

Matching and Efficiency in the Baseball Free-Agent System: An Experimental Examination

Haig R. Nalbantian, *National Economic
Research Associates, Inc*

Andrew Schotter, *New York University*

This article presents the results of an experimental study investigating the problem of allocating heterogeneous indivisible objects using market-like mechanisms. The object of study is the market for professional baseball players in their free-agent year. We investigate both the current free-agency system and a variant of the current system instituted informally by the teams and ruled illegal by arbitrators. We then propose and test a new alternative matching mechanism, which proves to have quite a few desirable characteristics.

I Introduction

This article presents the results of an experimental study of the problem of allocating heterogeneous indivisible objects or services using markets or market-like mechanisms. Examples include the problem of matching medical interns with hospitals, as studied by Roth (1984a), the problem of matching college students with dormitory rooms, or even the problem of matching high school graduates with colleges. In this article, we consider

The work on this article was sponsored by the C. V. Starr Center for Applied Economics and National Economic Research Associates, Inc. We are grateful to Louis A. Guth and Richard Rozek for valuable suggestions and commentary. We give special thanks to Vicky Myroni, who is responsible for writing all of the programs used, and to Ken Rogoza for his technical assistance.

[*Journal of Labor Economics*, 1995, vol. 13, no. 1]
© 1995 by The University of Chicago. All rights reserved.
0734-306X/95/1301-C001\$01.50

another application of this matching problem—the market for free agents in major league baseball.

Our work is inspired by two sources: first, the rapidly growing literature on the “matching problem” (see Roth and Sotomayor 1990), second, some current issues in the operation of baseball’s free-agency system.

The literature on matching is concerned with the development of algorithms that can be used to match people with people or people with objects and with the incentive properties of the mechanisms defined by these algorithms. Accordingly, the first objective of our study is to place some empirical meat on the sophisticated skeletal structure developed by Gale and Shapely (1962), Roth (1982, 1984*a*, 1984*b*, 1985), and others (namely, Shapely and Shubik 1972, Demange and Gale 1985, and Leonard 1983). We wish to determine whether the incentive properties claimed in theory can be observed in the lab. (Similar attempts were made by Harrison and McCabe [1990] and Olson and Porter [1994, in press].)

Our second objective is to investigate the market for free agents in the baseball industry, a market composed of heterogeneous and indivisible “goods” (the services of baseball players). We seek to compare and assess the efficiency and distributional properties of the current free-agency system (and a particular, “complete information” variant of that system) and those of an alternative allocation mechanism whose inspiration can be found in the matching literature.

Our article is organized as follows. In Section II we review, briefly, the history of the free-agency system. In addition, we outline the three alternative mechanisms of interest. This is followed by a review of some of the matching literature. We then present the most important results of our study. Section III describes the actual experiments conducted. A detailed presentation of our results is contained in Section IV, followed in Section V, by a summary and discussion.

II The Free-Agent System

Until the mid-1970s, professional baseball players did not own the rights to their services. These rights were owned exclusively by the teams they were playing for, so that when a player’s contract expired, he did not have the option of shopping around for the best offer. His only option was to sign a contract with the team he was on or sit out the year. Alternatively, he could hope that his team would sell his rights to another team or trade him. This situation was challenged in 1972 in the Curt Flood and Andy Messersmith cases. As a result of these cases, baseball players obtained the property rights to their own services, but only after they have played in the major leagues for 6 years. Under this system, players who have accumulated 6 years of service in “the majors,” and whose contracts have expired, can declare themselves free agents and negotiate with any team that is interested.

The first few years of free agency were tumultuous, characterized by bidding wars for superstars. The then huge contract signed by Catfish Hunter made headlines, and it appeared, at least for a while, that the players were becoming successful at capturing more of the rents available. By the 1980s team owners had become alarmed by the increases in salaries brought about by this new free-agent system, a 1986 suit alleging collusion on the part of the teams' owners was won by the players. The players asserted that instead of bidding against each other for the services of free agents, the teams had agreed not to bid for the players of any team except their own. The arbitrator in the Kirk Gibson case awarded damages to the players for the teams' refusal to deal, and the 1987 and 1988 free-agent markets were also contested as containing facilitating practices that allowed teams to keep salaries artificially low. There again, the arbitrators ruled in the players' favor, awarding sizable damages.

This history suggests that both sides of the industry would like a new mechanism to allocate free agents. The teams would like one that prevents the bidding wars they feel characterize the current system, while the players would like one that is less prone to collusion. This dissatisfaction led us to investigate three distinctly different mechanisms that might be used in the baseball industry. One, the current free-agency system (CFA), presents a laboratory version of what we feel are the salient characteristics of the free-agency system now in place in the major leagues. The next, a complete-information English auction (CIEA), incorporates an information modification of the CFA, which the team owners instituted on a voluntary basis in 1987 as a possible solution to what they felt were drawbacks in the current free-agency system. Finally, we investigate a simultaneous mechanism (SM), which is a generalization of the Walrasian mechanism of Demange and Gale (1985) and which uses the algorithm of Leonard (1983) to make its calculations. Let us explain these three mechanisms in turn.

A Mechanism Types

Free agency (CFA)—The current free-agency system (CFA) can be described as follows: by a given date all eligible players declare whether they are free agents or not. After that date any team is free to call any player and vice versa. The content of these negotiations is private information and cannot be verified. At any time a player is free to accept the latest offer made to him by any team, when he does, his participation in the market is over. Negotiations continue until either all players have agreed to a contract or time runs out. Payoffs are defined according to the terms of the contracts and whether or not a contract has been made.

Thus the current free-agency system constitutes a partial information sequential mechanism since information about the bids made by teams for players is not publicly available while the mechanism is being employed.

Complete information English auction (CILA)—Since the informational asymmetry existing in the current free-agency system can be expected to give an undue advantage to players, one may think of modifying the mechanism so that at any point in time all bids made by any team to any player are available for inspection by everyone. Such a system might be organized as follows. Players and teams sit by computer terminals that contain screens indicating the latest bids by all teams for all players. When a team wishes to bid it enters its bid into its terminal. Bids can be changed. When a player wishes to accept a bid, he enters its acceptance and his participation in the market is over. Bidding continues until all players have made a contract or until time runs out. Clearly such a mechanism is of the full-information sequential variety since all bids made are common knowledge to all participants.

Simultaneous mechanisms (SM)—A simultaneous mechanism might have the following description. On a given day, all teams and players submit bids to a central computer. The bids submitted by the teams would represent the maximum willingness-to-pay (“values”) that any team has for any player. Hence, each team enters a vector of bids, one bid for each player. The bids submitted by the players would represent their reservation prices (“costs”), namely, the minimum price they require in order to play on any given team. Once these bids are submitted, the computer would treat them as if they were the truthful values and costs of the teams and players, respectively. It would then match players and teams so as to maximize the sum of the surpluses generated by any such matching. In addition to matching the players and the teams, the computer would also indicate a *range* in which the salary of the player must be set. Teams and players would then negotiate their salaries within these ranges. Teams and players who fail to come to a negotiated agreement would be sent to arbitration. Teams and players who fail to make a match would remain unmatched.

The motivation for this type of mechanism comes from the matching literature, especially that part dealing specifically with the “marriage problem.” Hence, before we explain our experiments and their results, let us pause and quickly summarize the relevant aspects of this literature.

B. The Marriage Problem and Matching

Consider a set of men M and a set of women W . The men have complete binary preferences over both the women and the possibility of being unmatched, as do the women over the men. The “marriage problem” is to find a way to arrange monogamous marriages between the men and women so that the final outcome is stable. An outcome for the problem is a matching in which each man or woman is either matched with at most one member of the opposite sex or left single. In this context, an outcome of

the marriage problem is *individually rational* if it gives each person at least as much utility as he or she would have if left unmatched. An outcome is *stable* if it is individually rational and no man and woman can increase their utility by rejecting the person they were matched with and forming a match with each other.

Notice the properties of the marriage problem: Preferences are ordinal, no transferable utility exists, and matching is one-to-one. In this context, Gale and Shapley (1962) have shown that a nonempty set of stable matches always exists (i.e., the core of this market is nonempty). They present an algorithm for finding the set of stable outcomes. Further, Gale and Shapley (1962) have shown that among the set of stable outcomes, there is one (the *M*-optimal outcome) that is unanimously best for all men and one (the *W*-optimal outcome) that is unanimously best for all women. More interesting is the fact, as Knuth (1976) has shown, that the outcome that is *M*-optimal is the worst outcome for all women, while the opposite is true for the *W*-optimal outcome.

In light of the recent work on incentive compatibility, it is not surprising that there does not exist a matching mechanism that gives both sides of the market the incentive to truthfully reveal their actual preferences (i.e., truth telling is not a *dominant* strategy for all agents in the noncooperative game defined by any matching mechanism). However, Roth (1982) establishes that any mechanism that yields the *M*-optimal (*W*-optimal) outcome defines truth telling as a dominant strategy for men (women). Hence, we can get one side of the market to reveal truthfully.

Leonard (1983) and Demange and Gale (1985) have extended these results to situations in which preferences can be represented by continuous utility functions for which a medium of exchange exists with which to make side payments. Clearly such a generalization is needed if matching algorithms are to be applied to markets where people contract for dollars. The mechanism they use is quite simple: Men (women) (teams and players) submit bids indicating the maximum (minimum) they would be willing to pay (must be paid) to be matched with each other. This information is then taken and used to solve for that vector of competitive or Walrasian prices that is element-by-element the minimum. Such a minimum set of prices determines the *M*-optimal (*W*-worst) outcome for the market. (Shapley and Shubik [1972] established that such a set of prices exists and that the core of this market is nonempty.) If such a mechanism is used, then, as in the conventional marriage problem, it is a dominant strategy for the men to submit truthful values for the women.

The SM mechanism we will use is a direct application of the Leonard (1983) and Demange and Gale (1985) mechanisms. Preferences of teams for players are specified by a matrix

$$\mathbf{V} = \begin{array}{c} \text{t} \\ \text{e} \\ \text{a} \\ \text{m} \end{array} \begin{array}{c} 1 \\ 2 \\ \vdots \\ n \end{array} \begin{array}{c} \text{players} \\ 1 \quad 2 \quad \dots \quad k \end{array} \left[\begin{array}{cccc} & & & \\ & & & \\ & & v_{ij} & \\ & & & \end{array} \right],$$

where v_{ij} is the maximum willingness-to-pay (the value) of team i for player j . The preferences of players for teams are specified by a matrix

$$\mathbf{R} = \begin{array}{c} \text{P} \\ \text{l} \\ \text{a} \\ \text{y} \\ \text{c} \\ \text{r} \\ \text{s} \end{array} \begin{array}{c} 1 \\ 2 \\ \vdots \\ k \end{array} \begin{array}{c} \text{team} \\ 1 \quad 2 \quad \dots \quad n \end{array} \left[\begin{array}{cccc} & & & \\ & & & \\ & & r_{ji} & \\ & & & \end{array} \right],$$

where r_{ji} is the reservation wage or minimum that player j must be paid in order to agree to play for team i . Given these values and reservation wages, teams and players enter bids b_{ij} and asks c_{ji} that are then used to match teams with players and to establish price ranges for these matches. This is accomplished by solving the following pair of dual linear programming problems (which generalize Leonard [1983]) for the minimum set of prices consistent with a competitive equilibrium for this market taking bids and asks as given

PRIMAL PROBLEM 1

$$\max \sum x_{ij} (b_{ij} - c_{ji})$$

subject to

$$\sum_j x_{ij} \leq 1$$

and

$$\sum_i x_{ij} \leq 1,$$

where

x_{ij} = the intensity with which we match team i and player j ,
 b_{ij} = the bid entered by team i for player j , and
 c_{ji} = the bid entered by player j for team i

DUAL PROBLEM 2

$$\min \sum_{ij} (p_j - c_{ji})$$

subject to

$$M_i + (p_j - c_{ji}) \geq (b_{ij} - c_{ji}),$$

$$p_j \geq c_{ji},$$

and

$$\sum_i M_i + \sum_{ij} (p_j - c_{ji}) = \pi,$$

where

M_i = a fictitious dual variable representing the consumer surplus generated by a match of team i and player j at price p_j ,
 p_j = the price (salary) for player j , and
 π = the optimal value of the primal problem

Note that at the equilibrium, p_j is the price attached to player j , while M_i can be considered the maximum fee that team i would be willing to pay in order to participate in the market. Charging team i a fee less than this amount would have no incentive effects. Hence, changes in M_i have no incentive effects but merely shift the surplus between the team and the player. When $p_j = b_{ij}$, $M_i = 0$, since all surplus of the match is awarded to the seller. As p_j is moved below b_{ij} , the surplus of the team increases.

If programs 1 and 2 are used to process the bids entered by the players and teams, and if $M_i = b_{ij} - p_j$ so that the price of the match is the lowest competitive price supporting this outcome (i.e., determines the team-optimal core imputation), then it is still a dominant strategy for the teams to report truthfully. Such is not the case for the players, however, since obviously their bids will influence the price of the matches made. Hence, this set of programs determines a set of matches of teams and players and provides an incentive for truthful revelation at least on the team side of the market.

On a practical level, there are three immediate objections to this mechanism. First, players may object that it is not fair to them since it determines that price which is best for teams given any set of messages or bids. Another equally plausible mechanism would be one that yielded the highest set of prices or in which the imputations of the *players* were as high as possible (i.e., $M_i = 0$)—choosing to give the surplus to the teams is arbitrary. Second, the baseball industry has a history of bargaining for salaries, the participants may not be willing to accept salaries and team assignments that are prescribed by some mathematical maximization problem and its dual. Finally, because there is no role for bargaining here, the mechanism has substantially cut down the role of the sports agent. Eliminating their profits may ultimately lead to the rejection of this mechanism as politically unfeasible.

To answer these objections we, instead, tested a modified version of this mechanism in our experiments. In our version, subjects playing the role of teams (which we called U-type subjects) and subjects playing the role of players (which we called S-type subjects) submit their bids just as we described above. Using this information, programs 1 and 2 determine the optimal matches and tell each pair of subjects who are successfully matched the range in which their salary must be negotiated. This range is defined by program 2, and for each pair (i, j) it falls in the interval $[p_j, p_j + M_i]$. Note, however, that this range is simply the range defined by letting M_i vary from its minimum value of zero to its maximum of $b_{ij} - p_j$.

This mechanism preserves bargaining as well as a role for the sports agent. The drawback, of course, is that this mechanism (as well as the mechanism that uses the unrestricted program 2) does not make it a dominant strategy for any subject to tell the truth, so we have lost even the partial incentive compatibility properties discussed before. Still, it is neither a necessary nor sufficient condition that a mechanism be incentive compatible. Many mechanisms that lack this property perform quite well in efficiency terms. In fact, even inefficient mechanisms may be popular with the people who use them for a variety of political and sociological reasons.

C Some Preliminary Results

Our experiments were aimed at investigating three simple questions:

1. Which type of mechanism performs the best—that is, which is better able to capture a greater fraction of the potentially available gains from trade, and which is able to produce the most number of optimal matches?
2. Under which mechanism are prices the highest?
3. Which mechanism generates the highest profits for the teams, and which is most beneficial for the players?

On the basis of the experiments performed we have the following conclusions to offer:

1. Except for its tendency to yield no matches when extreme bids are entered, the SM mechanism employed in our experiments demonstrates

good performance characteristics, ones that are on par with the CFA and CIFA mechanisms. For example, while 14 out of 180 potential matching situations (7.7%) led to no matches, for the remaining 166, the mechanism was able to capture 97% of the available gains from trade. It did this by achieving optimal matches for 146 of the remaining matches. While average efficiencies were better under the CFA mechanism, where 94.8% of the potentially available gains from trade were captured as opposed to 89.4% for SM, the CFA mechanism generated a far greater number of mismatches (31 out of 150) than did SM (which had only 20 out of 180). Further, it appears that the frequency of no matches under SM can be accounted for by the "extreme" bids entered by these subjects, which misrepresent their true values and costs by amounts ranging from 56% to over 400%. The CIFA mechanism performed in a manner equivalent to the SM mechanism. It had the greatest fraction of no matches (14 out of 150 potential matches, or 9.3%). In addition, when it succeeded in matching subjects, it failed to make the optimal match in 14 out of 136 instances. Overall (including the no match data), it was able to capture 88.3% of the available gains from trade and 97.4% of the gains available when it was successful in matching subjects.

2. Prices tended to be highest under the CFA mechanism, with the SM mechanism being second and the CIFA mechanism yielding the lowest prices of all. In terms of the actual prices formed, the CFA mechanism yielded an average price of \$2.65, while the SM mechanism determined an average price of \$2.35 and the CIFA an average price of \$2.20. These differences proved to be statistically significant.

3. Since prices were lower in CIFA than in the SM and CFA experiments (in that order), one would expect that U-type (buyer) payoffs would be ranked in the same order (CIFA, SM, and then CFA), while the S-type (seller) payoffs' ranking would be opposite. This, in fact, was the case. Under CFA, average realized payoffs equaled \$1.87 and \$1.94 per round for U- and S-types, respectively, as compared to \$2.00 and \$1.72 for SM and \$2.07 and \$1.45 for CIFA.

In short, by looking at gross summary statistics, it would appear that the efficiency properties of all mechanisms were quite good with the CFA mechanism doing the best (in a statistically insignificant manner). In addition, while CIFA yielded the highest payoffs for U-type subjects, CFA was distinctly more advantageous for S-types.

In the remainder of this article, we fully describe the experiments that were performed (Sec. III) to investigate the properties of these three mechanisms and then present a full description of the results (Sec. IV). In Section IV we also discuss the results of some statistical tests performed on the data, while in Section V we present an analysis of what we think the implications of this study are for the design and implementation of a baseball player allocation system.

III. The Experiments and Experimental Design

Three sets of experiments were conducted, each aimed at replicating the salient features of a different allocation mechanism (Instructions are available from the authors on request)

The objective of the subjects in all three experiments was to try to match themselves with another subject in the experiment and determine a price for that match. While the manner in which this was done changed from experiment to experiment, the preferences induced on the subjects were identical. This allowed us to impute any differences in behavior and performance to the institutional rule or mechanism used in the experiment. In all of the main experiments reported here,¹ subjects were randomly assigned to be either one of two types called in the instructions U-types or S-types. The instructions also informed them that they could be matched with *at most* one subject of the opposite type and that their payoffs would depend on with whom they were matched and the price determined for the match. To induce preferences on the subjects, U-types were given a schedule informing them of the amount of money they would be paid if they were matched with any S-type subject, denoted as S_1 , S_2 , and S_3 . These three values were similar in that it was always true that each U-type valued one S-type at \$5, one at \$4.5, and one at \$4. However, no U-type subject knew the preferences of anyone but himself.

To induce preferences on the subjects, S-types were given a schedule informing them of the amount of money they would have to pay at the end of the experiment if they were matched with any U-type subject denoted as U_1 , U_2 , and U_3 . These three values were similar in that it was always true that each S-type always valued one U-type at \$5, one at \$1, and one at \$2.² However, no S-type subject knew the preferences of anyone but himself. In each round of the experiment, we would change these schedules, but these changes merely constituted a permutation of the indices attached to the following pair of matrices.

	Matrix 1 U-Type Preferences				Matrix 2 S-Type Preferences		
	U_1	U_2	U_3		S_1	S_2	S_3
S_1	4.5	4	5	U_1	5	2	1
S_2	5	4.5	4	U_2	1	5	2
S_3	4	5	4.5	U_3	2	1	5

¹ Some subsidiary experiments were performed as pilot experiments, and while we will not refer to them in the main body of this article, some reference to them will be made in footnotes.

² In the SM experiment, all values and costs for U-types and S-types were multiplied by a factor of 10. We will discuss the reason for this later.

These matrices define all of the information known to the experimenter in each round of the experiment. Looking down each column, we see the value (matrix 1) or cost (matrix 2) of each U-type (S-type) for subjects of the opposite type. Each subject knew only the column in the matrix relevant to himself but knew that U-types had values of either \$5, \$4.5, and \$4, while S-types had values of either \$5, \$1, or \$2. Note that with these parameters profitable matches could be formed between any S-type subject and any U-type subject and that the difference between the surplus generated by optimal matches and suboptimal matches was not great. This, we expected, would lead to a fair amount of competition between the subjects.

As we see, the optimal (surplus-maximizing) set of trades occurs when S-type subjects with a cost of \$5 were matched with U-type subjects with a cost of \$4.5. All of these matches generated a surplus (sum of the consumers plus producers surplus) of \$4, while any other match generated a surplus of only \$3. Hence, in every round of the experiment the set of optimal matches remained unique, although because we permuted the indices it was not always true that U_1 was matched with S_1 , U_2 with S_2 , and U_3 with S_3 .

Considerable thought went into the selection of these parameters. First, we did not want numbers in these matrices to be such that optimal matches were too obvious. For example, we did not want the optimal matches to be ones in which each person received his first choice. We wanted to disguise the equilibrium. Second, we wanted there to be competition for S-types. We created such competition by having the surplus generated by nonoptimal matches be almost as great as the surplus generated by optimal ones. For instance, in these matrices, nonoptimal matches generate 75% of the surplus generated by optimal matches. If nonoptimal matches were extremely unprofitable, the optimal matches might be too salient and our results suspect since it might be claimed that we achieved a high efficiency level simply because it was too obvious how subjects should bid. Olson and Porter (1994, in press) construct what they call a "contention index" to reflect the competitiveness of such matching matrices and employ two different treatments, high contention and low contention, in an effort to compare how their allocating mechanisms work in those two environments. In a set of pilot experiments we ran, we also examined the use of low-contention matrices, but our results were qualitatively identical to the ones we will present below using matrices 1 and 2. Hence, we do not present these results in this article but refer to them in footnote 3 below.³

³ In our pilot experiment, we ran an experiment with parameters generating the following preference matrices:

Holding the preferences depicted in matrices 1 and 2 constant across experiments allows us to impute the differences between experiments to the different sets of rules existing in each one and not to value or cost changes.

A. The CIA Experiment

The CIA experiment was quite simple. Students were placed in offices of economics professors in the Department of Economics at New York University. On the desk where they sat was a telephone, a list of telephone numbers, and a set of 10 envelopes, one for each round of the experiment. If a subject was a U-type subject, the telephone numbers given to him or her were those of the S-types. The opposite was true for S-type subjects. Each round began with subjects opening one envelope. In this envelope was a piece of paper indicating the subject's preference schedules for that period. After these envelopes were opened and the information recorded on worksheets, the subjects had 5 minutes within which to call subjects of the opposite type and try to negotiate a match and a match price. If such a contract was formed, its existence was announced publicly and those subjects were out of the market for the remainder of that round. If

Matrix 3 U-Type Preferences				Matrix 4 S-Type Preferences			
	U ₁	U ₂	U ₃		S ₁	S ₂	S ₃
S ₁	50	10	20	U ₁	10	20	30
S ₂	20	50	10	U ₂	30	10	20
S ₃	10	20	50	U ₃	20	30	10

As you can see from these matrices, the optimal matches are when S₁ is matched with U₁, S₂ with U₂, and so on. (In the experiment we permuted the rows and columns each period so no one could figure out that 1 was matched always with 1, 2 with 2, etc.) Optimal matches generate surpluses of 40, while nonoptimal matches generate surpluses of either 0 or -10. Consequently, in this matrix, non-optimal matches are much less good than optimal matches (actually, they are awful). The fear with matrices such as these is that they force optimal matches because the consequences of nonoptimal matches are so dire. We ran these parameters on eight groups of six subjects each in the SM institution and on four groups of six subjects each in the CFA institutional environment. In looking at the data from these experiments, we come to the same qualitative conclusions as we show later in this article (see Sec. IV below). While the SM mechanism always matched subjects optimally, there was a tendency for no matches to occur. Hence, when the no matches were excluded, the efficiency was 100%, while it was only 83% when no matches were included. The CFA mechanism generated an overall efficiency of 93%, strikingly close to the 94.8% achieved here. The only qualitative difference was the occurrence of a number of no matches, which, in this case, was an artifact of the negative surpluses available for some nonoptimal matches. Again, in the SM mechanism, when no matches occurred, there was a great deal of misstatement of preferences and strategizing. In summary, despite our fear that such matrices would result in almost perfect efficiencies, the qualitative results from these experiments are not greatly different from the ones reported on in our article.

a U-type subject was successful in making a match within the 5-minute time limit, his or her payoff was equal to the difference between the value of the S-type subject they were matched with and the price of that match. For S-type subjects who were successfully matched, the payoff was equal to the difference between the price of the match and the cost of the U-type subject they were matched with, as was indicated on their schedule. If a subject failed to be matched, his or her payoff was zero for that round. A subject's final payoff equaled the sum of his or her payoffs over the entire 10 rounds of the experiment.

B. The CIEA Experiment

The CIEA experiment was conducted as follows. Subjects were seated in a classroom with S-types in the first row and U-types in the rear. At their seats were a stack of 10 envelopes as well as a small chalkboard on which they would write messages. At the start of each round, the subjects would again open their envelopes and inspect their preferences for that round. They would then be given 5 minutes to complete their contracts. This was done as follows: in the front of the room was a blackboard with the following table on it:

S ₁	S ₂	S ₃	<u>Contracts</u>
U ₁ U ₂ U ₃	U ₁ U ₂ U ₃	U ₁ U ₂ U ₃	

When the experimental administrator said "begin," the U-type subjects could enter a bid for any player of the S-type they wanted. This would be done by writing the bid on their chalkboard and raising it above their head. The experimental administrator stationed in the front of the room would then write the bid under the S-type subject's column. For example, if subject U₂ wanted to bid \$1 for S₁, he or she would only have to write S₁-1 on their chalkboard. This bid would then be placed in the U₂ column under the heading for subject S₁. As bids were made they were recorded in the appropriate places on the board. The last bid made by a U-type subject for an S-type subject was the only one currently available and remained active either until accepted or until the U-type had one of his other bids accepted. S-type subjects could not make counteroffers but could accept bids by writing the word "accept" and the identity of the subject whose bid was being accepted on their chalkboard. When they did so, a contract was made, and the experimental administrator notified everyone by writing who formed it and its price on the blackboard. The experiment was conducted in total silence and hence avoided the hysteria of oral auctions. Payoffs were calculated in an identical manner as discussed in the CFA experiment. Note, however, that in this experiment all bids made for all S-types were common knowledge.

Table 1
Experimental Design

Experiment	Number of Groups	Matrix Preferences	Number of Rounds	Number of Subjects
1 CIA	5	1 and 2	10	30
2 CIIA	5	1 and 2	10	30
3 SM	6	1 and 2	10	36

NOTE - CIA = current free agency system, CIIA = complete information English auction, SM = simultaneous mechanism

C. The SM Experiment

In the SM experiment, subjects were seated at computer terminals. At the beginning of each round, their preference schedules were flashed on the screen. They were then prompted by the computer to enter a vector of bids, one for each subject of the opposite type. This information from all subjects was entered into the main file server of the network where programs 1 and 2 were solved.

Once the optimal matches are determined, subjects were matched and told that they had 5 minutes to determine a price for their match. The price could be anything in the closed interval $[p_i, p_i + M_i]$. Because price setting in this mechanism requires some bargaining, we did not want to disrupt the experiment after each round and allow subjects to bargain. Hence, we multiplied the payoffs in each round by 10 and told the subjects that one round would be randomly chosen at the end of the experiment as the round that would count. The matches and prices determined in this round would, by themselves, define payoffs for each of the S- and U-type subjects.⁴

D. Experimental Design

Our experimental design is described in table 1.

We conducted three experiments. In two, CIA and CIIA we had five independent groups each containing three S-type and three U-type subjects, each group performed the experiment for 10 rounds. For the SM experiment, we had six independent groups. Hence, all together we had 96 subjects involved in these experiments.⁵ As can be seen, since all experiments are identical except for the allocating rule, they furnish us with a ceteris paribus test for the influence of the mechanisms themselves, holding preferences constant.

⁴ This is why we multiplied all payoffs here by 10 in order to preserve an equivalent expected payoff between these subjects and those of the other experiments.

⁵ This does not include a set of pilot experiments we performed as well.

E Data Set

Given our experimental design, our experiment can be expected to yield the following data. In each round of each experiment there is a potential for at most 3 matches. Hence, in experiments like CFA and CIEA where we have 5 groups of subjects, there are at most 15 potential matches per round and 150 potential matches over the course of the 10-round experiment. For SM, since there are six groups and 10 rounds, we generate a total of 18 potential matches per round and 180 over the 10 rounds of the experiment. Attached to each match made is a price and payoffs for each U-type and S-type subject as well as an efficiency for that match, which we measure by the fraction of the gains from trade available from an optimal match that was captured by the match actually made. For example, from matrices 1 and 2 above we know that when an optimal match is made it generates a surplus of \$4⁶ as measured by the excess of the U-type subject's value over the S-type subject's cost. Given our data, any suboptimal match produces an excess of \$3, while a no match produces an excess of 0. Hence, the efficiency of an optimal match is $(\$4/\$4) \times 100\%$, while that of a suboptimal match is $(\$3/\$4) \times 100\%$ and a no match is $(0/\$4) \times 100\%$. In the CFA and CIEA experiments we had a price, payoff, and efficiency actually formed whenever a match was made. This led to 150 and 136 prices, payoffs, and (nonzero) efficiencies made over the course of each of the CFA and CIEA experiments respectively (there were 14 no matches in the CIEA experiment). In the SM experiment, we chose one and only one round in which to actually have a price negotiated in each experiment. Hence, we had only 17 prices actually formed (there could have been 18 but in one of the rounds selected there was one no match in one experiment). Despite this fact, we still had 166 (nonzero) efficiencies since they depend only on the matches made by the mechanism. For the SM mechanism, in addition to investigating the set of 17 prices that were actually negotiated, we constructed another hypothetical price for each match. We assume that when a match is made, the subjects split the fee M_i over which they bargain. This would lead to a price at the midpoint of the interval $[p_i, p_i + M_i]$. Let us call this the "split-the-difference" price.

IV Results

In terms of broad descriptive statistics, table 2 describes the results of our experiments. Let us interpret this table by investigating the efficiencies, prices, and payoffs generated by our experiments.

A Efficiencies

Table 2 provides two measures of efficiency. One has been described above and is called the "surplus measure" since it measures the fraction

⁶ This surplus is 40 in the SM experiment, where all values and costs are multiplied by 10.

Table 2
CFA-SM-CIEA Experiments Summary of Results

Statistic	CFA	SM ¹	SM ²	CIFA ¹	CIEA
A. Efficiency					
Surplus measure (%)	94.8	89.4	97.0	88.3	97.4
Numbers measure (%)	79.3	81.1	87.9	81.3	89.7
B. Prices (per round)					
Average negotiated price (\$)	2.65	2.22	2.35	2.00	2.21
Average split-the-difference price (\$)		1.92	2.09		
Mode negotiated price (\$)	2.50	2.20	2.20	2.50	2.50
Mode split-the-difference price (\$)		2.00	2.00		
C. Payoffs (per round)					
1. S-types (sellers)					
Negotiated price (\$)	1.94	1.72	1.82	1.45	1.61
Split-the-difference price (\$)		1.36	1.48		
2. U-types (buyers)					
Negotiated price (\$)	1.87	2.00	2.12	2.07	2.29
Split-the-difference price (\$)		2.21	2.39		

NOTE: See table 1 note for an explanation of abbreviations. SM¹ = SM including "no match" outcomes. SM² = SM excluding "no match" outcomes. CIFA¹ = CIFA including "no match" outcomes. CIEA = CIFA excluding "no match" outcomes.

of the available surplus or rents captured by our subjects. The other measure is called the "numbers measure" and measures the fraction of the total number of potential matches that were optimal. Let us look at our experiments one at a time.

1. *The SM experiment*—In the SM experiment, out of 180 potential matches, there were 14 (7.7%) "no matches," 20 (11.1%) suboptimal matches, and hence 146 (81.1%) optimal matches. By definition, this yielded a numbers efficiency of 81.1%. In terms of surplus efficiency, SM was successful in capturing 89.4% of the potentially available gains from trade or surplus, including the zero efficiencies generated by the 14 no matches. If we exclude these no matches and look only at the fraction of the surplus captured when matches were actually made, we see that SM was successful in capturing 97% of the surplus.

2. *The CFA experiment*—In the CFA experiment, out of 150 potential matches, there were no no matches (0%) and 31 suboptimal matches (20.6%). Of the potentially available matches, 119 were optimal, yielding a numbers efficiency of 79.3%. In terms of surplus efficiency, CFA was successful in capturing 94.8% of the potentially available surplus.

3. *The CIEA experiment*—In the CIEA experiment, out of 150 potential matches, there were 14 no matches (9.3%) and 14 suboptimal matches (9.3%). Of the potentially available matches, 122 were optimal, yielding a numbers efficiency of 81.3%. In terms of surplus efficiency, CIEA was successful in capturing 88.3% of the potentially available gains from trade or surplus, including the zero efficiencies generated by the 14 no matches. If we exclude these no matches and look only at the fraction of the surplus

captured when matches were actually made, we see the CILFA was successful in capturing 97.4% of the surplus.

To investigate these observations, we performed two sets of tests. First, we ran a round-by-round Mann-Whitney U -test on each pair of experiments to see whether there were significant differences in the mean efficiencies between these mechanisms taken pairwise. Since within an experiment observations are not independent, we created a sample of group means round by round by averaging the efficiencies generated by matches within any group and using these group means as a sample. We also performed this test by pooling all of these observations over the 10 rounds of the experiment. What we found was that in only one round was there a significant difference (at the 5% level) between the surplus efficiencies of any of the three mechanisms when compared bilaterally (i.e., CILFA vs CFA, CILFA vs SM, etc.). For the numbers efficiency measure, there were no significant differences in any round at the 5% level.

4. *No-match behavior* — Since the efficiency of the SM mechanism was dramatically affected by the frequency of no-match outcomes, it is of interest to discover how much of a deviation from one's true cost or value is needed to produce this situation. Of course, if all subjects merely bid their true values and costs in the experiment, then we would always observe 100% efficiencies. The extent to which we observe suboptimal outcomes is therefore evidence of the extent of misrepresentation on the part of our subjects. To study this misrepresentation behavior in the SM mechanism, we present table 3 and figures 1 and 2. In table 3, we see the mean deviation of bids by U-type and S-type subjects from their truthful bids and costs averaged over all 10 rounds of the experiment and conditional on their realized outcomes.

As we can see, there is a discrete difference between the type of misrepresentation which exists in instances that lead to no matches and that which occurs in those instances that lead to trades. For example, while the mean deviation for U-types with a value of \$50 for an S-type was \$15.4 for those bids leading to optimal matches, it was \$28.7 for those bids leading to no matches. For S-types, a similar situation existed. When S-types bid for U-types for whom they had a cost of \$20, they tended to raise their bid an average of \$6.52 above their cost (they bid on average \$26.52) for those bids leading to optimal matches. When their bid led to a no match, it was typically \$18.3 above their cost (\$38.3), representing an almost 300% increase. There seemed to be no difference between those bids leading to optimal and those leading to suboptimal matches. This leads us to think that, despite the substantial number of no matches, the SM mechanism is fairly robust to strategic manipulation in the sense that it took a very large misrepresentation to lead to a no-match outcome. Furthermore, since without no matches the SM mechanism performed extremely well, one might conclude that over time, when such misrepre-

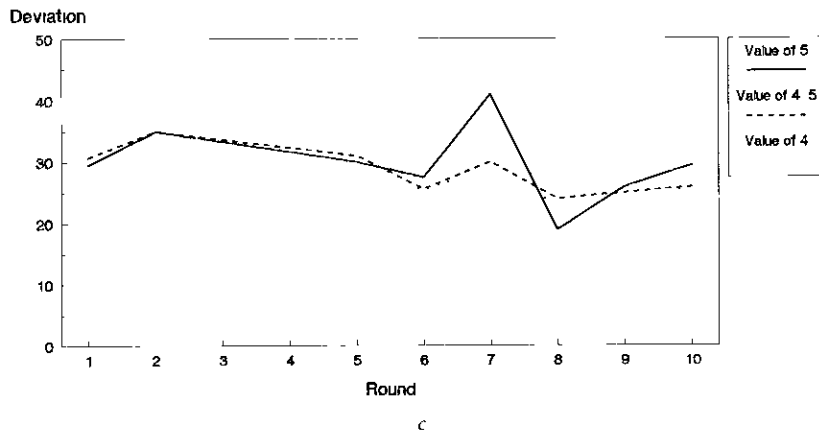
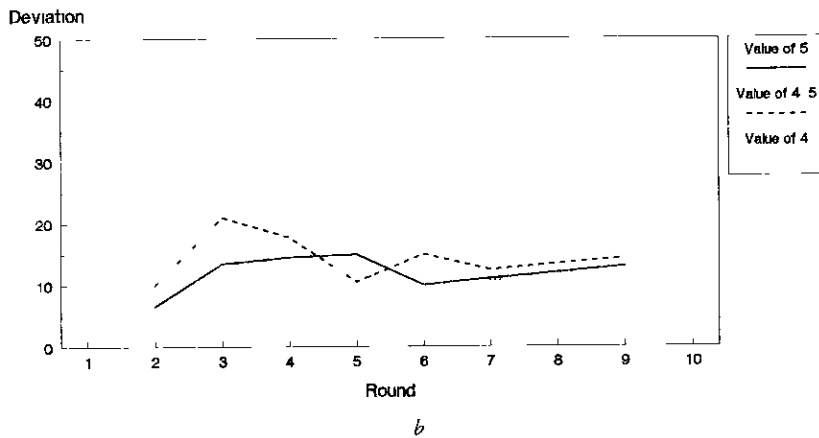
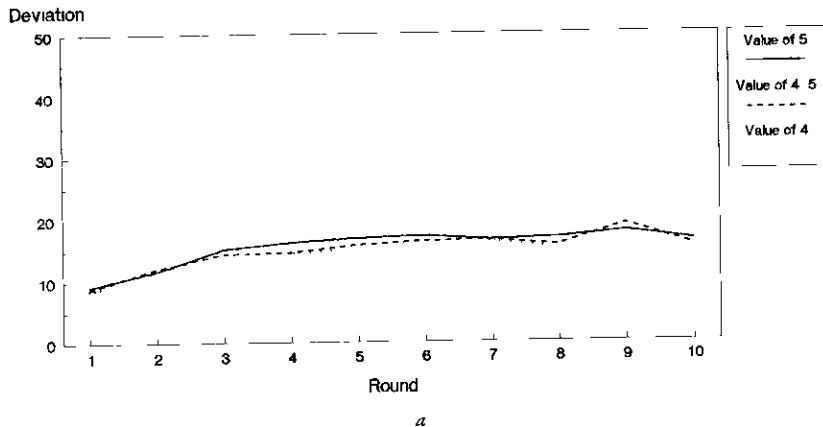
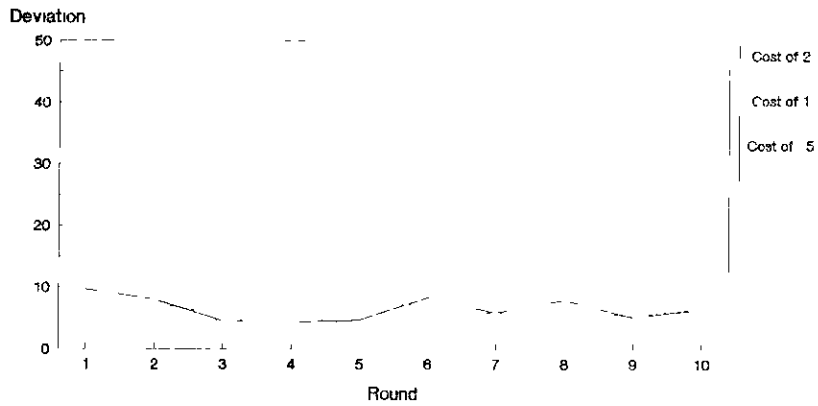
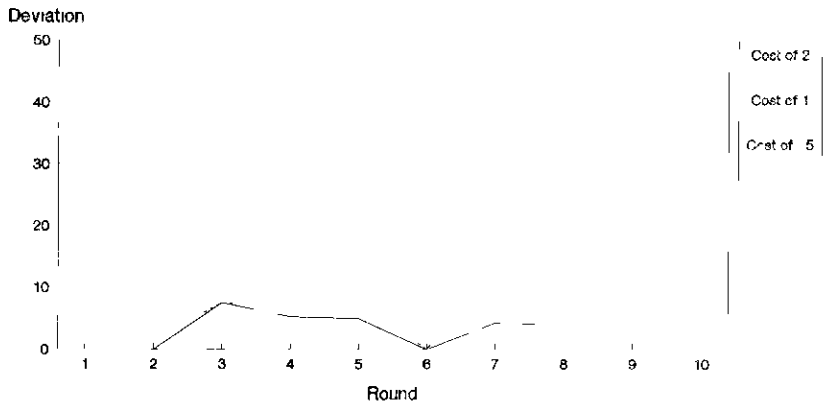


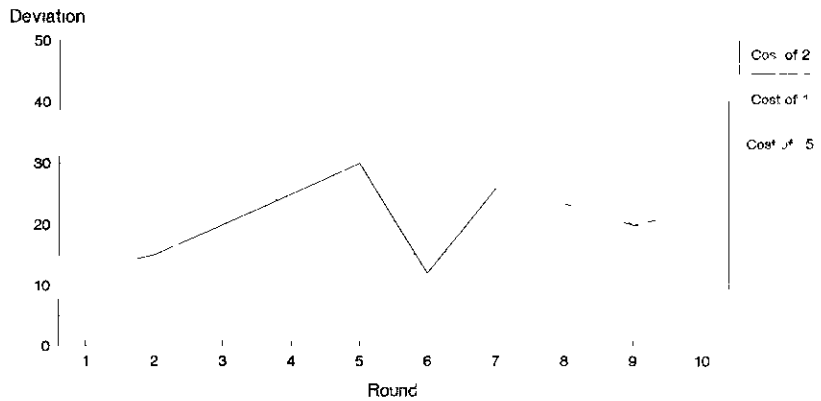
FIG. 1 —SM U-type misrepresentation behavior (average deviations) by round *a*, optimal matches, *b*, inefficient matches, *c*, no matches. Note —Deviation is defined as the value minus the actual bid average across all U-types in each round



a



b



c

FIG. 2 — SM S-type misrepresentation behavior (average deviations) by round: *a*, optimal matches; *b*, inefficient matches; *c*, no matches. Note: Deviation is defined as the bid value minus the actual bid average across all S types in each round.

sentations are discovered to be counterproductive, the efficiency of the SM mechanism would increase. We did not run our mechanism long enough to uncover this tendency, although we do note that 28% of the no matches did occur in round 1 of the experiment.

Figures 1 and 2 plot the mean deviation of bids from value (cost) round by round for U- and S-types. As we see, the U-types exhibited considerably larger deviations from their true values than did the S-types from their costs. More significant, perhaps, is the fact that as the experiment progressed the deviations of the U-types seemed to increase. This, we feel, is because the U-type players were capable of discovering the relationship between their bid and the fee M , over which they would eventually be negotiating. (In fact, the fee M was the difference between the price p for their match and their bid.) Hence, by lowering their bid, as long as it did not prevent their getting matched, they could lower the range over which the final price would be negotiated.

The pattern of bidding behavior for S-type subjects over 10 rounds appears relatively stable, with no significant trend in either direction. The average bid for high-cost matches (\$20) declined by 6% from round 1 to round 10, never exceeding the first-round value, the decline for middle-cost matches (\$10) was 8%, also remaining persistently below the first-round value. This suggests the presence of a modest learning effect, though the movement of values over all 10 rounds is more consistent with stable valuations. It is interesting that the average bid for lowest-cost matches (\$5, the optimal match) rose by 6% from round 1 to round 10, dropping below its initial level in only three of 10 rounds.

The excesses of average S-type bids across rounds over match costs are

Cost (\$C)	Average Bid (\$AB)	$AB - C$	% Excess
20	27	7.00	35
10	17.60	7.60	76
5	11.90	6.90	138

These bids are pooled over all subjects and all rounds regardless of whether they lead to matches or not.

In terms of absolute dollar values, S-types were quite consistent in demanding an approximately \$7.20 premium over actual costs regardless of match type. In percentage terms, the premium for the low-cost (optimal) match was substantially higher. In contrast, as we shall see below, U-types behaved consistently over all possible matches, deviating in their bids from actual values by roughly 35%.

The bidding behavior of U-types showed a pronounced downward trend over the course of 10 rounds (upward trend in deviations). The average bid for highest-value matches (\$50) dropped by 11% from round 1 to round 10, exceeding its initial level only once (in round 2) and declining

almost steadily from round to round, the decrease for middle value matches (\$45, the optimal match) was 9%, also remaining persistently below its initial level after the second round. Bids for lowest-value matches (\$40) behaved similarly, showing a 16% decline. Here, too, the initial bid level was exceeded only once, in the second round. It appears that after an initial testing period (rounds 1 and 2), U-type subjects adopted an aggressive bidding strategy of persistent bid reduction, which they implemented successfully as they gained experience with the matching mechanism.

The excesses of match values over average bids across rounds are

Value (\$V)	Average Bid (\$AB)	$\$V - \AB	% Excess
50	33.80	16.20	32
45	29	16	35
40	25.60	14.40	36

These bids are pooled over all subjects and all rounds regardless of whether they lead to matches or not.

There is no discernible difference in the way that U-types bid with respect to match types. Certainly, there is nothing to indicate in observed bidding behavior that U-types learned to discriminate between optimal and nonoptimal matches. Overall, it is clear that dollar-value deviations of U-type bids from valuations strongly exceeded deviations of S-type asks from costs. This differential pattern drove the bargaining process and had a visible impact on match prices and ultimate payoffs.

B. Prices

As table 2 indicates, prices tended to be highest when the CFA mechanism was used and considerably lower when the SM and CEA mechanisms were used. The mean price under the CFA mechanism was \$2.65, while it was \$2.35 for prices actually negotiated in the SM experiment. When we look at what the prices would have been in the SM experiment if we made the assumption that our subjects would "split the difference" and negotiate a price at the midpoint of their bargaining interval, we see that the mean price would have been even lower at \$2.09. The CEA experiment yielded an average price of \$2.21.

The same Mann-Whitney *U*-tests were performed on the price data as were performed on the efficiencies data. In this case, more definite conclusions can be drawn. For example, as table 4 indicates, the CFA mechanism generated prices that were significantly below those of the CEA and SM mechanisms. When we compare the mean prices formed in any round by the five groups of subjects each in the CFA and CEA experiments, we see that in seven of the 10 rounds there was a statistically significant difference (at the 10% level or below) between the mean price formed in the

Table 4
Price Data Mann-Whitney *U*-Tests for Differences in Means across Experiments

Round	z-Scores		
	CIA vs CIFA	CFA vs SM	CIFA vs SM
1	1.25 (.105)**	2.28 (.011)*	1.18 (.117)
2	.625 (.265)	.639 (.261)	.091 (.463)
3	.208 (.417)	1.36 (.085)**	.273 (.392)
4	1.25 (.105)**	1.00 (.157)	.091 (.463)
5	2.29 (.01)*	2.64 (.004)*	1.18 (.117)
6	1.46 (.07)**	1.91 (.027)*	.456 (.324)
7	.83 (.20)	1.55 (.060)**	.639 (.261)
8	1.34 (.089)**	1.73 (.041)*	.106 (.457)
9	1.46 (.071)**	2.09 (.017)*	.821 (.205)
10	1.67 (.047)*	1.73 (.041)*	.453 (.324)
Pooled data	4.00 (.000)*	5.80 (.000)*	1.48 (.069)**

NOTE.—See table 1 for an explanation of abbreviations. One-tailed probabilities are in parentheses.

* Significant at the 5% level or less.

** Significant at the 10% level or less.

CFA and CIFA experiments. Similarly, there was a statistically significant difference in eight of the 10 rounds between SM and CFA mechanisms. In no round was there a significant difference between the prices formed in the CIFA and SM experiments. In short, the CFA mechanism did, in fact, determine prices significantly higher than those formed by either the SM or CIFA mechanisms, while SM and CIFA were statistically indistinguishable.

The overall trend in prices over ten rounds of the CFA experiment appears to be stable. If we omit the first round where an initial price of \$7 was formed by one pair and consider instead the movement of prices between rounds 2 and 10, we see that prices actually rose a modest 2%. In contrast, the average price in round 10 was below the average of the first five rounds by 6%. The conclusion, then, is that CFA evidenced basic price stability over 10 rounds with, at most, a modest downward trend. The actual mean price formed in each round of the CFA experiment is presented in table 5.

Unlike the CFA experiments, there is a pronounced downward trend in prices over 10 rounds in the SM experiment. The average price fell by

11.5% from round 1 to round 10. In round 10, it was 8% below the average price of the first five rounds. This movement downward is consistent with the trend exhibited by U-types in this experiment to increasingly lower their bids as the experiment progressed. The mean of the “split-the-difference” price, round by round, is presented in table 6.

Finally, it appears that the prices formed in the CIEA experiment were stable over the experiment’s horizon. While from round 1 to round 10 the average price fell by 19%, from round 2 to round 10 the fall was only 3%. Table 7 presents the round-by-round mean price formed in the CIEA experiment.

Figures 3–5 show the distribution of prices formed in the three experiments. One thing worth noting is that the CFA mechanism seems to provide prices with much smaller variances than either of the other two mechanisms. In addition, it appears less prone to generate “low” prices or prices below \$1.50. More precisely, in the CFA experiment only one price was formed at the level of \$1.50 or below. In the SM experiment there were 29 (using the “split-the-difference” price), while in the CIEA experiment there were 27. From observing these experiments, however, it appears to us that the low prices in the SM experiment were a function of individual learning on the part of U-types about the impact of their bids on the second-stage bargaining range, while in the CIEA experiment it appeared that implicit collusion took place aided by the information available to the subjects. We say this because there were two groups in the CIEA experiment who seemed to be quite successful in keeping prices low and who did so in a fairly conscious manner.

C. Payoffs

Since prices were highest in the CFA experiment, second highest in the SM experiments, and lowest in the CIEA experiments, we might expect to find that the payoffs of U-type subjects have the opposite ranking. This is in fact true. In the CFA experiment, the mean payoff of U-type subjects was \$1.87 per round while it was \$2.00 in the SM experiment and \$2.07 in the CIEA experiment. For S-types, the ranking was just the opposite, with the CFA experiment generating a mean payoff of \$1.94 and the SM and CIEA experiments yielding payoffs of \$1.72 and \$1.45, respectively. While these differences appear substantial in many cases, their statistical significance was not always strong, as we will soon see. These mean payoffs include within them the fact that in the CIEA and SM experiments subjects were left several times without a match. In those cases, of course, their payoffs were zero. Hence, if we condition payoffs on whether a match was made, we see that payoffs were considerably higher for U-type subjects who were successfully matched in the SM and CIEA experiments than they were in the CFA experiment. For example, the mean payoffs per round for U-types were \$1.87, \$2.12, and \$2.29 in the CFA, SM, and CIEA

Table 5
Mean Price Round by Round in
the CFA Experiment

Round	Average Price (\$)
1	3.16
2	2.51
3	2.49
4	2.70
5	2.88
6	2.48
7	2.48
8	2.65
9	2.59
10	2.57

experiments, respectively, when we excluded the instances of no matches. These differences did not appear as strongly when we looked at S-types, however, where the means were \$1.94, \$1.82, and \$1.61 for the CFA, SM, and CIEA experiments, respectively. In conclusion, it appears that while considerable differences appeared in the payoffs to U-type subjects across our three experiments, those differences were even more significant in the SM and CIEA experiments for those subjects who successfully found matches.

Mann-Whitney *U*-tests were run to investigate whether there were significant differences, round by round, between the payoffs in our three experiments. In terms of statistical significance, when we include the zero payoffs that occur with no matches, there is not a significant round-by-round difference in the payoffs of U-type subjects across the three experiments. For instance, significant differences appear in only one round between the CFA and CIEA experiments and only three and four times in the comparisons between CFA and SM and CIEA and SM, respectively.

Table 6
Mean Split-the-Difference Price Round by
Round in the SM Experiment

Round	Average Split-the-Difference Price (\$)
1	2.24
2	2.22
3	2.11
4	2.26
5	1.99
6	2.02
7	2.10
8	1.92
9	1.85
10	1.99

Table 7
Mean Price Round by Round in
the CIEA Experiment

Round	Average Price (\$)
1	2.56
2	2.16
3	2.22
4	2.18
5	2.33
6	2.16
7	2.15
8	2.09
9	2.12
10	2.09

For S-types, there does appear to be a significant difference between the CFA and SM experiments since in eight of the 10 rounds the differences in means are significant.

The situation changes when we investigate the payoffs of subjects only in those situations when matches are made. In these circumstances, there is a significant difference in the payoff of U-types between the CFA experiment on the one hand and the CIEA and SM experiments on the other, with significant differences appearing in seven rounds in the comparison between CFA and CIEA and in eight rounds in the comparison between CFA and SM. With respect to S-type players, we again see the CFA experiment behaving differently from the SM experiment, with five rounds showing significant differences but no real difference in the other two comparisons we make.

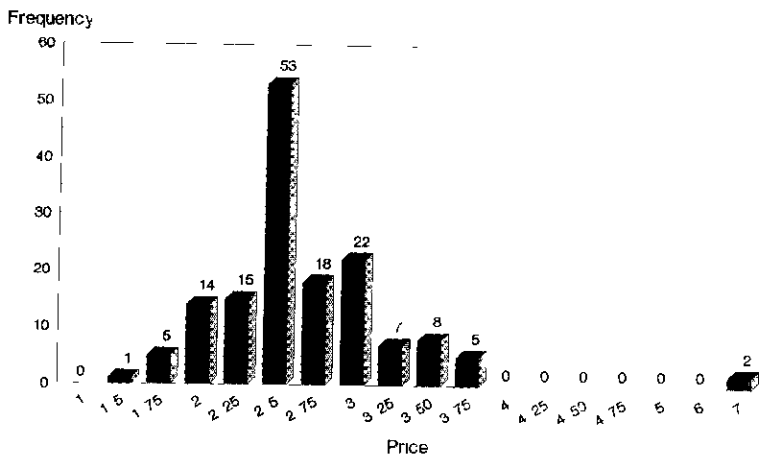
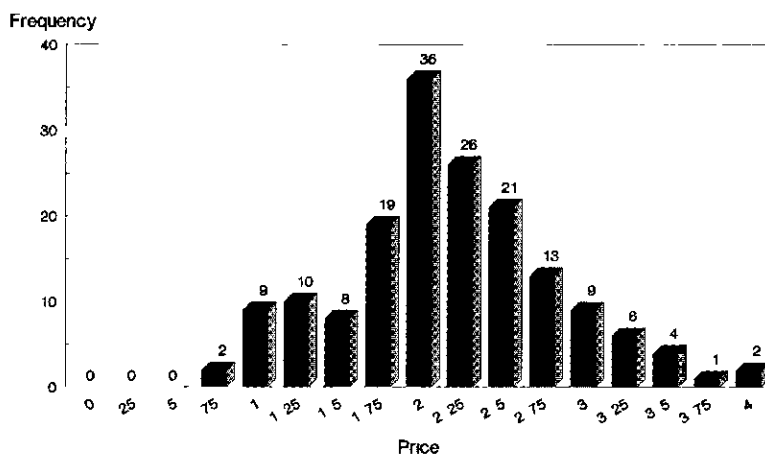
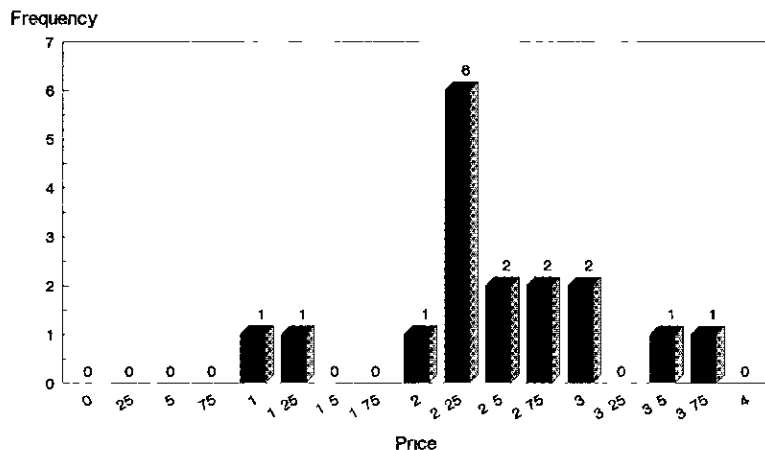


FIG. 3—CFA experiments: frequency of actual prices for all matches



a



b

FIG. 4—SM experiments frequency of prices (with half fee) a, all matches, b, for actual negotiated prices

V. Summary and Discussion

A Summary

Our experiments have uncovered the following results

1 Although visually it appears that the CFA mechanism achieved the greatest efficiencies, these differences are not borne out statistically. However, a mechanism's efficiency is influenced by two factors its ability to match people, and its ability to match people optimally. On these grounds, the three mechanisms differed. While the CFA mechanism was extremely successful in avoiding no-match outcomes (it determined none of them), it was less successful in matching people in an optimal (surplus-maximizing)

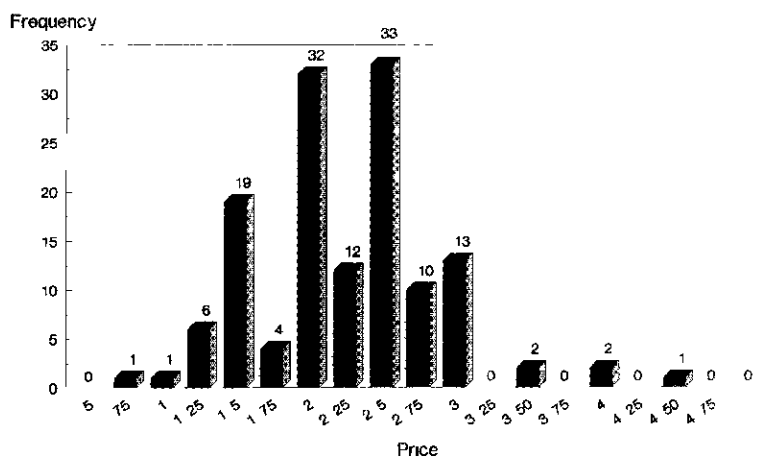


FIG. 5—CIEA experiments: frequency of actual prices for all matches

manner, failing to do so in 31 of 150 opportunities. The opposite was true of the SM mechanism. While it produced a considerable number of no matches (14 out of 180), it was relatively successful in avoiding suboptimal matches (only 20 out of 180). The CIEA mechanism suffered from both problems of determining an equal number of no and suboptimal matches; still, it surpassed CFA in its ability to avoid suboptimal matches.

These results are consistent with others found by Hoffman and Spitzer (1982) and Radner and Schotter (1989). In both of those experiments, it was found that face-to-face negotiation is extremely successful in avoiding no-trade or, in this case, no-match behavior.⁷ Since the CIEA mechanism is a voice-to-voice mechanism, which is similar to face-to-face bargaining, it appears that such behavior carries over to it. In Radner and Schotter (1989), the face-to-face mechanism used led to trades in almost all situations in which it was profitable to do so. Similar results were found by Hoffman and Spitzer, yet Roth (1979) reports that when bargainers communicate through a computer terminal by written messages to anonymous partners, a considerable number of no trades do occur. Hence the CIEA mechanism shares what appears to be a common property of face-to-face mechanisms—they are nonwasteful in their ability to avoid senseless no trades or no matches.

It is also interesting to note that in the only other experimental investigation of a telephone market we know of—the Hong and Plott (1982) experimental study of the pricing of barge traffic along the Mississippi, the mechanism is found to yield surplus efficiencies ranging between 83%

⁷ In both of those experiments, if a trade is consummated, it automatically is an optimal trade, so they always yielded what we would call 100% surplus efficiencies.

and 94%, with an overall average of 91%. These efficiencies are comparable to those found for the CFA experiment, which was itself a telephone market.

2. In the CIEA and CFA mechanisms, there appears to be a difference as to *when* during the 5-minute round we find trades taking place. Roth has commented elsewhere (Roth, Murnighan, and Schoumaker 1988) that when negotiations are made through indirect messages rather than by voice contact, and there is a time limit set for bargaining, trades seem to consummate at the very end of the trading period, almost at the last second. In contrast, Radner and Schotter (1989) noted that in their face-to-face bargaining experiment, transactions were consummated very early in the bargaining period, and there was certainly no tendency to exhibit any deadline effect. Similar results were found in the CIEA and CFA experiments, which were respectively message and face-to-face mechanisms. In the CIEA experiment, almost every round lasted the full 5 minutes allotted to it, with many trades taking place during the 10-second countdown time. In the CFA experiment, however, many if not most trades were made quickly, and few trading periods lasted the full 5 minutes.

3. Prices tended to be higher in the CFA experiments, followed by the SM and then the CIEA. The difference between the CFA mechanism and the others is statistically significant. The fact that the CFA mechanism yielded higher prices than the CIEA mechanism is, at first, somewhat surprising since from the results obtained on double oral auctions one is led to believe that when the strategy space of one side of a market is restricted so that they can only accept or reject bids but not make counteroffers, their payoffs should rise (see Smith [1982] for a discussion of this point). In our experiments, just the opposite occurred. We find that prices in the CIEA mechanism—a mechanism in which the S-types are relatively passive—were lower than those in the CFA mechanism, in which they were active. We attribute this anomaly to two factors. First, the double oral auction results may not be expected to carry over here since neither of these mechanisms is specifically of that type. Those previous results may be very institutionally sensitive. Second, from what we *directly* observed of the CIEA mechanism in action, we feel it is very susceptible to collusion on the part of U-types (buyers). Tacit collusion is made possible by the common knowledge of bids made and of the identities of those making them and by the opportunities available to punish those who appear to be defecting from the collusive strategy. These conditions were present in our CIEA experiments. Under the CIEA system, U-types could easily signal their intentions or bidding strategies through the bids they submitted, as these immediately became known to all market participants. Moreover, repeated participation in the market, as occurs in this multiround experiment, enhances the ability of U-types to punish others when they detect an attempt to raise price in any given round. In a number of instances,

there clearly existed among the U-types in our CFA experiment a “meeting of the minds” or an unspoken understanding about appropriate bidding strategies. The effect of this was to keep prices low. Evidence of collusion was hard to find in the other experiments we ran.

4. The fact that the mechanisms can be ranked in a statistically significant manner with respect to prices does not mean that they adhere to the same rankings when we look at payoffs. The reason for this is that the expected payoff from a mechanism must include in it the probability of being matched. Even though the SM and CIEA mechanisms yielded prices that were beneficial to the U-types, they produced a sufficient number of no matches and suboptimal matches (i.e., less profitable matches) so as to diminish the profitability of the mechanism for the U-type subjects. However, if we condition the subjects’ payoffs on whether or not they were matched, then we find that for U-type subjects *who were successfully matched*, both the SM and CIEA mechanisms were significantly better than the CFA mechanism. The opposite was true for S-types.

5. In the SM mechanism, the U-type subjects misrepresent their preferences far more than do the S-types. This is true despite the fact that the mechanism treats them symmetrically in the sense that misrepresentations by the S-type subjects determine the lower bound of the negotiation range, while those by the U-types determine the upper bound. Hence, we had no a priori grounds on which to expect such a difference. The fact that one appeared is consistent with the Radner and Schotter (1989) experiments, which found that in a symmetric sealed-bid simultaneous-move mechanism, buyers tended to shave their bids more than sellers. Those asymmetries could not be sufficiently explained by Radner and Schotter, likewise, we have no satisfactory explanation for our finding here.

B. Conclusions

While our results must be considered tentative at this point, we do feel they have shed considerable light on the behavioral and operational mechanics of the bidding systems studied. Our experiments have verified our suspicions about the way the current free-agency system works. They further suggest that some features of the formal matching mechanism (SM) might prove beneficial to baseball’s free agency. Some additional experiments are clearly in order, however, before definitive conclusions on that score can be reached.

References

- Demange, Gabrielle, and Gale, David. “The Strategy Structure of Two-sided Matching Markets.” *Econometrica* 53 (1985): 873–88.
- Gale, David, and Shapley, Lloyd. “College Admissions and the Stability of Marriage.” *American Mathematical Monthly* 69 (1962): 9–15.

- Harrison, Glenn, and McCabe, Kevin "Stability and Preference Distortion in Resource Matching: An Experimental Study of the Marriage Problem" Unpublished manuscript, Albuquerque: University of New Mexico, 1990
- Hoffman, Elizabeth, and Spitzer, Matthew "The Coase Theorem: Some Experimental Tests" *Journal of Law and Economics* 25 (1982): 73-98
- Hong, James, and Plott, Charles "Rate Filing Policies for Inland Water Transportation: An Experimental Approach" *Bell Journal of Economics* 13 (Spring 1982): 1-20
- Knuth, Donald E. *Marriages Stables*. Montreal: Les Presses de l'Université de Montreal, 1976
- Leonard, Herman B. "Elicitation of Honest Preferences for the Assignment of Individuals to Positions" *Journal of Political Economy* 93 (1983): 461-80
- Olson, Marc, and Porter, David "An Experimental Examination of the Assignment Problem" *Games and Economic Behavior* (1994), in press
- Radner, Roy, and Schotter, Andrew "The Sealed-Bid Mechanism: An Experimental Study" *Journal of Economic Theory* 48 (June 1989): 179-220
- Roth, Alvin "The Economics of Matching: Stability and Incentives" *Mathematics of Operations Research* 7 (1982): 617-28
- "The Evolution of the Labor Market for Medical Interns and Residents: A Case Study in Game Theory" *Journal of Political Economy* 92 (1984): 991-1016 (a)
- "A Misrepresentation and Stability in the Marriage Problem" *Journal of Economic Theory* 34 (1984): 383-87 (b)
- "The College Admissions Problem Is Not Equivalent to the Marriage Problem" *Journal of Economic Theory* 36 (1985): 277-88
- Roth, Alvin, and Malouf, M. "Game Theoretic Models and the Role of Information in Bargaining" *Psychological Review* 86 (1979): 574-94
- Roth, Alvin, Murnighan, Keith, and Schoumaker, Francois "The Deadline Effect in Bargaining: Some Experimental Evidence" *American Economic Review* 78 (1988): 806-23
- Roth, Alvin, and Sotomayor, Marilda. *Two-sided Matching: A Study in Game Theoretic Modeling and Analysis*. Cambridge: Cambridge University Press, 1990
- Shapley, Lloyd, and Shubik, Martin "The Assignment Game I: The Core" *International Journal of Game Theory* 1 (1972): 111-30
- Smith, Vernon "Reflections on Some Experimental Market Mechanisms for Classical Environments" In *Choice Models for Buyer Behavior*, ed. Leigh McAlister, pp. 13-47. Suppl. 1 to Research in Marketing Series. Greenwich, CT: JAI Press, 1982