

# GREQAM

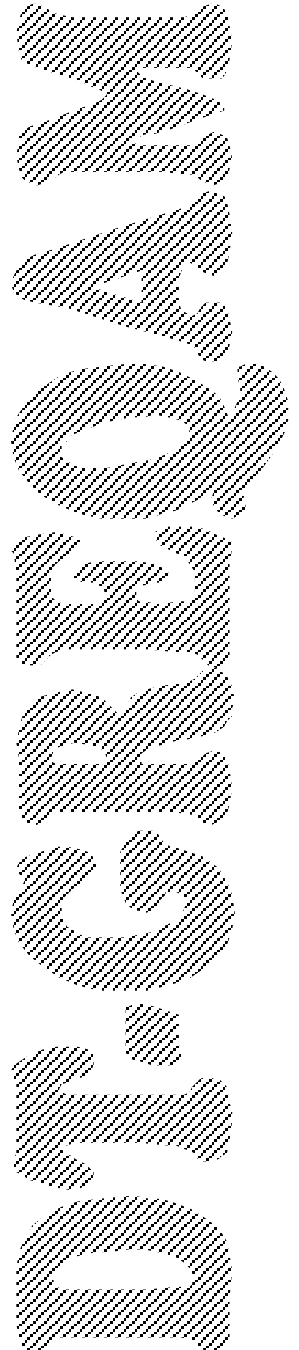
Groupement de Recherche en Economie  
Quantitative d'Aix-Marseille - UMR-CNRS 6579  
Ecole des Hautes Etudes en Sciences Sociales  
Universités d'Aix-Marseille II et III

**Document de Travail**  
**n°2008-31**

## **A MECHANISM FOR SOLVING BARGAINING PROBLEMS BETWEEN RISK AVERSE PARTIES**

**Emily TANIMURA**  
**Sylvie THORON**

**Juillet 2008**



# A mechanism for solving bargaining problems between risk averse players

Emily TANIMURA

*GREQAM*

*e-mail:* emily.tanimura@ehess.fr

Sylvie THORON \*

*GREQAM*

*e-mail:* thoron@univmed.fr

**Abstract:** We propose a mechanism for resolving bargaining problems. The mechanism allows two players to make a sequence of simultaneous propositions. At any step, as long as the players have not reached an agreement, they can choose to implement a lottery between the different propositions. In this aspect, the mechanism is similar to the so called final offers arbitration. However, contrary to the existing scheme, our mechanism is not compulsory. The history of the negotiation process is recorded and players can refuse an offer and go back in the process to a previous step. This generates an evolving sequence of status-quo points and results in a sequence of equilibrium offers of the two players that gradually converge towards each other. Our model assumes no discounting and complete information. Rather than time preferences, the main incentive to reach an agreement under our mechanism comes from risk aversion. Players have an incentive to avoid the uncertainty related to the lotteries that occur when offers do not result in an agreement. Rather than incomplete information, the process gradualism is driven by the necessity to make step by step concessions in order to generate evolving threat points. We show that under this mechanism, the unique subgame perfect equilibrium that does not use weakly dominated strategies coincides with a well-known static solution concept, the Raiffa solution.

**Keywords and phrases:** bargaining theory, Raiffa bargaining solution, risk aversion, final offers arbitration, chilling effect, gradualism.

## 1. INTRODUCTION

This paper proposes a mechanism to resolve bargaining disputes. This mechanism is not compulsory and leaves the bargaining parties with complete freedom of choice whilst offering them a framework for reaching an efficient

---

\*Corresponding author

agreement. It is a contribution to two different branches of literature. The first one concerns compulsory arbitration schemes. These interest practitioners as much as researchers since there are actually used to resolve real bargaining disputes. For example they are used to settle grievances in union-management contracts or between buyers and sellers in commercial contracts (for a variety of examples see Ashenfelter and Bloom (1984)). There are essentially two categories of arbitration mechanisms. In conventional compulsory arbitration (CCA), the arbitrator has the power to impose a settlement of his choice on the bargaining agents if their negotiations break down, that is if their demands are not compatible. It has been suggested that this mechanism produces a *chilling effect*: it erodes agents' incentive to make concessions in bargaining and their demands remain distant from each other. Stevens (1966) proposed a simple procedure, the final offer arbitration (FOA) designed to counteract this effect: the arbitrator chooses among the proposals the one she wants to impose. His idea was that: "... [FOA] generates just the kind of uncertainty about the location of the arbitration awards that is well calculated to recommend maximum notions of prudence to the parties and, hence, compel them to seek security in agreement" page 46. As a result of this theoretical proposal, this procedure has been applied to settle public sector labor disputes in several U.S. state jurisdictions (see Hebdon 1996). It turned out that this mechanism had to be implemented, meaning that the bargaining parties did not reach an agreement by themselves, contrary to Stevens' argument. Indeed, several models in the literature show that Stevens' intuition was wrong from two points of view.

First, Stevens thought that one quality of FOA was that it guarantees that the bargaining parties negotiate in good-faith. The simple threat of arbitration induces agents to reach agreement on their own. Crawford (1979) showed that this is not true. He assumes that both parties know with certainty the arbitrator's preferred outcome and concludes under this assumption that both arbitration mechanisms CCA and FOA support in fact the same outcome at the Nash equilibrium. Proposals are not compatible at the equilibrium of the conventional compulsory arbitration but are compatible with FOA. However, "...there exists a unique Nash equilibrium in FOA, which leads to the feasible final settlement considered most reasonable by the arbitrator, without regard to the bargaining agents' preferences". Page 135. Under the threat of arbitration, the parties are not actually involved in a real process of bargaining but are just trying to conform to the arbitrator's preferred outcome. A first feature of our mechanism is that arbitration consists of a pure lottery which chooses between the two parties' proposals with equal probability. This eliminates Crawford's objection to the FOA.

Following Stevens' argument, the second advantage of the FOA is that it eliminates the chilling effect. In Farber (1980, 1981) and Farber and Katz (1979), the parties are uncertain as to the arbitrator's preference. In this framework they show that the two different mechanisms lead to different results which depend on the parties' attitude toward risk. However, the so called chilling effect persists. More recently, the comparison of the two kinds of mechanism has been the subject of experiments (Dickinson (2006), Kriticos (2006)) which reinforce this last theoretical result. In our framework the chilling effect is eliminated. At the subgame perfect Nash equilibrium the two parties' proposals converge towards an efficient consensus.

One criticism which can be addressed to all of this literature is that the negotiation *process* is not described. The characterization of Nash equilibria in a static framework in order to describe a strategic negotiation may not capture important aspects of the process itself. These dynamic aspects may nevertheless be relevant for the outcome. In our model, the bargaining parties can make sequential proposals and can have recourse to the mechanism at each step.

Now, we turn to the other branch of literature to which our mechanism is related, the literature on bargaining theory. There are two types of approaches to bargaining theory, one of which is the axiomatic theory of bargaining. In this case, a bargaining problem between two parties has a very simple framework. It specifies a set of possible outcomes and a threat point which describes what the two players get in case of negotiation failure. The various solution concepts such as the best known Nash bargaining solution (1950) but also the Raiffa solution (1953) or the Kalai-Smorodinsky solution (1975) are all efficient. The other approach, which follows the Nash program, is then to implement the previous solutions in a strategic and sequential framework. Note that, when this is possible it is a justification of Coase's idea that negotiation must lead to an efficient outcome. A classical extensive form bargaining game is the Rubinstein alternating offers model with infinite horizon and discount factor (Rubinstein 1982). The unique subgame perfect equilibrium of this game coincides with the Nash bargaining solution prescribed by the axiomatic theory. The discount factor is essential in the Rubinstein model; players have an incentive to reach agreement because the "cake" to be shared shrinks over time. In Stahl's (1972) alternating offers model the horizon is finite with full disagreement. Sjöström (1991) uses Stahl's model and shows that the Raiffa solution is its unique subgame perfect Nash equilibrium. The main criticism addressed to these models is that at the equilibrium the bargaining parties reach the agreement immediately. The first proposal is accepted. This lack of delay in a sequential bargaining

game has been considered to be unrealistic. The main objective was then to explain which forces in reality drive the negotiating parties to *gradually* improve their offers to each other.

Indeed, Rubinstein's paper generated a wide literature whose purpose was to find ways to modify the model in order to obtain delay at the equilibrium. A first answer was found with the introduction of incomplete information (see the survey of Kennan and Wilson (1993)). In a recent paper Compte and Jehiel (2004) show that even with complete information, the simple threat of negotiation failure can generate gradualism in bargaining. The bargaining parties have the possibility to opt out of the negotiation process. Opting out is costly and results in a compromise partition which is a function of the previous proposals. This generates gradualism. They say in their discussion that "... [their] finding of gradualism is the dynamic counterpart of the chilling effect identified in [the first] literature". page 997.

In our framework there is complete information and the possibility of opting out as in Compte and Jehiel (2004), but no discounting. Our analysis is focused on the role of the parties' risk aversion in the negotiation which was pointed out by Stevens. In the mechanism we propose, the incentive to reach agreement is not based on the fact that utility is discounted over time but on the fact that incompatible offers lead to uncertainty of outcomes and players who are risk averse may prefer to modify their offers to obtain certain payoffs.

We specify a sequential bargaining game with simultaneous proposals. When the proposals are not compatible the bargaining parties have the possibility of opting out by implementing an arbitration scheme in the spirit of the FOA. A lottery draws with equal probability each party's proposal. However, there are two important differences. The first difference is that the resulting proposal cannot be imposed on the other party who has the right to reject it. The second difference is that the mechanism allows the parties to record the sequence of simultaneous but not compatible proposals. Thanks to this historical list of disagreements there may be several options when a party decides to opt out. Indeed, if a party rejects the proposal which has been drawn by the lottery, the mechanism goes back to implement a lottery with the previous proposals. If all the proposals are rejected the mechanism goes back until it reaches the original threat point. Therefore, the mechanism generates an evolving threat point. At the subgame perfect Nash equilibrium, the simultaneous proposals gradually converge towards the Raiffa solution.

## 2. Definitions, notation and setting of the model

In what follows, we will consider the following setting: two players  $J_1$  and  $J_2$  bargain over the partition of a pie of size  $K = 1$ , without loss of generality. The possible ways to split the pie are

$$\{(x, y) : x, y \in [0, 1] \text{ and } x + y \leq 1\}$$

in which  $x$  is the share of Player  $J_1$  and  $y$  those of Player  $J_2$ . Both players are assumed to be (weakly) risk averse in the sense that they have linear or concave utility functions, respectively  $U_1(x), U_2(y)$  that are strictly increasing on  $[0, 1]$  (this last assumption is for convenience and to exclude some degenerated cases). In terms of possible utility pairs  $(u, v) \in \mathbb{R}^2$ , the compact, convex bargaining set is thus delimited by the curves  $u = U_1(0), v = U_2(0)$  and the curve  $(U_1(x), U_2(1-x)), x \in [0, 1]$ . At the status quo, the players receive  $d_1$  and  $d_2$ . The game is sequential but there is no discounting. At any date  $t$ , the two players have to make simultaneous proposals  $(x_t, y_t)$ . If  $x_t = 1 - y_t$  we will say that an agreement has been concluded. In this case, the game ends. Player 1 gets  $x_t$  and Player 2,  $y_t = 1 - x_t$ . If the two proposals do not result in an agreement, the difference  $D_t =: x_t - (1 - y_t)$  measures the gap between the two proposals or the level of disagreement. Then, the two players have the possibility to implement a fair lottery between the propositions  $x_t$  and  $y_t$  in which each proposal is drawn with probability  $\frac{1}{2}$ . We will denote this lottery between  $x_t$  and  $y_t$  by  $\mathcal{L}_t$ . We will use the convenient notation  $U_1(\mathcal{L}_t) =: \frac{1}{2}U_1(x_t) + \frac{1}{2}U_1(1 - y_t)$ . If and only if both prefer to continue, they go to the following period and make another pair of simultaneous proposals  $(x_{t+1}, y_{t+1})$ . We adopt the convention that the payoffs induced by strategies where an agreement is never reached and where neither player ever accepts to implement a lottery are the threat point payoffs  $U_1(d_1)$  and  $U_2(d_2)$ . Consider a date  $t$  at which no player has decided to implement the lottery. At  $t$ , each player makes a proposition  $x_t$  ( $y_t$ ) and players then announce their decisions to implement the lottery or to proceed to a new proposal. Let us denote by  $h^t$  the  $t$ -history generated by all the propositions up to and including time  $t$ :  $(x_s)_{s \leq t}$  and  $(y_s)_{s \leq t}$ . The propositions  $x_t$  and  $y_t$  are then defined based on  $h^{t-1}$  and the decision to continue or to implement the lottery after the propositions  $x_t$  and  $y_t$  is conditioned on  $h^t$ . The part of the game described above is common to a model that we will first analyze rapidly and which can be seen as a formalization of Stevens FOA mechanism and to the mechanism we then propose. These mechanisms differ in what happens after a player decides to implement at the lottery.

### 2.1. Final offers arbitration without recursive structure

The following simple model can be seen as a formalization of Stevens FOA. Let  $T$  be the date when the lottery is implemented. If  $x_T$  ( $y_T$ ) is drawn and accepted, the game ends, and Player 1 gets  $x_T$  ( $1 - y_T$ ) and Player 2,  $1 - x_T$  ( $y_T$ ). If the proposition drawn by the lottery is rejected, then the game ends and player  $i$  receives  $d_i$ . The sequential game is represented in figure 1. An outcome of this game is a (possibly infinite) end date, denoted  $T$ , sequences  $(x_t)_{t \geq 0}$  and  $(y_t)_{t \geq 0}$  of proposals, the name of the player who decided to implement the lottery  $\mathcal{L}_T$  if the last simultaneous proposals did not result in an agreement:  $x_T \neq 1 - y_T$ , the outcome of the lottery  $\mathcal{L}_T$  and whether  $x_T$  (or  $y_T$ ) was accepted. It is readily observed that this mechanism does not result in agreement. Indeed, at equilibrium the unique pair of simultaneous proposals at each step is  $(1, 1)$ . Suppose it is not the case at the last period of the game or when at least one player decides to opt out. Then, when the game stops a lottery  $\mathcal{L}(x, y)$  is implemented, with  $x < 1$ . However, this cannot be the case at the equilibrium since Player 1 has an incentive to deviate and propose 1. Therefore at the equilibrium, when the game stops the pair of proposals is  $(1, 1)$ . Now consider a step before the game stops, when the two players continue. If the pair of proposals is  $(x, y)$ , with  $y < 1$ , Player 1 has an incentive to deviate: she can decide to propose 1 and to opt out.

### 2.2. The recursive model

In the previous model, rejection of a proposition in the first lottery  $\mathcal{L}_T$  entailed the end of the game and threat point payoffs  $U_i(d_i)$  to both players. Let us now consider a model where rejection of a proposition  $x_T$  (or  $y_T$ ) in  $\mathcal{L}_T$  will instead release the implementation of the lottery  $\mathcal{L}_{T-1}$  between the propositions  $x_{T-1}$  and  $y_{T-1}$  and rejection in this lottery releases the implementation of  $\mathcal{L}_{T-2}$  and so on. This way, if each player always rejects her partner's proposal drawn randomly, they can go back to the first degenerated lottery in which the status quo payoffs are applied. The sequential game can be represented by the extensive form in Figure 2.

Let  $T$  be the first date where at least one player decides to implement the lottery. For every  $1 \leq s \leq T$ , player 1 and 2 must decide whether or not to accept proposition  $y_s$  and  $x_s$  respectively if it is drawn. The decision to accept  $y_s$  (or  $x_s$ ) is a function of  $h^T$  and of the outcomes of the lotteries  $\mathcal{L}_T, \mathcal{L}_{T-1}, \dots, \mathcal{L}_{s+1}$ . This sums up all the relevant information on which the decision to accept or reject  $x_s$  can be conditioned. Indeed, if a player faces a

decision related to  $\mathcal{L}_s$ , it necessarily implies that the propositions drawn in the lotteries  $\mathcal{L}_T, \mathcal{L}_{T-1}, \dots, \mathcal{L}_{s+1}$  have been rejected.

We note that the recursive structure of the lotteries can be represented in its extensive form as a tree (see figure ). If the decision to implement the lottery was taken after  $T$  propositions, the tree has length  $T$ . The first branches from the root lead to two nodes at which  $x_T$  and  $y_T$  respectively are proposed. Each player intervenes at the nodes where his opponent's proposition is drawn and decides to accept or reject it. If there is rejection of  $x_T$  or  $y_T$  at level  $T$ , we follow one of the branches to either  $x_{T-1}$  or  $y_{T-1}$  which will be one level further from the root. Each node has two branches leading to successor nodes. Whenever the proposition at a node is rejected, each branch leading to a successor node at the level below is chosen with probability  $\frac{1}{2}$  leading to a node at the level below (i.e. one step further from the root).

An outcome of this game is a (possibly infinite) end date, denoted  $T$ , sequences  $(x_t)_{t \geq 0}$  and  $(y_t)_{t \geq 0}$  of proposals, the name of the player who decided to implement the lottery  $\mathcal{L}_T$  if the last simultaneous proposals did not result in an agreement:  $x_T \neq 1 - y_T$ , and the outcomes of the lotteries  $\mathcal{L}_T, \mathcal{L}_{T-1}, \dots, \mathcal{L}_s$  where  $\mathcal{L}_s$  is the first lottery in which the drawn proposition was accepted. As we have noted previously, the fact that the lotteries  $\mathcal{L}_T, \dots, \mathcal{L}_s$  were implemented necessarily implies that the player whose decision was not drawn rejected his partner's proposition in all the lotteries  $\mathcal{L}_T, \mathcal{L}_{T-1}, \dots, \mathcal{L}_{s+1}$ .

### 3. Characterizing the equilibria in the recursive model

In this section we will characterize the equilibria of the game. The proof will involve two phases. First we take the propositions  $x = (x_t)_{t \geq 1}$  and  $y = (y_t)_{t \geq 1}$  as given and we consider the extensive form game whose payoffs are determined by these propositions and by the outcomes of the lotteries. We examine the players' strategies concerning which propositions they will accept in each lottery and we show that subgame perfection imposes precise constraints on the propositions that are accepted. Having done so, we can then characterize the equilibrium propositions.

Suppose that some arbitrary sequence of propositions  $(x_t)_{t \geq 1}$  and  $(y_t)_{t \geq 1}$  are given. We define two recursive functions  $V_t^1$  and  $V_t^2$  such that  $V_t^i$  depends on  $(x_s)_{s \leq t}$  and  $(y_s)_{s \leq t}$  and show that player 1 accepts any proposition  $y_t$  such that  $U_1(1 - y_t) \geq V_{t-1}^1$  and player 2 accepts any proposition  $x_t$  such that  $U_2(1 - x_t) \geq V_{t-1}^2$ . The recursive functions  $V_t^1$  and  $V_t^2$  which depend of course on  $x$  and  $y$ , are in fact the expected payoffs that player 1 and 2 can ensure in the extensive form game where lottery  $\mathcal{L}_t$  is about to be implemented,



given that both players' acceptance strategies are subgame perfect. These functions are defined as follows:

$$V_0^1 = U_1(d_1), V_0^2 = U_2(d_2)$$

$$V_t^2 = \frac{1}{2} \left\{ \begin{array}{ll} U_2(y_t) & \text{if } U_1(1 - y_t) \geq V_{t-1}^1 \\ V_{t-1}^2 & \text{otherwise} \end{array} \right\} + \frac{1}{2} \max[U_2(1 - x_t), V_{t-1}^2].$$

$$V_t^1 = \frac{1}{2} \left\{ \begin{array}{ll} U_1(x_t) & \text{if } U_2(1 - x_t) \geq V_{t-1}^2 \\ V_{t-1}^1 & \text{otherwise} \end{array} \right\} + \frac{1}{2} \max[U_1(1 - y_t), V_{t-1}^1].$$

We have the following lemma:

**Lemma 3.1.** *Suppose that propositions  $x$  and  $y$  are given. Subgame perfection implies that player 1 accepts any proposition  $y_t$  such that*

$$U_1(1 - y_t) \geq V_{t-1}^1 \tag{1}$$

and player 2 accepts any proposition  $x_t$  such that

$$U_2(1 - x_t) \geq V_{t-1}^2, \tag{2}$$

where  $V_t^1$  and  $V_t^2$  are defined above. Moreover, if the lottery  $\mathcal{L}_t$  is implemented, player 1 and 2 can ensure an expected (before implementation) payoff of  $V_t^1$  and  $V_t^2$  respectively.

We will refer to propositions that must be accepted whenever players use subgame perfect strategies as *acceptable* propositions. We will now prove the lemma by induction. At  $t = 1$ , player 1 must accept any proposition  $y_1$  such that  $U_1(1 - y_1) \geq V_0^1 = U_1(d_1)$  since she gets  $U_1(d_1)$  if she refuses, the same is true for player 2. It follows that when player 1 is faced with the lottery  $\mathcal{L}_1$ , in the extensive form game, she is ensured of an expected payoff of  $V_1^1$  in the lottery between  $x_1$  and  $y_1$ . Indeed, with probability  $\frac{1}{2}$ ,  $y_1$  is drawn and she accepts this only if  $U_1(1 - y_1) \geq V_0^1$ . With probability  $\frac{1}{2}$ ,  $x_1$  is drawn and it is accepted if  $U_2(1 - x_1) \geq V_0^2$ . Thus player 1 is ensured of an expected utility  $V_1^1$ . Suppose that the statements are true at  $t - 1$ . Suppose that player 1 rejects  $y_t$ . She then faces the lottery  $\mathcal{L}_{t-1}$ . By the induction hypothesis, she is ensured of  $V_{t-1}^1$ . Thus she should accept  $y_t$  if and only if  $U_1(1 - y_t) \geq V_{t-1}^1$ . An analogous statement is true for player 2. Having shown this, we can show that player 1 is ensured of an expected utility  $V_t^1$

when she faces the lottery  $\mathcal{L}_t$  between  $x_t$  and  $y_t$ : With probability  $\frac{1}{2}$ ,  $x_t$  is drawn. It is accepted if  $U_2(1 - x_t) \geq V_t^2$ . If it is refused, player 1 faces the lottery  $\mathcal{L}_{t-1}$  where she is ensured of  $V_{t-1}^1$  by the induction hypothesis. With probability  $\frac{1}{2}$ ,  $y_t$  is drawn. Player 1 accepts if  $U_1(1 - y_t) \geq V_{t-1}^1$  otherwise she is again ensured of  $V_{t-1}^1$  by the induction hypothesis. Consequently, in the lottery  $\mathcal{L}_t$  she is ensured of

$$V_t^1 = \frac{1}{2} \left\{ \begin{array}{ll} U_1(x_t) & \text{if } U_2(1 - x_t) \geq V_t^2 \\ V_{t-1}^1 & \text{otherwise} \end{array} \right\} + \frac{1}{2} \max[U_1(1 - y_t), V_{t-1}^1].$$

This concludes the proof of the lemma.

Having shown this, it is obvious that the strategies must be such that they induce a common end date for the bargaining. If one player, say player 1 uses a strategy such that she ends the proposition phase at  $t$  while the other player's strategy is to continue beyond this date, then player 1 would benefit from deviating to a strategy where she proposes an  $x_{t+1}$  such that  $U_1(x_{t+1}) > V_t^1$  and  $U_2(1 - x_{t+1}) \geq V_t^2$ . Lemma 3.2 below shows that if at least one player is risk averse such an  $x_{t+1}$  always exists. Consequently, the strategies in every equilibrium must induce a common date  $T$  at which both players cease to make propositions and implement the lottery. We note that if the end date  $T$  was instead a pre-determined horizon after which no further propositions could be made, this would not affect the argumentation in what follows. We can use backward induction, starting from the end date  $T$  to determine the players' proposition strategies. As the lemma establishes, the expected utility of player 1 when she faces the lottery  $\mathcal{L}_T$  is  $V_T^1$ . If the lottery is implemented after the propositions  $x_T$  and  $y_T$ , player 1 must choose  $x_T$  to maximize  $V_T^1 = V_T^1(x_T, y_T, V_{T-1}^1, V_{T-1}^2)$ , and similarly for player 2. It is easy to see from the recursive definition of  $V_t^i$  that for every  $t$ ,  $V_t^1$  is an increasing function of  $V_{t-1}^1$  and a decreasing function of  $V_{t-1}^2$  and conversely for player 2. At each  $t$ , player 1 must thus choose  $x_t$  to maximize  $V_t^1$ . Let us show now that if players have strictly increasing utility functions and at least one player is risk averse, then the maximizing  $x_t$  and  $y_t$  respectively are unique and are such that  $V_t^1$  (or  $V_t^2$ ) is strictly greater than  $V_{t-1}^1$  (or  $V_{t-1}^2$ ).

**Lemma 3.2.** *If at least one player is risk averse in the sense that his utility function is strictly concave on  $[0, 1]$  and if both players utility functions are strictly increasing on  $[0, 1]$ , then there exists, for any history  $h^t \in H^t$ , a unique  $x_{t+1}$  that maximizes  $V_{t+1}^1$  and for which  $V_{t+1}^1 > V_t^1$  and there exists a unique  $y_{t+1}$  that maximizes  $V_{t+1}^2$  and for which  $V_{t+1}^2 > V_t^2$ .*

To show this, we note that at any date  $t > 1$ , we can write

$$\begin{aligned} V_t^1(x, y) &= \sum_{i=1}^{i=t} p_i U_1(x_i) + q_i U_1(1 - y_i) \\ V_t^2(x, y) &= \sum_{i=1}^{i=t} p_i U_2(1 - x_i) + q_i U_2(y_i) \\ &\text{with } \sum_{i=1}^{i=t} p_i + q_i = 1. \end{aligned}$$

The probabilities  $q_i$  and  $p_i$  may be zero if proposition  $x_i$  ( $y_i$ ) is rejected, or if both  $x_j$  and  $y_j$  were acceptable for some  $j > i$ . If we posit  $\tilde{x} = \sum_{i=1}^{i=t} p_i x_i + q_i(1 - y_i)$  we have

$$\begin{aligned} V_t^1(x, y) &\leq U_1\left[\sum_{i=1}^{i=t} (p_i x_i + q_i(1 - y_i))\right] = U_1(\tilde{x}) \\ V_t^2(x, y) &\leq U_2\left[\sum_{i=1}^{i=t} (p_i(1 - x_i) + q_i y_i)\right] = U_2(1 - \tilde{x}). \end{aligned}$$

Let us first assume that both players are risk neutral. In this case, the first inequalities are equalities. If both players have strictly increasing utility functions, there is no proposition that can make a player strictly better off and that is acceptable to his opponent. Indeed, if  $U_1(x) > V_t^1$ , necessarily  $x > \tilde{x}$  but then  $U_2(1 - x) < V_t^2$  and player 2 would not accept  $x$ . In fact,  $(V_t^1, V_t^2)_{t \geq 0}$  will be stationary for  $t \geq 1$ . Thus, when both players are risk neutral, there is indeterminacy since players will be indifferent between several alternatives: making a new proposition that is acceptable but not strictly utility improving or making a proposition that will be rejected. Player  $i$  may thus keep proposing  $1 - d_i$  for himself for an indeterminate number of periods and then implement a lottery between the incompatible propositions or players may reach agreement on the propositions  $x_2 = \frac{1+d_1-d_2}{2}$ ,  $y_2 = \frac{1+d_2-d_1}{2}$ . If, on the other hand, at least one player is risk averse, player  $i$  can always find a proposition (for example  $\tilde{x}$  and  $1 - \tilde{x}$  respectively) that makes him strictly better off and the other player no worse off than before. Moreover, if  $U_1$  and  $U_2$  are strictly increasing, then for every  $t$ , agent  $i$  has a unique proposition that maximizes  $V_t^i$ . It is the proposition that leaves the other player exactly his certainty equivalent of the lottery  $\mathcal{L}_{t-1}$ , that is the utility  $U_i(\mathcal{L}_{t-1})$ . In equilibrium, the propositions are thus uniquely defined by the following recursion:

$$\begin{aligned}
x_1 &= 1 - d_2, y_1 = 1 - d_1 \\
x_t &= \{x | U_2(1 - x) = U_2(\mathcal{L}_{t-1})\} \\
y_t &= \{y | U_1(1 - y) = U_1(\mathcal{L}_{t-1})\}.
\end{aligned} \tag{3}$$

Having established this, let us summarize the properties of the equilibrium strategies in the following propositions

**Proposition 1.** *In the recursive model, under the given assumptions ( $U_1, U_2$  strictly increasing, at least one risk averse player), the only subgame perfect equilibrium strategies are such that*

- *Conditional on the game history, both players stop making propositions and implement the lottery at a common date  $T = 1, 2, \dots$ .*
- *Before this date  $T$ , the proposition strategies are:  $x_1 = 1 - d_2, y_1 = 1 - d_1$  and for  $t \geq 1$  each player proposes the maximal  $x_t$  ( $y_t$ ) that verifies  $U_2(1 - x_t) \geq V_{t-1}^2$  ( $U_1(1 - y_t) \geq V_{t-1}^1$ ).*
- *In the lottery  $\mathcal{L}_t$ , player 1 accepts the proposition  $y_t$  if and only if  $U_1(1 - y_t) \geq V_{t-1}^1$ , and similarly for player 2.*

For every  $T \geq 1$ , there is thus a unique subgame perfect equilibrium where the lottery is implemented after  $T$  propositions. In each one of these, the strategies for the propositions that are made before  $T$  and the decisions about whether to accept them are identical and have a simple structure. Each player initially proposes  $1 - d_j$  for himself. At  $1 \leq t \leq T$  each player then proposes the  $x_t$  and the  $y_t$  defined by the recursive expression 3. In other words, they demand as much as they can, given that they need to leave to the other player the utility that corresponds to her certainty equivalent of the lottery between the previous propositions,  $x_{t-1}$  and  $y_{t-1}$ . In the following corollary, we compare the equilibria as a function of the date  $T$  where the lottery was implemented.

**Corollary 1.** *(Comparison of the equilibria) For every  $T > 0$ , there is a unique subgame perfect equilibrium where the lottery  $\mathcal{L}_T$  between  $x_T$  and  $y_T$  is implemented and the drawn propositions accepted. Comparing the equilibria as a function of  $T$  we have:*

- *$D_T = x_T - (1 - y_T)$  is decreasing in  $T$ .*
- *The expected equilibrium payoffs of the players,  $\frac{U_1(x_T) + U_1(1 - y_T)}{2}$  and  $\frac{U_2(y_T) + U_2(1 - x_T)}{2}$  respectively are increasing in  $T$ .*
- *For any  $T \geq 0$ , the players' strategies in the equilibrium where the players make exactly  $T$  propositions are weakly dominated by their strategies*

in the equilibrium where they make exactly  $T + K$  propositions, for any  $K > 0$ .

Note that the sequences of propositions  $(x_t)_{t \geq 0}, (y_t)_{t \geq 0}$  defined recursively by 3 are in fact the sequences defining the Raiffa bargaining process, so that  $\lim_{t \rightarrow \infty} x_t = \lim_{t \rightarrow \infty} (1 - y_t)$  and  $\lim_{t \rightarrow \infty} x_t$  and  $\lim_{t \rightarrow \infty} y_t$  are the shares of the pie received by the players in the Raiffa solution (see Raiffa 1952). Figure 3 shows a geometric representation of the Raiffa solution. In our bargaining mechanism, some of the equilibria "truncate" the process before it converges to the Raiffa solution. However, these equilibria are dominated by the ones where the process continues. In the following corollaries we give some properties of the equilibria and state their relation to the Raiffa solution.

**Remark 1.** *The game we have described does not have any predetermined end date. If we imposed such an end date  $T$  in the game, the unique subgame perfect equilibrium would coincide (in terms of propositions and acceptance strategies) with the equilibrium where the players coordinated themselves on the end date  $T$ .*

It follows from Proposition 1 that despite the presence of equilibria where the proposition process ends before payoffs are close to those of the Raiffa solution, the strategies used in these are weakly dominated. We summarize this in what follows.

**Corollary 2.** *For every  $\epsilon > 0$ , there is a  $\hat{T}$  such that for every  $T \geq \hat{T}$ :*

$$\left\| \left( \frac{U_1(x_T) + U_1(1 - y_T)}{2}, \frac{U_2(y_T) + U_2(1 - x_T)}{2} \right) - (e_R^1, e_R^2) \right\| \leq \epsilon \quad (4)$$

*Moreover, the strategies used in any equilibrium whose payoffs are not close to the Raiffa solution in the sense of (4) are weakly dominated by the strategies used in any of the equilibria where payoffs verify (4)*

Equilibrium refinement concepts such as iterative dominance solvability would thus select the equilibria that coincide with the Raiffa solution.

#### 4. Conclusion

In this paper we have suggested a mechanism for resolving bargaining conflicts. Without discounting and with complete information two parties with conflicting interests reach an agreement gradually. A first key point to understand this result is that our mechanism exploits the parties' risk aversion.

Like with a final offers arbitration scheme (FOA), our mechanism solves disagreement by means of a lottery between distant proposals. However, contrary to the FOA, it is not compulsory. The history of the negotiation process is recorded and players can refuse an offer drawn randomly and go back in the process to a previous step. We showed that at the equilibrium this is necessary for the players to make step by step concessions and to generate evolving threat points. Compte and Jehiel (2004) also proposed a bargaining game in which at the equilibrium status quo points evolve, which depend on the previous propositions of the players. However, in the setting of Compte and Jehiel, making generous offers would only improve the other players outside options, inducing a chilling effect. In our mechanism, this is not the case. In fact, by making concessions, that is, propositions that would be acceptable to the other player, a player can actually improve his own status quo outcome, more than the player who makes only unacceptable propositions. Compte and Jehiel also drew attention to the fact that imposing a deadline for negotiation would in some cases dramatically alter the players' willingness to make concessions, in such a way that players would choose to opt out immediately. This problem does not arise in our mechanism. Here players will begin to make concessions, even if an imposed deadline may prevent them from reaching full agreement. Consensus is reached gradually and the final offers which result in an agreement coincide with a well known static solution concept, the Raiffa solution. Gradualism and final agreement are salient properties of the consensus equilibrium only if at least one player is risk averse. If neither player is risk averse, the players' payoffs at equilibrium still coincide with the payoffs in the Raiffa solution but these payoffs may be achieved through implementation of the lottery and not by consensus. When at least one player is risk averse and if they can make as many proposals as they want, at the consensus equilibrium they never choose to implement the lottery but gradually move proposals towards each other to reach an agreement. There are also equilibria where agreement is not reached. However, these seem less likely since they use strategies that are weakly dominated by those in the consensual equilibrium. While we have studied the mechanism under the assumption that players have perfect information, it should be noted that the mechanism "arbitrator" does not need to know anything about the bargaining situation except information that is provided by the players. Moreover, the mechanism can be implemented in a discreet choice setting just as well as with a continuous bargaining domain, which is the setting considered here. From a theoretical point of view, the mechanism we propose has appealing properties. It is however more complicated than related ones that have been suggested in the literature on arbitration schemes,

due to the possibility of rejecting the arbitrator's choice in order to return to a previous phase in the negotiation. Nevertheless, the behavior at equilibrium involves a fairly simple strategy which can be summarized as asking for as much as one can, without making the other player prefer a lottery. It would be interesting to explore in an experimental setting whether individuals would be capable of finding these strategies. Other interesting extensions that we have not explored here is to consider the effects of a bias in the lotteries.

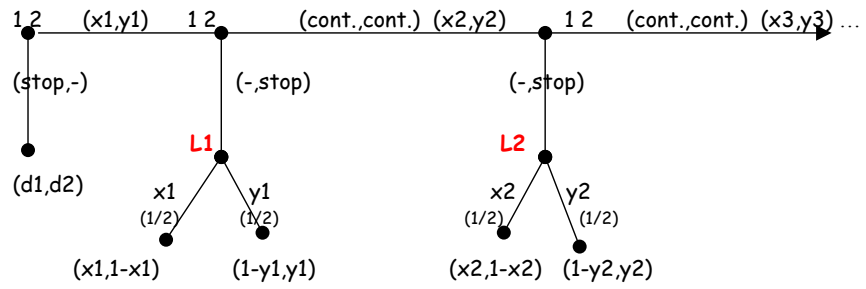
## References

- [1] Ashenfelter, O., and D. E. Bloom, (1984), "Models of Arbitrator Behavior: Theory and Evidence", *The American Economic Review*, 74 (1), 111-124.
- [2] Compte, O., and P. Jehiel, (2004), " Gradualism in Bargaining and Contribution Games ", *The Review of Economic Studies*, 71 (4), 975-1000.
- [3] Crawford, V. P., (1979), "On Compulsory-Arbitration Schemes", *Journal of Political Economy*, 87 (1), 131-159.
- [4] Dickinson, D.L. (2006), "The chilling effect of optimism: The case of final-offer arbitration", *The Journal of Socio-Economics*, 35, 17-30.
- [5] Farber, H.S., (1980), "An analysis of final-offer arbitration", *The Journal of Conflict Resolution*, 24 (4), 683-705.
- [6] Farber, H.S., (1981), "Splitting-the-Difference in Interest Arbitration," *Industrial and Labor Relations Review*, 35, 70-77.
- [7] Farber, H.S., and C., Katz, (1979), "Interest Arbitration, Outcomes, and the Incentive to Bargain: The Role of Risk Preferences," *Industrial and Labor Relations Review*, 33, 55-63.
- [8] Feuille, P., (1975), "Final Offer Arbitration and the Chilling Effect", *Industrial Relations*, 14 (3), 302-310.
- [9] Hebdon, R., (1996), Public sector dispute resolution in transition. In: Belman, D., Gunderson, M., Douglas, H. (Eds.), *Public Sector Employment in a Time of Transition*. Industrial Relation Research Association: University of Wisconsin-Madison, 85-125.
- [10] Kalai, E., and M., Smorodinsky, (1975), "Other Solutions to Nash's Bargaining Problem," *Econometrica*, 43, 513-518.
- [11] Kennan, J. and R., Wilson, (1993), "Bargaining with Private Information," *Journal of Economic Literature*, 31, 45-104.
- [12] Kriticos, A., (2006), "The Impact of Compulsory Arbitration on Bargaining Behavior - An Experimental Study", *Economics of Governance*,

7 (3), 293-315.

- [13] Moulin, H., (1984), "Implementing the Kalai-Smorodinsky bargaining solution", *Journal of Economic Theory*, 33, 32-45.
- [14] Nash, J. (1950), "The Bargaining Problem", *Econometrica*, 28, 155-152.
- [15] Raiffa, H., (1952), "Arbitration Schemes for Generalized Two-Person Games", in H.W. Kuhn and A.W. Tucker, eds., *Contributions to the Theory of Games II*, *Annals of Mathematical Studies*, No 28, Princeton: Princeton University Press, 361-387.
- [16] Rubinstein, A. (1982), "Perfect Equilibrium in a Bargaining Model", *Econometrica*, 50, 97-109.
- [17] Sjöström, T., (1991), "Stahl's bargaining model", *Economics Letters*, 36, 153-157.
- [18] Stahl I., (1972), *Bargaining Theory*, Stockholm School of Economics.
- [19] Stevens, C. M., (1966), " Is Compulsory Arbitration Compatible With Bargaining ? ", *Industrial Relations*, 5, 38-52.





imsart ver. 2006/03/07 file: bargain270608.tex date: July 1, 2008

FIG 1. The extensive form of the non recursive model

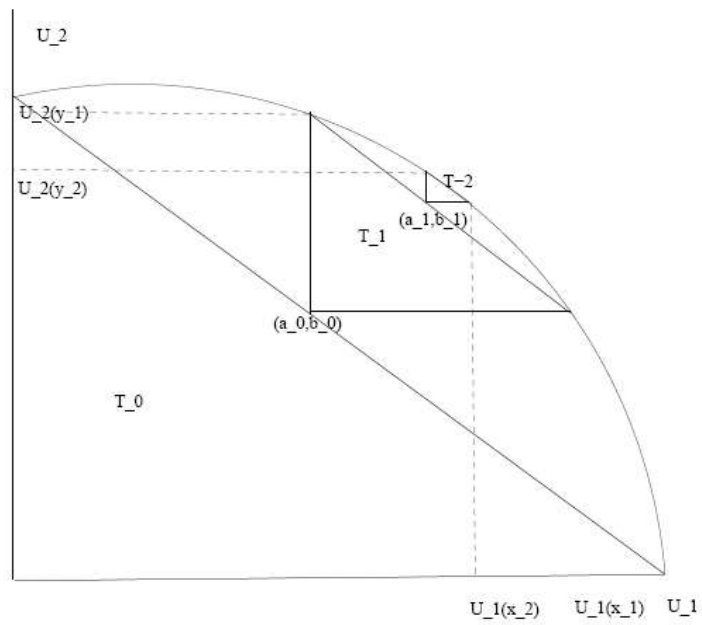


FIG 2. The extensive form of the recursive model

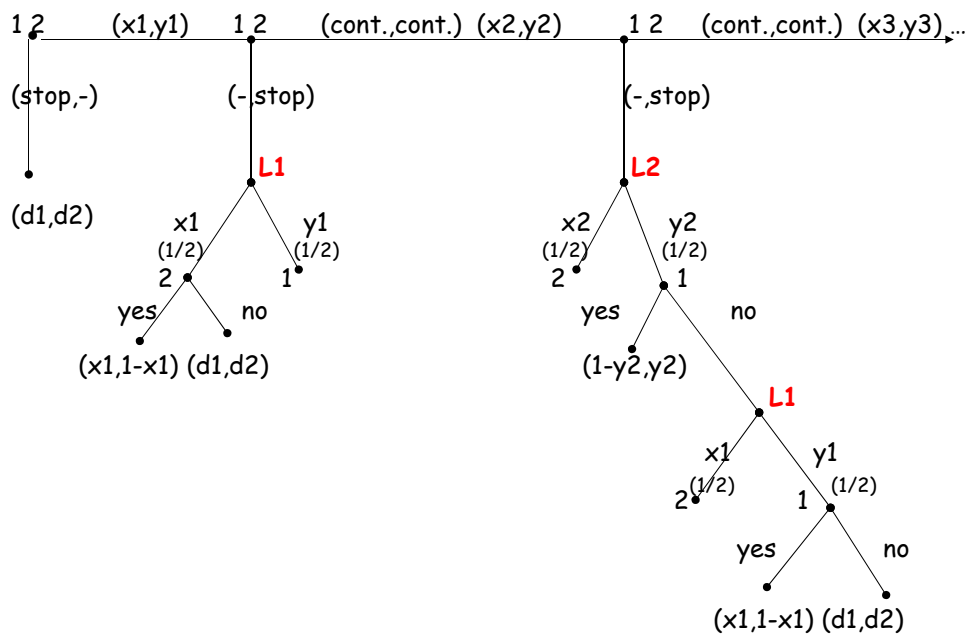


FIG 3. graphical representation of the Raiffa solution