Industrial Relations Section
Princeton University
Working Paper No. 213
August 1986

Longitudinal Analysis of Strike Activity*

David Card
Princeton University

August 1986

*I am grateful to Wayne Vroman for making available his contract data, and to Sheena McConnell and Joe Tracy for access to their data on strikes. Thanks to Rebecca Blank, George Jakubson and members of the NBER Labor Studies Group for comments and suggestions on an earlier draft.

Longitudinal Analysis of Strike Activity*

ABSTRACT

This paper presents evidence on two aspects of strike activity associated with the renegotiation of union contracts: the effects of contract characteristics on dispute probabilities; and the variation in strike activity over time within bargaining pairs. Cross-sectional and longitudinal estimation techniques show that strike probabilities are higher in summer and fall than winter and spring. Strike probabilities are also increased by increasing the length of time between negotiations, and reduced in limited wage reopening negotiations. Finally, strike probabilities are significantly affected by lagged strike outcomes. Relative to a peaceful settlement of the previous contract, strike probabilities are 10 percentage points higher following a strike of two weeks or less, and 5 to 7 percentage points lower following a strike of longer than two weeks.

David Card
Department of Economics
Princeton University
(609) 452-4045

Longitudinal Analysis of Strike Activity

I. Introduction

The growing availability of cross-sectional data on contract settlements in the union sector, together with recent advances in the game-theoretic analysis of bargaining, have generated new interest in the empirical analysis of strikes.^{1/} Most of this new research is focused on the role of strikes in revealing information to one party about the willingness of the other to participate in a match.^{2/} Direct tests of this imperfect-information paradigm, however, have proved elusive, since the key determinant of strike activity -- the value of a settlement to one of the parties -- is in principle unobservable. Tests of imperfect information strike models have therefore relied on indirect evidence in favor of the theory.^{3/}

In this paper I present a variety of evidence on the longitudinal structure of strike activity that may prove useful in evaluating information-based theories of strikes. In particular, I present evidence from a panel of negotiations for some 250 bargaining pairs on two sets of issues: the effect of contract characteristics on the likelihood of disputes; and the effect of past strike activity on the incidence and duration of subsequent disputes.

The main findings may be summarized as follows. First, the characteristics of collective bargaining agreements themselves appear to have a strong influence on the likelihood of disputes. Strike probabilities depend on the month in which the previous contract expires, on the dura-

tion of the previous agreement, and on the presence or absence of contract provisions limiting the range of negotiations to one or two issues. Second, strike probabilities depend on strike outcomes in previous negotiations. Strike probabilities are significantly higher following a short strike than a peaceful settlement, whereas strike probabilities are lower following a strike of two weeks or longer. The implications of these findings for further theoretical analysis of strike activity are briefly discussed in the concluding section of the paper.

II. Data Description

The data in this study represent strike outcomes for contract negotiations of 253 bargaining pairs during the period from 1955 to 1979. Contract terms and strike information were originally collected from the Bureau of Labor Statistics, Current Wage Developments (CWD) by Wayne Vroman.^{4/} Vroman's data set contains information on some 300 collective bargaining situations covering 1000 or more workers. For purposes of this paper, however, I have focused on bargaining pairs with complete information for at least seven consecutive agreements. The sample contains 2543 contracts, or an average of ten contracts per bargaining pair.

Table 1 presents an overview of the data. With the exception of contracts for seven bargaining pairs in transportation, communications, and public utilities, the contracts are drawn from the manufacturing sector.^{5/} The coverage within two-digit manufacturing industries is irregular, and reflects Vroman's original interest in "pattern"

bargaining and wage determination. The data set contains multi-establishment master contracts (such as the Autoworkers' agreements with General Motors, Ford, and Chrysler; the Electrical Workers' (IUE) agreements with General Electric and Westinghouse; and the Steelworkers' agreements with U.S. Steel and Anaconda) as well as smaller single-establishment agreements (such as the Machinist's agreements with Morse Chain). The average number of workers covered by each contract is 13,000.

Because bargaining pairs choose different contract lengths within and across industries, the sample is "unbalanced" in the sense that there are different numbers of contracts for each bargaining pair. For example, contracts in the textile industry tend to be short. As a result, there are more than 14 contracts per bargaining pair over the 24-year sample period. Most contracts in the transportation equipment industry, on the other hand, are for three years. As a result, there are approximately 8 contracts per bargaining pair in the sample from this industry.

In addition to fixed-duration noncontingent contracts, the sample contains a wide variety of alternative contract forms, including fixed-duration contracts with contingent cost-of-living wage adjustment formulas (approximately 25 percent of the sample), contracts with scheduled wage-reopening provisions (approximately 7.5 percent of the sample), and contracts with contingent wage reopening provisions (approximately 1 percent of the sample).^{6/} Since wage-reopenings place the parties at risk of a strike, I have defined each reopening as a new "contract".

The impact of this commitment to limited renegotiation on the probability of strike outcomes is analyzed in the next section.

Finally, most long term labor contracts can be reopened at any time with the mutual consent of the parties, and many contracts are automatically extended past their scheduled expiration date unless one party or the other files a formal notice of intent to terminate the contract and put the pair at risk of strike.^{7/} Approximately 1 percent of contracts in the sample were reopened earlier than three months before their scheduled expiration date, and another 6 percent ran more than three months past their scheduled expiration date before a new contract was reached or a strike was declared. While these facts introduce some ambiguity into the notion of a contract expiration date, I have adopted the convention of dating expirations by the earlier of the expiration date of the preceding contract and the actual renegotiation date of the next contract, as reported in Current Wage Developments.^{8/}

The third and fourth columns of Table 1 contain information on the probability and duration of strikes by industry. For most of the contracts in the sample the source of strike information is the contract listing in Current Wage Developments.^{9/} For contract expirations after 1970, however, strike information is also available from an exhaustive listing of major strikes assembled by Sheena McConnell and Joseph Tracy.^{10/} McConnell and Tracy's data include local-issue strikes associated with the ratification of multi-establishment master contracts, as well as more widespread disputes. For comparability with the CWD definition of contract strikes, however, I restricted attention to disputes

involving at least 70 percent of workers covered by each contract.^{11/} The McConnell-Tracy strike listings contribute an additional 29 strikes to the 114 strikes reported in CWD for contracts negotiated after January 1970. These figures suggest that some 20 percent of strikes associated with contract renegotiations are missing from the CWD listings. The added strikes are shorter than the strikes in CWD (27 versus 48 days) but are more-or-less evenly distributed over time and across industries. Assuming that 20 percent of actual strikes are randomly missing from the CWD listing prior to 1970, the true probability of strikes in the data set is 14.1 percent.

The simple averages in Table 1 show considerable variation across industries in both the probability and duration of strikes. Strikes are more likely in durable than nondurable manufacturing, and are most frequent in the rubber and non-electrical machinery industries.^{12/} There is a weak positive rank-order correlation across industries between strike probabilities and the conditional duration of strikes (.21) although the correlation falls to zero if the apparel industry is ignored. The longest strikes are in lumber and wood products, primary metals, and fabricated metals, while the shortest strikes are in apparel, petroleum refining, and instruments.

For purposes of a longitudinal analysis, it is convenient to work with a fixed number of contracts (or contract expirations) per bargaining pair. To this end, I have extracted a balanced sample of contracts based on the six most recent negotiations for each bargaining pair. Strike probabilities and durations for this sample of 1518

contracts are presented in the right-hand columns of Table 1. The pattern of probabilities and durations across industries is similar between the total sample and the balanced subsample, although strike probabilities are somewhat higher in the subsample. Part of the explanation for this increase in strike probabilities is the fact that the overall sample contains a larger fraction of contracts from the 1950's and early 1960's, when strike probabilities were relatively low. By the same token, since relatively fewer of the contract expirations in the balanced sample occurred before 1970 (39.7 percent versus 63.2 percent), under-reporting of strikes is less significant. Assuming that 20 percent of strikes prior to 1970 are unreported, the true probability of strikes in the balanced sample is 17.8 percent.

Figure 1 presents estimated annual strike probabilities for 1961-1979 based on the balanced sample of contracts. Actual probabilities and strike durations by year are recorded in Appendix Table 1. The figure shows both the actual probability of strikes associated with contract expirations in each year, and an adjusted probability that controls for the industry composition of contract expirations. The adjusted probabilities represent estimated year effects from a linear probability model that includes two-digit industry effects. Both series show very low strike probabilities in the early 1960's, followed by sharply higher strike probabilities from 1965 to 1970, and lower but widely-varying strike probabilities throughout the 1970's. Since the focus of this paper is on the longitudinal structure of strike activity, in the empirical analysis reported below I control for time-varying

aggregate strike propensities with a series of year-effects. Experiments with both ordinary and fixed-effect probability models suggest that the time-varying component of strike activity is well represented by a simple four-step function, with steps at 1965, 1971, and 1975. In particular, this function captures the relatively low strike probabilities observed in the early 1960's, and the relatively high strike probabilities from 1965 to 1970.^{13/}

III. The Effect of Contract Characteristics on Strike Incidence

In this section I analyze the effect of three characteristics of labor contracts on the probability of strikes: the seasonal timing of contract expirations; the explicit provision of a limited contract reopening; and the length of time between successive negotiations. Since these characteristics are determined by the bargaining parties with little or no outside interference, evidence that they effect strikes raises the interesting theoretical question of how the parties choose among alternative contract forms. The pattern of choices among contract characteristics, on one hand, and the pattern of resulting strike probabilities, on the other, provide useful data for evaluating alternative theories of strike activity.

From a statistical point of view, however, estimation of the effects of various contract characteristics on the probability of strikes is complicated by the nonrandom incidence of these characteristics across industries and bargaining pairs. The fact that strike probabilities are relatively low among contracts that expire in December, for example, does not imply that a bargaining pair who traditionally nego-

tiate in June could reduce the probability of a work stoppage by scheduling their negotiations in December. I use three alternative estimation strategies to control for heterogeneity in strike probabilities across bargaining pairs. The first strategy is to model heterogeneity across industries using a series of two-digit industry effects. The second estimation strategy is a conditional logistic regression -- the direct analogue of a fixed effects estimator for logistic probability models. The third strategy is a first-differenced version of the linear probability model. In each case, the estimator accounts for random under-reporting of strikes among expirations prior to 1970.

a. Seasonality of Expirations

Table 2 presents estimates of the effect of alternative expiration months on the probability of strikes. As a benchmark, the first column contains estimated month-effects from a logistic regression with no controls for heterogeneity. Actual strike probabilities by month and the monthly distribution of expirations are recorded in Appendix Table 2. In order to incorporate the assumption that 20 percent of strikes associated with expirations before 1970 are unreported, the following modified logistic likelihood is maximized:

$$(1) \prod_{i,t} p_{it}^{y_{it}} (1 - p_{it})^{1-y_{it}},$$

where $y_{it} = 1$ if the t^{th} negotiation for the i^{th} pair resulted in a strike, and $y_{it} = 0$ otherwise, and

$$(2) p_{it} = (1 - .2D_{it}) p_{it}^*$$

where p_{it}^* is the predicted probability of a strike in the t^{th} negotiation for the i^{th} pair in the absence of any under-reporting, and

D_{it} is an indicator variable for expirations prior to January 1970. In the case of the estimated model in the first column of Table 2,

$$\text{logit}(p^*_{it}) = \sum_{m=1}^{12} Z_{it}^m \beta_m ,$$

where $\text{logit}(x) = \log(x/(1-x))$, Z_{it}^m is an indicator variable equal to one if the t^{th} expiration for the i^{th} pair occurred in month m , and $(\beta_1, \dots, \beta_{12})$ is a vector of month effects.

The estimates show significantly different strike outcomes depending on the expiration month of the preceding contract. Average strike probabilities are relatively high in March, August, October, and November, and relatively low in January, April, May, and December. The likelihood-ratio test associated with the hypothesis that strike probabilities are constant across months is reported in the last row of the table, and is highly significant.

The second column of Table 2 introduces two-digit industry controls and a four-step function of time as additional covariates of strike activity.^{14/} Comparing the estimates in the first and second columns, there are only small differences between the estimated month effects. These results suggest that the monthly pattern of strike probabilities is not simply an artifact of the industry distribution of contract expirations. In fact, the monthly pattern of strike probabilities is very similar with and without industry controls.

The third column of Table 2 contains estimated month effects from a conditional logit model of strike incidence. The starting point for this model is a logistic specification of strike probabilities with fixed effects for each bargaining pair:

$$(3) \quad \text{logit}(p^*_{it}) = \alpha_i + x_{it}\beta ,$$

where p^*_{it} is the probability of a strike in the t^{th} negotiation for the i^{th} pair, and x_{it} is a vector of time-varying covariates (for example, year and month effects). The conditional logit model describes the likelihood of each pair's sequence of strike indicators

$(y_{i1}, y_{i2}, \dots, y_{i6})$ conditional on the total number of strikes observed for the pair $(\sum_t y_{it})$. Using (3) the conditional likelihood can be written as

$$P(y_{i1}, \dots, y_{i6} \mid \sum_t y_{it}) = \frac{\exp(\sum_t y_{it} x_{it} \beta)}{\sum_k \exp(\sum_t d_{kt} x_{it} \beta)} ,$$

where the index k in the denominator runs over alternative sequences of strike outcomes $d_k = (d_{k1}, d_{k2}, \dots, d_{k6})$ with $\sum_t y_{it}$ strikes in 6 contracts. The usefulness of the conditional likelihood approach arises from the observation that the conditional likelihood is independent of the fixed effect α_i . This is a consequence of the fact that the number of observed strikes is a sufficient statistic for α_i : conditioning on the number of strikes therefore eliminates the fixed effect from the likelihood.^{15/}

The assumption that strikes are under-reported in the early part of the data set introduces a serious complication for the conditional logit model. If the fixed-effects specification of true strike probabilities (3) is combined with the under-reporting model (2), then the conditional likelihood is no longer independent of the fixed effects. The fixed-effects specification can be combined with an approximation to the

under-reporting model, however, that leads to a workable conditional likelihood. In particular, combining (2) and (3), the implied probability of an observed strike in the fixed-effects model is

$$p_{it} = (1 - .2D_{it}) \frac{e^{\alpha_i + x_{it} \beta}}{1 + e^{\alpha_i + x_{it} \beta}} .$$

Consider the approximation

$$(4) \quad p_{it} \approx \frac{e^{\alpha_i + x_{it} \beta - \gamma D_{it}}}{1 + e^{\alpha_i + x_{it} \beta - \gamma D_{it}}} ,$$

where γ is chosen a priori to approximate the behavior of the true probability statement at the sample mean probabilities. Equation (4) is clearly amenable to a conditional likelihood formulation. As it happens, the approximation in (4) is also relatively precise, at least for values of the individual effects that give rise to predicted strike probabilities between 2 and 30 percent. I therefore use the approximate correction for under-reporting implied by equation (4) to estimate the conditional logit strike model.^{16/}

Conditional logit estimates of the month effects are presented in the third column of Table 2. These estimates are less precise than the corresponding conventional logit estimates, and indicate lower strike probabilities in February and March, and higher strike probabilities in August. The conditional logit estimates suggest that the monthly effects can be divided into just two "seasons": a low strike probability season from December to May, and a high strike probability season

from June to November. The likelihood ratio test for the hypothesis that strike probabilities are constant from December to May, and from June to November, has a marginal significance level of .38. The estimated month effect for June-November in this simple two-season model is .96, with a standard error of .28.

The fourth column of Table 2 presents estimated monthly effects from a first-differenced version of the linear-probability specification

$$(5) \quad p_{it}^* = \alpha_i + x_{it} \beta .$$

While this specification suffers from the well-known objection that the predicted probabilities can lie outside the unit interval, it is particularly convenient for handling fixed effects, since (ignoring under-reporting) it implies the linear regression equation

$$(6) \quad \Delta y_{it} = \Delta x_{it} \beta + \xi_{it} ,$$

where $\Delta y_{it} = y_{it} - y_{it-1}$ represents the change in strike outcomes between the $(t-1)^{st}$ and t^{th} negotiation, $\Delta x_{it} = x_{it} - x_{it-1}$ represents the vector of differences in the covariates, and ξ_{it} can be interpreted as a residual. Since Δy_{it} can take on only three values (-1, 0, 1), the residual ξ_{it} is conditionally heteroskedastic. Accordingly, the estimated standard errors in the fourth column of Table 2 have been corrected for heteroskedasticity using White's (1980) procedure.

The assumption that strikes prior to 1970 are under-reported can be easily accommodated within the linear probability specification.

The combination of equation (5) and the under-reporting model (2) implies that the actual probability of an observed strike outcome is

$$p_{it} = (1 - .2D_{it}) (\alpha_i + x_{it} \beta) ,$$

where, as before, $D_{it} = 1$ if the t^{th} expiration occurred prior to 1970. For consecutive expirations after 1970, $D_{it} = D_{it-1} = 0$ and the implied regression for Δy_{it} is (6). For consecutive expirations before 1970, $D_{it} = D_{it-1} = 1$ and the implied regression for Δy_{it} is

$$(6a) \quad 1.25 \Delta y_{it} = \Delta x_{it} \beta + \xi_{it} .$$

For consecutive expirations where the $(t-1)^{\text{st}}$ expiration occurs before 1970 and the t^{th} expiration occurs after 1970, the implied regression function is

$$(6b) \quad y_{it} - 1.25y_{it-1} = \Delta x_{it} \beta + \xi_{it} .$$

The estimates in the fourth column of Table 2 are obtained by combining these three regression equations, depending on the actual expiration dates of the contracts for each bargaining pair.

The estimated coefficient from the linear probability specification can be made roughly comparable to the estimates from the conditional logit model by scaling the linear probability coefficients by a factor of $[\bar{p}(1-\bar{p})]^{-1}$, where $\bar{p}$ is the sample average probability of strikes. Assuming an average strike probability of 15 percent, the implied scaling factor is 8. Scaled coefficients corresponding to the linear probability estimates are presented in the last column of Table 2.

The estimated month effects from the linear probability model are generally similar to the estimates from the conditional logit model, although the point estimates for March, May, August and December differ somewhat between the two specifications. Both models suggest that strike probabilities are significantly higher for contract expirations in summer and fall than in winter and spring. A test that the month effects from the linear probability specification are constant from December to May and from June to November has a marginal significance level of .37 (virtually identical to the significance level of the test on the conditional logit coefficients). Assuming a two-season model, the linear probability specification yields an estimated increase in the probability of strikes between June and November of .15, with a standard error of .05.

Both the conditional logit and differenced linear probability models therefore suggest an important distinction between contract expirations in winter and spring, on one hand, and summer and fall, on the other. The consistently higher incidence of strikes in September and October, compared to April and May, suggest that the seasonal pattern of strike probabilities is not related to weather patterns. A comparison of seasonal strike patterns among bargaining situations that can unambiguously be assigned to either Northern or Southern and Western states leads to a similar conclusion. There is no evidence that month effects for these two groups differ systematically from the overall effects in Table 2.^{17/}

b. Reopeners and Contract Duration

Table 3 presents estimates of the effects of two other contract

characteristics on strike outcomes: the length of time since the last negotiation; and limited reopener provisions. The estimates are obtained from models that exclude seasonal effects, although estimates from models that control for contract expiration month are very similar. The distribution of reopener provisions and contract lengths in the contract data is more-or-less independent of expiration month.

The first row of Table 3 presents the estimated change in the log-odds of a strike for negotiations conducted under a reopening clause of an ongoing agreement. A typical reopener clause specifies that the wage rate is to be renegotiated at some date in the future (usually the first or second anniversary date of the contract), with the proviso that other aspects of the contract will remain unchanged.^{18/} On the basis of information in the preceding contracts, 91 reopeners were identified, mostly in the textile, apparel, and chemical industries.^{19/}

The second row of the table presents the estimated effects of the length of the preceding contract (in months) on the log-odds of a strike. In cases where new contract or a strike was reached prior to the expiration date of the preceding contract, the length of the preceding contract is defined as the period of time between its negotiation date and the date of the next contract or strike.

The first column of the table contains the estimated effects of these two variables with no heterogeneity controls. Strike probabilities are significantly lower in reopening situations, and significantly higher following the expiration of a longer contract.^{20/} The second column introduces two-digit industry controls and a step-function of

time into the logistic regression. The estimated effect of a limited reopening is unchanged, while the effect of previous contract duration is reduced slightly. Finally, the right-hand columns of Table 3 present alternative fixed-effects estimators of the impact of reopening provisions and contract length on strike probabilities. The estimated contract length effects are similar to the estimates with industry-level controls, while the estimated effects of reopening provisions are smaller, particularly in the linear probability specification. The point estimates imply that strike probabilities are 9-11 percent lower in reopening situations, and about 6 percent higher for negotiations following a three-year, as compared to a two-year contract.^{21/}

These results suggest two conclusions. First, the commitment to limited negotiations implied by a reopener provision may actually reduce the probability of disputes--perhaps by reducing the number of issues on the bargaining table. Second, increases in contract length, while reducing the number of opportunities for disputes, lead to a higher probability of disputes in each negotiation. In fact, the estimates in Table 3 suggest that the expected number of disputes per period of time is approximately constant, whether bargaining occurs at one-, two- or three-year intervals.^{22/}

IV. The Longitudinal Structure of Strike Activity

This section presents and analyzes the pattern of strike incidence over time within bargaining pairs. Using information from six consecutive negotiations for each pair, I first describe the distribution of bargaining pairs over alternative strike histories. I then compare this

distribution to the distributions implied by three alternative models: a conventional logistic regression model; a conditional logit model; and a random-effects logit model in which the distribution of random effects is described by a discrete distribution. Finally, the random effects model is used to address the question of whether lagged strike activity effects the probability of strikes in subsequent negotiations.

Table 4 summarizes the distribution of bargaining pair over the 64 (2^6) possible strike histories. The first column of the Tables describes the relevant strike history, and the second column gives the number of bargaining pairs with that history. The first row of the Table, for example, pertains to the "000000" history, which represents the occurrence of no strikes in six negotiations. One hundred and eighteen pairs, or 47 percent of the sample, fall into this category. The second through seventh rows of the table contain the six alternative histories with a total of one strike in six negotiations. Bargaining pairs with one strike are rather unevenly divided among the alternatives, reflecting the influence of changing strike probabilities over time.^{23/} A total of 66 pairs, or 26 percent of the sample, report one strike in six negotiations.

The 9th through 23rd rows of Table 4 describe strike histories with a total of two strikes in six negotiations. Among the 16 possible outcomes with two strikes, 15 are present in the data, with the 000011 outcome being most frequent.^{24/} There is some indication from the relative frequencies of the various alternatives that strike probabilities are positively correlated over time. The six cells with consecutive strikes

(in rows 13-19 of the table) account for 50 percent of the observations with a total of two strikes.

For the sake of brevity, I have not displayed the individual histories with three, four, or five strikes in Table 4. The total numbers of bargaining pairs with these histories are presented in rows 25-27 of the table. Since strikes are relatively rare, the number of bargaining pairs with three or more strikes is low (31 pairs in all) and the individual cells are sparse. There are no pairs with five strikes, and only one pair (the Autoworkers and Allis-Chalmers) with six strikes. Among histories with three strikes the most likely individual cell is 010101, with three observations. Among histories with four strikes, only one cell has more than one observation -- the 011110 strike history with two observations.

The third column of Table 4 contains the predicted number of observations for each strike history based on a conventional logit model with two-digit industry effects and a four-step time function.^{25/} According to this model, consecutive strike outcomes for a given pair are independent, conditional on the industry and time effects. If $\hat{p}_{it}$ represents the predicted probability of a strike for pair i in negotiation t , then the probability that pair i has a sequence of strike indicators $(y_1, \dots, y_6)$ is

$$\prod_t \hat{p}_{it}^{y_t} (1-\hat{p}_{it})^{1-y_t},$$

and the expected fraction of the sample with a given history is

$$\frac{1}{N} \sum_{i=1}^N \prod_t \hat{p}_{it}^{y_t} (1-\hat{p}_{it})^{1-y_t},$$

where $N=253$ is the number of bargaining pairs in the sample.

Standard errors for the difference between the predicted and actual fractions of the sample with a given history can be constructed as follows. Let I_i represent a 64-element vector of indicators, whose j^{th} row is equal to 1 if bargaining pair i has strike history j , and 0 otherwise. If the probabilities p_{it} that the i^{th} pair has a strike in the t^{th} negotiation are known, then I_i has a multinomial distribution with mean vector Φ_i and covariance matrix Ω_i that are simple functions of p_{it} . Since the strike histories are exhaustive and mutually exclusive, the covariance matrix Ω_i is singular, with rank 63 if $p_{it} > 0$ for all t .

For an independent sample of N bargaining pairs, the fractions of the sample with each strike history are given by the rows of the vector

$$\bar{I} = \frac{1}{N} \sum_i I_i.$$

The expected value of $\bar{I}$ is

$$E(\bar{I}) = \frac{1}{N} \sum \Phi_i$$

and the covariance matrix of $\bar{I}$ is $N^{-1}\Omega$, where

$$\Omega = \frac{1}{N} \sum_i \Omega_i.$$

An appropriate standard error for the difference between the actual fraction of the sample with the j^{th} strike history and the expected fraction given the probabilities p_{it} is $N^{-\frac{1}{2}} \Omega(j,j)$, where $\Omega(j,j)$ represents the (j,j) element of Ω .

Using these expressions, an overall goodness-of-fit statistic for a

particular model of the probabilities p_{it} may be constructed from the vector of differences between the actual and predicted probabilities of each strike history. If the model is correct, then the quantity

$$N(\bar{I} - \frac{1}{N} \sum_i \phi_i)' \Omega^{-} (\bar{I} - \frac{1}{N} \sum_i \phi_i)$$

has an asymptotic chi-squared distribution with 63 degrees of freedom, where Ω^{-} denotes a generalized inverse of the matrix Ω .

These arguments break down when the probabilities p_{it} are derived from a set of estimated parameters, rather than specified a priori. Recently however, Moore (1977) and Heckman (1984) have presented a general expression for the covariance matrix of the difference in predicted and actual sample probabilities for a set of predetermined cells, when the predicted probabilities are based on sample-estimated parameters. Using these formulas, a corrected covariance matrix for the vector of differences in the predicted and actual probabilities of the various strike histories is presented in the Appendix, assuming that the predicted probabilities $\hat{p}_{it}$ are generated by a conventional logit model. The t-statistics reported in the third-column of Table 5 are based on this corrected covariance matrix. T-statistics based on the uncorrected covariance matrix are very similar.

Overall, the conventional logit model provides a remarkably good description of the distribution of strike histories. The goodness-of-fit statistic in the last row of the table is only marginally significant.^{26/} The t-statistics for the individual cells are also generally insignificant, with the notable exception of the 000000, 000011, and 111111 cells. The model underpredicts the fraction of the

sample with no strikes and places a relatively low likelihood on the occurrence of six strikes in six negotiations.

The fourth column of Table 4 presents the predicted numbers of observations with each strike history based on a conditional logit model with time effects. In the conditional logit model, the likelihood of each observation is conditioned on the total number of strikes for that observation. Rather than providing a complete model of the distribution of bargaining pairs across alternative strike histories, the conditional logit model describes the distribution of bargaining pairs within sufficiency classes (groups of histories with the same total number of strikes). For this reason, the conditional logit model "predicts" the number of observations with each strike total exactly.

In order to derive standard errors for the predicted number of observations with each history, observe that a bargaining pair with a given number of strikes (say T) can fall into any one of $k_T = \binom{6}{T}$ cells. Let I_i be a k_T - dimensional indicator vector whose j^{th} row is equal to 1 if the i^{th} pair has the j^{th} history, and zero otherwise. The probability that the i^{th} pair has the j^{th} history, conditional on the fact that i belongs to the T^{th} sufficiency class, can be calculated from an expression analogous to the expression for the conditional logit likelihood. Ignoring parameter estimation, then, it is straightforward to derive the expected value of I_i (say ϕ_i) and the covariance matrix of I_i , Ω_i . Among the N_T observations with T strikes in 6 negotiations, the expected fractions that fall into each of the K_T alternatives are $\frac{1}{N_T} \sum_i \phi_i$, and the covariance matrix of the difference between the actual and expected fractions is $N_T^{-1} \Omega_T$, where

$\Omega_T = \frac{1}{N_T} \sum_i \Omega_i$. There are four such vectors of expected sample fractions, and four associated covariance matrices, corresponding to the sufficiency classes with 1, 2, 3 and 4 strikes in 6 negotiations. (The sufficiency class with 5 strikes is empty). These expected sample fractions, and the associated variances, are used to compute the t-statistics in the fourth row of Table 4.^{27/}

Within the T th sufficiency class, an appropriate test statistic for the goodness-of-fit of the conditional logit model (ignoring parameter estimation) is

$$N_T \left[\frac{1}{N_T} \sum_{i \in T} (I_i - \Phi_i) \right], \quad \Omega_T^{-1} \left[\frac{1}{N_T} \sum_{i \in T} (I_i - \Phi_i) \right],$$

where Ω_T^{-1} denotes a generalized inverse of the matrix Ω_T . The sum of these statistics across sufficiency classes forms a test for the overall fit of the conditional logit model and is reported at the bottom of the table.

Relative to the conventional logit model, the conditional logit model provides a slightly better fit to the distribution of strike histories, although the 000011 cell is still a significant outlier. The conditional logit model also under-predicts the fraction of bargaining pairs with two strikes who have consecutive strikes (19 actual versus 14 predicted). Overall, however, the fixed-effects specification provides a relatively successful model of the distribution of strike histories.

The last column of Table 4 presents the predicted distribution of strike histories associated with an alternative random-effects logit

model. In this model, the probability of a strike in the t^{th} negotiation for the i^{th} bargaining pair is given by

$$p_{it}^*(\alpha_i) = \frac{e^{\alpha_i + x_{it}\beta}}{1 + e^{\alpha_i + x_{it}\beta}},$$

where x_{it} represents a vector of year indicators, and the individual effect α_i is assumed to be drawn from a discrete distribution with a small number of mass points (two in the model presented in Table 4).^{28/} The probability that the i^{th} pair has a sequence of strike indicators $(y_1, y_2, \dots, y_6)$ is given by

$$(7) \sum_k \left\{ \prod_t p_{it}^*(\alpha_k)^{y_t} (1-p_{it}^*(\alpha_k))^{1-y_t} \right\} \phi_k,$$

where ϕ_k represents the probability of the k^{th} value of the individual effect. The parameters β , the points of support α_k , and the probabilities ϕ_k are estimated jointly by maximizing the likelihood of the observed strike histories for each bargaining pair. Using equation (7) it is straightforward to derive expected sample fractions and associated standard errors for each of the 64 possible strike histories. In this paper I compute the standard errors of the expected sample fractions and the goodness-of-fit statistic for the overall table of strike histories without correcting for the estimated parameters in the random-effects model. Corrected covariances may be readily derived following the approach suggested by Moore (1977) and Heckman (1984), although experience with the conventional logit model suggests that the corrected and uncorrected covariances are fairly similar.

The results in the fifth column of Table 4 show that in spite of

its simplicity, the random-effects model provides a relatively accurate description of the distribution of strike histories. A random-effects model with two mass-points has 14 fewer parameter than the conventional logit model with industry effects, yet yields a significant improvement in the fit to the strike-history table. The only major outliers in the table are the 000011 cell and the 111111 cell. If a third mass-point is added to the model, the log-likelihood improves insignificantly (from -637.2 to -636.3), but the goodness-of-fit statistic for the strike table falls to 55.1, mainly as a result of an increase in the predicted probability of the 111111 cell.

Although the random-effects specification with 2 or 3 mass-points does not fit the strike-history table as well as the conditional logit model, the random-effects model has an important advantage over the conditional logit model. Specifically, the random-effects model can easily accommodate both individual heterogeneity and state-dependence in successive strike outcomes. The conditional logit model, by comparison, assumes that individual strike probabilities are independent, conditional on the individual effect.

The question of state-dependence in strike outcomes is addressed in Table 5 using the random-effects logit formulation. As a benchmark, the first column of the Table summarizes the goodness-of-fit of the random-effects model with no state-dependence. The second column of Table 5 summarizes a simple first-order model of state-dependence in strike probabilities. The probability of a strike in the t^{th} negotiation for the i^{th} pair (ignoring under-reporting) is given by

$$(8) \quad \text{logit}(p_{it}^*) = \alpha_i + x_{it}\beta + \delta y_{it-1} ,$$

where α_i represents the randomly distributed individual effect, x_{it} represents a vector of year indicators, and y_{it-1} represents the lagged strike outcome: $y_{it-1} = 1$ if there was a strike in the previous negotiation, and 0 otherwise. The parameter δ measures the extent of state dependence in strike outcomes. If $\delta > 0$, then strikes are more likely to follow one-another, while if $\delta < 0$ then the probability of a strike is reduced by an earlier dispute.

The estimates in the second column of the table show very little evidence of state-dependence.^{29/} The estimate of δ is .13, with a standard error of .23. Neither the log-likelihood nor the goodness-of-fit for the strike history table is significantly improved by allowing for state-dependence. In fact, the goodness-of-fit statistic is actually worse.^{30/}

Close inspection of the individual strike histories, however, suggests that the degree of state-dependence varies, depending on the severity of the preceding dispute. Some simple evidence is presented in Table 6, which describes the probability of strikes conditional on preceding strikes of various lengths. The first row of the table summarizes the overall likelihood of disputes for each previous strike outcome. Among contracts that settled peacefully in the previous negotiation, the probability of strikes is 13.9 percent. Among contracts preceded by a short strike (two weeks or less) the likelihood of a subsequent dispute is over 40 percent, while among contracts pre-

ceded by a long strike (longer than two weeks) the likelihood of a subsequent dispute is about 20 percent. A simple model of heterogeneity in strike probabilities, on the other hand, suggests that the probabilities in the 5 right-hand columns of Table 6 should all be equal.^{31/}

In order to further investigate state dependence in strike outcomes, I divided the preceding strikes into several categories, depending on their duration, and fitted a series of models of the form

$$(9) \quad \text{logit} (p_{it}^*) = \alpha_i + x_{it} \beta + \sum_k \delta_k y_{it-1}^k,$$

where y_{it-1}^k is an indicator variable for a strike in the k^{th} duration class in negotiation $t-1$. Several of these models are summarized in Table 5. The model in the third column of the table, for example, considers two categories of strikes: 1-14 days; and 15 days or longer. As the results in Table 6 suggest, the estimates of the effects of these two kinds of strikes are remarkably different. Whereas strikes of less than 15 days (some 30 percent of the total) significantly increase the probability of a subsequent strike, longer strikes reduce the probability of subsequent strikes. Further experiments with finer duration classes, reported in the fourth and fifth columns of the table, suggest that a two-state model of lagged strikes is adequate to describe the data.

Finally, to ascertain that these results are not a consequence of imperfect heterogeneity controls, I estimated random-effects models with three or four mass-points, and with grouped industry effects. The addition of one or two extra mass-points did not significantly improve the likelihood of the sample or alter the estimate of the lagged strike

effects, although it did improve the goodness-of-fit of the model to the strike-history table. The effect of adding industry-group variables is summarized in the last column of Table 5. Two-digit industries are grouped according to their estimated effects in a conventional logit model of strike incidence. The four groups include a low strike probability group (Apparel, Lumber and Wood), a moderately high strike probability group (Primary Metals, Fabricated Metals, Electrical Machinery, Transportation Equipment), a very high strike probability group (Rubber and Non-Electrical Machinery), and a base group (all other industries). The addition of the grouped industry variables reduces the point estimate of the effect of a longer strike in the previous negotiation from -.40 (with a standard error of .28) to -.82 (with a standard error of .39). The estimated effect of a short strike in the previous negotiation is only slightly reduced. These results confirm that lagged strike outcomes affect current strike probabilities, with shorter strikes increasing the likelihood of future disputes, and longer strikes reducing the likelihood of future disputes.

V. Longitudinal Analysis of Strike Duration and Incidence

In view of the impact of strike duration on subsequent strike outcomes, this section presents a more detailed analysis of the duration and incidence of strikes over time within bargaining pairs. The analysis is based on a model that describes strike outcomes in terms of only three possibilities: no strike; a strike for less than two weeks; and a strike for longer than two weeks. While this simple three-outcome model of strike activity is obviously incomplete, the results of the

last section suggest that the distinction between short and long strikes is very useful in longitudinal models of strike incidence. The three-outcome case is therefore a natural starting point for studying the joint determination of strike incidence and duration over time.

In order to describe the distribution of bargaining outcomes between no strikes, short strikes, and long strikes, consider a pair of indicator variables y_{it} and z_{it} , where $y_{it} = 1$ if a strike occurs in the t th negotiation for the i th bargaining pair, and 0 otherwise, and $z_{it} = 1$ if the strike lasts for longer than two weeks and 0 otherwise. Let p_{it} and q_{it} represent the probabilities of a strike and a long strike, respectively, ignoring under-reporting of strikes. Suppose that

$$(10a) \quad \text{logit}(p_{it}^*) = \alpha_i + x_{it} \beta + \delta_1(y_{it-1} - z_{it-1}) + \delta_2 z_{it-1}$$

$$(10b) \quad \text{logit}(q_{it}^*) = \gamma_i + x_{it} \theta + \rho_1(y_{it-1} - z_{it-1}) + \rho_2 z_{it-1},$$

where α_i and γ_i are individual effects and x_{it} is a vector of control variates (for example, year and/or industry effects). Finally, assume that y_{it} and z_{it} are independent, conditional on α_i , γ_i , x_{it} , y_{it-1} , and z_{it-1} .^{32/} The parameters δ_1 and δ_2 measure the effect of a short or long strike in negotiation $t-1$ on the probability of a strike in negotiation t .^{33/} The parameters ρ_1 and ρ_2 measure analogous effects on the probability of a strike continuing for more than 2 weeks.

Given values for α_i and δ_1 , and starting values for the strike outcomes, equations (10) can be used to calculate the likelihood of an

observed sequence of strike outcomes over time. The two equations describe a first-order Markov model of transitions between three alternative states: no strikes; strikes of less than 14 days; and strikes of more than 14 days. Following the approach in the previous section, I treat the joint distribution of α_i and δ_i as a discrete distribution with a small number (2 or 3) of mass points. I also condition the estimation and inference on the observed strike outcomes for each bargaining pair in the immediate pre-sample period.

The average transition matrix over all bargaining pairs and all negotiations is presented in Table 7. The columns report the observed probabilities of alternative strike outcomes, conditional on the outcome of the preceding negotiation. For example, 4.5 percent of all negotiations that follow a peaceful settlement result in a strike of two weeks or less, while 9.4 percent result in a strike of longer than two weeks. The fourth row of the table presents the conditional probability of a long strike, given that a dispute occurs. Although previous strike outcome have a strong impact on the likelihood of subsequent disputes, the distribution of strikes between short and long strikes is less affected by previous disputes. The table suggests that longer strikes are less likely if the previous contract settled peacefully, although the effect is relatively small.

Table 8 presents estimation results for the trinomial outcome model of strikes implied by equations (10a) and (10b). The first column of the Table presents estimates based on a two mass-point bivariate distribution of individual effects. Apart from individual effects, the only

covariates of strike activity are a step-function of time. For simplicity, the time effects in equations (10a) (the incidence equations) and (10b) (the duration equation) are restricted to be proportional:

$\theta = \xi\beta$. The estimated proportionality constant ξ is reported in the fifth row of the table.

The estimated effects of previous work stoppages on the probability of a strike are presented in the first two rows of the table. The estimates are very similar to the estimates in Table 5, with about the same level of precision. The estimated effects of short and long strikes in the preceding contract on the probability of a long strike in the current negotiation are reported in the third and fourth rows of Table 6. The point estimates suggest that a long strike is most likely if there was a long strike in the previous contract, and less likely of the previous contract settled peacefully, although the estimates are relatively imprecise.

The second column of the table presents estimation results for a three mass-point model of the distribution of individual effects. Overall, the results are very similar to the two mass-point specification, and the likelihood of the sample is not significantly improved.

The hypothesis that lagged strikes do not affect the probability of short or long strikes is addressed in the third column of the table. Comparing the likelihood and parameter estimates to those in the second column, there is very little evidence against the hypothesis. While lagged strike activity apparently influences the probability of subsequent strikes, the probability of continuing a strike beyond two weeks

is not much affected by previous strike outcomes. Some caution is nevertheless required in interpreting these results, since the number of strikes in the data set is small, and strike duration is apparently a noisy phenomenon.^{34/} A larger data set will probably be required to fully analyze the determinants of strike duration, and particularly the effects of lagged duration on current strike durations.^{35/}

The fourth column of the table presents estimates for a two mass-point specification that includes grouped industry effects as well as time covariates in equations (10a) and (10b). Here I have restricted the time effects and industry effects to be proportional in the two equations, but I have allowed different factors of proportionality for the two types of effects. While the time effects are highly significant in the duration equation, the industry effects, which play a strong role in the strike probability equation, have very little effect on duration.^{36/} As in Table 5, the addition of grouped industry effects increases the absolute value of the estimated effect of a longer strike in the preceding negotiation. Overall, the results in Table 8 for the incidence equation are very similar to the results in Table 5. The results for the duration equation, while relatively imprecise, suggest that previous strike outcomes do not affect the relative likelihood of short or long strikes.^{37/}

Summary and Conclusions

This paper has presented evidence on two aspects of strike activity associated with the renegotiation of union contracts: the effects of endogenous contract characteristics on the probability of disputes; and

the variation in strike activity over time within bargaining pairs. While the evidence is based on a small sample of bargaining pairs drawn mainly from the manufacturing sector, the findings suggest a number of avenues for further theoretical and empirical research.

First, there is fairly strong evidence that the characteristics of collective bargaining agreements, including their duration, expiration month, and reopening status, affect the probability of disputes. Both cross-sectional and longitudinal estimation techniques confirm that strike probabilities are higher when the preceding contract is longer, or when it expires in summer or fall. These same techniques show that negotiations conducted under a limited reopening clause of an on-going agreement are less likely to result in strikes. In view of these findings, further research on the choice between alternative contract forms, on one hand, and optimal strike probabilities, on the other hand, may prove useful in discriminating among alternative theories of strikes. The evidence on seasonality of strike probabilities, for example, calls into question the hypothesis that strikes are simply costly accidents. If strikes serve no economic purpose, it seems clear that bargainers should choose contract expirations in months when strike probabilities are lowest (assuming that the costs of strikes are constant across months). In this data set, however, expirations are actually concentrated in months when strike probabilities are highest.^{38/}

Second, longitudinal analysis of strike outcomes suggests that strike probabilities are significantly affected by preceding strike activity. In particular, relative to strike probabilities after a

peaceful settlement of the most recent contract negotiation, strike probabilities are some 10 percentage points higher if the contract was settled after a 1 to 14 day strike, and some 5 to 7 percentage points lower if the contract was settled after a longer strike. The effects of previous strike outcomes on subsequent strike durations are less precisely estimated and are not significantly different between negotiations that ended peacefully and those that ended in shorter or longer strikes.

The finding of state-dependence in strike outcomes suggests a number of possible explanations, including models of imperfect information where the parties' beliefs about each others' willingness to endure future strikes are affected by current strike outcomes. The differential effects of short and long strikes on subsequent strike probabilities should provide useful evidence for choosing between the alternative explanations.

Footnotes

^{1/} Among the most recent studies are papers by Tracy (1986), Kennan (1985), Kennan (1986b), Gramm (1986), Fudenberg, Levine, and Ruud (1983), and Gunderson, Kervin, and Reid (1986). Earlier studies on micro-data include Kennan (1980) and Farber (1978). There is also a substantial empirical literature devoted to aggregate strike activity. See in particular Ashenfelter and Johnson (1969), Pencavel (1970), Kaufman (1982), and Abbott (1985). A more complete bibliography, together with a survey of theoretical and empirical research, is presented in Kennan (1986a).

^{2/} See in particular the papers by Hayes (1984), Fudenberg, Levine and Ruud (1983), and Tracy (1986). The theoretical analysis in the latter two studies is closely related to the work of Crampton (1982).

^{3/} For example, Tracy (1986) tests the hypothesis that greater variation in firm-specific securities prices increases the likelihood of strikes.

^{4/} I am grateful to Wayne Vroman for making this data available.

^{5/} The nonmanufacturing bargaining pairs are: Class I Railroads and the Railroad Engineers; Class I Railroads and the Brotherhood of Railroad and Airline Clerks; Trucking Employers Incorporated and the Teamsters (National Master Freight Agreement); New York Shipping Association and the Longshoremen; Pacific Maritime Association and the Longshoremen; United Airlines and the Machinists; and ATT (Longlines Division) and the Communications Workers.

^{6/}A typical reopening clause states that the parties will open an ongoing agreement for purposes of renegotiating a limited number of contract issues (often only wages): see the Bureau of National Affairs (1985). Most reopener clauses specify a future date for reopening. A small number specify reopening contingent on a specific event (for example, discontinuation of wage and price controls; wage adjustments in contracts at other firms; inflation rates above a certain maximum).

^{7/}See the Bureau of National Affairs (1985). A small number of contracts in the textile industry specify an indefinite contract duration, although these contracts characteristically reopen at regular 12 or 24 month intervals.

^{8/}An interesting issue for further research is the question of when and why the parties continue to operate under the terms of the old contract.

^{9/}The original source of the strike information in CWD is apparently the contract report filed by companies at the request of the BLS, in connection with their ongoing analysis of wage changes in union contracts with 1000 or more workers. For contracts negotiated between 1960 and 1970 I checked the CWD strike information against contract information reported in the Monthly Labor Review's monthly summary of recent developments in industrial relations. I found only 3 examples of strikes reported in the Review that were not recorded in CWD.

^{10/}McConnell and Tracy's strike listings were assembled from three sources: a weekly BLS in-house newsletter entitled "Industrial Relations Facts"; a BLS data tape listing strikes recorded from

published newspaper reports; and a Bureau of National Affairs' data tape listing strikes from published media and other sources.

^{11/}The precise definition of a strike in Current Wage Developments is unclear. It is clear, however, that the BLS does not report local-issue strikes in CWD. The question of whether local-issue strikes should be treated differently from more general disputes in an important issue for further research.

^{12/}A comparison of the industry pattern of strikes among union contracts in U.S. and Canadian manufacturing shows a different rank-ordering of industry strike probabilities in the two countries. Using data for 1290 contracts in manufacturing from Labor Canada's Collective Bargaining tape, I tabulated average industry strike probabilities for 19 two-digit industries that are directly comparable to U.S. industries. The rank-order correlation coefficient of the industry probabilities in the two countries is .06.

^{13/}Susan Vroman (1986) has investigated the effects of a variety of macroeconomic variables on the probability of strikes in this data set. Her results suggest that unemployment rates, past inflation rates, and wage and price controls all effect the probability of strikes. She does not, however, compare the explanatory power of her probit regressions to regressions with unrestricted year effects. The likelihood ratio test statistic of the four-step time function against an unrestricted specification of the year effects in a logistic regression that includes 2-digit industry effects is 21.57 with 16 degrees of freedom. The probability value of this statistic is 16 percent.

^{14/}The main effect of the adjustment for unreported strikes is to change the estimated year effects in the specification for p^*_{it} , the probability of strikes in the absence of under-reporting. Appendix Table 3 presents estimated strike probability equations that include industry effects and time-period effects with and without the correction for under-reporting prior to 1970. Comparing these two specifications, the year effects for time periods before 1970 are .20 to .25 higher in the corrected model. The industry effects are similar between specifications.

^{15/}See Chamberlain (1980) and the references cited there. This sufficiency result depends critically on the functional form of the logistic distribution.

^{16/}Assuming that the average probability of strikes is 15 percent, the implied value for γ in equation (4) is .256. If the true probability of strikes is 2.5 percent, the probability of observed strikes with under-reporting is 2.0 percent and the probability predicted by equation (4) with $\gamma = .256$ is 1.97 percent. If the true probability of strikes is 30 percent, the probability of observed strikes with under-reporting is 24 percent and the probability predicted by equation (4) is 24.86 percent. As a check on the effects of this approximation, I used the same approximation to correct the ordinary logistic model with industry effects. A comparison of the estimated year and industry effects from the exact correction (equations (1) and (2)) and the approximate correction (equation (4) with $\alpha_i = \alpha$) is presented in the second and third columns of Appendix Table 3. The estimated industry and year effects

differ trivially between the two specifications.

^{17/}I fit separate first-differenced linear probability models of strike incidence to 61 bargaining pairs in Southern and Western states and 99 pairs in Northern states (the remaining bargaining situations are multi-state). Neither subsample rejected a two-season model based on December-May and June-November. The estimates of the June-November effects were .13 for the Northern pairs, and .23 for the Southern-Western pairs, with standard errors of .08 and .09, respectively.

^{18/}The courts have ruled, however, that a reopening agreement limited to "wages" does not preclude negotiation over "compensation" more broadly defined. See Meltzer (1977), pp. 712-718.

^{19/}I include as "reopenings" only those negotiations that are explicitly identified as reopeners in the preceding contract, and whose reopening dates are specified in the preceding contract.

^{20/}The probability of strikes among reopenings is 2 percent. Strike probabilities by the duration of the preceding contract are as follows: less than 18 months -- 7.11 percent; 18 to 29 months -- 9.93 percent; 30 to 41 months -- 20.91 percent; and 42 months and longer -- 29.73 percent.

^{21/}I have also estimated the effect of previous contract length by grouping contracts into one-year (less than 18 months), two-year (18 to 29 months), three-year (30 to 41 months) and longer (over 42 months) contracts. Relative to a one-year contract, the estimated effects and associated standard errors from a differenced linear probability model are: two-year contract -.01 (.03); three-year contract .06 (.03);

four-year or longer contract .26 (.09).

22/ For example, the average probability of strikes among two-year contracts is approximately 10 percent. The expected number of disputes per year on a two-year bargaining cycle is therefore .05. The estimates in Table 3 suggest that the probabilities of strikes in one- and three-year bargaining cycles are 4 and 16 percent, respectively. These probabilities imply .04 and .053 expected disputes per year, bargaining annually and every three years, respectively.

23/ The average expiration years corresponding to the six negotiations are 1964, 1966, 1969, 1972, 1974, and 1977.

24/ The bargaining pairs with this history are all unrelated (i.e., drawn from different industries, with different unions and firms).

25/ This model and all subsequent models are corrected for under-reporting of strikes using the adjustments described in the previous section.

26/ By comparison, the corrected goodness-of-fit statistic for a model with a constant probability of strikes across all bargaining pairs is 383.6 (with 63 degrees of freedom).

27/ The t-statistics are uncorrected for estimation of the parameters of the conditional logit model.

28/ This random-effects estimator was proposed by Card and Sullivan (1986) to describe the distribution of individual employment indicators over time. If the number of mass points is determined endogenously, rather than fixed a priori, the model can be interpreted as an application of the technique of nonparametric maximum likelihood--see Laird

(1978) and Heckman and Singer (1984).

29/ The estimates are conditional on the strike outcomes observed for each pair in the immediate pre-sample negotiation.

30/ Since the model is estimated by maximum likelihood, the estimates do not necessarily minimize the goodness-of-fit statistic for the strike table history. A similar situation arises when comparing the Pearson chi-squared statistic and the log-likelihood statistic for a simple contingency table of i.i.d. outcomes.

31/ The (Pearson) chi-squared test statistic for this restriction is 56.82, with four degrees of freedom.

32/ The assumption of independence between y_{it} and z_{it} is not particularly restrictive if the distribution of the individual effects α_i and γ_i is flexible.

33/ In the event of a short strike in $(t-1)$, $y_{it-1} = 1$ and $z_{it-1} = 0$, so $\text{logit}(p_{it}^*)$ is increased by δ_1 . In the event of a long strike in $t-1$, $y_{it-1} = 1$ and $z_{it-1} = 1$, so $\text{logit}(p_{it}^*)$ is increased by δ_2 .

34/ For example, the R^2 in a regression equation for log strike duration that includes two-digit industry effects and time variables is about 20 percent. The industry effects in such an equation are jointly significant at only about the 10 percent level.

35/ The number of consecutive strikes in the data set is relatively low: 58 (distributed among 42 bargaining pairs).

36/ This is consistent with results from a simple cross-sectional logit model of short and long strike durations. In such a model, none

of the industry effects are individually significant, and the probability value of the test for the hypothesis that the industry effects are zero is .12.

^{37/}Some additional evidence on the weak relation between lagged strike outcomes and strike durations is provided by the pattern of average strike durations by the number of strikes in 6 negotiations. These average durations are 52 days for 1 strike; 40 days for 2 strikes; 45 days for 3 strikes; 56 days for 4 strikes; and 16 days for 6 strikes.

^{38/}Approximately 60 percent of contract expirations occur in the five high probability strike months from June to November -- See Appendix Table 2.

References

- Abbott, Michael. "Specification Tests of Quarterly Econometric Models of Aggregate Strike Frequency in Canada." In R. Ehrenberg, ed. Research in Labor Economics 6 (1984): 177-250.
- Ashenfelter, Orley and George Johnson. "Bargaining Theory, Trade Unions, and Industrial Strike Activity." American Economic Review 59 (March 1969): 35-49.
- Bureau of National Affairs. Collective Bargaining and Negotiations (Part 2: Contract Clause Finder).
- Card, David and Daniel Sullivan. "Measuring the Effect of CETA Participation in Movements in and Out of Employment." Princeton University Industrial Relations Section Working Paper No. 206, February 1986.
- Chamberlain, Gary. "Analysis of Covariance with Qualitative Data." Review of Economic Studies 47 (January 1980): 225-238.
- Crampton, Peter. "Bargaining with Incomplete Information: An Infinite Horizon Model with Continuous Uncertainty." Stanford University Research Paper No. 680, May 1982.
- Farber, Henry S. "Bargaining Theory, Wage Outcomes, and the Occurrence of Strikes: An Econometric Analysis." American Economic Review 68 (June 1978): 261-71.
- Fudenberg, Drew, David Levine and Paul Ruud. "Strike Activity and Wage Settlements." UCLA Working Paper No. 249, January 1983.
- Gramm, Cynthia. "The Determinants of Strike Incidence and Severity."

- Industrial and Labor Relations Review 39 (April 1986): 361-76.
- Gunderson, Morley, John Kervin and Frank Reid. "Logit Estimates of Strike Incidence from Canadian Contract Data." Journal of Labor Economics 4 (April 1986): 257-76.
- Hayes, Beth. "Unions and Strikes with Asymmetric Information". Journal of Labor Economics 2 (January 1984): 57-83.
- Heckman, James. "The χ^2 Goodness-of-Fit Statistic for Models with Parameters Estimated from Microdata." Econometrica 52 (November 1984): 1543-47.
- Heckman, James and Burton Singer. "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data." Econometrica 52 (March 1984): 271-320.
- Kaufman, Bruce. "The Determinants of Strikes in the United States, 1900-1977." Industrial and Labor Relations Review 35 (October 1982): 473-90.
- Kennan, John. "Pareto Optimality and the Economics of Strike Duration." Journal of Labor Research 1 (Spring 1980): 77-94.
- _____. "The Duration of Contract Disputes in U.S. Manufacturing." Journal of Econometrics 28 (April 1985): 5-28.
- _____. "The Economics of Strikes." in Orley Ashenfelter and Richard Layard, eds., Handbook of Labor Economics North-Holland (forthcoming 1986a).
- _____. "Bargaining and Strikes with Private Information: Some Theoretical and Experimental Results." Unpublished Paper, February 1986b.

- Laird, Nan. "Non Parametric Maximum Likelihood Estimation of a Mixing Distribution." Journal of the American Statistical Association 73 (December 1978): 805-11.
- Meltzer, Bernard. Labor Law: Cases, Materials, and Problems (Second Edition). Boston: Little, Brown and Company, 1977.
- Pencavel, John. "An Investigation into Industrial Strike Activity in Britain." Economica 37 (August 1970): 239-56.
- Tracy, Joseph. "An Investigation into the Determinants of U.S. Strike Activity." American Economic Review 76 (June 1986): 423-36.
- Vroman, Susan. "Strike Activity and Unexpected Inflation." Unpublished Paper; June 1986.
- White, Halbert. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." Econometrica 48 (May 1980): 817-38.

APPENDIX

Derivation of Goodness-of-Fit Statistic
for Logistic Model Applied to Panel Data

Let $y_i = (y_{it} \dots y_{iT})$ represent a vector of observations on T binomial outcome variables for individual i , where $y_{it} = 1$ with probability p_{it} . Let x_{it} represent a vector of observed covariates for individual i in period t and let $x_i = (x_{i1}, x_{i2}, \dots, x_{iT})$. Let θ represent a k -dimensional vector of parameters, and let $f(y_i | x_i, \theta)$ represent the likelihood of the observed outcomes y_i , conditional on x_i and θ .

It is hypothesized that the T outcomes y_{it} are independently distributed conditional on x_{it} and θ , with $\log(p_{it}/(1-p_{it})) = x_{it} \theta$. In this case

$$(1) \quad f(y_i | x_i, \theta) = \prod_t f(y_{it} | x_{it}, \theta)$$

$$= \prod_t \frac{y_{it} e^{x_{it} \theta}}{1 + e^{x_{it} \theta}}.$$

Suppose that θ is estimated by maximum likelihood on a panel of individuals with T observations each. Let $\hat{\theta}_N$ represent the estimator based on N individuals (with fixed T). Under suitable regularity conditions on x_i ,

$$(2) \quad \sqrt{N} (\hat{\theta}_N - \theta) \overset{a}{\sim} N(0, \frac{1}{T} \Sigma_N)$$

where $\sum_N = -R_N^{-1}$

and $R_N = \frac{1}{NT} \sum_i \sum_t E \left[\frac{\partial^2 \log f(y_{it} | x_{it}, \theta)}{\partial \theta \partial \theta'} \right]$.

Consider the $J = 2^T$ possible outcomes for y_i indexed by $j = 1, 2, \dots, J$, with a representative outcome $d_j = (d_{j1}, d_{j2}, \dots, d_{jT})$. With $T = 2$, for example, the possible outcomes are $d_1 = (0, 0)$, $d_2 = (0, 1)$, $d_3 = (1, 0)$ and $d_4 = (1, 1)$. The goodness-of-fit test statistic is based on the difference between the actual number of individuals with a given outcome and the expected number under the model (1). Let $I_i(j) = 1$ if individual i has outcome j and let I_i represent the $J \times 1$ vector of indicator variables for individual i . Since the possible outcomes are exhaustive and mutually exclusive, $\sum_j I_i(j) = 1$ for all i .

Under (1), the probability that individual i is observed with the j th outcome is

$$(3) \quad E(I_i(j) | x_i, \theta) = \prod_t p_{it}^{d_{jt}} (1 - p_{it})^{1-d_{jt}} = \phi_i(j),$$

where $p_{it} = e^{x_{it}\theta} / (1 + e^{x_{it}\theta})$. If θ is known, then a goodness-of-fit statistic is readily available using the fact that the vector of indicators I_i is independently distributed across individuals with mean ϕ_i and covariance matrix

$$(4) \quad \Omega_i = E(I_i - \phi_i)(I_i - \phi_i)'$$

The elements of this covariance matrix have the simple form

$$\begin{aligned}\Omega_i(u,v) &= \Phi_i(u) (1 - \Phi_i(u)) , \quad u = v , \\ &= - \Phi_i(u) I_i(v) , \quad u \neq v .\end{aligned}$$

The vector of differences between the actual and expected sample fractions in each cell is

$$\frac{1}{N} \sum (I_i - \Phi_i) .$$

Let $\Omega_N = \frac{1}{N} \sum \Omega_i$ represent the average covariance matrix of I_i . Then (under suitable regularity conditions)

$$(5) \sqrt{N} \sum (I_i - \Phi_i) \sim N(0 , \Omega_N) .$$

The matrix Ω_N has rank $2^T - 1$. If Ω_N^- denotes a generalized inverse of Ω_N , the quantity

$$(6) \quad N \left(\frac{1}{N} \sum (I_i - \Phi_i)' \Omega_N^- \left(\frac{1}{N} \sum (I_i - \Phi_i) \right) \right)$$

is asymptotically distributed with $2^T - 1$ degrees of freedom.

If θ is unknown but the maximum likelihood estimator $\hat{\theta}_N$ is available then a straightforward application of the arguments in Heckman (1984) leads to the result that

$$(7) \quad \frac{1}{\sqrt{N}} \sum (I_i - E(I_i \mid x_i , \hat{\theta}_N)) \overset{a}{\sim} N(0 , \Delta_N)$$

where $\Delta_N = \Omega_N - \frac{1}{T} \alpha_N \sum_N \alpha_N'$,

and $\alpha_N = \frac{1}{N} \sum_i \frac{\partial \phi_i}{\partial \theta'}$. To motivate this expression observe that

$$\begin{aligned} I_i - E(I_i \mid x_i, \hat{\theta}_N) &= I_i - \phi_i - (E(I_i \mid x_i, \theta_N) - \phi_i) \\ &= I_i - \phi_i - \frac{\partial \phi_i}{\partial \theta'} (\hat{\theta}_N - \theta) + o(\theta_N - \theta)^2 . \end{aligned}$$

It follows that

$$\begin{aligned} \frac{1}{\sqrt{N}} \sum_i (I_i - E(I_i \mid x_i, \hat{\theta}_N)) &\approx \frac{1}{\sqrt{N}} \sum_i (I_i - \phi_i) \\ &+ \frac{1}{N} \sum_i \frac{\partial \phi_i}{\partial \theta'} \cdot \sqrt{N} (\hat{\theta}_N - \theta) . \end{aligned}$$

From the first-order conditions for the maximum likelihood estimator,

$$\sqrt{N} (\theta_N - \theta) \approx \frac{1}{T} \sum_N \frac{1}{\sqrt{N}} \sum_i \frac{\partial \log f(y_i \mid x_i, \theta)}{\partial \theta'} .$$

Combining these last two expressions,

$$(8) \sqrt{N} \sum_i (I_i - E(I_i \mid x_i, \hat{\theta}_N)) \sim \left[I, -\frac{1}{N} \sum_i \frac{\partial \phi_i}{\partial \theta'} \frac{1}{T} \sum_N \right] x$$

$$\sqrt{N} \sum_i \left[\frac{I_i - \phi_i}{\frac{\partial \log f(y_i \mid x_i, \theta)}{\partial \theta}} \right]$$

Note that

$$\begin{aligned} \frac{1}{N} \sum_i \text{var} \left[\frac{\partial \log f(y_i \mid x_i, \theta)}{\partial \theta'} \right] &= \frac{1}{N} \sum_i \text{var} \left[\sum_t \frac{\partial \log f(y_{it} \mid x_{it}, \theta)}{\partial \theta'} \right] \\ &= -\frac{1}{N} \sum_i \sum_t E \left[\frac{\partial \log f(y_{it} \mid x_{it}, \theta)}{\partial \theta \partial \theta'} \right] \\ &= T \sum_N^{-1} . \end{aligned}$$

Further,

$$\begin{aligned}
 & \frac{1}{N} \sum_i E \left[(I_i^* - \Phi_i) \frac{\partial \log f(y_i | x_i, \theta)}{\partial \theta'} \right] \\
 &= \frac{1}{N} \sum_i \left[\int_{y_i} I_i \frac{\partial \log f(y_i | x_i, \theta)}{\partial \theta'} \cdot f(y_i | x_i, \theta) dy_i \right] \\
 &= \frac{1}{N} \sum_i \left[\int_{y_i} I_i \frac{\partial f(y_i | x_i, \theta)}{\partial \theta'} dy_i \right] = \frac{1}{N} \sum_i \frac{\partial \Phi_i}{\partial \theta'} .
 \end{aligned}$$

Under suitable regularity conditions,

$$\frac{1}{\sqrt{N}} \sum (I_i - E(I_i \mid x_i, \hat{\theta}))$$

is therefore normally distributed with covariance matrix

$$\Delta_N = \begin{bmatrix} I & -\alpha_N \frac{\Sigma_N}{T} \end{bmatrix} \begin{bmatrix} \Omega_N & \alpha_N \\ \alpha_N' & T\Sigma_N^{-1} \end{bmatrix} \begin{bmatrix} I \\ -\frac{\Sigma_N}{T} \alpha_N' \end{bmatrix}$$

$$= \Omega_N - \frac{1}{T} \alpha_N \Sigma_N \alpha_N'.$$

An appropriate goodness-of-fit test accounting for the estimation of θ is

$$(9) \quad N \left[\frac{1}{N} \sum_i (I_i - E(I_i \mid x_i, \hat{\theta}_N)) \right]' \Delta_N^{-} \left[\frac{1}{N} \sum_i (I_i - E(I_i \mid x_i, \hat{\theta}_N)) \right]$$

where Δ_N^{-} represents a generalized inverse of the matrix Δ_N .

Table 1

Strike Characteristics By Industry

Industry	Number Pairs	All Available Contracts			Last 6 Contracts for Each Pair					Strike Duration (Days)
		Number Contracts	Strike Probability (Percent)	Strike Duration (Days)	Strike Probability (Percent)					
					Overall	1959-64	1965-69	1970-74	1975-79	
1. Food and Beverages	20	205	8.3	31.0	10.8	0.0	17.2	12.0	6.9	34.2
2. Tobacco	3	24	12.5	27.0	16.7	0.0	75.0	0.0	0.0	27.0
3. Textile Mills	13	185	4.9	20.2	7.7	0.0	11.1	15.8	2.6	13.8
4. Apparel	13	118	1.7	7.0	2.6	0.0	3.9	4.6	0.0	7.0
5. Lumber and Wood	9	88	6.8	79.7	3.7	0.0	5.6	5.9	0.0	59.0
6. Furniture	5	41	14.6	15.3	20.0	0.0	14.3	50.0	0.0	15.3
7. Paper	19	229	6.1	62.3	7.9	0.0	6.1	4.7	15.1	58.8
8. Printing	6	48	8.3	34.3	8.3	0.0	15.4	0.0	12.5	43.0
9. Chemicals	21	242	10.3	56.9	14.3	0.0	20.0	19.6	7.5	64.4
10. Petroleum	4	56	10.7	17.3	0.0	0.0	0.0	0.0	0.0	---
11. Rubber	7	73	27.4	56.6	40.5	16.7	35.7	58.3	40.0	59.8
12. Leather	5	45	0.0	---	0.0	0.0	0.0	0.0	0.0	---
13. Glass	22	215	8.8	42.7	12.1	0.0	24.4	7.5	9.4	40.7
14. Primary Metals	17	165	19.4	89.3	18.6	10.0	24.1	18.8	19.1	92.3
15. Fabricated Metals	14	131	19.9	76.5	25.0	29.4	28.6	28.0	14.3	40.9
16. Machinery (non-electrical)	14	124	29.8	33.1	38.1	21.4	40.0	42.9	41.2	28.3
17. Electrical Machinery	24	203	16.8	44.7	20.8	9.7	22.5	23.9	25.9	48.6
18. Transportation Equipment	24	200	20.5	42.3	24.3	14.3	32.5	16.3	33.3	42.1
19. Instruments	4	54	3.7	12.5	8.3	0.0	20.0	8.3	0.0	12.5
20. Miscellaneous Manufacturing	2	27	0.0	---	0.0	0.0	0.0	0.0	0.0	---
21. Transportation and Utilities	7	70	18.5	32.1	23.4	33.3	16.7	12.5	45.5	35.7
22. All Industries	253	2543	12.4	49.4	16.1	10.0	20.8	16.3	13.9	45.7

Note: Data source is described in text. Strikes include strikes from Current Wage Developments and strikes added from McConnell-Tracy strike listings.

Table 2
Seasonal Effects on Strike Incidence^{a/}

	Ordinary Logit ^{b/}		Conditional Logit ^{c/} with Year Effects	First Differenced Linear Probability with Year Effects	
	No Industry or Year Effects	Industry and Year Effects		Actual Estimates	Normalized Estimates
February	.79 (.36)	1.09 (.73)	.09 (.63)	.019 (.122)	.15 (.98)
March	1.24 (.32)	1.65 (.50)	.65 (.61)	.023 (.127)	.18 (1.02)
April	-.03 (.04)	-.02 (.14)	-.38 (.67)	-.089 (.108)	-.71 (.86)
May	-.19 (.36)	-.01 (.57)	-.00 (.67)	-.094 (.105)	-.75 (.84)
June	.30 (.30)	.70 (.50)	.65 (.61)	.071 (.089)	.57 (.71)
July	1.07 (.28)	1.08 (.39)	1.07 (.58)	.192 (.096)	1.54 (.77)
August	1.10 (.34)	1.07 (.37)	2.03 (.73)	.175 (.112)	1.40 (.90)
September	.61 (.34)	.48 (.42)	1.08 (.66)	.131 (.108)	1.05 (.86)
October	1.10 (.29)	.95 (.35)	.76 (.57)	.030 (.115)	.24 (.92)
November	1.30 (.33)	1.51 (.45)	1.09 (.59)	.114 (.124)	.91 (.99)
December	-1.65 (1.02)	-1.53 (.82)	-.83 (1.18)	-.020 (.105)	-.16 (.84)
Exclusion Test ^{e/} (probability value)	59.52 (.00)	53.62 (.00)	22.80 (.02)	21.97 (.02)	---

Notes: ^{a/} Based on last six contracts for each bargaining pair. All effects are relative to a January expiration.

^{b/} Estimated month effects from logit regression with and without controls for two-digit industry and time periods.

^{c/} Estimated month effects from conditional logit regression with approximate correction for unreported strikes prior to 1970. See text.

^{d/} Estimated month effects from first-differenced linear probability model. Standard errors and exclusion test statistic are corrected for heteroskedasticity. The normalized estimates represent the actual estimates multiplied by 8, and are directly comparable with the estimates obtained from the logit specifications.

^{e/} Likelihood ratio test for exclusion of month effects. The test statistic has 11 degrees of freedom. For the linear probability model, the test statistic is a Wald test that the month effects are all zero.

Table 3

Effects of Reopener Status and Previous
Contract Duration on Strike Incidence
 (standard errors in parenthesis)

	<u>Ordinary Logit</u>		<u>Conditional Logit</u> with Year Effects	<u>First-Differenced Linear</u> <u>Probability with Year Effects</u>	
	<u>No Industry</u> <u>or Year</u> <u>Effects</u>	<u>Industry</u> <u>and Year</u> <u>Effects</u>		<u>Actual</u> <u>Estimates</u>	<u>Normalized</u> <u>Estimates</u>
1. Reopener	-1.68 (.73)	-1.70 (.92)	-1.43 (.76)	-.086 (.036)	-.69 (.28)
2. Previous Contract Duration (Months)	.052 (.009)	.043 (.010)	.043 (.013)	.0049 (.0015)	.039 (.012)

Note: Models with year effects include a four-step time function. See notes to Table 2.

Table 4

Actual and Predicted Strike Histories:Last Six Contracts for each Pair

History	Actual Number Pairs	Predicted Number (absolute t-statistics) ^{a/}		
		Ordinary Logit with Industry Effects	Conditional Logit	Random Effects Logit (2 mass points)
<u>No Strikes</u>				
1. 000000	118	105.8 (3.07)	118 (--)	118.9 (.12)
<u>1 Strike</u>				
2. 100000	12	10.7 .43	8.1 (1.49)	8.9 (1.06)
3. 010000	8	13.3 (1.56)	9.9 (.66)	11.3 (1.01)
4. 001000	17	17.3 (.09)	15.6 (.42)	14.4 (.69)
5. 000100	9	13.9 (1.41)	10.8 (.59)	10.7 (.54)
6. 000010	9	12.0 (.95)	10.20 (.41)	9.2 (.06)
7. 000001	11	13.4 (.81)	11.49 (.41)	10.4 (.20)
8. Total 6 cases	66	80.7 (2.20)	66.0 (--)	64.9 (.15)
<u>2 Strikes</u>				
9. 110000	1	2.17 (.82)	1.6 (.49)	2.2 (.81)
10. 011000	5	3.6 (.80)	3.1 (1.11)	3.3 (.94)
11. 001100	3	4.2 (.61)	4.3 (.66)	3.4 (.23)
12. 000110	2	2.9 (.53)	2.6 (.39)	2.3 (.17)
13. 000011	8	2.7 (3.34)	2.4 (3.69)	2.3 (3.90)
14. 101000	2	2.8 (.48)	2.2 (.11)	2.5 (.32)
15. 010100	1	2.8 (1.08)	2.3 (.88)	2.4 (.93)
16. 001010	2	3.5 (.84)	3.4 (.81)	2.9 (.54)
17. 000101	2	3.2 (.71)	3.0 (.60)	2.6 (.35)
18. 100100	3	2.2 (.52)	1.7 (1.04)	2.0 (.74)
19. 010010	1	2.5 (.95)	2.0 (.75)	2.2 (.84)
20. 001001	4	4.0 (.02)	3.9 (.03)	3.3 (.39)
21. 100010	0	2.0 (1.43)	1.5 (1.23)	1.8 (1.34)
22. 010001	2	2.8 (.47)	2.3 (.21)	2.5 (.33)
23. 100001	2	2.2 (.15)	1.6 (.28)	2.00 (.00)
24. Total 16 cases	38	43.4 (1.06)	38 (--)	37.7 (.06)
<u>3 Strikes</u>				
25. Total 32 cases	24	17.1 (2.06)	24 (--)	21.8 (.51)
<u>4 Strikes</u>				
26. Total 16 cases	6	5.0 (.49)	6 (--)	8.0 (.71)
<u>5 Strikes</u>				
27. Total 6 cases	0	.9 (1.05)	0 (--)	1.6 (1.26)
<u>6 Strikes</u>				
28. 111111	1	.1 (3.21)	1 (--)	.1 (2.43)

29. Goodness-of-fit for table uncorrected for parameter estimation ^{b/} (probability value)		74.47 (.15)	49.35 (.75)	60.07 (.58)
30. Goodness-of-fit for table corrected for parameter estimation ^{c/} (probability value)		83.55 (.04)	---	---

- Notes to Table 4 -

- Notes: a/ Predicted number represents the expected number of pairs with a given history, conditional on the estimated parameters. The number in parentheses represents the absolute t-statistic associated with the test that the predicted and actual numbers of pairs are equal. See text for derivation of test statistics. T-statistics for ordinary logit are corrected for parameter estimation. T-statistics for other models are not corrected for parameter estimation.
- b/ Goodness-of-fit statistic for the overall table (64 elements) treating the estimated parameters as known constants.
- c/ Goodness-of-fit statistic for the overall table (64 elements) corrected for parameter estimation.

Table 5

Estimates of the Effect of Preceding Strike Activity on Strike Probabilities

	Random Effects Logit Model ^{a/}					
	2-mass points	2-mass points	2-mass points	2-mass points	2-mass points	2-mass points and 4-industry effects ^{b/}
	(1)	(2)	(3)	(4)	(5)	(6)
<u>Effect of Previous Strike:</u>						
1. 0-7 day strike	0.0 (---)	.13 (.23)	1.02 (.33)	.84 (.47)	1.02 (.33)	.94 (.33)
2. 8-14 day strike	0.0 (---)	.13 (---)	1.02 (---)	1.18 (.46)	1.02 (---)	.94 (---)
3. 15-60 day strike	0.0 (---)	.13 (---)	-.40 (.28)	-.41 (.29)	-.47 (.33)	-.82 (.39)
4. 61 day or longer strike	0.0 (---)	.13 (---)	-.40 (---)	-.41 (---)	-.28 (.42)	-.82 (---)
5. Log Likelihood	-637.21	-637.06	-629.79	-629.65	-629.71	-598.26
6. Goodness-of-fit for Strike Outcome Table ^{c/} (probability value)	60.07 (.58)	62.77 (.48)	59.76 (.59)	58.31 (.64)	59.28 (.61)	56.93 (.72)

Notes: ^{a/} All models include a four-step time function.

^{b/} Model includes four grouped industry effects.

^{c/} Goodness-of-fit to 64-element table of possible strike outcomes in six contract negotiations. The statistic does not account for estimation of the parameters in the model.

Table 6

Probability of Strikes Conditional on Preceding Strike Outcome

	Previous Strike Outcome					
	No Strike	1-7 Day Strike	8-14 Day Strike	15-28 Day Strike	29-42 Day Strike	43 Day or Longer Strike
1. Probability of Strike	13.9	40.0	47.5	24.4	18.5	19.0
2. Number of Contracts (percent of sample)	1291 (85.05)	35 (2.31)	40 (2.64)	41 (2.70)	27 (1.78)	84 (5.53)

Note: Sample cross-tabulation of strike outcomes by previous strike outcome.

Table 7

First-Order Transition Matrix for Strike Outcomes

	Previous Strike Outcome			All
	No Strike	Short Strike (≤ 14 days)	Long Strike (> 14 days)	
<u>Probability of:</u>				
1. No Strike	86.1	56.0	79.6	83.9
2. Short Strike (≤ 14 days)	4.5	10.7	4.6	4.8
3. Long Strike (> 14 days)	9.4	3.3	15.8	11.3
4. Conditional Probability of Long Strike	67.6	75.7	77.5	70.2
5. Number of Contracts (percent of total)	1291 (85.1)	75 (4.9)	152 (10.0)	1518 (100.0)

Table 8
Trinomial Outcome Model of Strike Probability and Duration
 (standard error in parentheses)

	Random Effects Logit Model ^{a/}			
	2 Mass Points	3 Mass Points	3 Mass Points	2 Mass Points and 4 Industry Effects ^{b/}
<u>Effect of Previous Strike on Strike Probability:</u>				
1. 0-14 day strike	1.08 (.36)	1.04 (.34)	1.11 (.35)	1.01 (.35)
2. 15 day or longer strike	-.44 (.29)	-.51 (.29)	-.49 (.29)	-.69 (.30)
<u>Effect of Previous Strike on Probability of Long Strike:</u>				
3. 0-14 day strike	.13 (.46)	.17 (.48)	.00 (---)	.00 (---)
4. 15 day or longer strike	.44 (.58)	.37 (.57)	.00 (---)	.00 (---)
5. Relative effect of time on probability of long strike	1.49 (.58)	1.53 (.57)	1.60 (.61)	1.45 (.53)
6. Log Likelihood	-771.34	-771.03	-771.34	-749.92
7. Goodness-of-fit for binomial strike outcome table ^{c/} (probability value)	56.17	55.68	55.21	55.85

Notes: ^{a/}All models include a four-step time function normalized to have a unit effect on the probability of a strike. The coefficient in row 5 gives the relative effect of this time function on the probability of a strike longer than 14 days.

^{b/}Model includes four grouped industry effects in probability of strike and probability of long strike equations. The industry effects are restricted to enter proportionately in the two equations. The estimated relative effect of the industry variables in the duration equation is .17 (with a standard error of .18).

^{c/}Goodness-of-fit to 64 element table of possible strike outcomes in six contract negotiations. The statistic does not account for estimation of the parameters in the model.

Appendix Table 2

Strike Probabilities by Month: Last Six Contracts for Each Pair

	Percent of Contracts	Strike Probability (Percent)	
		Adjusted	Adjusted for Industry Composition ^{a/}
January	6.5	9.2	7.7
February	5.5	18.1	20.5
March	6.9	24.8	27.6
April	7.6	8.7	8.5
May	10.3	7.7	9.2
June	15.6	11.8	15.3
July	13.4	21.7	20.3
August	5.6	22.4	20.3
September	7.9	15.0	11.8
October	11.0	22.2	18.7
November	6.3	26.0	25.8
December	3.5	1.9	2.7

Note: ^{a/} Estimated month effects from linear probability model that also includes two-digit industry effects.

Appendix Table 3

Ordinary Logit Parameter Estimates with
Alternative Corrections for Under-Reporting of Strikes^{a/}

	No Correction	Exact Correction ^{b/}	Approximate Correction ^{c/}
1. Constant	-1.22	-1.21	-1.23
2. Food and Beverages	-.91	-.95	-.91
3. Tobacco	-.26	-.32	-.24
4. Textile Mills	-1.25	-1.26	-1.23
5. Apparel	-2.50	-2.55	-2.50
6. Lumber and Wood	-2.11	-2.13	-2.09
7. Furniture	-.11	-.12	-.12
8. Paper	-1.29	-1.32	-1.29
9. Printing	-1.19	-1.23	-1.19
10. Chemicals	-.59	-.61	-.58
11. Petroleum	---	---	---
12. Rubber	.88	.91	.88
13. Leather	---	---	---
14. Glass	-.76	-.79	-.75
15. Primary Metals	-.17	-.18	-.15
16. Fabricated Metals	.14	.13	.14
17. Machinery (non-electrical)	.78	.78	.78
18. Electrical Machinery	-.10	-.10	-.09
19. Transportation Equipment	.17	.17	-.18
20. Instruments	-1.27	-1.33	-1.28
21. Miscellaneous Manufacturing	---	---	---
22. 1959-65	-.56	-.29	-.31
23. 1966-70	.39	.63	.59
24. 1971-75	-.14	-.13	-.14
25. Log Likelihood	-599.36	-599.26	-599.35

Notes: ^{a/} Estimated industry and time-period effects for logistic specification of strike probabilities (pooled cross-section time series data).

^{b/} See text. Correction for under-reporting of strikes based on equations (1) and (2) of text.

^{c/} See text. Correction for under-reporting of strikes based on equation (4) of text.