# Commitment in Alternating Offers Bargaining*

Topi Miettinen
Stockholm School of Economics

Andrés Perea
Maastricht University

This version: September 2010

**Abstract**

We extend the Ståhl-Rubinstein alternating-offer bargaining procedure to allow players, prior to each bargaining round, to simultaneously and visibly commit to some share of the pie. If commitment costs are small but increasing in the committed share, then the unique outcome consistent with common belief in future rationality (Perea, 2010), or more restrictively subgame perfect Nash equilibrium, exhibits a second mover advantage. In particular, as the smallest share of the pie approaches zero, the horizon approaches infinity, and commitment costs approach zero, the unique bargaining outcome corresponds to the reversed Rubinstein outcome $(\delta/(1+\delta), 1/(1+\delta))$, where $\delta$ is the common discount factor.

KEYWORDS: alternating offer bargaining, bargaining power, commitment, epistemic game theory, patience

JEL CODES: C73, C78, D84

**Topi Miettinen:** Stockholm School of Economics
P.O. Box 6501, Sveavägen 65, SE-113 83 Stockholm, Sweden.
topi.miettinen@hhs.se
Web: http://www2.hhs.se/SITE/homepages/topi.html

**Andrés Perea:** Maastricht University, Department of Quantitative Economics
P.O. Box 616, 6200 MD Maastricht, The Netherlands.
a.perea@maastrichtuniversity.nl
Web: http://www.personeel.unimaas.nl/a.perea/

*We would like to thank Tore Ellingsen and Jörgen Weibull for insightful comments.

## 1. Introduction

*"...it has not been uncommon for union officials to stir up excitement and determination on the part of the membership during or prior to a wage negotiation. If the union is going to insist on $2 and expects the management to counter with $1.60, an effort is made to persuade the membership not only that the management could pay $2 but even perhaps that the negotiators themselves are incompetent if they fail to obtain close to $2. The purpose ... is to make clear to the management that the negotiators could not accept less than $2 even if they wished to because they no longer control the members or because they would lose their own positions if they tried."*

In this quotation from his classic book, Schelling (1960) vividly illustrates that strategic commitment is often an essential feature of bargaining tactics and that parties of negotiations often have access to actions that commit them to some strategically chosen bargaining position. During the 50 years that have followed, however, there has been fairly little progress in game theoretic analysis of such commitments. Notable exceptions are Crawford (1982), Muthoo (1992, 1996), and more recently Ellingsen and Miettinen (2008, 2010) and Li (2010).

In this paper, we follow Muthoo (1992), Li (2010) and Ellingsen and Miettinen (2010) and consider the effect of commitment strategies in a dynamic complete information bargaining framework. We limit attention to the finite horizon alternating offer game (Ståhl, 1972; Rubinstein, 1982) although we do study the infinite horizon limit. We model parties who can, simultaneously and prior to each alternating offer, commit not to agree on any share smaller than specified in the commitment. Strategic commitment is assumed to incur small costs. These costs are increasing in the share to which the party commits reflecting the idea that more resources must be invested to build a credible commitment when the opportunity cost of turning down a deal is larger. After each round of bargaining, any prior commitments are relaxed and players may again choose a commitment to any share they wish.

Our work builds heavily upon the work of Rubinstein (1982) and Ellingsen and Miettinen (2008). The main insight of Rubinstein's pioneering work on bargaining is that, under complete information, equilibrium strategies are determined by the relative impatience of the bargaining parties. In equilibrium, the proposer makes an offer so that the responder is indifferent between accepting the offer and rejecting it, given the cost of waiting; and the responder accepts the offer. Thus there is an efficient immediate agreement with a first-mover advantage. Ellingsen and Miettinen (2008) recently illustrated how mutual attempts of aggressive incompatible commitment may be unavoidable in bilateral bargaining, if commitments can only be attempted prior to the negotiations and they are not certain to succeed. With certain commitments, it is better to remain flexible and sign on the other's offer than to commit oneself, especially if commitment incurs a small cost. Yet, a very aggressive commitment is the best-response to not committing at all. The opponent's aggressive commitment leaves little room of manoeuvre and thus even small loopholes in the opponent's commitment invite betting on being the only one to succeed in an aggressive commitment attempt despite the risk of an impasse.

In the present context we show that, in line with Rubinstein (1982) and contrary to Ellingsen and Miettinen (2008), the deal is always stroke immediately. However, contrary to Rubinstein's outcome, there is a second-mover advantage rather than a first-mover advantage! This is surprising, as both parties commit simultaneously at the beginning of every round and there are no exogenous asymmetries in the commitment technology. The intuition for the result is the following: commitments are short-lasting and there is no uncertainty about who has the initiative (pre-determined alternating offers structure). Thus, the first one to propose does not have to worry about committing before her proposal; whatever share the second-mover commits to, it is best for the first-mover to avoid any costs by refraining from committing. In equilibrium she proposes the second-mover the share he commits to and takes the residual share herself, provided that the residual makes her better off than waiting for the follow-up round. Knowing this, the second-mover will commit up to the share that makes the first-mover indifferent between having the residual share and waiting. Thus, the presence of symmetric commitment strategies entirely reverses the bargaining power of the parties! In the limit, where the cost of commitment and the smallest indivisible share of the pie approach zero, and where the number of rounds approaches infinity, the outcome approaches the reversed Rubinstein (1982) outcome, $(\delta/(1+\delta), 1/(1+\delta))$, where $\delta$ is the common discount factor.

We analyze the game using the concept of *common belief in future rationality* (Perea, 2010), meaning that both players always believe that the opponent will choose rationally now and in the future, that both players always believe that both players always believe that the opponent will choose rationally now and in the future, and so on. This concept is a typical *backward induction concept* as it assumes that players, throughout the game, only reason about the future, and not about the past. In fact, in generic games with perfect information, common belief in future rationality leads to the unique backward induction strategy for every player. The concept is weaker, but at the same time more basic, than subgame perfect equilibrium as no equilibrium condition is being imposed[1]. As Perea (2010) has shown, the strategies that may be chosen under common belief in future rationality can be computed by the algorithm of *backward dominance.* Since every subgame perfect equilibrium of the game survives the backward dominance procedure, it follows that the outcome of the bargaining game described above is also the unique subgame perfect equilibrium outcome of the game.

Our analysis contributes to the agenda, initiated by Schelling (1956), of carefully analyzing and understanding commitment institutions and mechanisms and their implications on the bargaining outcomes. Among the related works, Crawford (1982) and Muthoo (1992, 1996) have studied the effects of revocable commitments; in our model in contrast, commitments auto-

[1] Another non-equilibrium concept which has been developed for dynamic games is *extensive form rationalizability* (Pearce (1984), Battigalli (1997), Battigalli and Siniscalchi (2002)). However, extensive form rationalizability is a typical *forward induction concept,* which requires players also to critically reason about the opponents' past choices, and not only about the opponents' present and future choices, as common belief in future rationality does. In many dynamic games, extensive form rationalizability and common belief in future rationality lead to different strategy choices for the players .

matically vanish but parties can make a costly recommitment to any share they like after each round. Ellingsen and Miettinen (2008, 2010) analyze costly and long-lasting precommitment to offers and thus they cannot be freely adjusted after each round. Also, unlike in Ellingsen and Miettinen (2008, 2010), players do not commit directly to proposals in our model, but rather to veto any deal where their share is smaller than their commitment. In this respect the model resembles those of Muthoo (1992, 1996) and the endogenous commitment models analyzed in Fershtman and Seidmann (1993), Li (2007), and Miettinen (2010), in which yet, the smallest acceptable shares are determined by the bargaining history in some exogenously determined way rather than freely chosen by players.

Schelling also mentions reputation as an important means of pre-commitment. Myerson (1991) and Abreu and Gul (2000) analyze such reputation contexts where one party has incomplete information about the opponent's stubbornness not to accept anything less than an exogenously given share of the pie. The opponent can then use commitment tactics that exploit this incomplete information and strategically mimic stubbornness in order to force concessions from the other party. This induces delay and influences the final sharing.

Outside options bear a close relation to the current complete information alternating offer bargaining model. Compte and Jehiel (2002) show that exogenous outside options may altogether eliminate the strategic effects of reputation for stubbornness. It has also been shown that, when a party, by opting out, gets a payoff that is inferior to the equilibrium payoff he would obtain in the game without outside options, then these latter have no effect on the equilibrium outcomes (Binmore et al. 1989). In our setting deliberately chosen commitment strategies influence bargaining outcomes exactly because they are chosen so as to force concessions superior to those in the Rubinstein outcome.

The results closest in spirit to ours are perhaps Dixit's (1980) extension of the Spence-Dixit excess capacity model and Ellingsen's (1995) analysis of timing in oligopoly. Dixit shows that an incumbent firm, who nevertheless is presumed to play the role of the follower, can use the commitment, provided by an excess capacity investment, in seizing limited initiative back from the entrant. Ellingsen shows in a Cournot duopoly setting that, if one of the firms alone can choose to pile up investment later, that firm will endogenously end up in the Stackelberg follower position, whereas the firm who can only invest at present will become the leader.[2]

The paper is organized as follows. In Section 2, we set up the model and the bargaining procedure. In Section 3 we present the concept of common belief in future rationality, and the associated algorithm of backward dominance. In Section 4 we analyze the model with one round of bargaining. We will use it as a benchmark for our analysis of more than one round in Section 5. We also investigate the limit behavior of this outcome, when the commitment costs go to zero and the number of rounds goes to infinity. We conclude in Section 6.

[2] These are the only strategies surviving iterated elimination of weakly dominated strategies.

## 2. The Bargaining Procedure

There are two players, 1 and 2, who must reach an agreement about the division of one unit of some good. We assume that the smallest amount is $1/K$ for some integer number $K$. Let $X := \{0, 1/K, 2/K, ..., 1\}$. Hence, the set of possible divisions is given by

$$D := \{(x_1, x_2) : x_1, x_2 \in X \text{ and } x_1 + x_2 \leq 1\}.$$

Players 1 and 2 use the following bargaining procedure, which can last for at most $N$ rounds.

**Round 1:** At the beginning, both players simultaneously choose *commitment levels* $c_1, c_2 \in X$. The commitment levels become known to both players, and player 1 proposes a division $(x_1, x_2) \in D$ with $x_1 \geq c_1$. Subsequently, player 2 decides whether to accept or reject the proposal under the condition that he can only accept offers with $x_2 \geq c_2$. If he accepts, $(x_1, x_2)$ is the final outcome. If he rejects, the game moves to round 2.

**Round 2:** At the beginning, both players simultaneously choose new commitment levels $c_1, c_2 \in X$. Afterwards, player 2 proposes a division $(x_1, x_2) \in D$ with $x_2 \geq c_2$. Subsequently, player 1 decides whether to accept or reject $(x_1, x_2)$, under the condition that he can only accept offers with $x_1 \geq c_1$. If he accepts, $(x_1, x_2)$ is the final outcome. If he rejects, the game moves to round 3.

**Round 3:** This is a repetition of round 1. And so on.

This bargaining procedure goes on until an agreement is reached, or the process enters round $N+1$. In round $N+1$, a given division $(y_1, y_2) \in D$ is realized.

We assume that both players incur a cost for commitment, and that this cost is increasing in the amount to which the player commits. The reason for the latter is that the higher the amount to which the player commits, the more difficult it will be to stick to this commitment. More precisely, if player $i$ commits to an amount $c_i$, this will cost him $\lambda c_i$, where $\lambda$ is some small positive number. For convenience, we assume that $\lambda$ is the same for both players. We finally assume that both players discount future payoffs by a common discount factor $\delta$.

So, in view of all the above, the players' utilities are as follows: If the players reach an agreement on division $(x_1, x_2)$ in round $n$, then the utility for player $i$ is

$$\delta^{n-1} x_i - \lambda \left( c_i^1 + \delta c_i^2 + ... + \delta^{n-1} c_i^n \right),$$

where $c_i^k$ is the commitment level chosen at round $k$. If the game reaches round $N+1$, his utility would be

$$\delta^N y_i - \lambda \left( c_i^1 + \delta c_i^2 + ... + \delta^{N-1} c_i^N \right).$$

In order to avoid uninteresting indifferences, we assume that $\delta$ is such that a player is never indifferent between two outcomes that are realized at two different rounds. Note that for every open interval $(a, b)$ in $[0, 1]$, we can always find such a $\delta$ that lies in $(a, b)$, since there are only finitely many rounds, and finitely many divisions and commitment levels at every round. By choosing $\delta$ is this way, we guarantee that a player will never be indifferent between accepting and rejecting an offer. This, eventually, will lead to a unique outcome in the bargaining game, which makes our analysis more transparent.

Within our bargaining procedure above, the interpretation of the commitment levels is thus that the proposer commits to never offer less than his commitment level for himself, whereas the responder commits to reject any offer that would give him less than his commitment level. With this interpretation in mind, it makes intuitive sense that the cost of commitment is assumed to be increasing in the commitment level. A higher commitment level, namely, more heavily restricts the subsequent choice set of the player, and for higher commitment levels, makes it more tempting for this player to break his commitment. The higher cost of commitment for larger shares should in this way reflect the larger opportunity cost.

## 3. Common Belief in Future Rationality

The concept we use to analyze the game is *common belief in future rationality* (Perea, 2010). In this concept, we assume that a player always believes that his opponent will choose rationally now and in the future, that a player always believes that his opponent always believes that he will choose rationally now and in the future, and so on. It is thus a typical backward induction concept, as a player is only required to reason about the opponents' choices in the *future* of the game, and not about the opponents' *past* choices. In that sense, is differs considerably from the notion of *extensive form rationalizability* (Pearce, 1984; Battigalli, 1997; Battigalli and Siniscalchi, 2002), which is a typical *forward induction* concept, requiring players also to think critically about opponents' *past* choices, and where possible try to draw conclusions from these about possible future moves by this opponent. In many dynamic games, common belief in future rationality and extensive form rationalizability select different strategy choices for the players. In fact, in terms of strategy choices there is no general logical relationship between the two concepts – in some games the former can be more restrictive, in other games the latter can be more restrictive, and in yet some other games they may select completely opposed strategies for a given player.

For a formal definition of common belief in future rationality within an epistemic model the reader is referred to Perea (2010). In that paper, it is also shown that the strategies that can rationally be chosen under common belief in future rationality are characterized by an elimination procedure called *backward dominance*. For dynamic games with observed past choices[3], such as the game we consider in this paper, the backward dominance procedure works as follows: We start at the ultimate subgames, that is, those subgames after which the game is over. At each of those subgames, we restrict to strategies that reach this subgame, and

[3] That is, games that may include simultaneous moves, but where the players always observe which choices have been made by the opponents in the past.

apply iterated strict dominance (or, iterated elimination of strictly dominated strategies) to this restricted game.

We then move to penultimate subgames, that is, subgames after which either the game is over or an ultimate subgame starts. At each of those subgames, we restrict to strategies that reach this subgame and that have not been eliminated yet by the procedure. We then apply iterated strict dominance to these restricted games.

And so on, until we reach the beginning of the game. There, we restrict to strategies that have not been eliminated yet, and apply iterated strict dominance to this restricted game. The strategies that survive the final round of iterated strict dominance at the beginning of the game are said to survive the backward dominance procedure.

For dynamic games with observed past choices, subgame perfect equilibrium is a strict refinement of common belief in future rationality. That is, every strategy that is optimal in a subgame perfect equilibrium can also be chosen rationally under common belief in future rationality, but not vice versa (Perea, 2010). So, common belief in future rationality is weaker than subgame perfect equilibrium. At the same time, we believe it is a more basic concept than subgame perfect equilibrium, as no equilibrium condition is being imposed – we only require a player to believe that his opponent chooses rationally now and in the future, that he believes that his opponent reasons in this way as well, and so on. No other conditions are being imposed. If we apply common belief in future rationality to generic games with perfect information, we obtain the usual backward induction procedure, and hence the unique backward induction strategy for every player. This confirms the intuition that it is indeed a backward induction concept.

## 4. The Case of One Round

We start with the easiest case, namely when there is only one round of bargaining. For this case, we already encounter a surprising result: Under common belief in future rationality (and hence also under subgame perfect equilibrium) the proposer faces a first-mover *disadvantage,* rather than a first-mover advantage. Actually, we can say a little more, namely the proposer gets exactly what he would obtain as a responder in the procedure without commitment. So, introducing the possibility to commit reverses the outcome completely! All this is obtained under the assumption that the commitment costs are sufficiently small. More precisely, we require $\lambda < 1 - \delta$.

**Theorem 4.1.** *(Case of one round) Consider the procedure with only one round of bargaining, and suppose that $\lambda < 1 - \delta$. Then, under common belief in future rationality, player 1 chooses commitment level 0, player 2 chooses commitment level $1 - \overline{\delta y_1}$, player 1 proposes $(\overline{\delta y_1}, 1 - \overline{\delta y_1})$ and player 2 accepts.*

Here, $\overline{x}$ denotes the smallest number in $X$ larger than, or equal to, $x$. Remember that $(y_1, y_2)$ is the outcome if the proposal is rejected. So, player 1, the proposer, gets the minimal amount he would still accept, whereas player 2, the responder, gets all the surplus! Notice that in the classical bargaining procedure without commitment, this would be exactly the outcome when player 2 would be the proposer and player 1 the responder.

**Proof.** For every pair $(c_1, c_2)$ of commitment levels, the subgame that starts after $(c_1, c_2)$ is a game with perfect information. Hence, applying backward dominance to this subgame is the same as using backward induction. After every $(c_1, c_2)$, the backward induction outcome is as follows:

1. If $c_1 + c_2 > 1$, or $c_1 > 1 - \delta y_2$, then player 2 will reject any proposal by player 1. Hence, the outcome will be $(y_1, y_2)$, with utility $\delta y_1 - \lambda c_1$ for player 1, and utility $\delta y_2 - \lambda c_2$ for player 2.

2. If $c_2 > 1 - \delta y_1$, then player 1 does not want to make any offer that player 2 would accept. Hence, the outcome will be $(y_1, y_2)$, with utility $\delta y_1 - \lambda c_1$ for player 1, and utility $\delta y_2 - \lambda c_2$ for player 2.

3. Suppose that $c_1 + c_2 \leq 1$ and $\delta y_2 < c_2 < 1 - \delta y_1$. Then, the best that player 1 can do is to offer player 2 precisely $c_2$, which player 2 would accept. So, the outcome would be $(1 - c_2, c_2)$, with utility $1 - c_2 - \lambda c_1$ for player 1, and utility $c_2 - \lambda c_2$ for player 2.

4. Suppose that $c_1 < 1 - \delta y_2$ and $c_2 < \delta y_2$. Then, the best that player 1 can do is to offer player 2 exactly $\overline{\delta y_2}$, which player 2 would accept. So, the outcome would be $(1- \overline{\delta y_2}, \overline{\delta y_2})$, with utility $1- \overline{\delta y_2} - \lambda c_1$ for player 1, and utility $\overline{\delta y_2} - \lambda c_2$ for player 2.

It can easily be seen that this covers all possible cases. In Figure 1 we have depicted the backwards induction utilities for both players after every possible pair $(c_1, c_2)$. So, Figure 1 represents exactly the restricted game that the backward dominance procedure would consider at the beginning of the game. In order to finish the backward dominance procedure, we must apply iterated strict dominance to the game in Figure 1.

From player 1's utilities in Figure 1 it is easily verified that, for every $c_2$, player 1's utility is decreasing in his commitment level $c_1$. This means, however, that $c_1 = 0$ strictly dominates every other $c_1$ for player 1. So, we eliminate all $c_1 > 0$ for player 1, which leaves only $c_1 = 0$. But then, in the reduced game that remains, player 2's best choice is $c_2 = 1 - \overline{\delta y_1}$. Here, we use the assumption that $\lambda < 1 - \delta$. As we have seen above, the best that player 1 can do in this case is to propose $(\overline{\delta y_1}, 1 - \overline{\delta y_1})$, which player 2 would accept. So, by applying the backward dominance procedure, we obtain that player 1 chooses commitment level $c_1 = 0$, player 2 chooses $c_2 = 1 - \overline{\delta y_1}$, player 1 proposes $(\overline{\delta y_1}, 1 - \overline{\delta y_1})$ and player 2 accepts. This completes the proof. ■

Theorem 4.1 illustrates two points. First, by setting $y_1 = y_2 = 0$, one can see that in a single round ultimatum bargaining game, the second mover will reap the entire pie. Second, by setting $y_1 = \delta/(1 + \delta)$ and $y_2 = 1/(1 + \delta)$, we would effectively add a simultaneous move

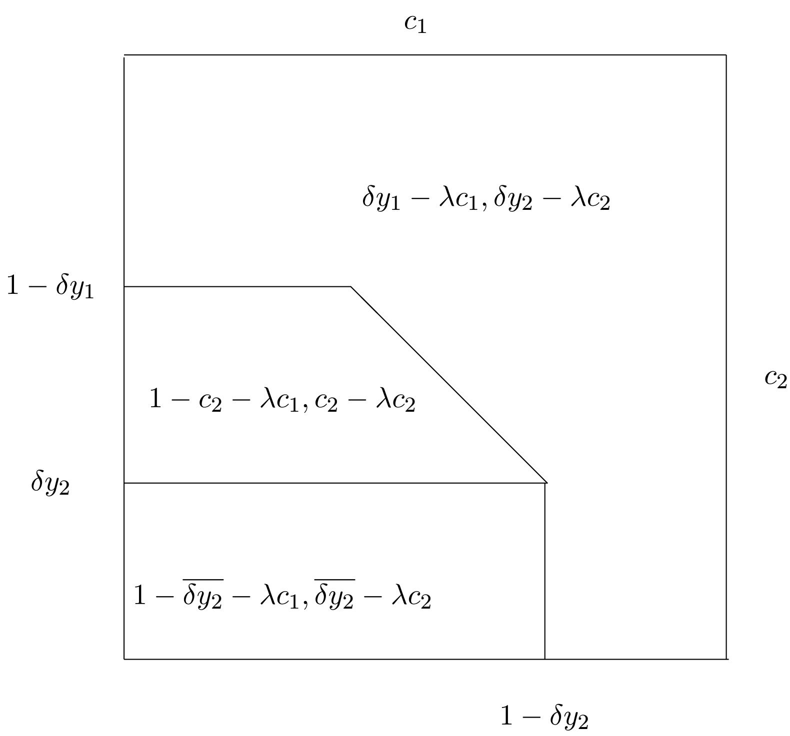


**Figure 1:** The case of one round: Backward induction utilities after every pair $(c_1, c_1)$

commitment stage to the alternating-offer protocol such that precommitments are valid only in the first round of bargaining. Our result shows that this would in fact put the recipient of the first offer in an even more advantageous position than where the proposer in the game without commitments is: the recipient of the first proposal commits to $1-\delta^2/(1+\delta) = \frac{1+\delta-\delta^2}{1+\delta}$ and leaves only $\delta^2/(1+\delta)$ to the first mover.

## 5. The Case of More Rounds

We now turn to the case of more than one round. Also in this case, common belief in future rationality leads to a unique outcome, where the proposer at round 1 faces a first-mover *disadvantage,* rather than a first-mover advantage. Actually, when the commitment cost $\lambda$ tends to zero, then the first proposer gets exactly what he would obtain as the first responder in the procedure without commitment, and vice versa. So, again, introducing the possibility to commit completely reverses the outcome as $\lambda$ tends to zero! As every subgame perfect equilibrium satisfies common belief in future rationality (Perea, 2010), it follows that this outcome is also the unique subgame perfect equilibrium outcome in the game.

**Theorem 5.1.** *(Case of more than one round) Suppose that the bargaining procedure consists of $N$ potential rounds, and that $\lambda < 1-\delta$. Let $p$ denote the proposer at round 1, and $r$ the responder at round 1. Then, common belief in future rationality leads to a unique outcome, namely at round 1 proposer $p$ commits to $c_p = 0$, responder $r$ commits to $c_r = x_r^N$, proposer $p$ proposes the division $(x_p^N, x_r^N)$ and responder $r$ accepts, where $x_p^N, x_r^N$ are recursively given by*

$$\begin{aligned} x_p^1 &= \overline{\delta y_p},\ x_r^1 = 1-\overline{\delta y_p}, \\ x_p^N &= \overline{\delta(1-\lambda)x_r^{N-1}},\ \text{and } x_r^N = 1-\overline{\delta(1-\lambda)x_r^{N-1}}, \end{aligned}$$

*for every $N \geq 2$.*

If we let $\lambda$ tend to zero, then the recursive equations above would exactly yield the outcomes for the players in the procedure without commitment, but with the roles of the proposer and responder reversed! If the size of the smallest slice $1/K$ is small, then the amounts $x_p^N$ and $x_r^N$ are approximately equal to

$$x_p^N \approx \frac{\delta(1-\lambda)+(-1)^n\delta^n(1-\lambda)^n}{1+\delta(1-\lambda)} + (-1)^{n-1}\delta^n(1-\lambda)^{n-1}y_p$$

and

$$x_r^N \approx \frac{1-(-1)^n\delta^n(1-\lambda)^n}{1+\delta(1-\lambda)} - (-1)^{n-1}\delta^n(1-\lambda)^{n-1}y_p.$$

Recall that $y_p$ is the amount that player $p$ would get at the end of the game, when all proposals have been rejected. These approximations are obtained by setting $\bar{x} = x$ in the recursive equations above, and solving them. If the number of rounds $N$ becomes very large, then

$$x_p^N \approx \frac{\delta(1-\lambda)}{1+\delta(1-\lambda)} \text{ and } x_r^N \approx \frac{1}{1+\delta(1-\lambda)}$$

which shows that there is a clear first-mover disadvantage. If in addition the commitment cost $\lambda$ would tend to zero, then in the limit we would obtain the reversed Rubinstein outcome

$$x_p^N \approx \frac{\delta}{1+\delta} \text{ and } x_r^N \approx \frac{1}{1+\delta}.$$

**Proof of Theorem 5.1.** We prove the statement by induction on the number of rounds. If $N = 1$, then the statement follows immediately from Theorem 4.1.

Now, assume that $N \geq 2$, and that the statement holds for the procedure with $N-1$ rounds. Let $p$ be the proposer at round 1, and $r$ the responder at round 1. Suppose that the proposal at round 1 would be rejected. Then, the subgame that starts at round 2 is a procedure with $N-1$ rounds, where $r$ is the first proposer and $p$ is the first responder. The commitment costs incurred at round 1 are sunk costs, and therefore do not affect the analysis in this subgame. By our induction assumption we know that in this subgame, common belief in future rationality (or, equivalently, the backward dominance procedure) leads to a unique outcome: player $r$ chooses commitment level $c_r = 0$, player $p$ chooses commitment level $c_p = x_r^{N-1}$, player $r$ proposes $x_p^{N-1}$ for himself and $x_r^{N-1}$ for player $p$, and player $p$ accepts. The corresponding utilities would be $x_r^{N-1} - \lambda x_r^{N-1} = (1-\lambda)x_r^{N-1}$ for player $p$ and $x_p^{N-1}$ for player $r$.

Let us now move to round 1, the beginning of the game. If we apply the backward dominance procedure, then we restrict to strategies that have not been eliminated yet, and perform iterated strict dominance within this restricted game. By our induction assumption, the strategies that have not been eliminated yet are such that, whenever the proposal at round 1 is rejected, then the discounted utility for $p$ is $\delta(1-\lambda)x_r^{N-1}$, and the discounted utility for $r$ is $\delta x_p^{N-1}$. By a similar argument as in the proof of Theorem 4.1, we can then conclude that the restricted game at round 1 is given by Figure 2. The only change compared to the proof in Theorem 4.1 is that we substitute $(1-\lambda)x_r^{N-1}$ for $y_1$, and substitute $x_p^{N-1}$ for $y_2$. In Figure 2, the first utility always corresponds to player $p$, and the second utility to player $r$.

From Figure 2, it can easily be concluded that proposer $p$'s utility is strictly decreasing in his commitment level $c_p$. Hence, all choices but $c_p = 0$ are strictly dominated for player $p$. So, we obtain a reduced game in which player $p$ only chooses $c_p = 0$. But then, using the assumption that $\lambda < 1-\delta$, we see that player $r$'s best choice is $c_r = 1-\overline{\delta(1-\lambda)x_r^{N-1}}$. So, the backward dominance procedure (and hence also common belief in future rationality) leads to a unique outcome, in which at round 1 player $p$ commits to $c_p = 0$, player $r$ commits to $c_r = 1-\overline{\delta(1-\lambda)x_r^{N-1}}$, player

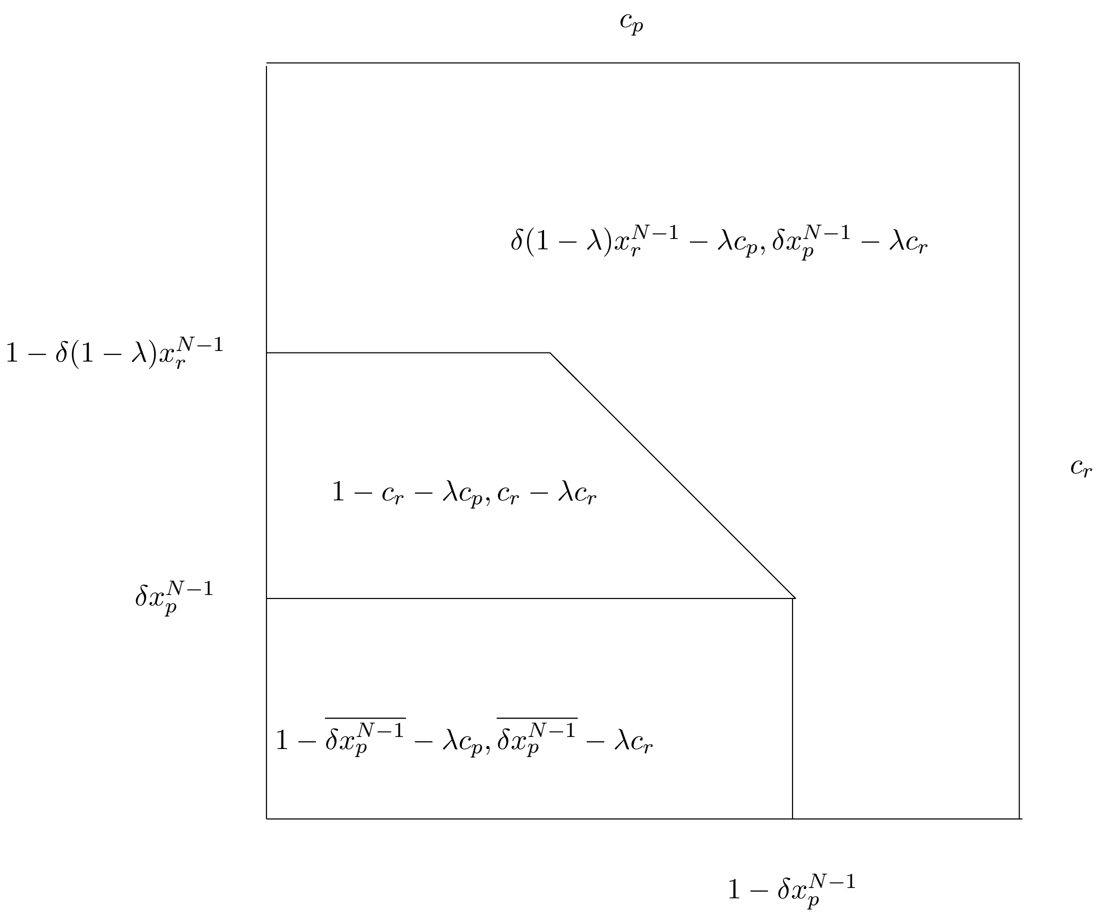


**Figure 2:** The restricted game at round 1

$p$ proposes $\overline{\delta(1-\lambda)x_r^{N-1}}$ for himself, player $p$ proposes $1-\overline{\delta(1-\lambda)x_r^{N-1}}$ for player $r$, and player $r$ accepts. Since $x_p^N = \overline{\delta(1-\lambda)x_r^{N-1}}$ and $x_r^N = 1-\overline{\delta(1-\lambda)x_r^{N-1}}$, the statement of the theorem follows for $N$ rounds. By induction on $N$, the statement holds for every $N$, and hence the proof is complete. ■

## 6. Concluding Remarks

### 6.1. Commitment Costs

In our model we have assumed that the commitment costs for both players are given by $\lambda c$, where $c$ is the amount committed to, and $\lambda$ is some fixed number less than $1-\delta$. In fact, we do not really need this specific functional form for the commitment costs. Instead, we could assume that the commitment costs are given by a more general function $\gamma(c)$, where $\gamma(0)=0$, the function $\gamma$ is strictly increasing in the commitment level $c$, and $\gamma(1)\leq 1-\delta$. The reader may verify that under these assumptions, common belief in future rationality would again lead to a unique outcome, in which the proposer at round 1 faces a first-mover disadvantage. The outcome can be computed by a recursive formula similar to the one used in Theorem 5.1. Also under these assumptions we would obtain the reversed Rubinstein outcome $(\delta/(1+\delta), 1/(1+\delta))$ if we let the number of rounds go to infinity, let the size of the smallest slice go to zero, and let the commitment costs go to zero. However, in the paper we have chosen the specific functional form $\lambda c$ for the commitment costs as to keep the presentation and the analysis as simple as possible.

### 6.2. Common Belief in Future Rationality

The reader may wonder why we have not chosen the more traditional concept of subgame perfect equilibrium, instead of common belief in future rationality, to analyze the game. There are two reasons.

First, common belief in future rationality is a more *basic* concept than subgame perfect equilibrium, as it does not impose any equilibrium condition. It only requires that a player always believes that his opponent will choose rationally now and in the future, that he always believes that his opponent always believes that he will choose rationally now and in the future, and so on. The concept of subgame perfect equilibrium also imposes these conditions, but in addition requires some equilibrium conditions that are harder to justify, and which are not assumed by common belief in future rationality.

Second, using common belief in future rationality as a concept makes our Theorem 5.1 stronger. Namely, common belief in future rationality is a weaker concept than subgame perfect equilibrium. In fact, every subgame perfect equilibrium satisfies common belief in future rationality, but not vice versa. Therefore, our Theorem 5.1 *implies* that the outcome described there is also the unique subgame perfect equilibrium outcome. However, the statement is stronger than this: We do not need the equilibrium condition to arrive at this outcome. Imposing only common belief in future rationality is already enough.

## References


[1] Abreu, D., F. Gul (2000): "Bargaining and Reputation", *Econometrica*, 68, 85-118.

[2] Battigalli, P. (1997): "On rationalizability in extensive games", *Journal of Economic Theory,* 74, 40-61.

[3] Battigalli, P. and M. Siniscalchi (2002): "Strong belief and forward induction reasoning", *Journal of Economic Theory,* 106**,** 356-391.

[4] Binmore, K., Shaked, A., Sutton, J. (1989): "An Outside Option Experiment" *Quarterly Journal of Economics*, 104, 753-770.

[5] Compte, O., P. Jehiel (2002): "On the Role of Outside Options in Bargaining with Obstinate Parties", *Econometrica* 70, 1477–1517.

[6] Crawford, V.P. (1982): "A Theory of Disagreement in Bargaining" *Econometrica*, 50, 607-637.

[7] Dixit, A. (1980): "The Role of Investment in Entry-Deterrence", *Economic Journal*, 90, 95-106.

[8] Ellingsen (1995): "On Flexibility in Oligopoly", *Economics Letters,* 48, 83-89.

[9] Ellingsen, T., T. Miettinen (2008): "Commitment and Conflict in Bilateral Bargaining", *American Economic Review* 98, 1629-1635.

[10] Ellingsen, T., T. Miettinen (2010): "Temporary Commitment in Dynamic Bargaining" Stockholm School of Economics, mimeo.

[11] Fershtman, C., D. Seidmann (1993): "Deadline Effects and Inefficient Delay in Bargaining with Endogenous Commitment." *Journal of Economic Theory* 60, 306–321.

[12] Li, D. (2007): "Bargaining with History-Dependent Preferences", *Journal of Economic Theory* 136, 695-708.

[13] Li, D. (2010): "Commitment and Compromise in Dynamic Bargaining", Chinese University of Hong Kong, mimeo.

[14] Miettinen, T. (2010): "History-dependent Reciprocity in Alternating Offer Bargaining", *Finnish Economic Papers,* 23, 1-15.

[15] Muthoo, A. (1992): "Revocable Commitment and Sequential Bargaining", *Eonomic Journal,* 102, 378-387.

[16] Muthoo, A. (1996): "A Bargaining Model Based on the Commitment Tactic", *Journal of Economic Theory,* 69, 134-152.

[17] Myerson, R. (1991). *Game Theory: Analysis of Conflict*, Harvard University Press, Cambridge, MA/USA.

[18] Pearce, D.G. (1984): "Rationalizable strategic behavior and the problem of perfection", *Econometrica* ,52, 1029-1050.

[19] Perea, A. (2010): "Belief in the Opponents' Future Rationality" Working paper, Available at: http://www.personeel.unimaas.nl/a.perea/

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

[21] Schelling, T. (1956): "An Essay on Bargaining." *American Economic Review* 46, 281-306.

[22] Schelling, T. (1960): *The Strategy of Conflict.* Harvard University Press, Cambridge, MA/USA.

[23] Ståhl, I. (1972). *Bargaining Theory.* Stockholm School of Economics, Stockholm, Sweden.