Three Procedures for Inducing Honesty in Bargaining

D. Marc Kilgour
Department of Mathematics
Wilfrid Laurier University
Waterloo, ON N2L 3C5
CANADA
mkilgour@wlu.ca

Steven J. Brams
Department of Politics
New York University
New York, NY 10003
USA
steven.brams@nyu.edu

Todd R. Kaplan*
Department of Economics
University of Haifa
Haifa 31905
ISRAEL
Dr@ToddKaplan.com

ABSTRACT

A bargaining procedure, or mechanism, is a set of rules for two bargainers to follow as they make offers in order to reach a mutually satisfactory agreement on, say, a price. The efficiency of a mechanism is the expected surplus it delivers to the bargainers, relative to the surplus that a social planner would deliver, or that the bargainers themselves might achieve if they truthfully revealed their reservation prices. A theoretical limit on this efficiency is known, as is a specific procedure that achieves this maximum. But this procedure induces players to make offers that do not truly reflect their reservation prices. This paper discusses three procedures that induce honest offers, although they necessarily fail to achieve maximum efficiency. Each procedure has its own characteristics and costs, and each may have some uses in particular circumstances.

Categories and Subject Descriptors

H.5.3 [**Group and Organization Interfaces**]: Organizational design; I.2.6 [**Learning**]: Knowledge Acquisition; J.4 [**Social and Behavioral Sciences**]: Economics

General Terms

Economics

Keywords

Bargaining, Incomplete information, Truth-telling mechanisms

1. INTRODUCTION

How to induce players to go to their "bottom lines" in bargaining is an age-old problem. In the context of sealed-bid auctions, the second-price, or Vickrey, auction [12] is a solution; the winner is the high bidder, but she pays only the second-highest bid, rendering what the winner pays independent of her own bid. The extension of this idea to Vickrey[12]-Clarke[5]-Groves[6] (VCG) mechanisms likewise induces honesty, because the settlement does not depend directly on what any player offers. Similarly, Brams and Kilgour [3] show how to solve the housemates problem using bidding and a specially designed procedure, the "gap procedure," which they argue motivates each housemate to bid honestly.[1] For bilateral bargaining, however, bidding procedures are difficult to adapt, mainly because of the constant-sum nature of the problem.

In fact, Myerson and Satterthwaite [9] prove that when two bargainers haggle over the price of some good or service, they inevitably miss some feasible trades because each—to maximize expected return—exaggerates his or her offer in opposite directions. Chatterjee and Samuelson [4] propose a simple mechanism, in which the bargainers' offers are averaged, and show that there is a simple symmetric equilibrium in which exaggeration is piecewise linear in the bargainers' *reservation prices*—the settlement prices that would make each bargainer indifferent between an agreement or none. In that game, the final price is an average of the offers if they *overlap* (i.e., if the buyer's offer does not fall below the seller's); otherwise, there is no agreement, and the players get nothing. Under certain conditions, this equilibrium achieves the maximum possible efficiency. Leininger et al. [7] show that the same game has an infinity of asymmetric equilibria different from the linear symmetric equilibrium of Chatterjee and Samuelson. For background information on mechanism design, see [10].

In this paper, we review two mechanisms proposed by Brams and Kilgour [2] that induce two bargainers to be truthful, i.e., to reveal their reservation prices. These include a "bonus procedure," in which a third party induces the bargainers to be truthful by paying them a bonus when their bids criss-cross, and a "penalty procedure" in which even overlapping offers do not *always* result in an agreement; the probability depends on the amount of overlap. Then we describe a new two-stage mechanism that induces truth-telling in the first of two stages [1]. The strategy of truthfully revealing one's reservation price to a referee in stage 1 weakly dominates any other strategy. If the 1st-

*Alternative address: School of Business and Economics, University of Exeter, Exeter, UK.

[1] In the housemates problem, each of n housemates is to be assigned a room in a house, and charged a rent for the room that does not exceed his or her willingness to pay. Moreover, the total of the assigned rents must equal the rent for the house. The gap procedure assumes that every housemate bids on every room, and allows for a simple feasibility calculation and an assignment that maximizes the total surplus.

TARK 2011, July 12-14, 2011, Groningen, The Netherlands.

stage offers overlap, there is the *potential* for an agreement, which is realized—at the mean of the 2^{nd}-stage offers—if *both* bargainers' offers fall in the overlap interval; if *only one* bargainer's offer falls in this interval, this offer becomes the agreement with probability $\frac{1}{2}$, but otherwise not; if *neither* bargainer's offer falls in the overlap interval, there is no agreement (with certainty).

As we will show, this new procedure, like the penalty procedure, is not as efficient as the Chatterjee-Samuelson, in part because of the random draw when exactly one offer falls in the overlap interval. But even if no agreement is *realized*, the new procedure does reveal—if it goes on to stage 2—that the reservation prices of the bargainers *allow* for a mutually profitable agreement. This information may be of benefit to others, although our bargainers must assign probability zero to any benefits to themselves. Another positive feature of our mechanism is that there is a positive probability of settlement even for extreme reservation values. In contrast, bargainers with extreme values have no incentive to bargain under the Chatterjee-Samuelson procedure, because they know in advance that there is no possibility of agreement.

2. INDUCING HONESTY

The context of our analysis is the possible sale of an object by a Seller to a Buyer. If they can agree on a price, p, the object will be transferred from Seller to Buyer, and the Seller will receive p as compensation. Of course, Seller prefers a higher p, whereas Buyer prefers a lower p. If they cannot agree on a p, there is no sale.

Both bargainers have a reservation price (true value) for the object. The Seller's reservation price, S, is modeled as the value of a random variable with cumulative distribution function F_S. The Buyer's reservation price, B, is modeled as the value of a random variable with cumulative distribution function F_B. Both F_S and F_B have support $[C, D]$. Both players' reservation prices are private information. If a sale takes place, the reservation prices determine the players' utilities. Specifically, if there is a sale at price p, Buyer will receive $B-p$ and Seller will receive $p-S$. If there is no sale, both players receive 0. Note that these assumptions imply that the utilities are quasi-linear and that both players are risk-neutral.

Definition 1. One strategy for a player (weakly) *dominates* another strategy for that player if, no matter what strategy is chosen by the opponent, the expected utility of the first strategy is never less, and sometimes more, than the expected utility of the second. A (Bayesian-Nash) *equilibrium* is a profile of strategies with the property that each player's strategy maximizes that player's expected utility, given that the opponent plays according to its specified strategy.

Brams and Kilgour [2] propose and analyze two mechanisms that induce the Buyer and the Seller to be truthful. In both of these mechanisms, Buyer and Seller simultaneously reveal their bids, $\widehat{B}(B)$ for Buyer and $\widehat{S}(S)$ for Seller. The referee then carries out a calculation that determines the bargainers' final utilities, based on who possesses the object, and what payments are made. Both procedures assume that, whenever $\widehat{B} \geq \widehat{S}$, the exchange price is $p = \frac{\widehat{B}+\widehat{S}}{2}$. Both procedures induce honesty in the sense that $\widehat{B}(B) = B$ and $\widehat{S}(S) = S$ are dominant strategies for the players, so they cannot do better than to be truthful. It follows that the strategy profile $(\widehat{B}(B), \widehat{S}(S)) = (B, S)$ forms an equilibrium.

In the Bonus Procedure, each player receives a bonus that depends only on the bids. However, the bonus is paid only if there is a sale; when there is no sale, there is no bonus. Note that the bonus must be paid by some outside party, who does not directly benefit from the sale (but must have some interest in learning the true values of B and S). There is a possibility of "taxing" Buyer and Seller prior to the transaction in order to cover the expected cost of the bonus, but this idea is difficult to implement as Buyer and Seller may be unwilling to pay it *after* they learn their reservation prices [2].

The Bonus Procedure specifies a bonus of $\frac{\widehat{B}-\widehat{S}}{2}$ to *each* bargainer whenever $\widehat{B} \geq \widehat{S}$. This amount is substantial because, providing bidding is honest, it doubles each player's expected utility whenever there is a sale. When $\widehat{B} \geq \widehat{S}$, the players are said to produce a *surplus* of $\widehat{B} - \widehat{S}$. When bidding is honest, the price $p = \frac{\widehat{B}+\widehat{S}}{2} = \frac{B+S}{2}$ implies that the surplus is split equally between the two bargainers. But, under the Bonus Procedure, whenever there is a sale each player effectively receives the entire surplus for himself or herself. Of course, under the Bonus Procedure the cost of inducing honesty is exactly the cost of the bonuses.

A second mechanism proposed and analyzed by Brams and Kilgour [2] is the Penalty Procedure. This procedure begins by determining whether a sale is feasible by considering the condition $\widehat{B} \geq \widehat{S}$. If this condition fails, there is no possiblity of a sale, and the procedure ends. If it holds, the price $p = \frac{\widehat{B}+\widehat{S}}{2}$ is determined. But whether the settlement is implemented is determined by a random device; in other words, settlements are implemented probabilistically. (Actually, independent devices can be used for Buyer and for Seller.)

Formally, the Penalty Procedure requires that if $\widehat{B} \geq \widehat{S}$, the Buyer's expected utility equal $(B-p)(\widehat{B}-\widehat{S})c_B$, where the constant c_B is chosen so that $(\widehat{B}-\widehat{S})c_B \leq 1$. In our context, this condition is simply that $c_B = \frac{1}{D-C}$. (Smaller positive values of c_B are possible, but we ignore them as they yield even less expected utility.) Similarly, the Penalty Procedure requires that if $\widehat{B} \geq \widehat{S}$, the Seller's expected utility be $(p-S)(\widehat{B}-\widehat{S})c_S$, where a suitable choice of c_S is $c_S = \frac{1}{D-C}$. To summarize, under the Penalty Procedure, if the bargainers "find" a deal, that deal is implemented with a probability that increases in the separation of their two bids.

It is useful to associate the Buyer's bid, $\widehat{B}$, with the interval $[C, \widehat{B}]$ of values of $\widehat{S}$ that the Buyer can accept. Similarly, the Seller's bid, $\widehat{S}$, is associated with the interval $[\widehat{S}, D]$ of values of $\widehat{B}$ that the Seller can accept. Then, when $\widehat{B} \geq \widehat{S}$, we say that the players' bids *overlap*. In particular, under the Penalty Procedure the implementation probability for a feasible settlement is proportional to the amount or degree of overlap.

It can be shown that, under the Penalty Procedure, truthful bidding, $\widehat{B}(B) = B$ and $\widehat{S}(S) = S$, is a dominant strategy for each player, and the strategy profile $(\widehat{B}(B), \widehat{S}(S)) = (B, S)$ is an equilibrium. However, the cost of inducing honesty, namely, the loss of surplus from settlements that are feasible but not implemented, can be severe—if $B > S$ but $B - S$ is small, then bidding under the Penalty Procedure will rationally be truthful, but even with exact honesty, the probability that any settlement is implemented will be small.

3. A NEW MECHANISM

The mechanism proposed in [1] is a two-stage procedure:

Stage 1. The players submit *reserves* to the referee: Seller submits $\widehat{S}$ and Buyer submits $\widehat{B}$. The reserves may or may not equal the corresponding reservation prices (i.e., the 1st-stage submissions are not necessarily truthful). If $\widehat{S} \leq \widehat{B}$, the overlap interval is $[\widehat{S}, \widehat{B}]$, and the procedure moves to stage 2. If $\widehat{S} > \widehat{B}$, the reserves do not overlap, there is no settlement, and the procedure ends.

Stage 2. The players submit *offers* to the referee: Seller submits s, and Buyer submits b. If both s and b fall in the overlap interval defined in stage 1, there is a sale at price $p = \frac{s+b}{2}$. If only one of s and b falls in the overlap interval, the name of one player is selected at random; if the selected player's offer is the one in the overlap interval, then it is sale price; if not, there is no sale. If neither offer is in the overlap interval, there is no sale.

Note that our notation for bids in the Bonus Procedure and the Penalty Procedure, $\widehat{B}$ and $\widehat{S}$, is the same as our notation for first-stage strategies, or *reserves*, in the present procedure. As we will see, players are motivated to be truthful in both cases. However, this is not the case with the second-stage strategies, or *offers*.

As usual, the new mechanism determines (i) whether there is a sale and (ii) if there is a sale, at what price. Each player has private knowledge of his or her own (true) reservation price (B or S) prior to stage 1, and uses this information to choose its strategy: ($(\widehat{S}, s)$ for Seller; $(\widehat{B}, b)$ for Buyer). Thus, a strategy for Seller is now a pair of functions $\widehat{S}(S)$ and $s(S)$ that give the values of Seller's strategic variables as a function of its reservation price. Similarly, Buyer's strategy can be thought of as two functions, $\widehat{B}(B)$ and $b(B)$. We assume that all four of these strategy functions are differentiable and increasing in the players' reservation prices.

To represent this mechanism, we use two functions,

$$t : \mathbb{R}^4 \longrightarrow [0, 1] \text{ and } p : \mathbb{R}^4 \longrightarrow \mathbb{R}$$

with the interpretation that $t(\widehat{S}, s, \widehat{B}, b)$ is the probability that an agreement is reached if the 1$^{\text{st}}$-stage reserves are $\widehat{S}$ and $\widehat{B}$, and the 2$^{\text{nd}}$-stage offers are s and b; similarly, $p(\widehat{S}, s, \widehat{B}, b)$ is the price. Note that both $\widehat{S}$ and s are functions of Seller's true reservation price S; we have written $\widehat{S}$ instead of $\widehat{S}(S)$, and s instead of $s(S)$, for notational simplicity. Observe that, if $t = 0$, the value of p is irrelevant. Using the functions t and p, we can describe our mechanism formally, as follows:

$$(t,p) = \begin{cases} (1, \frac{s+b}{2}) & \text{if } \widehat{S} \leq s, b \leq \widehat{B}, \\ (\frac{1}{2}, b) & \text{if } \widehat{S} \leq b \leq \widehat{B} < s, \\ (\frac{1}{2}, s) & \text{if } b < \widehat{S} \leq s \leq \widehat{B}, \\ (0,0) & \text{otherwise.} \end{cases} \quad (1)$$

We will assume below that players always choose 2$^{\text{nd}}$-stage strategies that are at least as "aggressive" as their 1$^{\text{st}}$-stage strategies, i.e., that $b \leq \widehat{B}$ and $s \geq \widehat{S}$. This assumption is innocuous because, by (1), the players' payoffs are certain to be 0 if it does not hold, whereas if it does hold each payoff is never less than 0 and exceeds 0 with positive probability.

We can construct a strategically equivalent mechanism by retaining stage 1 and replacing stage 2 by

Stage 2′. One player, Seller or Buyer, is chosen at random. If Seller is chosen, and if Seller's 2$^{\text{nd}}$-stage offer s satisfies $s \leq \widehat{B}$, then the transaction takes place at price $p = s$; if $s > \widehat{B}$, then there is no sale. Similarly, if the player chosen is Buyer, and if Buyer's 2$^{\text{nd}}$-stage offer b satisfies $\widehat{S} \leq b$, then the transaction takes place at price $p = b$; if $b < \widehat{S}$, there is no transaction.

We assume that the random selection of a player in stage 2′ is independent of the players' reservation prices. The equivalence of the two mechanisms arises because, if stage 2′ is followed, the players' expected utilities are exactly as in (1). For example, if $\widehat{S} \leq s, b \leq \widehat{B}$, then Seller's expected utility is

$$\frac{1}{2}(s - S) + \frac{1}{2}(b - S) = \frac{s + b}{2} - S = p - S,$$

where p is determined by the first condition of (1). A similar relation holds for Buyer. The verification is immediate if the 2$^{\text{nd}}$-stage offer of only one player, or none, falls in the overlap interval. Below, we will use the stage 2 and stage 2′ formulations interchangeably.

To analyze the new mechanism, we must refine our definition of truth-telling, as follows:

Definition 2. Seller's strategy $(\widehat{S}, s)$ is *truth-telling* if $\widehat{S}(S) = S$ for all $S \in [C, D]$. Buyer's strategy $(\widehat{B}, b)$ is *truth-telling* if $\widehat{B}(B) = B$ for all $B \in [C, D]$. A strategy profile $(\widehat{S}, s; \widehat{B}, b)$ is a *truth-telling equilibrium* if it is an equilibrium and both players' strategies are truth-telling.

We repeat that truth-telling refers to the players' reserve strategies (1$^{\text{st}}$ stage), not their offer strategies (2$^{\text{nd}}$ stage).

Definition 3. A cumulative distribution function for Buyer, $F_B(x)$ satisfies a *monotone hazard rate condition for Buyer* iff $\frac{d}{dx} \frac{F'_B(x)}{1-F_B(x)} \geq 0$ for all $x \in [C, D]$. A cumulative distribution function for Seller, $F_S(x)$, satisfies a *monotone hazard rate condition for Seller* iff $\frac{d}{dx} \frac{F'_S(x)}{F_S(x)} \leq 0$ for all $x \in [C, D]$.

If Buyer's and Seller's distributions satisfy appropriate monotone hazard rate conditions, the new procedure has a truth-telling equilibrium, as demonstrated in [1].

THEOREM 1. *Any strategy of Seller, $(\widehat{S}, s)$, is weakly dominated by the truth-telling strategy (S, s). Any strategy of Buyer, $(\widehat{B}, b)$, is weakly dominated by the truth-telling strategy (B, b). If $F_S(\cdot)$ and $F_B(\cdot)$ are strictly increasing and satisfy appropriate monotone hazard rate conditions, then there is a truth-telling equilibrium in which the players' 2nd-stage offers, s^* and b^* are the solutions of $1 - F_B(s) = (s - S)F'_B(s)$ and $F_S(b) = (B - b)F'_S(b)$, respectively. Moreover, there are no truth-telling equilibria other than $(S, s^*; B, b^*)$.*

PROOF (SKETCH). We use the procedure of stage 2′ to calculate Seller's expected utility. Seller knows the value of S and determines strategy $(\widehat{S}, s)$ to maximize this expected utility. The expectation, taken with respect to Buyer's value

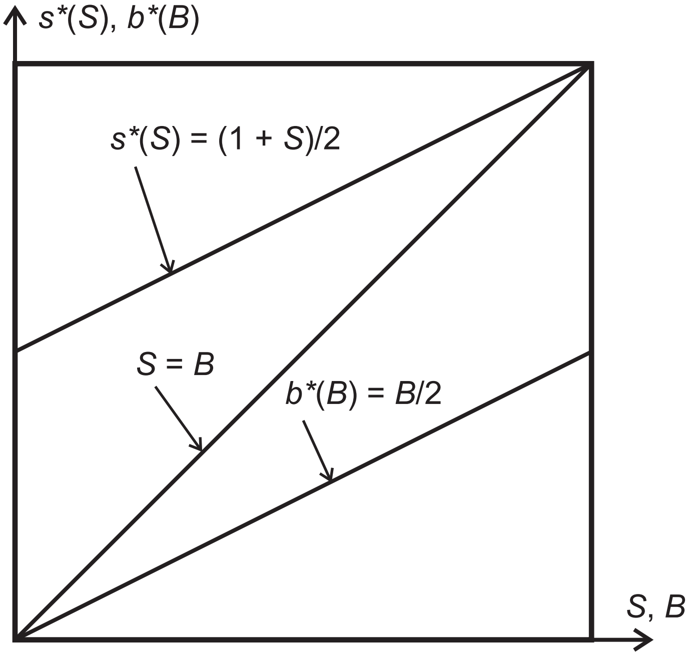

Figure 1: Offer Strategies: Example 1

B and the random selection of Seller or Buyer can be shown to equal

$$\frac{1}{2}\int_{b^{-1}(\widehat{S})}^{D}(b(B)-S)\,dF_B(B)+\frac{1}{2}\int_{\widehat{B}^{-1}(s)}^{D}(s-S)\,dF_B(B),$$

where the first integral is associated with the random selection of Buyer (so the price is b) and the second with the selection of Seller (so the price is s). Note that the first integral depends on $\widehat{S}$ but not s, and the second depends of s but not $\widehat{S}$. There is a similar expression for Buyer's expected utility.

The first integral above can be shown to be a strictly increasing function of $\widehat{S}$ if $\widehat{S} < S$ and a strictly decreasing function of $\widehat{S}$ if $\widehat{S} > S$. Therefore, it is maximized at $\widehat{S} = S$. A similar conclusion can be drawn for Buyer. The fact that $\widehat{B} = B$ can then be used to simplify the second integral given above, and the determination of the maximizing choice s^* is then straightforward. □

The proof that $(S, s^*; B, b^*)$ is an equilibrium thus relies on the fact that non-truth-telling strategies are weakly dominated by truth-telling strategies, so there must be a truth-telling equilibrium. Then the offer strategies are obtained by maximizing the players' expected utilities under the assumption of truth-telling. To understand why truth-telling dominates, note that each player benefits from maximizing the degree of overlap. In other words, a player gains by making the overlap interval as wide as possible, up to its reservation price—Seller from below and Buyer from above—in order to ensure, insofar as possible, that the 2^{nd}-stage bids, s and b, fall in the interval, thereby meeting a necessary condition for an agreement.

In fact, truthfully reporting one's reservation price in stage 1 is analogous to bidding one's reservation price in a Vickrey auction: Just as a player cannot win in a Vickrey auction without being the highest bidder, a bargainer cannot reach a settlement unless there is an overlap interval, leading to stage 2. In each case, a player goes to its bottom line for two reasons: (i) failing to do so in stage 1 could preclude a favorable outcome in stage 2 (or cause an unfavorable outcome) and (ii) once in stage 2, the outcome does not depend on what the player reported in stage 1.

A player's utility, if positive, does not depend on the reserves, $\widehat{S}$ and $\widehat{B}$, submitted in stage 1 but, instead, on its bid, s or b, in stage 2. The fact that a player's reserve is independent of its offer is why it can "afford" to be truthful in stage 1. In fact, a player cannot do worse by reporting its reservation price truthfully in stage 1, and may do better. In conclusion, under our mechanism each player has an incentive to report its reservation prices truthfully.

The story is different, however, in stage 2: Each player will have an incentive to shade its offer, depending on the distribution of the opponent's reservation price. In the next section, we illustrate, for two particular distribution functions, how much shading is optimal.

4. EXAMPLES

For our examples, we assume $C = 0$ and $D = 1$, so that $0 \leq S, B \leq 1$.

Example 1 Uniform distribution: $F_S(x) = F_L(x) = x$.

The players' optimal offers, $s^*(S) = \frac{1+S}{2}$ and $b^*(B) = \frac{B}{2}$, are shown in Figure 1. Each of these strategies halves the distance between the reservation price itself (shown as the 45° line) and the extreme endpoint, 1 for Seller and 0 for Buyer, of the bargaining range. For example, if Seller's reservation price is $S = \frac{1}{3}$, Seller would be willing to accept any price between $\frac{1}{3}$ and 1. If it happens that Seller's offer is selected as the price in stage 2, Seller must balance a low offer (almost certain to result in a sale, as it is likely to fall below Buyer's truthful reserve, but not offering much profit) against a high offer (relatively unlikely to result in a sale, but offering a great deal profit if it succeeds). The greatest net gain occurs when Seller uses the strategy $s^*(\frac{1}{3}) = \frac{2}{3}$, which lies at the midpoint. However, notice that, as a consequence, Seller never offers below $\frac{1}{2}$, and Buyer never offers above $\frac{1}{2}$.

The equilibrium determined above is based on weakly but not strictly dominant strategies. To see this, notice that if, for example, Seller's (true) reservation price is $S = \frac{3}{4}$, then at equilibrium Seller's 2nd-stage offer will be $s = s^*(\frac{3}{4}) = \frac{7}{8}$, and there will be a sale with probability $\frac{1}{2}$ if $B > \frac{7}{8}$; otherwise, there is no possibility of a sale. It follows that, in the 1st stage, Seller will be indifferent between reporting $\frac{3}{4}$ and, say, $\frac{5}{8}$ (provided the 2nd-stage bid remains $s = \frac{7}{8}$). Thus, truthful reporting is weakly but not strictly dominant in the 1st stage.

Figure 2 shows how the equilibrium strategies we have identified interact for all possible values of S and B. A sale occurs with certainty when $B < 2S$ and $B > \frac{1+S}{2}$; these two conditions define the region with darker shading in Figure 2. Notice that this is the region of small values of S and large values of B; the difference between S and B is so great that the offers s^* and b^* fall in the overlap interval. A transaction occurs with probability $\frac{1}{2}$ when $2S < B < \frac{1+S}{2}$ and when $\frac{1+S}{2} < B < \min\{2S, 1\}$, which are the two regions with lighter shading in Figure 2.

In the first of these lighter-shaded regions (lower left), $s^* > B$ but $b^* > S$, so there is a sale at $p = b^*$ when Buyer's name is drawn in stage $2'$, and no sale otherwise. Similarly, in the upper right region, $s^* < B$ and $b^* < S$, so there is a sale at $p = s^*$ when Seller's name is drawn in stage $2'$, and no sale otherwise. In the unshaded region—even the part of it that lies above and to the left of the the 45° line—there is never a sale.

It is instructive to compare the new mechanism with the Chatterjee-Samuelson procedure [4]. The Chatterjee-Samuelson procedure produces a transaction, for certain, if and only if $B \geq S + \frac{1}{4}$, which is the area above the dashed line in Figure 2. For this mechanism the shading in Figure 2 would be dark above and to the left of the dashed line, and unshaded below it. Thus it is clear that each of the two mechanisms sometimes results in sale when the other does not—the Chatterjee-Samuelson mechanism always produces a sale above the dashed line, when the new mechanism sometimes achieves a sale for certain, sometimes with probability 50%, and sometimes results in no sale. On the other hand, it is clear that the new mechanism sometimes results in a sale at times that the Chatterjee-Samuelson mechanism does not.

In particular, for the new mechanism, a sale is always possible even if one player has an extreme value (provided the opponent's value is also extreme. The Chatterjee-Samuelson procedure does not share this feature. For instance, if $S = 0.8$ and $B \geq 0.9$, a sale occurs with probability 0.5 under our mechanism, but the Chatterjee-Samuelson mechanism rules out any possibility of a transaction. In particular, if the Chatterjee-Samuelson procedure were to be applied, a Seller with $S = 0.8$ has no incentive to participate, as no transaction is possible for such an extreme resevation price. Thus, a positive feature of the new mechanism is the potential for trade at all possible values of S and all possible values of B.

Another means to compare mechanisms is to use the expected surpluses they produce. The surplus is the total increase in expected utility of Buyer and Seller after the transaction, if any. For an "ideal" procedure, which produces a settlement whenever the players' reservation prices overlap, the total surplus is

$$\int_0^1 \int_S^1 (B - S)\, dSdB = \frac{1}{6}.$$

Myerson and Satterthwaite [9] demonstrated that no mechanism can produce a larger surplus than the Chatterjee-Samuelson procedure, which gives

$$\int_0^{\frac{3}{4}} \int_{S+\frac{1}{4}}^1 (B - S)\, dBdS = \frac{9}{64}.$$

The surplus from the new mechanism is

$$\frac{1}{2}\int_0^1 \int_{\frac{1+S}{2}}^1 (B - S)\, dBdS + \frac{1}{2}\int_0^{\frac{1}{2}} \int_{2S}^1 (B - S)\, dBdS = \frac{1}{8},$$

which is $\frac{8}{9} = 88.9\%$ of the maximally possible surplus.

The Penalty Procedure of [2] can also be compared to the new mechanism, and the Chatterjee-Samuelson procedure, according to surplus. It can be shown that the most favorable constants result in a surplus of $\frac{1}{12}$, which is $\frac{16}{27} = 59.3\%$ of the maximally possible surplus. Clearly the Penalty Procedure is less efficient on average, but it is noteworthy that whenever a bargain is feasible, there is a small probability of achieving it using the Penalty Procedure.

Example 2 Power distribution: $F_S(x) = x^\alpha$, $F_B(x) = 1 - (1-x)^\beta$, for $\alpha, \beta > 0$.

(It is easy to verify that these distributions satisfy the monotone hazard rate conditions.) Buyer's optimal offer is $b^* = \frac{\alpha B}{1+\alpha}$ and Seller's is $s^* = \frac{1+\beta S}{1+\beta}$, in agreement with Example 1, which corresponds to $\alpha = \beta = 1$. For example, when $\alpha = \beta = 2$, $b^*(B) = \frac{2}{3}B$ and $s^*(S) = \frac{1}{3} + \frac{2}{3}S$, and when $\alpha = \beta = \frac{1}{2}$, $b^*(B) = \frac{1}{3}B$ and $s^*(S) = \frac{2}{3} + \frac{1}{3}S$.

5. CONCLUSION

We have compared three bargaining procedures that induce honesty on the part of bargainers. The Bonus Procedure and the Penalty Procedure have been known for some time, and a new and generally more efficient procedure has recently been discovered. The new procedure is a simple and elegant 2-stage mechanism that induces two bargainers to be truthful in reporting their reservation prices in the 1st stage; if these prices criss-cross, the referee reports that there is an overlap interval, and the bargainers make offers in a 2nd stage. The mean of these offers becomes the settlement if they both fall in the overlap interval. If only one offer does, it is implemented as the settlement price with probability $\frac{1}{2}$, whereas if neither offer does, there is no settlement.

We have also compared the three truth-inducing mechanisms with the Chatterjee-Samuelson mechanism, which is known to be the most efficient possible in the simple context in which we carried out our comparison. The loss of efficiency of the new mechanism is not as great as for the previously known mechanisms. Moreover, the truth-inducing mechanisms do have several positive features, including the possibility of a transaction even for extreme

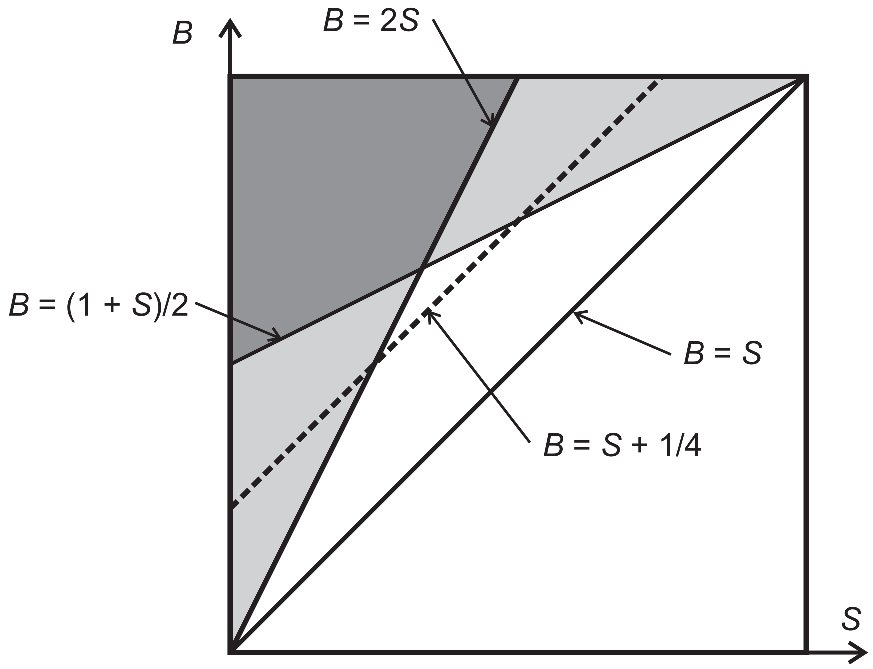

Figure 2: Conditions for a Sale: Example 1

reservation prices. Also, the truth-telling equilibria have some uniqueness properties [1].

The inefficiency of any mechanism stems from the fact that none of them can implement every feasible bargain with certainty. Even the Chatterjee-Samuelson mechanism, which maximizes efficiency without inducing truth-telling, sacrifices many feasible exchanges (see Figure 2). In the new mechanism, a 2$^{\text{nd}}$-stage offer, s or b, become the exchange price only it falls in the overlap interval—and only when the player who made the offer is chosen in the random draw. In the case of the new mechanism, for instance, randomizing the implementation of a single inside offer is the penalty one pays to render a player's reserve independent of its offer in the expected-payoff calculation, thereby making it optimal for the player to report truthfully its reservation price. This independence would be broken, and it would be suboptimal for a player truthfully to report its reservation price, if single inside offers were implemented with certainty.

Of course, a truth-telling mechanism can be based on the Chatterjee-Samuelson procedure and the Revelation Principle [8], but the strategies are not weakly dominant, and the details of the procedure. In Example 1, for instance, one would have a sale iff $\widehat{B} \geq \widehat{S} + \frac{1}{4}$, at price $p = \frac{\widehat{B}+\widehat{S}+\frac{1}{2}}{3}$ [11]. In contrast, using the Chatterjee-Samuelson mechanism directly implies that the players do not reveal their reservation prices except when they are extreme—Buyer when its reservation price is very low, and Seller when its reservation price is very high. By comparison, all three procedures discussed here include always truthfully reporting one's reservation price as a dominant strategy.

Under all of the mechanisms we have studied, the bargainers make simultaneous bids, and the probability of implementation of a settlement is a function of the *degree* of overlap, if any, in the bids: the greater the overlap, the higher this probability.[2] Similarly, the Bonus Procedure has a cost, but it requires an outside agent. Since its cost is comparable to the maximum achievable surplus, it seems appropriate to compare it to the Penalty Procedure. There is an important difference, however; for the Bonus Procedure, the outside agent pays the cost of inducing honesty; for the Penalty Procedure, and the new mechanism, it is the bargainers themselves who bear the cost.

Another feature of the Penalty Procedure is comparable to Chatterjee-Samuelson: The players never learn whether their failure to settle was because (i) their reservation prices did not criss-cross (as in stage 1), or (ii) they did criss-cross but probabilistic implementation prevented a settlement (as in stage 2). In principle, however, they could be told whether (i) or (ii) prevented a settlement; if (ii), they might be motivated to try again (as discussed below).

An advantage of the new mechanism is that the players *always* learn if stage 2 is reached and, therefore, that there is an overlap interval and the potential for a mutually profitable settlement. While our mechanism does not reveal the amount of regret—for example, how close the 2$^{\text{nd}}$-stage offers are to the overlap interval—we see no reason why the values of $\widehat{S} = S$, $\widehat{B} = B, s$, and b could not be revealed by the referee, making public the reason why implementation failed in stage 2.

If the optimality of shading one's "bottom line" in stage 2 is the reason that a settlement eluded the players, this outcome might motivate them, or a third party, to try to find a settlement—though, of course, under our model the

[2]This probability increases linearly in the overlap for the Penalty Procedure. For the new mechanism, as for the Chatterjee-Samuelson procedure, the probability increases as a step function, with value 0 for a small overlap, and 1 for a large enough overlap, and, in the case of the new mechanism, with value $\frac{1}{2}$ for an intermediate overlap.

players must assign probability 0 to this eventuality during the bargaining. But would they in good conscience walk away from the possibility of a mutually profitable settlement that they know exists?[3] Even if they would, perhaps other bargainers in comparable situations would be encouraged to try to reach a settlement by other means, or perhaps an external agent would be motivated to adopt the Bonus Procedure.

6. REFERENCES

[1] S. J. Brams, T. R. Kaplan, and D. M. Kilgour. A simple bargaining mechanism that elicits truthful reservation prices. *Preprint*, New York University.

[2] S. J. Brams and D. M. Kilgour. Bargaining procedures that induce honesty. *Group Decision and Negotiation*, 5(3):239–262.

[3] S. J. Brams and D. M. Kilgour. Competitive fair division. *Journal of Political Economy*, 109(2):418–443.

[4] K. Chatterjee and W. Samuelson. Bargaining under incomplete information. *Operations Research*, 31(5):835–851.

[5] E. Clarke. Multipart pricing of public goods. *Public Choice*, 8:19–33.

[6] T. Groves. Incentives in teams. *Econometrica*, 41:617–631.

[7] W. Leininger, P. B. Linhart, and R. Radner. The sealed-bid mechanism for bargaining with incomplete information. *Journal of Economic Theory*, 48(1):63–106.

[8] R. B. Myerson. Incentive compatibility and the bargaining problem. *Econometrica*, 47:61–73.

[9] R. B. Myerson and M. A. Satterthwaite. Efficient mechanisms for bilateral trading. *Journal of Economic Theory*, 29(2):265–281.

[10] N. Nisan. Introduction to mechanism design (for computer scientists). In N. Nisan, T. Roughgarden, E. Tardos, and V. Vazirani, editors, *Algorithmic Game Theory*, pages 209–241. Cambridge University Press, Cambridge, UK, 2007.

[11] R. Radner and A. Schotter. The sealed-bid mechanism: An experimental study. *Journal of Economic Theory*, 48:179–220.

[12] W. Vickrey. Counterspeculation, auctions, and competitive sealed bid tenders. *Journal of Finance*, 26:8–37, March 1961.

[3] One possible amendment to the new mechanism would involve forcing a settlement at the mean of the reservation prices if there is an overlap interval but no settlement in stage 2. But this is a different mechanism, and truth-telling would not be optimal for the players in stage 1. Moreover, our demonstration above that the optimal 2^{nd}-stage offers are the form $s = \frac{1+\widehat{S}}{2}$ and $b = \frac{\widehat{B}}{2}$ would no longer apply, so neither the reserves nor the offers would be related to the players' (truthful) reservation prices.