NBER WORKING PAPER SERIES

AN ANALYSIS OF THE SELECTION OF ARBITRATORS

David E. Bloom

Christopher L. Cavanagh

Working Paper No. 1938

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 1986

The authors would like to thank Orley Ashenfelter, two referees, and seminar participants at Harvard University, Princeton University, the NBER, and M.I.T. for helpful comments, This research was supported by NSF Grant SES-8309148. The research reported here is part of the NBER's research program in Labor Studies. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Working Paper #1938
June 1986

An Analysis of the Selection of Arbitrators

ABSTRACT

This paper analyses data on union and employer rankings of different panels of arbitrators in an actual arbitration system. A random utility model of bargainer preferences is developed and estimated. The estimates indicate that unions and employers have similar preferences, in favor of lawyers, more experienced arbitrators, and arbitrators who seem to have previously favored their side. Alternative rankings models, which are estimated to test whether bargainers rank arbitrators strategically, reveal no evidence of strategic behavior.

David E. Bloom
Department of Economics
Harvard University
Cambridge, MA 02138

Christopher L. Cavanagh
Department of Economics
Harvard University
Cambridge, MA 02138

I. Introduction

Arbitration is a rapidly-growing method for resolving disputes. It is used widely in the U.S. to resolve private disputes arising under collective bargaining agreements and commercial contracts, to resolve certain types of civil disputes, and to set wages in the public sector. Although arbitration has been applied in a wide range of settings and takes numerous forms, the central feature of virtually all arbitration mechanisms is that they involve a third party, i.e., an arbitrator or panel of arbitrators, hearing and deciding how a dispute is to be resolved. In this respect, arbitration may be viewed as a private sector analog to the court system with arbitrators performing similar functions to judges. Arbitration tends, however, to be a cheaper, quicker, and less formal method of dispute resolution than the court system. Arbitration systems also tend to provide disputing parties with greater latitude in choosing "judges" (i.e., arbitrators) than do court systems.[1]

Among persons who are regularly involved in arbitration, it is generally accepted that "good arbitrators" are the key to "good arbitration." However, there seems to be far less consensus about the meaning of the phrase "good arbitrators." For example, is a good arbitrator someone who favors your side or someone who is painstakingly fair in arriving at a decision? Is it someone who has extensive experience with the type of dispute at hand or someone with good common sense and the ability to analyze, interpret, and judge? Is it someone who tries to appease the parties by splitting decisions or someone who strictly "calls them as he or she sees them?"

Our purpose in this paper is to address these and other related questions by presenting an empirical analysis of the selection of arbitrators. We do this by analyzing a remarkable set of data on the preferences of unions and employers for different arbitrators under New Jersey's Fire and Police

Arbitration Law. According to this law, disputes over terms and conditions of employment involving New Jersey's organized public safety employees and the governments that employ them must be resolved by arbitration. Arbitrators, whose awards are binding by law, are chosen by the bargainers from a roster of roughly seventy names maintained by the New Jersey Public Employment Relations Commission (PERC). In cases in which the bargainers are unable to negotiate an agreement, PERC circulates a list of seven arbitrators and their resumes to the parties, each of which is instructed to veto three names and to rank, in order of their preferences, the remaining four names. PERC then appoints as arbitrator the individual who was not vetoed by either side and whose combined rank is highest; rank ties are broken randomly by PERC.

Using information on employer and union rankings of different panels of arbitrators, along with information on the characteristics of the arbitrators, we attempt to provide direct evidence on the following three issues:

(1) How similar are the preferences of unions and employers with respect to a given panel of arbitrators?

(2) What characteristics of arbitrators do unions and employers find desirable or undesirable? Do the unions and employers attach the same or different weights to specific characteristics? and

(3) Do unions or employers engage in strategic behavior in ranking arbitrators?

In proceeding this way, we also hope to shed light on three broader issues. First, there has recently developed in the academic literature a body of theoretical work on the subject of bargaining and arbitration.[2] The basic premise of most of this work is that arbitration is simply a mechanism for distributing income between conflicting interests. In contrast, institutional economists and labor practitioners place greater emphasis on arbitration as a mechanism for helping disputants identify and reach efficient outcomes. In their view, arbitrators are professional gatherers and processors of information

who play a highly constructive role in a bargaining process which is better treated as a cooperative attempt at problem-solving than as direct economic conflict. By analyzing the similarity of union and employer preferences for individual arbitrators and for different arbitrator characteristics, we suspect that much can be learned about the general issue of whether collective bargaining is primarily an institution of cooperation or conflict.

Second, one of the most important characteristics of arbitration systems is that they may be designed in different ways. Indeed, one key dimension in which arbitration systems differ involves the mechanism for selecting the arbitrator. Most mechanisms take account of the parties' preferences, either through a rank-four/veto-three system like New Jersey's, or by allowing each party to successively veto a name from an odd-numbered list of three or more arbitrators. Other arbitration systems appoint arbitrators on a purely rotating basis from a list agreed to in advance by potential disputants or established by a third party such as the state. A final system involves the appointment of a single individual or panel of individuals to arbitrate all disputes involving a particular set of parties and arising in a specified period of time. The key feature of all of these systems is that they guarantee the appointment of an arbitrator without requiring explicit agreement (or even face-to-face contact) by two parties who are unable to reach agreement on some other (substantive) matter. In addition, these systems all prescreen individuals before they are added to the master list of eligible arbitrators. However, the first two systems provide for an additional level of screening by the parties prior to the appointment of an arbitrator to hear a particular case. This additional level of screening is said to contribute to the legitimacy of the arbitrator and his award in the eyes of the parties. However, it can also contribute to delays in the arbitration process, which is one of the most frequently cited complaints

about arbitration. Thus, it is interesting to ask whether the appointment of arbitrators can be left entirely in the hands of the state or some impartial organization like the American Arbitration Association or whether it is important to take account of the parties' preferences on a case-by-case basis. We will address this question below when we analyze the strength of union and employer preferences for different members of a set of "prescreened" arbitrators.[3]

Third, our study raises important questions about the possibility of strategic behavior and its treatment in empirical analysis. In particular, it is well known that the outcomes of voting mechanisms can often be manipulated by the strategic misrepresentation of preferences. Although it is natural to address this problem by directly estimating a structural model of the underlying economic game, the complexity and dimensionality of the game may render this approach infeasible, as it does in our case. Thus, we develop some indirect tests that we think will let us "back our way" into strategic behavior if it is there. We suspect that the type of indirect approach we propose may be useful in a variety of game-theoretic settings in which the games are too complicated to solve.

In the next section of the paper we provide some institutional background on the selection of arbitrators and describe our data. In Section III we set out a simple random utility model that we use to represent the preferences of employers and unions for arbitrators with different characteristics. We also present the likelihood function we maximize to estimate the parameters of this model.[4] In Section IV we present a descriptive summary of the data. We also present estimates of the econometric model as well as two alternative models we estimate to test for the presence of strategic behavior. Section V summarizes and concludes the paper.

II. Institutional Background

Most of the data analyzed in this paper were drawn from the PERC's arbitration records. First, we collected information on the lists of arbitrators sent by PERC to disputing parties along with the preference rankings returned to PERC by the parties. We focused only on cases involving 1980 contract negotiations. That was the third year of operation of the New Jersey arbitration system and the third (and final) year in which PERC used its original master list of eligible arbitrators to form panels.[5] Thus, we felt that by 1980 the parties had reasonably good information about the arbitrators on which to base their preference rankings. In many cases, at least one of the parties did not strictly follow PERC's request for a preference ranking. Sometimes parties ranked more than four names on the list; other times the parties vetoed more than three names; in a few cases a party responded to PERC by saying that all seven names were equally acceptable; there were also a number of cases in which a party either failed to express its preferences to PERC or its preferences were simply not recorded in the PERC records.[6] Altogether we collected information on 193 arbitration panels. Of these, 75 are perfect in the sense that both parties ranked four names and vetoed three names. It should also be noted that many (indeed, most) of the cases for which PERC circulated arbitration panels did not end up being arbitrated. In other words, many disputants were able to reach voluntary settlements after the arbitration panel was circulated but before a binding arbitration award was rendered. This characteristic of the bargaining/arbitration process in New Jersey explains why some parties did not report their arbitrator preferences to PERC (i.e., either their case was settled before the due date for reporting their preferences or they expected that it would not end up in arbitration).

Second, we collected information on the characteristics of the 69 arbitrators on PERC's master list. This information was derived from a variety of sources including PERC's 1978 and 1979 interest arbitration records and awards, PERC's 1979 grievance arbitration records, and the arbitrators' resumes. In collecting this background information, we were guided by the institutional literature on the relevant characteristics of arbitrators and by conversations with labor relations practitioners.[7] Roughly speaking, the relevant characteristics of arbitrators seem to fall into four categories: (1) impartiality, (2) consistency, (3) training, and (4) experience.[8]

Impartiality refers to an arbitrator's lack of predisposition to rule in favor of one side or another. This characteristic is usually judged by considering an arbitrator's prior decisions. It appears to be the most important characteristic of an arbitrator since no party is likely to be satisfied with an arbitrator it perceives to be biased against its position. There has even been some debate over whether disputants prefer arbitrators who they perceive to be biased in their favor. On the one hand, such bias is desirable because it suggests a higher probability of receiving a favorable arbitration decision. But on the other hand, it damages the integrity of the institution of arbitration and does not promote the legitimacy and mutual acceptability of the arbitrator's award.

Consistency refers to the extent to which an arbitrator decides cases solely on their merits, without reference to his (or her) "box score" of previous decisions. The parties' concern with the consistency of arbitrators stems from their awareness that many arbitrators derive considerable income from arbitrating and therefore may have an incentive to "split their awards" in order to appear impartial and maintain their acceptability. Such a practice greatly threatens the institution of arbitration since it suggests that a

certain fraction of cases will be won by a given party, not on their merits, but simply because they are brought before an arbitrator. Consistency is most often judged by subjectively reviewing an arbitrator's previous awards to see that similar decisions were reached in similar cases.

A third dimension along which arbitrators differ is their training. Most labor arbitrators are lawyers, undoubtedly because legal training is well-suited to analyzing and judging the vast majority of labor disputes, i.e., disputes over the terms of existing contracts, also known as grievances. However, the New Jersey system involves disputes over the terms of new contracts, i.e., disputes of interest, with the most common and important issue in dispute being wages. Thus, one might expect that training in other areas, and especially in economics, might be particularly desirable to the parties.

Arbitration experience is another important characteristic of prospective arbitrators. Practitioners generally regard this characteristic as a measure of an individual's expertise as an arbitrator and usually insist upon an experienced arbitrator in cases involving complex or otherwise difficult issues. At a more theoretical level, it seems likely that experience is desirable because it reduces the uncertainty that risk-averse disputants have over the outcome of arbitration and thereby reduces its (indirect) cost.[9] Indeed, the acquisition of experience is generally regarded to be the most significant hurdle faced by aspiring labor arbitrators in the U.S.[10]

III. A Model of Arbitrator Selection

In this section we outline a simple model of arbitrator selection. We start by considering a situation in which an employer (E) and a union (U) are unable to reach agreement on a vector of contract items although they do agree (or are compelled by law) to have their dispute resolved by arbitration. A list of seven potential arbitrators along with information on their qualifications is

circulated among the two parties by an impartial organization. Each party is instructed to veto three names and rank the remaining four names in order of their preference. The individual who is not vetoed by either side and who has the lowest combined rank is appointed to hear the case; rank ties are broken by coin tosses.[11]

We now assume that each party has a preference ordering defined over the entire set of possible arbitration outcomes and that these preference orderings can be represented by well-behaved Von Neumann-Morgenstern utility functions. We also assume that each party has one set of prior beliefs about the distribution of arbitration outcomes for each of the seven potential arbitrators. Finally, we assume that these sets of prior beliefs depend on the arbitrator's qualifications, some of which are observed.

In this model, an arbitrator's characteristics determine a party's prior beliefs about the arbitrator's decision. These beliefs, in turn, determine the expected utility that a party associates with that arbitrator. It follows that there exist direct mappings of arbitrator characteristics into an expected utility for each party. For each potential arbitrator (i), we write these expected utilities as a linear function of the arbitrator's characteristics.

$$(1a) \qquad Y_{Ei} = X_i\beta_E + \epsilon_{Ei} \qquad (i = 1, 2, \ldots, 7)$$

$$(1b) \qquad Y_{Ui} = X_i\beta_U + \epsilon_{Ui} \qquad (i = 1, 2, \ldots, 7)$$

where Y_{Ei} and Y_{Ui} are the expected utilities the employer and union associate with arbitrator i; X_i is a vector of observed characteristics of arbitrator i; β_E and β_U are vectors of unknown (reduced-form) parameters characterizing the preferences of the employer and union for different arbitrator characteristics; and ϵ_{Ei} and ϵ_{Ui} represent random utility effects.

Given this framework, the simplest model one could adopt for arbitrator

selection would treat each party as ranking arbitrators in order of their expected utilities. In other words, each party's stated ranking is assumed to coincide with the ordering of its true preferences. More complicated models which account for each party's incentives to rank arbitrators some other way are possible, too. For now, however, we shall work with the simple model and defer consideration of alternative models to Section IV.C.

In order to make this model of arbitrator selection empirically tractable, it is necessary to treat the unobserved characteristics ϵ_{Ei} and ϵ_{Ui} (i= 1, ..., 7) as random variables. We do this by assuming that these random variables have independent extreme value distributions:

$$(2) \qquad \Pr(\epsilon_{pi} \leq x) = \exp(-e^{-x}) \qquad \begin{matrix}(p = E, U) \\ (i = 1, \ldots, 7)\end{matrix} \quad .$$

This distributional assumption is quite common in random utility models of this general form.[12] It is also quite convenient, both analytically and computationally, in the present application.

We derive the likelihood function for each party's rank choices by developing an expression for the probability of a particular ranking. For example, suppose the union ranked a list of seven arbitrators in the following order: (1, 2, 3, 4, Veto, Veto, Veto). The likelihood of this ranking is simply the probability that $(Y_{U1} > Y_{U2} > Y_{U3} > Y_{U4} > Y_{U5}, Y_{U6}, Y_{U7})$. The general formula for the probability of a particular ranking is given by

$$(3) \qquad \prod_{j=1}^{K} \mathrm{MNL}(A(j) \mid \text{Remaining choice set}, \beta)$$

where A(j) denotes the arbitrator who receives the j^{th} rank, K is the number of arbitrators that receive a rank (e.g., four, in the standard case), and MNL is the multinomial logit probability that a party with preference parameters β will most prefer arbitrator A(j) given the option of choosing from the 8-j least pre-

ferred arbitrators. This probability is defined as

$$(4) \qquad MNL(A(j) \mid \text{Remaining choice set, } \beta) = \frac{\exp(X_{A(j)}\beta)}{\sum_{j \epsilon \Phi_{8-j}} \exp(X_{A(j)}\beta)}$$

where Φ_{8-j} refers to the remaining choice set of 8-j least preferred arbitrators.

The probability in equation (3) is simply the product of the multinomial logit probabilities of ranking an arbitrator from among those arbitrators that have not already been ranked. Notice that this model is extremely amenable to dealing with the type of unbalanced data configuration we face. For example, in cases where more than four arbitrators are ranked, the likelihood of the ranking will simply consist of the product of more than four multinomial logit probabilities. Thus, this model makes efficient use of all of the preference ordering information we have available. It should also be noted that this likelihood function is globally concave. This property guarantees a unique maximum if the model is asymptotically identified and typically assures numerical stability in computing maximum likelihood estimates.

IV. Empirical Results

A. Descriptive Statistics

We begin our discussion of empirical results by presenting a descriptive summary of the data. Table I presents the distribution, across arbitration panels, of the number of arbitrators not vetoed by either side. If employer and union preferences tend to be in direct conflict, one would expect most panels to yield only a single jointly-ranked arbitrator. Alternatively, if employer and union preferences tend to be similar, one would expect most panels to yield four jointly-ranked names. As the first row of Table I makes clear, neither of these extremes appears to be true. Over 80 percent of the panels yielded two or three jointly acceptable arbitrators with an average overlap of nearly 2.5

Table I

Percent Distribution of Number of Arbitrators
Not Vetoed by Either Party*

	1	2	3	4	Average Overlap
A. All Cases (N=75)	9.3	44.0	37.4	9.3	2.47
B. Arbitrated Cases (N=29)	6.9	41.4	37.9	13.8	2.59
C. Negotiated Cases (N=46)	10.9	45.7	37.0	6.5	2.39
D. Independent Rankings	11.4	51.4	34.2	2.9	2.28

*These statistics, as well as those in Table II and Figure I, are based on the 75 cases in which both sides ranked four arbitrators and vetoed three.

arbitrators per panel. Observe also the last row of Table I that presents the distribution of the number of jointly-ranked names under the assumption that the parties' rankings of arbitrators are independent of each other. A chi-square test comparing the observed distribution (row A) to the independent-rankings distribution (row D) yields a test statistic of 11.2, which is statistically significant at the .05 level (3 degrees of freedom). This result suggests that the parties' rankings are not independent. Moreover, the biggest contribution to the test statistic comes from the proportion of lists with 4 overlapping names, providing some evidence of positive correlation in the parties rankings. In addition, it is worth noting that the distribution of jointly-ranked names is not statistically significantly different in the subset of 29 cases that ended up being arbitrated than in the subset of 46 cases that ended up being negotiated (i.e., a chi-square test for the equality of the true distributions yields a test statistic of 1.4, which is not in the 5 percent tail of a χ^2 distribution with 3 degrees of freedom).

Since Table I is not informative about the closeness of employer and union rankings within each panel of arbitrators, we have computed a rank correlation coefficient (ρ) for each of the 75 "perfect" cases

$$(5) \qquad \rho = \frac{1}{28} \sum_{i=1}^{7} R_{Ei}R_{Ui} - 4$$

where R_{Ei} and R_{Ui} are the ranks assigned by the employer and the union, respectively, to the i^{th} arbitrator on the panel.[13] In Figure I we plot the (observed) frequency distribution of this statistic. As can be seen from the plot, there are some cases in which the parties' rankings are very different (i.e., $\rho \leq -.5$), some cases in which they are very similar (i.e., $\rho \geq .5$), and many cases in which they seem to be uncorrelated. However, on balance, the plot seems to produce evidence of a slight tendency for the employer and union rankings to be positively correlated ($\bar{\rho} = .13$).[14]

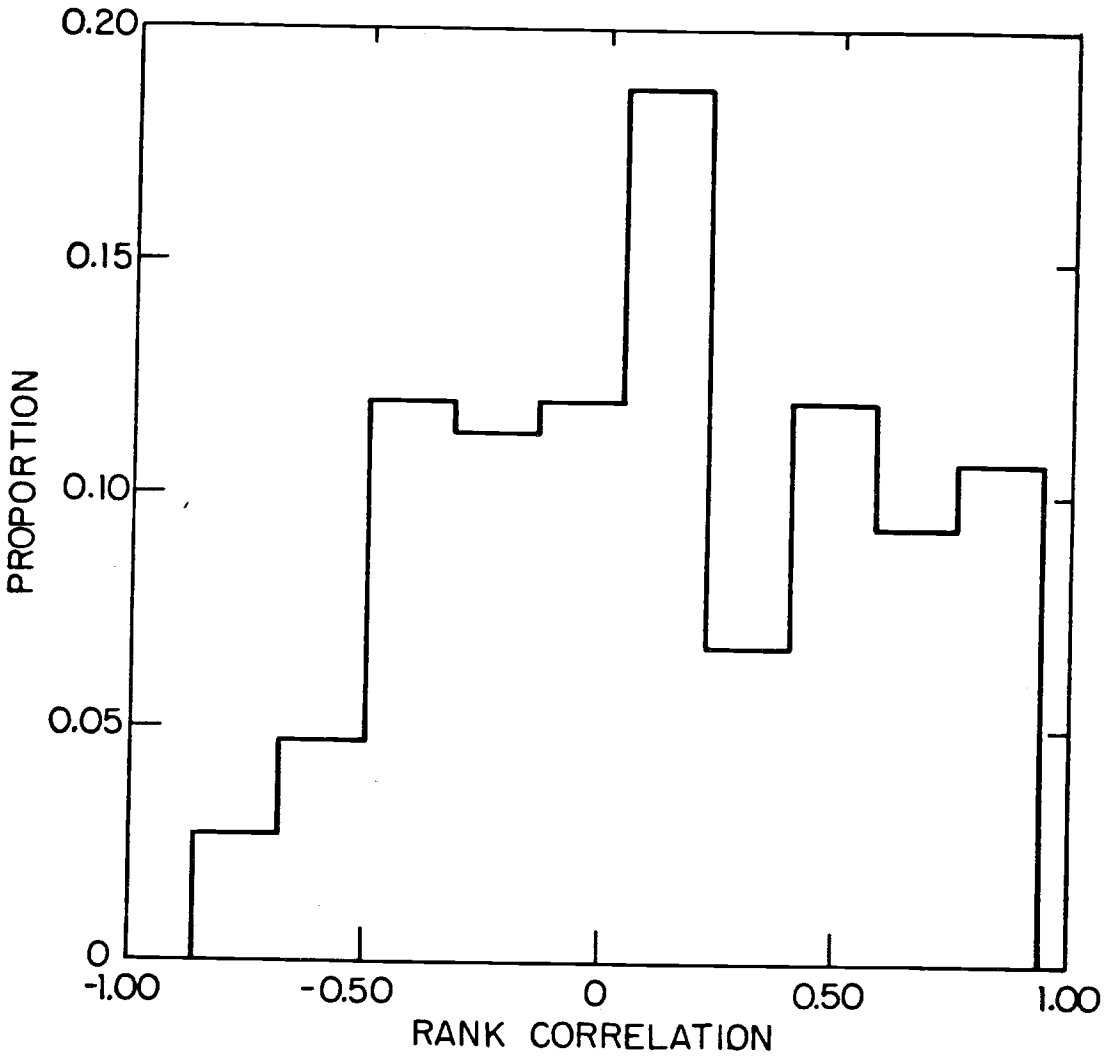

Figure I

Frequency Distribution of Rank Correlation Coefficient
Computed for Employer and Union Rankings

We now take a closer look at the similarity of employer and union rankings by constructing a two-way contingency table of these rankings and formally testing the hypothesis that they are independent. This contingency table, which reports the number of times that arbitrators listed on the 75 panels received each of twenty-five possible combinations of employer and union ranks, is presented in Table II. This table also presents (in parentheses) theoretical frequencies for each cell computed under the hypothesis that the rankings are independent.[15] The hypothesis of independence is tested using the familiar χ^2 statistic which has the value 36.4 for this table. Since this statistic has 15 degrees of freedom, we reject the hypothesis of independence at all conventional significance levels (e.g., the critical value of a χ^2_{15} random variable at the 99 percent level is 32.8).

Before leaving Table II, it is interesting to note that the (1, 1) cell makes the biggest contribution to the χ^2 statistic. In other words, the hypothesis of independence is rejected largely because the unions and the employers ranked the same individual first in 21 of the 75 panels they reviewed. Indeed, the observed frequencies are greater than the theoretical frequencies for all of the diagonal cells in Table II. This provides some further evidence that employer and union preferences tend to be at least moderately similar.

We now describe the characteristics of the New Jersey arbitrators. To begin with, 45 percent of the arbitrators have law degrees and 12 percent are Ph.D. economists; the remaining 43 percent are labor relations practitioners, some of whom have Ph.D.'s in labor relations or other areas. All of the arbitrators, of which there are only four women, have mediation or factfinding experience in public sector wage disputes. In addition, two-thirds of the arbitrators had rendered at least one interest arbitration award in the New

Table II

Two-Way Distribution of Observed Employer and Union Rankings
(Theoretical frequencies are reported in parentheses)

		Union					
		1	2	3	4	V	Total
EMPLOYER	1	21 (10.7)	10 (10.7)	11 (10.7)	5 (10.7)	28 (32.1)	75
	2	8 (10.7)	16 (10.7)	15 (10.7)	14 (10.7)	22 (32.1)	75
	3	7 (10.7)	12 (10.7)	14 (10.7)	7 (10.7)	35 (32.1)	75
	4	10 (10.7)	10 (10.7)	10 (10.7)	15 (10.7)	30 (32.1)	75
	V	29 (32.1)	27 (32.1)	25 (32.1)	34 (32.1)	110 (96.4)	225
Total		75	75	75	75	225	525

Jersey system during its first two years of operation, although only 12 percent of the arbitrators had rendered five or more awards. On average, the New Jersey arbitrators have about fifteen years of arbitration experience, with the range of experience being from four years to forty years. Nearly two-thirds of the arbitrators had been appointed to hear two or more grievance arbitration cases in the New Jersey public sector in 1979. About one half of the arbitrators are members of the National Academy of Arbitrators.

In Figure II we plot the average rankings received from the employers and the unions by the 65 arbitrators who received at least two rankings from each party.[16] In computing these averages we were able to use ranking information from most of the 193 panels by assigning vetoed arbitrators a rank equal to the median of the unassigned ranks. As Figure II makes clear, there is considerable dispersion across arbitrators in the average rankings they received from the parties. This provides further evidence that the parties prefer some arbitrators to others. Moreover, the average rankings in Figure II exhibit a fairly high positive correlation (ρ = .51), providing a further indication of similarity between the parties' preferences. Of course, one weakness of the average-rank statistics plotted in Figure II is that they do not control for the characteristics of the other arbitrators who appeared on the same lists. However, controlling for those characteristics is accomplished by our structural model of arbitrator selection, to which we now turn.

B. Estimation Results

Table III presents estimates of the employer and union preference parameters of the random utility model in equation (1). Positive signed coefficients indicate that a particular characteristic tends to increase the expected utility of an arbitrator and, therefore, the likelihood that the arbitrator receives a favorable rank. Negative coefficients indicate the

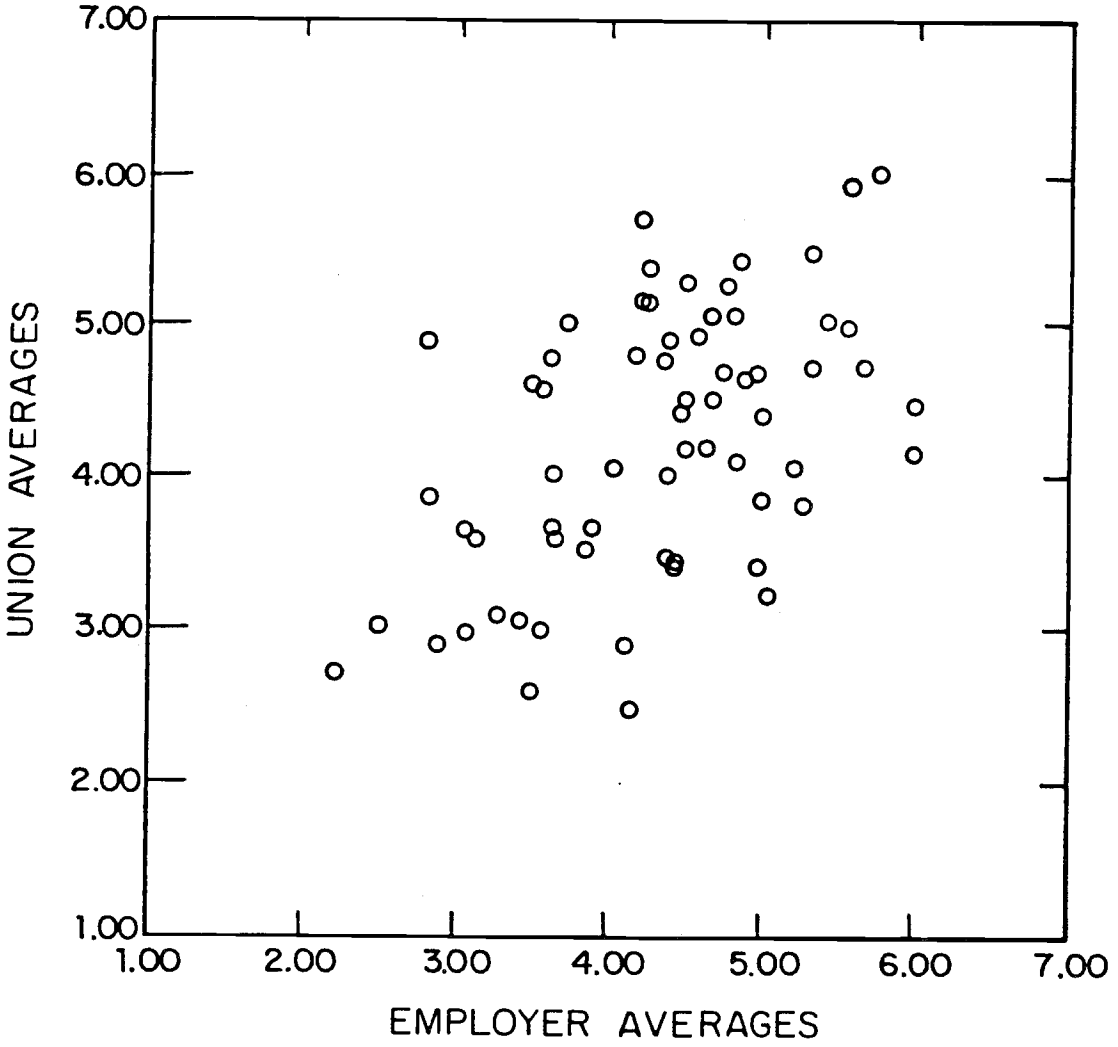

Figure II

Average Rankings of Individual Arbitrators,
by Employers and Unions

opposite. The estimates were obtained by maximizing the log of the likelihood function implied by equation (3) for a particular specification of arbitrator characteristics. The maximization was accomplished on a personal computer using the modified quadratic hill-climbing method (GRADX) proposed by Goldfeld and Quandt (1972). All of the rankings data available from the 193 arbitration panels were used in the estimation (i.e., we used 129 employer rankings and 160 union rankings). Since the estimates are maximum likelihood, standard tests of their significance can be performed based on their asymptotic normality.

In specifying the vector of arbitrator characteristics, we paid close attention to the institutional literature on arbitrator selection reviewed in Section II. Thus, we attempted to capture the training dimension of an arbitrator's characteristics by including dummy variables for lawyers and economists, with all other arbitrators comprising the reference category. We have no strong priors on the effect of these variables although it does seem likely that economists would be viewed as best able to resolve the wage and benefit disputes that are central to most negotiations in New Jersey. Our specification also includes variables reflecting an arbitrator's experience. Specifically, we include the number of grievance arbitration appointments each arbitrator received in 1979 and the number of conventional and final-offer arbitration awards each arbitrator rendered in New Jersey prior to 1980.[17] If experience is truly a desirable characteristic, we would expect both of these measures, which are derived from independent arbitration systems, to be positively associated with each parties' preferences for an arbitrator.

Finally, we attempt to control for an arbitrator's impartiality in several ways. First, we include a variable (FAVU) defined to be the difference between the number of final-offer cases decided in favor of unions and the number decided in favor of employers in the years 1978 and 1979. This variable is non-zero for

about 55 percent of the arbitrators and takes on values between -4 and 4. If this variable adequately measures the tendency of an arbitrator to be more sympathetic to one side than the other, and if the parties prefer arbitrators they perceive to be relatively sympathetic to their position, this variable should be associated with more favorable rankings from the union and less favorable rankings from the employer. Second, we construct a variable (FAVUHAT) which measures the expected difference between the number of final offer cases decided in favor of unions and employers. This measure is constructed using estimates of the unconstrained arbitrator decision-making functions reported in Ashenfelter and Bloom (1984, Table 2, columns 1 for 1978 and 1979) along with information on the final offers in the 1978 and 1979 cases. The inclusion of this variable in the empirical model refines the first measure of impartiality by controlling for an important subset of facts in the final-offer cases. Thus, under the assumption that the parties evaluate the record of each arbitrator's final-offer decisions in light of the quality of the final offers they were forced to choose between, our expectation is that the coefficient of FAVUHAT will be opposite in sign but have the same absolute magnitude as the coefficient of FAVU. Third, as a simple alternative measure of arbitrator bias, we include a variable (AVEDEV) defined as the average wage increase awarded by each arbitrator in their 1978 and 1979 cases. This variable, which incorporates information from both conventional and final-offer cases, is expressed as a deviation from the average of all awards rendered in 1978-79; it has the value zero for arbitrators who made no awards. If case facts are similar across bargaining disputes, this variable will provide a reasonable measure of each arbitrator's tendency to favor the union or the employer.

The estimates in Table III reveal a number of interesting features about employer and union preferences for arbitrators. First, both the employers and

Table III

Maximum Likelihood Estimates of Employer and Union Preference Functions, Complete Rankings Model (CR)*

Variable	Definition	Sample Average for 69 Arbitrators	Employer Preferences (1)	Employer Preferences (2)	Employer Preferences (3)	Employer Preferences (4)	Union Preferences (1)	Union Preferences (2)	Union Preferences (3)	Union Preferences (4)
LAWYER	Arbitrator has a law degree	.45	.204 (.114)	.177 (.121)	.174 (.113)	.101 (.112)	.158 (.100)	.109 (.105)	.194 (.099)	.219 (.098)
ECONOMIST	Arbitrator is a Ph.D. economist	.12	.478 (.205)	.437 (.214)	.433 (.207)	.410 (.208)	-.782 (.219)	-.865 (.225)	-.746 (.220)	-.744 (.223)
GRAPTS79	Number of grievance arbitration appointments in NJ public sector in 1979	3.99 (4.66)	.076 (.013)	.076 (.013)	.076 (.013)	.070 (.015)	.078 (.012)	.078 (.012)	.080 (.012)	.074 (.014)
NFOA	Number of final-offer arbitration (FOA) awards rendered in NJ in 1978-79	1.73 (2.56)	.076 (.022)	.060 (.032)	.058 (.020)		.005 (.019)	-.025 (.028)	.026 (.018)	
NTOT	Total number of conventional and final-offer awards rendered in NJ in 1978-79	2.23 (3.12)				.046 (.021)				.035 (.018)
FAVU	Number of FOA awards decided in favor of unions minus number of FOA awards decided in favor of employers, 1978-79	.22 (1.39)	-.077 (.032)	-.085 (.034)			.097 (.030)	.079 (.032)		
FAVUHAT	Predicted value of FAVU (See text)	.48 (.76)		.076 (.116)				.163 (.105)		
FAVDIFF	FAVU minus FAVUHAT	-.27 (1.17)			-.085 (.034)				.082 (.032)	
AVEDEV	Arbitrator's average percent wage increase, expressed as deviation from grand average of wage increases.	-.05 (.88)				-.046 (.054)				.108 (.049)
Log-likelihood ($\beta = 0$)			-769.1	-768.9	-768.9 (-827.4)	-772.8	-931.0	-929.7	-932.8 (-1018.0)	-933.6

*Estimated standard errors are reported in parentheses.

the unions tend to prefer individuals with law degrees to labor relations practitioners. However, employers prefer economists to both of these groups whereas unions prefer both of these groups to economists. Perhaps this somewhat surprising result is explained by the fact that economists are likely to be heavily influenced by efficiency considerations whereas lawyers are more likely to place greater emphasis on equity.

Second, the three measures of an arbitrator's experience (GRAPTS79, NFOA, and NTOT) typically have positive coefficients in both the employer and union equations. This indicates that employers and unions both prefer more experienced arbitrators to less experienced arbitrators, controlling for variables reflecting their training and impartiality. It is, however, possible to interpret the experience variables as controlling for arbitrator-specific characteristics which are not in the model (akin to fixed effects) since, to some extent, they reflect arbitrators' past popularity. Since it is not possible for us to distinguish empirically between these alternative interpretations of the experience variable, we conclude simply that our results are not inconsistent with the hypothesis that experience is a desirable characteristic, as emphasized in the institutional literature on arbitrator selection.

Third, our measures of arbitrator bias, i.e., their past tendency to favor unions, have negative coefficients in the employer equation and positive coefficients in the union equation, whether or not we control for the case facts. This indicates that employers tend to give poorer rankings to arbitrators who seem to have favored unions in the past whereas unions tend to prefer such arbitrators. This finding is consistent with our expectation that the parties do not like arbitrators whom they perceive to be biased against them. Note also that the coefficients of FAVUHAT are not significant at the

five percent level in either the employer or the union equations and that their inclusion in the equations does not substantially change the magnitudes of the coefficients of FAVU. In addition, in the employer equation the coefficient of FAVUHAT is opposite in sign and of roughly the same magnitude as the coefficient of FAVU. However, in the union equation the two coefficients have the same sign and, counter to our expectation, we reject the hypothesis that they sum to zero at the five percent level. Nonetheless, if we constrain the coefficients of FAVU and FAVUHAT to be equal and opposite in sign (i.e., by entering FAVDIFF), the measure of bias has the theoretically expected effect and is close in magnitude to the bias coefficient in the other models.

The last point worth noting about the results in Table III relates to the magnitudes of the coefficients. These are difficult to interpret since their scale is determined by our assumption that the ϵ_{Ei} and ϵ_{Ui} are drawn independently from extreme value distributions as defined in equation (2). However, one may gauge the magnitude of the coefficients by measuring their values in relation to the standard deviation of the ϵ's. The latter are fixed by our distributional assumption and equal $\Pi/\sqrt{6}$ ($\approx$ 1.28). Viewed in this way, we see that all of the coefficient estimates are relatively small in magnitude. For example, a one standard deviation change in the number of grievance arbitration appointments only represents a change of about one-fourth of a standard deviation of ϵ. Thus, while our estimates of the effect of arbitrators' characteristics on the parties' rankings are reasonably precise, they are also quite small.[18] Given the similarity of results from alternative specifications (reported in Bloom and Cavanagh, 1985a) that allow for nonlinearities and that include alternative measures of experience, we suspect that the large amount of noise we are observing indicates that the parties are relatively indifferent to many of the arbitrators in the New Jersey system.

C. Testing for Strategic Behavior

As noted earlier, a critical assumption of our analysis is that the observed rankings data reveal the true preferences of the parties. This would be true if each party perceived that it had no incentive to misstate its true preferences. However, it is not difficult to imagine a scenario in which the parties do have incentives to rank arbitrators strategically. For example, suppose the union's true preference ranking of a list of seven arbitrators is [1, 2, 3, 4, Veto, Veto, Veto] and the employer's true preference ranking is [1, 4, 3, 2, Veto, Veto, Veto]. Although not identical, these preference orderings do have important similarities: they both rank the first arbitrator as most preferred and they both veto the fifth, sixth, and seventh arbitrators. However, if either party has some information about the other party's preferences, it is quite likely that it will have an incentive to misstate its true preferences. For example, if each party expects the other to veto the last three arbitrators on the list, but has no idea how the other party will rank the first four arbitrators, they will both try to ensure the appointment of the first arbitrator by presenting identical rankings of [1, Veto, Veto, Veto] for the first four arbitrators. Thus, the two parties' revealed preferences might appear quite similar even though the only correspondence between the true and revealed preferences of each party would be the assignment of the most favorable rank to the first arbitrator. Other examples, in which similar preferences can lead to dissimilar rankings, are possible, too. This type of behavior is particularly distressing since it can lower both parties' welfare (see Bloom and Cavanagh, 1985b). Moreover, it suggests that the estimates presented so far in this section may not reflect the parties' true preferences.

A natural way to address this issue is to model the ranking of arbitrators as a game in which each party must choose a strategy, i.e., a ranking, given

its subjective opinion about the ranking of the other party. A natural definition of equilibrium strategies for this game would be the Bayesian Nash equilibrium which has the characteristic that neither party can increase its expected utility by unilaterally changing its strategy. In principle, one could solve for each party's Nash strategy and derive a likelihood function that could be maximized to estimate each party's preference parameters conditional on this strategic behavior. However, in practice, the dimensionality of the problem makes this infeasible unless substantial structure is placed on the problem, which we prefer not to do.[19] Instead, we will develop an empirical approach to this issue.

Our test for strategic behavior primarily involves the estimation of alternative random utility models that use different subsets of information available in the arbitrator rankings data. Under the null hypothesis that the data reveal the parties' true preferences, estimates of the parameters of these models should not be significantly different. We consider two alternative models. First, we consider a model in which each party assigns a rank of one to the arbitrator it truly prefers the most, but may rank the remaining six arbitrators strategically, i.e., not in the true order of their expected utilities. The likelihood function for this model is derived from expressions for the probability that $(Y_{A(1)} > Y_{A(2)}, Y_{A(3)}, Y_{A(4)}, Y_{A(Veto1)}, Y_{A(Veto2)}, Y_{A(Veto3)})$. This is simply a multinomial logit model (MNL). Second, we consider a model in which each party's ranking correctly distinguishes between the four most preferred arbitrators and the three least preferred arbitrators, although the ranking of the top four choices may be done strategically. The likelihood function for this model is derived from expressions for the probability that $(Y_{A(1)}, Y_{A(2)}, Y_{A(3)}, Y_{A(4)} > Y_{A(Veto1)}, Y_{A(Veto2)}, Y_{A(Veto3)})$. This model shall be referred to as the rank/veto model (RV).[20]

In Table IV we present estimates of the MNL model and the RV model for specification (1) in Table III. If the rankings data do not reflect the true preferences of the parties, then the three sets of estimates will in general have different probability limits, regardless of the true underlying model. This suggests that a specification test based on the differences between the estimated parameter vectors has power against a broad range of alternative models in which the rankings do not reflect true preferences. This test should be especially powerful against strategic behavior that is closely approximated by the MNL or the RV models.[21] In constructing this test we follow Hausman (1978) who suggests a statistic of the form

$$S = (\hat{\beta}^{I} - \hat{\beta}^{II})' \, (\hat{V}^{I} - \hat{V}^{II})^{-1} \, (\hat{\beta}^{I} - \hat{\beta}^{II})$$

where the $\hat{\beta}$ are estimated parameter vectors, the $\hat{V}$ are their estimated variance-covariance matrices, and the superscripts I and II index the alternative estimators. Hausman (1978) shows that, under the null hypothesis, S is asymptotically distributed as a chi-square random variable with degrees of freedom equal to the rank of $(\hat{V}^{I} - \hat{V}^{II})$.

Test statistics for the hypotheses $\beta^{CR} = \beta^{MNL}$ and $\beta^{CR} = \beta^{RV}$ are reported in the last row of Table IV. Since the critical value for these statistics at the 5 percent level is 11.07, none of the equality hypotheses are rejected with the single exception of $\beta^{CR} = \beta^{RV}$ for employers. Indeed, the closeness of the parameter estimates in Tables III and IV is a remarkable finding that provides strong evidence that the rankings data mainly reveal the parties' true preferences. Even the two estimated parameter vectors that give rise to the χ^2 statistic of 36.35 are close in magnitude and substantively no different. The high χ^2 value appears to be the result of the two sets of estimates having virtually identical estimated variances. In other words, the statistic reflects a precise measure of a small difference; it does not reflect a large

difference between the estimated β's. Overall, then, the results of this subsection provide no support for the hypothesis of strategic behavior by either party.

V. Conclusions

This paper opens up the empirical analysis of a new area in the literature on bargaining under arbitration: the selection of arbitrators. Our major substantive findings are that (1) employers and unions distinctly prefer some arbitrators to others, (2) employer and union preferences tend to be moderately similar to each other; (3) employers tend to prefer arbitrators with training in economics whereas unions prefer arbitrators with legal training and dislike economists, (4) both parties prefer arbitrators with greater experience, (5) there is evidence that the parties' preferences are affected by arbitrators' win-loss tallies under final-offer arbitration, and (6) there is no evidence that the parties rank arbitrators strategically.

Overall, the results suggest that New Jersey's prescreening procedure for establishing its master list of arbitrators works well. Even after controlling for the arbitrators' characteristics, there is much noise in the parties' preference orderings. Nevertheless, it still seems in the interest of the parties to have input into the appointment of an arbitrator since our results indicate that they have sensible preferences that they reveal accurately. Indeed, appointment mechanisms that account for the parties' preferences tend to result in higher welfare for both parties than mechanisms that appoint arbitrators on a rotating basis.

Of course, it has been argued that arbitration should be made costly so that bargainers have an incentive to settle their disputes voluntarily (see Stevens, 1966; Farber and Katz, 1979; Bloom, 1981). In this view, mechanisms that allow disputants to choose their arbitrator are undesirable because they

Table IV

Maximum Likelihood Estimates of Employer and Union Preference Functions, Multinomial Logit Model (MNL) and Rank/Veto Model (RV)*

Variable	Employer Preferences		Union Preferences	
	MNL	RV	MNL	RV
LAWYER	.386	.146	.088	.169
	(.226)	(.119)	(.193)	(.108)
ECONOMIST	.669	.355	-.574	-.814
	(.391)	(.218)	(.452)	(.229)
GRAPTS79	.103	.082	.066	.076
	(.025)	(.014)	(.022)	(.013)
NFOA	.090	.044	.012	.016
	(.037)	(.024)	(.034)	(.021)
FAVU	-.098	-.073	.166	.072
	(.057)	(.034)	(.058)	(.033)
Log-likelihood	-220.6	-405.6	-283.4	-472.2
$\chi^2[\beta^{CR}=\beta^{MNL};$ $\beta^{CR}=\beta^{RV}]$	4.07	36.35	3.01	3.57

*Estimated standard errors are reported in parentheses.

reduce the (indirect) costs of arbitration. However, this negative factor must be weighed against the increased legitimacy an arbitrator will have when the parties have been involved in his (or her) appointment.

Our empirical results seem to indicate that the New Jersey mechanism for selecting arbitrators works mainly as a safety net which allows the parties to filter out the least acceptable arbitrators who survived the prescreening. In other words, the extraordinary closeness of estimates derived from the complete rankings and the rank/veto model suggests that most of the information about the parties preferences comes from the vetoed arbitrators and not from the rank order of the non-vetoed arbitrators. Indeed, it may well be true that having seven arbitrators per panel is optimal given the degree of prescreening that takes place and the nature of the parties' preferences.

Finally, the similarity of union and employer preferences for different arbitrators suggests that collective bargaining functions more cooperatively than most existing models indicate. This finding represents potentially important input into the further development of models of employer-union interactions.

Footnotes

1. For some additional comparisons of arbitration and the court system, in which the advantages of arbitration are stressed, see the text of Chief Justice Burger's remarks to the American Arbitration Association and the Minnesota State Bar Association (Burger, 1985).

2. See, especially, the work of Crawford (1979, 1982) and Farber (1979, 1980). For a useful review of selected empirical analyses of some of the theoretical issues raised in these papers, see Ashenfelter (1985).

3. In many respects, the selection of an arbitrator is just one example of a general class of social choice problems in which two or more economic agents must collectively decide an intermediate or final outcome of some economic game. Voting for public officials, reaching committee decisions, choosing real estate appraisers (e.g., in cases of eminent domain), and determining the recipients of different honors and awards are all examples. But the examples which are most closely akin to the problem of selecting an arbitrator are the problems of judge and jury selection. In the case of judges, prescreening is substantial (e.g., all federal judges must be nominated by the President and confirmed by the Senate), although assignments are random except for the practice of "forum shopping" and of recusing in situations where there are conflicts of interest. On the other hand, prescreening is minimal in the process of jury selection since federal and state jury selection laws generally require that potential jurors be "selected at random from a fair cross-section of the community (P.L. 90-274, 82 Stat. 53). However, jury selection procedures offer opportunities to remove jurors both for cause and, although to a lesser extent, without cause. These procedures, known as voir dire, are modeled in Roth, Kadane, and DeGroot (1977).

4. Detailed derivations of all of the likelihood functions we use and their properties are presented in the Appendix to Bloom and Cavanagh, 1985a. This paper is available from the authors on request.

5. PERC revises its master list of eligible arbitrators every three years.

6. This unbalanced data configuration complicates the descriptive presentation of the data below. However, the econometric model developed in Section III is ideally suited to this type of problem and makes efficient use of all available information.

7. We relied particularly on material contained in MacDonald (1948) and Elkouri and Elkouri (1985), and on conversations with Ben Fischer (former head of the Arbitration Department of the United Steelworkers of America) and Richard Reilly (Regional Director of the American Arbitration Association, Boston Region).

8. The per diem that arbitrators charge is also a way in which they differ. Most ad hoc arbitrators, for example, presently charge between two hundred and six hundred dollars per day, plus expenses. Although little is known about the extent to which cost influences the selection of an arbitrator, it does drain the parties' funds and has been argued to be an important determinant of the use of arbitration in the case of financially small disputants (see Bloom, 1981). In the New Jersey system under study, PERC establishes a maximum per diem rate which is the rate charged by nearly all of the arbitrators on the master list. As a result, there is very little variation in arbitration fees across arbitrators. Thus, we do not include this variable in our empirical analysis.

9. This conclusion requires the assumption that each party is equally risk averse if negotiated outcomes are a possibility.

10. Recently, however, the high cost, long delays, and general shortage of

experienced arbitrators have caused some individuals to question the importance of experience. For example, a system known as expedited arbitration was adopted by labor and management in the basic steel industry in 1971. Under this system, unresolved employee grievances that do not require precedent-setting rulings are arbitrated by a rotating panel of young, inexperienced arbitrators (mostly lawyers) who decide the case for a relatively small fee within two weeks of the decision to arbitrate. Although there has yet to be an in-depth study of this system, its growing utilization in the basic steel industry, in other United Steelworker contracts, in some United Mineworker contracts, and in the U.S. Postal Service provides some evidence of its success.

11. In this section we ignore the possibility that a voluntary settlement can occur after the appointment of an arbitrator but prior to arbitration.

12. See, for example, McFadden (1982).

13. If all arbitrators on the panel were assigned a rank between one and seven by each party, ρ would always lie in the interval [-1.0, 1.0]. However, because only four arbitrators of the seven listed on each panel receive a rank, we were forced to assign a rank of 6.0 to the three vetoed arbitrators (i.e., the median of 5, 6, and 7). As a result, our estimates of ρ must lie in the interval [-.86, .93].

14. We have not worked out the distribution of ρ for the incomplete rankings case. Thus, we are unable to construct a formal test.

15. For example, the expected number of arbitrators vetoed by the union and ranked first by the employer is 32.1 (= 3/7 • 1/7 • 525).

16. Different arbitrators appeared on different numbers of panels because PERC generated the panels randomly and because a number of arbitrators requested that their names not be circulated actively throughout the entire year.

17. Final-offer arbitration is utilized to resolve about three-fourths of the bargaining disputes arising under the New Jersey statute. Under final-offer arbitration the arbitrator is constrained to render an award which consists of one or the other of the bargainers' final positions. Most of the remainder of the New Jersey cases are resolved by conventional arbitration in which the arbitrator renders a decision which consists of his or her best judgment of a fair settlement and which may be a compromise between the parties' final offers. For a more detailed description of the New Jersey statute see Bloom (1980). For an analysis of arbitrator decision-making under the different forms of arbitration in New Jersey see Ashenfelter and Bloom (1983, 1984).

18. Nevertheless, the arbitrator characteristics included in the model do have significant explanatory power when considered jointly, as judged by the difference between the maximized log likelihood and the log likelihood evaluated at $\beta = 0$.

19. In principle, each side has 840 distinct pure ranking strategies, of which at least several hundred are undominated. Clearly, solving for the Nash equilibrium strategies in such a game is computationally infeasible (see Bloom and Cavanagh, 1985b, for a description of equilibria in these and other related games).

20. It should be stressed that neither the MNL nor the RV model is presumed to be the correct structural model under strategic behavior. Both are, however, reasonable approximations to structural models that strategic behavior is likely to imply. For example, even when behaving strategically, bargainers are likely to give the top rank to one of their most preferred arbitrators. In this case, the MNL model would be a reasonable approximation to the underlying structural model. Alternatively, strategic behavior is unlikely to imply vetoing one's most preferred arbitrators. In this case, the RV model would be

a reasonable approximation.

21. In principle, the test also has power against a broad range of other specification errors relating to distributional assumption, functional form, and independence of the errors.

References

Ashenfelter, Orley. 1985. "Evidence on U.S. Experiences with Dispute Resolution Systems," Princeton University Industrial Relations Section Working Paper No. 185.

_____ and David E. Bloom. 1983. "The Pitfalls in Judging Arbitrator Impartiality by Win-Loss Tallies Under Final-Offer Arbitration." Labor Law Journal 34: 534-539.

_____. 1984. "Models of Arbitrator Behavior: Theory and Evidence." American Economic Review 74: 111-124.

Bloom, David E. 1980. "Customized 'Final-Offer': New Jersey's Arbitration Law." Monthly Labor Review 103: 30-33.

_____. 1981. "Is Arbitration Really Compatible with Bargaining?" Industrial Relations 20: 233-244.

_____ and Christopher L. Cavanagh. 1985a. "An Analysis of the Selection of Arbitrators," Harvard Institute of Economic Research Working Paper No. 1150.

_____. 1985b. "An Analysis of Alternative Mechanisms for Selecting Arbitrators." Unpublished paper.

Burger, Warren E. 1985. "Using Arbitration to Achieve Justice." Arbitration Journal 40: 3-6.

Crawford, Vincent P. 1979. "On Compulsory Arbitration Schemes." Journal of Political Economy 87: 131-160.

_____. 1982. "Compulsory Arbitration, Arbitral Risk and Negotiated Settlements: A Case Study in Bargaining Under Imperfect Information." Review of Economic Studies XLIX: 69-82.

Elkouri, Frank and Edna Asper Elkouri. 1985. How Arbitration Works. Washington, D.C.: Bureau of National Affairs, Inc.

Farber, Henry S. 1980. "An Analysis of Final-Offer Arbitration." Journal of Conflict Resolution 35: 683-705.

_____ and Harry C. Katz. 1979. "Interest Arbitration, Outcomes, and the Incentive to Bargain." Industrial and Labor Relations Review XXXIII: 55-63.

Goldfeld, Stephen M. and Richard E. Quandt. 1972. Nonlinear Methods in Econometrics. New York: North Holland, Inc.

Hausman, Jerry. 1978. "Specification Tests in Econometrics." Econometrica 46: 1251-71.

MacDonald, Lois. 1948. "Selection and Tenure of Arbitrators in Labor Disputes." Proceedings of First Annual Conference on Labor. Albany, New York: Matthew Bender and Company, 145-195.

McFadden, Daniel. 1982. "Econometric Models of Probabilistic Choice." In Charles Manski and Daniel McFadden, Editors, Structural Analysis of Discrete Data: With Econometric Applications. Cambridge, Mass.: M.I.T. Press.

Roth, Arthur, Joseph B. Kadane, and Morris H. DeGroot. 1977. "Optimal Peremptory Challenges in Trials by Juries: A Bilateral Sequential Process." Operations Research 25: 901-919.

Stevens, Carl M. 1966. "Is Compulsory Arbitration Compatible with Bargaining?" Industrial Relations 5: 38-52.

Appendix

1. The Econometric Model

In this appendix, we outline some of the main econometric results used in analyzing the rank-choice models of Section III.

Suppose each of T agents is faced with K options or alternatives. For each alternative i, each agent has a (random) utility which is a linear function of the observed attributes of that alternative and some unobserved attributes. Hence, each agent can compute the vector of utilities $(Y_{1t}, \ldots, Y_{Kt})$:

$$\begin{aligned} Y_{1t} &= X_{it}\beta_i + \varepsilon_{it} \\ &\vdots \\ Y_{Kt} &= X_{Kt}\beta_K + \varepsilon_{Kt} \end{aligned} \qquad t = 1, \ldots, T \tag{A.1}$$

where Y_{it} is the utility agent t attaches to alternative i; X_{it} is a vector of observed attributes associated with alternative i and agent t; $\beta = (\beta_1, \ldots, \beta_K)$ is a vector of unknown parameters characterizing the tastes of the agents; and ε_{it} represents the utility effects of unobserved attributes.

By treating the unobserved attribute terms -- ε_{it} i = 1, ..., K t = 1, ... T -- as random variables, the econometrician can view (A.1) as a random utility model which can potentially be used to draw inferences concerning the tastes $(\beta_1, \ldots, \beta_K)$ of the agents.

For our purposes, we suppose that the random variables ε_{it} have independent extreme value distributions -- that is,

$$\Pr[\varepsilon_{it} \leq x] = e^{-e^{-x}} \qquad \text{for } i = 1, \ldots, K \qquad t = 1, \ldots, T.$$

[In multinomial choice models, this distributional assumption gives rise to the multi-nomial logit model. See, for example, McFadden (1974). The extreme-

value distribution is also computationally convenient for the rank-choice models analyzed below.]

We consider the situation in which the econometrician observes, not the actual values of the utilities each agent attaches to each alternative, but the preference rankings of some of the alternatives by each of the agents. So, we suppose that for each t the econometrician observes a vector of integers $R_t(1), \ldots, R_t(K_t)$ where $R_t(1)$ indicates which of the alternatives $(1, \ldots, K)$ the t^{th} agent most prefers, $R_t(2)$ is the second most preferred alternative etc. and $K_t \in \{1, \ldots, K\}$ indicates how many alternatives we have rank information on for agent t. We will consider three different procedures for using such data to estimate the parameters $(\beta_1, \ldots, \beta_K)$. Before discussing the statistical methods, we introduce some additional notation. Let $b = (b_1, \ldots, b_K)$ be a list of vectors of the same dimensionality as $(\beta_1, \ldots, \beta_K)$,

$$S_t(i,b) = \exp(X_{it} b_i)$$

$$MNL_t(i_1, \mid i_1, \ldots, i_j, b) = \frac{S_t(i_1, b)}{\sum_{k=1}^{j} S_t(i_k, b)}$$

In the multi-nomial logit literature, $S_{it}(b)$ is referred to as the scale value of alternative i for agent t with parameters b and $MNL_t(i_1, \mid i_1, \ldots, i_j; b)$ is the multinomial logit probability that agent t with taste parameters b will choose alternative i, when given the option of choosing amongst alternatives $i_1, \ldots, i_j$.

Method 1. The simplest method for using the rank-choice data to estimate the parameters $\beta = (\beta_1, \ldots, \beta_K)$ is simply to ignore all rankings except the top ranking. The model then reduces to a multinomial logit model and inference can

be based on the log-likelihood function for the MNL model:

$$(A.2) \qquad L_1(b) = \sum_t \log MNL_t(R_t(1) \mid 1, \ldots, K; b).$$

Under suitable regularity conditions on the sequence of observed attributes, the b which maximizes $L_1(b)$ will be a consistent and aymptotically normal estimator of the true parameters β. However, this estimator will not, in general, be asymptotically efficient since it ignores information available in the additional rankings -- $R_t(2)$, etc.

Method 2. The second method fully utilizes the information available in the observed rankings. To calculate the likelihood function for the ranked data, we need to be able to calculate the probability that agent t with preference parameters b will have rank choice $R_t(1), \ldots, R_t(K_t)$. Thus, we need to calculate

$$P_t(R_t(1), \ldots, R_t(K_t); b) =$$

$$Pr_b[Y_{R(1)t} > Y_{R(2)t} > \ldots > Y_{R(K)t} > \text{all the remaining } Y_{it}].$$

This problem was addressed in Block and Marschak (1960) (see also Beggs, Cardell and Hausman (1981)). They show that

$$(A.3) \qquad P_t(R_t(1), \ldots, R_t(K_t); b) =$$

$$\prod_{i=1}^{K_t} MNL_t(R_t(i) \mid \{1, \ldots, K\} \setminus \{R_t(1), \ldots, R_t(i-1)\}; b),$$

where $\setminus$ denotes set difference. That is, the probability of interest is simply

the product of the MNL probabilities of choosing the i^{th} alternative given the option of those alternatives that have not already been chosen.

Using these calculations, we see that the log-likelihood function for the rank-choice data is given by

$$L_2(b) = \sum_{t=1}^{T} \sum_{i=1}^{K_t} \log MNL_t(R_t(i) \mid \{1, \ldots, K\} \setminus \{R_t(1), \ldots, R_t(i-1)\}; b).$$

Asymptotically efficient statistical procedures can be constructed using the maximum likelihood estimator based on this likelihood function. Inference based on $L_2(b)$ will, in general, be more efficient that inference based on $L_1(b)$. [Note also that L_2 is globally concave -- this guarantees a unique maximum if the model is asymptotically identified and typically assures numerical stability in calculating the MLE.]

Method 3. The idea underlying Method 1 is that a likelihood function which uses only a part of the available data will yield inefficient but consistent estimators. In general, any informational garbling will yield consistent but inefficient limited information maximum likelihood estimators, so long as the limited information model is identified. Another limited information model of interest in this paper is one in which the exact rankings of each agent are not recorded -- what is recorded is the distinction between which alternatives are among the agent's K_t most desirable options and which alternatives are among the $K-K_t$ least desirable. Of course, in the special case in which $K_t = 1$ for all t, this model is identical with the MNL model and Method 3 simplifies to Method 1. In general, this method bases inference on the log-likelihood

function

$$L_3(b) = \sum_{t=1}^{T} \log\Big[\sum_{\sigma} P_t(\sigma(R_t(1)), \ldots, \sigma(R_t(K_t)); b)\Big]$$

where the inner summation is over all permutations σ of the integers $\{R_t(1), \ldots, R_t(K_t)\}$ and the funtion $P_t(\cdot)$ is defined by (A.3). Hence, the summation over σ equals the probability that $R_t(1), \ldots R_t(K_t)$ are the top K_t choices of agent t in any possible order.

Each of the three methods yields an estimator which is consistent and asymptotically normal under fairly general regularity conditions. However, because methods 1 and 3 use only some of the information in the data they are inefficient but they will yield consistent estimates of β under some circumstances in which the data are not truly generated by the model. For example, if, for strategic reasons, agents rankings correctly reflect their top preferences but not this second, third, etc., then Method 1 will yield consistent estimates of β but Methods 2 and 3 will give inconsistent estimates. Similarly, if agents reorder their most preferred alteratives, then Method 3 will yield consistent estimates while 1 and 2 will not.

In general, if the ranking data does not reflect the true preference rankings of the agents then the three estimators will have different probability limits. These observations suggest that specification tests based on the differences between the Method 1 and 3 estimators and the Method 2 estimator will yield tests with power against a broad range of alternatives for which the ranking data does not reflect preference rankings.