COSTLY BARGAINING AND RENEGOTIATION*

by

Luca Anderlini
St John's College, Cambridge

and

Leonardo Felli
London School of Economics and Political Science

Contents:

Theoretical Economics Workshop
The Suntory Centre
Suntory and Toyota International Centres
for Economics and Related Disciplines
London School of Economics and Political Science
Houghton Street
London WC2A 2AE
Tel.: 0171-405 7686

Discussion Paper
No. TE/98/361
October 1998

* We are grateful to Vincent Crawford, Eric Maskin, Steven Matthews, Andrew Postlewaite, Larry Samuelson, Ilya Segal, and seminar participants at S.I.T.E. 1998 (Stanford) and the Stockholm School of Economics for helpful discussions and comments. The research work for this paper was carried out while both authors were visiting the Department of Economics at the University of Pennsylvania. Their stimulating hospitality is gratefully acknowledged. Of course, were are solely responsible for any remaining errors.

Abstract

We identify the inefficiencies that arise when negotiation between two parties takes place in the presence of transaction costs. First, for some values of these costs it is efficient to reach an agreement but the unique equilibrium outcome is one in which agreement is never reached. Secondly, even when there are equilibria in which an agreement is reached, we find that the model always has an equilibrium in which agreement is never reached, as well as equilibria in which agreement is delayed for an arbitrary length of time.
Finally, the only way in which the parties can reach an agreement in equilibrium is by using inefficient punishments for (some of) the opponent's deviations. We argue that this implies that, when the parties are given the opportunity to renegotiate out of these inefficiencies, the only equilibrium outcome which survives is the one in which agreement is never reached, regardless of the value of the transaction costs.

Keywords: Optimal bargaining costs; inefficient bargaining outcomes; renegotiation; imperfect recall.

JEL Nos.: C70, D23, D60, D80.

1. Introduction

1.1. Motivation

The Coase theorem (Coase 1960) is one of the cornerstones of modern economic analysis. It shapes the way economists think about the efficiency or inefficiency of outcomes in most economic situations. It guarantees that, if property rights are fully allocated, economic agents will exhaust any mutual gains from trade. Fully informed rational agents, unless they are exogenously restricted in their bargaining opportunities, will ensure that there are no *unexploited gains from trade.*

This view of the (necessary) exploitation of all possible gains from trade is at the centre of modern economic analysis. Economists faced with an inefficient outcome of the negotiation between two rational agents will automatically look for reasons that impede full and 'frictionless' bargaining between the agents.

In this and a companion paper (Anderlini and Felli 1997) we focus on the impact of *transaction costs* on the Coase theorem. In both cases, we show that, in a complete information world, transaction costs imply that the Coase theorem no longer holds in the sense that an efficient outcome is no longer guaranteed.

In the model which we analyze in this paper we find that for certain (transaction costs) parameter values only inefficient equilibria are possible, while for other parameter values both efficient and inefficient equilibria obtain. In the latter case we find that it is not possible to 'select' the efficient outcomes in a 'consistent' way: there are no equilibria of the model which guarantee an efficient outcome in every subgame.

Given the impossibility of selecting efficient outcomes by fiat, we proceed as follows. Keeping as given the friction introduced by the transaction costs, we expand the negotiation possibilities for the two agents — we build into the extensive form opportunities for the parties to 'renegotiate' out of inefficient outcomes. We find that, in this case, the *only* equilibrium outcome which survives is the most inefficient possible one: agreement is never reached and the entire surplus fails to materialize.

Intuitively, our results are closely related to a large body of literature in contract theory (Grossman and Hart 1986, Hart and Moore 1988, Maskin and Moore 1998, Dewatripont 1989, Segal 1998, Hart and Moore 1998) in which ex-post renegotiation possibilities harm ex-ante efficiency. We discuss the relationship between this paper and existing literature in further detail in Section 1.3 below.

1.2. Costly Bargaining

Our point of departure is the leading extensive form model of negotiation between two parties, namely an alternating offers bargaining game with complete information with potentially infinitely many rounds of negotiation in which the players discount the future at a strictly positive rate (Rubinstein 1982). We introduce transaction costs in the following way. Both parties, at each round of negotiation, must pay a positive cost to 'participate' to that round of the bargaining game. At each round, both parties have a choice of whether or not to pay their respective participation costs. Each round of negotiation takes place only if both parties pay their participation costs. If either player decides not to pay his participation cost, the negotiation is postponed until the next period.

The interpretation of the participation costs which we favour is the following. At the beginning of each period, both parties must decide (irrevocably for that period) whether to spend that period of time at the negotiation table, or to engage in some other activity which yields a positive payoff. The participation costs in our model can simply be though of as these alternative payoffs which the agents forego in order to engage in the negotiation activity for that period.

The first sense in which Coase theorem fails in our model is the following. There exists values of the participation costs such that it is efficient for the parties to reach an agrement (the sum of the costs is strictly smaller than the size of the surplus) and yet the *unique* equilibrium of the game is for the parties never to pay the costs so that an agreement is never reached (Theorem 2 below).[1]

Having established Theorem 2 below, we proceed to focus on the case in which the values of the participation costs are 'low enough' so that the parties will be able to reach an agreement in equilibrium. In this case the model displays a wide variety of equilibria: (efficient) equilibria with immediate agreement (Theorem 3 below), (inefficient) equilibria with an agreement with an arbitrarily long delay (Theorem 4

[1] Theorem 2 below is related to Proposition 14 in Anderlini and Felli (1997). The chief difference between the two is that here we consider players who discount the future at a strictly positive rate.

below), and (inefficient) equilibria in which an agreement is never reached (Theorem 1 below). Therefore, the Coase theorem fails in this case too in the sense that it is no longer *necessarily* the case that the outcome of the bargaining between the parties is Pareto efficient.

In the case in which the participation costs are such that there are both efficient and inefficient equilibria, a natural reaction is that it is just a matter to choose the right selection criterion to be able to isolate the efficient equilibria. If this were possible one would conclude that, in a sense, the Coase theorem does not fail in this setting for low enough transaction costs. In Section 5 below, we are able to show that this way of proceeding does not work in our model. The reason is that all equilibrium agreements are sustained by off-the-equilibrium-path inefficient strategies needed to punish the players for not paying their participation costs. Therefore, it is impossible to apply a selection criterion which implies efficiency in a consistent way across every subgame. The set of equilibria that survives any such selection criterion is empty in our model (Theorem 5 below).

The fact that inefficient equilibrium outcomes are possible in our model leads naturally to the question whether the source of the inefficiency and the failure of Coase theorem lies in the limited negotiation opportunities given to the parties. To address this question we proceed in the following way. We modify the extensive form of the game so as to allow the parties a chance to start a fresh negotiation whenever they are playing strategies that put them strictly within the Pareto frontier of their payoffs. We do this by modifying the extensive form of the game and transforming it in a game of imperfect recall. We assume that, at the beginning of each period, with strictly positive probability, the parties do not recall the past history play. This affords them a chance to 'renegotiate out' of inefficient punishments. The result is devastating for the equilibria in which agreement is reached. When the probability of 'forgetting' the history of play is above a minimum threshold (smaller than one), the unique equilibrium outcome of the modified game is for the parties never to pay the costs and therefore never to reach an agreement. This is true regardless of the size of the participation costs, provided of course that they are positive.

We view this as the most serious failure of Coase theorem in our model. If one expands the parties' opportunities to bargain the inefficiency becomes extreme. Agreement is never reached, whatever the size of the transaction costs.

1.3. Related Literature

There are several strands of existing literature to which our work is related. Purely for the sake of clarity, we distinguish between four main ones: the Coase theorem, the vast literature on inefficiencies and delays in bargaining, the large and fast growing literature on incomplete contracts, and the sizeable literature on renegotiation in finitely and infinitely repeated games.

It is clear that the original version of the Coase theorem (Coase 1960) explicitly assumes the absence of any transaction costs or other frictions in the bargaining process. Indeed, Coase (1992) describes the theorem as a provocative result that was meant to show how unrealistic is the world without transaction costs. It should, however, be noticed that, sometimes, subsequent interpretations of the original claim have strengthened it way beyond the realm of frictionless negotiation. It does not seem uncommon for standard microeconomics undergraduate texts to suggest that the Coase theorem should hold in the presence of transaction costs.[2].

Anderlini and Felli (1997) is a paper that can be viewed as a companion to the present one. There, we are concerned with the 'hold-up' problem generated by 'ex-ante' contractual costs in a stylized contracting model. The focus of our analysis in Anderlini and Felli (1997) is mainly the robustness of our inefficiency to changes in a number of assumptions. In particular, we concentrate on the nature of the costs payable by the parties to make the contracting stage feasible, and on the possibility that the parties may rely on an 'expanded contract' which includes contracting on the ex-ante costs themselves. By converse, in this paper we take it as given that the parties bargain according to the standard 'alternating offers' protocol, and that they have to pay their participation costs in order to negotiate at each round. The main focus of this paper is then to explore the effect of allowing the parties to 'renegotiate' out of inefficient punishments. We find that 'renegotiation opportunities' work against ex-ante efficiency. The (inefficient) punishments needed to sustain efficient equilibria

[2]For instance, an excellent textbook widely in use in the U.S. and elsewhere claims that, in its strongest formulation, the Coase theorem is interpreted as guaranteeing an efficient outcome whenever the potential mutual gains "exceed [the] necessary bargaining costs" (Nicholson (1989, p.726))

are no longer available to the players. The only equilibrium outcome which survives is a highly inefficient one. As we noted before, the fact that ex-ante efficiency is harmed by renegotiation possibilities is a common phenomenon in a variety of contracting models (Grossman and Hart 1986, Hart and Moore 1988, Maskin and Moore 1998, Dewatripont 1989, Segal 1998, Hart and Moore 1998). Renegotiation opportunities embedded in the bargaining process may imply that efficiency is not guaranteed whenever the parties to a contract need to bargain over the available surplus, both at the ex-ante and at the ex-post stage.

We are certainly not the first to point out that the Coase theorem no longer holds when there are 'frictions' in the bargaining process. There is a vast literature on bargaining models where the frictions take the form of incomplete and asymmetric information. Our analysis here shares the concern of many of these papers with the nature of the inefficiencies that may arise in equilibrium. With incomplete information, efficient agreements often cannot be reached and delays in bargaining may obtain (Admati and Perry 1987, Ausubel and Deneckere 1989, Chatterjee and Samuelson 1987, Fudenberg, Levine, and Tirole 1987, Grossman and Perry 1986, Hart 1989, Rubinstein 1985, Sobel and Takahashi 1983, Cramton 1992, among others).

The source of inefficiencies in these papers differs from what we find in our analysis here — we work in a complete information set-up, while the source of inefficiencies in these models is the strategic behaviour designed to conceal (one's own) information or to extract information (about the opposing player), or both. In the setting in which bargaining parties are asymmetrically informed it is true (with the exception of Admati and Perry (1987) and Cramton (1992)) that when the parties are given renewed opportunities to negotiate with each other (in the form of more frequent offers) efficiency is restored (Gul, Sonnenschein, and Wilson 1986). This is not true in our model. Our basic inefficiency result survives even if we allow the frequency of offers to increase without bound (Remark 3 below).

Fernandez and Glazer (1991) and Busch and Wen (1995) are both concerned with the possible inefficient equilibria of bargaining models with complete information. These papers highlight how delays may arise when the parties have the option to strike holding up the other party during negotiation. In this setting, however, inefficient delays may be observed in some of the multiple subgame perfect equilibria of the game. Some of the equilibria of the game instead are always efficient. The difference with our analysis is in the nature of the costs the parties incur. In some sense the nature of the costs considered in Fernandez and Glazer (1991) is the exact opposite to the one tackled here. In Fernandez and Glazer (1991) the parties may choose to pay a cost *not to* negotiate for a period (strike) while in our setting the parties have to pay the cost to negotiate for that period. In our setting the unique subgame perfect equilibrium of the game is inefficient in some cases, while in their setting an efficient equilibrium is always available.

Both Riedl (1997) and Dixit and Olson (1997) also focus on complete information models. In both of these papers, the parties make ex-ante decisions whether to participate or not in the bargaining process, much as we postulate here. Riedl (1997) does not concentrate, as we do here, on the possible inefficiencies which might arise. Dixit and Olson (1997) are concerned with a classical 'Coasian' public good problem, and consider both one-shot and repeated versions of the same model. While they find both efficient and inefficient equilibria, they highlight the inefficiency of the symmetric (mixed-strategy) equilibria of their model.

The vast and fast growing literature on incomplete contracts is not only related to the present paper, but in some sense is one of the motivating points of our research. In particular, we view our work as related to the many contributions that have concentrated on the possible causes of contractual incompleteness (Hart and Moore 1988, Anderlini and Felli 1994, MacLeod 1997, Maskin and Tirole 1998, Segal 1998, Hart and Moore 1998, among others). This literature focuses on the sources of inefficiencies that may arise in a relationship in which the parties attempt to exploit the potential gains from trade through a contract. The contract represents the tool though which the surplus 'materialises' while bargaining is the process through which the parties select such contract, and hence 'distribute the surplus' among them. In a sense, the set of feasible mechanisms or contracts defines the Pareto frontier faced by the parties in their process of bargaining, which is often left in the background by contracting models. If the Coase theorem hods, of course the parties will select a contract which corresponds to a point on this frontier.

The existing literature on the causes of contractual incompleteness has focused mainly on the constraints that apply to the set of feasible mechanisms and how these constraints modify the Pareto frontier faced by the contracting parties. Hart and Moore (1988) argue that in an environment in which the relevant state of nature is observable but not verifiable and the contracting parties cannot commit not to renegotiate, the Pareto frontier becomes constrained in such a way as to force the choice of a contract that lies strictly below the Pareto frontier which corresponds to the case in which the state of nature is verifiable. On the other hand, Maskin and Tirole (1998) argue that these constraints on the Pareto frontier in the presence of observable but not verifiable states of nature stem from the fact that message contingent mechanisms are ruled out. Once these mechanisms are allowed the frontier coincides with the one that would arise in a world in which the states of nature are verifiable. On the other hand, message contingent mechanisms are allowed in both Segal (1998) and Hart and Moore (1998). In both of these papers, the key constraints on the Pareto frontier stem from the increasing complexity of the environment faced by the parties coupled with their inability to commit not to renegotiate.

Anderlini and Felli (1994) and MacLeod (1997), on the other hand, focus on the constraints imposed on the Pareto frontier, by the fact that the contracting parties face 'writing constraints' in their contract technology and/or by the fact that they are 'boundedly rational' in their perception of the Pareto frontier.

The present paper takes a completely different perspective. Instead of considering the constraints on the Pareto frontier faced by the contracting parties, we focus on the bargaining process through which the contract is selected. In the presence of some forms of transaction costs in the bargaining process, the parties will not necessarily choose a mechanism on the Pareto frontier, but rather select an inefficient (sometimes extremely inefficient) and hence 'incomplete' contract.

The last strand of literature which must be mentioned here is the one on renegotiation in finitely and infinitely repeated games. Benoît and Krishna (1993), Farrell and Maskin (1989), Bernheim and Ray (1989) and Abreu, Pearce, and Stacchetti (1993) (among others) have looked at the effects of renegotiation in the context of finitely and infinitely repeated games. Some other contributions have addressed the issue of renegotiation in an implementation setting (Maskin and Moore 1998, Moore 1992, Aghion, Dewatripont, and Rey 1994). All of these papers focus on 'black-box' renegotiation, modelled as an equilibrium selection criterion, or (in the case of implementation) as an exogenously given function that whenever an inefficient allocation is selected, puts the parties back on the Pareto frontier. Within this literature, Rubinstein and Wolinsky (1992) have asked how the impact of renegotiation possibilities changes when a 'time dimension' is explicitly taken into account. Their main finding is that the addition of the time dimension greatly reduces the constraints on the set of outcomes which are implementable in a renegotiation proof way.

The approach we take in this paper is different in that we explicitly modify the extensive form of the bargaining game so as to include the renegotiation opportunities in the structure of the game. We further discuss the rationale behind our choice, and the relationship with black-box renegotiation, in Section 5 below.

1.4. Overview

The paper is organized as follows. In Section 2 we describe in detail our model of alternating offers bargaining with transaction costs. Section 3 contains our first inefficiency result and a characterization of the equilibria of the model described in Section 2. In Section 4 we investigate the robustness of the inefficient and of the efficient equilibria of our model to some basic changes in the description of the game. In Section 5 we show that it is impossible to consistently select the Pareto efficient equilibria of our game in a way which is consistent across subgames. Section 6 contains our model of renegotiation opportunities in the extensive form. Here, we present our second main result — namely the fact that the only equilibrium outcome of our game of imperfect recall is that agreement is never reached. Section 7 briefly concludes the paper. For ease of exposition all proofs are relegated in the Appendix.

2. The Model

We consider a bargaining game between two players indexed by $i \in \{A, B\}$. The game consists of potentially infinitely many rounds of alternating offers $n = 1, 2, \ldots$. Each player i has to pay a participation cost at round n denoted c_i (constant through time). We interpret this cost as the opportunity cost to player i of the time the player has to spend in the next round of bargaining.

In all odd periods $n = 1, 3, 5, \ldots$, player A has the chance to make offers, and player B the chance to respond. In all even periods $n = 2, 4, 6, \ldots$ the players' roles are reversed, B is the proposer, while A is the responder. Throughout the paper we refer to the odd periods as 'A periods' and to even periods as 'B periods'.

The size of the surplus to be split between the players is normalized to one. Any offer made in period n is denoted by $x \in [0, 1]$. This denotes A's share of the pie, if agreement is reached in period n. The discount factor of player i is denoted by $\delta_i \in [0, 1)$.[3]

To clarify the structure of each round of bargaining, it is convenient to divide each time period in three *stages*. In stage I of period n, both players decide simultaneously and independently, whether to pay the costs c_i. If both players pay their participation costs, then the game moves to stage II of period n. At the end of stage I, both players observe whether or not the other player has paid his participation cost. If one, or both, players do not pay their cost, then the game moves directly to stage I of period $n + 1$.

In stage II of period n, if n is odd, A makes an offer $x \in [0, 1]$ to B, which B observes immediately after it is made. At the end of stage II of period n, the game moves automatically to stage III of period n. If n is even, the players' roles in stage II are exactly reversed.

In stage III of period n, if n is odd, B decides whether to accept or reject A's offer. If B accepts, the game terminates, and the players receive the payoffs described in (1) below. If B rejects A's offer, then the game moves to stage I of period $n + 1$. If n is even, the players' roles in stage III are exactly reversed.

The players' payoffs consist of their shares of the pie (zero if agreement is never reached), minus any costs paid, appropriately discounted. To describe the payoffs formally, it is convenient to introduce some further notation at this point. Let (σ_A, σ_B) be a pair of strategies for the two players in the game we have just described, and consider the *outcome path* $\mathcal{O}(\sigma_A, \sigma_B)$ which these strategies induce. Let $\mathcal{N}_i(\sigma_A, \sigma_B)$

[3] Some of our results do in fact generalize to the case $\delta_A = \delta_B = 1$ (see in particular Remark 3 below). We concentrate on the case $\delta_i \in [0, 1)$ because we are interested in a set-up in which *delays* in bargaining yield *inefficient* outcomes. The reasons for this are two-fold. The first is that one of the types of inefficiencies which might arise in our model is precisely a delay in reaching agreement (see Theorem 4 below). The second is that our way of embedding 'renegotiation opportunities' into the extensive form of the game is best suited to a situation in which the only viable punishments are *inefficient* ones. We return to this point in Subsection 6.1 — see footnote 8 below.

be the number of periods in which player $i \in \{A, B\}$ pays his participation cost along the outcome path induced by (σ_A, σ_B). Let also $\mathcal{C}_i(\sigma_A, \sigma_B) = \mathcal{N} c_i$, so that $\mathcal{C}_i(\sigma_A, \sigma_B)$ is the total of participation costs which i pays along the outcome path induced by (σ_A, σ_B).

If the outcome path $\mathcal{O}(\sigma_A, \sigma_B)$ prescribes that the players agree on an offer x in period n, then the payoffs to A and B are respectively given by

$$\Pi_A(\sigma_A, \sigma_B) = \delta_A^n x - \mathcal{C}_A(\sigma_A, \sigma_B) \quad \text{and} \quad \Pi_B(\sigma_A, \sigma_B) = \delta_B^n (1 - x) - \mathcal{C}_B(\sigma_A, \sigma_B) \qquad (1)$$

while if the outcome path $\mathcal{O}(\sigma_A, \sigma_B)$ prescribes that the players never agree on an offer, then the payoff to player $i \in \{A, B\}$ is given by

$$\Pi_i(\sigma_A, \sigma_B) = -\mathcal{C}_i(\sigma_A, \sigma_B)$$

3. Subgame Perfect Equilibria

In this section we provide a full characterization of the set of subgame perfect equilibria of the alternating offer bargaining game described in Section 2 above.

We first show that the equilibrium in which the players do not ever pay the costs and enter the negotiation is always an SPE of the bargaining game.

Theorem 1: *Consider the alternating offers bargaining game with participation costs described in Section 2. Whatever the values of δ_i and c_i for $i \in \{A, B\}$, there exists an SPE of the game in which neither player pays his participation cost in any period, and therefore an agreement is never reached.*

We now proceed to characterize the necessary and sufficient conditions on the pair of costs (c_A, c_B) and the parties' discount factors (δ_A, δ_B) under which the parties are able to achieve an agreement.

Theorem 2: *Consider the alternating offers bargaining game with participation costs described in Section 2. The game has an SPE in which an agreement is reached in finite time if and only if δ_i and c_i for $i \in \{A, B\}$ satisfy*

$$\delta_A(1 - c_A - c_B) \geq c_A \quad \text{and} \quad \delta_B(1 - c_A - c_B) \geq c_B \qquad (2)$$

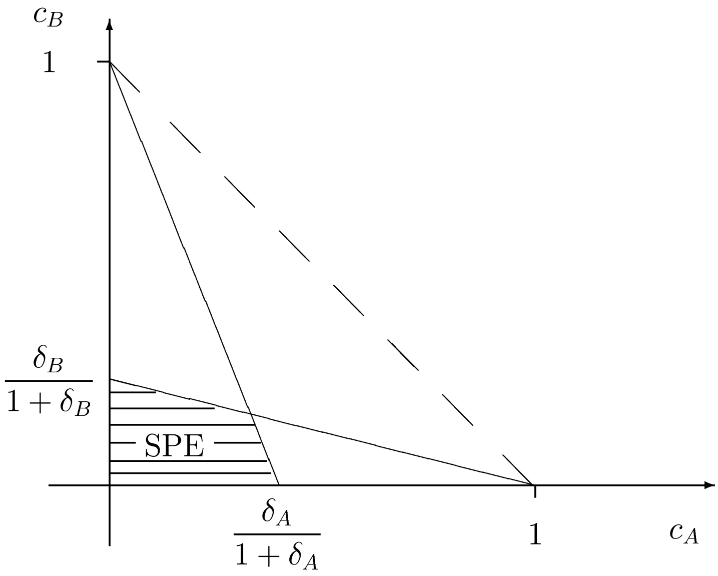

Figure 1: SPE with Agreement in Finite Time

For given δ_A and δ_B, the set of costs (c_A, c_B) for which an agreement is reached is represented by the shaded region in Figure 1.

Using Figure 1, it is also immediate to see that Theorem 2 supports our first inefficiency claim. The sum of the participation costs is less than the total available surplus anywhere below the dashed line in Figure 1. Given any pair of discount factors, there exist a region of possible participation costs such that the model has a unique, inefficient, SPE outcome. In Figure 1, for any pair (c_A, c_B) below the dashed line but outside the shaded area, the participation costs add up to less that one, but no agreement is ever reached. In Remark 3 below we make explicit the fact that this inefficiency result extends to the case in which the players do not discount the future.

We now proceed to give a more detailed characterization of the SPE with agreements of this game. We start by identifying the range of possible equilibrium shares of the pie in every possible subgame when agreement is immediate.

Theorem 3: *Consider the alternating offers bargaining game with participation costs described in Section 2, and assume that δ_i and c_i for $i \in \{A, B\}$ are such that (2) holds so that the game has some SPE in which an agreement is reached in finite time. Consider the subgames in which it is A's turn to make an offer to B (the A subgames from now on). Then there exists an SPE of the A subgames in which x_A*

is agreed immediately, if an only if

$$x_A \in [1 - \delta_B(1 - c_A - c_B), 1 - c_B] \tag{3}$$

Symmetrically, consider the subgames in which it is B's turn to make an offer to A (the B subgames from now on). Then there exists an SPE of the B subgames in which x_B is agreed immediately, if and only if

$$x_B \in [c_A, \delta_A(1 - c_A - c_B)] \tag{4}$$

Our next result both closes our characterization of the set of SPE payoffs, and supports our second inefficiency claim. Every sharing of the pie which can be supported as an immediate agreement can also take place with a delay of an arbitrary number of periods.

Theorem 4: *Consider the alternating offers bargaining game with participation costs described in Section 2, and assume that δ_i and c_i for $i \in \{A, B\}$ are such that (2) holds so that the game has some SPE in which an agreement is reached in finite time.*

Let any x_A as in (3) and any odd number n be given. Then there exists an SPE of the A subgames in which the (continuation) payoffs to the players are respectively given by

$$\Pi_A = \delta_A^n(x_A - c_A) \quad \Pi_B = \delta_B^n(1 - x_A - c_B) \tag{5}$$

Moreover, let any x_B as in (4) and any even number n be given. Then there exists an SPE of the A subgames in which the (continuation) payoffs to the players are respectively given by

$$\Pi_A = \delta_A^n(x_B - c_A) \quad \Pi_B = \delta_B^n(1 - x_B - c_B) \tag{6}$$

Symmetrically, let any x_B as in (4) and any odd number n be given. Then there exists an SPE of the B subgames in which the (continuation) payoffs to the players are as in (6).

Moreover, let any x_A as in (3) and any even number n be given. Then there exists an SPE of the B subgames in which the (continuation) payoffs to the players are as in (5).

4. Robustness of Equilibria

In this section, we carry out three 'robustness exercises' about the SPE of the game described in Section 2 which we have identified in Section 3.

Our first concern is the relationship between the set of SPE of our game with the set of SPE of a finite version of the same game. The unique SPE identified by Rubinstein (1982) of the same bargaining game when there are no participation costs has many 'reassuring' properties. Among these is the fact that if a version of the same game with a truncated time horizon is considered, the limit of the SPE of the finite games coincides with the unique SPE of the infinite horizon game. This is not the case in our bargaining model with participation costs. In fact when we truncate the time horizon to be finite in our model, the only possible SPE outcome is the one in which neither player ever pays his participation cost, and hence no agreement is reached.

The intuition behind Remark 1 below is a familiar 'backward induction' argument. No agreement is possible in the last period since the responder would have to get a share of zero if agreement is reached, and therefore he will not pay his participation cost in that period. This easily implies that no agreement is possible in the last period but one, and so on.

Let Γ^∞ represent the infinite horizon alternating offers bargaining game with participation costs described in Section 2. For any finite $N \geq 1$, let Γ^N represent the same game with time horizon truncated at N. In other words, in Γ^N, if period N is ever reached, the game terminates, regardless of whether an agreement has been reached or not. If no agreement has been reached by period N, the players' payoffs are zero, minus any costs paid of course. We can then state the following.

Remark 1: *Let any finite $N \geq 1$ be given. Then the unique SPE outcome of Γ^N is neither player pays his participation cost in any period and hence agreement is never reached.*

Trivially, Remark 1 implies that the only SPE outcome of Γ^∞ which is in fact the limit of any sequence of SPE outcomes of Γ^N, as N, grows is the one in which agreement is never reached.

Our next concern is the robustness of the SPE in which neither party ever pays his participation cost and hence no agreement is ever reached to the sequential payments of the participation costs. It is a legitimate concern to check whether this equilibrium is attributable to a simple 'coordination failure' or whether it depends on other feature of the structure of our alternating offers bargaining game with participation costs. It turns out that this SPE is indeed robust to the players paying their participation costs sequentially, before any offer is made and accepted or rejected. Let $\mathcal{S}$ be any sequence of the form $\{i_1, i_2, \ldots, i_n, \ldots\}$, where $i_n \in \{A, B\}$ for every n. Let $\Gamma(\mathcal{S})$ be the game derived from the one described in Section 2, modified as follows. In stage I of period n, player i_n first decides whether to pay his participation cost or not. Following i_n's decision, the other player observes whether i_n has paid his cost or not, and then decides whether to pay his own participation cost. The description of stages II and III of every period in $\Gamma(\mathcal{S})$ is exactly the same as for the original game described in Section 2. We are then able to state the following.

Remark 2: *Fix any arbitrary sequence $\mathcal{S}$ as described above. Then $\Gamma(\mathcal{S})$ always has an SPE in which neither player ever pays his participation cost, and hence agreement is never reached.*

Our last concern is with the robustness of our first inefficiency result to the case of no discounting. This is a natural question to ask since the case of $\delta_A = \delta_B = 1$ can be interpreted as the limit of our model in which the time interval between offers shrinks to zero. As we noted in the Introduction (Section 1.3), in some models of bargaining with asymmetric information (see for instance (Gul, Sonnenschein, and Wilson 1986)), efficiency is guaranteed in the limit as the delay between offers tends to zero. This is not so in our set-up.[4]

Remark 3: *Consider the alternating offers bargaining model with participation costs described in Section 2, but set $\delta_A = \delta_B = 1$. Then the model has an SPE with*

[4]Remark 3 which follows is in fact a re-statement of Proposition 14 in Anderlini and Felli (1997).

agreement in finite time if and only if

$$1 - c_A - c_B \geq \max\{c_A, c_B\} \tag{7}$$

Therefore, there exist a set of pairs of participation costs (c_A, c_B) for which the unique SPE outcome of the game is that agreement is never reached, but which are such that an agreement between the parties is efficient in the sense that $c_A + c_B < 1$.

5. Consistently Pareto Efficient Equilibria?

Theorems 3 and 4 tightly characterize the SPE payoffs of the alternating offers bargaining game described in Section 2, when the players agree in finite time on how to divide the available surplus.

On the other hand, Theorem 1 tells us that the game always also has an SPE in which no agreement is reached in finite time. In this SPE, neither player ever pays his participation cost and the players' payoffs are zero.

Thus all the subgames have both Pareto efficient equilibria, in which an agreement is reached immediately (see Theorem 3), and a 'highly' inefficient one in which the surplus is completely dissipated through an infinite delay (see Theorem 1). There are also SPE in which part of the surplus is dissipated since agreement takes place, but is delayed by a finite number of periods (see Theorem 4).

Recall that the game described in Section 2 is one of *complete* information. This makes the inefficient SPE of the game all the more 'surprising'. As we mentioned in Section 1.3 above, there is a vast literature on bargaining with incomplete information in which inefficient delays are found to be a common equilibrium phenomenon (Admati and Perry 1987, Ausubel and Deneckere 1989, Chatterjee and Samuelson 1987, Fudenberg, Levine, and Tirole 1987, Grossman and Perry 1986, Hart 1989, Rubinstein 1985, Sobel and Takahashi 1983, Cramton 1992).

A natural question to ask at this point, and one which is central to this paper, is whether the inefficient SPE of the alternating offers bargaining game with participation costs described in Section 2 can be 'ruled out'.

It is tempting to argue as follows. Since the game at hand is one of complete information, there are no possible strategic reasons for either player to delay agreement. Neither player can possibly hope to accumulate a 'reputation' which will help in subsequent stages of the game. Neither player can possibly gain information about the other player as play unfolds. Therefore, the players will somehow 'agree' to play an efficient equilibrium in which no delays occur. The players will in some way 'renegotiate' out of inefficient equilibria.

This line of reasoning, in our view, is flawed on at least two accounts. The first concerns the modelling of 'renegotiation' in a bargaining game. The second is that, in the game described in Section 2, once renegotiation possibilities are explicitly taken into account, the only SPE which survives is in fact the one in which an agreement is never reached. Therefore the SPE characterized by the most extreme form of inefficiency is the one that is robust to the introduction of renegotiation. Section 6 is entirely devoted to this claim.

The difficulty in taking into account renegotiation possibilities in a bargaining game stems from a simple observation. A bargaining game is, by definition, a model of how the negotiation proceeds between the two players. When they are explicitly modelled, clearly there should be no intrinsic difference between negotiation and renegotiation. Renegotiation is just another round of negotiation, which takes place (ex-post) if the original negotiation has failed to produce an efficient outcome. In short, in a model of negotiation, renegotiation possibilities should be explicitly taken into account in the *extensive form*, rather than grafted as a 'black box' onto the original model. This is what we do in Section 6 below.

In the remainder of this Section, we point out that a simple-minded 'black box' view of renegotiation does not work in the game described in Section 2.

Suppose that, in a 'Coasian' fashion we attempt simply to select for efficient outcomes in our bargaining game with participation costs. A minimal 'consistency requirement' for this operation is that we should recognize that each stage of the bargaining game at hand is in fact an entire negotiation game by itself. Therefore, if we believe that efficient outcomes should be selected simply on the grounds that they are efficient, we should now be looking for an SPE which yields an efficient outcome in *every subgame* of the bargaining game. It turns out that this is impossible.

We first proceed with the formal definition of a 'consistently Pareto efficient' SPE and with our next result, and then elaborate on the intuition behind it.

Definition 1: *An SPE* (σ_A, σ_B)*is called* Consistently Pareto Efficient *(henceforth CPESPE) if and only if it yields a Pareto-efficient outcome in every possible subgame.*

We now show that it is impossible to sigle out an SPE which is consistently Pareto efficient in the way we have just described.

Theorem 5: *Consider the alternating offers bargaining game with participation costs described in Section 2. The set of CPESPE for this game is empty.*

The intuition behind Theorem 5 can be outlined as follows. A CPESPE must yield an agreement in every period, regardless of the history of play that lead to that subgame.[5]

Recall that, except for the participation costs our bargaining game is the original alternating offers bargaining game analyzed by Rubinstein (1982). Once we impose that an agreement must be reached in every period, we can reason about our model in a way which closely parallels well known arguments that apply to the model with no participation costs. In particular (adapting the argument used by Shaked and Sutton (1984)) one can show the following. First of all if an SPE were to exist with agreement in every period, then there would be a *unique* share of the pie x_A which is sustainable in equilibrium in every A subgame, and a unique share of the pie x_B for every B subgame. Moreover, x_A and x_B have the following property. In stage III of every A subgame, B is exactly indifferent between accepting A's offer x_A and rejecting it, and, symmetrically, in stage III of every B subgame A is exactly indifferent between accepting x_B and rejecting it. Therefore, in stage I of every A period, B has an incentive not to pay his participation cost: by moving to the next period he earns a payoff which is larger by precisely c_B. Similarly in stage I of every B period, player A can gain c_A by not paying his cost and forcing the game to move to the next stage.

We conclude this section by recalling that various notions of 'renegotiation proofness' were developed by Benoît and Krishna (1993) (for finitely repeated games), and by Bernheim and Ray (1989), Farrell and Maskin (1989), Farrell and Maskin

[5]Notice that the definition of CPESPE does *not* imply that the *same* agreement must be reached irrespective of history. It only implies that some agreement must be reached in every period, whatever the history of play that lead the players to arrive at the subgame.

(1987) and Abreu, Pearce, and Stacchetti (1993) (among others) for infinitely repeated games.

Our bargaining game with participation costs, of course is neither a finite game nor a repeated game. Although none of the notions of renegotiation proofness readily applies to our model, some of the existing ones can be adapted to fit it. It turns out that the adapted form of these existing notions of renegotiation proofness have little bite in our set-up. This is chiefly because one can exploit the alternating offers nature of the game, and the possibility of (one-period) delays, to punish one player while rewarding the other. However, it should be noted that Theorem 5 implies that to sustain an agreement as an SPE outcome, *inefficient* punishments (off-the-equilibrium-path) are *necessary*. Clearly these must take the form of (off-the-equilibrium-path) delays of one period or more. Definition 1 above is designed to highlight this feature of any SPE involving an agreement in our model.

However, as we stated above, we do not believe that grafting a renegotiation 'refinement' onto a negotiation game is the correct way to proceed. We take Theorem 5 above simply to say that there is no way consistently to select efficient outcomes in our game. Its value lies mainly in clarifying that this is not possible, and in making explicit the 'sunk cost' nature of the intuition behind this fact.

On the basis of Theorem 5 the inefficient SPE of our game have to be granted equal dignity with the efficient ones at this stage of the analysis. In the next section, we proceed to incorporate renegotiation possibilities into the extensive form of the game, and to argue that in this case the SPE with no agreement in finite time is selected among the many possible ones.

6. Extensive Form Renegotiation

6.1. Modelling Renegotation

In this section we modify the bargaining game described in Section 2 in a way which, in our view, embeds into the extensive form the chance for the players to renegotiate out of inefficient outcomes.

We do this in a way which is designed to satisfy three, in our view critical, criteria. First of all, whenever the players find themselves trapped in an inefficient (punishment) phase of play, the extensive form has to give them at least a chance to 'break' out of this inefficient outcome path. Secondly, the possibility of renegotiation must be built into the extensive form as a possibility, rather than an obligation to start afresh and switch to an efficient equilibrium. This is because we want to ensure that our way of tackling the problem here is distinct from the 'black box' renegotiation which we discussed in Section 5 above. If the extensive form in some way 'forced' efficient play whenever an inefficient outcome path has started, there would be little difference between 'extensive form' renegotiation and 'black box' renegotiation. Our third criterion is closely related to the second one — the extensive form we study must be non-trivial in the sense that it must *allow* in principle for the outcome path both on- and off-the-equilibrium-path to be inefficient. If this were not the case, besides violating our second criterion, via Theorem 5, we would automatically know that the equilibria of the modified extensive form have little to do with the SPE of the original game. This is simply because, Theorem 5 tells us that there are no SPE of the original game which yield a Pareto efficient outcome in every subgame.

We modify the bargaining game described in Section 2 by transforming it into a game of *imperfect recall*.[6] At the end of each round of negotiation, we introduce a positive probability that the players might *forget* the previous history of play. It should be noticed that in the event of forgetfulness, we do allow the players to condition their future actions on time. In other words, the players forget the outcome path which has taken place so far, but are *not* constrained to play the same strategy starting at every 'forget' information set.[7]

We believe that this type of imperfect recall is a cogent way to model renegotiation opportunities in our model. Recall that, as we noted before, the crucial inefficient punishments in the bargaining game described in Section 2 are the ones which may be used to punish a player who has not paid his participation cost.[8] As all off-

[6]To our knowledge, bargaining games with imperfect recall have not been analysed before in any form (see footnote 11 below for further references on games with imperfect recall). Chatterjee and Sabourian (1997) analyse a bargaining game (with N players) in which the players have *bounded memory* because of *complexity* considerations.

[7]Notice that imposing that the players play the same strategy at every possible 'forget' information set would clash with the alternating offers nature of the bargaining protocol, which we want to preserve. The players need to know, at least, wheter n is odd or even in order to know whose turn it is to be a proposer in the bargaining.

[8]Viable punishments are necessarily inefficient because both players discount the future at a strictly positive rate, and punishments involve delayed agreements (see Theorem 5 and the discussion

the-equilibrium-path punishments these represent 'history dependent' switches in the behaviour of the players. The probability of forgetting the past history of play represents a chance to 'forgive and forget' for the payers. More specifically, given that the players know the 'date', even when they forget, they are able to infer something about the previos history of play, even when they find themselves at a 'forget' information set — namely that an agreement has not been reached so far. Crucially, however, they are unable to distinguish between the possible different 'reasons' for the failure to reach an agreement. There are, of course, three possible such reasons: failure to pay the participation costs, a deviation at the offer stage, and a deviation at the response stage of the previous bargaining rounds. When they forget, our players will be unable to punish (or reward) in different ways for these three types of behaviour. Notice further that one of these three types of deviations naturally implies a 'reward' in an alternating offers bargaining game. When the proposing player deviates to offer a share of the pie to the responder which is lower than what it 'should be', the responder must be 'rewarded' in the future with a payoff which is larger than the offer he rejected. The necessary reward in this case builds into the extensive form a robust reason to avoid punishments for all three types of deviation when the players forget the past history of play.

Theorem 6 below states that when the probability of forgetting the past history of play in each period is above a minimum cut-off value (strictly below one), then the only equilibrium outcome of our modified bargaining game with imperfect recall is for neither player to ever pay his participation cost, and hence that no agreement is ever reached. In our view, this confirms that, when renegotiation possibilities are introduced, regardless of the values of participation costs, the unique equilibrium outcome of our model is that an agreement is never reached. In the presence of transaction costs and renegotiation embedded in the bargaining procedure, the Coase theorem may fail in a very strong way: no agreement is ever reached, and the entire surplus fails to materialise.

which follows its statement). The picture changes considerably if it is possible to delay agreement without dissipating any of the surplus. In this case it would be possible to punish (one player) without resorting to inefficient outcomes. In this case the 'lack of memory' we are introducing would rule out efficient as well as inefficient punishments. Our way of modelling renegotiation opportunities is clearly less well suited to this case.

6.2. Bargaining With Imperfect Recall

The game which we analyse here is a modification of the game described in Section 2 above along the following lines. At the beginning of each period $n \geq 1$, we add an additional stage, stage O, in which Nature makes a chance move. Nature's draws are described by a sequence of i.i.d. random variables $\Delta = \{\mathcal{D}^1, \mathcal{D}^2, \ldots, \mathcal{D}^n, \ldots\}$. The realization of each of the $\mathcal{D}^n$ is denoted by d^n and takes one of two possible values: $d^n = \mathcal{F}$ (for 'forget') with probability p, and $d^n = \mathcal{R}$ (for 'recall') with probability $1 - p$.[9] For future reference, it is also convenient to define now the (sub)sequences of Δ starting in periods 1, 2, 3 and so on. For any $m \geq 1$ let $\Delta_m = \{\mathcal{D}_m, \mathcal{D}_{m+1}, \ldots\}$. Moreover, let μ be the (product) probability measure over the sequence Δ, and for any $m \geq 1$ let μ_m be the same over Δ_m. Finally, a typical realization of Δ will be denoted by $\sigma_N \in \Sigma_N$, while $\sigma_N^m \in \Sigma_N^m$ denotes the same for Δ_m, for every $m \geq 1$.[10]

The players do not observe the outcome of $\mathcal{D}^n$ until the *end* of period n, after the responder has accepted or rejected the proposer's offer in stage III of period n or either player has not paid his cost in stage I of period n. If the realization of $\mathcal{D}^n$ is $\mathcal{R}$, the game moves to period $n+1$ (unless, of course, an offer has been made and accepted in period n, in which case the game terminates) with all the nodes corresponding to different outcome paths within period n belonging to distinct information sets. If, on the other hand, the realization of $\mathcal{D}^n$ is $\mathcal{F}$, and the game has not terminated in period n, the players forget the previous history of play. In other words, in this case, for both players, all the nodes corresponding to stage I of period $n + 1$, via any possible history of play up to and including the whole of period n, are in the same information set. The description of the extensive form within stages I, II and III of each period is exactly the same as for the model described in Section 2 above.

We want to carachterize the Nash equilibria of this game of imperfect recall which satisfy *sequential rationality.* As it is well known, in general, in games of imperfect recall this can pose a variety of technical problems and questions of interpretation.[11]

[9]While independence of these random variables plays a role in the proof of Theorem 6 below, it is easy to show that the actual probability p could be made to depend on time without affecting our results.

[10]Notice the slight redundancy in our notation. In fact we have $\Delta = \Delta_1$, $\mu = \mu_1$, $\sigma_N = \sigma_N^1$, and $\Sigma_N^1 = \Sigma_N$. This is purely to save on some subscripts in later manipulations.

[11] Recently, Piccione and Rubinstein (1997) have sparked a debate on the interpretation of certain games of imperfect recall. We refer to their work and to the other papers in the special issue of

Luckily, in the case at hand matters seem to be considerably simpler than in the general case.

Given that we are dealing with a game of *incomplete information*, our equilibrium concept is (the weakest version of) *Perfect Bayesian Equilibrium* (PBE hereafter).[12]

We start by dealing with the beliefs of the players at every possible information set. Consider the possible information sets of each player at the beginning of period n.[13] Notice that since the information that the two players have about the past history of play is always exactly the same, and since the past history of play does not affect directly any future payoffs, their beliefs about which node they are actually at, within any information set, are in fact *irrelevant* for their decisions at this stage.

Consider now the players' beliefs about the future. Clearly, in equilibrium, we must require these to be correct insofar as the other player's strategy is concerned, conditional on any possible realization of the future draws by Nature. Moreover, since the draws by Nature are unaffected by any actions the players might take, and are i.i.d., the players' beliefs about future draws are entirely pinned down by the objective probability distribution μ over the sequence of random variables Δ.

Some further notation is necessary to define formally the PBE of our game with imperfect recall. Let $\mathcal{S}_i^j$ with $i, j \in \{A, B\}$ be the strategy set of player i in the original bargaining game described in Section 2, but modified, if necessary, so that player j is the proposer in the first period. Let also s_i be a typical element of $\mathcal{S}_i^A \cup \mathcal{S}_i^B$. It will be convenient to think of a generic strategy for player i in our game of imperfect recall as an object of the following sort.

$$\sigma_i = \{s_i^1, s_i^2, s_i^3 \ldots\} \tag{8}$$

where $s_i^n \in \mathcal{S}_i^A$ if n is odd and $s_i^n \in \mathcal{S}_i^B$ if n is even. In other words, we can think of

---

Games and Economic Behavior (1997) for further details and references. Here we simply notice that the game we are analysing does not exhibit 'absent-mindedness' in the sense which they specify.

[12]See Fudenberg and Tirole (1991)

[13]We are omitting from our discussion here the fact that, since the players move *simultaneously* when they decide to pay their participation costs their information sets will actually differ. This is only so because they are not informed of each other's decision of whether to pay or not. This is irrelevant to our argument since their beliefs about each other's decision of whether to pay or not the particpation cost will simply have to be correct in any Nash equilibrium.

player i as choosing a strategy in $\mathcal{S}_i^A$ for every odd period, and a strategy in $\mathcal{S}_i^B$ for every even period. We denote the set of all possible strategies for player i as in (8) by Σ_i

To define the outcome path induced by a pair (σ_A, σ_B) we now have to take into account the sequence σ_N of Nature's moves. Given σ_N, for every $n \geq 2$, let $f(\sigma_N, n)$ be the last period less than or equal to n in which the players forget. That is let[14]

$$f(\sigma_N, n) = \max\{m \quad | \quad d^m = \mathcal{F} \quad \text{and} \quad m \leq n-1\} + 1 \tag{9}$$

We are now ready to define the (continuation) outcome path(s) $\mathcal{O}(\sigma_A, \sigma_B, \sigma_N)$ induced by the triple $(\sigma_A, \sigma_B, \sigma_N)$. The outcome path in period 1 is simply determined by the strategies s_A^1 and s_B^1. Provided that the game has not terminated at some earlier stage, the outcome path in period $n \geq 2$ depends on what the previous moves by Nature have been. Recall that, at n, the last period in which the players forget is $\sigma_N(n)$. If at n we have that $f(\sigma_N, n) = n$ then the outcome path at n is determined simply by s_A^n and s_B^n. If at n we have that $f(\sigma_N, n) \leq n-1$ we have to consider the possibility of off-the-equilibrium-path behaviour between stage I of period $f(\sigma_N, n)$ and stage III of period n. In this case the outcome path in period n is determined by the strategies $s_A^{f(\sigma_N,n)}$ and $s_B^{f(\sigma_N,n)}$, *conditional* on the history of play between stage I of period $f(\sigma_N, n)$ and stage III of period $n-1$.

In Figure 2 we have depicted schematically (for the first three periods), how player i's 'overall' strategy, σ_i, depends on Nature's moves and on the chosen sequence of 'elementary' strategies $\{s_i^1, s_i^2, \ldots\}$. Each branch of the tree in Figure 2 is labelled with the 'elementary' strategy s_i^n which is 'active' at that point, given the previous sequence of Nature's moves.

Notice that the above construction, given a triple $(\sigma_A, \sigma_B, \sigma_N)$ yields a (continuation) outcome path starting at every possible information set in the game. Let $\mathcal{I}_i^n$ denote the set of all possible information sets for player i in period n, with typical element I_i^n and let $\mathcal{I}^n$ denote the union of $\mathcal{I}_A^n$ and $\mathcal{I}_B^n$, with tipycal element I^n. Notice further that every I^n contains enough information to determine the value of

[14]Notice that the term '+1' on the right-hand-side of (9) takes into account that when $d^m = \mathcal{F}$, the players effectively forget the past history of play from stage I of period $m+1$.

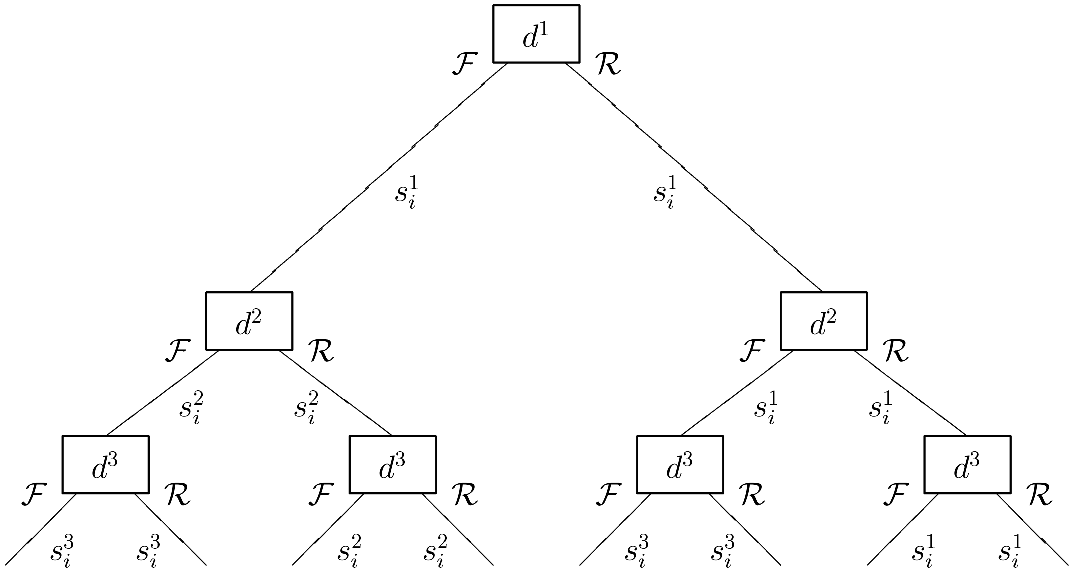

Figure 2: 'Overall' Strategies and 'Elementary' Strategies For Player i

$f(\sigma_N, n)$. In other words to each I^n there corresponds a *unique* value of $f(\sigma_N, n)$.[15] It follows that, using the construction above, for any n and for any I^n, we can define $\mathcal{O}(\sigma_A, \sigma_B, \sigma_N^n | I^n)$ as the (continuation) outcome path generated by the triple $(\sigma_A, \sigma_B, \sigma_N^n)$, conditional on the information set I^n having been reached.

The (continuation) outcome path $\mathcal{O}(\sigma_A, \sigma_B, \sigma_N^n | I^n)$ will either yield an agreement x^m in some period $m \geq n$ or no agreement at all. If an agreement is reached, we define the (continuation) payoff to players A and B to be given respectively by

$$\Pi_A(\sigma_A, \sigma_B, \sigma_N^n | I^n) = \delta_A^{m-n+1} x^m - \mathcal{C}_A(\sigma_A, \sigma_B, \sigma_N^n | I^n) \tag{10}$$

and

$$\Pi_B(\sigma_A, \sigma_B, \sigma_N^n | I^n) = \delta_B^{m-n+1}(1 - x^m) - \mathcal{C}_B(\sigma_A, \sigma_B, \sigma_N^n | I^n) \tag{11}$$

where $\mathcal{C}_i(\sigma_A, \sigma_B, \sigma_N^n | I^n)$ is the total of participation costs which player i pays along the outcome path $\mathcal{O}(\sigma_A, \sigma_B, \sigma_N^n | I^n)$. If $\mathcal{O}(\sigma_A, \sigma_B, \sigma_N^n | I^n)$ yields no agreement in finite

[15]This simply says that the palyers always know when was the *last* time that they forgot the history of play. If they did not, they would exhibit a form of 'absent-mindedness' in the sense of Piccione and Rubinstein (1997).

time, the players' payoffs are respectively given by

$$\Pi_A(\sigma_A, \sigma_B, \sigma_N^n | I^n) = -\mathcal{C}_A(\sigma_A, \sigma_B, \sigma_N^n | I^n) \tag{12}$$

and

$$\Pi_B(\sigma_A, \sigma_B, \sigma_N^n | I^n) = -\mathcal{C}_B(\sigma_A, \sigma_B, \sigma_N^n | I^n) \tag{13}$$

It is now easy to define the expected (continuation) payoffs for each player at every information set. For $i \in \{A, B\}$, they are given by

$$\Pi_i^E(\sigma_A, \sigma_B | I^n) = \sum_{\sigma_N^n \in \Sigma_N^n} \mu^n(\sigma_N^n) \Pi_i(\sigma_A, \sigma_B, \sigma_N^n | I^n) \tag{14}$$

We are now ready to define formally what is required for a PBE in our game of imperfect recall.

Definition 2: *A Perfect Bayesian Equilibrium (PBE) for our bargaining game with imperfect recall is a pair of strategies and a set of beliefs such that, at every information set, the strategies are optimal given beliefs and beliefs are obtained from equilibrium strategies and observed actions using Bayes' rule. Formally, in our game a PBE is a pair of strategies* (σ_A^*, σ_B^*) *as in (8) such that, for every* n *and for every* $I^n \in \mathcal{I}_A^n$

$$\sigma_A^* \in \underset{\sigma_A \in \Sigma_A}{\operatorname{argmax}}\ \Pi_A^E(\sigma_A, \sigma_B^* | I^n) \tag{15}$$

and symmetrically, for every n *and for every* $I^n \in \mathcal{I}_B^n$

$$\sigma_B^* \in \underset{\sigma_B \in \Sigma_B}{\operatorname{argmax}}\ \Pi_B^E(\sigma_A^*, \sigma_B | I^n) \tag{16}$$

We are now ready to state formally our last result.

Theorem 6: *Consider the alternating offers bargaining game with participation costs and imperfect recall described above. For any given pair of costs* (c_A, c_B) *there* *exists a* $\bar{p} < 1$*, which is independent of the discount factors* δ_A *and* δ_B *with the following property. Whenever the probability* p *that the players forget the past history of play in every period exceeds* $\bar{p}$*, then the unique PBE of the game is such that, along any possible realization of Nature's moves, both players never pay their participation cost in any period, and therefore an agreement is never reached.*

The intuition behind the proof of Theorem 6 is relatively simple to describe. In a sense, it is a rather more complex version of the 'sunk-cost' argument that provides the intuition of Theorem 5.

Suppose that an agreement x_B is reached in a period in which B is the proposer and A is the responder. The share x_B must satisfy several constraints. First of all, A's 'net' payoff, $x_B - c_A$, must be at least as much as what A gets if he does not pay his participation cost. This of course means that x_B must be at least as much as A's continuation payoff if he does not pay his cost, plus c_A.

The agreed share x_B must also be less than or equal to the continuation payoff which A gets if he rejects offers below x_B in stage III of the agreement period. This is because A must be better off by rejecting *any* offer below x_B rather than by accepting it.

Putting the above two facts together, tells us the following. The continuation payoff to A after he rejects offers below x_B cannot be smaller that A's continuation payoff if he does not pay his cost in stage I, plus c_A. But, when the players forget the history, these two continuation payoffs for A must in fact be the *same*. Clearly this cannot be the case for large enough p, whenever c_A is positive. For large enough p, when c_A is positive, A is better off by not paying his participation cost, thus moving the game into the next period.

We view Theorem 6 as the most serious indication that inefficiencies are 'pervasive' in our bargaining model with participation costs. In the original game which we described in Section 2, the no agreement equilibrium outcome, for low enough participation costs, was one of many possible ones. When the parties are given the possibility to renegotiate out of inefficient punishments, it is the only one which survives, for *any* positive values of the participation costs. In a bargaining game with positive participation costs, Coasian renegotiation opportunities destroy the efficient equilibria which a simple-minded interpretation of the Coase theorem would lead us to select among the many possible ones.

7. Conclusions

This paper shows that when negotiation takes place in the presence of transaction costs the Coase theorem does not necessarily hold. In particular we show that in an alternating offers bargaining game under perfect information, and with discounting, several types of inefficiencies may arise.

These inefficiencies should be viewed as pervasive for at leat two reasons. First of all, we have shown that it is impossible consistently to 'select' for efficient equilibria in our model. Secondly, and in our view more importantly, if the parties are given sufficient opportunities to renegotiate out of inefficient outcomes, the only outcome which survives in equilibrium is in fact the most inefficient possible one.

We conclude by noticing that, since we restricted ourselves to extensive forms of the bargaining game which follow the alternating offers 'protocol', and in which both players discount the future at a strictly positive rate, our arguments are clearly specific to this setting. Wether the inefficiencies we find are also specific to this setting is clearly a matter for future research. A natural question to ask is what happens if the identity of the proposer in a given period is 'endogenized'. An interesting way to achieve this would be to make the identity of the proposer depend on the amount of cost that each party pays at the beginning of each period. We conjecture that inefficiencies will arise in this setting as well.

APPENDIX

Lemma A.1: *Consider the alternating offers bargaining game with participation costs described in Section 2. Whatever the values of δ_i and c_i for $i \in \{A, B\}$, in any SPE of the game the payoffs to both players are non-negative.*

Simply notice that either player can guarantee a payoff of zero by playing a strategy which prescribes never to pay any of his participation costs. ■

Proof of Theorem 1: We simply display a pair of strategies (σ_A^0, σ_B^0) which constitute an SPE of the game and which yield the desired outcome path.

For all $i \in \{A, B\}$, the strategy σ_i^0 is described as follows. In stage I of any period σ_i^0 prescribes that i does not pay his participation cost, regardless of the previous history of play. In stage II of any period in which i is a proposer, σ_i^0 prescribes that i demands the entire pie for himself ($x = 1$ if $i = A$ and $x = 0$ if $i = B$), regardless of the previous history of play. In stage III of any period in which i is a responder, σ_i^0 prescribes that i accepts any offer $x \in [0, 1]$, regardless of the previous history of play. It is easy to check that these strategies constitute an SPE of the game, and therefore this is enough to prove the claim. ■

Lemma A.2: *Consider the alternating offers bargaining game with participation costs described in Section 2. Assume that δ_i and c_i for $i \in \{A, B\}$ are such that the game has an SPE in which an agreement is reached in finite time (see Theorem 2)*

Let x^L be the infimum and x^H the supremum of all possible equilibrium agreements over the entire set of SPE of the game. Let also x_i^L be the infimum and x_i^H the supremum of all possible equilibrium agreements over the set of SPE in which an agreement is reached with player i being the proposer (the set of i SPE). Both x_i^L and x_i^H are undefined if the set of i SPE is empty.

Then x_i^L and x_i^H are defined for all $i \in \{A, B\}$, and they satisfy $x_A^H = x^H \leq 1 - c_B$, $x_B^L = x^L \geq c_A$ as well as

$$x_B^H \leq \delta_A(x_A^H - c_A) \tag{A.1}$$

and

$$x_A^L \geq 1 - \delta_B(1 - x_B^L - c_B) \tag{A.2}$$

Proof: By Lemma A.1, in any SPE the payoffs to both players must be non-negative. The fact that it must be that $x^H \leq 1 - c_B$ and $x^L \geq c_A$ is now obvious since if the first inequality is violated B would get a negative payoff in some SPE and if the second inequality is violated, A would get a negative payoff in some SPE.

By hypothesis, the set of SPE which prescribe some agreement is not empty. Therefore, either the set of A SPE is not empty, or the set of B SPE is not empty, or both are not empty.

If the set of B SPE is not empty we must have that

$$x_B^H \leq \delta_A(x^H - c_A) \tag{A.3}$$

To see this, consider the subgame which starts in stage III of the agreement period. If A rejects B's offer at this stage, he will get a continuation payoff which is bounded above by $\delta_A(x^H - c_A)$.

Therefore, it must be that A's SPE strategy prescribes to accept *any* offer above $\delta(x^H - c_A)$. Therefore, in stage II of this period, B's equilibrium strategy cannot be one that offers any $x > \delta(x^H - c_A)$, since otherwise he could reduce his offer by a small amount and A would still respond by accepting the offer. Therefore B's offer must be some $x \leq \delta_A(x^H - c_A)$, and this is clearly enough to prove that (A.3) must hold in this case.

Notice next that (A.3) also implies the following. If the set of B SPE is not empty, then the set of A SPE is also not empty. This is because (A.3) says that $x_B^H < x^H$ so that it must be that case that $x_A^H = x^H$.

Using a completely symmetric argument to the one above, we can now argue that if the set of A SPE is not empty then we must have that

$$1 - x_A^L \leq \delta_B(1 - x_B^L - c_B) \tag{A.4}$$

which can be rewritten as (A.2) Using a symmetric argument again, we can then see that (A.4) implies that if the set of A SPE is not empty than it must be the case that the set of B SPE is not empty either. Indeed, it must be the case that $x_B^L = x^L$.

Since we have just argued that either the sets of A and B SPE are both empty or both not empty, and by hypothesis at least one is in fact not empty, (A.3) and (A.4) are enough to prove the claim. ■

Proof of The 'Only If' Part of Theorem 2: Using Lemma A.2, we know that if the set of SPE in which an agreement is reached is not empty we must have that

$$\begin{array}{ll} x_A^H \leq 1 - c_B & x_A^L \geq 1 - \delta_B(1 - c_A - c_B) \\ x_B^H \leq \delta_A(1 - c_A - c_B) & x_B^L \geq c_A \end{array} \tag{A.5}$$

Recalling that, by definition, $x_i^H \geq x_i^L$ for $i \in \{A, B\}$, (A.5) directly implies (2). This is clearly enough to prove the claim. ■

Lemma A.3: *Consider the alternating offers bargaining game with participation costs described in Section 2. Let $\{x^n\}_{n=1}^{\infty}$ be a sequence of numbers such that $x^n \in [c_A, 1 - c_B]$ for all n and such that for all odd n*

$$\delta_B(1 - x^{n+1} - c_B) \geq 1 - x^n \tag{A.6}$$

and for all even n

$$\delta_A(x^{n+1} - c_A) \geq x^n \tag{A.7}$$

Then there exists an SPE of the game (σ_A, σ_B) as follows.

i) If at any point in the previous history of play either or both players have not paid their participation costs, then the strategies (σ_A, σ_B) *revert to being the same as the strategies* (σ_A^0, σ_B^0) *described in the proof of Theorem 1 for the remainder of the game.*

ii) Unconditionally in stage I *of period* 1*, and in stage* I *of every period conditionally on the fact that i) above must not apply, both players pay their participation costs.*

iii) Provided that i) above does not apply, in stage II *of every period* n *the proposing player makes an offer* x^n *to the responding player.*

iv) Provided that i) above does not apply, in stage III *of every period* n *the responding player accepts all offers which leave him with a share of the pie at least as large as the offer* x^n*, and he rejects all other offers.*

v) If the responding player rejects any offer which he is supposed to accept according to iv) above, then strategies (σ_A, σ_B) *revert to being the same as the strategies* (σ_A^0, σ_B^0) *described in the proof of Theorem 1 for the remainder of the game.*

Proof: By Theorem 1, the strategies (σ_A, σ_B) constitute an equilibrium for any subgame following a history as in i).

We now concentrate on the subgames starting stage I of an odd period n, following a history to which i) does not apply (or the empty history if $n = 1$). The argument for the even periods is symmetric and we omit the details.

Consider then any such subgame. By deviating and not paying his cost each player would earn a continuation payoff of zero. Following the prescription of (σ_A, σ_B) both players earn a continuation payoff of at least zero in any subgame. Therefore neither player has an incentive to deviate in any of these subgames.

Next, consider the subgame following the one above, starting in stage II of an odd period n. Clearly player A does not want to deviate and offer an $x < x^n$ (the offer will be accepted and this will lower A's payoff). Suppose now that player A deviates and offers $x > x^n$. Then his continuation payoff is $\delta_A(x^{n+1} - c_A)$. Since (A.6) implies $x^n \geq x^{n+1}$, we have that $x^n > \delta(x^{n+1} - c_A)$. Therefore, this is not a profitable deviation for player A.

Move now to the subgame following the one above, starting in stage III of an odd period n. At this point, some offer x has been made by A. Suppose first that $x \leq x^n$. At this point B is supposed to accept the offer x, and hence gets a continuation payoff of $1 - x > 0$. If B rejects the offer his continuation payoff is zero. Therefore this is not a profitable deviation for B. Suppose now that A has made an offer $x > x^n$, which B is supposed to reject. If B rejects, his continuation payoff is $\delta_B(1 - x^{n+1} - c_B)$. If B accepts, his continuation payoff is $1 - x < 1 - x^n$. But, using (A.6), we know that $\delta_B(1 - x^{n+1} - c_B) \geq 1 - x^n$. It follows that accepting the offer x is not a profitable deviation for B.

Therefore, no player has a profitable deviation from the behaviour prescribed by (σ_A, σ_B)in any possible subgame. This is clearly enough to prove the claim. ■

Lemma A.4: *Consider the alternating offers bargaining game with participation costs described in Section 2. Let $\{x^n\}_{n=1}^{\infty}$ be a sequence of numbers in $[c_A, 1-c_B]$, satisfying (A.6) and (A.7) as in Lemma A.3. Then, for every n odd, every A subgame has an SPE in which agreement is reached immediately and the agreed share of the pie is x^n, and for every n even, every B subgame has an SPE in which agreement is reached immediately and the agreed share of the pie is x^n.*

Proof: The claim is immediate using the strategies described in the proof of Lemma A.3. ■

Proof of The 'If' Part of Theorem 2: It is enough to notice that if δ_i and c_i for $i \in \{A, B\}$ are such that (2) hold then (A.6) and (A.7) must hold when we set $x^n = 1-c_B$ for all odd n and $x^n = c_A$ for all even n. Therefore the game has an SPE with immediate agreement as in Lemma A.4. This is enough to prove the claim. ■

Proof of The 'If' Part of Theorem 3: Fix any x_A as in (3). Notice next that for such x_A, if we choose $x^1 = x_A$, $x^n = c_A$ for all even n, and $x^n = 1 - c_B$ for all odd $n \geq 3$, we have a sequence $\{x^n\}_{n=1}^{\infty}$ which satisfies (A.6) and (A.7) of Lemma A.3. By Lemma A.4, this is enough to prove the claim for the A subgames.

Symmetrically, now fix any x_B as in (4). Notice next that for such x_B, if we choose $x^n = 1-c_B$ for all odd n, $x^2 = x_B$ and $x^n = c_A$ for all even $n \geq 4$, we have a sequence $\{x^n\}_{n=1}^{\infty}$ which satisfies (A.6) and (A.7) of Lemma A.3. By Lemma A.4, this is enough to prove the claim for the B subgames. ■

Proof of The 'Only If' Part of Theorem 3: If an SPE for an A subgame (a B subgame) were to exist, with immediate agreement on a share x_A (a share x_B) outside the interval (3) (outside the interval (4)), we would have an immediate contradiction of A.5. ■

Proof of Theorem 4: We concentrate on the claim for the payoffs of the A subgames. The argument for the B subgames is symmetric and therefore the details are omitted.

Let any x_A as in (3) and any n odd be given. We now display a pair of strategies (σ_A^A, σ_B^A), which constitute an SPE of the A subgames, and which gives the players payoffs as in (5)

Up to and including period $n-1$ the strategies (σ_A^A, σ_B^A) are exactly the same as the strategies (σ_A^0, σ_B^0) of the proof of Theorem 1.

If any deviation from the prescribed outcome path is observed in any period $1, \ldots, n-1$, then the strategies (σ_A^A, σ_B^A) are again the same as the strategies (σ_A^0, σ_B^0) for the remainder of the game.

If no deviation from the prescribed outcome path is observed up to and including period $n-1$ then the strategies (σ_A^A, σ_B^A) from stage I of period n are the same as the strategies (σ_A, σ_B) of the proof of the 'if' part of Theorem 3. Thus, the strategies in this subgame are SPE by construction, and yield an agreement of x_A in period n.

Next, let x_B as in (3) and any n even be given. As before, we display a pair of strategies (σ_A^A, σ_B^A), which constitute an SPE of the A subgames, and which gives the players payoffs as in (6)

Up to and including period $n-1$ the strategies (σ_A^A, σ_B^A) are exactly the same as the strategies (σ_A^0, σ_B^0) of the proof of Theorem 1.

If any deviation from the prescribed outcome path is observed in any period $1, \ldots, n-1$, then the strategies (σ_A^A, σ_B^A) are again the same as the strategies (σ_A^0, σ_B^0) for the remainder of the game.

If no deviation from the prescribed outcome path is observed up to and including period $n-1$ then the strategies (σ_A^A, σ_B^A) from stage I of period n are the same as the strategies (σ_A, σ_B) of the proof of the 'if' part of Theorem 3, starting in period 2. Thus, the strategies in this subgame are SPE by construction, and yield an agreement of x_B in period n. This is clearly enough to prove our claim. ■

Proof of Remark 1: Let (σ_A^N, σ_B^N) be an SPE of Γ^N. We concentrate on the case in which N is odd. The details for the case of N even are symmetric and hence they are omitted. We start by showing that (σ_A^N, σ_B^N) must prescribe that in stage I of period N neither player pays his participation cost, and therefore that the continuation payoffs to both players from the beginning of period N must be both 0.

Consider stage II of period N. By subgame perfection it is clear that A must make an offer $x = 1$ to B at this stage. This because if B rejects A's offer at this stage he earns a continuation payoff of zero, and hence his strategy must be to accept any $x > 0$. Therefore B's continuation payoff in stage II of period N must be zero. It follows that if B pays c_B in stage I of period N his continuation payoff is $-c_B$. Clearly if he does not pay c_B at this stage he will earns a continuation payoff of zero. Therefore, (σ_A^N, σ_B^N) must prescribe that B does not pay his participation cost in stage I of period N, and hence that A doe not pay his cost either.

Once we know that the continuation payoffs for both players starting in stage I of period N are both zero we can move to stage I of period $N-1$. Repeating the argument in the previous paragraph, with the players roles exchanged, is now enough to show that (σ_A^N, σ_B^N) must prescribe that neither player pays his participation cost in stage I of period $N-1$.

Continuing backwards up to stage I of period 1 is now enough to prove the claim. ■

Proof of Remark 2: Given any sequence $\mathcal{S}$, it is immediate to check that the strategies (σ_A^0, σ_B^0) of the proof of Theorem 1 constitute an SPE of $\Gamma(\mathcal{S})$. Since these strategies induce the required outcome path, this is enough to prove the claim. ■

Proof of Remark 3: Simply notice that the proof of Theorem 2 is in fact valid for the case $\delta_A = \delta_B = 1$. Since (2) reduces to (7) when $\delta_A = \delta_B = 1$, this is clearly enough to prove the claim. ■

Proof of Theorem 5: Suppose, by way of contradiction, that the set of CPESPE is not empty. Notice every CPESPE must yield an agreement in every subgame, whenever this is reached. Otherwise, since the players discount the future at a positive rate, the outcome could not possibly be Pareto efficient in every possible subgame.

Let x_i^H and x_i^L for $i \in \{A, B\}$ be the supremum and the infimum respectively of the possible agreements in A periods and in B periods, taken over the set of all possible CPESPE.

The next few steps in the proof parallel closely the proof of the main result in Shaked and Sutton (1984).

Start with an A subgame. Since in stage III of such subgame B accepts any offer x below $\delta_B(1 - x_B^L - c_B)$, using subgame perfection we must have that

$$1 - x_A^L \leq \delta_B(1 - x_B^L - c_B) \tag{A.8}$$

Moreover, since in stage III of any A subgame B rejects any x such that $1 - x < \delta_B(1 - x_B^H - c_B)$ we must have that

$$1 - x_A^H \geq \delta_B(1 - x_B^H - c_B) \tag{A.9}$$

Using a symmetric argument for the B subgames we find that

$$x_B^H \leq \delta_A(x_A^H - c_A) \tag{A.10}$$

and

$$x_B^L \geq \delta_A(x_A^L - c_A) \tag{A.11}$$

Substituting (A.8) into (A.11) we now find that

$$x_B^L \geq \frac{\delta_A[1 - \delta_B(1 - c_B) - c_A]}{1 - \delta_A\delta_B} \tag{A.12}$$

Substituting (A.9) into (A.10) we also obtain that

$$x_B^H \leq \frac{\delta_A[1 - \delta_B(1 - c_B) - c_A]}{1 - \delta_A\delta_B} \tag{A.13}$$

so that clearly we must have

$$x_B = x_B^H = x_B^L = \frac{\delta_A[1 - \delta_B(1 - c_B) - c_A]}{1 - \delta_A \delta_B} \tag{A.14}$$

Symmetrically, substituting (A.10) into (A.9) and then (A.11) into (A.8) we also find out that

$$x_A = x_A^H = x_A^L = \frac{1 - \delta_B + \delta_A c_A}{1 - \delta_A \delta_B} \tag{A.15}$$

Finally, notice that (A.14) and (A.15) together imply that

$$x_B = \delta_A(x_A - c_A) \tag{A.16}$$

and

$$1 - x_A = \delta_B(1 - x_B - c_B) \tag{A.17}$$

Recall now that since an agreement must be reached in every subgame, it must be the case that both players pay their participation costs in stage I of every period. Consider now stage I of any A period. If player B pays his participation cost he gets a continuation payoff of

$$1 - x_A - c_B \tag{A.18}$$

while if B deviates and does not pays his participation cost he gets a continuation payoff equal to

$$\delta_B(1 - x_B - c_B) \tag{A.19}$$

but, using (A.17), it is immediate that the quantity in (A.19) exceeds the quantity in (A.18). Therefore B finds it profitable to deviate and not pay his participation cost in stage I of every A subgame.

Symmetrically, we can verify that in stage I of every B subgame, A will find it profitable to deviate and not pay his participation cost. This is because (A.16) implies that

$$x_B - c_A < \delta_A(x_A - c_A) \tag{A.20}$$

Therefore, we have concluded that in every CPESPE, both players would have an incentive to deviate from their equilibrium behaviour. This contradiction is clearly enough to prove the claim that the set of CPESPE is empty. ■

Proof of Theorem 6: Fix a pair of costs c_A and c_B Next, suppose, by way of contradiction, that for every $p \in (0, 1)$ there exist a PBE of the alternating offers bargaining game with participation costs and imperfect recall in which the parties reach an agreement for at least one of the realizations of the sequence of moves of Nature σ_N.

Let $n(\sigma_N)$ be the period in which this agreement is reached. We start by considering the case in which this period is even. The details for n odd are in fact symmetric. For the remainder of the proof, we denote $n(\sigma_N)$ simply by n for ease of notation. Moreover, all notation pertaining to the players information sets will be suppressed since the actual information set reached at the beginning of period n plays no role in our argument.

By our contradiction hypothesis, in period n both players pay their costs in stage I, B makes an offer x_B to A in stage II, and A accepts this offer in stage III.

Recall that the equilibrium beliefs of both players are consistent with equilibrium strategies and with Bayes' rule in every PBE of this game. Therefore, as we have argued in Section 6.2, at n, the players's beliefs about Nature future moves are simply given by μ_n.

For x_B to be an equilibrium offer it needs to be optimal for B to make such an offer. In other words, it must not be possible for B to make a lower offer $x < x_B$ that A accepts in stage III of period n. This implies that for any offer $x < x_B$, A must be at least as well off by rejecting x than by accepting it. This is the same as saying that the expected continuation payoff to A, $\Pi_A^E(x)$, if he rejects the offer must be at least as high as x. Therefore

$$x \leq \Pi_A^E(x) \qquad\qquad \forall x < x_B \tag{A.21}$$

which trivially implies that

$$x_B \leq \Pi_A^E = \sup_{0 \leq x \leq x_B} \Pi_A^E(x) \tag{A.22}$$

The term Π_A^E can bounded above focusing on whether d^n is equal to $\mathcal{F}$ or $\mathcal{R}$. With probability $(1-p)$ (corresponding to $d^n = \mathcal{R}$) A's continuation payoff is *at most* $\delta_A(1-c_A)$. This is because agreement can be reached at the earliest in period $n+1$, and A must pay his participation cost in period $n+1$ for this to be the case.[16]

With probability p, A's continuation payoff after he rejects in stage III of period n is what he obtains after the players forget the history of play (corresponding to $d^n = \mathcal{F}$). Let A's continuation payoff in this case be denoted by $\Pi_A^E(\mathcal{F})$. Therefore we can now write

$$\Pi_A^E \leq p\Pi_A^E(\mathcal{F}) + (1-p)\delta_A\,(1-c_A) \tag{A.23}$$

[16]In fact it can be shown that this upper bound cannot be achieved since A's equilibrium share of the pie can never be equal to 1.

and, using (A.22), we get

$$x_B \leq p\Pi_A^E(\mathcal{F}) + (1-p)\delta_A\,(1-c_A) \tag{A.24}$$

By our contradiction hypothesis that x_B is agreed at n, it must also be the case that it is optimal for A to pay the cost in stage I of period n. This implies that the equilibrium share x_B less the cost c_A needs to be higher than the expected continuation payoff to A if he does not pay his cost. Let this continuation payoff be denoted by $\hat{\Pi}_A^E$. We therefore have that

$$x_B - c_A \geq \hat{\Pi}_A^E \tag{A.25}$$

With probability $(1-p)$ (corresponding to $d^n = \mathcal{R}$) the continuation payoff to A after he does not pay his cost in stage I of period n is the payoff he gets if both players recall the history of the game at the end of period n. Clearly, this payoff must be at least zero.

With probability p the players forget the history of play (corresponding to $d^n = \mathcal{F}$). In this case the outcome path starting in stage I of period $n+1$ is independent of what happens during period n. In other words, in this case the continuation payoff to A if he does not pay his participation cost must be precisely $\Pi_A^E(\mathcal{F})$ as defined above. We can now conclude that

$$\hat{\Pi}_A^E \geq p\Pi_A^E(\mathcal{F}) \tag{A.26}$$

and therefore, using (A.25), we now have that

$$x_B \geq p\Pi_A^E(\mathcal{F}) + c_A \tag{A.27}$$

Putting together (A.24) and (A.27) yields

$$p\Pi_A^E(\mathcal{F}) + c_A \leq x_B \leq p\Pi_A^E(\mathcal{F}) + (1-p)\delta_A\,(1-c_A) \tag{A.28}$$

which trivially implies that it must be the case that

$$c_A \leq (1-p)(1-c_A) \tag{A.29}$$

Notice now that (A.29) is a contradiction unless

$$p \leq \frac{1-2c_A}{1-c_A} \tag{A.30}$$

Using a completely symmetric argument (the details are therefore omitted), it is possible to show that an agreement in any odd period n yields a contradiction unless

$$p \leq \frac{1 - 2c_B}{1 - c_B} \tag{A.31}$$

Let now

$$\bar{p} = \max\left\{\frac{1 - 2c_A}{1 - c_A}, \frac{1 - 2c_B}{1 - c_B}\right\} \tag{A.32}$$

and notice that since $c_A \in (0,1)$ and $c_B \in (0,1)$ we have that $\bar{p} < 1$. Since any agreement for any $p > \bar{p}$ yields a contradiction, this clearly enough to prove the claim. ■

References

Abreu, D., D. Pearce, and E. Stacchetti (1993): "Renegotiation and Symmetry in Repeated Games," *Journal of Economic Theory*, 60, 217–40.

Admati, A., and M. Perry (1987): "Strategic Delays in Bargaining," *Review of Economic Studies*, 54, 343–64.

Aghion, P., M. Dewatripont, and P. Rey (1994): "Renegotiation Design with Unverifiable Information," *Econometrica*, 62, 257–82.

Anderlini, L., and L. Felli (1994): "Incomplete Written Contracts: Undescribable Sates of Nature," *Quarterly Journal of Economics*, 109, 1085–1124.

——— (1997): "Costly Coasian Contracts," C.A.R.E.S.S. Working Paper 97-11, University of Pennsylvania.

Ausubel, L., and R. Deneckere (1989): "Reputation in Bargaining and Durable Goods Monopoly," *Econometrica*, 57, 511–32.

Benoît, J.-P., and V. Krishna (1993): "Renegotiation in Finitely Repeated Games," *Econometrica*, 61, 303–323.

Bernheim, B. D., and D. Ray (1989): "Collective Dynamic Consistency in Repeated Games," *Games and Economic Behavior*, 1, 295–326.

Busch, L.-A., and Q. Wen (1995): "Perfect Equilibria in Negotiation Model," *Econometrica*, 63, 545–65.

Chatterjee, K., and H. Sabourian (1997): "Multiperson Bargaining and Strategic Complexity," Economic Theory Discussion Paper 9733, Department of Applied Economics, University of Cambridge.

Chatterjee, K., and L. Samuelson (1987): "Bargaining with Two-Sided Incomplete Information: An Infinite Horizon Model with Alternating Offers," *Review of Economic Studies*, 54, 175–92.

Coase, R. H. (1960): "The Problem of Social Cost," *The Journal of Law and Economics*, 3, 1–44.

——— (1992): "The Institutional Structure of Production," *American Economic Review*, 82, 713–19.

Cramton, P. (1992): "Strategic Delay in Bargaining with Two-Sided Uncertainty," *Review of Economic Studies*, 59, 205–25.

Dewatripont, M. (1989): "Renegotiation and Information Revelation over Time: the Case of Optimal Labor Contracts," *Quarterly Journal of Economics*, 104, 589–619.

Dixit, A., and M. Olson (1997): "Does Voluntary Participation Undermine the Coase Theorem?," Princeton University, mimeo.

Farrell, J., and E. Maskin (1987): "Renegotiation in Repeated Games," Discussion Paper 1335, Harvard Institute of Economic Research, Harvard University.

——— (1989): "Renegotiation in Repeated Games," *Games and Economic Behavior*, 1, 327–360.

Fernandez, R., and J. Glazer (1991): "Striking for a Bargain Between Two Completely Informed Agents," *American Economic Revie*, 81, 240–252.

Fudenberg, D., D. Levine, and J. Tirole (1987): "Incomplete Information Bargaining with Outside Opportunities," *Quarterly Journal of Economics*, 102, 37–50.

Fudenberg, D., and J. Tirole (1991): "Perfect Bayesian Equilibrium and Sequential Equilibrium," *Journal of Economic Theory*, 53, 236–260.

Games and Economic Behavior (1997): *Special Issue on Imperfect Recall*, vol. 20, no. 1. New-York: Academic Press.

Grossman, S. J., and O. D. Hart (1986): "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, 94, 691–719.

Grossman, S. J., and M. Perry (1986): "Sequential Bargaining under Asymmetric Information," *Journal of Economic Theory*, 39, 120–54.

Gul, F., H. Sonnenschein, and R. Wilson (1986): "Foundations of Dynamic Monopoly and the Coase Conjecture," *Journal of Economic Theory*, 39, 155–90.

Hart, O. D. (1989): "Bargaining and Strikes," *Quarterly Journal of Economics*, 104, 25–43.

Hart, O. D., and J. Moore (1988): "Incomplete Contracts and Renegotiation," *Econometrica*, 56, 755–85.

——— (1998): "Foundations of Incomplete Contracts," *Review of Economic Studies*, forthcoming.

MacLeod, W. B. (1997): "Complexity, Contracts and the Employment Relationship," University of Southern California, mimeo.

Maskin, E., and J. Moore (1998): "Implementation and Renegotiation," *Review of Economic Studies*, forthcoming.

Maskin, E., and J. Tirole (1998): "Unforseen Contingencies, Property Rights, and Incomplete Contracts," *Review of Economic Studies*, forthcoming.

Moore, J. (1992): "Implementation, Contracts, and Renegotiation in Environments with Complete Information," in *Advances in Economic Theory: Sixth World Congress*, ed. by J.-J. Laffont, vol. I. Cambridge: Cambridge University Press.

Nicholson, W. (1989): *Microeconomic Theory: Basic Principles and Extensions.* Chicago: The Dryden Press. Fourth International Edition.

Piccione, M., and A. Rubinstein (1997): "On the Interpretation of Decision Problems with Imperfect Recall," *Games and Economic Behavior*, 20, 3–24.

Riedl, A. (1997): "Bargaining and Opting In Costs," Institute for Advanced Studies, Department of Economics, Vienna, mimeo.

Rubinstein, A. (1982): "Perfect Equilibrium in a Bargaining Model," *Econometrica*, 50, 97–110.

——— (1985): "A Bargaining Model with Incomplete Information about Time Preferences," *Econometrica*, 53, 1151–72.

Rubinstein, A., and A. Wolinsky (1992): "Renegotiation-Proof Implementation and Time Preferences," *American Economics Review*, 82, 600–614.

Segal, I. (1998): "Complexity and Renegotiation: A Foundation for Incomplete Contracts," *Review of Economic Studies*, forthcoming.

Shaked, A., and J. Sutton (1984): "Involuntary Unemployment as a Perfect Equilibrium in a Bargaining Game," *Econometrica*, 52, 1351–1364.

Sobel, J., and I. Takahashi (1983): "A Multistage Model of Bargaining," *Review of Economic Studies*, 50, 411–26.