working paper
department
of economics

WORKER PREFERENCES FOR UNION REPRESENTATION

Henry S. Farber*

Number 290

August 1981

massachusetts
institute of
technology

50 memorial drive
cambridge, mass. 02139

Department of Economics
Massachusetts Institute of Technology

WORKER PREFERENCES FOR UNION REPRESENTATION

Henry S. Farber*

Number 290 August 1981

ABSTRACT

A model of the determinants of worker preferences for union representation as distinct from their actual union status is developed and estimated using data from the Quality of Employment Survey. In order to implement the model, a pair of econometric issues were addressed. First, the worker preferences were available only for nonunion workers. After correcting for this censoring, it was found that preference for union representation was higher among the workforce in general than among the nonunion workforce. In addition, preferences for representation vary much more by worker characteristics among the workforce in general than they do among the nonunion workforce. This is undoubtedly due to sorting based on worker preferences. The second problem regarded proper estimation of the union-nonunion wage differential, which was hypothesized to be a positive determinant of worker's preferences for union representation. Three different measures were used and, while there was some variation between them, they all showed a similar relationship with worker preferences.

*This research was supported by the National Science Foundation under Grant No. SES-7924880. The author also received support as an Alfred P. Sloan Research Fellow.

I. Introduction

A relatively neglected area of research on labor unions is the determination of the union status of workers. In order to understand the process through which labor unions developed, what their future holds, and what their effects on workers, the workplace, and compensation are, it is crucial that a thorough understanding of this issue be gained. In most studies which develop structural models of the determination of the union status of workers (e.g., Ashenfetter and Johnson, 1972; Lee, 1978), it is argued that union status is determined strictly as a result of worker preferences for unionization. However, as Kochan and Helfman (1981) point out, this is only part of the story. A worker's preference for a unionized job will only translate into such a job if a unionized employer is willing to hire that worker, and it is likely that there will be excess demand for vacancies in existing union jobs.[1] Hence, the employer's criteria for selection of workers from the queue need to be modeled along with worker preferences in order to model adequately the determination of union status.

Abowd and Farber (1982) attempt an analysis of the determination of a worker's union status in which a distinction is drawn between worker preferences and employer choice criteria. However, their analysis is hampered by the fact that only the final outcome (union status) is observed, and it is impossible to determine whether nonunion workers did not want a union job, could not get a union job, or both. This difficulty is compuonded by the problem that a worker's current preference for a union job, given accrued seniority, may be different than it was at the time the worker took his current job. For example, a nonunion worker with ten years seniority may not want a union job even if one were offered at that point. However, the worker may have preferred a union job ten years earlier but was not offered one.

In this study, a rather unique data set is employed which can be used to identify for nonunion workers their preferences for unionization <u>holding seniority fixed</u>. Thus, one-half of the process through which worker union status is determined can be investigated. However, these data do not permit investigation of the employer selection process, and a satisfactory solution of the entire puzzle must await future research.

In the next section, a simple structural model of worker preferences for unionization is developed based on utility maximization by individual workers. Section III contains a discussion of the data set from the 1977 Quality of Employment Survey (QES) developed by the Survey Research Center at the University of Michigan. Particular attention is paid to the interpretation of the question, "If an election were held with secret ballots, would you vote for or against having a union or employees' association represent you?" The response to this question (VFU) serves as the basis for the analysis of this study.

Empirical implementation of the model developed in Section II is hampered by two problems. First, the crucial question (VFU) was asked only of nonunion workers. Analysis of the responses in this context is interesting in that insight can be gained into the characteristics of nonunion workers which lead them to desire union representation <u>conditional</u> on their nonunion status.[2] However, for an analysis of preferences for union representation <u>unconditional</u> on union status, the data are censored on the basis of an obviously related variable. An econometric model which takes account of this censoring and yields consistent estimates of unconditional preferences for union representation is developed in Section IV.

The second problem which hampers the empirical implementation of the model is that a crucial element of the structural model is the union-nonunion wage differential (ΔW) facing a worker. Potential problems of sample selection bias in estimating ΔW, while solved from a technical standpoint, are notoriously difficult to handle in a convincing fashion from a practical point of view.[3] For this reason, the analysis of the structural version of the model is deferred until a later section, and a reduced form version of the model is derived in Section IV which does not require estimates of ΔW. This reduced form model is then estimated in Section V both with and without accounting for the censored sample problem noted above. The estimates suggest that preferences for union representation among the nonunion workforce are relatively flat across most individual characteristics, while among the workforce as a whole there are sharp distinctions. This result is due to sorting in the sense that many workers who desire union representation on the basis of both their observed and their unobserved characteristics are already working on union jobs, leaving a group of nonunion workers whose preferences for union representation show little systematic variation with characteristics.

In Section VI, separate union and nonunion earnings functions are estimated by a number of different techniques in order to derive estimates of ΔW for the structural version of the union preference model. First, ordinary least-squares (OLS) estimates are derived. However, these are potentially biased and inconsistent due to the fact that the sample censoring of wage rates based on union status may be correlated with the wage differentials. A not completely satisfactory attempt to account for this sample selection bias is made by estimating the union and nonunion earnings functions by two additional methods. The first of these is to use OLS augmented by the hazard rate (inverse Mill's ratio) derived from a reduced form probit union status model. This technique is described by Lee (1979). The second method is to estimate a maximum likelihood switching regression model consisting of the union and nonunion earnings functions plus an equation explaining union status. The union-nonunion wage differentials implied by these various methods are then evaluated and compared but, while the results differ, it cannot be concluded that any particular measure is clearly superior.

In Section VII the structural version of the preferences for the union representation model is estimated both conditional on nonunion status and accounting for sample censoring. All three measures of ΔW derived in Section V are used due to the ambiguity concerning the correct measure, and the results are compared. The results are remarkably similar, particularly in light of the substantial differences in the estimated wage differentials. Overall, the analysis in this section confirms the predictions of the theory developed in Section II in that a positive, though insignificant, relationship is found between ΔW and worker preferences for union representation.

The final section of the paper contains a synthesis of the results along with conclusions which can be drawn from the analysis.

II. A Simple Model of Worker Preferences for Union Representation

At its simplest level, a worker's preference for union representation on a job versus no such representation can be modeled as a comparison by the worker of the utilities yielded to him by his job in each case. His preference will be for that case which yields him the largest utility. A worker's utility on the job is affected by many factors, including wages, fringe benefits, safety, job security, comfort, etc. In addition, there are subjective factors such as satisfaction with supervision, perceived fairness of treatment, equitable comparisons with others, perceived chances for promotion, etc. In the absence of explicit measures of most of these factors for each worker in both a union and a nonunion environment, it is argued that these utilities vary across workers as functions of their personal and occupational characteristics as well as the measurable characteristics of the union and nonunion jobs.

In order to examine this argument more carefully, assume, as is done below, that the only explicit job characteristic which can be measured for each worker in both a union and a nonunion context is the wage rate. More formally, it is argued that

$$(1) \quad V_u = V_u(Z, W_u)$$

and

(2) $V_n = V_n(Z, W_n)$

where V_u and V_n represent the worker's union and nonunion utilities respectively, Z represents a vector of personal and occupational characteristics, and W_u and W_n represent the worker's union and nonunion wage rate respectively. The worker's preference for union representation can be expressed by computing $y = V_u - V_n$. If this difference in utilities (y) is positive, then the worker will opt for union representation. If y is negative, then the worker will not opt for union representation. It is clear that this difference in utilities can be expressed as

(3) $y = y(Z, W_u, W_n)$,

where variations in Z measure variations in the difference between the worker's union and nonunion utilities.

Examples of such variations are not hard to come by. For instance, it is well known that fringe benefits are substantially more generous on union jobs in such dimensions as medical insurance, pensions, and vacation pay.[4] Those workers who place a greater value on these fringe benefits are more likely to have a positive y and hence desire union representation. For example, older workers are likely to value their potential pension benefits more than younger workers, while workers with young families are likely to value medical insurance relatively highly. In another dimension, the prevalence of layoff by inverse seniority rules in unionized establishments may lead workers with relatively more seniority to desire union representation due to the increased job security such seniority confers in union settings.[5]

In order to derive the empirical analog of this model, a specific functional form must be selected for y. This is

$$(4)\quad y = Z\gamma + \delta(\ln W_u - \ln W_n) + \varepsilon_1$$

where γ is a vector of parameters, δ is a paremeter expected to be positive, and ε_1 represents unmeasured components of the utility difference. Clearly, if y, Z, W_u, and W_n were observable for all workers, the parameters in equation (4) could be estimated using straightforward linear methods. However, this is not the case, and the discussion turns to an examination of the data and their limitations.

III. The Data

The data used are from the 1977 cross-section of the Quality of Employment Survey (QES) developed by the Survey Research Center of the University of Michigan. The QES contains data for approximately 1500 randomly selected workers (both union and nonunion) on their personal characteristics and job attributes.[6] The crucial bit of information for this study is the response to the question asked only of nonunion workers, "If an election were held with secret ballots, would you vote for or against having a union or employee's association represent you?" This variable is called "Vote-for-union?" or VFU. It is interpreted here as the preference of a worker for union representation on his current job. Thus, it holds all job characteristics fixed except those which the worker expects the union to affect. It is assumed that the worker's response is based on his current utility as compared with what the worker's utility is expected to be were the job to be unionized.[7]

A sample of workers was derived from QES by selecting those workers for whom the survey contained valid information on the variables listed in Table 1. Self-employed workers, managers, sales workers, and construction workers were deleted from the sample. The remaining sample contains 880 workers. Table 1 contains descriptions of the variables used in the study as well as their means and standard deviations for both the entire sample and the union and nonunion subsamples. The base group for the dichotomous variables are white, nonsouthern, single, male blue collar workers with twelve years of education. On average, the 37 percent of the sample who are unionized are slightly older, earn substantially more, have somewhat more experience, and are more likely to be male, married, nonwhite, nonsouthern, and in a blue collar occupation.

Thirty-eight percent of the nonunion sample expressed a preference for union representation (VFU = 1). It is unfortunate that the analogous question was not asked of union members (If an election were held by secret ballot, would you vote to continue union representation?). This would make information available for all workers about worker preferences for unionization of their current job at the current time.[8] The lack of this information poses some important problems of econometrics and interpretation. It is to this and related problems that the next section is addressed.

Table 1. Means (Standard Deviations) of Data
Quality of Employment Survey, 1977

Variable	Description (Dichotomous variables = 0 otherwise)	Combined Sample (n=880)	Union Sub-Sample (n=327)	Non-Union Sub-Sample (n=553)
U	= 1 if works on union job	.372	--	--
VFU	= 1 if desires union represent.	--	--	.376
Age	age in years	36.4 (12.9)	38.1 (12.5)	35.4 (13.0)
Exp	labor market experience in years	16.9 (12.4)	19.1 (12.4)	15.6 (12.2)
Sen	firm seniority in years	6.81 (7.46)	9.50 (8.22)	5.22 (6.45)
Fe	= 1 if female	.420	.324	.477
Marr	= 1 if married w/spouse present	.636	.703	.597
Marr*Fe	= 1 if Fe = 1 and Marr = 1	.200	.174	.215
NW	= 1 if nonwhite	.139	.162	.125
South	= 1 if worker resides in South	.353	.235	.423
Ed < 12	= 1 if <12 years education	.220	.257	.199
12<Ed<16	= 1 if >12 years & <16 years educ.	.213	.165	.241
Ed ≥ 16	= 1 if ≥16 years education	.199	.199	.199
ln(wage)	natural logarithm of wage	1.58 (.859)	1.83 (.983)	1.43 (.737)
Cler	= 1 if occupation is clerical	.208	.116	.262
Serv	= 1 if occupation is service	.153	.113	.177
Prof&Tech	= 1 if occupation is professional or technical	.230	.205	.244

IV. Econometric Issues and Options

A number of problems arise in the estimation of the parameters of the worker preference model specified in equation (4). The first problem is that y, which represents the differences in a worker's utility between union and nonunion status on the job, is not observed. All that is observed is the sign of y through the worker's response to the VFU question ($y > 0 \Leftrightarrow VFU=1$, $y < 0 \Leftrightarrow VFU=0$). The limited nature of the dependent variable implies that $Pr(VFU=1) = Pr(y > 0)$, yielding from equation (4) that

$$(5)\quad Pr(VFU = 1) = Pr(\varepsilon_1 > -Z\gamma - \delta(\ln W_u - \ln W_n)) .$$

If ε_1 is assumed to be distributed normally with zero mean and unit variance, then equation (5) implies a probit specification for a likelihood function.[9] The contribution of any individual to the log-likelihood function is

$$(6)\quad L_1 = VFU \ln\left(\Phi(Z\gamma + \delta(\ln W_u - \ln W_n))\right)$$

$$+ (1 - VFU) \ln\left(1 - \Phi(Z\gamma + \delta(\ln W_u + \ln W_n))\right)$$

where $\Phi(\cdot)$ represents a standard normal cumulative distribution function.[10]

If all of the elements of equation (6) were observed for all workers, it would be a straightforward exercise to maximize the appropriate likelihood function to obtain estimates of Z and γ. Unfortunately, our task is not so simple. As mentioned above, VFU is observed only for nonunion workers. If the question of interest is the estimation of a model of nonunion worker preferences for union representation conditional on their nonunion status and if the assumption is made that ε_1 has a standard normal distribution conditional on the workers being nonunion, then maximum likelihood estimation applied to the likelihood function implied by equation (6) over all of the nonunion workers in the sample will lead to consistent estimates of γ and δ. However, these estimates cannot be interpreted as those which describe a model applicable to all workers regardless of union status unless a particular (testable) restriction described below is imposed.

In order to continue the analysis, an empirical model of the determination of the union status of workers is required. A simple model is specified of the form

$$(7) \quad S = C\alpha + \varepsilon_2 \ ,$$

where S is an unobservable latent variable determining union status, C is a vector of worker and job characteristics, α is a vector of parameters, and ε_2 is a random component with a standard normal distribution which captures unmeasured aspects of the union status decision. If S is positive, then the worker works on a union job (U=1), and if S is negative then there is no union representation on the job (U=0). Thus, Pr(U=1) = Pr(S > 0), which implies that

$$(8) \quad \Pr(U=1) = \Pr(\varepsilon_2 > -C\alpha) \ .$$

Given the normality assumption regarding ε_2, the contribution to the log-likelihood function regarding union status is a probit of the form

$$(9)\quad L_2 = U \ln \left(\Phi(C\alpha)\right) + (1 - U) \ln\left(1 - \Phi(C\alpha)\right) .$$

In light of the introductory discussion of the process by which union status is determined through separate decisions by workers and employers, the behavioral underpinnings of this probit model are left deliberately vague. It is to be interpreted as a reduced form empirical relationship describing the union status of a worker. However, a note of caution is required. It is difficult (if not impossible) to think of a structural model of the determination of the union status of a worker where both the worker and employer make separate decisions which would have a reduced form which could be described as a simple univariate probit. In fact, this study is motivated, at least in part, by a desire to model the union status determination of workers in a manner which is consistent with separate worker and employer decisions in order to move away from the behaviorally naive structural model which is implicit in the simple probit model described here. Nonethless, we continue with the simple probit reduced form representation in order to continue the analysis in the hope that it is a reasonable approximation to a reduced form which would be derived from an appropriate behavioral model.

If the random components in the VFU and U functions (ε_1 and ε_2) are correlated (e.g, they have a standardized bivariate normal density function $b(\varepsilon_1,\varepsilon_2;\ \rho_{12})$), then estimation of the likelihood of VFU from equation (6) is incorrect if the goal is to estimate γ and δ for workers unconditional on their union status. In particular, ε_1 was assumed to have a normal distribution with zero mean unconditional on union status, but the ε_1 are observed only for nonunion workers. The condition for a worker being nonunion from equation (8) is that $\varepsilon_2 < -C\alpha$ so that ε_1 is observed only if $\varepsilon_2 < -C\alpha$, and the likelihood must be written in terms of this conditional distribution. Using Bayes' Rule and assuming joint normality of ε_1 and ε_2, the conditional distribution of ε_1 given $\varepsilon_2 < -C\alpha$ is

$$(10) \quad f(\varepsilon_1|\varepsilon_2 < -C\alpha) = \frac{\int_{-\infty}^{-C\alpha} b(\varepsilon_1,\varepsilon_2;\ \rho_{12})d\varepsilon_2}{\Phi(-C\alpha)} .$$

This conditional distribution is non-normal and involves the parameters α and ρ_{12}.

Two points are worth noting here. First, if ε_1 and ε_2 are uncorrelated and so that ρ_{12}=0, then $\varepsilon_1|\varepsilon_2 < -C\alpha$ is distributed as a standard normal and the likelihood function on VFU implied by equation (6) and estimated over the sample of nonunion workers can be interpreted correctly as that relevant to all workers unconditional on their union status. However, it seems likely that unmeasured determinants of the union status of a worker and of the worker's preference for union representation are correlated with each other so that it is unlikely that ρ_{12}=0. This potential restriction will be tested in succeeding sections. The second point to note is that where $\rho_{12}\neq 0$, the conditional interpretation given to the probit VFU likelihood function derived from equation (6) and the conditional likelihood function for VFU derived from the bivariate normal model in equation (10) are inconsistent with each other because $f(\varepsilon_1|\varepsilon_2<-C\alpha)$ is non-normal. Nonetheless, estimates from the simple probit VFU likelihood function will be interpreted as estimates of worker preferences for union representation conditional on being nonunion.

While the appropriate conditional likelihood function for VFU could be derived from equation (10), a much more efficient approach is to use information from the whole sample to write the joint likelihood of preferences for union representation and union status while noting that VFU is censored for union workers. For nonunion workers who desire union representation the appropriate contribution to the likelihood function is $\Pr\left(\varepsilon_1 > -Z\gamma - \delta(\ln W_u - \ln W_n),\ \varepsilon_2 < -C\alpha\right)$. Given the distributional assumption, this is

$$(11)\quad \Pr(\mathrm{VFU}=1,\ U=0) = \int_{\infty}^{-C\alpha}\int_{\kappa_1}^{\infty} b(\varepsilon_1,\varepsilon_2;\ \rho)\,d\varepsilon_1 d\varepsilon_2$$

where $\kappa_1 = -Z\gamma-\delta(\ln W_u - \ln W_n)$. Similarly, for nonunion workers who do not desire union representation, the appropriate contribution is $\Pr\left(\varepsilon_1 < -Z\gamma - \delta(\ln W_u - W_n),\ \varepsilon_2 < -C\alpha\right)$, which yields

$$(12)\quad \Pr(\mathrm{VFU}=0,\ U=0) = \int_{-\infty}^{-C\alpha}\int_{-\infty}^{\kappa_1} b(\varepsilon_1,\varepsilon_2;\ \rho)\,d\varepsilon_1 d\varepsilon_2\ .$$

Finally, for union workers no information regarding VFU is known, so that ε_1 is integrated out and the contribution of these workers to the likelihood function is a univariate normal CDF representing $\Pr(\varepsilon_2 > -C\alpha)$, which yields

$$(13)\quad \Pr(U=1) = \Phi(C\alpha)\ .$$

Using these results, the contribution of a worker to the structural log-likelihood function accounting for the sample censoring is

$$(14)\quad L = (VFU)(1 - U)\ \ln\big(\Pr(VFU=1,\ U=0)\big) + (1 - VFU)(1 - U)\ \ln\big(\Pr(VFU=0,\ U=0)\big) + U\ \ln\big(\Pr(U=1)\big)$$

where the relevant probabilities are defined in equations (11)-(13).

One more hurdle must be overcome before the model can be estimated. Two important variables required for all nonunion workers are those workers' union and nonunion wage rates. However, only the nonunion wage is observed, and a question arises as to how to handle this problem. The difficulty is that it is likely that the union-nonunion wage differential, and hence the union and nonunion wage rates, are important determinants of ultimate union status (in a structural version of the model) as well as of worker preferences for union representation. This raises potentially serious problems of selection bias in estimating a union and a nonunion earnings function which will be addressed in Section VI. In addition, the fact that the observed nonunion wage is likely to be correlated with union status makes it improper to estimate the model conditional on this wage rate. The correlation must be accounted for.

One approach toward solution of this problem is to specify union and nonunion individual earnings functions as

$$(15)\quad \begin{aligned} \ln W_u &= X\beta_u + \varepsilon_u \quad \text{and} \\ \ln W_n &= X\beta_n + \varepsilon_n \end{aligned}$$

where X is a vector of exogenous characteristics, β_u and β_n are vectors of parameters, and ε_u and ε_n are normally distributed unobserved elements. The difficulties stem from the possibility that ε_u and/or ε_n are correlated with ε_1 and ε_2. If the representations of ln W_u and ln W_n from equations (15) are substituted into equation (4), the resulting union-nonunion utility difference can be expressed as

$$(16)\quad y = Z\gamma + \delta X(\beta_u - \beta_n) + \varepsilon_3$$

where $\varepsilon_3 = \varepsilon_1 + \delta(\varepsilon_u - \varepsilon_n)$ and is distributed normally unconditional on union status.

Substituting into equation (6), the individual contribution to the log-likelihood function for VFU which is interpreted conditionally on nonunion status is

$$(17)\quad L_1 = \text{VFU}\ \ln\left(\Phi(Z\gamma + \delta X(\beta_u-\beta_n))\right) + (1-\text{VFU})\ \ln\left(1-\Phi(Z\gamma + \delta X(\beta_u-\beta_n))\right)$$

under the normalization that the variance of ε_3 is one.[11] Similarly, the relevant probabilities for the joint likelihood function defined in equation (14) can be rewritten as:

$$(18)\quad \Pr(\text{VFU}=1,\ U=0) = \int_{-\infty}^{-C\alpha} \int_{\kappa_3}^{\infty} b(\varepsilon_3,\varepsilon_2;\ \rho_{23})d\varepsilon_3,d\varepsilon_2$$

and

$$(19)\quad \Pr(\text{VFU}=0,\ U=0) = \int_{-\infty}^{-C\alpha} \int_{-\infty}^{\kappa_3} b(\varepsilon_3,\varepsilon_2;\ \rho_{23})d\varepsilon_3 d\varepsilon_2$$

where $\kappa_3 = -Z\gamma-\delta X(\beta_u-\beta_n)$ and ρ_{23} is the correlation between ε_2 and ε_3. The Pr(U=1) is unchanged.

The structure of equation (16) raises some serious identification problems because only for those elements of Z which are not included in X can the associated γ's be estimated. Similarly, only the product of $\delta(\beta_u-\beta_n)$ can be estimated, and then only for elements of X which are not included in Z. What can be identified is a reduced form version of equation (16) which is specified by substituting $Z^*\gamma^* = Z\gamma + \delta X(\beta_u-\beta_n)$ into equations (16) through (19). The vector Z* contains all of the variables which are in either Z or X, and γ^* is the vector of reduced form coefficients. This reduced form model is estimated in the next section.

The entire structural model can be identified and estimated by using data on the union and nonunion wage rates to derive estimates of β_u and β_n which can then be substituted into the model to estimate γ and δ conditional on these estimates of β_u and β_n. However, these is some question as to the best technique to estimate these vectors, and three different methods are used in Section VI. Finally, in Section VII the structural version of the model is estimated.

V. Estimation of the Reduced Form Model

Substitution of $Z^*\gamma^*$ for $Z\gamma + \delta X(\beta_u-\beta_n)$ in equation (17) yields the reduced form probit contribution to the log-likelihood function for VFU interpreted conditionally on nonunion status of

$$(20)\quad L_1 = VFU\ \ln\left(\Phi(Z^*\gamma^*)\right) + (1 - VFU)\ \ln\left(1 - \Phi(Z^*\gamma^*)\right) .$$

The vector Z* includes all variables which appear either in the structural union preference function (Z) or in the earnings functions (X). Those variables assumed to be in the earnings function are three dichotomous variables for different levels of education, labor market experience and its square, seniority with current employer and its square, and dichotomous variables for nonwhite, female, and Southern residence. The labor market experience measure is actual years worked for pay since age sixteen rather than the standard Age-Education-6. The variables assumed to be in the preference for union representation function (Z) include seniority with current employer and its square, and the dichotomous variables for nonwhite, female, and Southern residence. In addition, the Z function includes age, dichotomous variables for married with spouse present and the product of female and marital status, and three dichotomous variables for broad occupational groupings. The union of these sets of variables contains the sixteen variables plus a constant which make up Z*. All variables are defined in Table 1 along with their means and standard deviations. The base group for the sample consists of white, nonsouthern, single males with twelve years of education working in a blue-collar occupation.

Note that there are two sets of constraints implicit in this formulation. The first is that five variables are excluded from the structural VFU function while they are included in the earnings functions. These are the three education and two experience variables. The set of four overidentifying constraints is testable in the structural version, and such a test is performed in Section VII.[12] Second, five variables are also excluded from each earnings function, and these ten constraints are also theoretically testable. However, as discussed in the next section, difficulties in estimation and interpretation hinder the actual performance of an appropriate test.

The estimates derived for γ^* by maximizing the likelihood function implied by equation (20) over the 553 nonunion workers are contained in the first column of Table 2. At first glance the only variable which has a substantial effect on a nonunion worker's preference for union representation is race. No other variable is asymptotically significantly different from zero at conventional levels, and only six of the sixteen coefficients have estimates whose absolute values exceed their asymptotic standard errors.[13] While the relationship looks relatively flat, a likelihood ratio test of the hypothesis that all of the coefficients except the constant equal zero rejects the hypothesis at any reasonable level of significance.[14]

In order to investigate how sensitive nonunion worker preferences for union representation are in a number of dimensions, Table 3 contains values of $Pr(VFU=1|U=0)$ at the point estimates of the parameters contained in the first column of Table 2 for 30 year old single males with twelve years of education, ten years experience, five years seniority, and various occupations, race, and regions. It is clear that for any occupation and region, nonwhites are substantially more likely to desire union representation than are whites.[15] On the other hand, region has a trivial

Table 2: Estimates of Pr(VFU) and Pr(U)
Univariate and Bivariate Probit Models

Variables	Pr(VFU=1|U=0) Univariate (1)	Pr(U=1) Univariate (2)	Pr(VFU=1) Bivariate (3)	Pr(U=1) Bivariate (4)
Constant	-.151 (.271)	-.0852 (.253)	.444 (.358)	-.0933 (.254)
Ed < 12	.0530 (.172)	.0538 (.133)	.0641 (.150)	.0467 (.134)
12 < Ed < 16	-.190 (.150)	-.0942 (.130)	-.187 (.136)	-.120 (.129)
Ed ⩾ 16	.129 (.200)	.231 (.174)	.188 (.180)	.267 (.172)
Exp	.0142 (.0226)	.0151 (.0194)	.0169 (.0207)	.0147 (.0195)
Exp^2	-.000368 (.000398)	-.000316 (.000331)	-.000401 (.000371)	-.000336 (.000330)
Sen	.000255 (.0302)	.105 (.0232)	.0431 (.0362)	.110 (.0232)
Sen^2	-.00143 (.00129)	-.00224 (.000908)	-.00186 (.00111)	-.00239 (.000919)
NW	.827 (.180)	.295 (.141)	.787 (.192)	.303 (.142)
Fe	.221 (.197)	-.0513 (.167)	.187 (.178)	-.0394 (.168)
South	-.00481 (.118)	-.530 (.102)	-.218 (.167)	-.547 (.102)
Age	-.00469 (.0109)	-.0123 (.0104)	-.00936 (.00994)	-.0115 (.0105)
Marr	-.117 (.180)	.0777 (.142)	-.0436 (.169)	.0789 (.144)
Marr*Fe	-.273 (.243)	.00253 (.212)	-.260 (.219)	-.0331 (.213)
Cler	-.154 (.174)	-.738 (.151)	-.433 (.209)	-.749 (.151)
Serv	.121 (.181)	-.499 (.144)	-.128 (.221)	-.491 (.145)
Prof & Tech	-.209 (.193)	-.656 (.171)	-.446 (.216)	-.719 (.169)

(cont'd)

Table 2 (cont'd)

ρ_{23}		0	.778 (.366)
ln L	-337.8	-496.7	-833.1
n	553	880	880

(Numbers in parentheses are asymptotic standard errors.)

Table 3. Predicted Pr(VFU=1) for
Workers Varying by Race, Region, and Occupation

	Pr(VFU=1|U=0)		Pr(VFU=1)	
	Non-South	South	Non-South	South
Blue Collar				
white	.413	.411	.678	.596
nonwhite	.728	.726	.894	.848
Clerical				
white	.354	.352	.511	.425
nonwhite	.674	.673	.792	.725
Service				
white	.460	.458	.630	.546
nonwhite	.766	.765	.869	.839
Professional and Technical				
white	.334	.332	.398	.350
nonwhite	.654	.653	.702	.656

Computed for 30-year-old single males with 12 years education, ten years experience, and five years seniority. Pr(VFU=1|U=0) computed from estimates in column (1), Table 2. Pr(VFU=1) computed from estimates in column (3), Table 2.

effect, while the occupational variation is moderate.

The second column of Table 2 contains maximum likelihood estimates over the sample of 880 workers of the simple probit model of the union status of workers based on the likelihood function derived from equation (9) where the vector of variables (C) is the same set as Z*. These are consistent estimates of the reduced form empirical relationship described earlier and, given the ambiguity regarding its behavioral underpinnings, not much space will be allocated to evaluation of these results. Suffice it to say that nonwhites and nonsoutherners are more likely to be union members, as are younger workers and those with more seniority.[16] In addition, there are rather sharp occupational breaks which imply that blue-collar workers are most likely to be unionized, while clerical workers are least likely, holding other factors fixed. A likelihood ratio test of the hypothesis that all coefficients except the constant term are zero can be rejected at any reasonable level of significance.[17] This relationship regarding Pr(U) is re-estimated as a piece of each succeeding analysis and, as is expected, the results do not change substantially. Hence, no further discussion of Pr(U=1) will take place.

Estimation of the reduced form joint union preference-union status model proceeds by substituting $Z^*\gamma^*$ for $Z\gamma + \delta X(\beta_u - \beta_n)$ in equations (18) and (19), yielding

$$(21) \quad \Pr(VFU=1,\ U=0) = \int_{-\infty}^{-C\alpha} \int_{-Z^*\gamma^*}^{\infty} b(\varepsilon_3, \varepsilon_2;\ \rho_{23})\, d\varepsilon_3\, d\varepsilon_2$$

and

$$(22)\quad \Pr(VFU=0,\ U=0) = \int_{-\infty}^{-C\alpha} \int_{-\infty}^{-Z^*\gamma^*} b(\varepsilon_3, \varepsilon_2;\ \rho_{23}) d\varepsilon_3 d\varepsilon_2 \ .$$

Again, the Pr(U=1) is unchanged from equation (13). These expressions are combined with equation (14) to form the appropriate likelihood function. The maximum likelihood estimates of this model are contained in columns 3 and 4 of Table 2. Note that the estimates contained in the first two columns of this table relate to a constrained version of the joint model where $\rho_{23} = 0$. The estimated value of ρ_{23} is .778 with an asymptotic standard error of .366. The hypothesis that $\rho_{23} = 0$ can be rejected at the 5 percent level of significance using a two-tailed asymptotic t-test. The asymptotically equivalent likelihood ratio test can be performed by summing the log-likelihoods for the first two columns of Table 2 and comparing the constrained log-likelihood to the unconstrained value. Using this test, the hypothesis that $\rho = 0$ can be rejected at the 10 percent level of significance.[18] The positive value estimated for ρ_{23} suggests that unobserved factors which make workers more likely to work on union jobs also make these workers more likely to prefer union representation.

The estimates on the Pr(VFU=1) function unconditional on union status contained in the third column of Table 2 are much better determined than those for the conditional model (column 1). While only three of the sixteen coefficients are significantly different from zero at conventional levels, fully eleven of the sixteen coefficient estimates exceed their asymptotic standard error in absolute value. The effect of race on worker preferences for union representation is significantly different from zero at conventional levels, as are two of the three occupational variables. Both clerical and professional/technical workers are less likely than blue collar workers to prefer union representation. Southern workers are significantly less likely at the 10 percent level to prefer union representation than nonsouthern workers.

Table 3 contains values of Pr(VFU=1) at the point estimates of γ^* contained in the third column of Table 2 for thirty year old single males with twelve years of education, ten years experience and five years seniority. The hypothetical worker's occupation, race, and region are varied in order to investigate the sensitivity of Pr(VFU=1) to these factors. It is clear that race and occupation have large effects on workers' preferences for union representation. Nonwhites are substantially more likely to prefer union representation. Professional and technical workers are the least likely in terms of occupation to prefer union reprsentation, while blue collar workers are most likely. Southern workers are somewhat less likely than nonsouthern workers to desire union representation.

It is interesting to contrast the preferences of nonunion workers for union representation to workers' preferences unconditional on union status. The calculated probabilities contained in Table 3 facilitate this comparison. It is clear that workers in general are more likely to desire union representation than nonunion workers. For example, for the four types of blue collar workers listed in Table 3 (combinations of race and region), the probability that workers in general desire union representation is on average

37 percent higher than the probability that workers desire union representation conditional on being nonunion. The difference is positive but less pronounced for the other listed occupations. While this result is not unexpected, there is nothing in the specification which guarantees it, and finding this result is evidence of the "reasonableness" of the estimates.

Two other differences are that region plays a much greater role in determining overall preferences for unionization than it does among the nonunion workforce. In addition, the occupational distinctions are much greater among the workforce in general than among the nonunion workforce. These results are doubtless the result of "sorting," and it will be discussed in more detail below in the context of estimation of the structural model.

VI. Estimation of the Union-Nonunion Wage Differential

In order to estimate the parameters of the structural model, consistent estimates of the parameters of the union and nonunion earnings functions,

$$\ln W_u = X\beta_u + \varepsilon_u \quad \text{and}$$

(15)

$$\ln W_n = X\beta_n + e_n ,$$

must be derived. However, W_u is observed only for union workers, while W_n is observed only for nonunion workers. The reduced form empirical model which described the union status of workers was specified as

(7) $S = C\alpha + \varepsilon_2$,

where $Pr(U=1) = Pr(S > 0) = Pr(\varepsilon_2 > -C\alpha)$. It is straightforward to show that if the random component of the union status decision (ε_2) is correlated with the random components of earnings (ε_u and ε_n), then ordinary least squares (OLS) applied to the two equations in (15) separately will yield biased and inconsistent estimates for β_u and β_n. This so-called sample selection bias is due to the same sort of data censoring which was described above in relation to the missing data for union workers on their preferences for union representation. Given this problem, two alternative estimation procedures are developed.

The bias is introduced by the fact that the union and nonunion earnings functions are estimated only over their respective subsamples. This causes the expectations of ε_u and ε_n to vary by observation because they are only observed conditionally on union status. More formally, $E(\varepsilon_u|U=1) = E(\varepsilon_u|\varepsilon_2 > -C\alpha)$ and $E(\varepsilon_n|U=0) = E(\varepsilon_n|\varepsilon_2 < -C\alpha)$. If ε_u and ε_n are not independent of ε_2, then these conditional expectations vary with $C\alpha$. Assuming joint normality of ε_u, ε_n, and ε_2 results in

(23) $$E(\varepsilon_u|\varepsilon_2 > -C\alpha) = \sigma_u\rho_{2u}\frac{\phi(C\alpha)}{\Phi(C\alpha)}$$

and

(24) $$E(\varepsilon_n|\varepsilon_2 < -C\alpha) = -\sigma_n\rho_{2n}\frac{\phi(C\alpha)}{1 - \Phi(C\alpha)}$$

where σ_u and σ_n are the standard deviations of ε_u and ε_n respectively, ρ_{2u} and ρ_{2n} are the correlations between ε_2 and ε_u and between ε_2 and ε_n respectively, and $\phi(\cdot)$ is the standard normal density function. The quantities $\frac{\phi(\cdot)}{\Phi(\cdot)}$ and $\frac{\phi(\cdot)}{1 - \Phi(\cdot)}$ are the "hazard rates" or inverse "Mill's ratios" of union and nonunion status respectively. They will be called HR_u and HR_n.

The first approach to consistent estimation of β_u and β_n is a two-step procedure described in detail by Lee (1979). Write the earnings functions conditional on union status as

$$(25)\quad \ln W_u = X\beta_u + (\varepsilon_u | U = 1)$$

and

$$(26)\quad \ln W_n = X\beta_n + (\varepsilon_n | U = 0)\ .$$

The conditional error terms can be written as

$$(27)\quad (\varepsilon_u | U = 1) = E(\varepsilon_u | U = 1) + \theta_u$$

and

$$(28)\quad (\varepsilon_n | U = 0) = E(\varepsilon_n | U = 0) + \theta_n$$

where θ_u and θ_n are random components with zero mean. Substitution for the error terms in equations (25) and (26) yields

$$(29)\quad \ln W_u = X\beta_u + \lambda_u HR_u + \theta_u$$

and

$$(30)\quad \ln W_n = X\beta_n + \lambda_n HR_n + \theta_n$$

using the conditional expectations derived in equations (23) and (24) and the definitions of HR_u and HR_n. The parameters λ_u and λ_n represent $\sigma_u\rho_{2u}$ and $-\sigma_n\rho_{2n}$ respectively. If HR_u and HR_n are observed for union and nonunion workers respectively, then OLS can be applied to these conditional earnings functions, and consistent estimates of β_u, β_n, λ_u, and λ_n will be obtained.

Although HR_u and HR_n are not observed directly, they are strictly functions of $C\alpha$, and the latter can be estimated consistently from the simple probit model of union status derived earlier. The maximum likelihood estimates of α from the model are contained in the second column of Table 2. They were used to compute consistent estimates of HR_u and HR_n which can then be used to compute consistent estimates of β_u and β_n by OLS as described above.

The second and more efficient approach to consistent estimation of β_u and β_n is to derive the likelihood function of the switching regression model defined by the two earnings functions and the union status function, and to use the likelihood function to derive the maximum likelihood estimates of the parameters. The contribution to the appropriate log-likelihood function for an individual is[19]

$$(31)\quad L = U \ln\Big(\int_{-C\alpha}^{\infty} f_{2u}(\ln W_u - X_u\beta_u,\ \varepsilon_2)d\varepsilon_2\Big)$$

$$+ (1 - U)\ \ln\Big(\int_{-\infty}^{-C\alpha} f_{2n}(\ln W_n - X_n\beta_n,\ \varepsilon_2)d\varepsilon_2\Big)$$

where $f_{2u}(\cdot,\cdot)$ and $f_{2n}(\cdot,\cdot)$ represent the bivariate normal densities of ε_2 and ε_u and of ε_2 and ε_n respectively. The parameters of the model are β_u, β_n, α, ρ_{2u}, ρ_{2n}, σ_u^2, and σ_n^2.

The OLS estimates of β_u and β_n are contained in columns (1) and (3) of Table 4. These results are atypical in a number of respects. Although the average union-nonunion differential computed for these estimates $(\bar{X}(\beta_u-\beta_n))$

Table 4. Estimate of Union and Nonunion Earnings Functions

Variable	ln W_u (1)	ln W_u (2)	ln W_n (3)	ln W_n (4)
Constant	1.27 (.175)	1.38 (.331)	1.19 (.0872)	1.24 (.127)
Ed < 12	.0490 (.141)	.0413 (.140)	-.195 (.0796)	-.189 (.076)
12 < Ed < 16	.135 (.151)	.156 (.158)	.184 (.0737)	.177 (.0771)
Ed ⩾ 16	.562 (.143)	.573 (.144)	.337 (.0788)	.335 (.0746)
Exp	.0480 (.0177)	.0480 (.0173)	.0244 (.00847)	.0243 (.00802)
Exp^2	-.000923 (.000346)	-.000905 (.000344)	-.000470 (.000180)	-.000479 (.000171)
Sen	.0416 (.0255)	.0343 (.0311)	.0523 (.0144)	.0571 (.0166)
Sen^2	-.00150 (.000925)	-.00134 (.000992)	-.00167 (.000623)	-.00177 (.000619)
NW	-.337 (.147)	-.352 (.150)	-.175 (.0862)	-.162 (.0854)
Fe	-.429 (.114)	-.405 (.128)	-.354 (.0585)	-.371 (.0645)
South	.0261 (.123)	.0657 (.157)	.0130 (.0568)	-.00909 (.0726)
HR_u	---	-.112 (.283)	---	---
HR_n	---	---	---	-.0911 (.176)
SEE	.918	.904	.652	.617
n	327	327	553	553

HR_u = Hazard Rate (inverse Mills' ratio) for union workers computed from estimates in column (2) of Table 2.

HR_n = Hazard Rate (inverse Mills' ratio) for nonunion workers computed from estimates in column (2) of Table 2.

The numbers in parentheses are standard errors. These are asymptotic and corrected in columns (2) and (4).

is positive (.264), previous evidence and experience with other data sets suggests that the union earnings function (even conditional on union status) ought to be flatter in virtually every dimension.[20] In addition, the previous evidence suggests that the unexplained variance in earnings is larger in the nonunion sector than in the union sector. The common explanation is that unions standardize wage rates by attaching wages to jobs rather than to workers. This reduces a union employer's discretion to vary wges according to individual characteristics.[21] However, the estimates presented here do not indicate this standardization, and it suggests that caution be exercised in interpreting the results derived from these data.

Closer examination of the results and comparison with the results obtained with other data sets suggests that it is the estimates of β_u which are "odd" rather than those of β_n. One approach to solving this problem might be to use a more "representative" group of union workers. However, the well known difficulties involved with choice-based sampling preclude such an approach. Hence, the analysis continues with the current data.[22]

Consistent estimates of β_u, β_n, λ_u, and λ_n obtained by applying OLS to equations (29) and (30) are contained in columns (2) and (4) of Table 4. The asymptotic standard errors and the standard error of estimation (SEE) are corrected through use of the consistent estimates of σ_u^2 and σ_n^2 rather than those printed by the OLS program.[23] The first thing to note is that the estimate of both λ_u and λ_n have relatively large asymptotic standard errors so that, although the hypotheses that λ_u and λ_n are zero cannot be rejected at conventional levels of significance, it is not possible to determine the potential for selection bias very precisely. The point estimates of the parameter vectors β_u and β_n are virtually identical to those derived using OLS without including the "selectivity regressors." The predicted average union-nonunion wage differential is $\overline{X}(\beta_u-\beta_n) = .325$, which is somewhat higher than that derived using OLS without HR_u and HR_n.

The maximum likelihood estimates of the switching regression model defined in equation (31) are contained in Table 5. The estimates of β_u and β_n differ substantially from those contained in Table 4, but the results are, if anything, less intuitively appealing than the earlier estimates. The union earnings function still does not exhibit the sort of standardization of rates expected of it and, in addition, the average union-nonunion wage differential is $\overline{X}(\beta_u-\beta_n) = -1.01$. This large negative differential suggests that an average worker earns in a union job only 36 percent of what could be earned in a nonunion job. This, of course, does not accord with any reasonable view of the union-nonunion wage differential debate.

Another somewhat surprising aspect of the results, particularly given the lack of significance of λ_u and λ_n in the "selectivity regression" model, are the maximum likelihood estimates of ρ_{2u} and ρ_{2n}. These are estimated to be large (.841 and .833 respectively) with very small standard errors (.0272 and .0182 respectively). The hypothesis that both correlations are zero can be rejected at any reasonable level of significance using a likelihood ratio test.[24] Note that the estimated correlations are so close to each other as to be virtually identical. The hypothesis that $\rho_{2u} = \rho_{2n}$ cannot be rejected

Table 5. Maximum Likelihood Estimates of Switching Regression Model

Variable	$\ln W_u$ (1)	$\ln W_n$ (2)	Pr(U) (3)
Constant	.438 (.323)	1.50 (.130)	-.155 (.280)
Ed < 12	.143 (.288)	-.156 (.109)	.0282 (.149)
12 < Ed < 16	-.0327 (.300)	.130 (.111)	-.154 (.149)
Ed ⩾ 16	.367 (.294)	.359 (.105)	.293 (.172)
Exp	.0414 (.0270)	.0247 (.0115)	.0109 (.0198)
Exp^2	-.00101 (.000481)	-.000537 (.000232)	-.000258 (.000358)
Sen	.110 (.0480)	.0799 (.0197)	.103 (.0253)
Sen^2	-.00306 (.00183)	-.00215 (.000833)	-.00231 (.00101)
NW	-.313 (.247)	-.0488 (.120)	.231 (.167)
Fe	-.654 (.169)	-.453 (.0855)	.250 (.185)
South	-.265 (.173)	-.142 (.0868)	-.400 (.108)
Age	--	--	-.0114 (.0104)
Marr	--	--	-.00799 (.147)
Marr*Fe	--	--	.150 (.221)
Cler	--	--	-.314 (.156)
Serv	--	--	.137 (.153)
Professional & Technical	--	--	-.481 (.153)

(continued)

Table 5 (cont'd)

ρ_{2u} = .841 (.0272)	ρ_{2n} = .833 (.0182)	ln L = -1417.6
$\sigma_u 2$ = 1.30 (.101)	$\sigma_n 2$ = .564 (.0171)	n = 880

(numbers in parentheses are asymptotic standard errors) at any reasonable level of significance using an asymptotic t-test.[25] The identity of these correlations is what would be expected if the earnings function errors (ε_u and ε_n) for any individual were perfectly correlated with each other. However, in the absence of longitudinal data it is impossible to identify ρ_{un}, so this must remain conjecture.

Since the estimated wage differential (ΔW) is what is of importance for the model of worker preferences for union representation, it is interesting to compare the three sets of estimates of β_u and β_n with regard to their implications of ΔW. Toward this end, Table 6 contains the coefficients of ΔW ($\beta_u - \beta_n$) for each of the three sets of estimates along with their standard errors. As expected, the estimates from the OLS and OLS augmented with the "selectivity regressors" are very similar. The estimates for the maximum likelihood model differ somewhat from the first two sets, but the major difference is in the sharply negative constant term. This is what yields the large negative average differential mentioned above, and it suggests that the lower average differential estimated using the maximum likelihood estimates (MLE) is an "across-the-board" reduction rather than associated primarily with particular groups, although some groups (nonwhites, females, and southerners) do have lower estimates of $\beta_u - \beta_n$ using MLE than using the two OLS techniques.

Overall, none of the results presented here offers much help in choosing a "best" measure of ΔW to use in the structural estimation. The maximum likelihood estimates are theoretically the best, but the large negative differentials estimated using MLE are counterintuitive. In addition, the two

Table 6. Estimates of Determinants of ΔW ($\beta_u - \beta_n$)

	ΔW^a OLS (1)	ΔW^b OLS with Hazard Rate (2)	ΔW^c MLE (3)
Constant	.080 (.20)	.140 (.355)	-1.06 (.346)
Ed < 12	.244 (.162)	.230 (.159)	.299 (.292)
12 < Ed < 16	-.049 (.168)	-.021 (.176)	-.163 (.309)
Ed ≥ 16	.225 (.163)	.238 (.162)	.008 (.295)
Exp	.0236 (.0196)	.0237 (.0191)	.0167 (.0282)
Exp^2	-.000453 (.000390)	-.000426 (.000384)	-.000473 (.000508)
Sen	-.0107 (.0293)	-.0228 (.0353)	.0301 (.0488)
Sen^2	.00017 (.00112)	.00043 (.00117)	-.00091 (.00188)
NW	-.162 (.170)	-.190 (.173)	-.264 (.270)
Fe	-.075 (.128)	-.034 (.143)	-.201 (.185)
South	.0131 (.135)	.0566 (.173)	-.123 (.184)

[a]Computed from estimates contained in columns (1) and (3) of Table 4. Standard errors are in parentheses.

[b]Computed from estimates contained in columns (2) and (4) of Table 4. Corrected asymptotic standard errors calculated assuming no covariances between estimates of β_u and β_n are in parentheses.

[c]Computed from estimates contained in Columns (1) and (2) of Table 5. Asymptotic standard errors are in parentheses.

sets of "consistent" estimates are likely to be sensitive to distributional and other specification assumptions. Finally, the reduced form probit relationship for union status used in the analysis may be inadequate for the purpose of correcting for selection bias in wage equations due to its shortcomings outlined above. Given the lack of a clear guide to the right set of estimates of ΔW, the analysis continues using all three measures so that their performance can be compared.

VII. Estimation of the Structural Model

Given the estimates of ΔW derived in the last section, the structural version of the model of worker preference for union representation can be estimated. This allows estimation of the effects of individual characteristics on worker preferences after controlling for variation in the union-nonunion wage differential. Both the model conditional on nonunion status and the unconditional model are estimated.

Consistent estimates of the parameters of the structural version of the conditional model can be derived by maximizing the simple probit likelihood function derived from equation (17) over the sample of 553 nonunion workers. Unfortunately, the asymptotic standard errors derived from the matrix of second derivatives of the log-likelihood function are inconsistent in this case because they do not account for the fact that the predicted wage differentials are random variables themselves. While it is possible to derive corrected asymptotic standard errors for these estimates, a more straightforward technique is to use these consistent parameter estimates as starting values for one Newton step on the likelihood function relating to the overall model consisting of the two earnings functions, the worker preference function, and the union status function.

The contribution to this log-likelihood function is

$$(32)\quad L = U \ln\left(\int_{-C\alpha}^{\infty} f_{2u}\,(\ln W_u - X_u\beta_u,\ \varepsilon_2) d\varepsilon_2 \right)$$

$$+ (1 - U)\ \mathrm{VFU}\ \ln\left(\int_{-\infty}^{-C\alpha} \int_{\kappa_3}^{\infty} h(\varepsilon_3,\ \varepsilon_2,\ \ln W_n - X_n\beta_n) d\varepsilon_3 d\varepsilon_2 \right)$$

$$+ (1 - U)(1 - \mathrm{VFU})\ \ln\left(\int_{-\infty}^{-C\alpha} \int_{-\infty}^{\kappa_3} h(\varepsilon_3,\ \varepsilon_2,\ \ln W_n - X_n\beta_n) d\varepsilon_3 d\varepsilon_2 \right)$$

where $f_{2u}(\cdot,\cdot)$ represents the bivariate normal density function of ε_2 and ε_u, $h(\cdot,\cdot,\cdot)$ represents the trivariate normal density function of ε_3, ε_2, and ε_n, and the quantity $\kappa_3 = -Z\gamma - \delta X(\beta_u - \beta_n)$. For any individual, this contribution represents the joint probability density of observing their preference for union representation, union status, and wage rate. A critical element of this likelihood function is the covariance matrix of the four errors. This is

$$\mathrm{cov}\begin{pmatrix} \varepsilon_3 \\ \varepsilon_2 \\ \varepsilon_u \\ \varepsilon_n \end{pmatrix} = \begin{pmatrix} \sigma_3^{\,2} & \sigma_{23} & \sigma_{3u} & \sigma_{3n} \\ \sigma_{23} & \sigma_2^{\,2} & \sigma_{2u} & \sigma_{2n} \\ \sigma_{3u} & \sigma_{2u} & \sigma_u^{\,2} & \sigma_{un} \\ \sigma_{3n} & \sigma_{2n} & \sigma_{un} & \sigma_n^{\,2} \end{pmatrix} .$$

Of the ten unique elements of this covariance matrix, two (σ_{3u} and σ_{un}) do not appear in the likelihood function and hence are not estimable. Two elements (σ_3^2 and σ_2^2) are normalized to one in order to fix the scale of the probit parameters (C, γ, and δ). This leaves six elements of the covariance matrix which must be estimated (σ_{23}, σ_{3n}, σ_{2u}, σ_{2n}, σ_u^2, and σ_n^2).

As it is written, this model is not conditional on nonunion status. However, given the joint normality of the errors, the conditional model is equivalent to the constrained version of the joint model where $\rho_{23} = 0$. Imposing this constraint and taking one Newton step on the entire sample from the appropriate consistent estimates yields consistent and asymptotically efficient estimates with consistent asymptotic standard errors. These estimates of the Pr(VFU) function are contained in Table 7 for the three different measures of ΔW.[26] Examination of the point estimates of the structural parameter (δ), which is the coefficient of the wage differential, yields the result that a nonunion worker's union-nonunion wage differential has a positive effect on preference for union representation. However, the effect is not asymptotically significantly greater than zero at conventional levels of significances for any of the three measures of ΔW.[27] Given the large difference between the estimated ΔW derived from the MLE switching regression and the other two measures, it is interesting that they yield roughly the same result. This is likely due to the fact that the major differences in the three measures of ΔW lay in the constant term (see Table 6), and this would explain the differences in the constant term between the three preference models estimated using the three measures of ΔW. The relatively large constant in the ΔW_{MLE} model is due to the relatively large negative mean of ΔW_{MLE}.

The estimates of the other parameters are virtually identical across the three versions. Nonwhite nonunion workers have a much larger probability of preferring union representation after controlling for the wage effect of unions. However, this result must be interpreted with caution due to the fact that the union-nonunion wge differential is estimated to be smaller for nonwhites than for whites. This is contrary to previous evidence, which suggests that nonwhites receive a larger wage advantage than whites from unionization (Ashenfelter, 1972).

An interesting relationship is found between sex and marital status and the desire for union representation among nonunion workers. Using the estimates contained in the first column of Table 7, nonunion single females are significantly more likely at the twelve percent level to prefer union representation than are nonunion single males, and they are significantly more likely at the five percent level to prefer union representation than are nonunion married males. However, married nonunion females behave in the opposite manner. Their preferences for union representation is significantly less than that of nonunion single females at the two percent level. In addition, the preference of married nonunion females for union representation does not differ significantly from either single or married nonunion males.

The overidentifying restrictions embedded in the structural model can be tested by noting that the reduced form estimates contained in column (1) of Table 2 represent an uncontrained version of the structural model. The relevant likelihood ratio test has four restrictions (representing the five variables included in the earnings functions but excluded from the preference

Table 7. Two-Step Estimates of Structural Probit Likelihood Function on Pr(VFU=1|U=0) with Different Measures of ΔW

Variable	(1)	(2)	(3)
Constant	-.210 (.389)	-.00173 (.602)	.333 (.537)
Sen	.00847 (.0534)	.0116 (.0605)	-.0155 (.0434)
Sen^2	-.00155 (.00213)	-.00160 (.00217)	-.000875 (.00174)
NW	.919 (.323)	.957 (.332)	.973 (.264)
Fe	.297 (.246)	.289 (.262)	.348 (.239)
Marr	-.104 (.188)	-.0985 (.188)	-.0880 (.188)
Marr*Fe	-.272 (.251)	-.272 (.252)	-.313 (.250)
South	-.0194 (.189)	-.0389 (.270)	.0392 (.158)
Age	-.00882 (.00732)	-.00945 (.00781)	-.00531 (.00618)
Cler	-.134 (.173)	-.134 (.174)	-.139 (.177)
Serv	.117 (.187)	.105 (.188)	.114 (.194)
Prof & Tech	-.186 (.183)	-.186 (.185)	-.0848 (.175)
ΔW_{OLS}	.701 (.775)	--	--
ΔW_{HR}	--	.737 (.794)	--
ΔW_{MLE}	--	--	.532 (.487)
ln L	-338.5	-338.7	-338.7

(cont'd)

Notes: Estimates computed by taking one Newton step on full four equation likelihood function from initial consistent estimates derived assuming $\rho_{23} = 0$.

Initial consistent estimates of γ and δ were derived by maximizing the likelihood function in equation (17). Consistent estimates of α are contained in column (2) of Table 2. An initial consistent estimate of ρ_{3n}, which only appears in the full likelihood function, was derived by grid search using consistent estimates of the other parameters. Initial consistent estimates of the other parameters were derived as follows:

Column 1: Initial consistent estimates of β_u, β_n, σ_u^2 and σ_n^2 were computed from the estimates in columns (1) and (3) of Table 4. The parameters ρ_{2u} and ρ_{2n} were constrained to zero.

Column 2: Initial consistent estimates of β_u, β_n, σ_u^2, σ_n^2, ρ_{2u}, and ρ_{2n} were computed from estimates in columns (2) and (4) of Table 4.

Column 3: Initial consistent estimates of β_u, β_n, σ_u^2, σ_n^2, ρ_{2u}, and ρ_{2n} were computed from estimates in columns (1) and (2) of Table 5.

The values of the log-likelihood function are based on the initial consistent estimates of $Pr(VFU=1|U=0)$.

These estimates are asymptotically efficient. The numbers in parentheses are asymptotic standard errors.

$n = 880$

function less one for ΔW). For none of the three measures of ΔW can the constrained structural model be rejected at conventional levels of significance.[28]

Consistent estimates of the structural version of the union preference model unconditional on union status can be derived based on the likelihood function defined by equations (18), (19), (13), and (14). However, as in the case described above, the estimated asymptotic standard errors are inconsistent due to the randomness of the predicted wage differentials. Asymptotically efficient two-step estimates of the parameters of the structural model are derived with corrected standard errors by taking one Newton step on the log-likelihood function defined in equation (32) from the initial consistent estimates without the constraint that $\rho_{23} = 0$. These estimates of the parameters of the Pr(VFU) and Pr(U) functions for the three measures of ΔW are contained in Tables 8 and 9.

The estimates contained in Table 7 when combined with the estimates of Pr(U=1) contained in the second column of Table 2 relate to a constrained version of the joint model where $\rho_{23} = 0$. The point estimates, contained in Table 8, for ρ_{23} are all significantly different from zero at conventional levels so that the constrained model can be rejected. The estimated positive correlation suggests that workers who are more likely for unobserved reasons to desire union representation are also more likely to work on a union job.

The central hypothesis of the structural model is that workers with high union-nonunion wage differentials will be more likely to desire union representation. While the point estimates of the coefficients of ΔW are

Table 8. Two-Step Estimates of Structural Bivariate Probit Likelihood Function on Pr(VFU) with Different Measures of ΔW

Variable	(1)	(2)	(3)
Constant	.405 (.474)	.545 (.681)	.874 (.558)
Sen	.0547 (.0593)	.0530 (.0677)	.0213 (.0403)
Sen^2	-.00200 (.00215)	-.00194 (.00227)	-.00139 (.00159)
NW	.890 (.358)	.923 (.373)	1.04 (.280)
Fe	.260 (.240)	.246 (.264)	.246 (.241)
Marr	-.0284 (.172)	-.0353 (.171)	-.0499 (.164)
Marr*Fe	-.260 (.223)	-.256 (.223)	-.285 (.224)
South	-.247 (.228)	-.259 (.298)	-.139 (.159)
Age	-.0131 (.00714)	-.0133 (.00760)	-.00767 (.00600)
Cler	-.425 (.201)	-.412 (.197)	-.352 (.164)
Serv	-.146 (.214)	-.136 (.209)	-.0591 (.174)
Prof & Tech	-.412 (.201)	-.404 (.200)	-.225 (.164)
ΔW_{OLS}	.790 (.811)	--	--
ΔW_{HR}	--	.815 (.834)	--
ΔW_{MLE}	--	--	.597 (.513)
ρ_{23}	.814 (.323)	.771 (.308)	.654 (.0953)
ln L	-834.0	-834.2	-834.6

(cont'd)

Notes: Estimates computed by taking one Newton step on full four equation likelihood function from initial consistent estimates derived from maximizing the likelihood function defined by equations (18), (19), (13), and (14), and using the appropriate estimate of β_u, β_n, and the covariance parameters. An initial consistent estimate of ρ_{3n}, which appears only in the full likelihood function, was derived by grid search using consistent estimates of the other parameters. The numbers in parentheses are asymptotic standard errors.

See notes to Table 7 for the sources of the consistent estimates of β_u, β_n, and the covariance parameters.

Ln L is computed based on the initial consistent estimates.

n = 880

Table 9. Two-Step Estimates of Structural Bivariate Probit Likelihood Function of Pr(U) with Different Measures of ΔW

Variable	(1) ΔW_{OLS}	(2) ΔW_{HR}	(3) ΔW_{MLE}
Constant	-.0898 (.258)	-.0972 (.257)	-.161 (.226)
Ed < 12	.0856 (.136)	.0865 (.136)	.114 (.146)
12 < Ed < 16	-.0782 (.130)	-.0687 (.129)	-.0828 (.133)
Ed ⩾ 16	.270 (.180)	.280 (.181)	.198 (.171)
Exp	.0151 (.0195)	.0150 (.0194)	.0141 (.0173)
Exp^2	-.000319 (.000333)	-.000329 (.000326)	-.000362 (.000321)
Sen	.111 (.0238)	.112 (.0240)	.116 (.0243)
Sen^2	-.00244 (.000935)	-.00250 (.000940)	-.00265 (.000944)
NW	.297 (.147)	.297 (.148)	.303 (.155)
Fe	-.0376 (.174)	-.0459 (.174)	-.439 (.150)
South	-.550 (.103)	-.552 (.103)	-.555 (.0973)
Age	-.0125 (.0103)	-.0121 (.0104)	-.0108 (.00715)
Marr	.0738 (.148)	.0666 (.148)	.0638 (.126)
Marr*Fe	-.0276 (.218)	-.0215 (.217)	.0121 (.165)
Cler	-.746 (.154)	-.747 (.154)	-.640 (.132)
Serv	-.493 (.151)	-.486 (.151)	-.358 (.131)
Prof & Tech	-.709 (.171)	-.725 (.172)	-.578 (.140)

Note: Estimates of Pr(VFU), ρ_{23}, and the value of ln L are in Table 8. See notes to Tables 7 and 8.

n = 880

positive for all three measures, in no case are they significantly greater than zero at conventional levels.[29] However, for the ΔW_{MLE} version the coefficient is significantly greater than zero at the .13 level.

The estimates of the parameters of the Pr(VFU) function are similar across the three measures of ΔW. Thus, in order to facilitate the discussion, only the parameters derived using ΔW_{OLS} will be examined explicitly.

Southern workers are significantly less likely to desire union representation than are nonsouthern workers at the fifteen percent level. This contrasts with the result that southern nonunion workers are no less likely than nonsouthern nonunion workers to desire union representation. This could explain in part the relatively low level of unionization which coexists with currently comparable levels of new organization in the two regions.[30] Specifically, the relatively more numerous nonsouthern workers who desire union representation are already union members, leaving in the nonunion sector a group of workers who are less likely to desire union representation and hence comparable to their nonunion southern brethren.

Older workers are significantly less likely to desire union representation than are younger workers after controlling for seniority. Marital status has an insignificant effect on male worker's preferences for union representation. On the other hand, single females are significantly more likely to desire union representation than are either males or married females. Married females are indistinguishable from males on this basis.

Sharp occupational distinctions arise in worker preferences for union representation. Clerical and professional and technical workers are significantly less likely to desire union representation than are either blue collar or service workers. Again, these distinctions do not exist conditional on nonunion status, and the same sorting argument which was made above for southern versus nonsouthern workers can be made here.

The overidentifying restrictions used to identify the coefficient of ΔW can be tested by noting that the structural model is a constrained version of the reduced form model whose estimates are contained in Table 2. A likelihood ratio test of these four overidentifying restrictions fails to reject the constrained model at reasonable levels of significance.[31]

Overall, the results concerning the structural model are mixed. For all three measures of ΔW, the effect of the wage differential on worker preferences for union representation is positive but not significantly greater than zero at conventional levels. However, this may be due more to imprecision in estimating ΔW rather than a problem with the structural specification itself. Evidence of sorting was found in a number of dimensions, including region, occupation, and age. Among nonunion workers, little distinction in preferences for union representation could be found along these dimensions. However, after correcting for the sample censoring on union status, differences in prefrences were defined quite sharply along these dimensions.

VIII. Summary and Conclusions

A model of the determination of worker preferences for union representation was developed which led to the hypothesis of a positive relationship between a worker's preference in this area and the worker's union-nonunion wage differential. A distinction was drawn between the observed union status of workers and their current preferences for union representation, which was based on costs of job mobility and the existence of queues for union jobs. A rather unique set of data, the Quality of Employment Survey, was used because it contained a question the response to which indicated directly a worker's preference for union representation. Unfortunately, this particular bit of information was available only for nonunion workers.

A pair of econometric issues were raised. One had to do with techniques for handling the censored nature of the union preference information. The second had to do with appropriate techniques for estimating the central explanatory variable, the union-nonunion wage differential.

The censored data problem was handled by developing a reduced form empirical model to explain union status and hence the censoring under the assumption of joint normality of latent variables determining union status and preference for union representation. The union preference function was estimated using both the model conditional on nonunion status and, by accounting for the censored data, the model unconditional on union status. A comparison of the general nature of the results both yields some insight into the determination of the extent and locus of unionization and has important implications for prospects for organizing currently unorganized workers.

Overall, worker preferences for unionization among nonunion workers are rather flat in that there is little variation across workers with different characteristics.[32] On the other hand, a number of relatively sharp delinations in worker preferences for union representation along such dimensions as region and occupation occur in the model unconditional on union status. This suggests that many workers with those characteristics (both observed and unobserved) which make them likely to desire union representation are, in fact, union workers. The nonunion workers who are left are relatively homogeneous in their lack of interest in union representation. This interpretation is supported by the positive correlation estimated between the unobservable factors affecting preference for union representation and actual union status. In terms of the prospects for union organizing, this sorting suggests both that current nonunion workers will be less receptive to organizing efforts and that effective targeting of campaign efforts on the basis of gross characteristics such as region will not be terribly useful in light of the flatness of preferences.

The problem of the appropriate estimates of the union-nonunion wage differential (ΔW) arose because, as must be true in the absence of longitudinal data, only one wage or the other is observed for any individual. Apart from OLS applied separately to the two subsamples, two techniques were used to derive "consistent" estimates of ΔW under the assumption of joint normality of $\ln W$ and the latent variable determining union status. One technique (Mill's ratio or selectivity regressors) gave results similar to the OLS estimates. The other technique (maximum likelihood switching regression) gave vastly different and unreasonable results in that a large negative average differential was predicted. The sensitivity of these techniques to sample and specification is well known, and as a result the analysis of the structural model was carried out using all three measures.

The results of the estimation were remarkably similar across all three measures of ΔW. The central hypothesis was weakly supported in that the effect of ΔW on the propensity to desire union representation was positive in all cases, though only significantly so in one case. This weakness may be due to problems in estimating ΔW rather than to problems with the structural model.

In closing, two cautions are necessary. First, all of the results presented here were derived under the assumption of joint normality largely for computational convenience. The results may be sensitive to alternative distributional assumptions. Second, as was discussed earlier, the reduced form empirical probit model used to explain union status and hence to correct for sample censoring has rather ambiguous behavioral underpinnings. Indeed, part of the reason for carrying out this study was to improve our understanding of union status determination. Some progress has been made but more is yet to be done.

FOOTNOTES

1. There is reason to believe that there are advantages to union employment which are not offset completely by union dues and initiation fees paid by workers. This results in an excess demand for vacancies in existing union jobs. See Abowd and Farber (1982) for a more detailed discussion of this point.

2. See Farber and Saks (1980) for an analysis which focuses on the preferences of nonunion workers for union representation.

3. See Freeman and Medoff (1981) for a convincing discussion of the problems with standard sample selection correction techniques in the union wage effect context.

4. Freeman (1981) presents evidence on the relationship between unionization and fringe benefits.

5. It must be cautioned that these examples are not meant to imply specific empirical hypotheses. Any particular personal characteristic can be correlated with these utilities in a number of dimensions. The effect of these characteristics in equation (3) is the net effect of all of these dimensions.

6. See Quinn and Staines (1979) for a detailed description of the survey design.

7. The question of how workers form their expectations about what unions do is interesting and important. However, it is left to future research. Kochan (1979) presents an analysis of worker perceptions of unions based on the QES.

8. It is fallacious to argue that since union workers are in fact union workers voluntarily, they desire union representation. While it is true that they desired union representation when they took the job in the sense that it was part of a package of job characteristics which was preferred to any other package, the accumulation of seniority can reduce mobility so that a union worker may desire to retain his job but eliminate unionization. This does not mean that the worker will desire to quit.

9. The assumption of unit variance is a normalization required by the dichotomous nature of VFU in order to fix the scale of γ and δ.

10. The identity that $\Phi(a) = 1 - \Phi(-a)$ is used in deriving this expression.

11. Note that this is a different normalization than the one used above ($Var(\varepsilon_1) = 1$). This will result in a different scaling for the parameters, but the initial scaling was arbitrary to begin with.

12. The full set is not testable because X must contain at least one variable which is not contained in Z in order to identify δ. Thus, the test carried out below embodies only four restrictions.

13. This represents a level of significance of 32 percent with a two-tailed test or 16 percent with a one-tailed test using an asymptotic t-test.

14. The constrained log-likelihood is -366.1, while the unconstrained log-likelihood is -337.8. The test statistic is $-2(-366.1 - (-337.8)) = 56.8 > 34.3 = \chi^2_{.005}(16)$.

15. Care must be taken in interpreting these results because there is not a complete analysis of variance. In other words, a complete set of

interaction variables was not included. The analysis is done because of the nonlinearity inherent in the relationship between $Z^{*}\gamma^{*}$ and $Pr(VFU=1|U=0)$ and the resulting difficulty in interpreting the parameter estimates from probit models.

16. The estimated effect of seniority may be more the result of union status than an explanatory factor. It is well known that seniority is higher on union jobs through lower quit rates. See Table 1 for union and nonunion means on seniority as well as Freeman (1980) for an analysis of the relationship between union status and quit rates.

17. The constrained log-likelihood is -580.6, while the unconstrained log-likelihood is -496.7. The test statistic is $-2(-580.6 - (-496.7)) = 167.8 > 34.3 = \chi^2_{.005}(16)$.

18. The constrained log-likelihood is -834.5, while the unconstrained log-likelihood is -833.1. The test statistic is $-2(-834.5 - (-833.1)) = 2.8 > 2.71 = \chi^2_{.1}(1)$.

19. See Lee (1979) for a more detailed discussion of this likelihood function.

20. This can be verified using samples from the Panel Study of Income Dynamics (PSID), the National Longitudinal Survey (NLS), and the Current Population Survey (CPS). See, for example, Bloch and Kuskin (1978) and Abowd and Farber (1982).

21. See Webb and Webb (1920) for a classic discussion of the standard rate.

22. See Cosslett (1981) and Manski and Lerman (1977) for discussions of the choice-based sampling problem.

23. The technique for deriving the consistent estimates of σ_u^2 and σ_n^2 is described by Lee (1979). Briefly, the estimated residuals in each sector are regressed on a constant and the appropriate hazard rate multiplied by the estimated $C\alpha$. The estimated constant terms are consistent estimates of the residual variances.

24. The constrained log-likelihood, derived from the OLS estimates of β_u and β_n and the simple probit estimates of α, is -1470.0. The unconstrained log-likelihood is -1417.6. The test statistic is $-2(-1470.0 - (-1417.6)) = 104.8 > 10.6 = \chi^2_{.005}(2)$.

25. The quantity $\rho_{2u} - \rho_{2n} = .008$ with an asymptotic standard error of .0313. The t-statistic is .256, which is marginally significant only at the 60 percent level.

26. See the note to Table 7 for sources of the initial consistent estimates of the parameters in each of the three cases.

27. It is interesting that examination of the inconsistent asymptotic standard errors derived from the initial consistent estimates suggests that for all three measures of ΔW the effect of the wage difference on Pr(VFU) is significantly greater than zero at the .06 level. Since these estimated standard errors would be correct under the assumption that the estimated differentials were in fact the actual differentials, this implies that the lack of precision in estimation of ΔW is what is causing the relatively large standard errors on δ.

28. The unconstrained model has a log-likelihood of -337.8. The three versions of the constrained model have log-likelihoods computed using

the initial consistent estimates of -338.5, -338.7, and -338.7. The test statistic is minus twice the difference in the log-likelihoods, which yields values of 1.4, 1.8, and 1.8. The critical value of a χ^2 distribution with four degrees of freedom at the .75 level of significance is 1.92. This test is not strictly valid due to the unaccounted-for randomness of ΔW. However, the results are suggestive.

29. Once again, the inconsistent standard errors were small enough to allow rejection of the hypothesis that the coefficient of ΔW equals zero at the five percent level of significance. This suggests that it is the imprecision in the estimation of ΔW which is the cause of the relatively large standard errors. See footnote 27.

30. Evidence from the U.S. Bureau of Labor Statistics (1975, 1978) and from the National Labor Relations Board (1974) indicate that 1.2 percent of nonunion workers in the south were eligible to vote in NLRB-supervised representation elections in 1974. Outside the southern region, only 0.9 percent of nonunion workers were eligible to vote in such elections. (Eligibility refers to working in a potential bargaining unit where an election was held.) Of those workers who voted, 46 percent of workers in the south voted for union representation compared with 50 percent of nonsouthern workers. Similarly, union representation rights were won in 46 percent of the southern elections and in 51 percent of the nonsouthern elections. In both regions approximately .3 percent of nonunion workers were newly organized in 1974 as a result of NLRB-supervised elections.

31. The unconstrained model has a log-likelihood of -833.1. The three versions of the constrained model have log-likelihoods computed from the initial consistent estimates of -834.0, -834.2, and -834.6. The likelihood ratio test statistics are 1.8, 2.2, and 3.0. The critical value of a χ^2 distribution with four degrees of freedom at the 50 percent level is 3.36. This test is not strictly valid due to the unaccounted-for randomness of ΔW. However, the results are suggestive.

32. An exception to this is that nonunion nonwhites are substantially less likely than nonunion whites to desire union representation.

REFERENCES

Abowd, John M. and Farber, Henry S. "Job Queues and the Union Status of Workers." Industrial and Labor Relations Review (1982), forthcoming.

Ashenfelter, Orley. "Racial Discrimination and Trade Unionism." Journal of Political Economy 80, No. 3, Pt. 1 (May/June 1972): 435-64.

_______ and Johnson, George E. "Unionism, Relative Wages, and Labor Quality in U.S. Manufacturing Industries." International Economic Review 13, No. 3 (October 1972): 488-508.

Bloch, Farrell E. and Kuskin, Mark S. "Wage Determination in the Union and Non-Union Sectors." Industrial and Labor Relations Review 31, No. 2 (January 1978): 183-92.

Cosslett, Stephen R. "Maximum Likelihood Estimator for Choice-Based Samples." Econometrica 49, No. 5 (September 1981): 1289-1316.

Farber, Henry S. and Saks, Daniel H. "Why Workers Want Unions: The Role of Relative Wages and Job Characteristics." Journal of Political Economy 88, No. 2 (April 1980): 349-369.

Freeman, Richard B. "The Exit-Voice Tradeoff in the Labor Market, Unionism, Job Tenure, Quits, and Separations." Quarterly Journal of Economics 94 (June 1980): 643-673.

_______. "The Effect of Trade Unionism on Fringe Benefits." Industrial and Labor Relations Review 34, No. 4 (July 1981): 489-509.

_______ and Medoff, James L. "The Impact of Collective Bargaining: Illusion or Reality?" (July 1981) mimeo.

Kochan, Thomas A. "How American Workers View Labor Unions." Monthly Labor Review (April 1979): 23-31.

Kochan, Thomas A. and Helfman, David E. "The Effects of Collective Bargaining on Economic and Behavioral Job Outcomes." Working Paper No. 1181-81, Alfred P. Sloan School of Management, M.I.T. (January 1981).

Lee, Lung-Fei. "Unionism and Wage Rates: Simultaneous Equations Model with Qualitative and Limited Dependent Variables." International Economic Review 19 (1978): 415-433.

_______. "Identification and Estimation in Binary Choice Models with Limited (Censored) Dependent Variables." Econometrica 47, No. 4 (July 1979): 977-996.

Manski, Charles F. and Lerman, Steven R. "The Estimation of Choice Probabilities from Choice Based Samples." Econometrica 45, No. 8 (November 1977): 1977-1988.

National Labor Relations Board. Annual Report, 1974. Washington: Government Printing Office, 1974.

Quinn, Robert P. and Staines, Graham L. The 1977 Quality of Employment Survey: Descriptive Statistics, with Comparison Data from the 1969-70 and the 1972-1973 Surveys. Ann Arbor, Mich.: Institute for Social Research, 1979.

U.S. Bureau of Labor Statistics. Directory of National Unions and Employee Associations, 1975. Bulletin No. 1937. Washington: Government Printing Office, 1975.

_______. Handbook of Labor Statistics, 1978. Bulletin No. 2000. Washington: Government Printing Office, 1978.

Webb, Sidney and Webb, Beatrice. Industrial Democracy. 1920. Reprint ed. New York: Augustus M. Kelley, 1965.