominates, and player 1 decreases his investment in offensive technology to reduce the intensity of the contest over γ.

Similarly, the FOC with respect to d_1 is

$$-\frac{\partial f(.,.)}{\partial b_2}\frac{\partial u(\phi_1,d_1,o_2)}{\partial d_1}b_2'(P)^-P+(1-\gamma)\frac{\partial u(\phi_1,d_1,o_2)}{\partial d_1}-\alpha\frac{\partial u(\phi_1,d_1,o_2)}{\partial d_1}b_2'(P)^-=c_d,$$

which, using (**??**) becomes:

$$(1-\gamma)\frac{\partial u(\phi_1,d_1,o_2)}{\partial d_1}+(1-\alpha)\frac{\partial u(\phi_1,d_1,o_2)}{\partial d_1}b_2'(P)^-=c_d. \tag{4}$$

Again, remember that whenever (**??**) is violated, then $b_2'(P)^- = 0$. It follows the above equation also characterizes player 1 optimal defensive investment when (**??**) does not hold.

In this case, whenever the concession game has an interior solution, it *increases* the benefit of making a defensive investment. The intuition is the reverse of what discussed in the previous section. A defensive investment decreases the ex-post surplus to be shared in the contest and therefore the incentive of both players to perform monetary payments. Hence, by making a defensive investment, player i can decrease b_{-i} and obtain a higher surplus share during the negotiation.

Following similar steps, we can derive the two FOCs for player 2. The FOC for o_2 is

$$-(1-\gamma)\frac{\partial u(\phi_1,d_1,o_2)}{\partial o_2}+(1-\alpha)\frac{\partial u(\phi_1,d_1,o_2)}{\partial o_2}b_1'(P)=c_o, \tag{5}$$

and the FOC with respect to d_2 is:

$$\gamma\frac{\partial u(\phi_2, d_2, o_1)}{\partial d_2} + (1-\alpha)\frac{\partial u(\phi_2, d_2, o_1)}{\partial d_2} b_1'(P) = c_d. \tag{6}$$

To summarize the above observations: whenever (**??**) and (**??**) hold, the contest for γ can be used to discourage offensive investments. The reason is that offensive investments increase the peace dividend and therefore the intensity of the "fight" in the concession game. This logic however implies that the contest for γ may increase the incentives to make defensive investments. If instead the contest for γ has corner solutions, (**??**) and (**??**) do not hold. In this case $b_1'(P) = b_2'(P) = 0$, and the contest for γ does not affect the choice of optimal investment.

An important case is when (**??**) and (**??**) hold at $b_1 = b_2 = 0$. In this case, $b_1'(P)^+$ and $b_2'(P)^+$ are positive but $b_1'(P)^-$ and $b_2'(P)^-$ are zero. In this case the contest for γ can be used to discourage offensive investments, without affecting the incentives to make defensive investments. As already discussed $b_1'(P)^+$ and $b_2'(P)^+$ can be made arbitrarily large for any positive change in P, which implies that the concession contest can be used to eliminate offensive investments. The level of defensive investment is the determined by $\gamma = f(0, 0)$.

At this point, it is useful to make a side comment relative to the above derivations: up until now, they do not depend in any way on the specific functional form assumed or on the fact that investments are purely offensive and purely defensive. This implies the following remark.

Remark 1. *The fact that the arbitration procedure can eliminate all investments that increase the peace dividend while generating zero wasteful concessions generalizes to all functional forms for $\underline{u}_1$, $\underline{u}_2$ and all types of wasteful investments aimed at manipulating $\underline{u}_1$, $\underline{u}_2$.*

Knowing this, we now turn to deriving the most efficient arbitration procedure. An important preliminary result is that it is never efficient to generate concessions in equilibrium.

Lemma 1. *For given investment levels, the welfare-maximizing $f(.,.)$ is such that $b_1(P) = b_2(P) = 0$.*

The proof of the above lemma considers a $f(.,.)$ and the resulting equilibrium offensive and defensive investments. It then shows that if equilibrium concessions are positive, it is possible to construct a different $f(.,.)$ that generates zero equilibrium concessions while inducing either the same or lower levels of investments.[15]

The above lemma implies that, at the welfare maximizing arbitration procedure, we have $b_1'(P)^- = b_2'(P)^- = 0$: if a player makes a defensive investment that decreases the peace dividend, the future equilibrium concessions will be constant at zero. By (**??**) and (**??**) and using the specific functional form for $u(.,.,)$ we get:

$$d_1(\phi_1, o_2, \gamma) = \max\left\{\log\left(\frac{(1-\gamma)\phi_1}{c_d}\right) - o_2, 0\right\} \tag{7}$$

$$d_2(\phi_2, o_1, \gamma) = \max\left\{\log\left(\frac{\gamma\phi_2}{c_d}\right) - o_1, 0\right\}. \tag{8}$$

where $d_i(\phi_i, o_{-i}, \gamma)$ is player i optimal defensive investment.

Depending on the shape of $f(.,.)$, instead, $b_1'(P)^+$ and $b_2'(P)^+$ can take any positive value (zero included). By (**??**) and (**??**), when they are sufficiently large, offensive investment will be zero; whey they are zero, offensive investments will reach their maximum levels. Hence, for given level of defensive investments, the arbitration procedure can induce any level of offensive investments such that:

$$o_1(\phi_2, d_2, \gamma) \in \left[0, \max\left\{\log\left(\frac{\gamma\phi_2(2-e^{-d_2})}{c_o}\right), 0\right\}\right] \tag{9}$$

$$o_2(\phi_1, d_1, \gamma) \in \left[0, \max\left\{\log\left(\frac{(1-\gamma)\phi_1(2-e^{-d_1})}{c_o}\right), 0\right\}\right] \tag{10}$$

Given this, the optimal arbitration procedure can be represented as a γ, o_1 and o_2 such that total waste is minimized, that is:

$$\min_{\gamma\in[0,1];\ o_1\in[0,o_1^{NE}(\gamma)];\ o_s\in[0,o_2^{NE}(\gamma)]} \{c_d(d_1(\phi_1, o_2, \gamma) + d_2(\phi_2, o_1, \gamma)) + c_o(o_1 + o_2)\}$$

where $o^N E_1(\gamma)$ and $o^N E_2(\gamma)$ are the level of offensive investments in the Nash equilibrium of the investment game in case $b_1'(P)^+$ and $b_2'(P)^+$ are zero so that offensive investments are at their maximum (which we derive below). The key observation is that the objective function is linear in o_1 and o_2 (see **??** and **??**). Hence, the welfare minimization problem has two possible corner solutions:

1. when $c_d > c_o$ it is optimal to set the level of offensive investment at its maximum level. This is achieved by setting $\gamma = f(b_1, b_2)$ for all b_1, b_2 so that $b_1'(P)$ and $b_2'(P)$. Because this case is equivalent to not having a concession game, we refer to it as the "no concession game" case.

[15] This is another result that does not depend on the specific functional form we assumed.

Fig. 2: Benefit of player 1's offensive investment for different values of γ.

Fig. 3: Benefit of player 1's defensive investment for different values of γ.

2. when $c_d < c_o$ it is optimal to set the level of offensive investments to zero. The arbitration procedure is such that equilibrium concessions are zero:

$$\frac{\partial f(0,0)}{\partial b_1} = -\frac{\partial f(0,0)}{\partial b_2} = \frac{1}{S - \phi_1(2 - e^{d_1^*}) - \phi_2(2 - e^{d_2^*})},$$

where d_1^* and d_2^* are the equilibrium levels of defensive investments. At the same time the arbitrator sets both $\frac{\partial^2 f(0,0)}{\partial b_1^2}$ and $\frac{\partial^2 f(0,0)}{\partial b_2^2}$ arbitrary large so to eliminate offensive investments. The equilibrium level of defensive investments then depends on $\gamma = f(0,0)$.

We now consider each case separately.

3.3.1 Case 1: $c_d \geq c_o$, no concession game.

In this case:

$$o_1(\phi_2, d_2, \gamma) = \max\left\{\log\left(\frac{\gamma\phi_2(2 - e^{-d_2})}{c_o}\right), 0\right\}$$

$$o_2(\phi_1, d_1, \gamma) = \max\left\{\log\left(\frac{(1-\gamma)\phi_1(2 - e^{-d_1})}{c_o}\right), 0\right\}$$

Note that $o_1(\phi_2, d_2, \gamma)$ and $d_2(\phi_2, o_1, \gamma)$ are both increasing in γ, while $o_2(\phi_1, d_1, \gamma)$ and $d_1(\phi_1, o_2, \gamma)$ are both decreasing in γ. Intuitively, as the share of the peace dividend received increases, a player's payoff depends more and more on the opponent's outside option rather than on his own outside option. In the limit case in which the entire peace dividend is allocated to player i, the final payoff for both players only depends on player $-i$'s outside option. As a consequence, the incentive to degrade the opponent and make an offensive investment increases with the share of the peace dividend received. Similarly, as the share of the peace dividend received *decreases*, a player's payoff depends more and more on his own outside option rather than on his opponent's. It follows that, as the share of the peace dividend received decreases, the incentive to make a defensive investment increases. See Figures **??** and **??** for an illustration.

Note that $o_1(\phi_2, d_2, \gamma)$ and $d_2(\phi_2, o_1, \gamma)$ are best response of each other, and $o_2(\phi_1, d_1, \gamma)$ and $d_1(\phi_1, o_2, \gamma)$ are best responses to each other. There are therefore two separate games. The first one is a "fight over player 2's outside option" in which player 1 makes an offensive investment and player 2 makes a defensive investment. In this game, the two best responses are increasing in γ. The other game is a "fight over player 1's outside option" in which player 1 makes a defensive investment and player 2 makes an offensive investment. In this game, the two best responses are decreasing in γ.

Putting the best responses together, we can characterize the Nash equilibrium of the game (its proof follows by simple algebra and is omitted).

Lemma 2. *If $c_d \geq c_o$, the Nash equilibrium of the investment game is*

$$o_1^{NE}(\gamma) = \max\left\{\log\left(\frac{\gamma\phi_2}{c_o}\right), 0\right\}, o_2^{NE}(\gamma) = \max\left\{\log\left(\frac{(1-\gamma)\phi_1}{c_o}\right), 0\right\}$$

and $d_1 = d_2 = 0$.

In equilibrium there is no investment in defensive technology. Remember that defensive investment is decreasing in offensive investment. When the cost of offensive investment is low relative to the cost of defensive investment, each player will make a large investment in offensive technology and, in equilibrium, drive the incentive to invest in defensive technology of the other player to zero.

Note that if $1 - \frac{c_o}{\phi_1} \leq \frac{c_o}{\phi_2}$, then at $\gamma \in [1 - \frac{c_o}{\phi_1}, \frac{c_o}{\phi_2}]$ defensive investments are zero and full efficiency is achieved. If instead $\frac{c_o}{\phi_2} \leq 1 - \frac{c_o}{\phi_1}$, efficiency is achieved at the γ that solves

$$\min_{\gamma \in [\frac{c_o}{\phi_2}, 1-\frac{c_o}{\phi_1}]} \left\{\log\left(\frac{\gamma\phi_2}{c_o}\right) + \log\left(\frac{(1-\gamma)\phi_1}{c_o}\right)\right\} = \min_{\gamma \in [\frac{c_o}{\phi_2}, 1-\frac{c_o}{\phi_1}]} \left\{\log\left(\gamma(1-\gamma)\right) + \log\left(\frac{\phi_1\phi_2}{c_o^2}\right)\right\}.$$

Note that $\gamma(1-\gamma)$ is a concave function. The solution is therefore one of the two extremes. It is easy to check that the minimum is reached at $\gamma = 1 - \frac{c_o}{\phi_1}$. The next proposition summarizes these observations.

Proposition 1. *Suppose $c_d \geq c_o$. When*

$$c_o \geq \frac{\phi_1\phi_2}{\phi_1 + \phi_2}, \tag{11}$$

then any

$$\gamma^\star \in \left[1 - \frac{c_o}{\phi_1}, \frac{c_o}{\phi_2}\right]$$

drives wasteful investment to zero.

If instead (**??**) *is violated, waste is minimized when*

$$\gamma^\star = 1 - \frac{c_o}{\phi_1} > 1/2.$$

Player 1's offensive investment is strictly positive, while all other investments are zero. Total waste is $c_o \log\left(\left(1 - \frac{c_o}{\phi_1}\right)\left(\frac{\phi_2}{c_o}\right)\right)$.

Condition (**??**) implies that, for given $\phi_1 + \phi_2$, the distribution of initial power is sufficiently uneven. In this case, there are values of γ that eliminate the players' incentives to invest. Instead, whenever the distribution of initial power levels is sufficiently equal so that (**??**) is violated, an unconditional γ is unable to eliminate wasteful investment. This is similar to well know results in tournament theory, in which total waste increases the more similar are the opponents.

More interesting, the above proposition highlights *efficiency may be incompatible with fairness.* If the distribution of power is uneven so that **??** holds, $\gamma^\star = \frac{1}{2}$ eliminates all investments if and only if $\phi_1 \leq 2c_o$. If instead **??** holds but $\phi_1 > 2c_o$, then achieving efficiency requires giving a larger share of the peace dividend to the strongest player (player 1). If the distribution of power is relatively even so that (**??**) is violated, then simple algebra shows that $\gamma^\star > 1/2$. Intuitively, for given share of the peace dividend received, the weak player has a stronger incentive to reduce the opponent's outside option by making offensive investments. The optimal sharing rule compensates for this by allocating a larger share of the peace dividend to the strongest player, therefore making the players' payoff less dependent on the strongest player's outside option.

3.3.2 Case 2: $c_o \geq c_d$, the concession game.

Suppose now that the concession game is structured so to eliminate offensive investments. Given this, for given γ defensive investment is

$$d_1(\phi_1, o_2 = 0, \gamma) = \max\left\{\log\left(\frac{(1-\gamma)\phi_1}{c_d}\right), 0\right\}$$

$$d_2(\phi_2, o_1 = 0, \gamma) = \max\left\{\log\left(\frac{\gamma\phi_2}{c_d}\right), 0\right\}$$

It is easy to check that whenever

$$\gamma \in \left[1 - \frac{c_d}{\phi_1}, \frac{c_d}{\phi_2}\right] \tag{12}$$

there are values of γ for which both defensive investments are zero. Otherwise, again, the value of γ that maximizes welfare is the one that solves

$$\min_{\gamma \in \left[\frac{c_d}{\phi_2}, 1-\frac{c_d}{\phi_1},\right]} \gamma(1-\gamma)$$

This immediately leads to the following proposition.

Proposition 2. *Suppose $c_o \geq c_d$. When*

$$c_d \geq \frac{\phi_1 \phi_2}{\phi_1 + \phi_2}, \tag{13}$$

then any

$$\gamma^\star \in \left[1 - \frac{c_d}{\phi_1}, \frac{c_d}{\phi_2}\right]$$

drives wasteful investment to zero.

*If instead (**??**) is violated, waste is minimized when*

$$\gamma^\star = 1 - \frac{c_d}{\phi_1} > 1/2.$$

Player 2's defensive investment is strictly positive, while all other investments are zero. Total waste is $c_d \log\left(\left(1 - \frac{c_d}{\phi_1}\right)\left(\frac{\phi_2}{c_d}\right)\right)$.

The above proposition is identical to Proposition **??**, with the only difference that the relevant cost here is the cost of making defensive investments (while in Proposition **??** is the cost of making offensive investments).

Beside the small difference, the message is the same. When the distribution of power is sufficiently uneven, then full efficiency may be achievable; but if the distribution of power is too similar, then there will be positive investment at all possible γ. Also, to minimize total defensive investment, the sharing rule may need to be biased toward the strongest player. This intuition is similar to the one already discussed. The strongest player has the strongest incentive to make defensive investments. The optimal sharing rule compensates for this by allocating a larger share of the sharing rule to the strongest player, therefore making the solution less sensitive to this player's outside option.

4 Conclusions

We proposed a novel arbitration scheme. At its core lies the concession game, in which the peace dividend is split as a function of the players' concessions—costly actions that are beneficial to the other player (but also socially wasteful). We allow the players to make pre-negotiation wasteful investments aimed at manipulating the outcome of the conflict that will ensure if the negotiation fails.

Our main result is that the concession game can be used to eliminate all investments that make the conflict more destructive. This result does not depend on the specific functional form we assumed, nor on the specific types of investment we consider. It is of relevance because most of the literature studying pre-negotiation wasteful investments studies precisely these types of investments (for example, investment in arms, see the literature review for more details), which, we show can be always fully eliminated.

By considering only offensive and only defensive investments and a specific functional form (chosen mostly for convenience as it delivers tractable closed form solutions) we then derive conditions under which using the concession game to eliminate offensive investments is optimal. We also derive the efficient equilibrium split of the peace dividend. We show that the first best can be achieved if and only if the player's distribution of power is sufficiently uneven—where power here is a player's payoff in a conflict absent any investment. If the two player's powers are instead sufficiently similar, then there will be strictly positive wasteful investment in equilibrium. We also highlight a tradeoff between efficiency and fairness: reducing pre-negotiation investments may require to allocate more than half of the peace surplus to the strongest player.

Our main assumption is that, unlike the arbitrator, the players are fully informed. Under this assumption, for given pre-negotiation investments, the arbitration procedure can easily achieve an efficient outcome. This allowed us to focus exclusively on inefficiencies arising before the start of the negotiation. However, several authors have drawn a connection between pre-bargaining wasteful investments and inefficiencies arising within the negotiation (see **?**, **?**, **?**, **?**, **?**). For example, an arms build up prior to the negotiation may increase the chance that an agreement is found and therefore increase the efficiency of the negotiation, either because it makes war more costly or because it reduces the asymmetry of information between players. On the other hand, a military mobilization may decrease the probability of reaching an agreement and the efficiency of the negotiation because it generates a *hands-tying effect*: a decrease in the cost of starting a war that operates as a public commitment device. Extending our model to the case in which inefficiencies during the negotiation stage are present, and are affected by inefficiencies arising before the negotiation is left for future work.

Mathematical derivations

Proof of Lemma **??**. Suppose $b_1(P) > 0$ and $b_2(P) > 0$. The four FOCs determining the level of offensive and defensive investments are

$$-\gamma\frac{\partial u(\phi_2, d_2, o_1)}{\partial o_1} + (1-\alpha)\frac{\partial u(\phi_2, d_2, o_1)}{\partial o_1}b_2'(P) \leq c_o$$

$$(1-\gamma)\frac{\partial u(\phi_1, d_1, o_2)}{\partial d_1} + (1-\alpha)\frac{\partial u(\phi_1, d_1, o_2)}{\partial d_1}b_2'(P) \leq c_d$$

$$-(1-\gamma)\frac{\partial u(\phi_1, d_1, o_2)}{\partial o_2} + (1-\alpha)\frac{\partial u(\phi_1, d_1, o_2)}{\partial o_2}b_1'(P) \leq c_o,$$

$$\gamma\frac{\partial u(\phi_2, d_2, o_1)}{\partial d_2} + (1-\alpha)\frac{\partial u(\phi_2, d_2, o_1)}{\partial d_2}b_1'(P) \leq c_d,$$

Note how $b_i'(P)$ decreases player i's incentive to make an offensive investment but increases player i's incentive to make defensive investments.

Suppose instead that $b_1(P) = b_2(P) = 0$. Assume furthermore that (**??**) and (**??**) hold at $b_1(P) = b_2(P) = 0$. The four FOCs determining the level of offensive and defensive investments are

$$-\gamma\frac{\partial u(\phi_2, d_2, o_1)}{\partial o_1} + (1-\alpha)\frac{\partial u(\phi_2, d_2, o_1)}{\partial o_1}b_2'(P)^+ \leq c_o$$

$$(1-\gamma)\frac{\partial u(\phi_1, d_1, o_2)}{\partial d_1} \leq c_d'(d_1)$$

$$-(1-\gamma)\frac{\partial u(\phi_1, d_1, o_2)}{\partial o_2} + (1-\alpha)\frac{\partial u(\phi_1, d_1, o_2)}{\partial o_2}b_1'(P)^+ \leq c_o,$$

$$\gamma\frac{\partial u(\phi_2, d_2, o_1)}{\partial d_2} \leq c_d,$$

Note how, in this case, $b_i'(P)^+$ decreases player i's incentive to make an offensive investment but has no impact on player i's incentive to make defensive investments.

Remember that $b_1'(P)^+$ and $b_2'(P)^+$ depend on $f(.,.)$. Te shape of $f(.,.)$ can therefore be set so that offensive investment is lower when $b_1(P) = b_2(P) = 0$ then when $b_1(P) > 0$ and $b_2(P) > 0$. Furthermore, the left hand side of the FOCs for the defensive investments are lower under $b_1(P) = b_2(P) = 0$ than $b_1(P) > 0$ and $b_2(P) > 0$ for any level of $b_1'(P)$ and $b_2'(P)$. That is, it is possible to move from $b_1(P) > 0$ and $b_2(P) > 0$ to $b_1(P) = b_2(P) = 0$, while decreasing all investments.

Finally, note that if $b_1(P) > 0$ and $b_2(P) > 0$ and all investments are already zero, it is always possible to set the equilibrium concessions to zero by manipulating $\frac{\partial f(b_1,b_2)}{\partial b_1}$ and $\frac{\partial f(b_1,b_2)}{\partial b_2}$, while at the same time maintaining $b_1'(P)$ and $b_2'(P)$ (and with it the incentives to make offensive and defensive investments) constant. This is welfare increasing because it eliminates concessions (which are socially costly) while maintaining all investments at zero. $\square$