A Welfare Analysis of Arbitration*

Wojciech Olszewski†

April 2010

Abstract The paper compares conventional and final-offer arbitration from a welfare perspective. By an existing contract, one party is supposed to make a payment to another party; the amount of this payment is supposed to depend on the state of the world. This state is, however, unknown. Two scenarios are considered. Under one of them, one party may have a private signal about the state which cannot be credibly revealed to the other party. The ranking of the two arbitration procedures then depends on the assumptions regarding the arbitrator. If the arbitrator's ability to recognize private signals is high, conventional arbitration dominates final-offer arbitration in the sense that the probability of filing a request for arbitration may be lower under the former form of arbitration than under the latter form. If the arbitrator's ability recognize the validity of private signals is low, final-offer arbitration dominates conventional arbitration in a quite similar sense. Under the second scenario, both parties believe that their opponents have wrong signals about the state of the world. In that case, conventional arbitration approximates the existing contractual arrangement better than final-offer arbitration.

*The author would like to thank the editor and two referees for very helpful comments and suggestions.

†Department of Economics, Northwestern University, Evanston IL 60208, wo@northwestern.edu

1 Introduction

Arbitration by third-party neutrals has become an often-used method of conflict resolution. It is frequently prescribed to resolve labor-management disputes when the labor unions are legally prohibited from striking, as in public services such as police and fire protection. The use of arbitration in the settlement of disputes under existing contracts includes: buyers and sellers in commercial contracts, baseball players and club owners, and divorce settlements.[1]

Various compulsory-arbitration schemes are, or have been, in use in many states, yet other schemes have also been proposed. The two schemes used most often are *conventional arbitration*, where the arbitrator is free to impose any settlement that she wishes; and *final-offer arbitration*, where the arbitrator is constrained to choose one of the final (also called last) offers of the disputing parties, without any possibility of compromise. Final-offer arbitration is used in the major league baseball, while conventional arbitration prevails in bargaining under commercial contracts; Lester (1984) discusses in detail the use of both procedures in public-service collective bargaining across different states.

Historically, the welfare analysis of both these forms of arbitration dates back to very early days of arbitration. A number of observers of early experience with conventional arbitration (see Stevens (1966), Feuille (1975), and Feigenbaum (1975)) reported that the arbitrators had a tendency of splitting the difference between the positions of the parties. In the literature, this was named the *"chilling" effect* in bargaining. The chilling effect implied - according to these observers - an excessive reliance on arbitration. And final-offer arbitration was proposed by Stevens (1966) as a remedy designed to counteract the chilling effect and reduce the reliance on arbitration.

The reliance on arbitration, or the percentage of conflicts that end up with an arbitration award, is the welfare criterion most emphasized in the existing literature.

[1] Lester (1984) and Najita and Stern (2001) summarize actual experience with binding arbitration of collective bargaining in public services, and Dworkin (1986) provides a review of highlights and controversies surrounding salary arbitration in baseball.

This reflects the view that a quality system makes parties reach a settlement without using the system. In particular, parties typically incur a cost of arbitration, and it is commonly believed that an agreement for which both parties are responsible is likely to be better for their future relations that one imposed by a binding decision of another party. Other welfare criteria include the fairness (or appropriateness) of the arbitrator's awards or the settlements reached by parties themselves, and their freedom from biases. Again, see Lester (1984) for a discussion of the aims that were to be achieved by state laws providing for arbitration of negotiating impasses.

A number of researchers analyze theoretically or experimentally the reliance on arbitration under different arbitration procedures in a setting with symmetric information regarding the arbitrator's decision. (See Crawford (1979), Farber and Katz (1979), Farber (1980a) and (1980b), Brams and Merrill III (1983), Ashenfelter et al. (1992), Dickinson (2004), and Deck and Farmer (2007).) However, most empiricists, such as Farber (1980a) or Lester (1984), emphasize the differences in parties' information as the major explanation for the failure of collective bargaining and relying on third-party decisions.

Chatterjee (1981) and Samuelson (1991) are probably the earliest attempts at comparing different arbitration procedures under incomplete, asymmetric information. More recently, Farmer and Pecorino (1998) and (2003) (see also Deck and Farmer (2003)) study a game-theoretic model in which one party has private information about the expected outcome in arbitration . This information can potentially be shared with the other party. However, the privately-informed party may wish instead to take advantage of its information in the process of submitting final offers, which may impede settlement in collective bargaining (in comparison to conventional arbitration).[2]

I also assume here that the differences in parties' information (or beliefs) about the expected outcome of arbitration are the major cause of the failure of collective bargaining. However, I question the premise that the strategic concealment of in-

[2] Farmer and Pecorino also argue that the impediments in collective bargaining disappear if bargaining is allowed to take place after final offers have been submitted to the arbitrator.

formation plays the key role. After all, in the early stages of most conflicts, parties negotiate in the hope of resolving the conflict by themselves.

The key innovation of the present model is introducing a new kind of information (or beliefs) that negotiating parties may have. I distinguish two types of signals that may not be shared with other parties, and which seem to be common in practice. First, arguments that are persuasive to one party may not be persuasive to the other party.[3] In such a case, the latter party may not take the arguments of the former party into account, especially in a world in which arguments can also be made up. In addition, parties may have noncommon prior beliefs. Intuitively, each party may believe that the other party is simply wrong.

The key insight of this paper is that the welfare ranking of different arbitration procedures depends crucially on whether the failure of face-to-face bargaining is caused by the presence of arguments which are persuasive to one party but are not persuasive to the other party, or the existence of non-common prior beliefs.

Suppose that the arbitrator has a certain ability to recognize whether the arguments are real and fake. If this ability is low, and the failure of negotiations is caused by unpersuasive arguments, then final-offer arbitration dominates conventional arbitration. If this ability is high, conventional arbitration is better. Conventional arbitration dominates final-offer arbitration also under noncommon prior beliefs, independently of the arbitrator's ability of recognizing real and fake arguments.

To be more precise, consider the presence of unpersuasive arguments first. If the arbitrator's ability of recognizing arguments is low, the chance of rejecting a settlement payment in face-to-face negotiations may be lower in the final-offer arbitration game. The intuition for this result relies on the fact that the arbitrator can interpret final offers as signals. The arbitrator may believe, in the final-offer arbitration game, that a party whose argument is real (and not fake) will make a final offer that would

[3]For example, I think that my paper provides valuable insights; otherwise, I would not bother writing it. A reader may disagree. And each of us may well be unable to persuade the other party.

Then, no matter what the actual value of my insights is, we face a situation in which parties fail to share their signals (or beliefs).

not be optimal if it were not interpreted as a signal about that party's argument. This belief of the arbitrator reduces the party's payoff to filing an arbitration request, and makes the party more willing to accept lower settlement payments.

If the arbitrator's discriminatory ability is high, the chance of rejecting a settlement payment in face-to-face bargaining may be lower in the conventional arbitration game. The intuition is that privately-informed parties, in the final-offer arbitration game, can extract informational rents, as their opponent have to make final offers with an inferior information regarding the arbitrator's decisions. Thus, the privately-informed parties demand higher settlement payments.

Under non-common prior beliefs, the chance of filing a request for arbitration is the same under both arbitration procedures, because if parties believe that their opponents are wrong, and that the arbitrator will find out that their opponents are wrong, then the parties are unable to reach an agreement in face-to-face negotiations, no matter which arbitration procedure is in use.

However, suppose that by an existing contract, one party is supposed to make a payment to the other party; and the amount of this payment is supposed to depend on the state of the world. Then, the outcomes of conventional arbitration approximate the existing contractual arrangement better than the outcomes of final-offer arbitration. That is, conventional arbitration dominates final-offer arbitration in terms of the appropriateness of the arbitrator's awards or their freedom from biases. Indeed, when each party has a signal in favor of the outcome that it prefers, and believes both that its opponent has a wrong signal, and that the arbitrator will find out that its opponent has a wrong signal, then some sort of dutch-book argument applies. Parties make "exaggerated" final offers trying to take advantage of their opponents being (in their opinion) wrong, and the "exaggerated" final offers move the arbitration outcome away from the existing contractual arrangement.

These results have been summarized in Table 1.

Let us finally try to relate the present analysis to the existing evidence. Dworkin (1986) reports that in the major league baseball, in which final-offer arbitration is in use, most cases are settled short of arbitration; cases in which an arbitration request ends up being filed are often characterized by widely divergent salary positions between the parties (with a similar number of winners on each side). The case of the major league baseball is particularly interesting for the present paper, because there seems to be more room for noncommon prior beliefs of baseball players and club owners than for possibly unpersuasive arguments. The version of my model with non-common prior beliefs indeed predicts widely divergent final offers. The model makes no prediction regarding the frequency of cases that are settled short of arbitration, and suggests that under conventional arbitration, the salaries would reflect better the value of players for their teams.[4]

The rich data sets from the public-service sector cannot unfortunately be clearly related to the predictions of the model. For example, Lester (1984) reports that conventional arbitration awards averaged around 29 percent of total negotiations in 1960s and 1970s in Philadelphia, while in the New York State, where conventional arbitration was also in use, awards declined from 15.9 percent of total negotiations to 9.4 percent between 1975-1976 and 1982-1983. For comparison, the average of final-offer arbitration awards to total negotiations in Michigan was 16.4 percent in the period 1973-1974 through 1976-1977, and 6.7 percent for the six-year period 1977-1978 through 1982-1983, with the figure for 1982-1983 rising to 12 percent.

2 Model

There are two equally likely states of the world, a and b, two risk-neutral interacting agents, A (female) and B (male), and an arbitrator. According to an existing contract, agent A is supposed to make a payment of \$1 to agent B, contingent on state b, but no payment is due in state a. I analyze the following two information structures.

The first is that of possibly *unpersuasive arguments (UA)*: Under this scenario,

[4] One should, however, be cautious in referring to this evidence, because (as reported by Dworkin) players who lose their cases in arbitration often gain substantial salary increases afterwards; this suggests that, in practice, collective bargaining in baseball may depart slightly from the final-offer scheme.

agent B knows the state of the world, but agent A knows it only with probability ξ. With the remaining probability of $1-\xi$, agent A has fifty-fifty prior about the state. Whether A knows the state or not is common knowledge. Similarly, the arbitrator knows the state of the world only with some probability. More precisely, the arbitrator knows the state whenever agent A knows it; and in addition, the arbitrator knows the state with probability η contingent on A not knowing it. Whether the arbitrator knows the state is the arbitrator's private information.

One possible interpretation of this information structure, which I will refer to in the paper, is that if the state of the world is b, then agent B learns an argument in support of state b; but he learns no argument, if the state of the world is a.[5] Agent B presents his argument to agent A. It is not a decision problem of agent B whether to present his argument to A. Arguments are presented at an early (unmodelled) stage of face-to-face negotiations, when agents may not yet have been seriously considering the possibility of resolving their conflict by arbitration. At this stage, one may ask why only one party learns about, or obtains a signal about the state of the world. I discuss this issue at the very end of this section.

If B learns an argument, he knows the state is b; if he learns no argument, he infers correctly that the state is a. In state b, agent A finds the argument presented to her by agent B convincing with probability ξ, in which case she learns that the state of the world is b. With the remaining probability of $1-\xi$, agent A does not find B's argument convincing, in which case B's argument does not affect her fifty-fifty prior about the state. Whether A finds B's argument convincing becomes common knowledge.

In state a, agent B makes up an argument that the state is b, which he presents to A. With probability ξ, agent A recognizes that the argument is fake, in which case she concludes that the state of the world is a. With the remaining probability of $1-\xi$, she does not recognize that the argument has been made up. In this case, B's

[5] The assumption that agent B learns an argument in support of state b with probability 1 and 0 in states b and a, respectively, is made for the sake of simplicity; alternatively, it could be assumed that B learns an argument in support of state b with probabilities $p > 1/2$ and $q < 1/2$, respectively.

argument does not affect A's fifty-fifty prior about the state. Whether A recognized that B's argument had been made up becomes common knowledge.[6]

If the parties end up with arbitration, agent B presents his (real or fake) argument to an arbitrator. If the argument was persuasive to agent A in state b, the arbitrator finds it persuasive as well. Additionally, still in state b, the arbitrator finds the argument persuasive with a probability $\eta > 0$ contingent on A finding it unpersuasive.

In state a, the arbitrator recognizes that the argument is fake whenever A recognizes that it has been made up. Additionally, the arbitrator recognizes that the argument is fake with probability η contingent on A not recognizing that it has been made up. Figure 1(a) outlines this interpretation of unpersuasive arguments in a diagram.

The other scenario analyzed is that of possibly *noncommon prior beliefs (NB)*: In each state, there is a probability ξ that both parties know the state; with the complementary probability, the two parties disagree on the state, being convinced that the right state is the one advantageous to them. While their disagreement is common knowledge, they also believe that the arbitrator will share their own belief. As under UA scenario, the arbitrator knows the state whenever the parties know it, and in addition, the arbitrator learns the state with probability η contingent on the disagreement.

In the interpretation, in state b, agent A finds agent B's argument convincing with probability ξ, in which case she learns that the state of the world is b. With the remaining probability of $1 - \xi$, agent A believes that the state of the world is a, and that B is mistaken in believing that the state is b. Agent A also believes that any third party will agree that the state is a. At the same time, agent B believes both that the state of the world is b, and that A is mistaken in believing that the state is a. Agent B also believes that any third party will agree that the state is b.[7]

[6] That is, the probability that in state a agent A fails to recognize that the argument has been made up is equal to the probability that in state b agent A does not find the argument presented by agent B convincing. This assumption is made solely for the sake of simplifying statistical inference; it is not fundamental to the analysis.

[7] Again, the assumption about the beliefs of both agents regarding a third party is made solely for

In state a, agent B cannot make up arguments (or equivalently, A recognizes fake arguments with probability 1). Both agents learn that the state of the world is a with probability ξ. With the remaining probability of $1-\xi$, agent B incorrectly believes that he learned an argument, and that the state of the world is b. Agent A does not find this argument convincing, and believes that the state of the world is a. Agent B believes that A is mistaken in believing that the state is a. Agent B also believes that any third party will find his argument convincing and thus will believe that the state is b. At the same time, agent A believes both that B is mistaken in believing that the state is b, and that any third party will believe that the state is a. Each party knows his or her opponent's beliefs.

If A does not find B's argument convincing, the arbitrator learns the state with probability η. And with the complementary probability, the arbitrator assumes that both states of the world are equally likely. In contrast, when both agents learn which state the world is in, the arbitrator also learns about the state. Figure 1(b) outlines this interpretation of noncommon prior beliefs in a diagram.

Of course, the arbitrator will also be allowed to make statistical inferences from equilibrium strategies. For example, if under UA, agent B rejects the settlement payment only when the state of the world is b, then the arbitrator concludes that the state is b from the fact that the arbitration request has been filed, even if she is unable to recognize the argument directly.

Under both UA and NB, the timing of the game is the same: In period 1, agent A offers a nonnegative settlement payment to agent B who can accept or reject this payment. In the former case, the payment is made and the game ends in period 1. In the latter case, agent B files a request for arbitration. An arbitrator then decides on the payment that A has to make to B, and the arbitrator's decision is enforceable. Figure 2 outlines this timing in a diagram. At this stage, one may ask why the uninformed party, rather than the informed party, makes an offer of a settlement payment. I discuss this issue at the very end of this section.

simplicity; alternatively, it could be assumed that each agent finds it more likely that third parties share her own belief regarding the state of the world than they share the belief of her opponent.

The two information structures: unpersuasive arguments and noncommon prior beliefs are modelled in a very symmetric fashion in order to emphasize that the differences in welfare rankings of different arbitration procedures come from the differences in the kind of information (or beliefs) that negotiating parties may have. However, they will be addressed separately, in different sections of the paper, and can be thought about as independent models.

The arbitration takes place in period 2 and imposes a legal cost of c on each agent. This cost includes the explicit costs of arbitrators' fees, stenographic expenses, and renting a room for a hearing, and also includes the implicit cost of getting involved in an arbitration procedure.[8] It is reasonable to assume that

$$c < \frac{1}{2}.$$

Following Farber (1980b), I assume that the arbitrator is statistically objective but makes systematic mistakes.[9][10] More precisely, the arbitrator would find it right

[8] In practice, the legal cost may not be the same for both parties. Under the labor-union-management arbitration law in most states, the explicit costs are shared equally, but there are exceptions (e.g., Pennsylvania). However, the assumption that the legal costs are the same is made solely for simplicity, and it is not fundamental to the analysis.

One may also consider the scenario in which the arbitrator does not perform any independent investigation, but the parties strategically expend resources on convincing the arbitrator to adopt their views. Lester (1984) argues, however, that arbitrators' duties regarding the method of collecting evidence are often quite precisely specified, so this scenario seems to be rare in practice.

[9] The assumption that the arbitrator's decision is uncertain has firm support from the empirical literature. (See Ashenfelter and Bloom (1984), Farber and Bazerman (1986), or Ashenfelter (1987).)

The assumption that the decision is unbiased is usually motivated by the rules for selecting arbitrators. Typically, an arbitrator is selected at least in part by a mutual agreement of the parties to a dispute; since parties can view the arbitrator's record in related arbitration cases, biased arbitrators are unlikely to be selected. (See Bloom and Cavanagh (1986) for an analysis of selecting arbitrators.)

[10] Some other ingredients of the present model are also known from the existing literature. For example, Gibbons (1988) studied final-offer arbitration as a signaling game, and Dewatripont and Tirole (1999) studied the possibility of forging information (making up arguments) in a different setting.

if A had to make the payment of

$$\pi = p_b(1+c) + (1-p_b)(-c) + \widetilde{\varepsilon}, \tag{1}$$

where p_b denotes the arbitrator's belief that the state of the world is b at the time in which she makes the decision regarding π, and $\widetilde{\varepsilon}$ is a noise term. This formula assumes that A pays a share of the total legal cost $2c$ proportional to p_b - the fraction of \$1 the arbitrator thinks is objectively justified. Since agent B had to pay the legal cost c himself, A must pay him $p_b c + (1-p_b)(-c)$ to make his cost share equal to $(1-p_b)2c$.[11]

I assume that $\widetilde{\varepsilon}$ is distributed symmetrically and unimodally around 0, which implies that $E\widetilde{\varepsilon} = 0$. I also assume that ε has a differentiable density $f(\varepsilon)$ which satisfies the monotone hazard-rate condition

$$\frac{d}{d\varepsilon}\left(\frac{f(\varepsilon)}{1-F(\varepsilon)}\right) > 0, \tag{2}$$

where F stands for the cumulative distribution. Given the assumption that f is symmetric around 0, the monotone hazard-rate condition is equivalent to the log-concavity of F, i.e., to

$$\frac{d}{d\varepsilon}\left(\frac{f(\varepsilon)}{F(\varepsilon)}\right) < 0. \tag{3}$$

The monotone hazard-rate assumption is satisfied for many probability distributions of interest, e.g., uniform or normal. Together, the monotone hazard-rate assumption and log-concavity will guarantee the existence of best responses in the final-offer arbitration game. Moreover, the two assumptions will guarantee that the best responses are characterized by the first-order conditions.

I compare two forms of arbitration: conventional and final-offer arbitration. Under *conventional arbitration*, the arbitrator imposes payment π, given by formula (1). The noise term $\widetilde{\varepsilon}$ will be disregarded in most of the analysis of conventional arbitration, since the agents are assumed to be risk-neutral. However, it will become essential

[11] The assumption is convenient but inessential for the analysis. It may not be satisfied in practice. (See, for example, Lester (1984).)

to the analysis of Section 4. Let us now consider *final-offer arbitration*. In this kind of arbitration, each agent suggests a payment (makes a final offer) of π_A and π_B, respectively, and the arbitrator selects the final offer that is closer to the payment the arbitrator thinks is objectively justified, given by formula (1). However, as I explain later, the value p_b is determined in equilibrium, and so it is typically different for the two arbitration procedures.

Agent A has to pay π_A if

$$\mid \pi_A - \pi \mid < \mid \pi_B - \pi \mid ,$$

and she has to pay π_B if the opposite inequality holds. Assume that the arbitrator tosses a fair coin in the case of equality.

In practice, the arbitrator can elicit final offers under conventional arbitration as well, and this is what usually happens there. The difference between the two schemes is that under conventional arbitration, the arbitrator is not committed to choosing one of the final offers.[12] Indeed, the set of equilibrium outcomes would not alter if we added an extra stage to the model of conventional arbitration in which parties would be sending cheap-talk messages to the arbitrator after filing the request for arbitration.

I will now explain how the value p_b is determined. If the arbitrator finds the argument of agent B convincing, which can happen only in state b, then $p_b = 1$. Similarly, if the arbitrator recognizes that the argument is fake (i.e., learns that the state of the world is a), then $p_b = 0$. In the alternative case in which the arbitrator does not find the argument of B persuasive and does not recognize that the argument is fake, she makes a statistical inference, so that the value of p_b is determined by: equilibrium strategies, the settlement payment offered to agent B, and - under final-offer arbitration - the final offers.

[12]It is interesting to note here that in the first years of arbitration in New Jersey (i.e., in the period between 1977 and 1983), a number of cases were prepared and mediated under final-offer arbitration. However, after the parties had already revealed their positions, the cases were converted into conventional arbitration to facilitate the drafting of the agreement and to avoid having to select a winner (see Comparisons and Conclusions in Lester, 1984).

To conduct a welfare analysis of the two forms of arbitration, I apply two criteria. The first criterion reflects the fact that a request for arbitration is costly as each agent must pay the legal cost c. Perhaps more importantly, negotiated settlements are typically better for the future relations between the disputing parties than settlements imposed by a third party. I will therefore compare the probabilities of filing an arbitration request in the equilibria of the two arbitration games. If this probability of filing an arbitration request is higher under one form of arbitration than another, then I will say that the *deadweight-loss* of that form of arbitration exceeds the deadweight loss of the other form.[13]

Denote by w_A and w_B the total payoff of agents A and B, respectively. Note that $w_B = -w_A$ is equal to the settlement payment made by A in period 1, if the game ends in period 1; and $w_B = \pi - c$ and $w_A = -\pi - c$ if the game ends in period 2 and the arbitrator orders A to make payment π to B. Note also that the comparison of the probabilities of filing an arbitration request is equivalent to the comparison of the expected values of $w_A + w_B$; and since agents are risk-neutral, $w_A + w_B$ represents the sum of the expected utilities of both agents.

The second criterion reflects the fact that one may wish to depart as little as possible from the existing contract, namely, $w_B = 1$ and $w_A = -1$ in state b and $w_A = w_B = 0$ in state a. Agents may want to do this for a number of reasons. Suppose, for example, that agent B is to make an unobservable investment that will benefit agent A in state b; and B is to be compensated by A, contingent on state b. In such a case, one may wish the compensation to be enforceable in order to give agent B appropriate incentives for making this unobservable investment. In the present

[13] I follow here most of the literature on arbitration, which focuses on minimizing the reliance on arbitration. Of course, one may criticize this focus by saying that if that were really the goal, no form of arbitration would do better than threatening both parties with some very bad outcome if they did not agree on a settlement.

However, there are many good arguments which counter this criticism; one of them claims that if the threat of a very bad outcome were the reason for reaching a settlement, then settlements negotiated under that condition would probably no longer be better for the future relations between the disputing parties than settlements imposed by a third party.

paper, I do not model any specific reason for minimizing any departure from the existing contract. Instead, I will simply compare the expected loss caused by such a departure,

$$E[l(|w_A + 1|, |w_B - 1|) \mid b] + E[l(|w_A|, |w_B|) \mid a],$$

where l stands for a loss function which increases in both arguments, and which takes the value 0 at 0. If this expected loss is lower under one form of arbitration than another, then I will say that form is more *outcome-accurate* than the other form.

I have made a number of assumptions only for the sake of simplicity of exposition, including the assumptions that states a and b are equally likely and that parties face identical legal costs c. Other assumptions are necessary to make the analysis tractable, including the assumption that only one party can learn an argument in support of its case, and the assumption that it is the uninformed party (rather than the informed party) who makes an offer of a settlement payment. If an offer of a settlement payment were made by a privately informed party, the model would contain a signalling game in period 1, since the settlement payment offered by the privately informed party could be used as a signal about this party's private information. The analysis would then be rather intractable, as this signalling game from period 1 would be followed, in the final-offer arbitration game, by the signalling game of period 2 in which the final offer of any privately informed party is also used as a signal about this party's private information.

Actually, multiple signalling stages is not the only problem with the analysis of games in which an informed party makes offers of settlement payments. We will return to this issue in Sections 3.3 and 3.4.

3 Unpersuasive Arguments

3.1 Conventional Arbitration

I shall now characterize the equilibrium outcomes of the conventional arbitration game. First, I summarize what can happen at the beginning of period 1. There are the following possibilities: (1) both agents know the state of the world; (2) agent B knows that the state is b, but agent A does not find B's argument persuasive; (3) agent B knows that the state is a, but agent A does not recognize that the argument presented to her has been made up. The analysis of possibility (1) is straightforward. Since both agents have the same expectation regarding the arbitrator's decision, agent A offers the settlement payment that makes agent B exactly indifferent between accepting and rejecting the offer. (If the offer that makes B indifferent is negative, then A offers a settlements payment of 0.) It follows that agent A offers 0 in state a and 1 in state b. Agent B accepts the offer made and the game ends in period 1.

Claim 1 *If* agent B *has an argument and* agent A *finds that argument of* B *persuasive, then* A *offers a settlement payment of* \$1, *which* B accepts. *If* B *makes up his argument and* A *recognizes that the argument is fake, then* A *offers no payment, and which* B *accepts.*

Cases (2) and (3) have to be analyzed simultaneously, since agent A cannot distinguish between the two cases. Recall that in both cases agent B knows the state of the world; on the other hand, agent A does not know the state and believes that both states are equally likely. Agent B who knows that the state of the world is a (call his type a) is willing to accept a strictly lower settlement payment than is agent B who knows that the state of the world is b (call his type b). This is so, because there is a probability $\eta > 0$ that the arbitrator will learn about the state of the world even when agent A does not. In Appendix A, I prove a slightly stronger claim:

Claim 2 *Suppose that agent* B *has an argument and agent* A *does not find the argument of* B *persuasive, or* B *makes up an argument but* A *does not recognize that* B*'s argument is fake. In equilibrium, there are only the following two possibilities: both types of* agent B *accept the settlement payment with probability* 1, *or* type a *accepts the settlement payment with some probability* $q \in [0, 1]$ *and* type b *rejects this payment with probability* 1.

First, I will characterize the equilibria in which type a is indifferent between accepting and rejecting the settlement payment and type b rejects this payment with probability 1. Denote by β the probability assigned by the arbitrator to the event that B's type is b when the arbitrator does not find the argument of agent B convincing but cannot tell whether that the argument has been made up. For every $\beta \in [1/2, 1]$, denote by s_β the expected payoff of type a contingent on filing a request for arbitration, given the value of β. If type a files such a request, then he faces the lottery that yields

$$-2c \text{ with probability } \eta \text{ and } \beta(1+c) - (1-\beta)c - c \text{ with probability } 1-\eta \tag{4}$$

(in this section, we disregard the noise term $\widetilde{\varepsilon}$), i.e.,

$$s_\beta = (1-\eta)\beta - 2c\left[\eta + (1-\eta)(1-\beta)\right]. \tag{5}$$

In equilibrium, type a accepts the settlement payment with probability q. So, by the Bayes rule, the arbitrator must believe (contingent on the rejection of the settlement payment) that B's type is b with probability β, given by

$$\beta = \frac{1}{2-q}. \tag{6}$$

If type a is to be indifferent between accepting and rejecting the settlement payment, this payment must be equal to the s_β given by (5) for β given by (6).

Notice that s_β increases with β. Depending on c and η, it can happen that s_β is negative (even for all $\beta \in [1/2, 1]$). However, the settlement payment must, by assumption, be nonnegative. Define S as the interval (possibly empty or degenerate) consisting of all nonnegative s_β. In Appendix A, I show that:

Claim 3 *In every equilibrium such that type a is indifferent between accepting and rejecting the settlement payment and type b rejects this payment with probability* 1*, the settlement payment offered by* agent A *has to maximize her payoff over the interval S, under the assumption that any s_β from this interval will be accepted by type a with probability q given by (6).*

Roughly speaking, this claim follows from the assumption that agent A makes take-it-or-leave-it offers in period 1; the actual argument is slightly more subtle as an off-equilibrium offer of s_γ need not be accepted by type a with probability q given by (6) for $\beta = \gamma$ contingent on the rejection of s_γ.

In Appendix A, I prove the following characterization of other equilibria in which type a accepts the settlement payment with some probability $q \in [0,1]$ and type b rejects this payment with probability 1:

Claim 4 (a) *Every equilibrium such that both types of* agent B *reject the settlement payment with probability* 1 *is outcome-equivalent to an equilibrium with the settlement payment equal to* $s_{1/2}$, *in which* type a *is indifferent between accepting and rejecting this settlement payment.*

(b) *In every equilibrium such that type* a *accepts the settlement payment with probability* 1 *and* type b *rejects the settlement payment with probability* 1, type a *must be indifferent between accepting and rejecting this payment or the settlement payment must be equal to* 0.

Moreover, if the equilibrium settlement payment is equal to 0 *and type* a *is not indifferent between accepting and rejecting this payment, then* $s_1 < 0$.

I will now turn to the existence of equilibria in which type a accepts the settlement payment with some probability $q \in [0,1]$ and type b rejects this settlement payment with probability 1. Notice that such an equilibrium exists only if agent A's payoff is no lower than the payoff to offering $s = 1$ that is accepted with probability 1 by both types of agent B. This is so because any offer higher than 1 guarantees that both types of agent B accept the settlement payment.

On the other hand, one can easily verify that if the constraint from the previous paragraph is satisfied, then for every $s_\beta \in S$ that maximizes agent A's payoff over this interval (assuming that the rejection of any s_β results in the arbitrator's belief that B's type is b with probability β), there exists an equilibrium in which A offers the settlement payment s_β. In this equilibrium, (i) the arbitrator believes that B's type is b with probability $1/2$ contingent on the rejection of any settlement payment $0 \leq s < s_{1/2}$, and no type accepts any payment $0 \leq s < s_{1/2}$; (ii) the arbitrator believes that B's type is b with probability γ contingent on the rejection of any settlement payment $s_\gamma \in S$; type a accepts such a payment s_γ with the probability q given by (6) in which β is replaced with γ, and type b rejects this payment s_γ with probability 1; (iii) the arbitrator believes that B's type is b with probability 1 contingent on the rejection of the settlement payment $s > s_1$; type a accepts any such offer, and type b accepts only offers $s \geq 1$.

Similarly, it is easy to see that if $s_1 < 0$, then A's payoff to offering no settlement payment (accepted by type a and rejected by type b) is no lower than the payoff to offering $s = 1$ that is accepted with probability 1 by both types of agent B. Therefore, there exists an equilibrium with the settlement payment equal to 0, which is accepted with probability 1 by type a and rejected with probability 1 by type b.

Recall that s_β denotes the expected payoff of type a of agent B contingent on filing a request for arbitration, given that the arbitrator believes that B's type is b with probability β. Let $U^A(s_\beta)$ denote the expected payoff of agent A to making the offer s_β, given that the offer is rejected by type b of agent B and accepted by type a of agent B with the probability such that β is the probability that B's type is b contingent on the offer being rejected. I have shown that:

Claim 5 *There exists an equilibrium in which* type a *accepts the settlement payment with some probability* $q \in [0,1]$ *and* type b *rejects this settlement payment with probability* 1 *if and only if* $s_1 < 0$ or

$$\max_{s_\beta \in S} U^A(s_\beta) \geq -\, 1.$$

To find the settlement payment s_β that maximizes A's payoff over the interval S, one can compare the marginal cost and benefit of an increase in the parameter β. Such an increase implies an increase in the probability with which type a of agent B accepts the settlement payment; and agent A saves then on the legal cost; as well as, she extracts the premium that B is willing to pay in order to avoid the legal cost. This means that the marginal benefit of an increase in β is

$$\frac{1}{2}(2c)\frac{dq}{d\beta} = \frac{c}{\beta^2}. \tag{7}$$

An increase in β also implies that A has to make a higher settlement payment. This higher payment has to be made both when B accepts the settlement payment and also, because of an increase in the arbitrator's belief that B's type is b, when she rejects the settlement payment. Therefore, the marginal cost of such an increase is

$$\frac{ds_\beta}{d\beta} = (1-\eta)(1+2c) \tag{8}$$

(with the derivative having been derived from (5)).

The marginal cost is therefore constant, and the marginal benefit is decreasing in β. The equilibrium s_β can therefore be determined by making the two equal, i.e.,

$$\beta = \sqrt[2]{\frac{c}{(1-\eta)(1+2c)}}, \tag{9}$$

unless we have one of the corner solutions. If the right-hand side exceeds 1, and then $\beta = 1$. If the right-hand side falls below 1/2, then $\beta = 1/2$ (or if it falls below β^* such that $s_{\beta^*} = 0$, then $\beta = \beta^*$).

I will now turn to equilibria in which both types of agent B accept the equilibrium settlement payment with probability 1. In any equilibrium with this property, the settlement payment s must satisfy two constraints. First, it must belong to the interval $[r_0, 1]$ where

$$r_0 := \eta - 2c(1-\eta).$$

Indeed, type b would not accept any lower settlement payment than r_0, because when he rejects the settlement payment, he would at worst face the lottery that yields

$$1 \text{ with probability } \eta \text{ and } -2c \text{ with probability } 1-\eta$$

(in this section, we disregard the noise term $\widetilde{\varepsilon}$) whose expected value is equal to r_0. On the other hand, both types of agent B must accept the settlement payment of 1, as they cannot expect any higher payoff by filing a request for arbitration.

Second, agent A's payoff to making the offer s, assuming that it will be accepted by both types, cannot be lower than the payoff to making any offer $s_\beta \in S$ with $s_\beta < s$, when $S \neq \emptyset$ (i.e., when $s_1 \geq 0$), assuming that the rejection of s_β results in the arbitrator's belief β that agent B's type is b. When $s_1 < 0$ (i.e., $S = \emptyset$), agent A's payoff to making the offer s cannot be lower than the payoff to offering no settlement payment, assuming that it will be accepted by type a and rejected by type b. This follows from an analogous argument to that used in the analysis of equilibria in which type a accepts the settlement payment with some probability $q \in [0, 1]$ and type b rejects this payment with probability 1.

On the other hand, one can easily verify that there exists an equilibrium in which A offers a settlement payment s satisfying the above two constraints, and both types of agent B accept this payment with probability 1. Consider the case of $S \neq \emptyset$ (or $s_1 \geq 0$). The case of $s_1 < 0$ (i.e., $S = \emptyset$) is analogous. In this equilibrium, the arbitrator believes that: (i) B's type is a (i.e., is b with probability 0), contingent on the rejection of s or any higher offer; (ii) she believes that B's type is b with probability 1/2, contingent on the rejection of any offer both lower than s and lower than $s_{1/2}$; (iii) she believes that the probability is β, contingent on the rejection of any offer $s_\beta \in S$ with $s_\beta < s$; and (iv) she believes that the probability is 1, contingent on the rejection of any offer lower than s but higher than s_1. In period 1, agent B responds to A's offer as follows: (i) the settlement payment of s (or any higher one) gets accepted by both types of agent B; (ii) settlement payments both lower than s and lower than $s_{1/2}$ get rejected by both types of agent B; (ii) the payment $s_\beta \in S$ with $s_\beta < s$ gets accepted by type a with probability q given by (6) and rejected by type b with probability 1; and (iv) settlement payments lower than s but higher than s_1 get accepted by type a and get rejected by type b.

I have thus proved that:

Claim 6 *Given an $s \geq 0$, there exists an equilibrium in which both types of* agent *B accept the settlement payment s if and only if*

$$s \in [r_0, 1]$$

and: (i) $s_1 \geq 0$ and

$$\max_{s_\beta \in S, s_\beta \leq s} U^A(s_\beta) \leq -s;$$

or (ii) $s_1 < 0$ and

$$-\frac{1}{2}(1+2c) \leq -s.$$

It follows from Claims 5 and 6 that: the conventional arbitration game has an equilibrium for any set of parameters of the model. Indeed, it follows from the fact that the conditions $\max_{s_\beta \in S} U^A(s_\beta) < -1$ and $\max_{s_\beta \in S, s_\beta \leq s} U^A(s_\beta) > -s$ (for $s = 1$) are mutually exclusive.

Remark 1 (a) Of course, the equilibria (even within each class: that described in Claim 5 or that described in Claim 6) are typically not unique, because multiple systems of beliefs may support any given equilibrium outcome. In addition, in the cases in which the settlement payment is rejected in equilibrium by both types of agent B, there are typically multiple settlement payments which would also be rejected by both types of agent B. Therefore, I will focus in the following material on equilibrium outcomes.

Within each of the two classes of equilibria, if for given parameters of the model this class is nonempty, the equilibrium outcome is unique. It follows from (9) in the case of equilibria described in Claim 5 in which the settlement payment is accepted by type a with some probability and rejected by type b with probability 1; and it follows from the definition of equilibrium outcome in the case of equilibria described in Claim 6 in which the settlement payment is accepted by both types of agent B.

For some sets of parameters (e.g., those studied in Section 3.4), there exist equilibria of only one type (that described in Claim 5 or that described in Claim 6), but for other sets of parameters (e.g., those studied in Section 3.3), the equilibrium outcome is not unique, i.e., there exist equilibria of both types.

(b) In the case of multiplicity, one may wonder whether some equilibria cannot be refined away by some standard arguments, e.g., the Cho-Kreps (1987) criterion or Banks and Sobel's (1987) divinity. These criteria are not applicable directly, because the action of the arbitrator is with probability η determined by the type of agent B.

However, it would be consistent with the spirit of these criteria to require that if a settlement payment is rejected, the belief that B's type is b cannot fall down compared to the prior; intuitively, type b has always a better reason to reject the offer than type a, because there is a chance that the arbitrator will find out the type of agent B. The equilibria supporting some equilibrium outcomes in which the settlement payment is accepted by both types of agent B, including ones described in this section, fail this requirement; whereas the equilibrium outcomes in which the settlement payment is accepted by type a with some probability and rejected by type b with probability 1 are still supported by equilibria satisfying this requirement, e.g., by the equilibria described in this section.

It will turn out that no equilibrium outcome in the limit cases, when $\eta \approx 0$ or $\eta \approx 1$, can be refined away by this sort of arguments.

3.2 Final-Offer Arbitration

As is usually the case in signalling games, the continuation game beginning in period 2 has multiple equilibria, including separating, pooling, and a number of kinds of hybrid equilibria. This multiplicity makes it difficult to obtain any insight into a general value of the parameter η. Thus, I will analyze only the two polar cases - where η close to 0 and where η close to 1 - for which some insight can be obtained. The former case approximates symmetric-information settings which have been studied extensively in the existing literature. In this case, agent B has superior private information about the state of the world, but this information tells him rather little about the expected outcome in arbitration. The latter case approximates a standard game of asymmetric information, similar to ones studied by Farmer and Pecorino (1998) and (2003). However, the present model will provide novel insights into each of the two cases. Additionally, the two polar cases exhibit strategic effects, which are present for any value of η, but which become dominant (and therefore easier to describe) only for the extreme values.

Two initial observations apply to any value of the parameter η. First, equilibria exist (see Appendix B for the proof):

Proposition 1 *Under the additional assumption that final offers have to belong to* *an interval $[-C, 1+C]$, where the value of C can be arbitrarily large but exogenously given, there exists an equilibrium of the two-stage final-offer arbitration game.*

Second, the equilibria of the final-offer arbitration game have the same form as the equilibria of the conventional arbitration game. More specifically, either the settlement payment is accepted by type a with some probability $q \in [0, 1]$ and rejected by type b with probability 1 or both types accept the settlement payment with probability 1. This follows from arguments similar to the ones used in my analysis of the conventional arbitration game.[14]

Remark 2 To prove Proposition 1, I apply a fixed-point theorem, which requires the compactness of the action space. This is the reason for making the additional assumption that the final offers have to belong to a bounded interval. It is an open question if this additional assumption can be disregarded.

The assumptions of the present setting do not guarantee the *single-crossing property* in the final-offer continuation game that begins in period 2. Thus, the equilibria can be of a form rather different from the form of the equilibria of signalling games studied in the existing literature.

A single-crossing property can be obtained under some additional assumption of the cdf of the noise term $\widetilde{\varepsilon}$. One can then show the existence of equilibria without assuming that the final offers have to belong to a bounded interval, and these equilibria can be characterized by methods similar to ones used in the existing literature on signalling games.

3.3 The case of small η

Consider first the conventional arbitration game. It follows from (9) (and the comment following (9)) that for every η close enough to 0, only two kinds of equilibria are possible: one such that the settlement payment is rejected with probability 1 by

[14]The only difference is that under final-offer arbitration there exist equilibria in which type a accepts with probability 1, and type b rejects with probability 1, yet type a is not indifferent between accepting and rejecting.

both types of agent B, and one such that it is accepted with probability 1 by both types of agent B. Furthermore, it follows from Claim 5 that there always exists an equilibrium in which both types of agent B reject the settlement payment; indeed, $s_{1/2} \approx 1/2 - c > 0$, $s_1 \approx 1$ and $\max_{s_\beta \in S} U^A(s_\beta) = -(1+2c)/2 > -1$. And it follows from Claim 6 that for an interval of settlement payments s there exists an equilibrium in which both types of agent B accept the settlement payment of s; indeed, $r_0 \approx -2c$ and $\max_{s_\beta \in S, s_\beta \leq s} U^A(s_\beta) \approx -(1+2c)/2 < -s$ when $s < (1+2c)/2$.[15]

The final-offer arbitration game also has multiple equilibria. First, observe that every equilibrium outcome of the conventional arbitration game can be achieved, in the limit as $\eta = 0$, in an equilibrium of the final-offer arbitration game. More precisely, suppose that in every continuation game beginning in period 2 in which B's type is b with probability β, agents play the pooling equilibrium in which the arbitrator believes that B's type is b with probability β independently of his final offer.[16] In this equilibrium, the arbitrator makes no inference from the final offer of agent B.[17] Agent A therefore chooses π_A to minimize and agent B chooses π_B to maximize

$$F^\beta\left(\frac{\pi_A+\pi_B}{2}\right)\pi_A + \left[1 - F^\beta\left(\frac{\pi_A+\pi_B}{2}\right)\right]\pi_B,$$

where F^β is the *cdf* of $\beta(1+c)+(1-\beta)(-c)+\widetilde{\varepsilon}$. The offers are thus jointly determined

[15] Notice that for any $s > s_{1/2} = 1/2 - c$ the equilibrium outcome in which both types of agent B accept the settlement payment of s satisfies the "divinity" condition discussed in Remark 1.

Indeed, in the construction of equilibria from the paragraph preceding Claim 6, one can replace the belief that B's type is b with probability 0 contingent on the rejection of s (or any higher offer) with the belief that B's type is b with probability 1/2. The condition $s > s_{1/2} = 1/2 - c$ guarantees that both types of agent B still have an incentive to accept this settlement payment.

[16] The result that the equilibrium outcomes of the conventional arbitration game can be achieved (in the limit as $\eta = 0$) in equilibria of the final-offer arbitration game typically does not hold under risk-aversion, because under the two procedures parties are exposed to different types of risk in period 2.

In the early literature on arbitration (summarized in Farber (1980a)), this different type of risk was perceived as the main source of the difference in welfare consequences of the two procedures.

[17] This guarantees that the equilibrium satisfies the "divinity" condition discussed in Remark 1.

by the following two first-order conditions:

$$\frac{\pi_B - \pi_A}{2} = \frac{F^{\beta}\left(\frac{\pi_A + \pi_B}{2}\right)}{f^{\beta}(\frac{\pi_A + \pi_B}{2})},$$

and

$$\frac{\pi_B - \pi_A}{2} = \frac{1 - F^{\beta}\left(\frac{\pi_A + \pi_B}{2}\right)}{f^{\beta}(\frac{\pi_A + \pi_B}{2})}.$$

It follows immediately from these first-order conditions that $\pi_A < \pi_B$ and that $\beta(1+c) + (1-\beta)(-c)$ is the middle of the segment $[\pi_A, \pi_B]$. With probability $(1-\eta)/2$ (which is close to 1/2), each of the two final offers is chosen by the arbitrator; and with probability η the offer of A is chosen in state a and the offer of B is chosen in state b. Thus, the payoff of B tends to $-c + \beta(1+c) + (1-\beta)(-c)$ as η tends to 0 (while the total payoff of A tends to $-c - \beta(1+c) - (1-\beta)(-c)$).

Since these also are the payoffs in the continuation game beginning in period 2 in which B's type is b with probability β under conventional arbitration, every equilibrium outcome of the two-stage conventional arbitration game can be achieved, in the limit as $\eta = 0$, in an equilibrium of the final-offer arbitration game, in which agents anticipate a pooling equilibria in period 2 contingent on the rejection of any settlement payment in period 1.

Now consider another pooling equilibrium of the continuation game beginning in period 2. Suppose that if agent B's type is b with probability β, then B responds optimally to his opponent's offer assuming that F^0 is the *cdf* of the arbitrator's peak points; Agent A best responds to her opponent's offer assuming that F^{β} is the *cdf* of the arbitrator's peak points. Notice that this is indeed an equilibrium, if for any out-of-equilibrium offer of B, the arbitrator believes that his type is b with probability 0.[18]

[18] Notice that the divinity-type argument from Remark 1 does not refine away this equilibrium. Indeed, this would now be type a of agent B who has a better reason of making a lower final offer than type b, because of the chance that the arbitrator will find out the type of agent B.

The offers in this equilibrium are jointly determined by the following two first-order conditions:

$$\frac{\pi_B - \pi_A}{2} = \frac{F^{\beta}\left(\dfrac{\pi_A + \pi_B}{2}\right)}{f^{\beta}(\dfrac{\pi_A + \pi_B}{2})}; \tag{10}$$

$$\frac{\pi_B - \pi_A}{2} = \frac{1 - F^{0}\left(\dfrac{\pi_A + \pi_B}{2}\right)}{f^{0}(\dfrac{\pi_A + \pi_B}{2})}. \tag{11}$$

I shall now describe an equilibrium of the two-stage game in which agents anticipate pooling equilibria described by (10)-(11) in period 2, and the settlement payment is accepted with some probability by type a and rejected with probability 1 by type b. Since this settlement payment must make type a indifferent between accepting and rejecting, it has to be equal to

$$t_{\beta} := -2c\eta + (1-\eta)\left\{\pi_A F^{\beta}\left(\frac{\pi_A + \pi_B}{2}\right) + \pi_B\left[1 - F^{\beta}\left(\frac{\pi_A + \pi_B}{2}\right)\right]\right\}, \tag{12}$$

where π_A and π_B are jointly determined by (10) and (11).

The equilibrium value of β can be derived in a manner similar to (9) (see also the comment following (9)), i.e., by comparing the marginal cost and benefit of an increase in parameter β. The marginal benefit is given by (7), and the marginal cost is given by

$$\frac{dt_{\beta}}{d\beta} = \frac{d\left\{(1-\eta)\pi_A F^{\beta}\left(\dfrac{\pi_A + \pi_B}{2}\right) + (1-\eta)\pi_B\left[1 - F^{\beta}\left(\dfrac{\pi_A + \pi_B}{2}\right)\right]\right\}}{d\beta}.$$

Lemma 1 *For any* $\beta \geq 1/2$,

$$\frac{dt_{\beta}}{d\beta} < \frac{ds_{\beta}}{d\beta}.$$

Recall that s_{β} denotes the settlement payment in the conventional arbitration game which makes type a indifferent between accepting and rejecting, assuming that the rejection of s_{β} results in the arbitrator's belief β that B's type is b.

Proof. See Appendix A ■

The comparison of the marginal costs and benefits of an increase in β implies that the equilibrium β in the final-offer arbitration game can be higher than that in the conventional arbitration game (see Figure 3(a)). In other words, the deadweight-loss in final-offer arbitration can be lower than that in conventional arbitration.

Not much can be said in terms of the outcome accuracy of the two equilibria. Final-offer arbitration can, but need not, benefit agent A at the expense of agent B, but this is independent of the state of the world.

The following proposition summarizes the discussion of this section:

Proposition 2 *For every η close enough to* 0*, both games have exactly two kinds of equilibria.*

(a) In one of them, the settlement payment is rejected with probability 1 *by both types of agent B under conventional arbitration. This is also the outcome of some equilibria of the final-offer arbitration game. However, for some parameters of the model, the final-offer arbitration game has also equilibria in which the settlement payment is accepted with a positive probability by type a (while it is rejected by type b). In every equilibrium in which the settlement payment is accepted with a positive probability, the deadweight-loss of final-offer arbitration falls below that of conventional arbitration.*

(b) In the other kind of equilibria, the settlement payment is accepted with probability 1 *by both types of* agent B*, under both conventional arbitration and final-offer arbitration, which implies that the deadweight-losses of the two procedures are equal.*

Remark 3 Under conventional arbitration, the equilibrium outcomes described in (a) and (b) are by definition unique (see also Remark 1). Similarly, the equilibrium outcome described in (b) is unique under final-offer arbitration. In turn, the equilibrium outcome described in (a) is not unique under final-offer arbitration for the parameters of the model such that the settlement payment is (in some equilibrium) accepted with a positive probability by type a (while it is rejected with probability 1 by type b).

In this latter case, I conjecture that there is a continuum of pooling equilibrium outcomes which yield an interval of probabilities, ranging from 0 to a positive number, such that the equilibrium settlement payment is accepted by type a (while it is rejected with probability 1 by type b). These equilibria can be constructed in a similar manner to the equilibria constructed in this section, because one may prescribe agent B respond optimally to his opponent's offer assuming that F^{γ}, $\gamma \in [0, \beta]$, (instead of F^0) is the *cdf* of the arbitrator's peak point.

I have no characterization of all equilibrium outcomes of the form described in (a) under final-offer arbitration. However, I argue that such a characterization is not essential for the present analysis. Proposition 2 should be interpreted in the following way: Under each arbitration procedure, there exist equilibria in which parties resolve the conflict without filing any arbitration request. Under each procedure, there also exist equilibria of another kind, in which type b of agent B never accepts the settlement payment. In the conventional arbitration game, these equilibria take the worst possible form. Namely, type a of agent B never accepts the settlement payment either, which generates the highest possible deadweight loss. So, the equilibria of this latter kind in the final-offer arbitration game can only generate no higher deadweight loss. And Proposition 2 shows that some of them generate a strictly lower deadweight loss.

The basic and rough intuition behind part (a) can be explained as follows: The final offer of a privately-informed party contains a signal about the party's type. The signalling may be costly. This cost makes lower settlement payments (compared to ones under conventional arbitration) acceptable for some privately-informed parties. The signalling cost may be positive as long as there is an uncertainty regarding the privately-informed party's type. Therefore, the deadweight-loss of final-offer arbitration may fall below that of conventional arbitration even when the private information is virtually negligible.

One may wonder the assumption that only the uninformed party makes offers of settlement payments is essential for the conclusions of this section. If the informed party was supposed to make offers, and these offers were later observed by the arbitrator, it could face a signalling cost also under conventional arbitration. This is true, but the signalling cost would have an ambiguous impact on the deadweight loss. Indeed, consider the conventional arbitration game for $\eta = 0$ in which this is the informed party who makes offers of settlement payment. This game has a continuum of equilibria in which both types of agent B make the offer of settlement payment that makes agent A indifferent between accepting and rejecting it, and agent A accepts it with an arbitrary probability. Both types of agent B have an incentive to make this offer provided that any lower offer results in the arbitrator believing that B's type is a.

3.4 The case of large η

Consider first the conventional arbitration game. It follows from (5) that $s_1 \approx -2c < 0$ (that is, $S = \emptyset$) for every η close enough to 1. Therefore, by Claims 4 and 5, the conventional arbitration game has an equilibrium in which the settlement payment is equal to 0, type a accepts this settlement payment with probability 1, and type b rejects it with probability 1; and this is the only equilibrium in which type b rejects the settlement payment. Furthermore, by Claim 6, there is no equilibrium in which the settlement payment is accepted by both types of agent B; indeed, $r_0 \approx 1$ and $-\frac{1}{2}(1 + 2c) > -s$ for all $s \in [r_0, 1]$.

Consider now the final-offer arbitration game in the limit case when $\eta = 1$. Suppose that B's type is b with probability β in the continuation game beginning in period 2. In equilibrium, type a best responds to A's final offer, knowing that F^0 will be the cdf of the arbitrator's peak point; and type b best responds to A's final offer, knowing that F^1 will be the cdf of the arbitrator's peak point. On the other hand, agent A best responds to B's final offers, assuming that his type (and the state of the world) is b with probability β and that the arbitrator will know the state of the world when making the decision. That is, π_A, $\pi_{B,a}$, and $\pi_{B,b}$ in the continuation game beginning in period 2 where B's type is b with probability β can be uniquely determined by the following set of first-order conditions:

$$\frac{\pi_{B,b}-\pi_A}{2} = \frac{1-F^1\left(\dfrac{\pi_A+\pi_{B,b}}{2}\right)}{f^1(\dfrac{\pi_A+\pi_{B,b}}{2})}, \tag{13}$$

$$\frac{\pi_{B,a}-\pi_A}{2} = \frac{1-F^0\left(\dfrac{\pi_A+\pi_{B,a}}{2}\right)}{f^0(\dfrac{\pi_A+\pi_{B,a}}{2})}, \tag{14}$$

$$\begin{aligned} 0 \;=\; & (1-\beta)\left\{F^0\left(\frac{\pi_A+\pi_{B,a}}{2}\right)-\frac{\pi_{B,a}-\pi_A}{2}f^0(\frac{\pi_A+\pi_{B,a}}{2})\right\} \\ & +\beta\left\{F^1\left(\frac{\pi_A+\pi_{B,b}}{2}\right)-\frac{\pi_{B,b}-\pi_A}{2}f^1(\frac{\pi_A+\pi_{B,b}}{2})\right\}. \end{aligned} \tag{15}$$

The two-stage game has an equilibrium in which the settlement payment is accepted with some probability by type a and rejected with probability 1 by type b; consequently, the probability that agent B is of type b contingent on the rejection of this payment is equal to some $\beta \geq 1/2$. This equilibrium β can be determined by comparing the marginal benefit and marginal cost of an increase in the parameter β. The marginal benefit is given by (7), and the marginal cost is given by:

$$\begin{aligned} & \frac{1}{2}\frac{d\left\{\pi_A F^0\left(\dfrac{\pi_A+\pi_{B,a}}{2}\right)+\pi_{B,a}\left[1-F^0\left(\dfrac{\pi_A+\pi_{B,a}}{2}\right)\right]\right\}}{d\beta} \\ & +\frac{1}{2}\frac{d\left\{\pi_A F^1\left(\dfrac{\pi_A+\pi_{B,b}}{2}\right)+\pi_{B,b}\left[1-F^1\left(\dfrac{\pi_A+\pi_{B,b}}{2}\right)\right]\right\}}{d\beta}. \end{aligned} \tag{16}$$

Lemma 2 The marginal cost given by *(16) is strictly greater than* 0 *for* $\beta > 1/2$, *and is equal to* 0 *for* $\beta = 1/2$.

Proof. See Appendix A. ■

By Lemma 2 taken together with (8), the marginal cost of an increase in β under final-offer arbitration is greater than that under conventional arbitration. Thus, the equilibrium β is no higher, and can be lower, under final-offer arbitration compared to conventional arbitration (see Figure 3(b)).

There is no equilibrium in the two-stage game, in which the settlement payment is accepted with probability 1 by both types of agent B. Indeed, the settlement payment that could be accepted by both types of agent B has to be at least \$1, which is the payoff of type b in the continuation game which begins in period 2 and in which B is of type b with probability 1. It is easy to see that the payoff of type b is even higher in the continuation game which begins in period 2 and in which B is of type b with a probability $\beta < 1$. However, a settlement payment of \$1, even if it is accepted by both types of agent B, makes the payoff of agent A lower than that to offering no settlement payment, which must be accepted by type a.

Summarizing, and applying the upper hemi-continuity of the set equilibrium outcomes of the final-offer arbitration game, I obtain the following proposition:

Proposition 3 *For every η close enough to 1, the conventional arbitration game has a unique equilibrium outcome. The final-offer arbitration game has only one sort of equilibrium outcomes.*

(a) Under conventional arbitration, the payment of \$0 is offered in period 1. This payment is accepted with probability 1 by type a and is rejected with probability 1 by type b.

(b) *Under final-offer arbitration the settlement payment can be equal to \$0 or it can be positive, depending on the equilibrium outcome. When the settlement payment is equal to \$0, it is accepted with probability 1 by type a and rejected with probability 1 by type b. When the settlement payment is positive, it is accepted by type a only with a probability $q < 1$ (while it is rejected with probability 1 by type b). That is, the deadweight-loss of final-offer arbitration never falls below, and may exceed, that of conventional arbitration; in addition, final-offer arbitration is never more, and may be less, outcome-accurate than conventional arbitration.*

Remark 4 I have no characterization of all equilibrium outcomes of the final-offer arbitration game. I conjecture that for any given set of the parameters of the model, the equilibrium outcome is unique. To prove this conjecture, it would suffice to show that the marginal cost curve and the marginal benefit curve from Figure 3(b) intersect at most once. This seems quite plausible, but the calculations are slightly involved.

However, such a characterization does not seem to be essential for the insight regarding the welfare properties of the two arbitration procedures. Proposition 3 should be interpreted in the following way: Under conventional arbitration, the equilibrium outcome is unique, that is, only type b files an arbitration request. Under final-offer arbitration type b also files an arbitration request in every equilibrium. So, the equilibria of the final-offer arbitration game cannot generate a lower deadweight loss. And Proposition 3 shows that these equilibria generate a strictly higher deadweight loss for some parameters of the model.

The intuition can be explained as follows: Final-offer arbitration allows the privately-informed party to extract some informational rent at the expense of the uninformed party, because the uninformed party makes its final offer with inferior information about the arbitrator. The rent of the privately-informed party of type a increases with the uninformed party's belief that it is type b. This reduces the uninformed party's willingness to offering higher settlement payments, since settlement payments are accepted more willingly by type a; and this in turn raises the belief that the privately-informed party's type is b, contingent on rejection. Consequently, higher settlement payments raise the rent of type a of the privately-informed party.

On the other hand, the rent of type b of the privately-informed party decreases with the uninformed party's belief that its type is b. Thus, higher settlement payments reduce the rent of type b of the privately-informed party. However, this is the rent of type a, not the rent of type b, that matters, when only type a can accept the settlement payment.

Again, one may wonder whether the assumption that only the uninformed party makes offers of settlement payments is essential for the conclusions of this section. For instance, the informational rents that the informed party derives from final-offer arbitration may exists only because the informed party gets no opportunity to reveal this information before the arbitration stage. Unfortunately, the model with the informed party making the offers has multiple equilibria, even under conventional arbitration. Notice, however, that for $\eta = 1$ there is an important difference between transmitting information through a final offer, and through a settlement payment. In the former case, the message will be verified by the arbitrator, whereas in the latter it may not be verified if the other party accepts the offer.

4 NonCommon Prior Beliefs

The analysis of this scenario is relatively simple. If agent B has no argument and he does not (incorrectly) believe that he has one, or if he has an argument and agent A finds her argument convincing, then the arbitrator also learns that the state of the world is a or b, respectively. In both cases, agents have identical, correct beliefs about the arbitrator's decision. If B has an argument and A does not find it convincing, then the agents differ in their beliefs about the arbitrator's decision; agent A believes that the arbitrator will believe that the state is a, and agent B believes that the arbitrator will believe that the state is b, while the arbitrator either learns the state the world is in (with probability η), or believes that both states are equally likely. Similarly, if B has no argument but he incorrectly believes that he has one, the agents differ in their beliefs about the arbitrator's decision.

If both agents have the same belief about the arbitrator's decision, then A offers the settlement payment that makes B indifferent between accepting and rejecting the offer. (If the offer that makes B indifferent is negative, then A offers the settlement payment of 0.) By rejecting A's offer, B ends up with the payoff of $-2c+\widetilde{\varepsilon}$ when all agents know that the state is a, and the payoff of $1+\widetilde{\varepsilon}$ when all agents know that the state is b. Thus, the settlement payments in the two cases are 0 and 1, respectively. These settlement payments get accepted by B.

If the agents have different beliefs about the arbitrator's decision, agent A is not willing to offer more than 0, and agent B is not willing to accept less than 1. Thus, B rejects the settlement payment offered by A. In the conventional arbitration game, the arbitrator imposes on agent A the payment $-c+\widetilde{\varepsilon}$ or $1+c+\widetilde{\varepsilon}$ (in state a and b, respectively) with probability η, and a payment of $1/2+\widetilde{\varepsilon}$ with the complementary probability.

The following proposition summarizes the discussion on conventional arbitration:

Proposition 4 *The conventional arbitration game has a unique equilibrium outcome.*

(a) In this equilibrium outcome, if agent B has no argument and he does not incorrectly believe that he has one, the game ends in period 1*, and no payment is made.*

(b) If B has an argument and agent A finds the argument of B convincing, the game also ends in period 1*. Then, the payment made by A is equal to* 1*.*

(c) If each agent believes that his or her opponent is mistaken, the game ends in period 2*. Then, with probability η, the random payment made by A is equal to $-c+\widetilde{\varepsilon}$ in state a and $1+c+\widetilde{\varepsilon}$ in state b, and with probability $1-\eta$, is equal to $1/2.\ +\widetilde{\varepsilon}$.*

In the final-offer arbitration game, if the play reaches period 2, agents A and B choose their final offers π_A and π_B, respectively, to maximize

$$F^A\left(\frac{\pi_A+\pi_B}{2}\right)(-\pi_A)+\left[1-F^A\left(\frac{\pi_A+\pi_B}{2}\right)\right](-\pi_B), \tag{17}$$

and

$$F^B\left(\frac{\pi_A+\pi_B}{2}\right)\pi_A+\left[1-F^B\left(\frac{\pi_A+\pi_B}{2}\right)\right]\pi_B, \tag{18}$$

respectively, given the final offer of his or her opponent. If the agents have the same belief about the decision, $F^A=F^B$ is the cumulative distribution of the random variable $p_b(1+c)+(1-p_b)(-c)+\widetilde{\varepsilon}$, where p_b denotes the arbitrator's belief that the state of the world is b at the time in which the arbitrator makes her decision. Notice that $p_b=0$ if B has no argument, and $p_b=1$ if B has an argument which is recognized by A. If the agents have different beliefs about the arbitrator's decision, $F^A=F^0$ is the cumulative distribution of $-c+\widetilde{\varepsilon}$, and $F^B=F^1$ is the cumulative distribution of $(1+c)+\widetilde{\varepsilon}$; this happens when each agent believes that his or her opponent is mistaken.

It is easy to see that in any equilibrium $\pi_A\leq\pi_B$. Moreover, by (2) and (3) from Section 2, π_A and π_B are determined by the first-order conditions:

$$\frac{\pi_B - \pi_A}{2} = \frac{F^A\left(\dfrac{\pi_A + \pi_B}{2}\right)}{f^A(\dfrac{\pi_A + \pi_B}{2})}, \tag{19}$$

$$\frac{\pi_B - \pi_A}{2} = \frac{1 - F^B\left(\dfrac{\pi_A + \pi_B}{2}\right)}{f^B(\dfrac{\pi_A + \pi_B}{2})}. \tag{20}$$

Note that by (3), the right-hand side of (19) is nondecreasing in π_A, and the left-hand side of (19) is decreasing in π_A. Since the left-hand side is equal to 0 for $\pi_A = \pi_B$, and the right-hand side is positive, the equilibrium π_A satisfies (19). Similarly, by (2), the right-hand side of (20) is nonincreasing in π_B, and the left-hand side of (20) is increasing in π_B. Since the left-hand side is equal to 0 for $\pi_B = \pi_A$, and the right-hand side is positive, the equilibrium π_B satisfies (20).

The following proposition summarizes the discussion on final-offer arbitration:

Proposition 5 *The final-offer arbitration game has a unique equilibrium outcome.*

(a) In this equilibrium outcome, if agent B has no argument and he does not incorrectly believe that he has one, the game ends in period 1*, and no payment is made.*

(b) If B has an argument and agent A finds the argument of B convincing, the game also ends in period 1*. Then, the payment made by A is equal to* 1*.*

(c) If each agent believes that his or her opponent is mistaken, the game ends in period 2*. The equilibrium final offers are jointly determined by (19) and (20).*

Having characterized the equilibria of the two arbitration games, I can now compare their welfare properties. The two equilibrium outcomes differ only in the case in which each agent believes that his or her opponent is mistaken. I claim that, under some mild assumptions about the density function f and the legal costs c, conventional arbitration is more outcome-accurate than final-offer arbitration. To state my result I need the following definition. Recall that if f is the density of $\widetilde{\varepsilon}$, then $f^{1/2}$ stands for the density of $1/2 + \widetilde{\varepsilon}$. The density function f is said to be up to δ in an interval $[-z, z]$ when

$$\left| \int_{-\infty}^{\infty} l(|-x-c+1|, |x-c-1|) f^{1/2}(x) dx - \int_{1/2-z}^{1/2+z} l(|-x-c+1|, |x-c-1|) f^{1/2}(x) dx \right| \leq \delta$$

and

$$\left| \int_{-\infty}^{\infty} l(|-x-c|, |x-c|) f^{1/2}(x) dx - \int_{1/2-z}^{1/2+z} l(|-x-c|, |x-c|) f^{1/2}(x) dx \right| \leq \delta.$$

Note that, under conventional arbitration,

$$E[l(|w_A+1|, |w_B-1|) \mid b] = \int_{-\infty}^{\infty} l(|-x-c+1|, |x-c-1|) f^{1/2}(x) dx$$

and

$$E[l(|w_A|, |w_B|) \mid a] = \int_{-\infty}^{\infty} l(|-x-c|, |x-c|) f^{1/2}(x) dx.$$

In other words, a density function is up to δ in an interval $[-z, z]$ if its tails (i.e., realizations to the left of $-z$ and to the right of z) can be disregarded for the sake of reaching an approximate solution. It seems reasonable to assume that the tails will not matter much for $z = 1/2$.

Proposition 6 *Suppose that $c \leq 1/2$. There exists $\delta > 0$ such that if the density function f of the noise term $\widetilde{\varepsilon}$ is up to δ in the interval $[-1/2, 1/2]$, then conventional arbitration is more outcome-accurate than final-offer arbitration.*

Proof. See Appendix A. ■

Here is the intuition behind this result: When agents believe that their opponent is wrong, they are unlikely to reach any agreement on a settlement payment. Each of them simply believes that the arbitrator will support his or her claim. In such a case, the form of arbitration does not much affect their willingness to negotiate an agreement.

However, final-offer arbitration allows agents to make "strong" claims against their opponent and (as they believe) thereby take advantage of their opponent's being wrong. This in turn makes the arbitrator take more extreme decisions, and thereby reduces the accuracy of arbitration.

The logic behind Proposition 6 is particularly easy to see in the limit case when there is no noise ($\widetilde{\varepsilon} \equiv 0$) in the arbitrator's decision. If agents have different beliefs about the arbitrator's decision, agent A expects $-c$ to be the peak of the arbitrator's preferences, and agent B expects the peak to be at $1 + c$. Thus, if the play reaches period 2, agent A's final offer never exceeds $-c$, while agent B's final offer never falls below $1 + c$. This in turn implies that A's final offer that best responds to B's final offer never exceeds

$$-c - [(1 + c) - (-c)] = -1 - 3c,$$

while B's final offer that best responds to A's final offer never falls below

$$1 + c + [(1 + c) - (-c)] = 2 + 3c.$$

Continuing this reasoning, one concludes that the final offers will tend to $-\infty$ and $+\infty$, respectively.

The difference in outcome accuracy is particularly large for large values of η and small noise $\widetilde{\varepsilon}$. In this case, the arbitrator makes with high probability almost correct decisions under conventional arbitration, but is committed to choosing one of the two very extreme final offers under final-offer arbitration.

The deadweight-loss is identical under both forms of arbitration since under both scenarios the parties end up with arbitration if and only if each agent believes that his or her opponent is mistaken. Note, however, that this result relies on risk neutrality. If parties were risk-averse, one would like not only to minimize the probability of filing an arbitration request, but also to minimize the risk included in the arbitrator's decision. Since final-offer arbitration typically exposes the parties to more risk, its deadweight loss would typically exceed that of conventional arbitration. Again, it is particularly easy to see in the limit case, in which there is no noise in the arbitrator's decision. Then the arbitrator imposes one of the payments: $-c$, $1 + c$, or $1/2$ in the conventional arbitration game, and the arbitrator randomizes between the final offers that tend to $-\infty$ and $+\infty$ in the final-offer arbitration game.

5 Comparative Statics, Extensions

5.1 Quality of Arbitration

In discussing the role of the quality of arbitration, a number of earlier papers have modelled a lower quality as a mean-preserving spread of the noise term ε. The typical result is that arbitration of lower quality can (paradoxically) result in more efficient outcomes by encouraging more collective bargaining by risk-averse parties.

The present setting offers two insights. Under noncommon prior beliefs, a lower quality (modelled as a mean-preserving spread) leads to more accurate outcomes even for risk-neutral agents in the final-offer arbitration game. The result is particularly clear in the limit case, in which there is no uncertainty about the arbitrator's decision. Then, as the quality of arbitration becomes almost perfect, the final offers tend to $-\infty$ and $+\infty$, which moves the outcome away from the most accurate one. This result turns out to be quite general, as it requires only some mild assumptions about the density function f.

The parameter η can be used as another measure of the quality of arbitration, under which, it may well happen that arbitration of lower quality results in less efficient outcomes. Consider the conventional arbitration game with attention restricted to the equilibrium in which type a randomizes between accepting and rejecting and type b rejects the settlement payment. Then a higher η implies a higher β (see (9)), i.e., a higher quality of arbitration reduces its deadweight-loss. By construction, a higher quality of arbitration will typically lead to a higher outcome-accuracy as well.

5.2 Future Research

This paper compares only two forms of arbitration, although they are the two most common forms. In practice, other forms of arbitration have also been (or are) in use. In public-service disputes, some states allow parties to choose the form of arbitration: final-offer arbitration applies if (and only if) one of the parties refuses conventional arbitration. In Iowa, an arbitration procedure involves three tiers: mediation, fact-finding with recommendations, and final-offer arbitration, during which the arbitrator must choose one of three offers, since the fact finder's recommendation is included as a separate offer. In some proposed schemes (see Crawford (1981)), each party makes two offers, the arbitrator selects the party who made the better offers (the winner), and then the other party chooses between the offers made by the winner. More generally, the question of optimal mechanism design remains to be investigated.[19]

One may also wish to explore the potential role of other signalling or screening procedures, combined with or as an alternative to final offers, e.g., "burning money" or the possibility of making another settlement payment offer after filing an arbitration request but before the arbitrator's decision. These modifications lead to a number of interesting questions, but unfortunately, the models of final-offer arbitration become intractable after introducing these modifications.

Finally, it should be emphasized that the present model is "one-dimensional" in that the conflict involves just one issue. In practice, collective bargaining is often "multi-dimensional" (in that it involves several issues). Two forms of final-offer arbitration have been developed for dealing with such multi-dimensional situations. In *package arbitration*, parties make a final offer, specifying their position on each issue, and the arbitrator selects one of the offers. In *issue-by-issue arbitration*, the arbitrator makes a separate decision on each issue. See Crawford (1981) and Lester (1984) for a survey discussion, and Çelen (2003) for an attempt at formal analysis of the two forms of final-offer arbitration.

6 Conclusions

The paper offers a comparison between conventional and final-offer arbitration from a welfare perspective, under the assumption that the differences in parties' information (or beliefs) about the outcome of arbitration are the major cause of the failure of collective bargaining. It is demonstrated that the outcome of this comparison depends on the nature of these differences. If some arguments are convincing for one party,

[19] Brams and Merrill III (1986) can be viewed as an attempt at addressing this question.

but may not be convincing to the other party, then the ranking of the two arbitration procedures depends on the assumptions regarding the arbitrator. If the arbitrator's ability to recognize the validity of arguments is high, conventional arbitration dominates final-offer arbitration: The probability of filing a request for arbitration is (weakly) lower under conventional arbitration. If the arbitrator's ability to recognize the validity of arguments is low, final-offer arbitration dominates conventional arbitration. If parties may believe that their opponents are simply wrong, conventional arbitration approximates the pre-existing contractual arrangements better than final-offer arbitration.

7 Appendix A

Proof of Claim 2: Recall that type a is willing to accept a strictly lower settlement payment than type b. Therefore, the following three responses to the equilibrium offer of the settlement payment are possible: (a) both types accept with probability 1; (b) type b accepts with a probability $q \in (0, 1)$ and type a accepts with probability 1; (c) type b rejects with probability 1 and type a accepts with a probability $q \in [0, 1]$.

To prove the claim, I have to eliminate possibility (b). Indeed, if (b) happened in equilibrium, then the arbitrator would know in period 2 that agent B's type is b. The settlement payment in period 1 would, therefore, have to be equal to 1 to make B indifferent between accepting and rejecting it. Suppose now that A offers (in period 1) a settlement payment higher by a little ν than the equilibrium offer. Both types of agent B accept this higher payment. They must do so because otherwise they would face the arbitrator's decision in period 2 and the arbitrator at best would assume that B's type is b; however, the payoff of this decision is equal to 1 for type b, and is even lower for type a, whose made-up argument can be detected by the arbitrator. Since offering the slightly higher settlement payment means that A does not have to pay the legal cost and pays even less in expectation (since she does not have to pay B's legal cost), she strictly prefers offering the higher settlement payment provided that ν is sufficiently small.

Proof of Claim 3: Suppose first that some s_γ yields a higher payoff to agent A than the settlement payment s actually offered by A in an equilibrium (assuming that the rejection of any s_γ results in the arbitrator's belief that agent B's type is b with probability γ). Then the only case in which A has an incentive to offer s will be when the off-equilibrium offer of s_γ does not result in the arbitrator believing that B's type is b with probability γ, contingent on the rejection of s_γ. If the rejection of any s_γ resulted in a higher belief (than γ), then type a would have to reject the settlement payment of s_γ, which in turn implies that the rejection of s_γ could not result in the higher belief. (It would instead result in the arbitrator's belief that B's type is b with probability $1/2$.) If it resulted in a lower belief, then type a would have to accept the settlement payment of s_γ, and the lower belief would be possible only if type b accepted it as well. This, however, would make the payoff of A even higher compared to the case in which the rejection of any s_γ results in the arbitrator's belief that B's type is b with probability γ.

Proof of Claim 4: (a) Observe that the settlement payment offered by agent A cannot be higher than $s_{1/2}$. As a result, it is easy to see that an identical equilibrium outcome obtains when the settlement payment is equal to $s_{1/2}$. In this case, both types of agent B reject all offers of a settlement payment no higher than $s_{1/2}$, but respond to the offers higher than $s_{1/2}$ as in the original equilibrium.

(b) If type a strictly preferred to accept a positive settlement payment, then A would strictly prefer to offer a settlement payment lower by a little ν than the equilibrium offer. By accepting such a payment, type a would obtain a higher payoff than she would if she had to face the arbitrator's decision in period 2 contingent on the event that B's type is b, because then she would have to pay the legal cost. It would therefore be accepted by type a, which would make the payoff of agent A strictly higher, contingent on type b rejecting this offer. It could happen that the alternative settlement payment gets accepted by type b, but this would make the payoff of agent A even higher.

To demonstrate the very last assertion, notice that if $0 < s_{1/2}$ and type b rejects the settlement payment of 0, then type a strictly prefers to reject this payment. If

$s_{1/2} \leq 0 \leq s_1$, then the settlement payment of 0 can be offered in equilibrium only when type a is indifferent between accepting and rejecting this payment.

Proof of Lemma 1: By definition,

$$
\begin{aligned}
\frac{dt_\beta}{d\beta} &= \frac{d\left\{(1-\eta)\pi_A F^\beta\left(\frac{\pi_A+\pi_B}{2}\right)+(1-\eta)\pi_B\left[1-F^\beta\left(\frac{\pi_A+\pi_B}{2}\right)\right]\right\}}{d\beta} \\
= &\ \frac{\partial\left\{(1-\eta)\pi_A F^\beta\left(\frac{\pi_A+\pi_B}{2}\right)+(1-\eta)\pi_B\left[1-F^\beta\left(\frac{\pi_A+\pi_B}{2}\right)\right]\right\}}{\partial\pi_A}\cdot\frac{d\pi_A}{d\beta}+ \\
&\ \frac{\partial\left\{(1-\eta)\pi_A F^\beta\left(\frac{\pi_A+\pi_B}{2}\right)+(1-\eta)\pi_B\left[1-F^\beta\left(\frac{\pi_A+\pi_B}{2}\right)\right]\right\}}{\partial\pi_B}\cdot\frac{d\pi_B}{d\beta}+ \\
&\ \frac{\partial\left\{(1-\eta)\pi_A F^\beta\left(\frac{\pi_A+\pi_B}{2}\right)+(1-\eta)\pi_B\left[1-F^\beta\left(\frac{\pi_A+\pi_B}{2}\right)\right]\right\}}{\partial\beta}
\end{aligned}
$$

The first term on the right-hand side equals 0 due to the first-order condition (10). The second term equals

$$
\begin{aligned}
&(1-\eta)\left\{\frac{\pi_A-\pi_B}{2}f^\beta(\frac{\pi_A+\pi_B}{2})+\left[1-F^\beta\left(\frac{\pi_A+\pi_B}{2}\right)\right]\right\}\frac{d\pi_B}{d\beta} \\
= &(1-\eta)\left[1-2F^\beta\left(\frac{\pi_A+\pi_B}{2}\right)\right]\frac{d\pi_B}{d\beta},
\end{aligned}
$$

again due to the first-order condition (10). Finally, one can be compute directly that the third term equals

$$
(1-\eta)(1+2c)\left(\pi_B-\pi_A\right)f^\beta(\frac{\pi_A+\pi_B}{2}) = (1-\eta)(1+2c)2F^\beta\left(\frac{\pi_A+\pi_B}{2}\right),
$$

again due to the first-order condition (10). It follows from the first-order conditions (10) and (11) that

$$
\frac{d\pi_B}{d\beta} < (1+2c)
$$

and

$$
F^\beta\left(\frac{\pi_A+\pi_B}{2}\right) < \frac{1}{2},
$$

and the latter inequality is strict unless $\beta = 1/2$. This can be easily seen in Figure 4: For any $\beta \geq 1/2$, the curve $\frac{\pi_B - \pi_A}{2} = \left[1 - F^0\left(\frac{\pi_A + \pi_B}{2}\right)\right] / f^0(\frac{\pi_A + \pi_B}{2})$ is obtained by moving the curve $\frac{\pi_B - \pi_A}{2} = \left[1 - F^\beta\left(\frac{\pi_A + \pi_B}{2}\right)\right] / f^\beta(\frac{\pi_A + \pi_B}{2})$ to the left. So the intersection of $\frac{\pi_B - \pi_A}{2} = \left[1 - F^0\left(\frac{\pi_A + \pi_B}{2}\right)\right] / f^0(\frac{\pi_A + \pi_B}{2})$ and $\frac{\pi_B - \pi_A}{2} = F^\beta\left(\frac{\pi_A + \pi_B}{2}\right) / f^\beta(\frac{\pi_A + \pi_B}{2})$ is at $\frac{\pi_A + \pi_B}{2} < \frac{1}{2}$, as the intersection of the latter curve with the curve $\frac{\pi_B - \pi_A}{2} = \left[1 - F^\beta\left(\frac{\pi_A + \pi_B}{2}\right)\right] / f^\beta(\frac{\pi_A + \pi_B}{2})$ is at $\frac{\pi_A + \pi_B}{2} = \frac{1}{2}$.

Thus, for any $\beta \geq 1/2$,

$$\begin{aligned} \frac{dt_\beta}{d\beta} &< (1-\eta)(1+2c)\left[1 - 2F^\beta\left(\frac{\pi_A + \pi_B}{2}\right)\right] + (1-\eta)(1+2c)2F^\beta\left(\frac{\pi_A + \pi_B}{2}\right) \\ &= (1-\eta)(1+2c) = \frac{ds_\beta}{d\beta} \end{aligned}$$

by (8).

Proof of Lemma 2: By calculations, the marginal cost given by (16) can be expressed as

$$\begin{aligned} &\frac{1}{2}\left\{f^0\left(\frac{\pi_A + \pi_{B,a}}{2}\right)\left(\frac{\pi_A - \pi_{B,a}}{2}\right) + F^0\left(\frac{\pi_A + \pi_{B,a}}{2}\right)\right\} \\ &+\frac{1}{2}\left\{f^1\left(\frac{\pi_A + \pi_{B,b}}{2}\right)\left(\frac{\pi_A - \pi_{B,b}}{2}\right) + F^1\left(\frac{\pi_A + \pi_{B,b}}{2}\right)\right\}, \end{aligned}$$

and further, by the first-order conditions (13) and (14), as

$$F^0\left(\frac{\pi_A + \pi_{B,a}}{2}\right) + F^1\left(\frac{\pi_A + \pi_{B,b}}{2}\right) - 1. \tag{21}$$

Combining (13)-(15), I obtain that:

$$2(1-\beta)F^0\left(\frac{\pi_A + \pi_{B,a}}{2}\right) + 2\beta F^1\left(\frac{\pi_A + \pi_{B,b}}{2}\right) - 1 = 0, \tag{22}$$

and so (21) equals 0 for $\beta = 1/2$. It follows from (13) and (14) that

$$F^0\left(\frac{\pi_A + \pi_{B,a}}{2}\right) > F^1\left(\frac{\pi_A + \pi_{B,b}}{2}\right),$$

and so it follows from (22) that for $\beta > 1/2$, (21) must be greater than 0.

Proof of Proposition 6: It follows from the definitions of f^B, F^B, f^A, F^A, and (19) and (20) that

$$\frac{\pi_A + \pi_B}{2} = \frac{1}{2}.$$

Indeed, since f is symmetric and unimodal around 0, f^A and f^B are symmetric and unimodal around $-c$ and $1+c$, respectively. As a result, I obtain that:

$$f^A(\frac{\pi_A + \pi_B}{2}) < f^B(\frac{\pi_A + \pi_B}{2}) \text{ and } 1 - F^B\left(\frac{\pi_A + \pi_B}{2}\right) < F^A\left(\frac{\pi_A + \pi_B}{2}\right)$$

if $\dfrac{\pi_A + \pi_B}{2} > \frac{1}{2}$, and so the right-hand side of (19) exceeds the right-hand side of (20). Conversely, the right-hand side of (20) exceeds the right-hand side of (19) if $\dfrac{\pi_A + \pi_B}{2} < \frac{1}{2}$.

The assumption that f^B is symmetric and unimodal around $1+c$ also implies that

$$f^B(\frac{1}{2}) \leq \frac{1}{1+2c}.$$

Otherwise, f^B would exceed $1/(1+2c)$ on the interval $\left[\frac{1}{2}, \frac{3}{2} + 2c\right]$ whose length is equal to $1+2c$.

Since

$$\int_{1/2}^{1+c} f^B > \left(\frac{1}{2} + c\right) \cdot f^B(\frac{1}{2}),$$

$$F^B\left(\frac{1}{2}\right) < \frac{1}{2} - \left(\frac{1}{2} + c\right) \cdot f^B(\frac{1}{2}).$$

Thus, by (20),

$$\begin{aligned}
\frac{\pi_B - \pi_A}{2} &> \frac{1 - \left[\frac{1}{2} - \left(\frac{1}{2} + c\right) \cdot f^B(\frac{1}{2})\right]}{f^B(\frac{1}{2})} = \\
&= \frac{\frac{1}{2}}{f^B(\frac{1}{2})} + \left(\frac{1}{2} + c\right) \geq \frac{\frac{1}{2}}{\frac{1}{1+2c}} + \left(\frac{1}{2} + c\right) = 1 + 2c.
\end{aligned}$$

Since $\frac{\pi_A+\pi_B}{2}=\frac{1}{2}$,

$$\pi_A < -\frac{1}{2} - 2c \text{ and } \pi_B > \frac{3}{2} + 2c. \tag{23}$$

Suppose first that $\eta = 0$. I conclude that

$$\begin{aligned}
& \int_0^1 l(|-x-c+1|, |x-c-1|) f^{1/2}(x)dx = \\
& \int_0^{1/2} l(|-x-c+1|, |x-c-1|) f^{1/2}(x)dx + \\
& \int_{1/2}^1 l(|-x-c+1|, |x-c-1|) f^{1/2}(x)dx \\
< \ & \frac{1}{2} l(1-c, 1+c) + \frac{1}{2} l\left(\max\left\{\frac{1}{2}-c, c\right\}, \frac{1}{2}+c\right) \\
< \ & \frac{1}{2} l(|-\pi_A - c + 1|, |\pi_A - c - 1|) + \frac{1}{2} l(|-\pi_B - c + 1|, |\pi_B - c - 1|).
\end{aligned}$$

The first inequality follows from $c < \frac{1}{2}$ and the assumption that the loss function l is increasing in both variables, while the second inequality follows from (23) and the assumption that the loss function l is increasing in both variables.

Similarly,

$$\begin{aligned}
& \int_0^1 l(|-x-c|, |x-c|) f^{1/2}(x)dx = \\
& \int_0^{1/2} l(|-x-c|, |x-c|) f^{1/2}(x)dx + \int_{1/2}^1 l(|-x-c|, |x-c|) f^{1/2}(x)dx \\
< \ & \frac{1}{2} l\left(\frac{1}{2}+c, \max\left\{c, \frac{1}{2}-c\right\}\right) + \frac{1}{2} l(1+c, 1-c) \\
< \ & \frac{1}{2} l(|-\pi_A - c|, |\pi_A - c|) + \frac{1}{2} l(|-\pi_B - c|, |\pi_B - c|).
\end{aligned}$$

This completes the proof due to the assumption that f is up to a small enough δ in the interval $[-1/2, 1/2]$.

For $\eta > 0$, the result follows now from from the result for $\eta = 0$ and (23). Indeed, the assumption that f is up to a small enough δ in the interval $[-1/2, 1/2]$ and (23) imply that conventional arbitration is more outcome-accurate than final-offer arbitration also when the arbitrator recognizes the state of the world.

8 Appendix B

Proof of Proposition 1: I shall first show that there exists an equilibrium of the continuation game beginning in period 2 under an additional exogenous constraint on final offers. Suppose that the final offers have to be chosen from a finite grid of the interval $[-C, 1+C]$:

$$\pi_A, \pi_B \in \left\{ -C + \frac{k}{n}(1+2C) : k = 0, ..., n \right\} \tag{24}$$

for some $n = 1, 2, ...$ Under this additional assumption, the existence of equilibria follows from Kakutani's fixed point theorem. Indeed, one can simply take a fixed point of the correspondence whose domain (and range) consist of the mixed actions π_A of agent A, the mixed actions of each type of agent B (denoted by $\pi_{B,a}$ and $\pi_{B,b}$, respectively), and the arbitrator's beliefs contingent on each offer from the grid. The correspondence assigns to every triple π_A, $\pi_{B,a}$, $\pi_{B,b}$, and to the arbitrator's beliefs, the actions that are best responses to the opponent's action(s) and to the arbitrator's optimal decision given the beliefs; the correspondence also assigns the beliefs determined by the Bayes rule to any offer that is used with positive probability, and the set of all possible beliefs to any offer that is used with probability 0.

I shall now show that condition (24) is dispensable. This relies on the following "limit" argument. For any given n, the equilibrium mixed actions π_A^n, $\pi_{B,a}^n$ and $\pi_{B,b}^n$ are probability (Borel) measures on $[-C, 1+C]$. The space of all (Borel) probability measures is metrizable and compact in the weak*-topology. One can therefore assume (passing to a subsequence if necessary) that these mixed actions converge (in the weak*-topology) as n goes to ∞ to some measures π_A, $\pi_{B,a}$, and $\pi_{B,b}$. For details, see the Riesz Theorem (Dudley 1989, Theorem 7.4.1), the Banach-Alaoglu Theorem (Rudin 1973, Theorem 3.15) and Rudin 1973, Theorem 3.16.

By Dudley (1989), Theorem 10.2.2, $\pi_{B,a}$ and $\pi_{B,b}$ determine conditional probabilities $\mu(a \mid \pi)$ and $\mu(b \mid \pi)$ with $\mu(a \mid \pi) + \mu(b \mid \pi) = 1$, unique up to a set of final offers π that are used with probability 0 by both $\pi_{B,a}$ and $\pi_{B,b}$. The convergence of $(\pi_{B,a}^n, \pi_{B,b}^n)$ to $(\pi_{B,a}, \pi_{B,b})$ implies that for almost every π (i.e., except a set of final offers π that are used with probability 0 by both $\pi_{B,a}$ and $\pi_{B,b}$) and every $\varepsilon > 0$, there exists a $\delta > 0$ such that if: (i) n is large enough, (ii) $\left|\pi' - \pi\right| < \delta$, and (iii) π' is of the form (24), then $\mu^n(a \mid \pi')$ and $\mu^n(b \mid \pi')$ differ by at most ε from $\mu(a \mid \pi)$ and $\mu(b \mid \pi)$, respectively. One can now easily verify that π_A, $\pi_{B,a}$ and $\pi_{B,b}$ together with $\mu(a \mid \pi)$ and $\mu(b \mid \pi)$ (say, $\mu(a \mid \pi) := 1$ and $\mu(b \mid \pi) := 0$ for final offers π that are used with probability 0) is an equilibrium of the continuation game beginning in period 2.

Thus, for every value β of the probability that agent B is of type b, the continuation game beginning in period 2 has an equilibrium. It is easy to see that the correspondence which assigns to every β the set of equilibrium strategies π_A, $\pi_{B,a}$, and $\pi_{B,b}$ is upper hemi-continuous (assuming, of course, that the space of strategies π_A, $\pi_{B,a}$, and $\pi_{B,b}$ is equipped with the weak*-topology). This in turn implies that the correspondence which assigns to every β the set of equilibrium payoff vectors in the continuation game beginning in period 2 is also upper hemi-continuous.

Thus, there also exists a triple β, E_β, and t_β which maximizes agent A's payoff (in the two-stage game) over all triples β, E_β, and t_β consisting of a belief $\beta \geq 1/2$, an equilibrium E_β of the continuation game beginning in period 2 given this belief β, and a settlement payment t_β which makes type a of agent B indifferent between accepting t_β and filing an request for arbitration (anticipating E_β). Denote by β^* the β at which the maximum is reached, and by u this maximum payoff of A. Now, take any equilibrium for $\beta = 1$ (say, E_1), and consider the minimum settlement payment that makes type b of agent B indifferent between accepting the payment and filing a request for arbitration (anticipating E_1). Denote this settlement payment by v.

If $u \geq -v$, then the two-stage game has an equilibrium in which A offers the settlement payment t_{β^*}. This payment is accepted by type a of agent B with the probability q given by (6) from Section 3 for $\gamma = \beta^*$, and is rejected with probability 1 by type b of agent B. Agents anticipate the equilibrium E_{β^*} contingent on the rejection of this settlement payment. Off the equilibrium path, agents anticipate any E_β contingent on the rejection of the settlement payment t_β for any other $\beta \neq \beta^*$; furthermore, t_β is accepted by type a of agent B with the probability q given by (6) for $\gamma = \beta$, and is rejected with probability 1 by type b of agent B. Agents anticipate

E_1 contingent on the rejection of any settlement payment higher than all t_β; any of these settlement payments is accepted with probability 1 by type a, and accepted or rejected with probability 1 by type b, depending on whether this type prefers this settlement payment or the payoff in the two-stage game contingent on filing a request for arbitration (anticipating that E_1will be played in period 2). Finally, agents anticipate $E_{1/2}$ contingent on the rejection of any settlement payment lower than all t_β, and both types of agent B reject these payments with probability 1.

If $u \leq -v$, then the two-stage game has an equilibrium in which A offers the settlement payment v. This payment is accepted by both types of agent B with probability 1. Off the equilibrium path, the play is defined in the same way as for $u \geq -v$.

9 References

Ashenfelter O. (1987), "Arbitrator Behavior," *American Economic Review* **77**, 342-346.

Ashenfelter O. and D.E. Bloom (1984), "Models of Arbitrator Behavior: Theory and Evidence," *American Economic Review* **74**, 111-124.

Ashenfelter O., J. Currie, H. Farber, and M. Spiegel (1992), "An Experimental Comparison of Dispute Rates in Alternative Arbitration Systems," *Econometrica* **60**, 1407-1433.

Banks J.S. and J. Sobel (1987), "Equilibrium selection in signaling games," *Econometrica* **55**, 647-661.

Bloom D.E. and C.L. Cavanagh (1986), "An Analysis of the Selection of Arbitrators," *American Economic Review* **76**, 408-422.

Brams S.J. and S. Merrill III (1983), "Equilibrium Strategies for Final-Offer Arbitration: There is No Median Convergence," *Management Science* **29**, 927-941.

Brams S.J. and S. Merrill III (1986), "Binding versus Final-Offer Arbitration: A Combination is Best," *Management Science* **32**, 1346-1355.

Çelen B. (2003), "Final Offer Arbitration with Multiple Issues," mimeo, New York

University.

Chatterjee K. (1981), "Comparison of Arbitration Procedures: Models with Complete and Incomplete Information," *IEEE Transactions on Systems, Man Cybernetics,* SMC-**11**, 101-109.

In-Koo Cho and D. M. Kreps (1987), "Signaling games and stable equilibria," *Quarterly Journal of Economics* **102**, 179-221.

Crawford V.P. (1979), "On Compulsory Arbitration Schemes," *Journal of Political Economy* **87**, 131–160.

Crawford V.P. (1981), "Arbitration and Conflict Resolution in Labor-Management Bargaining," *American Economic Review, P&P* **71**, 205–210.

Deck C. and A. Farmer (2003), "Bargaining with Asymmetric Uncertainty: Arbitration Mechanisms Compared," mimeo, the University of Arkansas.

Deck C. and A. Farmer (2007), "Bargaining over an Uncertain Value: Arbitration Mechanisms Compared," *Journal of Law, Economics, and Organization* **23**, 547-579.

Dewatripont M and J. Tirole (1999), "Advocates", *Journal of Political Economy* **107**, 1-39.

Dickinson D. (2004), "A Comparison of Conventional, Final-Offer, and Combined Arbitration for Dispute Resolution," *Industrial and Labor Relations Review* **57**, 288-301.

Dudley R.M. (1989): *Real Analysis and Probability*, Wadsworth and Brook.

Dworkin J.B. (1986), "Salary Arbitration in Baseball: An Impartial Assesment After Ten Years," *Arbitration Journal* **41**, 63-69.

Farber H.S. and H. Katz (1979), "Interest Arbitration, Outcomes, and the Incentive to Bargain," *Industrial and Labor Relations Review* **33**, 55-63.

Farber H.S.(1980a), "Does Final-Offer Arbitration Encourage Bargaining?" in the Proceedings of the 33rd Annual Meeting in Industrial Relations.

Farber H.S. (1980b), "An Analysis of Final-Offer Arbitration," *Journal of Conflict Resolution* **24**, 683-705.

Farber H.S. and M.H. Bazerman (1986), "The General Basis of Arbitrator Behavior: An Empirical Analysis of Conventional and Final-Offer Arbitration," *Economet-*

rica **24**, 1503-1528.

Farmer A. and P. Pecorino (1998) "Bargaining with Informative Offers: An Analysis of Final Offer Arbitration," *Journal of Legal Studies* **27**, 415-32.

Farmer A. and P. Pecorino (2003) "Bargaining with Voluntary Transmission of Private Information: Does the Use of Final Offer Arbitration Impede Settlement?" *Journal of Law, Economics and Organization* **19**, 64-82.

Feigenbaum C. (1975) "Final-Offer Arbitration and the Chilling Effect," *Industrial Relations* **14**, 302-310.

Feuille P. (1975), "Final-Offer Arbitration: Better Theory than Practice," *Industrial Relations* **14**, 311-317.

Gibbons R. (1988), "Learning in Equilibrium Models of Arbitration," *American Economic Review* **78**, 896-912.

Lester R.A. (1984): *Labor Arbitration in State and Local Government (An Examination of Experience in Eight States and New York City)*, Industrial Relation Section, Firestone Library, Princeton University, Princeton, New Jersey.

Najita J.M. and J.L. Stern (2001): *Collective Bargaining in the Public Sector: The Experience of Eight States*, M.E. Sharpe, Armonk, New York.

Rudin W. (1973): *Functional Analysis*, McGraw-Hill.

Samuelson W.F. (1991), "Final-Offer Arbitration under Incomplete Information," *Management Science* **37**, 1234-1247.

Stevens C. (1966), "Is Compulsory Arbitration Compatible with Bargaining?" *Industrial Relations* **5**, 38-52.