# Alternating-Offer Games with Final-Offer Arbitration

Kang Rong*

School of Economics, Shanghai University of Finance and Economic (SHUFE)

August, 2012

**Abstract** I analyze an alternating-offer model that integrates the common practice of having an arbitrator determine the outcomes if both players' offers are rejected. I assume that the arbitrator uses final-offer arbitration (as in professional baseball). I find that if the arbitrator does not excessively favor one player, then the unique subgame-perfect equilibrium always coincides with the subgame-perfect equilibrium outcome in Rubinstein's infinite-horizon alternating-offer game. However, if the arbitrator sufficiently favors the player making the initial offer, then delay occurs in equilibrium.

**Keywords:** Alternating-offer game; Final-offer arbitration; Rubinstein equilibrium; Delay in bargaining.

*E-mail: *rong.kang@shufe.edu.cn*. I thank Dan Bernhardt for his patient guidance and helpful advice. I thank Steven Williams, Bart Taub, Mattias Polborn, Paola Manzini, and participants at the UIUC Research Seminar for helpful comments. All errors are mine.

# 1 Introduction

The industrial relations literature features two types of well-known arbitration procedures: conventional arbitration, and final-offer arbitration. Conventional arbitration is an arbitration process in which an arbitrator is free to impose any settlement as the arbitration outcome. In contrast, in final-offer arbitration, the arbitrator must choose one player's final offer as the arbitration outcome. Final-offer arbitration was first proposed by Stevens (1966) and has been widely used in the public-sector to settle labor disputes and in major league baseball to resolve salary disputes (Chelius and Dworkin 1980; Wilson 1994).

Although final-offer arbitration is a common dispute resolution mechanism, it has received little attention in the bargaining literature. The purpose of my paper is to explore how introducing final-offer arbitration affects players' equilibrium strategies and bargaining outcomes. My paper addresses the following questions: when does the introduction of final-offer arbitration have an impact on the equilibrium of the bargaining game? and if final-offer arbitration has an impact, what is the impact and which player benefits? I show that if the arbitrator does not excessively favor one player, then the equilibrium of the game is *unaffected* by the specific details of the arbitrator's preference, and both players obtain Rubinstein equilibrium payoffs. In all other cases, the equilibrium of the game depends on the arbitrator's preference and the player favored by the arbitrator obtains a payoff higher than his Rubinstein equilibrium payoff. In addition, I show that delay in reaching agreement emerges when the arbitrator is excessively biased toward the player who makes the first offer: the bias encourages the player making the first offer to make an unattractive demanding offer in order to get "closer" to the threat of having allocations determined by the biased arbitrator.

My basic framework builds on Yildiz (2011).[1] Two players sequentially make offers. If a player's offer is accepted by his opponent, then that offer is the bargaining outcome and the

[1] In contrast to my current paper, Rong (2012) considers the symmetric arbitration solution (an axiomatic arbitration solution), rather than final-offer arbitration.

game ends. If, instead, both players' offers are rejected by their opponents, an arbitration stage is reached. In contrast to Yildiz (2011), who assumes that the arbitrator chooses the offer that yields the higher Nash product as the arbitration outcome, I consider a general class of final-offer arbitration rules. In particular, the arbitrator's *ideal settlement* can be any point on the Pareto frontier of the bargaining set and the arbitrator chooses his preferred offer, i.e., the offer closest to his ideal settlement as the arbitration outcome.

Player 1 makes the first offer. I find that, (i) if the arbitrator is "balanced" (i.e., the arbitrator does not excessively favor one player),[2] then the unique subgame perfect equilibrium (henceforth SPE) outcome of the game coincides with the equilibrium outcome of Rubinstein's infinite-horizon alternating-offer game (Rubinstein 1982); (ii) if the arbitrator sufficiently favors Player 1, then the unique SPE of the game is such that Player 1 makes an offer that Player 2 *rejects*, and Player 2 makes a counteroffer that Player 1 accepts; the equilibrium payoff received by Player 1 exceeds what he would obtain from the Rubinstein equilibrium; and (iii) if the arbitrator sufficiently favors Player 2, then the unique SPE of the game is such that Player 1 makes a more generous offer than the Rubinstein equilibrium offer that Player 2 accepts immediately.

In the game that Yildiz (2011) considers, the unique SPE outcome coincides with the Rubinstein equilibrium outcome. This result might lead one to believe that the arbitrator's preference in Yildiz (2011) is special to the extent that it may be the only preference for which the SPE outcome coincides with the Rubinstein equilibrium outcome. However, my analysis shows that there exists a broad class (depending on the discount factor) of arbitrator's preferences for which the unique SPE outcome coincides with the Rubinstein equilibrium outcome. What is special about the arbitrator's preference in Yildiz (2011) is that it belongs to that class for *all* discount factors.

One implication of my analysis is the following *irrelevance* result: as long as the arbitrator is not too biased toward a player, then the unique equilibrium of the arbitration game is

[2] An arbitrator is "balanced" if the arbitrator's ideal settlement is close "enough" to the Rubinstein equilibrium outcome. The measure of closeness depends on the discount factor.

unaffected by the arbitrator's preferences. In reality, when the arbitrator uses final-offer arbitration, people may have concerns about the arbitrator's fairness. However, my irrelevance result shows that outcomes are unaffected if the arbitrator has some bias, as long as this bias is not too large. In other words, there is a wide range of arbitrator's preferences under which the equilibrium of the arbitration game is independent of the arbitrator's preferences. Within this range, the precise choice of the arbitrator becomes unimportant.

Another implication of my analysis is that even when players have complete information, arbitration can lead to delay.[3] In equilibrium, delay always occurs when the arbitrator sufficiently favors Player 1. The intuition is as follows. If Player 1 demands more than the Rubinstein equilibrium outcome, then Player 2 will reject Player 1's offer to exploit "time delay". That is, Player 2 will make a counteroffer, which Player 1 accepts in order to avoid the time cost of going to arbitration, and Player 2 is better off by making such a counteroffer than accepting Player 1's initial offer.[4] If, instead, Player 1 demands less than the Rubinstein equilibrium outcome, then the offer is accepted by Player 2 immediately. If the arbitrator *sufficiently* favors Player 1 (for a given discount factor), then Player 1 prefers to demand more than the Rubinstein equilibrium outcome. This is because Player 1 can demand a payoff that is close to the arbitrator's ideal settlement, which is *sufficiently* higher than Player 1's Rubinstein equilibrium payoff. Such a demand is supported by the threat of the biased arbitrator, which implies that Player 2's counteroffer cannot be far away from Player 1's offer. As a result, the benefit that Player 1 can exploit from the biased arbitrator exceeds the cost incurred from the delayed agreement. Therefore, the equilibrium features delay.

[3]Delay in equilibrium within the framework of complete information also occurs in Manzini and Mariotti (2001), Manzini and Mariotti (2004), Ponsatí and Sákovics (1998) and Rubinstein (1982). However, the mechanism for delay is different. It arises in those models due to the existence of multiple equilibria.

[4]In particular, assume that Player 1 makes the offer $(x_1, f(x_1))$ with $x_1 > x^R$, where $x_1$ is Player 1's own demand, $x^R$ is Player 1's Rubinstein equilibrium payoff and $f(\cdot)$ is the Pareto frontier function. Then, rather than accept Player 1's offer, Player 2 can always obtain a higher payoff by rejecting Player 1's offer and making the counteroffer $(\delta x_1, f(\delta x_1))$ (depending on the arbitrator's preference, Player 2 may make a more ungenerous counteroffer in equilibrium), which Player 1 will accept. Notice that $\delta f(\delta x_1) > f(x_1)$, as long as $x_1 > x^R$.

Delay in equilibrium does not arise in Yildiz (2011) because it is not profitable for Player 1 to demand more than the Rubinstein equilibrium outcome. The reason is that in Yildiz (2011), if Player 1 demands more than the Rubinstein equilibrium outcome, then the optimal counteroffer by Player 2 is at most marginally more generous than the Rubinstein equilibrium outcome, because the arbitrator's ideal settlement, the Nash bargaining solution, is close to the Rubinstein equilibrium outcome. For Player 1, the benefit of making a demand that exceeds the Rubinstein equilibrium outcome is thus less than the time cost incurred by reaching a delayed agreement.

Manzini and Mariotti (2001) consider an alternating-offer model that also involves arbitration. They assume that an arbitrator can be called in whenever a player has just rejected an offer and both players agree to move to arbitration. My result contrasts with theirs in the sense that they show that the Rubinstein equilibrium can arise only if the arbitration outcome sufficiently *favors* one player, while my result shows that the Rubinstein equilibrium arises only if the arbitration rule does *not excessively favor* a player. The difference between the results is due to the following: (i) In Manzini and Mariotti (2001), the arbitration outcome is *exogenously* given and *both* players have veto power on the event that the dispute goes to arbitration. Therefore, if the arbitrator is too biased toward one player, arbitration (as a background threat) becomes a non-binding threat and the alternating-offer game in Manzini and Mariotti (2001) yields the same equilibrium outcome as Rubinstein's alternating-offer game. (ii) In my model, the arbitration outcome is *endogenous* in the sense that it depends on players' offers (in other words, players negotiate during the arbitration). In addition, *neither* player has veto power if an arbitrator is called in, once the offers of both players are rejected by the opponents. In my model, it turns out that the Rubinstein equilibrium plays a surprising role: any offer that is less generous than the Rubinstein equilibrium will be rejected by Player 2 (see also footnote 4). As a result, when the final-offer arbitration rule is sufficiently "balanced" in the sense that the arbitrator's ideal settlement is close "enough" to the Rubinstein equilibrium outcome, it is not profitable for Player 1 to demand more than the Rubinstein equilibrium outcome because of the time cost of delay.[5] Player 1 thus makes the Rubinstein equilibrium offer, and Player 2 accepts it immediately.

This paper is organized as follows. Section 2 introduces some notation that helps define the "alternating-offer arbitration game". Section 3 studies the arbitration game that uses the final-offer arbitration rule. Concluding remarks are offered in Section 4.

# 2 The model

There are two players, Players 1 and 2, who are expected utility maximizers. Let $S \subset R^2$ denote the bargaining set, which includes all possible bargaining outcomes, measured in terms of expected utility level. I use $(x_1, y_1) \in S$ to denote Player 1's offer and use $(x_2, y_2) \in S$ to denote Player 2's offer, where $x$ represents Player 1's utility payoff and $y$ represents Player 2's utility payoff. I normalize the disagreement point of $S$ to $(0, 0)$. I assume that $(x, y) \geq (0, 0)$ for any $(x, y) \in S$, and that there is at least one point $(x, y) \in S$ such that $(x, y) >> (0, 0)$. The bargaining set $S$ is assumed to be convex, compact and strictly comprehensive. The bargaining set $S$ is *comprehensive* if $(x', y') \in S$ whenever $(0, 0) \leq (x', y') \leq (x, y)$ and $(x, y) \in S$. The bargaining set $S$ is *strictly comprehensive* if $S$ is comprehensive and there exists a $(x'', y'') \in S$ such that $(x'', y'') >> (x, y)$ whenever $(x, y) \in S$ and $(x', y') \in S$ with $(x', y') \geq (x, y)$ and $(x', y') \neq (x, y)$.

The (weak) Pareto frontier of the bargaining set $S$ is defined as $PF = \{p \in S : q >> p \Rightarrow q \notin S\}$.[6] Define $b_i = max\{U_i : (U_1, U_2) \in S\}$ to be Player $i$'s maximal possible utility payoff from the bargaining set. Define $f : x \to max\{y|(x, y) \in S\}$ for $x \in [0, b_1]$. The function $f$ is well-defined because $S$ is compact. In addition, the assumption that $S$ is convex and strictly comprehensive implies that $f$ is concave, continuous, and strictly decreasing on

[5]More particularly, for Player 1, making a demand that is excessively higher than the Rubinstein equilibrium offer is not supported by the balanced arbitrator and is thus not profitable. Making a demand that is only slightly higher than the Rubinstein equilibrium offer is also not profitable due to the time cost of delay.

[6]$PF$ depends on $S$. However, I fix $S$ in this paper, and I omit the dependency on $S$ in notation whenever there is no confusion.

$[0, b_1]$ with $f(0) = b_2$ and $f(b_1) = 0$ (see Figure 1). Note that $(x, y) \in PF$ if and only if $y = f(x)$.

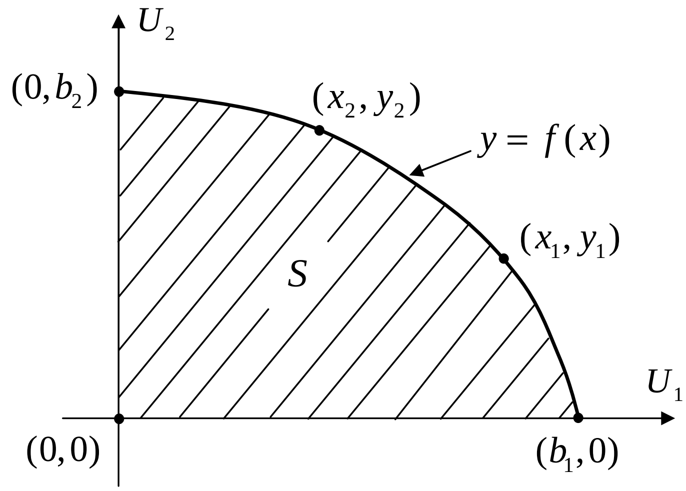


Figure 1: The bargaining set.

I assume that there is an arbitrator who is informed about players' utilities.[7] Define $\mathcal{B} = \{((x_1, y_1), (x_2, y_2)) \,|\, (x_1, y_1) \in S\,,\ (x_2, y_2) \in S\}$. An *arbitration rule* is a function $h : \mathcal{B} \to S$. I write $h((x_1, y_1), (x_2, y_2)) = (h_1((x_1, y_1), (x_2, y_2)), h_2((x_1, y_1), (x_2, y_2)))$, where $h_i((x_1, y_1), (x_2, y_2))$ is the arbitration outcome of Player $i$. A *final-offer arbitration rule* is any arbitration rule where $h((x_1, y_1), (x_2, y_2)) \in \{(x_1, y_1), (x_2, y_2)\}$ for any $((x_1, y_1), (x_2, y_2)) \in \mathcal{B}$. The final-offer arbitration rule $h$ is assumed to be known to both players.

I assume that players have a common discount factor $\delta \in (0, 1)$. The unique intersection point of the curve $y = \delta f(x)$ and the curve $y = f(\frac{1}{\delta}x)$ is denoted by $(\delta x^R(\delta), f(x^R(\delta)))$. That is, $f(x^R(\delta)) = \delta f(\delta x^R(\delta))$ (see Figure 2). The point $(x^R(\delta), f(x^R(\delta)))$ is the outcome of the unique SPE of Rubinstein's infinite-horizon alternating-offer game (Rubinstein 1982).[8] Since $\delta$ is fixed in most parts of the paper, I write $(x^R(\delta), f(x^R(\delta)))$ as $(x^R, f(x^R))$ whenever it does not create confusion.

[7]This is a standard assumption in the law and economics literature (Deck and Farmer 2007).

[8]In Rubinstein's model, if Player 1 makes the Rubinstein equilibrium offer $(x^R(\delta), f(x^R(\delta)))$, then Player 2 is indifferent between accepting the offer and rejecting the offer. This is because if Player 2 accepts the offer, then his payoff is $f(x^R(\delta))$; if Player 2 rejects the offer, then at the next stage Player 2 will make the offer $(\delta x^R(\delta), f(\delta x^R(\delta)))$ in equilibrium, which Player 1 will accept, giving Player 1 a discounted payoff of $\delta f(\delta x^R(\delta)) = f(x^R(\delta))$.

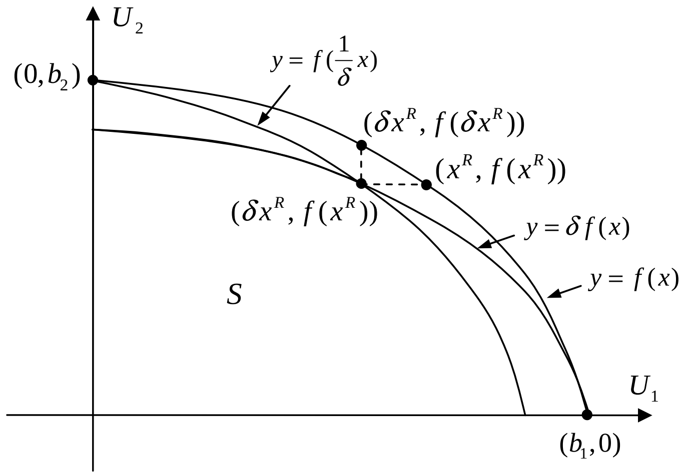


Figure 2: Definition of $(x^R, f(x^R))$.

I define the *alternating-offer arbitration game* (or simply, the *arbitration game*), which generalizes the game considered in Yildiz (2011), as the following three-stage procedure:

**Stage 1:** Player 1 makes an offer $(x_1, y_1) \in S$. Player 2 decides whether to accept the offer, ending the game with $(x_1, y_1)$, or to reject the offer, moving the game to the next stage;

**Stage 2:** Player 2 makes an offer $(x_2, y_2) \in S$. Player 1 decides whether to accept the offer, ending the game with $(x_2, y_2)$, or to reject the offer, moving the game to the arbitration stage;

**Stage 3:** An arbitrator decides the final outcome using arbitration rule $h$, i.e., $h((x_1, y_1), (x_2, y_2))$ is the arbitrated outcome.

The next section analyzes the arbitration game where $h$ is a final-offer arbitration rule.

# 3 Final-offer arbitration game

## 3.1 Notation and assumptions

In this subsection, I introduce notation and assumptions regarding the final-offer arbitration rule $h$.

For any $(x_1, y_1) \in S$, define the set $V(x_1, y_1) = \{(x_2, y_2) \in S | h((x_1, y_1), (x_2, y_2)) =$

$(x_2, y_2)\}$ and $V_x(x_1, y_1) = \{x_2 \in [0, b_1] | (x_2, y_2) \in V(x_1, y_1)\}$. Thus, $V(x_1, y_1)$ is the collection of $(x_2, y_2)$ that is chosen as the arbitration outcome when the final offers of the two players are $(x_1, y_1)$ and $(x_2, y_2)$, and $V_x(x_1, y_1)$ is the set of Player 1's payoffs in the set $V(x_1, y_1)$. The set $V_x(x_1, y_1)$ is nonempty because $(x_1, y_1) \in V(x_1, y_1)$ and $x_1 \in V_x(x_1, y_1)$. The set $V_x(x_1, y_1)$ is bounded because $S$ is bounded. For any $(x_1, f(x_1))$, define $g(x_1) = \min\{x_2 | x_2 \in V_x(x_1, f(x_1))\}$. Note that $g(x_1)$ is well defined if $V_x(x_1, f(x_1))$ is closed (Condition 1 below).

Figure 3 illustrates the sets $V(x_1, f(x_1))$ and $V_x(x_1, f(x_1))$ and the function $g(x_1)$ for the case of *Nash final-offer arbitration rule*. The Nash final-offer arbitration rule is the final-offer arbitration rule such that, among the final offers submitted by the two players, the offer that yields the higher Nash product is chosen as the arbitration outcome. This is the rule that Yildiz (2011) considers.

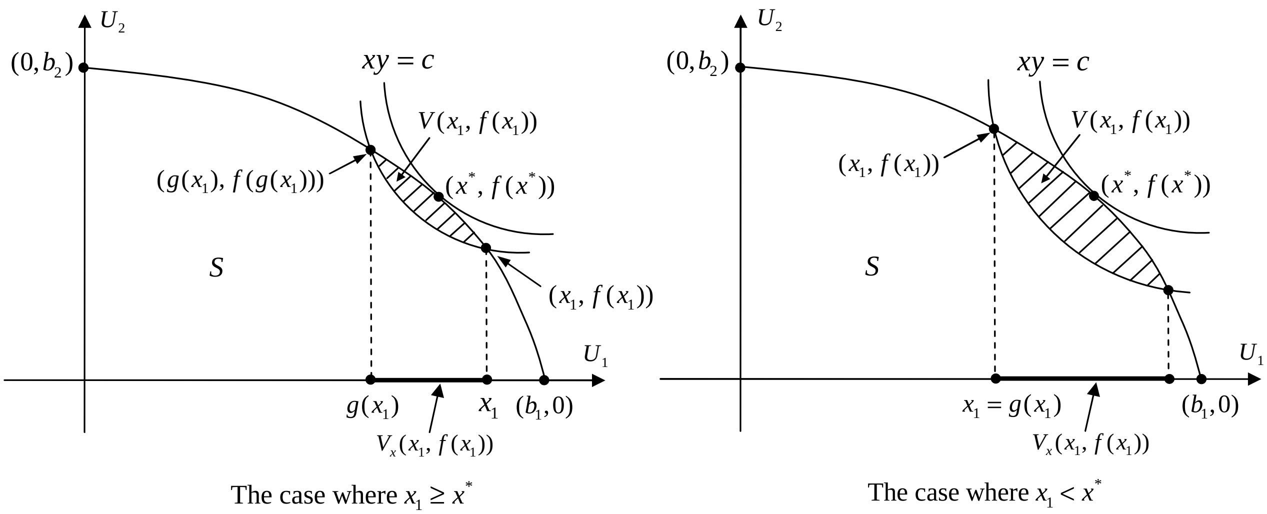


Figure 3: $V(x_1, f(x_1))$, $V_x(x_1, f(x_1))$ and $g(x_1)$ under the Nash final-offer arbitration rule.

A final-offer arbitration rule $h$ is *regular* if there exists a continuous, strongly monotone, and quasiconcave function $u : S \to R$ such that $h((x_1, y_1), (x_2, y_2)) = (x_2, y_2)$ if and only if $u(x_2, y_2) \geq u(x_1, y_1)$.[9] The function $u$ can be regarded as the arbitrator's utility function. It can be easily verified that if the final-offer arbitration rule $h$ is regular, then the set

[9]I am indebted to an anonymous referee for suggesting this definition. Strong monotonicity of $u$ requires that $u(x', y') > u(x, y)$ whenever $(x', y') \geq (x, y)$ and $(x', y') \neq (x, y)$. Quasiconcavity of $u$ requires that the set $\{(x, y) \in S : u(x, y) \geq \overline{u}\}$ be convex for any $\overline{u} \in R$.

$V_x(x_1, f(x_1))$ and the function $g(x_1)$ satisfy the following three properties:

**Condition 1 (Closedness).** *The set* $V_x(x_1, f(x_1))$ *is a closed interval for any* $x_1 \in [0, b_1]$.

**Condition 2 (Continuity).** *The function* $g(x_1)$ *is continuous in* $x_1$ *for* $x_1 \in [0, b_1]$.

**Condition 3 (Relative Fairness).** *There exists an* $x^* \in [0, b_1]$ *such that, (i) for* $x_1 \in [0, x^*]$, *we have* $g(x_1) = x_1$; *(ii) the function* $g(x_1)$ *is decreasing on* $(x^*, b_1]$.

Condition 1 ensures that the function $g$ is well-defined. Condition 2 ensures that the arbitration curve (defined below) is continuous.

The point $(x^*, f(x^*))$ in Condition 3 can be regarded as the arbitrator's *ideal settlement*. Roughly speaking, Condition 3 implies that (i) if Player 1's offer is *generous* to Player 2 (i.e., $x_1 \leq x^*$), then Player 2's offer will not be chosen as the arbitration outcome if Player 2 demands more than the amount that Player 1 offers him; (ii) if Player 1's offer is *ungenerous* to Player 2 (i.e., $x_1 > x^*$), then the arbitrator will allow Player 2 to make an offer where Player 2's own demand exceeds the ideal settlement, and Player 2's offer is still chosen as the arbitration outcome. Moreover, this "tolerance" for Player 2's high demand increases as Player 1's offer becomes more ungenerous.

One can show that the Nash final-offer arbitration rule is regular. Under the Nash final-offer arbitration rule, the utility function that the arbitrator maximizes is $u(x, y) = xy$, and the arbitrator's ideal settlement is the Nash bargaining solution outcome.

Let $\Sigma$ denote the set of all *regular* final-offer arbitration rules. The remainder of this section focuses on regular final-offer arbitration rules.

I now define two curves. The *arbitration curve* is defined as the curve $x = g(f^{-1}(y))$ where $y \in [0, f(x^*)]$ (see Figure 4). Using Condition 3 (ii) and the fact that $f$ is strictly decreasing, one can show that the arbitration curve must be increasing on $[0, f(x^*)]$.

The other curve, the *discounted Pareto frontier*, is defined as follows (see Figure 5):

$$y = \begin{cases} \delta f(x) & if \quad 0 \leq x \leq \delta x^R \\ f(\frac{1}{\delta} x) & if \quad \delta x^R < x \leq \delta b_1. \end{cases}$$

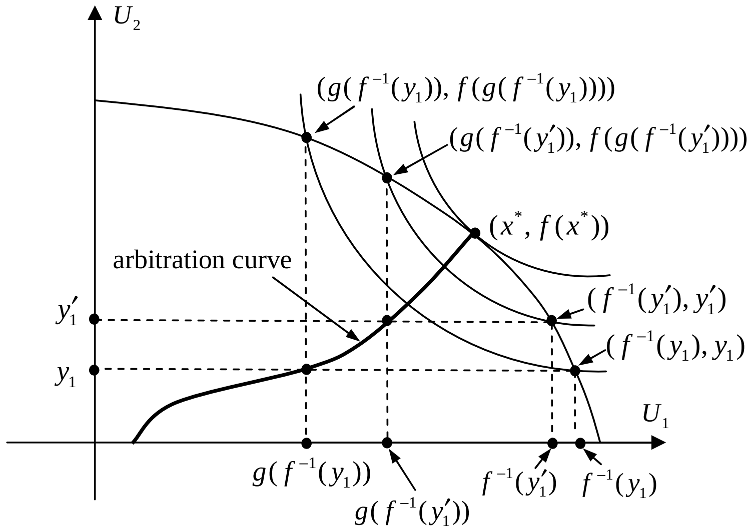


Figure 4: The arbitration curve.

If the arbitration curve and the discounted Pareto frontier intersect, then there must be a unique intersection point. Denote this point by $(\hat{x}(\delta), \hat{y}(\delta))$ (see Figure 5). In the remainder of the paper, I write $(\hat{x}(\delta), \hat{y}(\delta))$ as $(\hat{x}, \hat{y})$ whenever $\delta$ is fixed. If there is no intersection point between those two curves, then the arbitration curve must intersect the $X$-axis at a point that is to the right of the point $(\delta b_1, 0)$ (using the fact that the arbitration curve is continuous). In this case, define $(\hat{x}, \hat{y}) = (\delta b_1, 0)$.

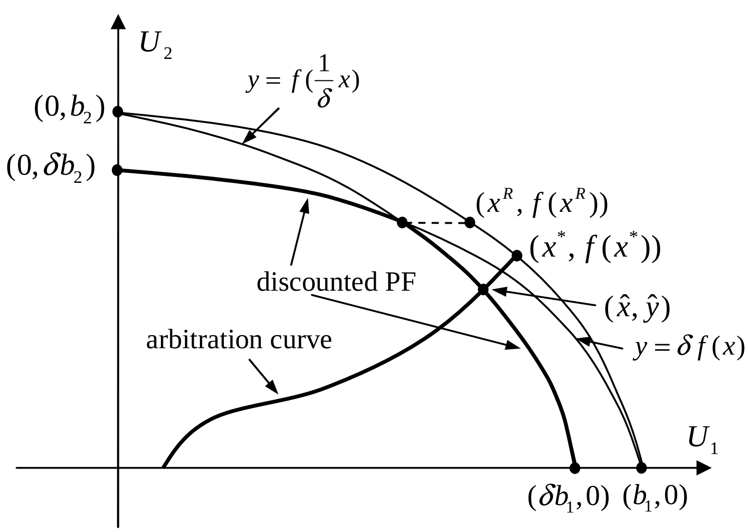


Figure 5: Definition of $(\hat{x}, \hat{y})$.

I show in Theorem 1 that the relationship between $\hat{x}$ and $x^R$ is the key to identifying the

SPE of any arbitration game that uses a regular final-offer arbitration rule.

### 3.2 Characterization of the equilibrium

I make the following two tie-breaking rules to simplify the analysis.

**Tie-breaking rule 1:** If a player is indifferent between acceptance and rejection, he accepts.

**Tie-breaking rule 2:** If a player is indifferent between two options that he can offer his opponent, he chooses the option that yields his opponent a higher payoff.

The following result characterizes the SPE of the arbitration game that uses arbitration rule $h \in \Sigma$.

**Theorem 1.** *In the arbitration game with arbitration rule $h \in \Sigma$, we have:*

*(i) (Rubinstein equilibrium.) If $\delta x^R \leq \hat{x} \leq \frac{1}{\delta} x^R$, then the outcome of the unique SPE is that Player 1 makes the offer $(x^R, f(x^R))$ and Player 2 accepts it;*

*(ii) (type-II arbitration-driven equilibrium.) If $\frac{1}{\delta} x^R < \hat{x} \leq \delta b_1$, then the outcome of the unique SPE is that at Stage 1, Player 1 makes the offer $(\frac{1}{\delta}\hat{x}, f(\frac{1}{\delta}\hat{x}))$, which Player 2 rejects, and at Stage 2, Player 2 makes the offer $(\hat{x}, f(\hat{x}))$, which Player 1 accepts;*

*(iii) (type-I arbitration-driven equilibrium.) If $0 \leq \hat{x} < \delta x^R$, then the outcome of the unique SPE is that Player 1 makes the offer $(f^{-1}(\delta f(\hat{x})), \delta f(\hat{x}))$ and Player 2 accepts it.*

Proof: See Appendix 1. □

A final-offer arbitration rule $h$ is *balanced* (or, an arbitrator is *balanced*) if $\hat{x} \in [\delta x^R, \frac{1}{\delta} x^R]$.[10] Roughly speaking, the balancedness of an arbitration rule requires that the arbitrator's ideal settlement be close enough to the Rubinstein equilibrium outcome.[11] Ac-

[10] The definition of "balanced" is distinct from the definition of "balanced" in Manzini and Mariotti (2001). In both papers, an arbitration rule is balanced if and only if the arbitration rule does not excessively favor a player. However, the measure of "favorness" is different.

[11] This is because the arbitrator's ideal settlement is close to $(\hat{x}, \hat{y})$, and $(\hat{x}, \hat{y})$ is close to the Rubinstein equilibrium outcome by balancedness. The measure of "closeness" depends on the discount factor.

cording to Theorem 1 (i), if a regular final-offer arbitration rule is balanced, then the equilibrium outcome is such that Player 1 offers $(x^R, f(x^R))$, which Player 2 accepts. This result follows from the following two facts. First, one can show that Player 1's offer is accepted by Player 2 if and only if Player 1's demand is less than the Rubinstein equilibrium payoff (Lemma 9 (i) in Appendix 1), so the best offer that Player 1 can make and Player 2 accepts is $(x^R, f(x^R))$. Second, if Player 1 makes a demand that is higher than the Rubinstein equilibrium offer, then Player 1's offer is rejected by Player 2, and Player 2 makes a counteroffer that Player 1 accepts (Lemma 9 (i) and Lemma 8 in Appendix 1). When the final-offer arbitration rule is balanced in the sense that the arbitrator's ideal settlement is close enough to the Rubinstein equilibrium outcome, Player 2's counteroffer is at most marginally more generous than the Rubinstein equilibrium outcome. For Player 1, the extra benefit of making an offer that will be rejected by Player 2 is thus less than the time cost incurred by reaching a delayed agreement. As a result, when the arbitrator is balanced, Player 1 makes the offer $(x^R, f(x^R))$ that Player 2 accepts immediately.

For the class of balanced final-offer arbitration rules, the details of the final-offer arbitration rule are irrelevant to the equilibrium outcome. That is, letting $\Sigma^* = \{h | h \in \Sigma \text{ with the corresponding } \hat{x} \in [\delta x^R, \frac{1}{\delta} x^R]\}$, we have:

**Corollary 2.** *(Irrelevance Result) For any $h \in \Sigma^*$ and $h' \in \Sigma^*$, the arbitration game that uses rule $h$ yields the same equilibrium outcome as the arbitration game that uses rule $h'$.*

To understand the significance of the above irrelevance result, we can imagine that there is a stage before the arbitration game. At this pre-arbitration stage, an arbitrator is chosen from a pool, in which different arbitrators may have different preferences over the two players' offers. My irrelevance result shows that there is a pool of arbitrators, in which the choice of the arbitrator does not matter for the equilibrium of the arbitration game. Moreover, this pool of arbitrators is reasonable and sufficiently wide in the sense that it includes all arbitrators who are *not too biased* toward a player. However, as the discount factor increases, this pool shrinks. When the discount factor approaches 1, the only arbitrator that belongs to this pool *must* be the arbitrator whose ideal settlement is the Nash bargaining solution outcome (see Appendix 3.1 for the robustness of the Nash final-offer arbitration rule).

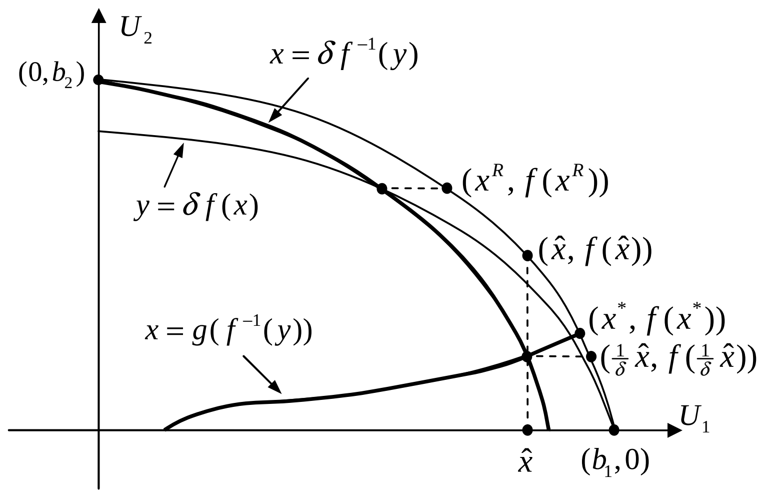


Figure 6: The equilibrium with delay.

If $\frac{1}{\delta}x^R < \hat{x} \leq \delta b_1$, then the unique SPE is a type-II arbitration-driven equilibrium, which is an equilibrium with delayed agreement. One may wonder why it is optimal for Player 1 to make an offer that will be rejected by Player 2, as opposed to making an offer that will be accepted by Player 2. The reason is as follows. On one hand, if Player 1 makes an offer that demands more than the Rubinstein equilibrium payoff, then the offer will be rejected by Player 2, and Player 2 will make a counteroffer that Player 1 accepts (Lemma 9 (i) and Lemma 7 in Appendix 1). However, if the arbitrator sufficiently favors Player 1 ($\frac{1}{\delta}x^R < \hat{x} \leq \delta b_1$), then Player 2's counteroffer will be close enough to Player 1's initial offer, as long as Player 1's demand is not excessively higher than the arbitrator's ideal settlement. One can show that if Player 1 makes the offer $(\frac{1}{\delta}\hat{x}, f(\frac{1}{\delta}\hat{x}))$, then Player 2's counteroffer $(\hat{x}, f(\hat{x}))$ is the most favorable counteroffer that Player 1 could possibly obtain (see Figure 6).[12] On the other hand, if Player 1 makes an offer that demands less than the Rubinstein equilibrium payoff, then the offer will be accepted by Player 2 (Lemma 9 (i)

[12]The reason is that if Player 2 rejects Player 1's offer $(x_1, y_1)$ at Stage 1, then at the next stage, Player 2 makes the counteroffer $(\min\{\delta x_1, g(x_1)\}, f(\min\{\delta x_1, g(x_1)\}))$, which Player 1 will accept (Lemma 7 in Appendix 1). Thus, Player 1 can "control" Player 2's optimal counteroffer by varying $x_1$. The most favorable counteroffer to Player 1 occurs when $x_1$ is such that $\delta x_1 = g(x_1)$, i.e., $x_1 = \frac{1}{\delta}\hat{x}$ (see Figure 6).

in Appendix 1). Thus, the maximum payoff that Player 1 can obtain by making an offer that will be accepted by Player 2 is the Rubinstein equilibrium payoff. Since $\delta\hat{x} > x^R$, it is optimal for Player 1 to make an offer that will be rejected by Player 2. Therefore, delay in equilibrium occurs.

Notice that even if Player 1 makes the offer $(x^*, f(x^*))$, which is the arbitrator's ideal settlement, Player 2's equilibrium action is to reject Player 1's offer and make the counteroffer $(\delta x^*, f(\delta x^*))$. Moreover, Player 1 will accept the counteroffer in order to avoid the time cost of going to arbitration. However, in equilibrium, Player 1 will not make the offer $(x^*, f(x^*))$, because he can obtain a more favorable counteroffer by making the offer $(\frac{1}{\delta}\hat{x}, f(\frac{1}{\delta}\hat{x}))$ since $\hat{x} > \delta x^*$.

Finally, if the final-offer arbitration rule is sufficiently biased in favor of Player 2, then the equilibrium is an equilibrium with immediate agreement. In addition, Player 1's equilibrium offer is more generous than the Rubinstein equilibrium offer. This is because if Player 1 makes an offer that is *not* more generous than the Rubinstein equilibrium offer, then Player 2 will reject the offer and make a counteroffer, in which Player 2's demand is sufficiently higher than the Rubinstein equilibrium offer. Such a counteroffer is supported by the biased arbitrator, and Player 1 has to accept it. Player 1 is thus better off to make a more generous offer than the Rubinstein equilibrium offer, which Player 2 accepts immediately.

## 3.3 Kalai-Smorodinsky final-offer arbitration

This section studies the final-offer arbitration rule in which the arbitrator's ideal settlement is the Kalai-Smorodinsky solution outcome.

**Definition 3.** *(Kalai and Smorodinsly 1975) The Kalai-Smorodinsky (KS) solution outcome* $(x^{KS}, f(x^{KS}))$ *is the intersection point of the Pareto frontier with the line connecting* $(0,0)$ *and* $(b_1, b_2)$*, i.e.,* $\dfrac{f(x^{KS})}{x^{KS}} = \dfrac{b_2}{b_1}$.

Suppose the arbitrator's utility function is $u^{KS}$, where $u^{KS} : S \to R$ is a continuous, strongly monotone, and quasiconcave function with

$$u^{KS}(x,y)=\begin{cases}(\frac{y/x}{b_2/b_1})^{-1} & if \quad \frac{y/x}{b_2/b_1}\geq 1;\\ \frac{y/x}{b_2/b_1} & if \quad \frac{y/x}{b_2/b_1}<1.\end{cases}$$

for $(x, y) \in PF$.[13] Notice that $u^{KS}$ is maximized at the Kalai-Smorodinsky solution outcome. We have:

**Theorem 4.** *Suppose the final-offer arbitration rule is such that* $h((x_1, y_1), (x_2, y_2)) = (x_2, y_2)$ *if and only if* $u^{KS}(x_2, y_2) \geq u^{KS}(x_1, y_1)$. *Then, as long as* $\frac{f(\frac{1}{\delta}x^R)}{\frac{1}{\delta}x^R} \leq \frac{b_2}{b_1} \leq \frac{f(x^R)}{\delta x^R}$, *the unique SPE outcome of the arbitration game is* $(x^R, f(x^R))$.

Proof: See Appendix 3.2. □

Theorem 4 implies that as long as the line that connects the origin to $(b_1, b_2)$ is above the line that connects the origin to $(\frac{1}{\delta}x^R, f(\frac{1}{\delta}x^R))$ and is below the line that connects the origin to $(\delta x^R, f(x^R))$, then the unique SPE outcome of the alternating-offer arbitration game is $(x^R, f(x^R))$ (see Figure 7).

The condition in Theorem 4 is a sufficient condition for the Rubinstein equilibrium outcome in the $KS$ final-offer arbitration game. More generally, Appendix 3.2 provides the sufficient and necessary condition for the Rubinstein equilibrium outcome and the conditions for other types of equilibrium (type-I arbitration-driven equilibrium and type-II arbitration-driven equilibrium).

## 3.4 The role of the discount factor

This subsection analyzes how the equilibrium payoffs of players change as the discount factor changes.

[13] I do not restrict the function $u^{KS}$ for $(x, y)$ inside the bargaining set, as long as $u^{KS}$ is continuous, strongly monotone, and quasiconcave inside the bargaining set. It is unnecessary because players make offers on the Pareto frontier in equilibrium. Therefore, only the arbitration rule defined on the Pareto frontier matters for the equilibrium outcome.

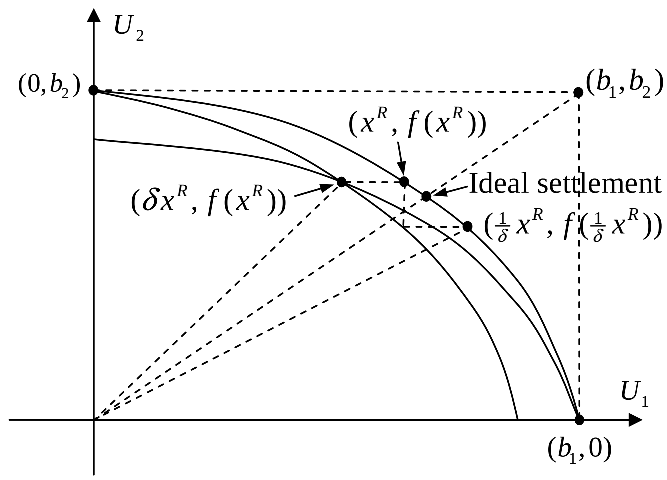


Figure 7: Sufficient condition for the Rubinstein equilibrium outcome in the KS final-offer arbitration game.

I first consider the case where $x^* > x^N$, where $(x^N, f(x^N))$ is the Nash bargaining solution outcome. The following result characterizes the SPE of the game for $\delta$ close to 1.

**Theorem 5.** *Assume that $x^* > x^N$. Let $\overline{\delta}$ be the unique $\delta \in (0,1)$ that satisfies $\hat{x}(\delta) = \frac{1}{\delta}x^R(\delta)$. Let $\overline{\overline{\delta}}$ be the largest $\delta \in [0, \overline{\delta}]$ that satisfies $\hat{x}(\delta) = \delta x^R(\delta)$. We have (i) if $\overline{\delta} < \delta < 1$, then the unique SPE of the alternating-offer arbitration game is a type-II arbitration-driven equilibrium, in which the agreement is delayed, and (ii) if $\overline{\overline{\delta}} \leq \delta \leq \overline{\delta}$, then the unique SPE of the alternating-offer arbitration game is the Rubinstein equilibrium.*

*Sketch of proof:* The threshold discount factor $\overline{\delta}$ exists and is unique due to the following facts (see Figure 8): (i) $\hat{x}(\delta)$ is increasing in $\delta \in (0,1)$; (ii) $\frac{1}{\delta}x^R(\delta)$ is strictly decreasing in $\delta \in (0,1)$;[14] (iii) as $\delta$ approaches 1, $\hat{x}(\delta)$ approaches $x^*$ and $\frac{1}{\delta}x^R(\delta)$ approaches $x^N$ where $x^N < x^*$; and (iv) as $\delta$ approaches 0, $\frac{1}{\delta}x^R(\delta)$ goes to infinity.

Threshold $\overline{\overline{\delta}}$ is also well-defined because there is at least one point ($\delta = 0$) at which $\hat{x}(\delta) = \delta x^R(\delta)$.[15] In addition, since the curve $\delta x^R(\delta)$ is below the curve $\frac{1}{\delta}x^R(\delta)$ for all

[14]See Appendix 2 for the proof.

[15]Figure 8 illustrates the case where the curve $\hat{x}(\delta)$ intersects with the curve $\delta x^R(\delta)$ exactly once, besides at $\delta = 0$. Notice that depending on the shape of the curve $\hat{x}(\delta)$, the curve $\hat{x}(\delta)$ can intersect with the curve $\delta x^R(\delta)$ multiple times besides at $\delta = 0$. Thus, we may have either the Rubinstein equilibrium or type-I arbitration-driven equilibrium for a given $\delta \leq \overline{\delta}$, depending on whether $\hat{x}(\delta) \geq \delta x^R(\delta)$ or $\hat{x}(\delta) < \delta x^R(\delta)$. However, for $\overline{\overline{\delta}} \leq \delta \leq \overline{\delta}$, we must have $\hat{x}(\delta) \geq \delta x^R(\delta)$.

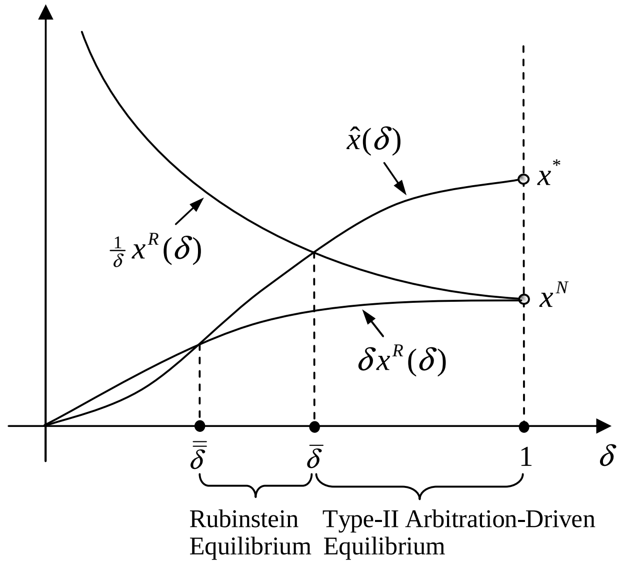


Figure 8: Determination of $\overline{\delta}$ and $\overline{\overline{\delta}}$.

$\delta \in [0, 1)$, it must be true that as $\delta$ decreases from 1 to 0, the curve $\hat{x}(\delta)$ first intersects with the curve $\dfrac{1}{\delta}x^R(\delta)$, and then intersects with the curve $\delta x^R(\delta)$. So, we must have $\overline{\overline{\delta}} < \overline{\delta}$.

Theorem 5 then follows from the above analysis and Theorem 1. □

Figure 9 illustrates the equilibrium payoffs received by players. Figure 9 reveals that the equilibrium payoff of Player 1 is $x^R(\delta)$ for $\overline{\overline{\delta}} \leq \delta \leq \overline{\delta}$ and is $\delta\hat{x}(\delta)$ for $\overline{\delta} < \delta < 1$. Note that when $\delta = \overline{\delta}$, Player 1's payoff obtained from the Rubinstein equilibrium is the same as that obtained from the type-II arbitration-driven equilibrium, i.e., $x^R(\overline{\delta}) = \overline{\delta}\hat{x}(\overline{\delta})$.

As $\delta$ increases, the equilibrium payoff of Player 1 strictly decreases. This implies that Player 1's payoff is smaller when both players are more patient.

In contrast, the equilibrium payoff of Player 2 is $f(x^R(\delta))$ for $\overline{\overline{\delta}} \leq \delta \leq \overline{\delta}$ and is $\delta f(\hat{x}(\delta))$ for $\overline{\delta} < \delta < 1$. Note that $f(x^R(\overline{\delta})) > \overline{\delta} f(\hat{x}(\overline{\delta}))$.

Player 1's equilibrium payoff is continuous in $\delta$ for any $\delta \in [\overline{\overline{\delta}}, 1)$. However, Player 2's equilibrium payoff is discontinuous at $\delta = \overline{\delta}$. At $\delta = \overline{\delta}$, the equilibrium of the game switches from the Rubinstein equilibrium to a type-II arbitration-driven equilibrium. The total equilibrium payoff of the two players shrinks at $\overline{\delta}$. This is because the Rubinstein equilibrium is an equilibrium with immediate agreement, whereas type-II arbitration-driven

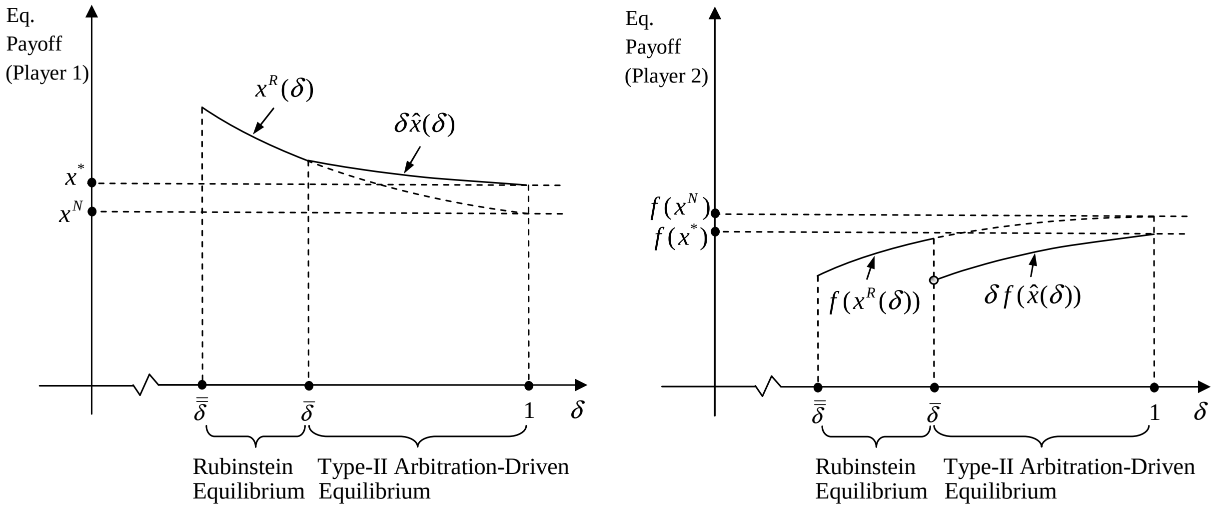


Figure 9: Players' equilibrium payoffs as functions of the discount factor.

equilibrium is an equilibrium with delayed agreement. Player 1, as the player who first makes an offer, is "immune" to the switch between equilibria. However, Player 2 is vulnerable and is subject to a strict payoff loss at $\delta = \bar{\delta}$.

Manzini and Mariotti (2001) also obtain a discontinuity result for the equilibrium payoff. In their game, the equilibrium payoffs of both players are "semidiscontinuous" within a range of discount factors. However, the mechanisms behind the appearance of discontinuity are very different. In their game, the discontinuity appears due to a multiplicity of equilibria. In my game, the discontinuity appears as a result of delay in bargaining.

Finally, when $x^* \leq x^N$, it must be true that $\hat{x}(\delta) < \frac{1}{\delta} x^R(\delta)$ for any $\delta \in (0, 1)$. Thus, for any $\delta \in (0, 1)$, the unique SPE of the arbitration game is either a type-I arbitration-driven equilibrium or the Rubinstein equilibrium. In both types of equilibrium, the agreement is reached immediately. The equilibrium payoffs of both players are thus continuous in the discount factor.

# 4 Conclusion

This paper studies a finite-horizon alternating-offer model that involves final-offer arbitration. I find that there exists a wide range of arbitrator preferences (i.e., the set of balanced final-offer arbitration rules), under which the unique equilibrium outcome of the arbitration game is unaffected by the specific details of the arbitrator's preference. Within this range, the equilibrium outcome coincides with the Rubinstein equilibrium outcome. Outside this range, delay in equilibrium might arise.

A crucial feature of the arbitration game considered in this paper is that the arbitration outcome depends on players' offers. This dependency distinguishes my model from the outside option literature[16] in terms of equilibrium strategies and equilibrium outcomes. The differences include (i) in my model, Player 2's optimal counteroffer depends on Player 1's initial offer, so that Player 1 can control Player 2's counteroffer by varying his own offer; (ii) the condition that yields the Rubinstein equilibrium outcome in my model is different from that obtained by Manzini and Mariotti (2001), and (iii) delayed agreements can occur in my model even though there is always a unique equilibrium.

## Appendix 1: Proof of Theorem 1

I use the following four lemmas (Lemma 6, Lemma 7, Lemma 8 and Lemma 9) to prove Theorem 1.

**Lemma 6.** *If the final offer arbitration rule $h$ is regular (i.e., $h \in \Sigma$), then*

*(i) for any $x_2 \in [g(x_1), x_1]$, we have $h((x_1, f(x_1)), (x_2, f(x_2))) = (x_2, f(x_2))$;*

*(ii) for any $x_2 \in [0, g(x_1))$, we have $h((x_1, f(x_1)), (x_2, f(x_2))) = (x_1, f(x_1))$.*

*Sketch of Proof:* The lemma follows from the definition of $g(x_1)$, and the strong monotonicity of the arbitrator's utility function. Notice that since $x_1 \in V_x(x_1, f(x_1))$, it must be true that $g(x_1) \leq x_1$ and the interval $[g(x_1), x_1]$ is well-defined. $\square$

[16] See the joint outside option model considered by Manzini and Mariotti (2001, 2004) and the unilateral outside option model considered by Binmore et al. (1989), Ponsatí and Sákovics (1998) and Shaked (1994). In the outside option literature, the outside options of players are exogenously given.

According to Lemma 6, for any given offer made by Player 1 $(x_1, f(x_1))$, the best counteroffer that Player 2 could make on the Pareto frontier and the arbitrator would choose is $(g(x_1), f(g(x_1)))$.

The following lemma characterizes Player 2's best counteroffer at Stage 2, generalizing a result in Yildiz (2011).

**Lemma 7.** *In the arbitration game where $h \in \Sigma$, if Player 1 offers $(x_1, y_1) \in PF$ at Stage 1 and Player 2 rejects it, then at Stage 2, in any equilibrium subgame, Player 2 makes the offer* $(\min\{\delta x_1, g(x_1)\}, f(\min\{\delta x_1, g(x_1)\}))$ *and Player 1 accepts it.*

*Proof:* I first show that it is never optimal for Player 2 to make an offer that is strictly inside the bargaining set (i.e., not on the Pareto frontier). To do this, I establish the following two facts. First, it is never optimal for Player 2 to make an offer that is rejected by Player 1. Second, for any Player 2's offer $(x_2, y_2) \notin PF$ that Player 1 would accept, the offer $(x_2, y_2)$ must be strictly dominated by the offer $(x_2, f(x_2))$, which is on the Pareto frontier.

To establish the first point, notice that if Player 2's offer is rejected by Player 1, then it must be true that Player 1's offer is the arbitration outcome. Thus, Player 2 would be strictly better off if he offers $(x_2, y_2) = (x_1, y_1)$, which will be accepted by Player 1 immediately.

To establish the second point, suppose that $(x_2, y_2) \notin PF$ would be accepted by Player 1. Then, it must be true that $h((x_1, y_1), (x_2, y_2)) = (x_2, y_2)$ or $x_2 \geq \delta x_1$ (or both). Since the arbitrator's utility function is strongly monotone, $h((x_1, y_1), (x_2, y_2)) = (x_2, y_2)$ implies that $h((x_1, y_1), (x_2, f(x_2))) = (x_2, f(x_2))$. Thus, if Player 2 offers $(x_2, f(x_2))$, it would also be accepted by Player 1. But then Player 2 obtains a strictly higher payoff by offering $(x_2, f(x_2))$.

It follows that Player 2 never makes an offer that is strictly inside the bargaining set. Without loss of generality, I now restrict Player 2's offer to be on the Pareto frontier. We have the following two possibilities.

(i) Player 2 offers $(x_2, y_2) \in PF$ with $x_2 < \min\{\delta x_1, g(x_1)\}$. If Player 1 accepts the offer, then his payoff is $\delta x_2$ (the payoff is measured at Stage 1). If Player 1 rejects the offer, the game moves to the arbitration stage. Given that $x_2 < g(x_1)$, we must have $h((x_1, y_1), (x_2, y_2)) = (x_1, y_1)$ (by Lemma 6 (ii)). Thus, Player 1's payoff is $\delta^2 x_1$. Since $x_2 < \delta x_1$, we have $\delta x_2 < \delta^2 x_1$, so Player 1 will reject Player 2's offer and Player 2's payoff is $\delta^2 y_1$.

(ii) Player 2 offers $(x_2, y_2) \in PF$ with $x_2 \geq \min\{\delta x_1, g(x_1)\}$. If Player 1 accepts, his payoff is $\delta x_2$. If Player 1 rejects the offer, the game moves to the arbitration stage. If $(x_2, y_2)$ is the arbitrated outcome, then Player 1's payoff is $\delta^2 x_2$, which is less than $\delta x_2$. If instead, $(x_1, y_1)$ is the arbitrated outcome, then Player 1's payoff is $\delta^2 x_1$. In this latter case, since $(x_1, y_1)$ is the arbitrated outcome, we must have either $x_2 < g(x_1)$ or $x_2 \geq x_1$ (by Lemma 6). If $x_2 < g(x_1)$, noting that $x_2 \geq \min\{\delta x_1, g(x_1)\}$, we must have $x_2 \geq \delta x_1$, which implies that $\delta^2 x_1 \leq \delta x_2$. If $x_2 \geq x_1$, then we have $\delta^2 x_1 \leq \delta x_2$. In each of these cases, Player 1 obtains a higher payoff by accepting Player 2's offer. Thus, Player 1 will accept Player 2's offer and Player 2's payoff is $\delta y_2$.[17]

In summary, if Player 2 offers $(x_2, y_2) \in PF$ with $x_2 < \min\{\delta x_1, g(x_1)\}$, then his equilibrium payoff is $\delta^2 y_1 \leq \delta y_1 = \delta f(x_1) < \delta f(\min\{\delta x_1, g(x_1)\})$, where the last inequality follows from the fact that $\min\{\delta x_1, g(x_1)\} \leq \delta x_1 < x_1$.[18] If Player 2 offers $(x_2, y_2) \in PF$ with $x_2 \geq \min\{\delta x_1, g(x_1)\}$, then his equilibrium payoff is $\delta y_2$, which is maximized at $(x_2, y_2) = (\min\{\delta x_1, g(x_1)\}, f(\min\{\delta x_1, g(x_1)\}))$ with the corresponding payoff for Player 2 being $\delta f(\min\{\delta x_1, g(x_1)\})$. Comparing these two cases, it is obvious that Player 2's optimal counteroffer is $(\min\{\delta x_1, g(x_1)\}, f(\min\{\delta x_1, g(x_1)\}))$. Moreover, Player 1 will accept the offer $(\min\{\delta x_1, g(x_1)\}, f(\min\{\delta x_1, g(x_1)\}))$. $\square$

Three factors determine Player 2's best counteroffer at Stage 2: (i) the discount factor $\delta$; (ii) the final-offer arbitration rule $h$; and (iii) Player 1's initial offer $(x_1, f(x_1))$[19].

Define $\tilde{x}(x_1) = \min\{\delta x_1, g(x_1)\}$. Define the *(optimal) counteroffer curve* (of Player 2)

[17] Here, we used tie-breaking rule 1.

[18] Note that the last inequality is strict because $x_2 < \min\{\delta x_1, g(x_1)\}$ implies $x_1 > 0$.

[19] The dependency of Player 2's counteroffer on Player 1's initial offer is a key feature of my arbitration game. This dependency is absent in Rubinstein's infinite-horizon alternating-offer game.

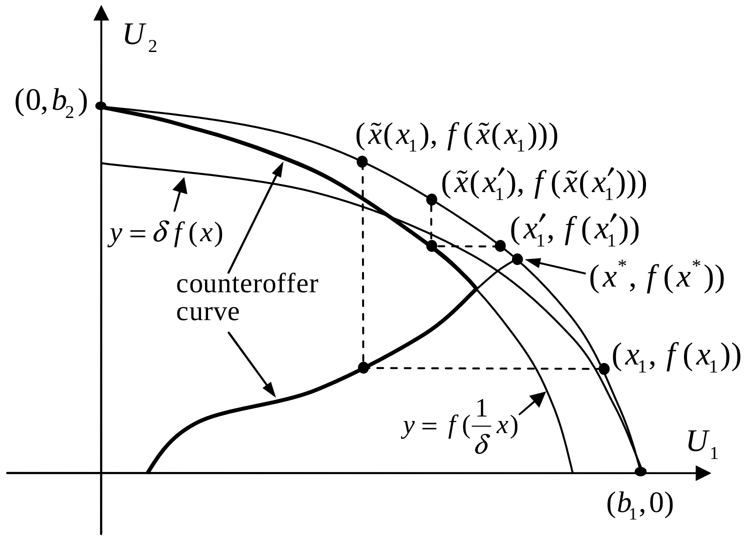


Figure 10: The counteroffer curve of Player 2.

as the curve $x = \tilde{x}(f^{-1}(y))$ where $y \in [0, b_2]$ (see Figure 10). The counteroffer curve is the collection of points $(\tilde{x}(x_1), f(x_1))$ as $x_1$ varies from 0 to $b_1$. The counteroffer curve can be used to determine Player 2's optimal counteroffer for any given Player 1's offer, if Player 2 chooses to reject Player 1's offer. See Figure 10.

The next lemma generalizes Lemma 7. It characterizes Player 2's optimal counteroffer when Player 1's offer $(x_1, y_1)$ is not on the Pareto frontier. Its proof is similar to that of Lemma 7 and is omitted.

Let $g(x_1, y_1) = \min\{x_2 | x_2 \in V_x(x_1, y_1)\}$ and $\tilde{x}(x_1, y_1) = \min\{\delta x_1, g(x_1, y_1)\}$.[20] We have:

**Lemma 8.** *In the arbitration game with $h \in \Sigma$, if Player 1 made an offer $(x_1, y_1)$ at Stage 1 and Player 2 rejected it, then at Stage 2, in any equilibrium subgame, Player 2 makes the offer $(\tilde{x}(x_1, y_1), f(\tilde{x}(x_1, y_1)))$ and Player 1 accepts it.*

The following lemma characterizes the necessary and sufficient conditions for Player 1's offer $(x_1, y_1) \in PF$ to be accepted by Player 2.

**Lemma 9.** *In the arbitration game with $h \in \Sigma$, if Player 1 made an offer $(x_1, y_1) \in PF$ at Stage 1, then in equilibrium:*

[20]Notice that $g(x_1) = g(x_1, f(x_1))$ and $\tilde{x}(x_1) = \tilde{x}(x_1, f(x_1))$.

*(i) If $\delta x^R \leq \hat{x} \leq \delta b_1$, then Player 1's offer $(x_1, y_1)$ is accepted by Player 2 if and only if $0 \leq x_1 \leq x^R$;*

*(ii) If $0 \leq \hat{x} < \delta x^R$, then Player 1's offer $(x_1, y_1)$ is accepted by Player 2 if and only if $0 \leq x_1 \leq f^{-1}(\delta f(\hat{x}))$.*

*Proof:*

*(i) Suppose $\hat{x}$ is such that $\delta x^R \leq \hat{x} \leq \delta b_1$.*

Refer to Figure 11. There are two cases.

(a) At Stage 1, Player 1 makes an offer $(x_1, f(x_1))$ with $x_1 > x^R$. If Player 2 accepts the offer, then his payoff is $f(x_1)$. If Player 2 rejects the offer, then at Stage 2, from Lemma 7, he offers $(\tilde{x}(x_1), f(\tilde{x}(x_1)))$ that Player 1 accepts; Player 2's payoff is $\delta f(\tilde{x}(x_1))$. We have $\delta f(\tilde{x}(x_1)) \geq \delta f(\delta x_1) > f(x_1)$, where the first inequality follows from the fact that $\tilde{x}(x_1) = \min\{\delta x_1, g(x_1)\} \leq \delta x_1$ and the second inequality follows from the fact that $x_1 > x^R$ (see Figure 11). So, Player 2 rejects the offer $(x_1, f(x_1))$ and makes the counteroffer $(\tilde{x}(x_1), f(\tilde{x}(x_1)))$, which Player 1 accepts.

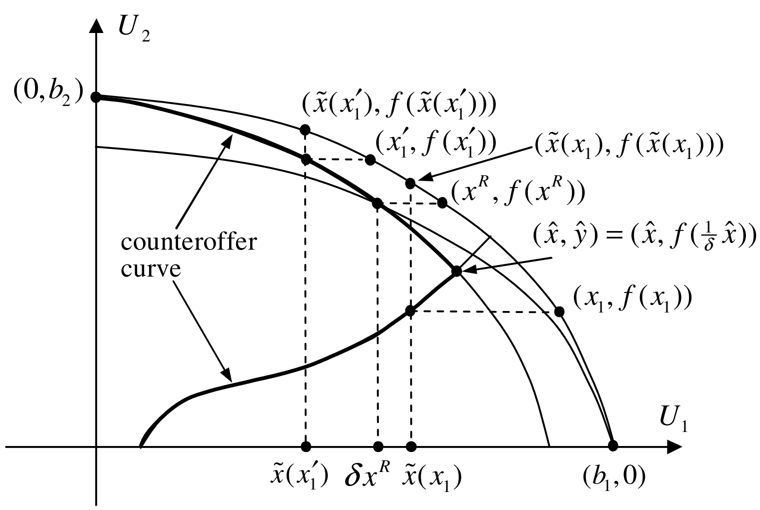


Figure 11: The case of $\delta x^R \leq \hat{x} \leq \delta b_1$.

(b) At Stage 1, Player 1 makes an offer $(x_1', f(x_1'))$ with $0 \leq x_1' \leq x^R$. If Player 2 accepts the offer, then his payoff is $f(x_1')$. If Player 2 rejects the offer, then at Stage 2, from Lemma 7, he offers $(\tilde{x}(x_1'), f(\tilde{x}(x_1')))$, which Player 1 accepts. Thus, Player 2's payoff is $\delta f(\tilde{x}(x_1')) = \delta f(\delta x_1') \leq f(x_1')$ (see Figure 11). Consequently, Player 2 accepts $(x_1', f(x_1'))$.

(ii) *Suppose $\hat{x}$ is such that* $0 \leq \hat{x} < \delta x^R$.

Refer to Figure 12. There are two cases.

(a) At Stage 1, Player 1 makes an offer $(x_1, f(x_1))$ with $x_1 > f^{-1}(\delta f(\hat{x}))$. If Player 2 accepts the offer, then his payoff is $f(x_1)$. If Player 2 rejects the offer, then at Stage 2, from Lemma 7, he offers $(\tilde{x}(x_1), f(\tilde{x}(x_1)))$ that Player 1 accepts. Thus, Player 2's payoff is $\delta f(\tilde{x}(x_1))$. We have $\delta f(\tilde{x}(x_1)) \geq \delta f(\hat{x}) > f(x_1)$, where the first inequality follows from the fact that $\tilde{x}(x_1) = \min\{\delta x_1, g(x_1)\} = g(x_1) \leq \hat{x}$ (see Figure 12) and the second inequality follows from the fact that $x_1 > f^{-1}(\delta f(\hat{x}))$. Therefore, Player 2 rejects the offer $(x_1, f(x_1))$ and makes the counteroffer $(\tilde{x}(x_1), f(\tilde{x}(x_1)))$, which Player 1 accepts.

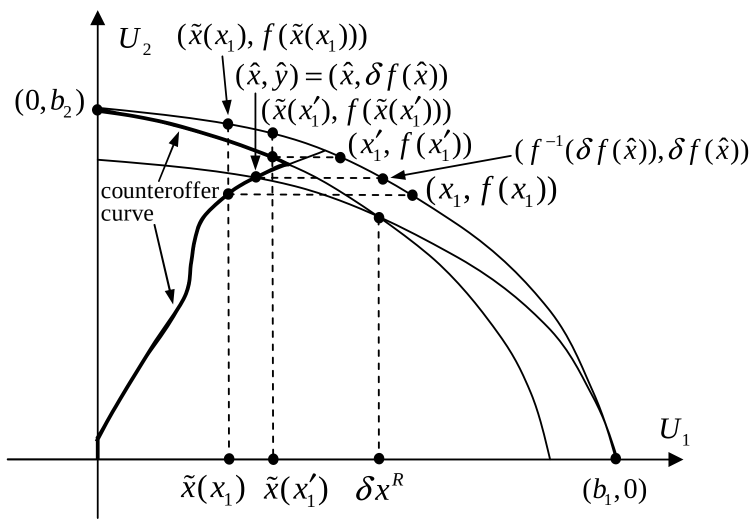


Figure 12: The case of $0 \leq \hat{x} < \delta x^R$.

(b) At Stage 1, Player 1 makes an offer $(x_1', f(x_1'))$ with $0 \leq x_1' \leq f^{-1}(\delta f(\hat{x}))$. If Player 2 accepts the offer, then his payoff is $f(x_1')$. If Player 2 rejects the offer, then at Stage 2, from Lemma 7, he offers $(\tilde{x}(x_1'), f(\tilde{x}(x_1')))$, which Player 1 accepts; Player 2's payoff is $\delta f(\tilde{x}(x_1'))$. Since $\delta f(\tilde{x}(x_1')) \leq f(x_1')$ (see Figure 12), Player 2 will accept the offer $(x_1', f(x_1'))$. $\square$

Now, we can state the proof of Theorem 1.

**Proof of Theorem 1:**

I first show that it is never optimal for Player 1 to make an offer that is strictly inside the bargaining set (i.e., not on the Pareto frontier). In particular, I show that if Player 1 offers $(x_1, y_1) \notin PF$, then the offer is strictly dominated by the offer $(f^{-1}(y_1), y_1)$.

Suppose $(x_1, y_1)$ would be accepted by Player 2. Then, it must be true that $y_1 \geq \delta f(\tilde{x}(x_1, y_1))$. Since the arbitrator's utility function is strongly monotone, we have $g(f^{-1}(y_1), y_1) > g(x_1, y_1)$ and $\tilde{x}(f^{-1}(y_1), y_1) > \tilde{x}(x_1, y_1)$. So, $y_1 \geq \delta f(\tilde{x}(x_1, y_1))$ implies that $y_1 \geq \delta f(\tilde{x}(f^{-1}(y_1), y_1))$. Thus, if Player 1 makes the offer $(f^{-1}(y_1), y_1)$, then it will also be accepted by Player 2. Player 1 thus obtains a strictly higher payoff by offering $(f^{-1}(y_1), y_1)$.

Suppose $(x_1, y_1)$ would be rejected by Player 2. Then, Player 1 must obtain a payoff of $\delta\tilde{x}(x_1, y_1)$. If Player 1 makes the offer $(f^{-1}(y_1), y_1)$, then it might be accepted by Player 2, or rejected by Player 2. If Player 2 accepts the offer, then Player 1 obtains a payoff of $f^{-1}(y_1) > x_1 \geq \delta\tilde{x}(x_1, y_1)$, where the last inequality follows from the fact that $\delta\tilde{x}(x_1, y_1) \leq \delta(\delta x_1) \leq x_1$. If Player 2 rejects the offer, then Player 1 obtains a payoff of $\delta\tilde{x}(f^{-1}(y_1), y_1)$, which is strictly greater than $\delta\tilde{x}(x_1, y_1)$.

In all the cases analyzed above, Player 2 is strictly better off by making the offer $(f^{-1}(y_1), y_1)$. I thus showed that Player 1 never makes an offer that is strictly inside the bargaining set. In the remainder of the proof, I assume that Player 1 can only make an offer on the Pareto frontier. There are three cases.

(i) Suppose $\hat{x}$ is such that $\delta x^R \leq \hat{x} \leq \frac{1}{\delta}x^R$.

If Player 1 makes an offer $(x_1, f(x_1))$ with $x_1 > x^R$, then Player 2 will reject it and make the counteroffer $(\tilde{x}(x_1), f(\tilde{x}(x_1)))$, which Player 1 will accept (by Lemma 9 (i) and Lemma 7). Player 1's payoff is thus $\delta\tilde{x}(x_1)$, which is maximized at $x_1 = \frac{1}{\delta}\hat{x}$ with corresponding payoff $\delta\tilde{x}(\frac{1}{\delta}\hat{x}) = \delta\hat{x}$.[21] If Player 1 makes an offer $(x_1', f(x_1'))$ with $0 \leq x_1' \leq x^R$, then Player 2 will accept it (by Lemma 9 (i)). Player 1's payoff is $x_1'$, which is maximized at $x_1' = x^R$ with

[21]This is true except when $\hat{x} = \delta x^R$, where the maximum of $\delta\tilde{x}(x_1)$ does not exist. However, in this case, one can show that for Player 1, any offer $(x_1, f(x_1))$ with $x_1 > x^R$ must be strictly dominated by the offer $(x^R, f(x^R))$, and thus Player 1 never makes an offer $(x_1, f(x_1))$ with $x_1 > x^R$.

corresponding payoff $x^R$. Since $\delta\hat{x} \le x^R$, Player 1's optimal strategy is to offer $(x^R, f(x^R))$ at Stage 1.[22] Player 2 accepts $(x^R, f(x^R))$ immediately.

(ii) Suppose $\hat{x}$ is such that $\frac{1}{\delta}x^R < \hat{x} \le \delta b_1$.

If Player 1 makes an offer $(x_1, f(x_1))$ with $x_1 > x^R$, then Player 2 will reject the offer $(x_1, f(x_1))$ and make the counteroffer $(\tilde{x}(x_1), f(\tilde{x}(x_1)))$, which Player 1 will accept (by Lemma 9 (i) and Lemma 7). Player 1's payoff is thus $\delta\tilde{x}(x_1)$, which is maximized at $x_1 = \frac{1}{\delta}\hat{x}$ with corresponding payoff $\delta\tilde{x}(\frac{1}{\delta}\hat{x}) = \delta\hat{x}$. If Player 1 offers $(x'_1, f(x'_1))$ with $0 \le x'_1 \le x^R$, then Player 2 will accept it (by Lemma 9 (i)). Player 1's payoff is $x'_1$, which is maximized at $x'_1 = x^R$ with corresponding payoff $x^R$. Since $\delta\hat{x} > x^R$, Player 1's optimal strategy is to offer $(\frac{1}{\delta}\hat{x}, f(\frac{1}{\delta}\hat{x}))$. Player 2 will reject the offer and make the counteroffer $(\tilde{x}(\frac{1}{\delta}\hat{x}), f(\tilde{x}(\frac{1}{\delta}\hat{x}))) = (\hat{x}, f(\hat{x}))$, which Player 1 will accept.

(iii) Suppose $\hat{x}$ is such that $0 \le \hat{x} < \delta x^R$.

If Player 1 makes an offer $(x_1, f(x_1))$ with $x_1 > f^{-1}(\delta f(\hat{x}))$, then Player 2 will reject the offer $(x_1, f(x_1))$ and make the counteroffer $(\tilde{x}(x_1), f(\tilde{x}(x_1)))$, which Player 1 will accept (by Lemma 9 (ii) and Lemma 7). Player 1's payoff is thus $\delta\tilde{x}(x_1)$, which is at most $\delta\hat{x}$ (see Figure 12). If Player 1 makes an offer $(x'_1, f(x'_1))$ with $0 \le x'_1 \le f^{-1}(\delta f(\hat{x}))$, then Player 2 will accept the offer (by Lemma 9 (ii)). Player 1's payoff is $x'_1$, which is maximized at $x'_1 = f^{-1}(\delta f(\hat{x}))$ with corresponding payoff $f^{-1}(\delta f(\hat{x}))$. Since $f^{-1}(\delta f(\hat{x})) > \delta\hat{x}$ (see Figure 12), Player 1's optimal strategy is to offer $(x_1, f(x_1))$ where $x_1 = f^{-1}(\delta f(\hat{x}))$. Player 2 will accept the offer immediately. □

# Appendix 2: Other proof

**Proof of the statement that** $\dfrac{1}{\delta}x^R(\delta)$ **is strictly decreasing in** $\delta \in (0,1)$ **and** $\delta x^R(\delta)$ **is strictly increasing in** $\delta \in (0,1)$**:**

[22] If $\delta\hat{x} = x^R$, Player 1 is indifferent between offering $(x^R, f(x^R))$ and offering $(\frac{1}{\delta}\hat{x}, f(\frac{1}{\delta}\hat{x}))$. Using tie-breaking rule 2, Player 1 must offer $(x^R, f(x^R))$.

I first show that $\frac{1}{\delta}x^R(\delta)$ is strictly decreasing in $\delta \in (0, 1)$. It is sufficient to show that $x^R(\delta)$ is strictly decreasing in $\delta \in (0, 1)$.

Notice that $x^R(\delta)$ satisfies $\delta f(\delta x^R(\delta)) = f(x^R(\delta))$. Differentiating with respect to $\delta$ and rearranging terms yields:

$$(f'(x^R(\delta)) - \delta^2 f'(\delta x^R(\delta)))x^{R'} = f(\delta x^R(\delta)) + \delta x^R(\delta) f'(\delta x^R(\delta)). \tag{1}$$

Using the fact that $f$ is concave and strictly decreasing, we have:

$$f'(x^R(\delta)) - \delta^2 f'(\delta x^R(\delta)) < f'(x^R(\delta)) - f'(\delta x^R(\delta)) \leq 0. \tag{2}$$

Notice that the Nash bargaining solution $(x^N, f(x^N))$ satisfies $f'(x^N) = -\frac{f(x^N)}{x^N}$, i.e., $f(x^N) + f'(x^N)x^N = 0$. Since $\delta x^R f(\delta x^R) = x^R f(x^R)$, the two points $(\delta x^R(\delta), f(\delta x^R(\delta)))$ and $(x^R(\delta), f(x^R(\delta)))$ must lie on the curve $xy = c$ with the same constant $c$. Thus, it must be true that $\delta x^R(\delta) < x^N$ for any $\delta \in (0, 1)$. Then,

$$f(\delta x^R(\delta)) + \delta x^R(\delta) f'(\delta x^R(\delta)) > f(x^N) + f'(x^N)x^N = 0 \tag{3}$$

where the inequality follows from the facts that $\delta x^R(\delta) < x^N$ and that $f$ is a concave and strictly decreasing function.

Now, using equations 1, 2, 3, we have $x^{R'} < 0$, i.e., $x^R$ is strictly decreasing in $\delta \in (0, 1)$. We thus proved that $\frac{1}{\delta}x^R(\delta)$ is strictly decreasing in $\delta \in (0, 1)$.

Finally, $\delta x^R(\delta)$ is strictly increasing in $\delta \in (0, 1)$. This follows from the following two facts: (i) the point $(\delta x^R(\delta), f(\delta x^R(\delta)))$ and the point $(x^R(\delta), f(x^R(\delta)))$ lie on the same "indifference curve" $xy = c$, and (ii) $x^R(\delta)$ is strictly decreasing in $\delta \in (0, 1)$. $\square$

# Appendix 3: Additional results

In this appendix, I present additional results regarding the Nash final-offer arbitration game and the KS final-offer arbitration game.

## Appendix 3.1: Robustness of the Nash final-offer arbitration rule

Yildiz (2011) shows that the unique SPE outcome of the arbitration game that uses the Nash final-offer arbitration rule is $(x^R, f(x^R))$. Here, we provide a simple proof of this result based on Theorem 1.

**Theorem 10.** *(Yildiz 2011) In the alternating-offer arbitration game where $h$ is the Nash final-offer arbitration rule, the unique SPE outcome is $(x^R, f(x^R))$ for any given $\delta < 1$.*

*Proof:* Given that $x^R f(x^R) = \delta x^R f(\delta x^R)$, we have $g(x^R) = \delta x^R$. That is, $g(f^{-1}(f(x^R))) = \delta x^R$, which implies that the point $(\delta x^R, f(x^R))$ is on the arbitration curve. Since $(\delta x^R, f(x^R))$ is also on the discounted Pareto frontier, $(\delta x^R, f(x^R))$ is exactly the intersection point of the arbitration curve and the discounted Pareto frontier and we must have $\hat{x}(\delta) = \delta x^R$. Thus, the Nash final-offer arbitration rule is balanced for any discount factor. According to Theorem 1 (i), the unique SPE outcome of the arbitration game that uses the Nash final-offer arbitration rule is $(x^R, f(x^R))$. $\square$

Although for a *given* $\delta$, there exists a class of final-offer arbitration rules under which the arbitration game yields an SPE outcome of $(x^R, f(x^R))$, the following theorem shows that if we require the arbitration game to yield an SPE outcome of $(x^R, f(x^R))$ for *any* $\delta < 1$, then we must have $(x^*, f(x^*)) = (x^N, f(x^N))$ where $(x^N, f(x^N)) = \arg\max\{xy|(x, y) \in S\}$. That is, the arbitrator's ideal settlement must be the Nash solution outcome.

**Theorem 11.** *In the arbitration game where $h \in \Sigma$, if the SPE outcome is $(x^R, f(x^R))$ for any $\delta < 1$, then we must have $(x^*, f(x^*)) = (x^N, f(x^N))$.*

*Sketch of Proof:* The theorem follows from the following facts: (i) as $\delta$ approaches 1, both $\delta x^R$ and $\frac{1}{\delta}x^R$ converge to $x^N$ (see Binmore et al. (1986)); (ii) as $\delta$ approaches 1, $\hat{x}$ converges to $x^*$; and (iii) if the SPE outcome of the arbitration game is $(x^R, f(x^R))$ for any $\delta < 1$, then we must have $\delta x^R \leq \hat{x} \leq \frac{1}{\delta}x^R$ for any $\delta < 1$. $\square$

## Appendix 3.2: Additional results for the KS final-offer arbitration game

**Theorem 12.** *Suppose the final-offer arbitration rule is such that* $h((x_1, y_1), (x_2, y_2)) = (x_2, y_2)$ *if and only if* $u^{KS}(x_2, y_2) \geq u^{KS}(x_1, y_1)$ *(*$u^{KS}$ *is defined in Section 3.3). Then,*

*(i) if* $\dfrac{\sqrt{\delta f(\frac{1}{\delta^2}x^R) f(\frac{1}{\delta}x^R)}}{\frac{1}{\delta}x^R} \leq \dfrac{b_2}{b_1} \leq \dfrac{f(x^R)}{\delta x^R}$*, then the unique SPE outcome of the final-offer arbitration game is* $(x^R, f(x^R))$*.*

*(ii) if* $\dfrac{b_2}{b_1} < \dfrac{\sqrt{\delta f(\frac{1}{\delta^2}x^R) f(\frac{1}{\delta}x^R)}}{\frac{1}{\delta}x^R}$*, then the unique SPE of the final-offer arbitration game is a type-II arbitration-driven equilibrium. That is, at Stage 1, Player 1 makes the offer* $(\frac{1}{\delta}\hat{x}, f(\frac{1}{\delta}\hat{x}))$ *that Player 2 rejects, and at Stage 2, Player 2 makes the counteroffer* $(\hat{x}, f(\hat{x}))$ *that Player 1 accepts. Moreover,* $\hat{x}$ *is determined by the equality that* $\dfrac{\hat{x}/f(\hat{x})}{b_1/b_2} = \dfrac{f(\frac{1}{\delta}\hat{x})/\frac{1}{\delta}\hat{x}}{b_2/b_1}$*.*

*(iii) if* $\dfrac{b_2}{b_1} > \dfrac{f(x^R)}{\delta x^R}$*, then the unique SPE of the final-offer arbitration game is a type-I arbitration-driven equilibrium. That is, Player 1 offers* $(f^{-1}(\delta f(\hat{x})), \delta f(\hat{x}))$ *that Player 2 accepts. Moreover,* $\hat{x}$ *is determined by the equality that* $\dfrac{\hat{x}/f(\hat{x})}{b_1/b_2} = \dfrac{\delta f(\hat{x})/f^{-1}(\delta f(\hat{x}))}{b_2/b_1}$*.*

*Sketch of proof:* First, $\hat{x} \geq \delta x^R$ is equivalent to $g(x^R) \geq \delta x^R$. The condition $g(x^R) \geq \delta x^R$ (i.e., the arbitrator weakly prefers $(x^R, f(x^R))$ over $(\delta x^R, f(\delta x^R))$) is satisfied if and only if (i) $\dfrac{f(x^R)/x^R}{b_2/b_1} < 1$, $\dfrac{f(\delta x^R)/\delta x^R}{b_2/b_1} > 1$, and $\dfrac{f(x^R)/x^R}{b_2/b_1} \geq (\dfrac{f(\delta x^R)/\delta x^R}{b_2/b_1})^{-1}$, or (ii) $\dfrac{f(x^R)/x^R}{b_2/b_1} \geq 1$. The condition (i) is equivalent to $\dfrac{b_2}{b_1} > \dfrac{f(x^R)}{x^R}$, $\dfrac{b_2}{b_1} < \dfrac{f(\delta x^R)}{\delta x^R}$, and $\dfrac{b_2}{b_1} \leq \dfrac{f(x^R)}{\delta x^R}$ (for the last inequality, using the fact that $\delta f(\delta x^R) = f(x^R)$), which is equivalent to $\dfrac{f(x^R)}{x^R} < \dfrac{b_2}{b_1} \leq \dfrac{f(x^R)}{\delta x^R}$. The condition (ii) is equivalent to $\dfrac{b_2}{b_1} \leq \dfrac{f(x^R)}{x^R}$. Thus, the condition (i) or (ii) is satisfied if and only if $\dfrac{b_2}{b_1} \leq \dfrac{f(x^R)}{\delta x^R}$.

Second, $\hat{x} \leq \frac{1}{\delta}x^R$ is equivalent to $g(\dfrac{1}{\delta^2}x^R) \leq \dfrac{1}{\delta}x^R$. The condition $g(\dfrac{1}{\delta^2}x^R) \leq \dfrac{1}{\delta}x^R$ is equivalent to $\dfrac{b_2}{b_1} \geq \dfrac{\sqrt{\delta f(\frac{1}{\delta^2}x^R)f(\frac{1}{\delta}x^R)}}{\frac{1}{\delta}x^R}$ (the proof here is similar to the first step and is thus omitted).

Theorem 12 then follows from the above analysis and Theorem 1. $\square$

Now, we can prove Theorem 4.

**Proof of Theorem 4:**

By Theorem 12, if $\dfrac{\sqrt{\delta f(\frac{1}{\delta^2}x^R)f(\frac{1}{\delta}x^R)}}{\frac{1}{\delta}x^R} \leq \dfrac{b_2}{b_1} \leq \dfrac{f(x^R)}{\delta x^R}$, then the unique SPE outcome of the KS final-offer arbitration game is $(x^R, f(x^R))$. A sufficient condition for $\dfrac{\sqrt{\delta f(\frac{1}{\delta^2}x^R)f(\frac{1}{\delta}x^R)}}{\frac{1}{\delta}x^R} \leq \dfrac{b_2}{b_1} \leq \dfrac{f(x^R)}{\delta x^R}$ is $\dfrac{f(\frac{1}{\delta}x^R)}{\frac{1}{\delta}x^R} \leq \dfrac{b_2}{b_1} \leq \dfrac{f(x^R)}{\delta x^R}$. This follows from the fact that $f(\frac{1}{\delta^2}x^R) \leq f(\frac{1}{\delta}x^R)$ and $\delta < 1$. $\square$

# References


K. Binmore, A. Rubinstein, A. Wolinsky, *The Nash bargaining solution in economic modelling*, Rand Journal of Economics 17 (1986), 176-188.

K. Binmore, A. Shaked, J. Sutton, *An outside option experiment*, Quarterly Journal of Economics 104 (1989), 753-770.

J. R. Chelius, J. B. Dworkin, *An economic analysis of final-offer arbitration as a conflict resolution device*, Journal of Conflict Resolution 24 (1980), 293-310.

C. A. Deck, A. Farmer, *Bargaining over an uncertain value: arbitration mechanisms compared*, Journal of Law, Economics, and Organization 23 (2007), 547-579.

E. Kalai, M. Smorodinsky, *Other solutions to Nash's bargaining problem*, Econometrica 43 (1975), 513-518.

P. Manzini, M. Mariotti, *Perfect equilibria in a model of bargaining with arbitration*, Games and Economic Behavior 37 (2001), 170-195.

P. Manzini, M. Mariotti, *Going alone together: joint outside options in bilateral negotiations*, Economic Journal 114 (2004), 943-960.

J. Nash, *The bargaining problem*, Econometrica 18 (1950), 155-162.

C. Ponsatí, J. Sákovics, *Rubinstein bargaining with two-sided outside options*, Economic Theory 11 (1998), 667-672.

K. Rong, *An axiomatic approach to arbitration and its application in bargaining games*, The B.E. Journal of Theoretical Economics (Contributions), forthcoming (2012).

A. Rubinstein, *Perfect equilibrium in a bargaining model*, Econometrica 50 (1982), 97-110.

A. Shaked, *Opting out: bazaars versus "hi tech" markets*, Investigaciones Ecónómics 18 (1994), 421-432.

C. Stevens, *Is compulsory arbitration compatible with bargaining?* Industrial Relations 5 (1966), 38-52.

R. Wilson, *Negotiation with private information: litigation and strikes*, Nancy L Schwartz Memorial Lecture (1994), Northwestern University.

M. Yildiz, *Nash meets Rubinstein in final-offer arbitration*, Economics Letters 110 (2011), 226-230.