AN EXPERIMENTAL COMPARISON OF DISPUTE RATES
IN ALTERNATIVE ARBITRATION SYSTEMS

Orley Ashenfelter
Janet Currie
Henry S. Farber
Matthew Spiegel

AN EXPERIMENTAL COMPARISON OF DISPUTE RATES
IN ALTERNATIVE ARBITRATION SYSTEMS

ABSTRACT

This paper reports the results of a systematic experimental comparison
of the effect of alternative arbitration systems on dispute rates. The key
to our experimental design is the use of a common underlying distribution of
arbitrator "fair" awards in the different arbitration systems. This allows
us to compare dispute rates across different arbitration procedures where we
hold fixed the amount of objective underlying uncertainty about the
arbitration awards.
     There are three main findings. First, dispute rates are inversely
related to the monetary costs of disputes. Dispute rates were much lower in
cases where arbitration was not available so that the entire pie was lost in
the event of dispute. Second, contrary to conventional wisdom, the dispute
rate in a final-offer arbitration system is at least as high as the dispute
rate in comparable conventional arbitration system. Third, dispute rates are
inversely related to the uncertainty costs of disputes. Dispute rates were
lower in conventional arbitration treatments where the variance of the
arbitration award was higher and imposed greater costs on risk-averse
negotiators. Our results can also be interpreted as providing tentative
evidence that the negotiators were risk-averse on average.

Orley Ashenfelter
Department of Economics
Princeton University
Princeton, NJ   08544

Janet Currie
Department of Economics
University of California
Los Angeles, CA 90024

Henry S. Faber
Department of Economics, E52-252f
Massachusetts Institute of Technology
Cambridge, MA 02139

Matthew Spiegel
414 Uris Hall
Columbia University
New York, NY 10027

I. Introduction

This paper reports the results of a systematic experimental comparison of the effect of alternative arbitration systems on dispute rates. Three types of arbitration are considered. The first is conventional arbitration (CA), where the arbitrator is free to impose any settlement he or she prefers. The second is final-offer arbitration (FOA), where the arbitrator is constrained to impose one or the other of the specially-formulated last offers of the parties without any ability to compromise. The third is tri-offer arbitration (TRI), where the arbitrator is constrained to impose one of three outcomes: one or the other of the last offers of the parties or the recommendation of a neutral fact-finder.[1]

These mechanisms are commonly used to settle disputes that arise in the negotiation of labor agreements in the public sector (Lester, 1984). Conventional arbitration is also used to settle many commercial disputes, and it has been experimented with as an alternative to trials in civil litigation in a number of jurisdictions. Perhaps the best-known application of FOA is in the determination of the wage in disputed contracts for major league baseball players, but the earliest recorded use of FOA that we are aware of is in ancient Athens in the trial of Socrates. I. F. Stone (1988, pp. 186-189) describes how, in the penalty phase of the trial, the jury (of 500!) was not permitted to compromise between the punishment suggested by the prosecution (death) and the punishment suggested by the defense (a fine).[2] Tri-offer arbitration is used to settle labor disputes in the public sector

---

[1] The fact-finder is essentially an arbitrator who makes a recommended outcome at an earlier stage in the negotiation process.

[2] Stone argues that had Socrates wanted to avoid execution he would have proposed a stiffer penalty for himself to give the jury a reasonable alternative.

in Iowa.

Since arbitration systems are so precisely structured and since the
same class of individuals act as arbitrators in the different schemes, it is
natural to examine dispute rates in a controlled laboratory environment. The
stumbling block to this approach in the past has been the difficulty in
handling the inherent three-party nature of arbitration systems.

A primary innovation in this paper is to simulate arbitrator behavior
by utilizing the emerging results from field studies (e.g., Ashenfelter and
Bloom, 1984; Ashenfelter, 1987) which indicate that acceptable arbitrator
decisions contain a random component. The key idea is that since the parties
play a role in the selection of the arbitrator who will decide their dispute,
arbitrators who are known to favor one or the other of the parties are
eliminated. This selection process creates incentives for arbitrators to
maintain characteristics that make them (statistically) exchangeable with
other arbitrators.[3] While different arbitrators will make different awards
in the same case, exchangeability requires that there be no predictable
differences. Analytically this reduces a three-party problem to a two-party
problem because actual negotiations are broken into two parts: (a) the
parties mutually select the arbitrator, and (b) the parties present their
cases to the arbitrator selected.

Statistical exchangeability of arbitrators implies that the arbitration
decisions may be modeled as being based on random draws from a fixed
distribution, and this is how we simulated arbitrator behavior in the
bargaining experiments reported below.[4] The result is that an apparently

---

[3]Bloom and Cavanagh (1986) analyze data on the selection process of
arbitrators for New Jersey public sector labor disputes. Their findings are
consistent with this argument.

[4]In CA the arbitration award was simply the draw from a distribution. In FOA

complex bargaining problem may be simplified and easily implemented in the laboratory. Bargainers are first given a list of previous arbitration awards in a form consistent with what they would see in the field for their particular type of arbitration. The bargainers then either negotiate settlements or arrive at an impasse, and impasses are resolved by resorting to an arbitrator in the context of a particular arbitration system. The key is that both the information on previous awards and the arbitrators' decisions in the different arbitration systems are based on draws from a common underlying distribution. In this sense the arbitrator's behavior is entirely controlled when the outcomes under the various arbitration systems are compared.

Our approach contrasts with previous experimental research (Notz and Starke, 1978; Grigsby and Bigoness, 1982; Neale and Bazerman, 1983) where the arbitrator's behavior is not precisely controlled. In this earlier research the subjects are told simply that if no agreement is reached an arbitrator will be appointed to make a binding decision. In CA this decision is stated to be anything the arbitrator wants to do. In FOA this decision is stated to be one or the other of the final offers. No further information is given to the negotiators about the likelihood of various arbitration awards.

This earlier research ignores the facts 1) that negotiators in most field settings obtain substantial information about likely arbitration awards by observing past awards and 2) that much of this information has a common basis across different forms of arbitration. Thus, experimental subjects in these studies have no basis for generating expectations about arbitrator

---

the award was the offer closest to the draw from the same distribution. In TRI, a fact-finder's recommendation was drawn from the distributions, offers were formulated, and the award was the option closest to a new draw from the same distribution. This view is elaborated in the next section.

behavior that are consistent with arbitrator behavior in the field.
Moreover, subjects expectations may be influenced by the form of arbitration,
in which case there can be no meaningful comparisons of bargaining outcomes
across different arbitration schemes from these earlier experiments.

In the next section of the paper we present a unified framework for
characterizing arbitrator behavior and show how it may be applied to
conventional, final-offer, and tri-offer arbitration. The design of the
experiments is described in section III, and section IV contains the central
results on dispute rates. The implications of our results for a simple model
of negotiator behavior are presented in section V, and section VI contains an
analysis of the experimental evidence on offers and negotiated agreements.
Section VII concludes.

II. A Unified Framework for Arbitrator Behavior

Under one simple model of arbitrator behavior, the arbitrator examines
the facts of the particular case and determines a fair or acceptable award
independent of the last offers of the negotiators (Farber and Katz, 1979).
However, the parties are unable to forecast perfectly what a particular
arbitrator will think is appropriate in any specific case. More formally,
suppose that two parties (A and B) are bargaining over the determination of
some variable Y. Party A prefers greater values of Y while party B prefers
smaller values of Y. Let the arbitrator's notion of the appropriate value be
Y_f. The parties' uncertainty about Y_f is summarized by the probability
density function $g(Y_f)$.

What differs across arbitration mechanisms is how Y_f is translated into
an award. In conventional arbitration, the arbitrator simply imposes Y_f so
that

$$(\text{II.1}) \qquad Y = Y_f.$$

In final-offer arbitration, because the arbitrator is not free to impose Y_f, the arbitrator selects the final offer closest to Y_f (Farber, 1980). Given that party A offers Y_A and party B offers Y_B, the arbitrator selects party B's offer if and only if

$$(II.2) \qquad |Y_f - Y_B| < |Y_A - Y_f|.$$

Since party A prefers greater values of Y while party B prefers smaller values, $Y_A > Y_B$ in disputed cases. Thus, the probability that B's offer is selected (P_B) is the probability that Y_f is less than the average of the final offers. This is

$$(II.3) \qquad P_B = G[(Y_A + Y_B)/2]$$

where $G[Y_f]$ is the cumulative distribution function of Y_f.

The analysis of tri-offer arbitration is more complicated but is analogous to the analysis of final-offer arbitration. The arbitrator now has a choice from among three alternatives: the final offers of the parties and the recommendation of the fact-finder (Y_r). Consistent with the model of final-offer arbitration, the arbitrator chooses the alternative that is closest to his/her notion of a fair award. Since the fact-finder's report is known prior to the parties making their final offers, the final offers are conditional on the particular recommendation made by the fact-finder. Three possible configurations of the offers and Y_r are possible, and the choice probabilities of the arbitrator depend on the particular configuration.

Where Y_r is some intermediate value (defined more precisely below), the final offers will bracket the fact-finder's recommendation so that $Y_B < Y_r < Y_A$. This is the usual configuration, and we call this case #1.[5] The arbitrator's choice probabilities in this case are

_____

[5]We will see that case #1 is predicted to occur in the large majority of cases.

$$P_A = 1 - G[(Y_r + Y_A)/2]$$

(II.4) $$P_B = G[(Y_r + Y_B)/2]$$

$$P_r = G[(Y_r + Y_A)/2] - G[(Y_r + Y_B)/2]$$

where P_A, P_B, and P_r are the probabilities that the arbitrator selects A's offer, B's offer, and the fact-finder's recommendation respectively. If the fact-finder's recommendation is unusually high, party A will prefer to make an offer that is lower than Y_r so that $Y_B < Y_A < Y_r$. We call this case #2. The arbitrator's choice probabilities in this case are

$$P_A = G[(Y_r + Y_A)/2] - G[(Y_A + Y_B)/2]$$

(II.5) $$P_B = G[(Y_A + Y_B)/2]$$

$$P_r = 1 - G[(Y_r + Y_A)/2].$$

Finally, if the fact-finder's recommendation is unusually low, party B will prefer to make an offer that is higher than Y_r so that $Y_r < Y_B < Y_A$. We call this case #3. The arbitrator's choice probabilities in this case are

$$P_A = 1 - G[(Y_A + Y_B)/2]$$

(II.6) $$P_B = G[(Y_A + Y_B)/2] - G[(Y_r + Y_B)/2]$$

$$P_r = G[(Y_r + Y_B)/2].$$

This analysis shows how arbitration awards in CA, FOA, and TRI rely on a common underlying construct: Y_f. The distribution of Y_f (G) is invariant to the type of arbitration scheme, and we build this common distribution into our experiments.[6] This allows us to focus on the effects of the structural differences across the different schemes without the confounding factor of different arbitrator behavior across schemes. Whenever an arbitration award is required in our experiment, a random draw from the distribution of Y_f is

---

[6]Ashenfelter and Bloom (1984) and Farber and Bazerman (1986) present evidence from the field and the laboratory respectively that is consistent with the view that arbitrators' fair awards are drawn from the same distribution in CA and FOA.

made. This value is then used to determine the actual arbitration award. In CA, the random draw on Y_f becomes the award. In FOA, the parties formulate final offers, and the offer closest to the random draw on Y_f is selected. In TRI, a fact-finder's recommendation is drawn from the same distribution as Y_f, the parties formulate final offers, and the option closest to the random draw on Y_f is selected.

III. Design of the Experiments

Our experiments consisted of a series of repeated pie-splitting games. Each subject negotiated with the same opponent for twenty rounds. Ten pairs (the control group) bargained without an arbitration system for their disputes throughout all twenty rounds. One-hundred-thirty-one subject pairs first bargained ten rounds without an arbitration system and then ten rounds with an arbitration system. Seventy-nine of these pairs bargained with conventional arbitration, twenty-six pairs bargained with final-offer arbitration, and twenty-six pairs bargained with tri-offer arbitration.[7] The seventy-nine pairs assigned to the CA treatment were split into three roughly equal groups, each of which had its arbitration awards drawn from a distribution with a different variance.

The initial bargaining experiments were conducted in the summer of 1984 with five pairs of bargainers in each treatment. Further experiments were conducted in the winter of 1988 with approximately twenty (randomly assigned) pairs in each treatment.[8] The experiments used Plato software at the

---

[7]Two additional pairs bargained in the tri-offer condition but did not complete all twenty rounds due to a computer malfunction. They are not included in the analysis.

[8]Only five additional pairs were added to the control group.

University of Arizona, and subjects were recruited from students at the
University, a subject pool that has been used extensively for laboratory
experiments in economics. Upon arrival at the laboratory, a subject was
seated at a computer terminal and given instructions containing the basic
information necessary to send, receive, and accept offers. Subjects did not
know the identity of their opponents, who were seated at different computer
terminals some distance away. They did know that they would be bargaining
with the same opponent for all twenty rounds. Precautions were built into
the software to ensure that subjects could not accept offers by accident.

The bargaining protocol in each round was deliberately left
unstructured, as is naturally the case in the field. Each party's last offer
was posted on their screen and on their opponent's screen at all times. An
offer consisted of a number between 100 and 500. The subjects were given a
schedule revealing the cash value of various settlements to them, but they
were not given the schedule of their opponent. In fact, their opponent's
schedule was identical except that odd numbered parties desired high
outcomes, while even numbered parties desired low outcomes. These schedules
had a linear relationship between the settlement and the cash values. An
agreement at 300 split the pie, which contained $1.20, in half. The least a
subject could earn was $.15 per round, and the most $1.05. All subjects were
also given $3.00 simply for participating in the experiment. Thus, if the
parties agreed to split the pie in all twenty rounds, each would earn fifteen
dollars ($3.00 + 20($.60) for a one to one-and-one-half hour experiment. We
reckoned these stakes to be more than double the hourly wage available to our
subjects.

Each bargaining round was limited to five-and-one-half minutes. How
disputes were handled varied by treatment group. While players always knew
how much time was left in the round and what round they were in, they were

not told the total number of rounds.

A. The Control Group

The control group pairs bargained in the absence of an arbitration system throughout the experiment. Their payoffs were zero in any round in which they did not agree. Compared to the arbitration treatments, the control groups paid a high cost for disputing.

B. Conventional Arbitration

Conventional arbitration pairs were told before starting that an arbitration scheme would be introduced after the first ten rounds and that their actions would have no bearing on the arbitrator's decisions. They were only given information about the arbitration scheme after round ten. In the rounds without arbitration, failure to reach a settlement resulted in a zero pay-off, just as for the control group.

After the first ten rounds the pairs were told that an arbitrator would dictate their payoffs if they failed to reach agreement. They were not given any specific information about how the arbitrator would decide, but they were given a list of what they were told were the arbitrator's last one-hundred decisions. These awards as well as the actual awards where the pairs failed to reach agreement were generated randomly from a normal distribution with mean 350 and the standard deviation (12.5, 25, or 50) appropriate for their treatment.[9] The pairs were not told anything specific about the distribution. The presentation of information about the arbitrator's behavior as a list of past awards was designed to be similar to the kind of information about arbitrator behavior that is available in field settings. The subjects were free to study the new information for as long as they

---

[9]Note that the arbitrator was biased in favor of the odd-numbered party because the mean of the distribution was greater than the midpoint of the bargaining range (300). This bias was intended to move the bargaining pairs away from mechanical 50-50 divisions of the pie.

wished. They were instructed to press a key when they were ready to
continue. When the negotiators failed to agree, the negotiators were
informed of the arbitration decision immediately.

Different CA treatments were selected to provide evidence of the
relationship between the variance of the arbitration award and the dispute
rate. The *only* feature of the experiment that differed across the three CA
treatments was the variance implicit in both the list of one hundred awards
each pair was given and the actual awards when the parties failed to agree.
Any systematic difference in dispute rates across the CA treatments can then
be attributed to differences in the variance and may provide evidence that
the negotiators are not risk neutral.

B. Final-Offer Arbitration

The FOA pairs were also told before starting that an arbitration scheme
would be introduced after the first ten rounds, and that their actions would
have no bearing on the arbitrator's decisions. They were only given
information about the arbitration scheme after round ten. In the rounds
without arbitration, failure to reach a settlement resulted in a zero
pay-off, just as for the control groups and the CA treatments.

After the first ten rounds the pairs were told that henceforth an
arbitrator would dictate a division if they failed to agree. The negotiators
were told that the arbitrator would choose one of their final offers, but
they were not told specifically how the arbitrator would make this choice.
Instead, they were given a list of one-hundred pairs of final offers and the
decisions of eighty arbitrators (the offer selected) for each of these
one-hundred pairs. The pairs of final offers were generated from a uniform
distribution over the 100-500 range. The arbitrators' decisions were
generated by drawing a random number from a normal distribution with mean 350
and standard deviation 50 and then choosing the final offer closest to this

draw.   Information for each of 100 pairs of final offers was given to the
subjects in the following format:

```
offers:          selections:
  250    x x    xxxx x                 x   x . . . .
  400    x x xxxx    x xxxxxxxxxxxxxxx xxx   . . . .
```

where the x indicates the offer selected. In this particular example, most
arbitrators selected the higher final offer (400 = one standard deviation
above the mean) rather than the lower final offer (250 = two standard
deviations below the mean).

The pairs were not shown the draws from the normal distribution that
were the basis for the selections, and the negotiators were left to infer the
arbitrator's decision rule from the lists of offers and choices. Once again,
this form of presentation of information was designed to be similar to the
kind of information about arbitrator behavior that is available in field
settings. As in the CA treatments, the subjects were free to study the new
information for as long as they wished. They were instructed to press a key
when they were ready to continue.

When the negotiators failed to agree after five and one-half minutes,
each side submitted a final offer, and the negotiators were informed of the
arbitration decision immediately. The offer selected was the one closest to
a draw from a normal distribution with mean 350 and standard deviation fifty.

B. Tri-Offer Arbitration

The TRI pairs were also told before starting that an arbitration scheme
would be introduced after the first ten rounds, and that their actions would
have no bearing on the arbitrator's decisions. They were only given
information about the arbitration scheme after round ten. In the rounds
without arbitration, failure to reach a settlement resulted in a zero
pay-off, just as for the control group and the other treatments.

After the first ten rounds the pairs were told that henceforth an arbitrator would dictate a division if they failed to agree. The negotiators were told that the arbitrator would choose one of the final offers or a fact-finder's recommendation, but they were not told specifically either how the fact-finder's recommendation would be generated or how the arbitrator would make the choice. Instead, they were given two charts. The first contained a list of one-hundred earlier fact-finder's recommendations. This was presented in the same format as the one-hundred earlier arbitration decisions were presented in the conventional arbitration treatment. The second chart contained one-hundred final offer triples (pairs of final offers and a fact-finder's recommendation) and the decisions of eighty arbitrators (the option selected) for each of these one-hundred triples.

The fact-finder's recommendations were generated from the same normal distribution (mean=350, standard deviation=50) that was used to generate arbitration awards. The pairs of final offers were generated from a uniform distribution over the 100-500 range. The arbitrators' decisions were generated by drawing a random number from the normal distribution with mean 350 and standard deviation 50 and then choosing the option closest to this draw.

Information for each of 100 pairs of final offers was given to the subject in the following format:

```
offers:          selections:
 250    x x    x x  x              x   x  . . . .
 325      x x xx    x xx  xxxxx x xx x x    . . . .
 425    x   x   x x    xx      x x    x  x . . . .
```

where the x indicates the offer selected. In this particular example, most arbitrators selected the intermediate option (325 = one-half standard deviation below the mean) rather than the lower option (250 = two standard

deviations below the mean) or the higher option (425 = one and one-half standard deviations above the mean.

The pairs were not shown the draws from the normal distribution that were the basis for the selections, and the negotiators were left to infer the arbitrator's decision rule from the lists of options and choices. Once again, this form of presentation of information was designed to be similar to the kind of information about arbitrator behavior that is available in field settings. As in the other treatments, the subjects were free to study the new information for as long as they wished. They were instructed to press a key when they were ready to continue.

When the negotiators failed to agree after 2.75 minutes, they were presented with a fact-finder's recommendation. If they had not reached agreement after five and one-half minutes, each side submitted a final offer, and the negotiators were informed of the arbitration decision immediately. The option selected was the one closest to a draw from a normal distribution with mean 350 and standard deviation fifty.

The key to this design is the use of a common underlying distribution of arbitrator "fair" awards in the high-variance CA treatment and the FOA and TRI treatments. In all three cases this distribution was normal with mean 350 and standard deviation 50. Other than the arbitration scheme (how the draw from the common normal distribution was translated into an award) and the style of information presented to the pairs about previous awards, the treatments were identical.[10] This allows us to compare dispute rates across

---

[10]It is possible that the style of presentation matters. For example, the FOA and TRI treatments had an arbitration scheme that was potentially more interactive and interesting from the subjects point of view. If the subjects were intrigued by this, they might have had disputes simply to "play the game." This is not explored further here.

different arbitration procedures where we hold fixed the amount of objective underlying uncertainty about the arbitration awards.

IV. Dispute Rates

The central results of the experiment with regard to dispute rates are contained in table 1. Ignoring the control group, every treatment had substantially lower dispute rates in the first ten rounds, where the entire pie was lost in the event of disagreement, than in the last ten rounds, where the pie was divided by an arbitrator in the event of disagreement. Chi-squared tests of equality of the dispute rates across early (1-10) and late (11-20) rounds carried out within treatment groups reject equality at any reasonable level of significance (p-value < 10^{-7}) in all five cases.

Figure 1 presents the round-by-round dispute rates for the five arbitration treatments. This illustrates graphically the lower dispute rates in the first ten rounds for these treatments. In addition, it shows no clear patterns of dispute rates either within the first ten rounds or within the last ten rounds. Contingency table analysis of the relationship between dispute rate and round shows no significant relationship for any treatment within the first ten rounds or within the second ten rounds.[11]

The dispute rates for the five arbitration treatments in the first ten rounds were not significantly different from each other (p-value of chi-squared statistic = .398). The average dispute rate was 11.8 percent. The equality of dispute rates across the different arbitration treatments in the early rounds suggests that comparisons of dispute rates across treatments in the last ten rounds will provide meaningful information about how the different arbitration schemes affect dispute rates.

---

[11]No p-value for any of these tests is less than .45.

Table 1

Dispute Rates by Treatment

(cell size)

TREATMENT	# pairs	Rounds 1-10	Rounds 11-20
Control no arbitration	10	.04 (100)	.04 (100)
Low-Variance Conventional Arbitration	29	.138 (290)	.414 (290)
Medium-Variance Conventional Arbitration	25	.104 (250)	.316 (250)
High-Variance Conventional Arbitration	25	.096 (250)	.284 (250)
Final-Offer Arbitration	26	.108 (260)	.381 (260)
Tri-Offer Arbitration	26	.139 (260)	.323 (260)
All Treatments	141	.112 (1410)	.340 (1410)

Figure 1
DISPUTE RATES BY ROUND
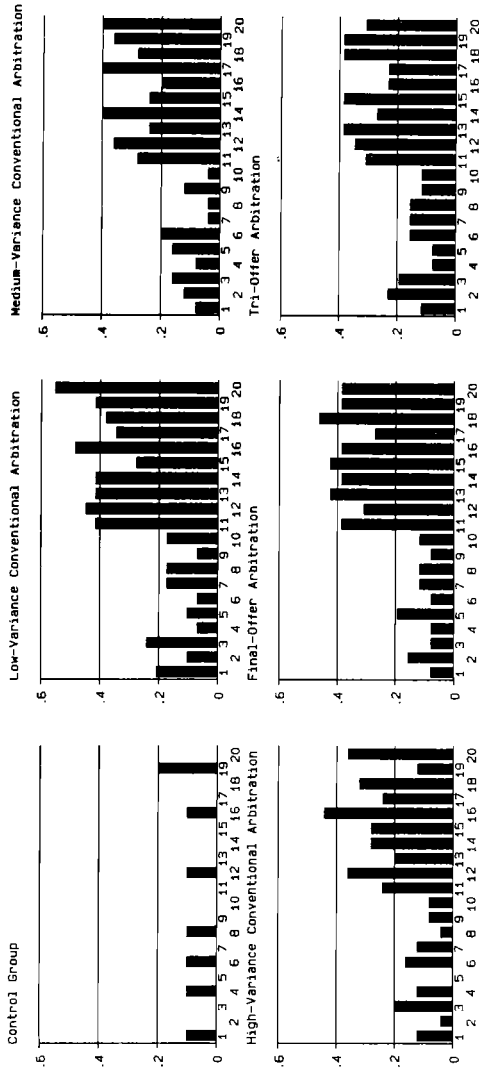
The control group had a dispute rate in the early rounds of four percent, and this was significantly lower than that of the arbitration treatments taken together in the early rounds (p-value of chi-squared statistic = .018). This was not expected, and it may be that the behavior of the arbitration treatments was affected by the knowledge that there would be arbitration after ten rounds. Another potential explanation is that the control group happens (by chance) to be composed of individuals who are less dispute prone than those in the arbitration treatments. These considerations suggest that the control group may be less useful than it would otherwise be. As a result we will make use of the early rounds for the arbitration treatments to provide a baseline dispute rate.

While we intend to be agnostic about a particular model of disputes, a common implication of virtually all theories of disputes is that the probability of disagreement is inversely related to the cost of disagreement. Our finding of higher dispute rates in the late rounds is consistent with that fact that arbitration is a less costly dispute settlement mechanism than destruction of the pie (which is what happened in the early rounds). An alternative explanation for our finding is that the dynamics of the experiment "naturally" led to higher dispute rates in the later rounds. This alternative is inconsistent with our finding that the dispute rate of the control group was exactly the same in the late rounds as in the early rounds. Thus, there is nothing inherent in the later rounds, independent of the introduction of arbitration, that caused the increase in dispute rates in the arbitration treatments.

Our experimental results are consistent with field results in comparable settings. The evidence from the field on dispute rates in collective bargaining by public sector employees where arbitration is used to settle disputes is in line with our experimental dispute rates of 28 to 41

percent in the last ten rounds for the arbitration treatments.[12] Evidence
from the field on dispute rates in collective bargaining where strikes are
used to settle disputes show strike rates of around 15 percent.[13] This is
slightly greater than our experimental dispute rates of 10 percent in the
first ten rounds for the arbitration treatments and 4 percent for the control
group.

A. The Role of Variance: Comparison of the CA Treatments

The difference in relative dispute rates in the last ten rounds among
the three CA treatments are clear on inspection of figure 1. The
high-variance (HVCA) treatment had the lowest dispute rate (28.4%) while the
low-variance (LVCA) treatment had the highest dispute rate (41.4%). The
medium-variance (MVCA) treatment was in between (31.6%). A chi-squared test
rejects equality of these dispute rates at conventional levels (p-value =
.004). This rejection is entirely due to the relatively high dispute
rate in the LVCA treatment. There was no significant difference between the
dispute rates for the MVCA and HVCA treatments (p-value = .435).

Given the repeated-measures design of our experiment, it is natural to
ask whether individual heterogeneity in dispute proneness plays any role in
our findings. Table 2 contains results of some analyses designed to shed
some light on this issue. The first two columns contain OLS and probit
analyses respectively of dispute probabilities with dummy variables for the
CA treatments as regressors. The base group is the low-variance treatment,
and these analyses mimic the simple contingency table analysis described

---

[12]Field results are reported by Lester (1984), Ashenfelter and Bloom (1984),
Babcock (1988), Ashenfelter, Dow, and Gallagher (1985), Currie (1989), and
Currie and McConnell (1989).

[13]Field results are reported by Tracy (1986), McConnell (1989), and Currie and
McConnell (1989).

Table 2

Analyses of Dispute Probabilities in

Conventional Arbitration Treatments

Rounds 11 - 20

coefficient estimate
(standard error)
[average derivative of probit]

VARIABLE	(1) OLS	(2) PROBIT	(3) OLS	(4) PROBIT	(5) RE PROBIT
Constant	.414 (.0277)	−.218 (.0742)	.264 (.032)	−.639 (.0939)	−.343 (.237)
Medium Variance	−.0978 (.0407)	−.261 (.111) [−.0976]	−.0609 (.0393)	−.179 (.114) [−.0620]	−.395 (.341) [−.126]
High Variance	−.130 (.0407)	−.353 (.112) [−.132]	−.0842 (.0394)	−.236 (.117) [−.0820]	−.530 (.338) [−.170]
Average Dispute Rate, Rounds 1-10 (for pair)	---	---	1.09 (.130)	3.00 (.395) [1.04]	---
Log Likelihood		−501.8		−470.5	−411.1
R^2	.014		.094		
n=790					

Note: The numbers in parentheses are standard errors. The numbers in square
brackets are the average derivatives of the probit function with respect to
each variable. The derivative for variable j is the average over sample
observations i of $\hat{\beta}_j \phi(X_i \hat{\beta})$. The base group is the low-variance treatment.
The mean of the average dispute rate in the first ten rounds is .114
(sd=.125). The estimated variance of the individual effects in the random
effects probit model is 1.168 (se = .324) The log likelihood for the
constrained model with only a constant is −507.3.

above. Columns 3 and 4 of the table contain OLS and probit analyses
respectively that include the average dispute rate in the first ten rounds
for each pair as an additional regressor. This is a measure of how dispute
prone each pair is, and, if there is persistent heterogeneity, this variable
will have a positive coefficient in the analysis of dispute probabilities in
the later rounds. In both the OLS and the probit analyses, there is a large
and significant effect of the early dispute rate on the probability of a
dispute in the later rounds. The dispute rate in the high-variance treatment
remains significantly larger than the dispute rate in the low-variance
treatment.

Another approach is to specify a probit model of dispute probabilities
that includes an error-components structure. Consider a model where there is
a dispute in round j for pair i if

$$(IV.1) \qquad Y_{ij} = X_i\beta + \theta_i + \varepsilon_{ij} > 0$$

where θ_i is a normally distributed individual-specific fixed factor with zero
mean and variance σ_θ^2, X_i is a vector of treatment dummies, β is the vector of
treatment effects to be estimated, and ε_{ij} is a $N(0,1)$ random factor that is
independently distributed across all i and j. The probit model in column 2
of table 2 is the special case of this model where θ_i is zero for all
pairs $(\sigma_\theta^2 = 0)$.

The estimates in column 5 of table 2 are for an unrestricted random-
effects probit model where the parameters to be estimated include the σ_θ^2 as
well as the vector of treatment effects (β). The contribution to the
log-likelihood function for the ith pair is

$$(IV.2) \qquad L_i = \ln\left[\int \left([\Phi(X_i\beta+\theta_i)] \right)^{D_i} \cdot \left([1-\Phi(X_i\beta+\theta_i)] \right)^{10-D_i} \cdot f(\theta)d\theta \right]$$

where $f(\theta)$ is the density function of the fixed factor (in our case $N(0,\sigma_\theta^2)$)
and D_i is the number of cases out of ten where pair i had a dispute. This
likelihood function is based on the likelihood of observing a particular

number of disputes for pair i conditional on the pair's value of θ_i. We then integrate over the distribution of θ to derive the unconditional likelihood.[14]

The estimates of the random-effects probit model in column 5 are somewhat less precise than the estimates of the simple probit in column 2. The estimated high-variance treatment effect is significantly negative at the .06 level using a single-tailed test suggesting that dispute rates are lower in the high-variance treatment even after accounting for the correlation of outcomes within pairs. The estimates also suggest, like the estimates in columns 3 and 4, that there is substantial heterogeneity across pairs in dispute proneness. The estimate of the variance of the fixed effect (θ) is 1.168 (se = .324) which is about the same size as the variance of the random component (ε) for each case (σ_ε^2 is normalized to one in the probit). These findings are, of course, conditional on the untested distributional assumptions.

The only difference among the three CA treatments is that the variance in the list of earlier arbitration decisions given to the subjects varied systematically across treatment groups. This variance sould only affect dispute rates only if the parties are not risk neutral. The direction of the results we have obtained, however, suggests that the parties are, on average, risk averse. Because uncertainty imposes a cost on risk-averse bargainers, there will be more disputes where the variance is lower, and this is what we find. We can interpret variation in θ across pairs in the error components model as variation in attitudes toward risk, with high values of θ representing less risk aversion. These ideas are discussed more formally in the next section.

---

[14]Note that computationally this requires numerical integration in a single dimension rather than the evaluation of a ten-dimensional normal probability.

B. <u>The Role of Mechanisms: Comparison of the Three High Variance Treatments</u>

The dispute rates in the HVCA, FOA, and TRI treatments in the last ten rounds are presented in table 1 and figure 1. The FOA treatment had a dispute rate of 38.1% while the HVCA treatment had a dispute rate of only 28.4% in the last ten rounds. The TRI treatment had a dispute rate of 32.4% in the last ten rounds. A chi-squared test of equality of these dispute rates marginally rejects equality at conventional levels (p-value = .064). However, we more convincingly reject the hypothesis that the HVCA and the FOA treatments had equal dispute rates (p-value=.02).

Table 3 contains results for the comparison of the HVCA, FOA, and TRI treatments in order to determine how sensitive our findings are to inter-pair heterogeneity in dispute proneness. The first two columns contain OLS and probit analyses respectively of dispute probabilities with dummy variables for the FOA and TRI treatments as regressors. The base group is the HVCA treatment, and these analyses mimic the simple contingency table analysis described above. Columns 3 and 4 contain OLS and probit analyses respectively that include the average dispute rate in the first ten rounds for each pair as an additional regressor. As in table 2, in both the OLS and the probit analyses, there is a large and significant effect of the early dispute rate on the probability of a dispute in the later rounds. The dispute rate in the FOA treatment remains significantly larger than the dispute rate in the HVCA treatment.

Estimates of the random effects probit model derived above are contained in column 5 of table 3 and are substantially less precise than the estimates of the simple probit model in column 2. The hypothesis that the dispute rate in the FOA treatment effect equals the dispute rate in the HVCA treatment cannot be rejected (p-value=.38). The estimates again suggest that there is substantial heterogeneity across pairs in dispute proneness. The

Table 3

Analysis of Dispute Probabilities in

High-Variance Treatments

Rounds 11 - 20

coefficient estimate
(standard error)
[average derivative of probit]

VARIABLE	(1) OLS	(2) PROBIT	(3) OLS	(4) PROBIT	(5) RE PROBIT
Constant	.284 (.0302)	-.571 (.0841)	.173 (.0307)	-.896 (.0949)	-1.27 (.242)
Final-Offer	.0968 (.0416)	.268 (.115) [.0938]	.0832 (.0395)	.236 (.119) [.0776]	.278 (.316) [.0562]
Tri-Offer	.0391 (.0416)	.112 (.117) [.0392]	-.0102 (.0398)	-.0383 (.113) [-.0126]	.0159 (.372) [.00321]
Average Dispute Rate, Rounds 1-10 (for pair)	---	---	1.16 (.126)	3.22 (.392) [1.06]	---
Log Likelihood		-485.6		-448.6	-382.6
R^2	.0071		.106		
n=770					

Note: The numbers in parentheses are standard errors. The numbers in square
brackets are the average derivatives of the probit function with respect to
each variable. The derivative for variable j is the average over sample
observations i of $\hat{\beta}_j \phi(X_i \hat{\beta})$. The base group is the high-variance conventional
arbitration treatment. The mean of the average dispute rate in the first ten
rounds is .114 (sd=.128). The estimated variance of the individual effects
in the random effects probit model is .999 (se = .308). The log-likelihood
for the constrained model with only a constant is -488.2.

estimate of the variance of the fixed effect is .999 (se = .308) which is almost precisely the same size as the variance of the random component for each case (σ_ε^2 is normalized to one in the probit). These findings are once again conditional on the untested distributional assumptions.

Our finding that dispute rates are at least as high in FOA as in CA is particularly interesting in light of the fact that FOA was developed in response to criticisms that CA led to a "chilling" of bargaining and very high dispute rates (Stevens, 1966; Feuille, 1975). It is generally argued that FOA will lead to fewer disputes, but the evidence seems not to bear this out. We turn now to a discussion of some theoretical considerations that can shed some light on this apparent contradiction.

V. Implications of the Results for a Simple Model of Negotiator Behavior

The evidence presented in the previous section, that dispute rates in the CA treatments vary inversely with uncertainty, suggests that, on average, our subjects behave as if they were risk averse. A natural question is whether our finding of a higher dispute rate in the FOA than in the comparable CA treatment can also be explained by risk aversion.

A simple model (Farber and Katz, 1979) can be used to relate the risk preferences of the negotiators and the variance of arbitration decisions to the costs of disagreement. For the purposes of interpreting relative dispute rates across different arbitration mechanisms, a parametric model of negotiator behavior is necessary.

Assume that the negotiation is over some variable (Y), measured as a deviation from some benchmark level. Recall that negotiator (A) wants as high a value for Y as possible, and the other negotiator (B) wants as low a value as possible. Some important quantities are:

Y_f Arbitrator's notion of a fair award.

Y_A A's final offer

Y_B B's final offer

$\overline{Y}$ average of final offers: $(Y_A + Y_B)/2$

Y_r fact-finder's recommendation

Since our experiment drew arbitrator's fair awards from a normal distribution, we assume that the underlying distribution of (statistically-exchangeable) arbitrator's fair awards is normal so that $Y_f \sim N(\mu, \sigma^2)$.

With regard to the preferences of the negotiators, we assume that they both have utility functions that exhibit constant absolute risk aversion (CARA):

$$(V.1) \qquad U_A[Y] = 1 - \exp[-\delta(Y-\mu)]$$

and

$$(V.2) \qquad U_B[Y] = 1 - \exp[-\delta(\mu-Y)].$$

The coefficient of absolute risk aversion, $\delta = -U''(x)/U'(x)$, is assumed the same for both negotiators, and the utility functions treat deviations in outcomes from the mean of the arbitrator's distribution (μ) symmetrically.[15] The utilities are equal at $Y=\mu$, and any deviation from this yields a utility for A that is the same as the utility for B from a deviation of opposite sign with the same magnitude.

A. Conventional Arbitration

Where conventional arbitration was the dispute settlement mechanism in our experiment, the arbitrator simply imposes a randomly drawn award (Y_f). The negotiators have no control over the award through their offers so that the risk preferences of the negotiators are not relevant to the distribution

---

[15]Given random assignment of subjects as A or B, it is appropriate to assume that preferences are symmetric.

of arbitrated outcomes. The expectation and variance of the arbitration award are simply the mean (μ) and variance (σ^2) of the normal distribution from which the awards are selected.

Assuming the CARA utility functions in V.1 and V.2 and that Y_f is distributed $N(\mu,\sigma^2)$, the expected utility of each of the bargainers from using arbitration is[16]

$$(V.3) \qquad E(U_A) = E(U_B) = 1 - \exp[\frac{1}{2}\delta^2\sigma^2].$$

The parties have identical expected utilities because they have identical risk preferences, their utility functions are symmetric around the mean of the distribution of Y_f, and the distribution of Y_f is symmetric.

The total risk premium of the negotiators is computed in the following way. A negotiator would prefer any settlement with certainty that yielded a higher utility than the expected utility from arbitration. The minimum certain settlement that party A would accept (Y_{mA} = A's certainty-equivalent) is the solution for Y_{mA} of

$$(V.4) \qquad U_A(Y_{mA}) = E(U_A).$$

Similarly, the maximum certain settlement that party B would accept (Y_{mB} = B's certainty-equivalent) is the solution for Y_{mB} of

$$(V.5) \qquad U_B(Y_{mB}) = E(U_B).$$

The difference between the certainty equivalents defines the range of settlements that both parties would prefer to arbitration (the contract zone). With our assumptions of normality and CARA preferences, the total cost of disagreement due to uncertainty (the total risk premium, TRP) is simply the product of the coefficient of absolute risk aversion (δ) and the variance of the underlying distribution of arbitrators' awards (σ^2),

---

[16]This follows directly from the moment generating function of a normal distribution.

$$(V.6) \qquad TRP_{CA} = Y_{mB} - Y_{mA} = \delta\sigma^2.$$

A useful benchmark case is the situation where the parties are risk neutral ($\delta=0$). Under CA, risk neutral parties will be indifferent between a negotiated outcome of μ and an arbitration decision that has expected value of μ. The total risk-premium is zero and there is no uncertainty cost of using arbitration.

So long as the parties are risk averse, this simple model implies that the parties will always settle their disputes. This is not very plausible. We prefer to interpret this model as implying that the probability of a settlement is greater the greater is the total risk premium. This is equivalent to assuming that the parties attitudes toward risk contain an idiosyncratic component that varies randomly across bargaining rounds. Given the distribution of these idiosyncratic components, the parties will be more likely to reach a voluntary agreement the greater is the size of the total risk premium. Since the size of the total risk premium is positively related to the variance in the arbitrator decisions that the parties face, this implies that the settlement rates in our CA experiments should be greater in the high variance treatment groups, which is precisely what we find.

This analysis also has implications for other aspects of the parties bargaining behavior. The contract zone in (V.6) is centered on μ so long as the parties have the same attitudes toward risk. Although we remain agnostic about particular bargaining models, most models will imply that agreements negotiated in the shadow of the arbitration system should not differ systematically from the center of the contract zone μ. Since $\mu=350$ in our CA experiments, we expect the average negotiated settlement to be 350. We test this hypothesis below.

B. Final-Offer Arbitration

Where final-offer arbitration is the dispute settlement mechanism, the arbitrator chooses the final offer that is closest to a randomly drawn value (Y_f). Since the final offers are determined by the parties, the distribution of outcomes cannot be computed without an assumption about how the preferences of the negotiators are translated into the final offers.

Since in our experiments Y_f has a normal distribution, the probability that the arbitrator chooses B's offer is

$$(V.7) \qquad Pr(ch\ Y_B) = \Phi([\overline{Y}-\mu]/\sigma) = 1 - Pr(ch\ Y_A)$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function. The expected utility of Party i is a probability-weighted average of i's utility at each of the offers.

In the benchmark case of risk neutrality the Nash equilibrium pair of offers, which has the property that neither negotiator can make themselves better off by changing their offer, are

$$(V.8) \qquad Y_A = \mu + \frac{1}{2}[\sigma/\phi(0)]$$

and

$$(V.9) \qquad Y_B = \mu - \frac{1}{2}[\sigma/\phi(0)]$$

where $\phi(0)$ is the standard normal density function evaluated at zero (approximately 0.3989). These offers are equidistant from the mean of the Y_f distribution so that the offers will be equally likely to be selected.[17] As under CA, the parties will be indifferent between a negotiated outcome of μ and taking their chances with an arbitrator who selects either Y_A or Y_B according to which is the closest to (the randomly drawn) Y_f. Thus, as with CA, the total risk-premium is zero and there is no uncertainty cost of using

---

[17]This is strictly a result of the facts that the utility functions of the two negotiators treat deviations from the mean of the (symmetric) distribution identically and that the negotiators are equally risk averse.

arbitration when the bargainers are risk neutral.

Under risk neutrality this simple model has strong implications for our experimental data on FOA. In particular, since $\mu=350$ and $\sigma=50$ in our FOA experiments, the difference between the offers should be approximately $2.5\sigma = 125$, and the offers should be approximately 412.5 and 287.5. We will test these hypotheses below.

Assuming the CARA utility function in equation (V.1), the expected utility of party A is

$$(V.10) \qquad EU_A = 1 - \Phi[(\bar{Y}-\mu)/\sigma]\exp[-\delta(Y_B-\mu)] - \{1-\Phi[(\bar{Y}-\mu)/\sigma]\}\exp[-\delta(Y_A-\mu)].$$

Similarly, the expected utility of party B is

$$(V.11) \qquad EU_B = 1 - \Phi[(\bar{Y}-\mu)/\sigma]\exp[\delta(Y_B-\mu)] - \{1-\Phi[(\bar{Y}-\mu)/\sigma]\}\exp[\delta(Y_A-\mu)].$$

The Nash equilibrium pair of offers in this case (Farber, 1980), which has the property that neither negotiator can make themselves better off by changing their offer, is

$$(V.12) \qquad Y_A = \mu + (1/2\delta)\cdot\ln[1+\delta\sigma/\phi[0]]$$

and

$$(V.13) \qquad Y_B = \mu - (1/2\delta)\cdot\ln[1+\delta\sigma/\phi[0]].$$

Note that under risk aversion A's offer is strictly less than under risk neutrality and B's offer is strictly greater than under risk neutrality. The difference between the optimal offers is

$$(V.14) \qquad Y_A - Y_B = \frac{1}{\delta}\cdot\ln[1+\delta\sigma/\phi[0]]$$

Given the known value of $\sigma=50$ in our FOA experiment, the difference between the offers in our experiment can be interpreted as evidence on the risk preferences of the negotiators.

The total risk premium of the negotiators is computed in a manner analogous to CA. A negotiator would prefer any settlement with certainty that

yielded a higher utility than the expected utility from arbitration. With
our assumptions of normality and CARA preferences and our calculation of the
equilibrium offers, the total cost of disagreement due to uncertainty (the
total risk premium, TRP) is

$$(V.15) \qquad TRP_{FOA} = \frac{2}{\delta}\ln\left(\frac{1}{2}(1+\delta\sigma/\phi[0])^{-.5} + \frac{1}{2}(1+\delta\sigma/\phi[0])^{.5}\right).$$

As with CA and for reasons of symmetry, the center of this contract zone is
μ. We should therefore expect the negotiated settlements to center around μ
(= 350 in our FOA experiment) when the parties bargain in the shadow of FOA.
We will also test this hypothesis below.

It is now straightforward to examine the relative costs of disagreement
by comparing TRP_{FOA} with TRP_{CA} in the situation where the underlying
distributions have the same variance. Total risk premia under FOA and HVCA
for selected levels of risk aversion are contained in table 4. This table
and figure 2 illustrates that, at very low levels of risk aversion and for
our experimental level of uncertainty ($\sigma=50$), the cost of disagreement is
higher for FOA than for CA while at higher levels of risk aversion, the cost
of disagreement is higher for CA than for FOA.

The intuition for why the total risk premium in FOA rises at a sharply
declining rate is that risk-averse negotiators in FOA can mitigate their risk
by making conservative offers. The same negotiators in CA have no such
option so that the risk premium rises linearly with risk aversion.[18] To
illustrate the convergence of the offers, table 4 contains the calculated
difference between the offers for selected levels of risk aversion (δ) and
figure 3 contains a plot of the calculated difference in the offers in FOA as

---

[18]Farber's (1981) model of CA allows the negotiators to mitigate the risk to
some extent in that procedure. However, Bazerman and Farber's (1985)
analysis of arbitrator behavior in CA suggests that the opportunities for
this are quite limited as an empirical matter.

Table 4

Difference Between Optimal Offers and Total Risk Premia

Various Values of Coefficient of Absolute Risk Aversion

CARA	CA	FOA		TRI	
	TRP	Diff	TRP	Diff	TRP
0	0	125	0	150	0
.002	5	112	6.25	136	4.17
.004	10	102	10.25	128	7.58
.01	25	81	16.07	124	14.9
.02	50	63	18.5	108	22.1

CARA = coefficient of absolute risk aversion

TRP = total risk premium

Diff = Difference between optimal offers

Note: The figures for CA and FOA are derived from formulas given in the text. The figures for TRI are based on calculations from a Monte-Carlo simulation with sample size 10,000. The underlying distribution of arbitrators' fair awards was normal with mean 350 and standard deviation 50 in all cases.

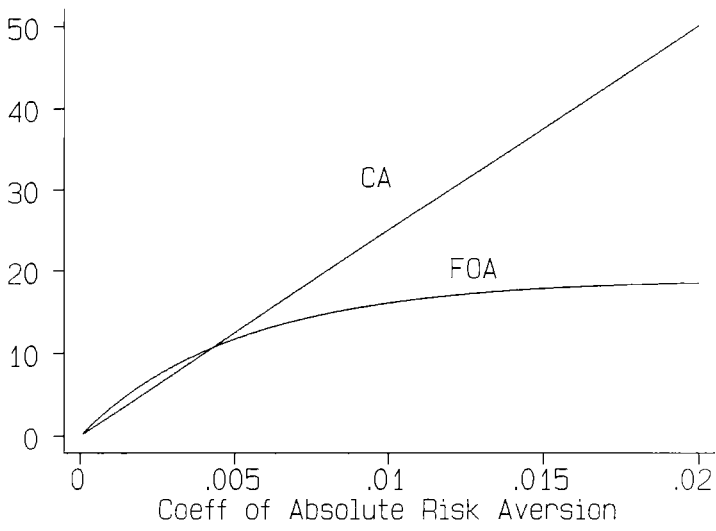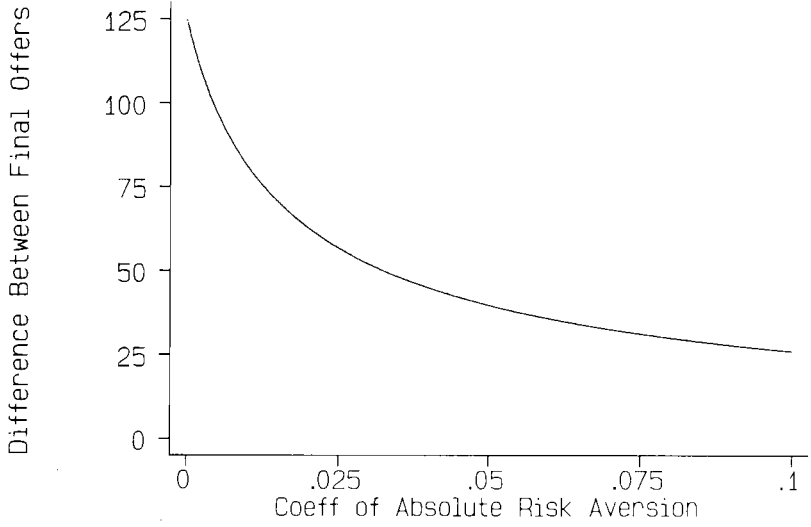Figure 2
Total Risk Premia in HVCA and FOA

Figure 3
Difference Between Equilibrium Offers in FOA

a function of δ. The predicted difference between the offers in the special case where the negotiators are risk-neutral is 125. As the parties are increasingly risk-averse, the difference between the offers falls from this value to 112 where $\delta=.002$ and further to 62 where $\delta=.02$. The difference between offers in FOA is 100 at the value of δ (.0043) where the CA and FOA total risk premium are equal.

C. Tri-Offer Arbitration

The analysis of tri-offer arbitration is more complicated. The arbitrator now has a choice from among three alternatives: the final offers of the parties and the recommendation of the fact-finder. The parties formulate their offers so as to maximize their expected utilities, and the arbitrator chooses the alternative that is closest to his/her randomly drawn value (Y_f). Once again, the distribution of outcomes cannot be computed without using specific assumptions about the risk preferences of the negotiators.

In the benchmark case of risk neutrality, the Nash equilibrium pair of offers depends on the recommendation of the fact-finder. Party A wants to maximize and party B wants to minimize

$$(V.16) \qquad E(Y) = Pr(\text{ch } Y_B) \cdot Y_B + [Pr(\text{ch } Y_A)] \cdot Y_A$$
$$+ [Pr(\text{ch } Y_r)] \cdot Y_r$$

The three possible configurations of the facts and offers were described in section II along with the relevant probabilities. Assuming that Y_f has a normal distribution, numerical analysis of the determination of optimal offers and the resulting expected wage in each of the three regimes yields the result that the optimal offers are of case #1 $(Y_B<Y_r<Y_A)$ if

$$(V.17) \qquad \mu - 1.32\sigma < Y_r < \mu + 1.32\sigma.$$

The offers are of case #2 $(Y_r<Y_B<Y_A)$ if

$$(V.18) \qquad Y_r < \mu - 1.32\sigma,$$

and the offers are of case #3 $(Y_B < Y_A < Y_r)$ if

(V.19) $Y_r > \mu + 1.32\sigma.$

Assuming that Y_r is drawn from the same normal distribution as Y_f, the offers will bracket the fact-finder's recommendation approximately eighty percent of the time while cases #2 and #3 will each occur in about ten percent of the cases.

A Monte-Carlo simulation of the arbitration process was carried out by drawing 10,000 fact-finder's recommendations from a normal distribution with the standard deviation of 50 used in our TRI experiment and computing the optimal offers and the arbitrator's choice probabilities in each case. The difference between the offers was computed to be 147.5 (2.95σ), which is larger than the 125 (2.5σ) difference in FOA.

As with FOA, these computations are sensitive to the presence of risk aversion. The expected utilities of the negotiators are

(V.20) $E(U_A) = \Pr(\text{ch } Y_B) \cdot U_A(Y_B) + [\Pr(\text{ch } Y_A)] \cdot U_A(Y_A)$
$+ [\Pr(\text{ch } Y_r)] \cdot U_A(Y_r)$

and   (V.21) $E(U_B) = \Pr(\text{ch } Y_B) \cdot U_B(Y_B) + [\Pr(\text{ch } Y_A)] \cdot U_B(Y_A)$
$+ [\Pr(\text{ch } Y_r)] \cdot U_B(Y_r)$

where U_A and U_B are defined in equations V.1 and V.2.

The Nash equilibrium pair of final offers will be closer together on average than in the risk-neutral case. Consider, for example, the constant absolute risk aversion utility functions in equations (V.1) and (V.2) where δ = 0.01. Assuming again that Y_f has a normal distribution with $\sigma=50$ as in our TRI experiment, numerical analysis of the determination of optimal offers and the resulting expected wage in each of the three regimes yields the result

that the optimal offers are of case #1 $(Y_B < Y_r < Y_A)$ if[19]

(V.22) $\mu - 97.26 < Y_r < \mu + 97.26.$

The offers are of case #2 $(Y_r < Y_B < Y_A)$ if

(V.23) $Y_r < \mu - 97.26,$

and the offers are of case #3 $(Y_B < Y_A < Y_r)$ if

(V.24) $Y_r > \mu + 97.26.$

Comparing this with equations (V.17)-(V-19) and noting that $1.32\sigma = 66$, it is clear that risk-averse negotiators will have optimal offers that bracket the fact-finder's recommendation in a larger fraction of the cases than risk-neutral negotiators. Assuming that Y_r is drawn from the same normal distribution as Y_f, the offers will bracket the fact-finder's recommendation approximately ninety-five percent of the time while cases #2 and #3 will each occur in about 2.5 percent of the cases.

The same 10,000 randomly-selected fact-finder's recommendations used in the risk-neutral case, were used to compute the optimal offers and the arbitrator's choice probabilities in each case. The difference between the offers was computed to be 124 (2.48σ), about 17 percent smaller than the difference of 2.95σ computed for risk-neutral negotiators in TRI. Compared with FOA, on the other hand, the offer difference is substantially larger than the the 81.2 (1.62σ) difference in FOA for negotiators with the same risk preferences.

Table 4 contains calculated differences in optimal offers for a range of risk aversions from $\delta=.002$ through $\delta=.02$ computed using the 10,000 randomly generated fact-finder's recommendations. It is clear that the

---

[19] The boundary values for the fact-finder's recommendation in determining the regime are not simply multiples of the standard deviation of Y_f as they are in the risk-neutral case.

difference between the optimal offers is always substantially larger in TRI than in FOA for negotiators with the same risk preferences. Intuitively, the negotiators in TRI present more extreme offers on average because the fact-finder's report is generally in the middle of the distribution and is available to the arbitrator as an option.

Table 4 also contains calculated total risk premia for a range of risk aversions computed using the 10,000 randomly generated fact-finder's recommendations, and these can be used to compare the costs of arbitration in TRI with the costs of arbitration in the other mechanisms. At reasonable levels of risk aversion, TRP_{TRI} is close to TRP_{FOA} and substantially lower than TRP_{CA}. As in FOA, risk-averse negotiators in TRI are mitigating the risk by making offers that are closer together. The risk in TRI is also reduced by the availability of the fact-finder's recommendation as an option for the arbitrator.

D. Comparison of Mechanisms

This analysis shows that the relationship between the risk premia under CA, FOA and TRI cannot be determined theoretically without a priori information on the risk preferences of the negotiators. It is useful to ask whether $\delta=.0043$, the degree of risk aversion where CA becomes riskier than FOA, is small or large.[20] Consider the total risk premium (TRP) in CA, where the parties cannot mitigate the risk. Where $\delta=.002$, the TRP is 5.0, and this rises to 10.7 where $\delta=.0043$. These values do not seem very large when compared with the standard deviation of the arbitration award ($\sigma=50$). Thus, negotiators with $\delta=.0043$ would each be willing to give up 5 units, only about one-tenth of the standard deviation of the arbitration award, to avoid the

---

[20]TRI becomes less risky than CA at an even lower degree of risk aversion (δ approximately .002).

risk. If δ=.02, the TRP in conventional arbitration is exactly 50, and each
negotiator would be willing to give up one-half of the standard deviation of
the award to avoid the risk. All we mean to conclude from this calculation
is that δ=.0043 does not seem like a very large degree of risk aversion, and
it would not be surprising to find that negotiators exhibited more than this
degree of risk aversion.

Given our interpretation of the evidence on relative dispute rates in
the three CA treatments, that negotiators behave on average as if risk
averse, our experimental finding of higher (or no lower) dispute rates in FOA
and TRI than in the comparable CA treatment is consistent with our simple
model of negotiator behavior. For reasonable degrees of risk aversion, the
costs of disagreement due to uncertainty may simply be higher in CA than in
FOA or TRI. Thus, the ability of the parties to mitigate their risk under
FOA or TRI may lead to higher dispute rates.

VI. Negotiator Behavior: Evidence from the Offers and Negotiated Settlements

The simple model of negotiator behavior that we have outlined above
provides an explanation for some of the variability in dispute rates that we
found in our experiments. As we have seen, direct tests of this model of
negotiator behavior are difficult to implement with data from dispute rates
alone. However, this model also has a rich set of implications for the
offers that the parties submit to arbitration in FOA and TRI, and for the
negotiated settlements upon which the parties agree in CA, FOA, and TRI.
Here we explore these data to determine the extent to which they are
consistent with our model of negotiator behavior.

A. The Parties Offers

Under conventional arbitration there are no binding offers made to the
arbitrator. Although our experimental subjects proposed offers to each other

in the course of reaching agreements, these offers are a form of "cheap talk" under CA since the bargainers paid no penalty for making outrageous demands. Under FOA and TRI, however, the parties submitted final offers that might well be accepted and implemented through the arbitration process. Outrageous demands are costly in these systems because they may lead the arbitrator to accept an opponent's proposal.

Sample statistics describing the 96 pairs of offers proposed by our experimental subjects in FOA are contained in Table 5. Our simple model of negotiator behavior implies that these offers should be symmetrically positioned around the center of the distribution of potential arbitrator awards regardless of whether the parties are risk neutral or risk averse. Since $\mu=350$ in our experiments, these offers should therefore be symmetrically placed around 350.

To provide a benchmark for assessing whether the parties are behaving in a way that is risk averse, Table 5 also contains the predicted values for the parties offers under the assumption of risk neutrality. Risk averse bargainers should be submitting offers that are more conservative than the risk neutral predictions. (That is, if the parties are risk averse, Party A's offers should be lower, and Party B's offers should be higher, than the risk neutral predictions.) Note, however, that if risk preferences vary among our subjects, it is the least risk averse subjects who will be most likely to engage in disputes that require arbitration. Thus, the parties for whom we record final offers are likely to be among the least risk averse.

As the data in Table 5 indicate, the parties' final offers are both somewhat lower than the risk neutral predictions. As a result, the difference in the mean offers is virtually identical to what would be predicted under risk neutrality, while the mean of their position is lower than would be predicted. Although the mean offers are statistically

Table 5

Observed and Calculated Offers

Final-Offer Arbitration

	Mean Offer (s.e.)	Calculated Offer under risk neutrally	Variance (s.e.)
Party A	396.2 (4.81)	413	2218. (319.)
Party B	269.5 (8.31)	288	6626. (954.)
Average	332.8 (4.73)	350	
Difference	127. (9.73)	125	

Note: The sample size is 96. The predicted offers are calculated as described in the text under the assumption of risk neutrality, and they are the same for all observations. The reported standard errors are the standard deviation of the mean.

significantly different from the predicted values (p-values < .001), it is unclear how great the substantive importance of these differences are in terms of foregone payoffs to the subjects.[21] There is no strong evidence from these offers favoring the hypothesis that the parties are risk averse. This is consistent with the hypothesis that it is the more risk averse bargainers who are more likely to settle their disputes.

Sample statistics describing the 84 pairs of offers made by the parties in tri-offer arbitration are contained in Table 6. In our simple model of negotiator behavior, the parties optimal offers should vary with the recommendation made by the fact-finder, except where the fact-finder's proposal is very extreme. As a benchmark, table 6 provides the mean difference of the parties' offers from the optimal offers for risk neutral parties. Risk averse bargainers should make offers that are conservative compared to the risk neutral offers. (That is, Party A's offer should be less than, and Party B's offer should be greater than, the risk neutral offer.) As with the FOA offers, the parties offers under TRI are both less than the risk neutral offers. Only one of these differences is statistically significant, and it is unlikely that either difference would lead to a substantial change in the parties monetary payoffs. As with FOA, the parties offers under tri-offer arbitration do not indicate strong evidence of risk aversion. This is also consistent with the hypothesis that it is the least risk averse parties who carry their disputes to arbitration.

Under tri-offer arbitration there is a simple way to examine the sensitivity of the parties offers to the fact-finders recommendation. The parties optimal offers become more extreme in tri-offer arbitration as the fact-finders recommendations become more extreme $(Y_B < Y_r < Y_A)$ until a point is

---

[21]See Harrison (1989) for a discussion of this issue.

Table 6

Observed and Calculated Offers

Tri-Offer Arbitration

	Mean Offer (s.e.)	Mean Difference Offer – Pred. offer Risk Neutrality
Party A	387.1 (8.16)	-34.8 (8.11)
Party B	269.2 (7.43)	-9.39 (7.38)
Factfinder	349.2 (4.91)	
Average of A and B (observed)	328.2 (6.10)	
Difference between A and B (observed)	118.0 (9.74)	
Difference between A and B (predicted, risk neutrality)	145.2 (2.47)	

Note: The sample size is 84. The predicted offers are calculated as described in the text using the actual fact-finder's recommendations under the assumption of risk neutrality. The reported standard errors are the standard deviation of the mean.

reached where the fact-finders recommendation is more extreme than one of the offers ($Y_r < Y_B < Y_A$ or $Y_B < Y_A < Y_r$). When the parties are risk neutral it is a straightforward matter to tabulate the order that the optimal offers should take given the fact-finder's recommendations in our experiment. Table 7 contains the results of this tabulation and a comparison of these results against the actual ordering of the offers. Although there is a significant correlation between the actual and predicted offers, the presence of a considerable number of offers that are off the diagonal of this table indicates that the correlation is far from perfect. The most notable prediction errors are in the bottom left-hand corner of the table, indicating, as did Table 6, that the parties offers are typically lower than expected.

There may be a relatively simple explanation for why the mean of the parties offers in both FOA and TRI are centered below the point $\mu=350$ that is expected in our simple model and for why the tri-offer cases where $Y_B < Y_A < Y_r$ occur more frequently than is expected in our simple model. Our experiments were designed so that the bargainers would be likely to negotiate settlements prior to the implementation of the arbitration treatments around the point 300, which was the mid-point of the feasible range of bargains (from 100 to 500) in the pre-arbitration periods. However, we deliberately centered the distribution of potential arbitration decisions around the point 350 in order to determine whether the parties would respond to this change. It appears that the bargainers may have come to expect agreements around the point 300 in the pre-arbitration periods and that they found it difficult to react to the change in the bargaining environment that resulted from the implementation of the arbitration systems. We explore this hypothesis in more detail below.

Table 7

Tabulation of Ordering of Offers and Predicted Ordering of Offers

Tri-Offer Arbitration

Risk Neutrality

Predicted Ordering

Observed Ordering	$Y_B < Y_r < Y_A$	$Y_r < Y_B < Y_A$	$Y_B < Y_A < Y_r$
$Y_B < Y_r < Y_A$	35	4	5
$Y_r < Y_B < Y_A$	6	5	0
$Y_B < Y_A < Y_r$	20	0	5

Note: Four cases where the offers were ordered $Y_A < Y_B < Y_r$ are deleted from this table.   The sample size is 80.

Y_A = A's offer

Y_B = B's offer

Y_r = fact-finder's recommendation

B. *Negotiated Settlements*

Summary statistics describing the settlements negotiated by the bargainers in our experiments are contained in Table 8. In our experiments the center of the bargaining range in the pre-arbitration periods was 300. Since the bargaining pairs were assigned randomly we expected the negotiated settlements in the pre-arbitration periods (rounds 1-10) to center around 300. As Table 8 indicates, in most cases the negotiated settlements settled slightly below 300 in the pre-arbitration periods, and in four of the six cases these differences are statistically significant.

Since the distribution of potential arbitration awards was centered around 350, we expected that the negotiated settlements would be higher under the arbitration regimes (rounds 11-20) than under the pre-arbitration regime (rounds 1-10). The results in table 8 indicate that negotiated settlements averaged significantly less than 350 in four of the five arbitration treatments. The average settlement in the control group remained significantly less than 300.

A comparison of the mean of the negotiated settlements in the pre-arbitration and arbitration regimes is complicated by the fact that the incidence of disputes is far greater in the arbitration periods and not evenly distributed across bargaining pairs. The final column of Table 8 contains estimates of the arbitration treatment effects that control for individual differences among the bargaining pairs by including in a regression a dummy variable for each pair. These treatment effects represent only the within-pair effects of the arbitration treatments. The results indicate that the arbitration treatments did increase the values of the negotiated settlements, and four of the five estimated treatment effects are statistically significant at conventional test levels. In no case, however, is the size of the estimated treatment effect as large as our simple model

Table 8

Average Negotiated Settlements by Treatment

(standard error)

[cell size]

TREATMENT	# pairs	Pre-Treatment (Rounds 1-10)	Treatment (Rounds 11-20)	Treatment Effect
Control no arbitration	10	276.8 (8.30) [96]	258.5 (8.91) [96]	-19.5 (9.78)
Low-Variance Conventional Arbitration	29	263.7 (5.92) [250]	277.4 (6.66) [170]	19.2 (6.96)
Medium-Variance Conventional Arbitration	25	296.3 (7.00) [224]	307.6 (7.82) [171]	11.7 (7.05)
High-Variance Conventional Arbitration	25	307.4 (7.06) [226]	356.7 (6.91) [179]	46.6 (6.94)
Final-Offer Arbitration	26	283.8 (7.32) [232]	279.7 (8.46) [161]	16.5 (7.22)
Tri-Offer Arbitration	26	273.7 (6.41) [224]	305.4 (8.77) [154]	41.8 (7.31)
All Treatments	131	283.9 (2.88) [1252]	301.3 (3.39) [931]	22.7 (3.04)

Note: The standard errors reported in the first two columns are the standard deviations of the sample mean. The estimated treatment effect (standard error) in the last column is the parameter (standard error) on a dummy variable for rounds 11-20 from a regression for each treatment that includes pair fixed effects.

would have predicted.

These data indicate that the parties negotiated settlements did respond to the incentives created by the introduction of the arbitration mechanisms. By altering the relative bargaining strength of the parties at the same time as we introduced the arbitration mechanisms we may have attempted to accomplish too much with these experiments. Further experimentation might usefully shed some light on this issue.

VII. Conclusions

Our experiment yielded three concrete findings. First, dispute rates are inversely related to the monetary costs of disputes. Dispute rates were much lower in cases where arbitration was not available so that the entire pie was lost in the event of a dispute. This confirms the empirical importance of the so-called "chilling effect" on bargaining that has been conjectured is produced by the adoption of arbitration systems. Second, the dispute rate in a final-offer arbitration system is at least as high as the dispute rate in a comparable conventional arbitration system. Contrary to the usual argument, we find no evidence that final-offer arbitration eliminates the chilling effect. Third, dispute rates are inversely related to the uncertainty costs of disputes. Dispute rates were lower in conventional arbitration treatments where the variance of the arbitration award was higher and imposed greater costs on risk-averse negotiators.

The experiments can also be interpreted as providing tentative evidence that the negotiators were risk-averse on average. The three conventional arbitration treatments differed only in the variance of the arbitration awards implicit in the information given to the negotiators, and we found that the dispute rates varied inversely with the variance. Using the result that dispute rates are inversely related to the cost of disagreement, the

findings on dispute rates must mean that the costs of disagreement are higher where the variance of the arbitration award is higher. This can only be true if the negotiators are risk averse on average. On the same basis, a moderate degree of risk aversion is also consistent with our tentative finding of higher dispute rates in final-offer arbitration than in conventional arbitration.

Our analysis of the final offers in the FOA and TRI treatments concluded that these offers were roughly consistent with risk neutral behavior. Since disputes (and, hence, offers) are more likely to be observed where there is less risk aversion, this is additional tentative evidence of some average degree of risk aversion.

Finally, where it is possible to make comparisons, we find general agreement between the dispute rates in our experiment and dispute rates found in the field in comparable settings. This suggests that the design of our experiments provides a useful way to study disputes in a laboratory environment.

REFERENCES

Ashenfelter, Orley. "A Model of Arbitral Behavior", American Economic Review, 77 (May 1987).

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

Ashenfelter, Orley; James Dow; and Daniel Gallagher "Factfinder and Arbitration Behavior under an Appellate System", unpublished draft, Princeton University, November 1985.

Babcock, Linda. "The Role of Incomplete Information in Bargaining under the Threat of Final-Offer Arbitration", unpublished draft, Carnegie-Mellon University, November 1987.

Bazerman, Max and Henry S. Farber, "Arbitrator Decision Making: When are Final Offers Important?" Industrial and Labor Relations Review, 60 (October 1985), 76-89.

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

Chatterjee, Kalyan and William Samuelson. "Bargaining Under Incomplete Information," Operations Research 31 (1983), 175-192.

Currie, Janet. "Who Uses Interest Arbitration: The Case of British Columbia's Teachers 1947-1981", Industrial and Labor Relations Review 33 (October 1989), 363-379.

Currie, Janet and Sheena McConnell. "Strikes and Arbitration in the Public Sector: Can Legislation Reduce Dispute Costs?", unpublished draft, UCLA, August 1989.

Farber, Henry S. "An Analysis of Final-Offer Arbitration," Journal of Conflict Resolution, 24 (December 1980), 683-705.(a)

Farber, Henry S. "Splitting-the-Difference in Interest Arbitration," Industrial and Labor Relations Review, 35 (April 1981), 70-77.

Farber, Henry S. and Max H. Bazerman, "The General Basis of Arbitrator Behavior: An Empircal Analysis of Conventional and Final-Offer Arbitration," Econometrica, 54 (November 1986), 1503-1528.

Farber, Henry S. and Max H. Bazerman, "Divergent Expectations as a Cause of Disagreement in Bargaining: Evidence From a Comparison of Arbitration Schemes," Quarterly Journal of Economics 54 (February 1989), 99-120.

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

Feuille, Peter, "Final Offer Arbitration and the Chilling Effect," Industrial Relations, XIV (October 1975), 302 -310.

Gibbons, Robert, "Learning in Equilibrium Models of Arbitration," American Economic Review, 78 (December 1988), 111-124.

Grigsby, David and William Bigoness, "The Effects of Third Party Intervention on Pre-Intervention Bargaining Behavior," Journal of Applied Psychology, 76 (October 1982), 549-54.

Harrison, Glenn, "Theory and Misbehavior of First-Price Auctions," American Economic Review, 79 (September 1989), 749-762.

Kochan, Thomas A. and Jean Baderschneider, "Determinants of Reliance on Impasse Procedures: Police and Firefighters in New York State." Industrial and Labor Relations Review, XXXI (July 1978), 431-440.

Lester, Richard. Labor Arbitration in State and Local Government, Princeton, New Jersey. Princeton University Industrial Relations Section, 1984

Magenau, John M, "The Impact of Alternative Impasse Procedures on Bargaining: A Laboratory Experiment", Industrial and Labor Relations Review, 36 (April 1983), 378-388.

McConnell, Sheena. "Strikes, Wages, and Private Information", American Economic Review, 79 (September 1989), 801-815.

Neale, Margaret A., and Max H. Bazerman, "The Role of Perspective Taking Ability in Negotiating Under Different Forms of Arbitration," Industrial and Labor Relations Review, 36 (April 1983), 378-388.

Notz, W. W., and F.A. Starke, "Final Offer versus Convnetional Arbitration as Means of Conflict Management," Administrative Science Quarterly, 23 (June 1978), 189-203.

Stevens, Carl, "Is Compulsory Arbitration Compatible with Bargaining?", Industrial Relations, V (February 1966), 38-52.

Stone, I.F. The Trial of Socrates. Boston, Little, Brown & Co. 1988.

Tracy, Joseph S. "An Investigation into the Determinants of U.S. Strike Activity," American Economic Review, 76 (June 1986), 423-436.