Unionism and Wage Rates: A Simultaneous Equations Model with Qualitative and Limited Dependent Variables
Author(s): Lung-Fei Lee
Source: *International Economic Review*, Vol. 19, No. 2 (Jun., 1978), pp. 415-433
Published by: Wiley for the Economics Department of the University of Pennsylvania and Institute of Social and Economic Research, Osaka University
Stable URL: http://www.jstor.org/stable/2526310
Accessed: 20-03-2017 14:56 UTC

INTERNATIONAL ECONOMIC REVIEW
Vol. 19, No. 2, June, 1978

UNIONISM AND WAGE RATES: A SIMULTANEOUS EQUATIONS MODEL WITH QUALITATIVE AND LIMITED DEPENDENT VARIABLES*

BY LUNG-FEI LEE[1]

1. INTRODUCTION

A large number of studies have been made on the impact of labor unions on wage rates of workers. These studies generally have found positive and significant effects of unionism on wage rates. More recently, a few authors have studied the simultaneous effects between unionism and wages. Ashenfelter and Johnson [1972] used aggregated U.S. Manufacturing Industries Data and found that unionism had no significant impact on wage rates, but that the wage rate had a significant effect on the extent of unionism. Since the data they used were quite limited, they could only conclude that the magnitude of the effects of unionism on wage rates was uncertain. Schmidt and Strauss [1976] reached similar conclusions using microeconomic data. Their mixed logit approach, however, is not based on choice behavior and does not fit into a traditional econometric framework.

This study extends recent investigations of the joint determination of the extent of unionism and the effects of unions on wage rates, using microeconomic data from the Survey of Economic Opportunity Sample of 1967. Economic considerations suggest that the propensity to join a union depends on the net wage gains that might result from trade union membership. The explicit inclusion of this interdependence between the wage gain equation and the union membership equation in the model represents our point of departure from the previous work of Ashenfelter and Johnson and Schmidt and Strauss. The model is a variant of a traditional simultaneous equations model with a binary qualitative variable (union membership) and limited dependent variables. In Section 2, the conceptual framework of the model is discussed. Properties of the data are presented in Section 3. Section 4 briefly discusses estimation methods, and empirical estimates are presented in Section 5. Finally, in Section 6, conclusions are drawn.

* Manuscript received August 18, 1976; revised April 11, 1977.
[1] I would like to express my gratitude to G. S. Maddala and Sherwin Rosen for their supervision and encouragement during the writing of this paper. Thanks also to John Hause, Lawrence Olsen, Norman J. Simler and Christopher A. Sims for valuable comments and to Professor S. Rosen, Mr. Robert B. Litterman and two anonymous referees for suggestions on improving the presentation and correcting my English. Any errors remaining are solely my own. Financial support for SOC–76–04356 to the University of Florida is gratefully acknowledged.

2. THE UNION-MEMBERSHIP AND WAGE EQUATIONS MODEL

Since our purpose is to study the interaction between the labor unions and wage rates, the population that we are interested in consists primarily of full time workers. There are two options available to each worker: he can become a member of a union or not, depending on his own decision and the decision of the union. Each worker faces two wage rates, the union wage and the nonunion wage. Given the degree of union organization, however, choice of union membership is not free. Union initiation, fees, dues and other requirements have to be paid or met, and there are certain restrictions on entry (G. S. Becker [1959]). With these factors and some other taste factors, the individual becomes a union member or not, partially by his choice and partially by the selectivity of the labor union. With union or nonunion status determined, the workers' wage rates are set according to their socio-economic status and the job they perform.

Let W_{ui} and W_{ni} be the union and nonunion wages for individual i and let ρ_i be the reservation wage which summarizes his specific preferences. Individual i is assumed to join the union if

$$\frac{W_{ui} - W_{ni}}{W_{ni}} > \rho_i$$

i.e., he joins the union if the percentage union wage differential exceeds his reservation wage. The reservation wage, ρ_i, summarizes his receptivity for the labor union, and it can thus be either positive or negative. More specifically, we assume that ρ_i is a function of the characteristics of the person and the costs of becoming a union member. As a first order approximation,

$$\rho_i = \alpha X_i + \beta C_i + \varepsilon_{1i}$$

where X_i is a vector of individual characteristics, C_i is an index which summarizes the monetary costs and nonpecuniary costs of becoming a union member, and ε_1 is the error reflecting unobservable random factors. We assume throughout the model, that ε_{1i} is normally distributed with zero mean and variance $\sigma_{1\varepsilon}^2$.

Union organizers sell their services to union members. They organize workers, bargain with the employers, and must cover the costs of providing union services to members. A successful labor union may raise the wage rate along the demand curve for labor covered by union contracts to a certain level above that of the nonunionized sector (cf. Lewis [1959]). Variables such as the industrial concentration ratio are important. The average cost of unionizing is expected to be less if the size of establishments or firms is large. Also it is expected to be greater if unionized employees are more dispersed geographically. Costs are also greater for employees with large turnover rates. Hence unionism of full time male workers is likely to be more successful than unionism of the female and casual workers (Lewis [1959]). The costs, C_i, of union services may not accurately be reflected by union dues and fees. For certain kinds of monopolistic unions high admission fees or dues are used to restrict the number of applicants, as well as to discriminate against or show favoritism toward some applicants. However, trade unions have not used initiation fees extensively because they are not solely interested in pecuniary income. If the trade union uses only price rationing to accept its members, union dues, admission fees, or initiation fees are good indices for the cost C_i. It is, however, generally believed that most of the trade unions restrict entry through the use of non-price rationing methods which may be channeled through the unionized firm's employment policies. The firms may become more selective about their employees. If non-price rationing is also used by the union, dues and other fees do not provide a good proxy for C_i. The explicit fees charged are only a part of the total cost charged by the union to workers. Since there are no sample data that reflect this kind of rationing, and since data on union fees are scanty and unreliable, one possible method is to use an unobservable variables approach and assume that the explicit and implicit costs C_i for individual worker i are partitioned into two parts: an observable function of characteristics of the laborers and attributes of industries, and an unobservable residual. The measurable attributes of industries reflect the cost to the union of organizing the workers, while characteristics of the laborers reflect the selectiveness of employers' employment policy. That is, assume

$$C_i = \gamma_1 + \gamma_2 X_i + \gamma_3 Z_i + \varepsilon_{2i}$$

where X_i is a vector of characteristics of the workers i and Z_i is a vector of attributes of industry where the worker is employed. ε_{2i} captures the unobserved variables and is assumed to be $N(0, \sigma^2_{2\varepsilon})$. Thus the individual i joins the union if

$$\frac{W_{ui} - W_{ni}}{W_{ni}} > (\alpha + \beta\gamma_2)X_i + \beta\gamma_1 + \beta\gamma_3 Z_i + \varepsilon_{1i} + \beta\varepsilon_{2i}$$

This criterion may be written in the form of a probit model: if $I_i^* > 0$ worker i is in the union, otherwise not, where

$$I_i^* = \delta_0 + \delta_1\left(\frac{W_{ui} - W_{ni}}{W_{ni}}\right) + \delta_2 X_i + \delta_3 Z_i - \varepsilon_i.$$

In many of the studies of unionism specific characteristic variables and the unionism dummy variables are put in a wage regression implying that the effects of personal characteristics on the wage rates are independent of unionization status. Some authors have relaxed this assumption to allow some interactions (e.g., Rosen [1970]). Our model specifies complete interactions in the wage equation: one wage equation for union workers and one for nonunion workers. These are

$$W_{ui} = \theta_{u0} + \theta_{u1} X_{ui} + \theta_{u2} Z_{ui} + \varepsilon_{ui}$$

$$W_{ni} = \theta_{n0} + \theta_{n1} X_{ni} + \theta_{n2} Z_{ni} + \varepsilon_{ni}$$

where W_{ui}, W_{ni} are the union and nonunion wage rates for the individual i, X_i is a vector of personal characteristics and Z_{ui}, Z_{ni} are the attributes of the unionized and nonunionized industries respectively. ε_u, ε_n are random residuals which are assumed to be $N(0, \sigma_u^2)$ and $N(0, \sigma_n^2)$ respectively.

3. THE LIST OF EXPLANATORY VARIABLES IN THE MODEL

The natural logarithm of the wage rate is used as the dependent variable in the wage equations. Since $\log W_{ui} - \log W_{ni}$ is approximately equal to $\frac{W_{ui} - W_{ni}}{W_{ni}}$, we can write our model in terms of this variable, which simplifies estimation of the model and is consistent with the theory. The model to be estimated consists of three equations,

$$\log W_{ui} = \theta_{u0} + X_{ui}\theta_{u1} + Z_{ui}\theta_{u2} + \varepsilon_{ui}$$

$$\log W_{ni} = \theta_{n0} + X_{ni}\theta_{n1} + Z_{ni}\theta_{n2} + \varepsilon_{ni}$$

$$I_i^* = \delta_0 + \delta_1(\log W_{ui} - \log W_{ni}) + \delta_2 X_i + \delta_3 Z_i - \varepsilon_i$$

where $\varepsilon_u \sim n(0, \sigma_u^2)$, $\varepsilon_n \sim n(0, \sigma_n^2)$ and $\varepsilon \sim n(0, \sigma_\varepsilon^2)$.

The main data source is the Survey of Economic Opportunity (SEO). Only 1967 data are used since the 1966 data do not identify the union status. The 1967 survey does not report union dues or any union fees, but does have a 3 digit industrial classification variable, which allows construction of industry concentration ratios and a proxy variable for union power in the industry to which each worker belongs. Thus these survey data are superior to the other available survey data. However, the SEO data still are not rich enough for our purpose, since many relevant variables about the union are not available. We therefore must look for other relevant sources to complete the list of variables. In L. W. Weiss' Appendix [1966], there are some variables such as industrial concentration and the industrial percentage union coverage for the year 1965 which are useful proxy variables for our study.

The model has been estimated using those workers whose occupation is classified as operatives, that is semi-skilled workers. For skilled workers, union restrictions on entry are believed to be much stronger. Some unions require a certain number of years of specific training, such as apprenticeship, etc. Since more specific data are not available, there is a difficulty in using these data for skilled workers. Available data are more suitable for the study of semi-skilled workers since these types of restrictions are less relevant. Nonskilled laborers also could be studied using these data, but minimum wage regulations may have effects on the wage rate so that the model would have to be modified to deal with this additional truncation.

The exogenous variables consist primarily of socioeconomic, personal and individual characteristics. The socioeconomic indicators include regional location, city size, educational level, market experience and weeks worked per year. Personal characteristics include race, sex and health limitations. Education is assumed to have two opposing effects on the union status of workers. More educated individuals may expect more gain from individual bargaining and thus they are less likely to want to join a union. Once a union has been established, however, wage rates will rise and firms will be able to attract more highly educated persons. It is well known that age earnings profiles have strictly concave shapes (Becker [1964]). In our empirical estimation, we prefer to use market experience as an independent variable rather than age. Market experience is defined as age minus years of highest grade completed minus 6. To catch the nonlinearity of the earning profile, a second order term is also used.

Market experience, race, sex, health limitation and the weeks worked variables in the union status equation may represent the unionized firm's selectiveness, the union's cost to unionize the worker, and the individual's attitudes towards unions.

The industries in our sample include mining, construction, manufacturing (durable and nondurable goods), transportation, communication, utilities and sanitary services. The industrial concentration ratio is included in the union status equation since it is expected that it will be easier to organize the firms that are more concentrated. It is also expected that total union converage of the industry has a strong positive effect on the union wage (Rosen [1969]). Extent of unionism by industry is taken from Weiss [1966] and must be matched to individuals in the SEO data so that it is likely that there is a measurement error in this variable.

More completely, the exogenous variables X_{ui}, X_{ni}, X_i, Z_{ui}, Z_{ni} and Z_i used in the estimation consist of the following variables;

$X_{ui1} = X_{ni1} = X_{i1}$ = Northern-Eastern Region dummy variables; 1 for N.E., 0 for others (N.E.).

$X_{ui2} = X_{ni2} = X_{i2}$ = Northern-Central Region dummy variable; 1 for N.C., 0 for others (N.C.).

$X_{ui3} = X_{ni3} = X_{i3}$ = Southern Region dummy variable; 1 for S., 0 for the others (S.).

$X_{ui4} = X_{ni4} = X_{i4}$ = In SMSA dummy variable; 1 for SMSA size 500,000 or more, 0 for the others (UR_1).

$X_{ui5} = X_{ni5} = X_{i5}$ = Outside SMSA dummy variable; 1 for outside SMSA, 0 elsewhere (UR_2).

$X_{ui6} = X_{ni6} = X_{i6}$ = Highest Grade Completed, grade from 0 to 4, dummy variable; 1 for 0 to 4 grades, 0 for others (ED_1).

$X_{ui7} = X_{ni7} = X_{i7}$ = Highest Grade Completed, grade from 5 to 7, dummy variable; 1 for grade 5 to 7, 0 for the others (ED_2).

$X_{ui8} = X_{ni8} = X_{i8}$ = Highest Grade Completed, grade from 9 to 11, dummy variable; 1 for 9 to 11 grades, 0 for the others (ED_3).

$X_{ui9} = X_{ni9} = X_{i9}$ = Highest Grade Completed, grade 12, dummy variable; 1 for grade 12, 0 for the others (ED_4).

$X_{ui10} = X_{ni10} = X_{i10}$ = Highest Grade Completed, grade 13 or more, dummy variable; 1 for grade at least 13, 0 for the others (ED_5).

$X_{ui11} = X_{ni11} = X_{i11}$ = Years of Market Experience (ME).

$X_{ui12} = X_{ni12} = X_{i12} =$ Square of years of market experience (ME_2).
$X_{ui13} = X_{ni13} = X_{i13} =$ Race dummy variable; 1 for the white worker, 0 for the nonwhite (RACE).
$X_{ui14} = X_{ni14} = X_{i14} =$ SEX dummy variable; 1 for male worker, 0 for the female worker (SEX).
$X_{ui15} = X_{ni15} = X_{i15} =$ Health Limitation dummy variable; 1 for worker with health limitation, 0 for worker without health limitation (HLT).
$X_{i16} =$ Weeks worked from 1 to 26 dummy variable; 1 for the workers in this category, 0 for those workers not in this category (WK_1).
$X_{i17} =$ Weeks worked from 48 to 52 dummy variable; 1 for those workers in this category, 0 for the others (WK_2).
$Z_{ui1} = Z_{ni1} =$ Mining industry dummy variable; 1 for the mining industry, 0 for the others (IND_1).
$Z_{ui2} = Z_{ni2} =$ Construction industry dummy variable; 1 for the construction industry, 0 for the others (IND_2).
$Z_{ui3} = Z_{ni3} =$ Durable manufacturing industry dummy variable; 1 for the durable goods manufacturing, 0 for the others (IND_3).
$Z_{ui4} = Z_{ni4} =$ Nondurable manufacturing industry dummy variable; 1 for nondurable manufacturing, 0 for the other industries (IND_4).
$Z_{ui5} =$ Extent of union organization: percentage of union coverage in the industry (U).
$Z_{i1} =$ Industrial Concentration Ratio (CCR).
$W_{ui} =$ Union hourly wage rate, measured in cents.
$W_{ni} =$ Nonunion hourly wage rate, measured in cents.
$I_i =$ Union status of individual i; 1 for workers in labor union, 0 for workers not in the labor union.

4. THE ESTIMATION PROCEDURE

In this model, we observe the exogenous variables, the union status variable I_i and the limited dependent variables W_{ui} or W_{ni}. The observed wage rate depends on the worker's status, i.e., we observe

$$W_{ui} \text{ when } I_i = 1 \text{ and}$$
$$W_{ni} \text{ when } I_i = 0, \text{ but never both.}$$

Thus we have a simultaneous equations model involving qualitative and limited dependent variables.

In this model, the wage equations $\log W_{ui}$, $\log W_{ni}$ cannot in general be consistently estimated by ordinary least squares using the observed wage rates. The trouble occurs since

$$E(\varepsilon_u|I_i = 1) \neq 0 \quad \text{and} \quad E(\varepsilon_n|I_i = 0) \neq 0.$$[2]

Hence, other consistent estimation procedures have to be found. The estimation procedure we propose is a modified least squares procedure for the estimation of the wage equations. The idea of the procedure is to find the expression for the means $E(\varepsilon_u|I_i=1)$ and $E(\varepsilon_n|I_i=0)$ and adjust the error terms so that they will have zero means. This becomes a nonlinear least squares equation and the estimation of the wage rate equations proceeds in two stages. This kind of two stage estimation procedure was suggested by Amemiya [1974] for the Tobit type of limited dependent variables models. In our model, the truncation is more complicated than the Tobit type of models and hence the estimation procedure has to be modified. The first stage in our estimation is by Probit and the second stage is by ordinary least squares. The details of the estimation procedure, its theoretical properties and the identification conditions appear in Lee [1976] chapters 2 and 3. Here we will only describe briefly the procedures. A similar procedure has also been suggested in Heckman [1976].

Substituting the wage equations $\log W_{ui}$ and $\log W_{ni}$ into I_i^*, we have a typical probit model,

$$I_i^* = \gamma_0 + \gamma_1 X_i' + \gamma_2 Z_i' - \varepsilon^*$$

where X_2' includes all socioeconomic and personal characteristic variables and Z_i' all the observable attributes of the union organization and industrial variables. Thus we can estimate γ_0, γ_1 and γ_2 by probit analysis and obtain consistent estimates $\hat{\gamma}_0$, $\hat{\gamma}_1$ and $\hat{\gamma}_2$ after normalizing $\sigma^2_{\varepsilon^*}=1$.

Conditional on union status, the union wage equation is

$$\log W_{ui} = \theta_{u0} + X_{ui}\theta_{u1} + Z_{ui}\theta_{u2} + \sigma_{1\varepsilon^*}\left(-\frac{f(\Psi_i)}{F(\Psi_i)}\right) + \eta_u$$

where $E(\eta_u|I_i=1)=0$, $\Psi_i=\gamma_0+\gamma_1 X_i'+\gamma_2 Z_i'$. F is the cumulative distribution of a standard normal random variable and f is its density function. Conditional on nonunion status, the nonunion wage equation is

$$\log W_{ni} = \theta_{n0} + X_{ni}\theta_{n1} + X_{ni}\theta_{n2} + \sigma_{2\varepsilon^*}\left(\frac{f(\Psi_i)}{1 - F(\Psi_i)}\right) + \eta_n$$

where $E(\eta_n|I_i=0)=0$.[3]

The parameters (θ_{uj}) can be estimated consistently by regressing the observed union wage $\log W_{ui}$ on X_{ui}, Z_{ui} and $(-f(\hat{\Psi}_i)/F(\hat{\Psi}_i))$; where $\hat{\Psi}_i=\hat{\gamma}_0+\hat{\gamma}_1 X_i'+\hat{\gamma}_2 Z_i'$;

[2] The means $E(\varepsilon_u|I_i=1)=0$ and $E(\varepsilon_n|I_i=0)=0$ occur only in very special situations. For example, if $\varepsilon_u=\varepsilon_n$ and ε is independent of ε_u and ε_n, then the means will be zero and the ordinary least squares will give consistent estimates.

[3] The term $\sigma_{1\varepsilon^*}\left(-\frac{f(\psi_i)}{F(\psi_i)}\right)$ is, in fact, the mean $E(\varepsilon_u|I_i=1)$ and $\sigma_{2\varepsilon^*}\left(\frac{f(\psi_i)}{1-F(\psi_i)}\right)$ is the mean $E(\varepsilon_n|I_i=0)$.

(θ_{nj}) can be estimated consistently in a similar fashion.

Denote $\operatorname{cov}(\varepsilon_u, \varepsilon_n, \varepsilon^*) = \Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{1\varepsilon^*} \\ \sigma_{21} & \sigma_2^2 & \sigma_{2\varepsilon^*} \\ \sigma_{1\varepsilon} & \sigma_{2\varepsilon^*} & 1 \end{bmatrix}$

It has been shown in Lee [1976],

$$\operatorname{var}(\eta_u | I_i = 1) = \sigma_1^2 + \sigma_{1\varepsilon^*}^2(\Psi_i)\left(-\frac{f(\Psi_i)}{F(\Psi_i)}\right) - \sigma_{1\varepsilon^*}^2\left(\frac{f(\Psi_i)}{F(\Psi_i)}\right)^2$$

and

$$\operatorname{var}(\eta_n | I_i = 0) = \sigma_2^2 + \sigma_{2\varepsilon^*}^2(\Psi_i)\left(\frac{f(\Psi_i)}{1 - F(\Psi_i)}\right) - \sigma_{2\varepsilon^*}^2\left(\frac{f(\Psi_i)}{1 - F(\Psi_i)}\right)^2$$

The correct asymptotic covariance matrix of the OLS estimate of (θ_{uj}) is very complicated. In practice, we use the approximation

$$(\tilde{X}_u'\tilde{X}_u)^{-1}\tilde{X}_u'V_u\tilde{X}_u(\tilde{X}_u'\tilde{X}_u)^{-1}$$

where the i-th row of $\tilde{X}_u$ is $\left(1, X_{ui}, Z_{ui}, -\frac{f(\hat{\Psi}_i)}{F(\hat{\Psi}_i)}\right)$ and V_u is a diagonal matrix,

$$V_u = \begin{vmatrix} \operatorname{var}(\eta_{u1} | I_1 = 1) & & & & 0 \\ & \ddots & & & \\ & & \operatorname{var}(\eta_{ui} | I_i = 1) & & \\ & & & \ddots & \\ 0 & & & & \operatorname{var}(\eta_{un} | I_n = 1) \end{vmatrix}$$

Similarly, the matrix

$$(\tilde{X}_n'\tilde{X}_n)^{-1}\tilde{X}_n'V_n\tilde{X}_n(\tilde{X}_n'\tilde{X}_n)^{-1}$$

is used for the covariance matrix of (θ_{nj}) where the i-th row of $\tilde{X}_n$ is $(1, X_{ni}, Z_{ni}, f(\hat{\Psi}_i)/(1-F(\hat{\Psi}_i)))$ and

$$V_n = \begin{vmatrix} \operatorname{var}(\eta_{n1} | I_1 = 0) & & & & 0 \\ & \ddots & & & \\ & & \operatorname{var}(\eta_{ni} | I_i = 0) & & \\ & & & \ddots & \\ 0 & & & & \operatorname{var}(\eta_{nn} | I_n = 0) \end{vmatrix}$$

The covariance Σ can also be estimated consistently. The details again appear in Lee [1976].

A generalized least squares procedure may be desirable to increase the efficiency of the estimates. However, in our data, the matrix $(\tilde{X}_u'V_u^{-1}\tilde{X}_u)$ was found to be ill-conditioned due to the high correlation between the variables

$$1\Big/\left(\sigma_2^2 + \sigma_{2\varepsilon}^2(\Psi_i)\left(\frac{f(\Psi_i)}{1 - F(\Psi_i)}\right) - \sigma_{2\varepsilon}^2\left(\frac{f(\Psi_i)}{1 - F(\Psi_i)}\right)^2\right)$$

and

$$-\frac{f(\Psi_i)}{F(\Psi_i)}\Big/\mathrm{var}\,(\eta_{ui}I_i = 1)\,.$$

The same was true for

$$1/\mathrm{var}\,(\eta_{ui}\,|\,I_i = 0) \quad \text{and} \quad \frac{f(\Psi_i)}{1 - F(\Psi_i)}\Big/\mathrm{var}\,(\eta_{ni}\,|\,I_i = 1)$$

Thus generalized least squares was not used. This difficulty also tends to occur in the linear probability model (see Domencich and McFadden [1975]).

In this model industrial classification appears in the wage equations but not in the union status equation. Thus the structural parameters of union status, the δ_i's, can be estimated up to a positive proportional constant by applying the probit maximum likelihood procedure with exogenous variables X_i, Z_i and the estimated wage gain $\log \hat{W}_{ui} - \log \hat{W}_{ni}$. The consistency of this two stage probit estimation has been proved in Lee [1976]. The covariance matrix is approximated as if the estimated wage gain is the exact exogenous variable.

5. EMPIRICAL RESULTS

The total sample which is used in the estimation consists of observations on 3720 operatives, of which 1795 are not in the union and 1925 are in the union.

Maximum likelihood probit estimates of the coefficients in the reduced form union status equation are presented in Table 1. Since the likelihood function of the probit model is well behaved on the parameter space, the Newton Raphson iteration procedure converges rapidly. Estimates of the coefficients of the wage equations are then obtained by using the procedure described in Section 4 and are presented in Table 1 and Table 2. The t-values reported in Tables 1 and 2 are derived from the OLS estimates with homoscedastic residuals. The correct value should be derived from the covariance matrix $(\tilde{X}'\tilde{X})^{-1}\tilde{X}'V\tilde{X}(\tilde{X}'\tilde{X})^{-1}$ of the OLS estimates. But the estimated values of $\mathrm{var}(\eta_{ui}|I_i=1)$, $i=1,\ldots, n$ range from 0.070 to 0.085, which are approximately equal to the estimated variance of 0.078. In the nonunion wage equation, the estimated $\mathrm{var}(\eta_{ni}|I_i=0)$, $i=1,\ldots, n$ ranges from 0.100 to 0.1168 which is close to the estimated residual variance 0.1112 in that equation. Therefore the t-values derived from OLS estimates are good approximations. The regional effects are normalized on the west. Negative signs on the coefficients of N.E., N.C. appear in all the wage equations estimates throughout the study, so that observed wage rates are higher in the western region than the others in this data. All other coefficients of the regional and district variables have the expected signs. Compared with the other regions the operatives in the South receive the lowest wage rates in both union and nonunion sectors. The operatives living in large SMSA's receive higher wage rates than those living outside metropolitan areas. The absolute difference tends to be bigger for the unionized sector than the nonunionized sector. This difference is thus partly due to different supply and demand conditions and differences in the costs of living and partly due to union effects. Education levels and market experience all have the expected signs: the wage rate increases with education level and market experience. The negative and significant coefficient of ME_2 confirms the strict concavity of the earning profile. The absolute wage difference between high educated and low educated operatives tends to be much larger for the nonunionized group. Thus the rate of returns on education is larger in the nonunionized sector than the unionized sector and the high educated operatives may prefer to join the nonunionized sector. In the unionized sector, market experience has a larger effect on wage increment. This may be due to seniority rules of the labor unions. The coefficients of the race, sex and health limitations have the expected signs. Male operatives receive higher wages than female operatives. White operatives receive higher wages than nonwhite operatives. The black and white wage difference is less in the unionized sector than in the nonunionized sec-

TABLE 1[4]
THE UNION WAGE EQUATION ESTIMATES
(SELECTIVITY BIAS ADJUSTED)

Exogenous Variable	Coefficients	T-Values
Constant	4.431	27.129 **
N. E.	−0.083	−3.369 **
N. C.	−0.007	−0.240
S	−0.172	−5.422 **
UR_1	0.067	3.279 **
UR_2	−0.092	−3.667 **
ED_1	−0.108	−2.666 **
ED_2	−0.033	−1.330 Δ
ED_3	0.052	2.60 **
ED_4	0.111	5.168 **
ED_5	0.139	4.112 **
ME	0.016	7.526 **
ME_2	−0.0002	−5.418 **
RACE	0.095	6.367 **
SEX	0.317	14.915 **
IND_1	0.223	4.034 **
IND_2	0.169	3.722 **
IND_3	0.034	1.477 Δ
IND_4	0.018	0.722
U	0.662	6.168 **
HLT	−0.055	−2.105 **
Selectivity Variable	−0.168	−1.914 *

Selectivity Variable $= -f(\gamma_0+\gamma_1 X_i+\gamma_2 Z_i)/F(\gamma_0+\gamma_1 X_i+\gamma_2 Z_i)$
Residual variance $=0.078$
$R^2=0.311$
Observations $=1925$

[4] In the tables, ** indicates significance under 0.025 level of significance; *significance under 0.05 level of significance and a Δ significance under 0.1 level of significance.

TABLE 2

THE NONUNION WAGE EQUATION ESTIMATES
(SELECTIVITY BIAS ADJUSTED)

Exogenous Variable	Coefficients	T-Values
Constant	4.754	71.220 **
N. E.	−0.091	−2.975 **
N. C.	−0.074	−2.238 **
S	−0.139	−4.427 **
UR_1	0.039	1.610 Δ
UR_2	−0.067	−2.933 **
ED_1	−0.049	−1.187
ED_2	−0.016	−0.526
ED_3	0.087	3.380 **
ED_4	0.157	5.979 **
ED_5	0.282	6.747 **
ME	0.012	5.468 **
ME_2	−0.0002	−4.788 **
RACE	0.186	10.205 **
SEX	0.267	13.501 **
IND_1	0.120	1.915 *
IND_2	0.130	2.433 **
IND_3	0.053	1.474 Δ
IND_4	0.058	1.594 Δ
HLT	−0.088	−2.961 **
Selectivity Variable	0.136	3.152 **

Selectivity Variable$=f(\gamma_0+\gamma_1X_i+\gamma_2Z_i)/(1-F(\gamma_0+\gamma_1X_i+\gamma_2Z_i))$
Residual variance$=0.1112195$
$R^2=0.291$
Observations$=1795$

tor. While there is evidence that blacks may be discriminated against entering labor unions, they do better in unionized sectors once they are members. The wage difference between male and female operatives is larger in the unionized sector so female operatives may prefer the nonunionized sector. Operatives in the mining and construction industries earn more than those in other industries. Union coverage in the industry has a positive and significant effect on the wage of unionized operatives. At the bottom of Tables 1 and 2, we report the coefficient of the selectivity variable which is $-f(\Psi_i)/F(\Psi_i)$ for the union wage equation and $f(\Psi_i)/(1-F(\Psi_i))$ for the nonunion wage equation. The estimated truncation effect is positive for all observations for the union and the nonunion sectors. The positive truncation means that we only observe the upper sections of the union wage distributions given fixed socioeconomic and personal characteristics of the individual and the attributes of the industries. The positive value results from the individual's selection of the relevant sector that pays him better than the average operatives with the same characteristics and under the same working circumstances. More explicitly, the term truncation effect for the union wage equation refers to the term $\sigma_{1\varepsilon*}(-f(\Psi_i)/F(\Psi_i))$ and $\sigma_{2\varepsilon*}f(\Psi_i)/(1-F(\Psi_i))$

for the nonunion wage equation. They are $E(\varepsilon_u|I_i=1)$ and $E(\varepsilon_n|I_i=0)$ respectively. In general, the truncation effects need not be positive or negative. The signs are determined by the second moments of the disturbances ε_u, ε_n and ε. In this model,

$$\varepsilon^* = (\varepsilon - \delta_1(\varepsilon_u - \varepsilon_n))/\sigma^*$$

where σ^* is the variance of $(\varepsilon - \delta_1(\varepsilon_u - \varepsilon_n))$.

$$\sigma_{1\varepsilon^*} = E(\varepsilon_u\varepsilon^*) = \frac{1}{\sigma^*}\{E(\varepsilon_u\varepsilon) - \delta_1 E(\varepsilon_u^2) + \delta_1 E(\varepsilon_u\varepsilon_n)\}$$

and

$$\sigma_{2\varepsilon^*} = E(\varepsilon_n\varepsilon^*) = \frac{1}{\sigma^*}\{E(\varepsilon_n\varepsilon) - \delta_1 E(\varepsilon_n\varepsilon_u) + \delta_1 E(\varepsilon_n^2)\}$$

Under special assumptions, however, the signs of the truncation effects can be determined analytically. For example, if ε_u, ε_n and ε are independent,

$$\sigma_{1\varepsilon^*} = -\delta_1 E(\varepsilon_u^2)/\sigma^*$$

$$\sigma_{2\varepsilon^*} = \delta_1 E(\varepsilon_n^2)/\sigma^*$$

As δ_1 is expected to be positive, we see that $E(\varepsilon_u|I_i=1)$ and $E(\varepsilon_n|I_i=0)$ are both positive.

But in our model, the independence assumption is not likely to be valid because of unmeasurable common factors in the equations. The estimated positive truncation effect may be due to the fact that the variance of $E(\varepsilon_u^2)$ and $E(\varepsilon_n^2)$ dominate the other covariances.

The estimated wage equation allows comparison of wage differences between the two sectors. Wages are predicted using the estimates, i.e.,

$$\hat{\mu}_{ui} = \log W_{ui} = \hat{\theta}_{u0} + X_{ui}\hat{\theta}_{u1} + Z_{ui}\hat{\theta}_{u2}$$

$$\mu_{ni} = \log W_{ni} = \hat{\theta}_{n0} + X_{ni}\hat{\theta}_{n1} + Z_{ni}\hat{\theta}_{n2}.$$

Define

$$\frac{100}{\#(\delta)} \sum_{i \in \delta} (e^{\hat{\mu}_{ui}} - e^{\hat{\mu}_{ni}})/e^{\hat{\mu}_{ni}}$$

to measure the average percentage increment of the wage rate for the union sector compared with the nonunion sector for subgroup δ in the whole sample we used. $\#(\delta)$ is the number of workers in subgroup δ.[5]

[5] A more accurate measure can be defined as follows.

$$\mu_{ui} = \log W_{ui} = \theta_{n0} + X_{ui}\theta_{u1} + Z_{ui}\theta_{u2} + \varepsilon_u.$$

$$\mu_{ni} = \log W_{ni} = \theta_{n0} + X_{ni}\theta_{n1} + Z_{ni}\theta_{n2} + \varepsilon_n.$$

(*Continued on next page*)

The estimated average percentage increment for the whole sample is 15.68. More detailed estimates of average increments for different groups of operatives are reported in Tables 3 and 4. The differential is larger for male operatives than the female operatives. The highest differential is for the nonwhite males. A more disaggregated analysis appears in Table 4. Except for the older females, the nonunion sector pays females more than the union sectors. The table also shows obvious effects of seniority for union members. The wage differential appears higher for those with elementary or high school education but it is less for those with some college training.

Table 5 presents the percentage union-nonunion wage differential for various

TABLE 3

RACE-SEX (AVERAGE PERCENTAGE WAGE DIFFERENTIALS).

	White	Nonwhite
Male	16.239	28.451
Female	2.782	12.690

TABLE 4

MARKET EXPERIENCE-EDUCATION (AVERAGE PERCENTAGE WAGE DIFFERENTIALS)

	Market Experience	Highest Grade Completed			
		0–7	8	9–12	13 or more
	0 — 10	1.314	6.600	8.561	−2.225
White-Male	10+ — 25	12.059	15.328	13.906	3.732
	25 or more	20.032	30.463	24.613	9.807
	0 — 10	−12.646	−4.017	−5.656	−25.842
White-Female	10+ — 25	−7.054	−0.891	0.335	−14.685
	25 or more	6.796	14.070	9.532	−0.159
	0 — 10	1.674	11.498	16.959	12.434
Nonwhite-Male	10+ — 25	18.836	24.189	27.861	10.315
	25 or more	37.086	44.362	35.618	28.821
	0 — 10	13.321	7.916	4.115	−4.030
Nonwhite-Female	10+ — 25	12.617	7.291	10.158	3.827
	25 or more	20.350	25.812	24.030	14.396

(*Continued*)

W_{ui}, W_{ni} are lognormally distributed and

$$E(W_{ui}|X_{ui}, Z_{ui})=\exp\{\theta_{u0}+X_{ui}\theta_{u1}+Z_{ui}\theta_{u2}+1/2\sigma_1^2\}$$

$$E(W_{ni}|X_{ni}, Z_{ni})=\exp\{\theta_{n0}+X_{ni}\theta_{u1}+Z_{ni}\theta_{n2}+1/2\sigma_2^2\}$$

thus the more accurate measure is

$$\frac{100}{\#(\zeta)}\sum_{i\in\delta}\left(\frac{\exp\{\hat\theta_{u0}+X_{ui}\hat\theta_{u1}+Z_{ui}\hat\theta_{u2}\}\exp\{1/2(\hat\sigma_1^2-\hat\sigma_2^2)\}-\exp\{\hat\theta_{n0}+X_{ni}\hat\theta_{n1}+Z_{ni}\hat\theta_{n2}\}}{\exp\{\hat\theta_{n0}+X_{ni}\hat\theta_{n1}+Z_{ni}\hat\theta_{n2}\}}\right)$$

But in our estimates, we have $\hat\sigma_1^2=0.092767$ and $\hat\sigma_2^2=0.119948$ and $\exp\{1/2(\hat\sigma_1^2-\hat\sigma_2^2)\}\approx 1+1/2(\hat\sigma_1^2-\hat\sigma_2^2)=1-0.01359=0.986$. Thus the measure proposed above is almost the same.

TABLE 5

UNION POWER–INDUSTRIES (AVERAGE PERCENTAGE WAGE DIFFERENTIALS)

Industry Union Coverage	Industries				
	Mining[6]	Construction	Manufacturing Durable Goods	Manufacturing Non-Durables	Transportations, Communication, Utilities and Sanitary Services
0.0 — 0.5	−9.257		0.321	−12.143	−0.561
0.5+ — 0.8		27.758	18.100	11.016	27.897
0.8+ — 1.0	43.369		37.626	28.037	36.163

TABLE 6

THE UNION STATUS EQUATION ESTIMATES (THE STRUCTURAL FORM ESTIMATES)

	Max. Likelihood Est. Coefficient	Standard Error
Constant	−0.654	0.145 **
N. E.	0.227	0.076 **
N. C.	0.197	0.077 **
S	−0.296	0.077 **
UR_1	0.129	0.063 **
UR_2	−0.174	0.067 **
ED_1	−0.269	0.116 **
ED_2	−0.098	0.084
ED_3	0.079	0.072
ED_4	0.119	0.074 Δ
ED_5	0.258	0.119 **
ME	0.0020	0.0022
RACE	0.166	0.054 **
SEX	0.093	0.055 *
WK_1	−0.372	0.115 **
WK_2	−0.017	0.073
CCR	0.365	0.132 **
HLT	−0.185	0.087 **
$\ln \hat{W}_u - \ln \hat{W}_n$	2.455	0.205 **

industries. In all cases, the union-nonunion wage differential markedly increases with the extent of organization, similar to the results of Rosen [1969] who used more aggregated data. Unionism has the strongest effects on wage increments for operatives in construction and mining industries, a result which is compatible with other studies.

The most interesting advantage of using this approach is that we can estimate the structural form of the union status equation. These estimates are reported in Table 6. The most powerful factor determining unionization status is the

[6] In the Mining Industry, there are four subindustries. The industry union coverage is 0.94 for Metal Mining, 0.81 for Coal Mining, 0.13 for Crude Pet. & Natural Gas and 0.30 for Non-Metalic Mining industry.

union-nonunion wage differential. The estimates for region and city size variables are all significant. They reveal that Southern workers have different attitudes toward unions than their counterparts in other regions. The operatives in the northern regions have the strongest receptivity towards unions. Also, operatives living within large SMSA's have a greater propensity to join unions than those outside SMSA's, probably reflecting both tastes and opportunities. The estimated effect of industrial concentration shows that it is easier to unionize concentrated industries. The education, market experience, race and sex variables catch the tastes of individuals towards unionism and firm productivity selection effects. For example, more highly educated operatives may not like labor unions, but, when the wage differential has been controlled, the firm's selectivity effect on education dominates any possible taste factors. The estimates show that union firms select more educated workers, presumably on productivity grounds. The same factors appear to operate with respect to whites compared with nonwhites. Finally, the estimates show that more casual workers, consisting mostly of females, young operatives, workers with less weeks worked and those with health problems are less likely to be trade union members. These may reflect the high costs to unionize those workers.

In Table 7, the estimates give the net effects for the various factors on union

TABLE 7

THE REDUCED FORM ESTIMATES OF THE UNION STATUS EQUATION

Exogeneous Var.	Coefficients	Standard Error
Constant	−1.633	0.202 **
N. E.	0.242	0.076 **
N. C.	0.364	0.077 **
S	−0.398	0.077 **
UR_1	0.204	0.063 **
UR_2	−0.240	0.067 **
ED_1	−0.396	0.117 **
ED_2	−0.147	0.084 *
ED_3	−0.007	0.072
ED_4	0.0024	0.074
ED_5	−0.095	0.115
ME	0.013	0.0065 **
ME^2	−0.00004	0.00013
RACE	−0.061	0.052
SEX	0.251	0.054 **
WK_1	−0.378	0.115 **
WK_2	−0.029	0.073
IND_1	0.516	0.184 **
IND_2	−0.115	0.151
IND_3	−0.098	0.088
IND_4	0.030	0.095
U	1.793	0.145 **
CCR	0.433	0.142 **
HLT	−0.101	0.088

status. The effects are results of cost and benefit considerations, firms, selectiveness and personal tastes towards unions. A comparison of the estimates in Tables 6 and 7 shows some differences. The union firms tend to select more highly educated workers but the higher rate of return on education in the nonunion sector will induce higher educated workers in to the nonunion sector. The net effect is that the most educated operatives will be in the nonunion sector. While the union firms select more experienced workers and market experience has higher returns in the union sector, experienced workers would tend to be in the unionized sector. The nonwhites might be discriminated against by union organizations, but the benefits received in the union sector would attract the nonwhites. While male operatives are selected by union firms and the wage increments are larger than females, male operatives are more likely to be in unions.

OLS estimates using the subsamples, union workers and nonunion workers, without adjusting the truncation effect are reported in Tables 8, 9, 10 and 11 for purposes of comparison with the preferred, truncation-adjusted estimates. The overall estimate of the percentage union-nonunion wage differentials for the whole sample is 17.458% which is 2% higher than the "correct" estimate. Com-

TABLE 8

THE UNION WAGE EQUATION ESTIMATES
(Selectivity Bias Unadjusted)

Exogeneous Variables	Coefficients	T-Values
Constant	4.724	83.249 **
N. E.	−0.106	−4.978 **
N. C.	−0.041	−1.958 *
S	−0.130	−5.664 **
UR_1	0.049	2.703 **
UR_2	−0.066	−3.130 **
ED_1	−0.070	−1.988 **
ED_2	−0.020	−0.820
ED_3	0.053	2.634 **
ED_4	0.109	5.112 **
ED_5	0.147	4.341 **
ME	0.014	7.392 **
ME_2	−0.0002	−5.209 **
RACE	0.1004	6.808 **
SEX	0.290	18.126 **
IND_1	0.182	3.576 **
IND_2	0.191	4.351 **
IND_3	0.044	1.984 **
IND_4	0.022	0.903
U	0.475	10.835 **
HLT	−0.046	−1.802 *

Resid. Var. = 0.078
R^2 = 0.310
Observations = 1925

TABLE 9

THE NONUNION WAGE EQUATION ESTIMATES
(Selectivity Bias Unadjusted)

	Coefficients	T-Values
Constant	4.841	79.456 **
N. E.	−0.076	−2.513 **
N. C.	−0.040	−1.284 Δ
S	−0.174	−5.946 **
UR_1	0.057	2.427 **
UR_2	−0.088	−3.992 **
ED_1	−0.080	−2.011 **
ED_2	−0.030	−1.010
ED_3	0.088	3.399 **
ED_4	0.162	6.143 **
ED_5	0.277	6.616 **
ME	0.014	6.256 **
ME^2	−0.0002	−4.974 **
RACE	0.183	10.037 **
SEX	0.295	16.662 **
IND_1	0.102	1.631 Δ
IND_2	0.110	2.057 **
IND_3	0.035	0.983
IND_4	0.034	0.940
HLT	−0.099	−3.354 **

Resid. Var. = 0.1116
R^2 = 0.287
Observations = 1795

TABLE 10

RACE-SEX (Average Percentage Wage Differentials)

	White	Nonwhite
Male	15.809	26.854
Female	10.728	18.572

paring Tables 10 and 11 with Tables 3 and 4, the two most obvious differences are the estimates of the percentage wage differentials for female operatives and young operatives. The unadjusted estimates in Table 10 overestimate the percentage differentials for female operatives. It is about 8 percent more for white females and 6 percent for black females. This occurs because the wage difference of sex is underestimated in the union wage equation and it is overestimated in the nonunion wage equation. The percentage wage differentials are also overestimated for the young operatives. On the other hand, the differentials for the most experienced male operatives are underestimated. Finally, in Table 12, the estimates underestimate the effect of unionism in the highest union coverage category and overestimate the effect in the lowest union coverage category. Also OLS gives 6 percentage higher for the construction industries.

TABLE 11

MARKET EXPERIENCE–EDUCATION (Average Percentage Wage Differentials)

	Market Experience	Highest Grade Completed			
		0–7	8	9–12	13 or more
White-Male	0 — 10	15.498	13.922	11.795	1.823
	10+ — 25	18.802	16.176	13.161	3.434
	25 or more	23.923	22.960	17.949	5.723
White-Female	0 — 10	4.127	8.751	7.923	−12.541
	10+ — 25	7.452	10.308	8.549	−5.887
	25 or more	15.963	16.330	11.827	2.941
Nonwhite-Male	0 — 10	24.935	20.859	19.288	13.978
	10+ — 25	29.692	26.635	22.800	9.401
	25 or more	37.926	34.621	27.538	17.893
Nonwhite-Female	0 — 10	39.505	22.777	14.724	8.856
	10+ — 25	28.712	17.238	14.749	3.122
	25 or more	27.430	26.675	22.073	14.418

TABLE 12

UNION POWER–INDUSTRIES (Average Percentage Wage Differentials)

	Industries				
Industry Union-Coverage	Mining	Construction	Manufacturing Durable Goods	Manufacturing Nondurable Goods	Transportations, Communication, Utilities and Sanitary Services
0.0 — 0.5	4.98		9.480	0.713	0.621
0.5+ — 0.8		34.193	17.369	15.996	21.710
0.8+ — 1.0	40.390		29.302	25.917	28.289

6. CONCLUSIONS

In this paper we investigate the interactions between unions and wage rates. Data on semi-skilled laborers in several industries are analyzed in a simultaneous equations model with limited dependent and qualitative endogenous variables. A two stage method of estimation is employed.

In this study we find results in two directions. Unionism does have a significant effect in raising wage rates. On the other hand, union membership is determined mainly by wage gains and the selectiveness of employers' employment policy. Unionized firms tend to select highly productive workers. This finding explains why unionized and nonunionized firms coexist in the same product market. Their unit labor costs do not differ by enough to allow the nonunionized firms to drive the unionized firms out.

University of Minnesota

REFERENCES

AMEMIYA, TAKESHI, "Multivariate Regression and Simultaneous Equation Models When the Dependent Variables are Truncated Normal," *Econometrica,* 42, (November, 1974), 999–1012.

ASHENFELTER, O., AND G. E. JOHNSON, "Unionism, Relative Wages and Labor Quality in U. S. Manufacturing Industires," *International Economic Review,* 13, (October, 1972), 488–508.

BECKER, G. S., "Union Restrictions on Entry," in, P. D. Bradley, ed., *The Public Stake in Union Power*, (Charlottesville, Va: University of Virginia Press, 1959).

———, *Human Capital,* (New York: Columbia University Press, 1964).

DOMENCICH, T. A., AND DANIEL MCFADDEN, *Urban Travel Demand,* (Amsterdam: North Holland Publishing Co., 1975).

HECKMAN, J., "The Common Structure of Statistical Models of Truncation, Sample Selection, and Limited Dependent Variables and a Simple Estimator for Such Models," *Annals of Economic and Social Measurement,* 5, (1976), 475–492.

LEE, L. F. "Estimation of Limited Dependent Variables by Two Stage Method," unpublished Ph.D. Thesis, Department of Economics, University of Rochester, (1976), Chapters 2 and 3.

LEWIS, H. G., "Competitive and Monopoly Unionism," in, P.D. Bradley, ed., *The Public Stake in Union Power,* (Charlottesville, Va: University of Virginia Press, 1959).

ROSEN, S., "Unionism and the Occupational Wage Sturcture in the United States," *International Economic Review,* 11, (June, 1970), 269–285.

———, "Trade Union Power, Threat Effects and Extent of Organization," *Review of Economic Studies,* 36, (April, 1969), 185–191.

SCHMIDT, P., AND R. STRAUSS, "The Effect of Unions on Earnings and Earnings on Unions: A Mixed Logit Approach," *International Economic Review,* 17, (February, 1976), 204–212.

WEISS, L. W., "Concentration and Labor Earnings," *American Economic Review,* 56, (March, 1966), 96–117.