THE INTERINDUSTRY STRUCTURE OF UNIONISM, EARNINGS, AND EARNINGS DISPERSION

BARRY T. HIRSCH*

This study examines the relationship between unionism and earnings dispersion within U.S. manufacturing and nonmanufacturing industries. The author hypothesizes not only that unionism narrows earnings dispersion, as others have shown, but also that the dispersion in earnings, reflecting the degree of worker homogeneity, influences the level of unionism. Estimation of a three-equation model, using 1970 three-digit industry data, provides evidence regarding the simultaneous determination of unionism, earnings, and earnings dispersion within U.S. industries. The estimated equalizing effects of unionism on within-industry earnings distributions are found to be significant both in the manufacturing and nonmanufacturing sectors, the size of these estimates increasing after accounting for simultaneity. In addition, the dispersion in earnings does appear to affect the level of unionism, although the evidence on this point is ambiguous.

Most empirical studies of unionism have focused on estimates of union/nonunion wage differentials. For some time, it has been realized that unionism and earnings are determined simultaneously and that union/nonunion wage differentials estimated by single-equation techniques are likely to be biased. Thus, simultaneous equation techniques are now typically, although not universally, used to estimate union wage effects. Recently, increased attention has been given to the effects of unions on wage dispersion, with most of the evidence indicating that unions significantly narrow wage differentials. There

---

*The author is an associate professor at the University of North Carolina at Greensboro. Research support was provided by UNC-Greensboro. Lee Edlefsen, Solomon Polacheck, James Ragan, and Michael Sattinger provided helpful comments. Terry Seaks provided assistance with the cubic spline program and helpful discussion.


are several reasons, however, for also believing that earnings dispersion will affect the likelihood of unionism. Yet no attempt appears to have been made to integrate an earnings-dispersion equation into a union-wage model in order to explore the possible simultaneous determination of unionism and earnings inequality.

Here we present a model in which unionism, earnings, and earnings dispersion are simultaneously determined. The model is estimated using a cross-section of three-digit Census of Population industries for 1970 in both manufacturing and nonmanufacturing. We examine not only the effect of unionism on earnings dispersion, but also the effect of earnings dispersion on the level of unionism.

### Unionism and Earnings Dispersion

Perhaps the most complete treatment of the relationship between unionism and wage dispersion is provided in the recent paper by Richard Freeman. As he points out, institutional labor economists noted long ago the well-established trade union goal of rate standardization. Union policies have attempted to equalize wage rates across establishments for similar jobs and within establishments by lowering white-collar/blue-collar differentials and by establishing rates of pay based more on job category than on merit or personal characteristics.

Why unions have adopted wage standardization as a goal is an interesting question. Freeman cites some relevant literature by institutional economists. In addition, a median voter model applied to unionism may help explain support by a majority of workers for wage leveling, as long as mean earnings are greater than the median. Supporting this view is the evidence from Farber and Saks regarding individual voting in NLRB representation elections that suggests that support for unionism is inversely related to one's position in the intrafirm earnings distribution. Worker demand for rate standardization may also be an attempt to lower at least some forms of risk. Unionism provides a collective voice through which workers can bargain for measurable, objective standards and pay based on well-defined job categories. Overall, many workers may desire less intrafirm wage dispersion across and within job classifications.

Utilizing detailed microdata, Freeman finds that unions significantly decrease the dispersion of wages in more highly unionized establishments in both the manufacturing and nonmanufacturing sectors. The dispersion in earnings is reduced in part because wage-determining characteristics have smaller effects in the union sector and because less dispersion exists within given cells of workers with similar characteristics. Freeman concludes that these within-sector effects of unionism that narrow differentials outweigh union wage effects that increase across-industry wage dispersion so that, on net, unionism reduces wage inequality.

A recent paper by Quan has also attempted to estimate union effects on the size

---


5Estimates of flatter, but higher, earnings profiles for union members have been found in Farrell E. Bloch and Mark S. Kuskin, "Wage Determination in the Union and Nonunion Sectors," *Industrial and Labor Relations Review*, Vol. 31, No. 2 (January 1978), pp. 183–92, and in several more recent studies. A paper by Solomon W. Polachek ("Unionization of the White Collar Worker," unpublished paper, University of North Carolina, April 1981) uses longitudinal data and finds the change into union status associated with faster wage growth. This result appears inconsistent with the cross-sectional findings of union wage leveling. However, the change in union status is treated by Polachek as exogenous, even though selectivity bias in this type of analysis is likely to be large. Also see Wesley Mellow, "Unionism and Wages: A Longitudinal Analysis," *Review of Economics and Statistics*, Vol. 63, No. 1 (February 1981), pp. 43–52.
distribution of earnings. After estimating the incidence of unionism by earnings class and assuming equal percentage union wage effects across classes, he makes a comparison of the Gini coefficients of union and non-union earnings. With one of his two micro-data sets he finds greater equality among union earnings, though the magnitude of the difference is small.

While both these papers support the belief that unionism acts to decrease earnings dispersion, neither addresses directly the possibility that earnings dispersion and unionism are simultaneously determined. Freeman does examine the possibility that the lower dispersion in earnings in the union sector is due to less variability in worker characteristics among unionized establishments, rather than to direct union effects on the wage structure. He concludes, however, that differences in worker characteristics explain only a "moderate" amount of the difference and that union effects on the wage structure are primarily responsible for the narrower dispersion among union wages. While Freeman does an excellent job in adjusting for worker and firm characteristics, he does not test directly for simultaneity between unionism and earnings dispersion.

While this recent literature has stressed the wage-narrowing effects of unionism, there are also reasons for believing that the earnings structure may directly affect the degree of unionism. The determinants of unionization have previously been examined within an economic framework in which equilibrium levels are determined by the interaction of demand and supply. This framework has been used to analyze changes in unionism over time and differences across industries, states, and SMSAs. Union membership is regarded as an asset providing a flow of pecuniary and nonpecuniary services to utility-maximizing workers. After specifying union demand and supply functions, and assuming observed unionism is at an equilibrium level, a reduced-form union equation can be estimated. Empirical specification of such a model requires identification and measurement of those variables expected to influence the demand for or supply of unionism.

We believe there are good reasons for expecting equilibrium levels of unionization also to be affected by the dispersion in earnings. Note that unionism has important public good characteristics and that the equilibrium (though generally not efficient) level of a public good is determined by the preferences of the median voter. Most of the services unions provide (or, for unorganized workers, propose to provide) are nonrival, and exclusion from these services is difficult. For instance, all workers within a bargaining unit are covered by a single contract; thus, workers in similar circumstances receive roughly similar treatment, and generally none can be excluded from contract provisions. Other types of union services, such as assistance in handling grievances, may of course have less of these public-good aspects. Union coverage can be treated then as offering to workers a single package of expected services that—for workers with preferences near those of the median voter—are likely to be close to their preferred package, while expected benefits are likely to be small (or negative) for workers with preferences far from the median.

Assume that the less the earnings dispersion within an industry, the more homogeneous are workers in their characteristics and preferences. If this is correct, then it is likely that, other things equal, unionism will be greater in industries with less inequality in earnings. For any package of services a union can offer close to the preferences of the median voter, support for the union will be greater the more highly concentrated are preferences around the median. Stated differently, for a given level of demand for union services, a union can

---

6Quan, "Union Effects on the Size Distribution of Earnings."


8Existence of an equilibrium assumes preferences are single peaked. Because unions are likely to represent a package of issues, a cyclical majority (nonexistence) is a distinct possibility.
more easily construct a package that will satisfy a majority the less dispersed are preferences around the median. Thus unionization should be more likely, more effective, and more stable politically the more homogeneous are a group of workers.

In addition, the benefits and costs of unionism may be related directly to the dispersion in worker preferences and characteristics. Similar workers are easily substitutable and are likely to be in a relatively weak individual bargaining position (they face a more elastic demand); thus, they may value collective bargaining more highly. We should also expect lower costs the more homogeneous are workers due to economies of scale in both the provision of union services (grievance procedures, institutionalized work rules, and so forth) and in negotiations with management. Management may also find economies of scale in bargaining with a union relative to the commonality in “voice” among its employees.

While all the above arguments lead to an expected negative relationship between unionism and earnings dispersion, other factors may lead to a positive relationship. As discussed previously, workers may demand unions as a means to reduce some elements of risk or to decrease dispersion in compensation, both across and within job categories, leading to a positive effect of earnings dispersion on the extent of unionism. Because unions tend to represent the average (median) worker while the “market” will more closely represent the marginal worker, differences in preferences between average and marginal workers (as between existing and new workers, for example) may affect the likelihood of unionism. If relatively larger differences were to exist between average and marginal workers where earnings dispersion is greater, demand for union-induced wage equalization would be greater, leading to a positive unionism-inequality relationship (and vice versa). Where differences between average and marginal workers are small, worker demand for altering the market-determined compensation structure should be less. If all workers were identical, for instance, union wage standardization would bring about little if any change from the market-determined wage dispersion.

In addition, Farber and Saks and others have found that individuals are significantly more likely to vote for union representation if they believe both that employees have been treated unfairly and that unions ensure greater fairness. If these sentiments are more likely the greater the inequality in earnings, then a positive unionism-inequality relationship is again possible. Thus, if worker demand for union-induced wage equalization is large (and union election campaign literature suggests it often is), this demand may offset the negative unionism—earnings-dispersion relationship outlined above.

For these reasons, the relationship between unionism and the dispersion in earnings is a worthwhile empirical question. Of interest is not only whether and how earnings dispersion affects unionization, but also whether, after accounting for their simultaneous determination, the estimated equalizing effects of unions on earnings dispersion are less or greater than single equation estimates would suggest. After outlining a simultaneous model in which earnings, earnings dispersion, and unionism are endogenous, these relationships will be examined.

---

9Farber and Saks, “Why Workers Want Unions,” p. 364. Greg J. Duncan and Frank P. Stafford, “Do Union Members Receive Compensating Wage Differentials,” *American Economic Review*, Vol. 70, No. 3 (June 1980), pp. 355–71, have recently stressed the role of worker complementarity in production, a condition that gives rise to unions as a mechanism to deal with the attendant public-good problems in such a workplace. Their reasoning suggests a negative correlation between unionism and the dispersion in productivity and earnings. Finally, Polachek, “Unionization of the White Collar Worker,” has suggested another reason for believing there may be a systematic relationship between earnings dispersion (or the steepness of the earnings profile) and the extent of unionism. He argues that worker-firm bargaining associated with the implicit sharing of costs and benefits from specific on-the-job training acts as a substitute for unionism. He presents evidence showing that workers with more specific on-the-job training, as measured by the slope of the earnings profile, are less likely to be unionized and concludes that this helps explain why white-collar workers are less likely to be union members. For a number of reasons, while interesting, this hypothesis is not fully convincing.
Unionism, Earnings, and Earnings Dispersion: Specification

Earnings equation. Most empirical work on the wage effects of unions has focused on the proportionate wage differential between union and nonunion workers. Starting with the identity that the log of the geometric mean wage is a weighted average of union and nonunion wages, and postulating a vector C of j observable industry and worker characteristics that determine wages in the absence of unionism, we obtain:

\[ \ln W_i = f(C_{ij}) + U_i \ln(1 + m) + e_i. \]

Here, \( \ln W_i \) is the log of the geometric mean wage in industry \( i \), \( U_i \) is the proportion unionized in industry \( i \), and \( m \) is the proportionate wage differential between union and nonunion members. Estimates of \( m \) are obtained from a regression coefficient on \( U_i \) by \( m = \exp(\alpha) - 1 \). As is well known, OLS estimation is not strictly appropriate unless \( \text{Cov}(U_i, e_i) = \text{Cov}(C_{ij}, e_i) = 0 \), and \( U_i \) and variables in \( C_{ij} \) are exogenous. Here unionism is treated as endogenous to the model (other estimation problems are addressed later).

The earnings equation builds upon the now familiar human-capital semi-log earnings function. Beginning with the individual earnings function and allowing for an effect of union membership on (weekly) wages:

\[ \ln E = a + rS + r't + r''t^2 + \gamma (\ln WW) + U \ln(1 + m), \]

where \( \ln E \) = log of annual earnings, \( r \) = the average rate of return to schooling, \( S \) = years of schooling, \( t \) = years of experience, normally proxied by \( Age - S - 5 \), \( r' \) and \( r'' \) = parameters mapping the log earnings-experience profile, \( \ln WW = \) natural log of weeks worked, \( \gamma = \) elasticity of annual earnings with respect to weeks worked (assumed a constant), \( U = \) union membership, and \( m = \) percentage union/nonunion (weekly) wage differential (assumed constant). Treating \( r, r', \) and \( r'' \) as random parameters not constant across individuals, and keeping in mind that for any two variables, \( x \) and \( y \), with a nonzero correlation \( \rho \):

\[ (\bar{x}y) = (\bar{x}) (\bar{y}) + \text{Cov}(x, y) = (\bar{x}) (\bar{y}) + \rho (x, y) \sigma (x) \sigma (y), \]

the mean value of the earnings equation becomes:

\[ \left( \ln E \right)_i = a + rS + [\rho (r, S) \sigma (r)] \sigma (S) + r't + [\rho (r', t) \sigma (r')] \sigma (t) + r''t^2 + [\rho (r'', t^2) \sigma (r'')] \sigma (t^2) + \gamma (\ln WW) + \ln(1 + m)U. \]

For estimation \( \hat{\theta} = \hat{\theta}(2) + \sigma^2 (t), \) while the \( \sigma (t^2) \) term is deleted.

In addition to variables deriving directly from the human-capital model, the earnings function includes control variables such as region and, where available, measures of industry concentration and average firm size. We will forgo discussion of such variables at this point.

By way of summary, the industry earnings equation takes the following form:

\[ \left( \ln E \right)_i = a_0 + a_1S_i + a_2 \sigma (S)_i + a_3 \bar{t}_i + a_4(t^2)_i + a_5 \sigma (t)_i + a_6(\ln WW)_i + a_7U_i + a_8CON_i + a_9\text{SIZE}_i + a_1\text{REGION}_i + u_i, \]

where unionism is endogenous, \( (\ln E)_i = \) the log of the geometric mean of wages in industry \( i \), \( CON_i = \) the fraction of industry shipments accounted for by the four leading firms in 1972, \( \text{SIZE}_i = \) the average firm size in industry \( i \), \( \text{REGION}_{ij} = j \) variables measuring the percentage of the employed labor force who reside in the North Central, Northeast, West, and South (the latter will be the reference group), \( u_i = \) random error term with zero mean and constant variance, and other variables are as defined previously (a data appendix provides the definition and source for the variables). Based on the framework outlined above, we unambiguously predict \( a_1, a_5, a_6, \) and \( a_7 > 0, \) and \( a_4 \).
< 0, while the signs of $a_2$ and $a_5$ will depend on the correlation between the parameters and variables of the earnings function. Further discussion will follow the presentation of results.

The earnings function presented above differs from those estimated in other earnings-unionism studies primarily by the use of the log of the geometric mean of earnings as implied by theory, and in the more explicit derivation and treatment of the form indicated by the human-capital model (in particular, the inclusion of $\sigma(S)$ and $\sigma(t)$ in the model).

Earnings variation. Previous studies examining union wage effects have not treated the determination of earnings variation simultaneously with the determination of the union-earnings relationship. Extending upon the human-capital earnings function presented earlier, an earnings-dispersion model can be easily derived. We begin by deleting the $t^2$ term from the equation to avoid cumbersome calculations and temporarily delete the unionization variable $U_i$. Thus, we start with the individual earnings function:

$$\ln E = b + rS + r^*t + \gamma (\ln WW),$$

where $\ln E$ is log of annual earnings, $r^*$ is the average slope of the earnings profile, $WW$ is weeks worked per year and $\gamma$, the elasticity of annual earnings with respect to weeks worked, is assumed constant across individuals. Taking the variance of both sides of the equation, noting that $\sigma^2(xy) = \sigma^2(x)\sigma^2(y) + \rho^2\sigma^2(x) + \sigma^2(x)\sigma^2(y)$, we obtain:

$$\sigma^2(\ln E) = \sigma^2(S) + \gamma^2\sigma^2(WW) + \sigma^2(r)\sigma^2(t) + \sigma^2(r^*)t^2,$$

where $\sigma^2(\ln E)$, the variance of the logarithm of earnings, is a mean invariant measure of relative earnings dispersion. Note that we have assumed the covariances among the explanatory variables to be zero, due to our inability to measure these easily with grouped industry data.

Integration of unionization into the earnings-dispersion model presents some difficulties. Taking the variance of $[\ln(l + m)]$, where $m$ is assumed constant, $[\ln(l + m)]^2\sigma^2(U)$ is obtained, indicating that increased variance across firms in the degree of unionism would increase wage dispersion within an industry. Unfortunately, we are unable directly to measure interfirm variations in unionism within an industry. In addition, we have assumed that the covariances between unionism and earnings dispersion in the absence of information on $\sigma^2(U)$, $Cov(S, U)$, $Cov(t, U)$, $Cov(\ln WW, U)$, and so forth. Based on the literature summarized earlier, however, we expect to find a negative effect of unionism on earnings dispersion.

To summarize, if we rearrange the terms in the inequality model, include an endogenous unionization variable, and add control variables for the percentage black (BLACK) and region (REGION), we obtain:

$$\sigma^2(\ln E_i) = b_0 + b_1 \bar{S}_i + b_2 \sigma^2(S)_i + b_3 \sigma^2(t)_i + b_4 \sigma^2(\ln WW)_i + b_5 \sigma^2(U)_i + b_6U_i + b_7 BLACK_i + b_8 REGION_{ij} + v_i,$$

where $v_i$ is a random error term with zero mean and constant variance. Because we are unable to hold constant the covariances between the variables, predicted signs are ambiguous. However, we expect $b_1$, $b_2$, $b_3$, $b_4$, and $b_5 > 0$ and $b_6 < 0$. Finally, $b_7 > 0$ would be consistent with findings elsewhere.

Unionism equation. It has long been recognized that industries and workers that become highly unionized are likely to have different characteristics than those with little unionization. The use of OLS to measure the effect of unionism on wages is therefore likely to lead to biased estimates and thus many studies now use a simultaneous-equations approach to estimate the determination of wages and unionism. As mentioned earlier, recent studies have also examined the effect of unions on wage dis-

\[12\text{We do know that } \sigma^2(1 + m)] U, \text{ where } m \text{ is assumed constant, } [\ln(1 + m)]^2 \sigma^2(U) \text{ is obtained, indicating that increased variance across firms in the degree of unionism would increase wage dispersion within an industry. Unfortunately, we are unable directly to measure interfirm variations in unionism within an industry. In addition, we have assumed that the covariances between unionism and earnings dispersion in the absence of information on } \sigma^2(U), \text{ Cov}(S, U), \text{ Cov}(t, U), \text{ Cov}(\ln WW, U), \text{ and so forth. Based on the literature summarized earlier, however, we expect to find a negative effect of unionism on earnings dispersion. To summarize, if we rearrange the terms in the inequality model, include an endogenous unionization variable, and add control variables for the percentage black (BLACK) and region (REGION), we obtain:}\]
persion and have found that unions narrow wage dispersion. However, only a few studies have suggested, and none have treated explicitly, the possibility that earnings dispersion affects unionization even though such effects are probable. In this section, a model of union determination is outlined in which earnings dispersion is treated as a determinant of unionism.

Adopting the demand-supply framework now common in the literature, we also include in the reduced-form unionism equation those variables expected to affect the benefits or costs of unionization. The level of earnings in an industry is expected to have a positive income effect on the demand and equilibrium level of unionism as long as union services are a normal good. Previous evidence from manufacturing industries, individuals, states, and SMSAs suggests that this relationship is statistically significant. The potential wage gain from unionization, typically measured by the proportional difference between union and nonunion wages, is not included here.13

Blue-collar (or production) workers are more likely to be unionized than other workers, due partly to differences in organizing costs, tastes, and expected wage gains from unionism. Blue-collar workers are likely to have less identification with management, lower probability of self-employment, and greater homogeneity. Each of these factors should lead to greater demands for a union "voice." Duncan and Stafford have recently shown that unionism is more likely in jobs characterized by less flexible working conditions, where employees work with machines, and where there is greater interde-

13Such a measure cannot be observed directly with grouped data. Also, as pointed out by a referee, the potential union gain is a potential loss for the company; so greater firm resistance may be likely. Thus, the effect on the likelihood of unionization is indeterminate. For a treatment of the effects of the potential wage gain on union membership, see Lung-Fei Lee, "Unionism and Wage Rates: A Simultaneous Equations Model with Qualitative and Limited Dependent Variables," International Economic Review, Vol. 19, No. 2 (June 1978), pp. 415–43, and Gregory M. Duncan and Duane E. Leigh, "Wage Determination in the Union and Nonunion Sectors: A Sample Selectivity Approach," Industrial and Labor Relations Review, Vol. 34, No. 1 (October 1980), pp. 24–34.

14Blue-collar jobs are more likely to be characterized by these conditions and, thus, more likely to be unionized. The percentage of female workers in an industry is likely to be negatively related to the percent unionized. Industries with a high percent of females are more likely to contain jobs with less permanent attachment, making the benefits from unionism less and the costs of organizing greater. Unionism is also less likely in those industries with the most flexible working conditions, those same industries likely to have a higher percentage of female workers. Moreover, Freeman and Medoff have found that women covered by a collective bargaining agreement are less likely to be union members than are men.15

A number of studies using individual data have found that, ceteris paribus, blacks are more likely to be union members than are whites, evidence consistent with the finding that the union wage effect is relatively higher for blacks.16 There is little evidence from industry data, however, showing that the racial composition of an industry is an important determinant of its level of unionism. On the basis of preliminary results, the race variable was excluded from the unionism function.

The effect of the age structure on unionism is ambiguous. To the extent that older workers have fewer years over which to accumulate nonpension benefits and that unions flatten the age-earnings profile, older workers will be less likely to prefer union membership. Farber and Saks report
evidence from union certification elections that indicates that older workers are less likely to vote for unions.\textsuperscript{17} On the other hand, if senior workers have relatively strong job attachment and low mobility, their expected benefits from unionism may be high (particularly nonwage benefits from seniority rules, grievance procedures, and health and pension benefits), while organizing costs may be low. Or, unionism may be higher in older industries, which in turn may have older work forces. Empirical evidence on the effects of age is inconclusive, most studies finding age to be an insignificant determinant of unionism.\textsuperscript{18} We believe an experience variable is preferable to age, given that the average length of a working life varies little with years of schooling and age at entry.

Region may be related to unionism due to differences in tastes and differences in worker or labor market characteristics not captured by the model’s other variables. While differences across areas in unionization are quite large, evidence in Hirsch suggests that the partial effects of region per se may be small.\textsuperscript{19} We will account for the percent of an industry’s labor force in the four broad census regions.

Both industry product market concentration and establishment size can potentially affect the demand and supply of union services, though the empirical evidence on the effect of each, holding the other constant, is limited.\textsuperscript{20} Studies by Ashenfelter and Johnson and by Lee, utilizing simultaneous-equation techniques, have found concentration to be positively associated with unionism. Descriptive analysis by Rose indicates that election-unit size is negatively correlated with union wins in NLRB elections. Kochan’s results are inconclusive but suggest that worker attitudes toward unionism are less favorable in very small or large firms.\textsuperscript{21} We include both a four-firm concentration ratio variable (CON) and an average firm-size variable (SIZE) in the unionism equation for manufacturing industries, where such data are available. A positive relationship between CON and unionism is likely to reflect lower organizing costs and lower labor demand elasticities in highly concentrated industries. A positive relationship between firm size and unionism probably also reflects lower organizing costs, plus higher expected benefits from union membership in larger firms.\textsuperscript{22}

To summarize, interindustry differences in unionism are expected to be related to earnings dispersion and those variables systematically related to the benefits and costs of unionism. The following specification reflects the discussion above.

\[
\bar{U}_i = c_0 + c_1 \ln E_i + c_2 \ln \bar{E}_i + c_3 BC_i + c_4 \text{FEM}_i + c_5 \text{CON}_i + c_6 \text{SIZE}_i + c_7 \text{REGION}_i + w_i,
\]

where \(\bar{U}_i\) = proportion of workers who are of Kentucky, 1982, examine in depth the effects of industry characteristics and find that concentration, capital intensity, and size increase the likelihood of union membership.

\textsuperscript{17}Farber and Saks, “Why Workers Want Unions.”
\textsuperscript{18}An exception is Antos, Chandler, and Mellow, “Sex Differences in Union Membership,” who find older workers more likely to be union members.
\textsuperscript{19}Hirsch, “The Determinants of Unionization.”
\textsuperscript{20}In an analysis of the determinants of unionism in British industries, Bain and Elsheikh find establishment size and unionism to be positively related, despite a very small sample size; they find no relationship with sales concentration, however. See G. S. Bain and Farouk Elsheikh, “An Inter-Industry Analysis of Unionisation in Britain,” \textit{British Journal of Industrial Relations}, Vol. 17, No. 2 (July 1979), pp. 137–57. Kahn, “Unionism and Relative Wages,” includes both variables in a unionism equation (within a simultaneous framework), but neither is statistically significant. Barry T. Hirsch and Mark C. Berger, “Union Membership Determination in Manufacturing: The Effects of Concentration, Capital Intensity, and Firm Size,” unpublished paper, University
union members (or, alternatively, covered by collective bargaining) in industry \(i\), \(\sigma^2 \ln E_i\) is endogenous and equals the variance of the natural log of earnings in industry \(i\), \(\ln E_i\) is endogenous and equals the natural log of the geometric mean of earnings in industry \(i\), \(BC_i\) = percent blue-collar workers in industry \(i\), \(FEM_i\) = percent female workers in industry \(i\), \(w_i\) = a random error term with zero mean and constant variance, and other variables are as previously defined. As discussed above, we expect \(c_2, c_3, c_5, c_6, c_7 > 0\) and \(c_4 < 0\); \(c_1\) is a priori indeterminate.

Data

The unit of observation for the empirical work in the paper is the three-digit Census of Population industry. Data on union membership (U-MEM) and collective bargaining coverage (U-COV) among private sector workers are provided for three-digit census industries in Freeman and Medoff.\(^{23}\) Their data on union membership were derived from the 1973–75 Current Population Survey (CPS) tapes, while data on the extent of collective bargaining coverage—which measures the proportion of workers in establishments where a majority of workers are covered by union-management agreements—were derived from the 1968–72 Expenditures for Employee Compensation (EEC) establishment surveys conducted by the Bureau of Labor Statistics. Freeman and Medoff provide data for 209 three-digit industries. The sample is reduced here to 187 industries.\(^{24}\) While the time periods of the unionism variables do not coincide precisely with the period of the 1970 Census of Population, relative differences across industries in unionism were almost certainly stable over these years. For the 187 industries in our sample (76 in manufacturing and 111 in nonmanufacturing) the unweighted mean of the union membership variable is .24 (.36 and .16 respectively) while the unweighted mean for collective bargaining coverage is .30 (.46 and .19 respectively). Despite differences in definition, time period, and source, the two measures are highly correlated, the first-order simple correlation being .89 (.76 and .88 respectively).

All other variables, except the percent of blue-collar workers, industry concentration, and average firm size, were derived from the Census of Population. Description of the source and derivation of variables is provided in the Appendix. All variables from the census are for male workers only, but the unionism variables are calculated from a sample of male and female workers.

Although derivation of most of the variables is straightforward and requires no elaboration, discussion of the earnings and earnings dispersion variables is in order. The census provides grouped earnings data with the number of persons in each of thirteen earnings cells, the open-ended category being $25,000 and over. Median earnings is provided, but not the mean or total earnings in the industry. In order to derive measures of the geometric mean and of earnings dispersion, a “micro” distribution was simulated from the grouped data. We first estimated the mean of the open-ended category by assuming the upper tail was characterized by the Pareto distribution. Estimates were made from a regression within each industry of the log of the proportion of persons above each earnings level (the highest six earnings cells were used) on the log of earnings. The slope coefficient on the log of earnings is an estimate of Pareto’s distribution parameter, \(\alpha\), and the mean earnings of the open-ended category was obtained by \((\alpha / \alpha - 1)25,000\).\(^{25}\)

---

\(^{23}\)Freeman and Medoff, “New Estimates of Private Sector Unionism,” Table 2, pp. 155–61.

\(^{24}\)Deleted were twelve industry groupings in which workers were “not specified” and sample size from the CPS was 25 or less, two industries for which sample sizes were 11 and 6, four industries for which estimation of earnings distributions were difficult due to a large percentage of individuals in the open-ended earnings interval, and four additional industries in which data were not provided for collective bargaining coverage.

\(^{25}\)In addition to the four industries excluded because of the large share of persons in the open-ended category, two industries (security and brokerage firms, legal services) had the mean value of the open-ended category assigned. For a discussion of the Pareto distribution see F. A. Cowell, Measuring Inequality: Techniques for the Social Sciences (New York: Halsted Press, 1977).
A Lorenz curve was then constructed in order to simulate the micro distribution by using a cubic spline technique modified from a published routine by Greville. The cubic spline method has been used to calculate Gini ratios and has been found to produce accurate results even where there are few data points. Gini coefficients were calculated after integration of the area under the constructed Lorenz curve. In order to simulate a micro distribution, each industry’s earnings distribution was divided into 100 intervals (a compromise between accuracy and cost), each interval containing the number of persons and amount of earnings implied by the constructed Lorenz curve. From this generated data and the assumption that all persons within these narrow classes have identical earnings, we calculated the log of the geometric mean of earnings and the variance of the log of earnings. The simple correlation of the log of the geometric mean is .966 with the log of the median and .847 with the log of the mean. The correlation between $\sigma^2(\ln E)$ and the Gini coefficient is .927.

Variables measuring the distribution of schooling, age and experience, and weeks worked were also estimated from grouped data. It was assumed that all individuals were at the midpoints of their respective intervals; thus, all standard deviations and variances are underestimated. This presents no problem for estimation if relative differences across industries are accurately measured.

**Empirical Results**

Complete regression results for the earnings, earnings dispersion, and unionism equations for manufacturing and nonmanufacturing using 3SLS are presented in Tables 1, 2, and 5, respectively, while partial results on the endogenous variables using OLS, 2SLS, and 3SLS are presented in Tables 3, 4, and 6. Estimated coefficients of the exogenous variables showed little sensitivity to estimation technique. The relative trade-offs between equation-by-equation (2SLS) and systems (3SLS) approaches are well known, though no clear criterion exists for choosing between them. Because some fairly high cross-equation correlations between residuals were found to exist, we believe the 3SLS estimates are the most reliable.

The OLS earnings functions obtained $R^2$s of .96 and .89 in manufacturing and nonmanufacturing, respectively, the earnings dispersion equations .87 and .75, and the unionism equations .64 and .57. No overall measures of goodness of fit are generally presented for 2SLS or 3SLS results. McElroy has shown, however, within the context of seemingly unrelated regressions (SUR), how to derive an “$R^2$” goodness-of-fit measure that is a monotonic transform of the appropriate asymptotic $F$ statistic. A similar generalized least squares (GLS) procedure can be used for 2SLS and 3SLS to derive goodness-of-fit measures bounded by [0,1] and corresponding to the approximate $F$ test on all nonintercept parameters. The 2SLS “$R^2$”s are virtually identical to those reported for OLS, while the weighted “$R^2$”s for the 3SLS systems reported in Tables 1, 2, and 5 are .92 in manufacturing and .89 in nonmanufacturing.

The OLS earnings functions obtained $R^2$s of .96 and .89 in manufacturing and nonmanufacturing, respectively, the earnings dispersion equations .87 and .75, and the unionism equations .64 and .57. No overall measures of goodness of fit are generally presented for 2SLS or 3SLS results. McElroy has shown, however, within the context of seemingly unrelated regressions (SUR), how to derive an “$R^2$” goodness-of-fit measure that is a monotonic transform of the appropriate asymptotic $F$ statistic. A similar generalized least squares (GLS) procedure can be used for 2SLS and 3SLS to derive goodness-of-fit measures bounded by [0,1] and corresponding to the approximate $F$ test on all nonintercept parameters. The 2SLS “$R^2$”s are virtually identical to those reported for OLS, while the weighted “$R^2$”s for the 3SLS systems reported in Tables 1, 2, and 5 are .92 in manufacturing and .89 in nonmanufacturing.

Earnings equation and union wage effects. Regression results from the earnings functions, presented in Table 1, are consistent with expectations. The coefficients on mean schooling suggest an average rate of return to schooling of about 10 or 11 percent in 1969. These estimates are higher than regression estimates of the rate of return from individual earnings functions using

---


27Complete OLS and 2SLS results are available from the author on request.

Table 1. Earnings Function: 3SLS Results
(asymptotic t-ratios in parentheses)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Manufacturing</th>
<th>Nonmanufacturing</th>
</tr>
</thead>
<tbody>
<tr>
<td>U-MEM</td>
<td>.166</td>
<td>.021</td>
</tr>
<tr>
<td></td>
<td>(2.18)</td>
<td>(.12)</td>
</tr>
<tr>
<td>S</td>
<td>.111</td>
<td>.099</td>
</tr>
<tr>
<td></td>
<td>(7.62)</td>
<td>(7.11)</td>
</tr>
<tr>
<td>σ (S)</td>
<td>-.015</td>
<td>-.132</td>
</tr>
<tr>
<td></td>
<td>(-.54)</td>
<td>(-3.97)</td>
</tr>
<tr>
<td>t</td>
<td>.064</td>
<td>.104</td>
</tr>
<tr>
<td></td>
<td>(2.09)</td>
<td>(3.98)</td>
</tr>
<tr>
<td>t²</td>
<td>-.0009</td>
<td>-.0016</td>
</tr>
<tr>
<td></td>
<td>(-1.49)</td>
<td>(-3.11)</td>
</tr>
<tr>
<td>σ (t)</td>
<td>-.042</td>
<td>-.040</td>
</tr>
<tr>
<td></td>
<td>(-2.47)</td>
<td>(-1.95)</td>
</tr>
<tr>
<td>InWW</td>
<td>1.685</td>
<td>1.082</td>
</tr>
<tr>
<td></td>
<td>(5.79)</td>
<td>(3.68)</td>
</tr>
<tr>
<td>CON</td>
<td>-.043</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>(-1.12)</td>
<td>—</td>
</tr>
<tr>
<td>SIZE</td>
<td>-.010</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>(-1.29)</td>
<td>—</td>
</tr>
<tr>
<td>BC</td>
<td>—</td>
<td>−.0020</td>
</tr>
<tr>
<td>REGION</td>
<td>included</td>
<td>included</td>
</tr>
<tr>
<td>constant</td>
<td>.854</td>
<td>3.413</td>
</tr>
</tbody>
</table>

Dependent variable is the mean of log earnings, InE. U-MEM is endogenous. In manufacturing, n=76; in nonmanufacturing, n=111.

The percent blue-collar (production) workers is included as a control variable in the nonmanufacturing equation (its coefficient is close to zero in manufacturing). In manufacturing, average firm size (measured by sales) is negatively but not significantly related to earnings. Freeman and Medoff found a significant effect of firm size on earnings using microdata, while Masters has found plant (employment) size to be significant with aggregate data. Concentration is found to be negatively (though not significantly) related to earnings. Most previous studies have not found concentration to have a significant effect on earnings, however, after controlling in detail for personal characteristics.

The effect of unions on wage rates has been the focus of substantial study. We measure the proportionate union wage effect \( m = \exp(\alpha) - 1 \), where \( \alpha \) is the coefficient on \( U \), and multiply by 100 for the percentage effect. Estimated union (weekly) wage effects in manufacturing are 3, 11, and 18 percent using OLS, 2SLS, and 3SLS, respectively, while in nonmanufacturing they are 19, 7, and 2 percent, respectively. Almost all these estimates have large standard errors.

Despite the fact that, until recently, most lower as schooling level increases, and earnings profiles are flatter in industries with older work forces. As expected, the log of earnings is significantly related to the log of weeks worked, and in manufacturing the elasticity of annual earnings with respect to weeks worked is significantly greater than unity.

1970 census data (these run about 7 percent), but are quite similar to estimates of the internal rate of return for this period. The coefficients on \( t \) and \( t^2 \) reflect the concavity of the earnings-experience profile, the profile being steeper in nonmanufacturing industries. The coefficients from the 3SLS regressions indicate that the earnings profile peaks at 35.6 years of experience (51.5 years of age) in manufacturing and 32.5 (49.2) in nonmanufacturing.

Inclusion of the standard deviations of schooling and experience, \( \sigma (S) \) and \( \sigma (t) \), adds significant explanatory power to the earnings equations, though \( \sigma (S) \) is not significant in the manufacturing regressions. The negative sign on each indicates that \( \rho (r,S), \rho (r^*,t) < 0 \), implying that the rate of return to schooling is on average 29Chiswick, Income Inequality, pp. 110–11, provides a discussion regarding the determination of \( \gamma \). 
estimates of union wage effects were based
on grouped industry data that were probably
inferior to the data used here (previous
studies have used older data for manufac-
turing only, with fewer observations, less
precise measures of unionism, and fewer
controls for personal characteristics), we do
not regard our estimates as reliable. Not only
are the standard errors of these estimates
large, but it was found that the union coef-
ficients were sensitive to the functional
form of the dependent variable (estimates
from the theoretically appropriate log of the
geometric mean were higher than from the
arithmetic mean or median), to the choice of
an earnings or wage variable (even when
controlling for hours and weeks worked on
the right-hand side), and to the estimation
procedure.32 Inclusion of $\sigma(S)$ and $\sigma(t)$
in the earnings function and the joint estima-
tion of the earnings and earnings-disper-
sion equations reduced estimated union
wage effects.

As suggested some time ago by Rosen, and
more recently by Freeman and Medoff and
by Lewis, wage rates may be affected not
only by union membership or coverage, but
also by the extent of organization in an in-
dustry.33 While these separate effects cannot
be identified with industry data (thus mak-
ing interpretation of the unionism coef-
ficient difficult), their existence suggests
the possibility of nonlinear wage effects
with respect to the extent of industry cover-
age. Unfortunately, the estimation of non-
linear union wage effects from our data was
found to be highly imprecise.34 Finally, the
existence of interactions between unionism
and other explanatory variables, while not
statistically significant with our data, again
makes precise estimation of the average
union wage effect difficult.

The problems encountered here in esti-
mating the average union/nonunion differen-
tial from grouped data suggest that the
use of individual data might be preferable,
although analyses using such data also
exhibit sensitivity to both specification and
estimation technique. Industry data are
better suited, however, for examining inter-
industry differences in unionism and earn-
ings dispersion. Fortunately, other results of
our model were rarely found to be sensitive
to the changes in specification outlined
above.

Earnings variation equation. Empirical
results from the earnings dispersion model
are presented in Table 2. As predicted, earn-
ings inequality is positively and signifi-
cantly related to employment dispersion,
the coefficients on $\sigma(\ln WW)$ being highly
significant in both manufacturing and non-
manufacturing. The effects of the schooling
and experience distributions are obscured
somewhat by the inability to hold constant
the covariances between schooling, experi-
ence, and weeks worked. Consistent with
the human-capital model, earnings in-
equality is positively and significantly re-
lated to the level of schooling, ceteris par-
ibus.

Counter to the prediction from the hu-
man-capital model, however, the variance
in schooling is negatively related to earn-
ings inequality in nonmanufacturing,
while unrelated in manufacturing. This
probably results from our inability to hold constant
the covariances between schooling and experience, which was shown to be sig-
ificant in previous work. Because of the
secular increase in schooling, earnings

32Depending on specification, moving from OLS to
2SLS or 3SLS caused estimates of $m$ both to decrease
and increase. Freeman and Medoff, “The Impact of
Collective Bargaining,” critically review the conflicting
results from studies using simultaneous-equation
techniques.

33Sherwin Rosen, “Trade Union Power, Threat
Effects and the Extent of Organization,” Review of
185–96; Freeman and Medoff, “The POW Relation-
ship”; and H. Gregg Lewis, “Interpreting Unionism
Coefficients in Wage Equations,” unpublished paper,
Duke University, August 1980.

34Nonlinear union wage effects were estimated
within a simultaneous model by using the procedure
described in Harry H. Kelejian, “Two-Stage Least
Squares and Econometric Systems Linear in Para-
eters but Nonlinear in the Endogenous Variables,”

66, No. 334 (June 1971), pp. 573–74. In addition,
wage effects were estimated using a piecewise method
similar to Rosen, “Trade Union Power.” Some evi-
dence of an increasing wage effect with coverage was
found for manufacturing; no such pattern was found
in nonmanufacturing. In all cases, standard errors of
the estimates were large.
Table 2. Earnings Dispersion Equation: 3SLS Results.  3
(asymptotic t-ratios in parentheses)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Manufacturing</th>
<th>Nonmanufacturing</th>
</tr>
</thead>
<tbody>
<tr>
<td>U-MEM</td>
<td>-.147</td>
<td>-.716</td>
</tr>
<tr>
<td></td>
<td>(-1.63)</td>
<td>(-2.76)</td>
</tr>
<tr>
<td>32</td>
<td>.0029</td>
<td>.0045</td>
</tr>
<tr>
<td></td>
<td>(3.18)</td>
<td>(5.11)</td>
</tr>
<tr>
<td>σ²(S)</td>
<td>-.0003</td>
<td>-.0235</td>
</tr>
<tr>
<td></td>
<td>(-.03)</td>
<td>(-2.47)</td>
</tr>
<tr>
<td>2t</td>
<td>-.0005</td>
<td>.0004</td>
</tr>
<tr>
<td></td>
<td>(-2.68)</td>
<td>(2.23)</td>
</tr>
<tr>
<td>σ²(t)</td>
<td>.0050</td>
<td>.0009</td>
</tr>
<tr>
<td></td>
<td>(7.30)</td>
<td>(.97)</td>
</tr>
<tr>
<td>σ²(InWW)</td>
<td>1.966</td>
<td>2.619</td>
</tr>
<tr>
<td></td>
<td>(4.06)</td>
<td>(5.75)</td>
</tr>
<tr>
<td>BLACK</td>
<td>.012</td>
<td>.005</td>
</tr>
<tr>
<td></td>
<td>(4.01)</td>
<td>(.98)</td>
</tr>
<tr>
<td>REGION included</td>
<td></td>
<td></td>
</tr>
<tr>
<td>constant</td>
<td>-.736</td>
<td>-.259</td>
</tr>
</tbody>
</table>

3Dependent variable is variance of log earnings, σ²(InE). U-MEM is endogenous. In manufacturing, n=76; in nonmanufacturing, n=111.

differentials between highly schooled young workers and less highly schooled older workers are smaller than they would otherwise be. In manufacturing, the level of experience is associated with lower earnings dispersion while the variance of experience is positively associated with dispersion. In nonmanufacturing, both are positively related to earnings dispersion within industries. The percentage black employment is found to be positively related to earnings inequality within manufacturing industries, but not in nonmanufacturing.

As predicted, unionism is found to be negatively and significantly associated with intra-industry dispersion in earnings in both the manufacturing and nonmanufacturing sectors. These findings lend strong support to those of Freeman regarding the equalizing effects of union wage-standardization policies. Moreover, unionism is found to decrease within-industry dispersion even after accounting for the endogeneity of the level of unionism.

In order to examine the robustness of this important finding, identical models were estimated using, alternatively, union coverage (U-COV) as a measure of unionism and the Gini coefficient (GINI) as a measure of earnings dispersion. These OLS, 2SLS, and 3SLS partial results are summarized in Table 3. Use of the Gini as a measure of inequality strongly reinforces our findings. Unionism, as measured by U-MEM or U-COV, is found significantly to decrease earnings dispersion in both manufacturing and nonmanufacturing. The Gini coefficient, while not derived from the human-capital inequality model, may be the preferable inequality measure. It is less sensitive to assumptions about the mean of the open-ended earnings interval (estimated by the Pareto distribution) than is the variance of the log, the coefficient of variation over all 187 industries for GINI being .231, while that for σ²(InE) is .479. Similar results are obtained using U-COV as a measure of unionism, except that U-COV is not found to be significantly related to earnings dispersion in manufacturing when using σ²(InE) as the inequality measure.

The results summarized in Table 3 also suggest, particularly for the nonmanufacturing sector, that failure to account for the simultaneous determination of unionism and earnings dispersion results in too low an estimate of the equalizing effect of unionism. Thus the important findings previously reported by Freeman may actually underestimate the true effect of unionism.

As an additional test of the union equalizing hypothesis, the effect of unionism on the proportion of workers in each earnings interval is examined. If unions raise and flatten their members' earnings profiles, we expect unionism to increase significantly the proportion of workers at some earnings interval(s) toward the (upper) middle of the distribution. Thus we should observe a single-peaked pattern of coefficients on unionism; that is, a positive coefficient at

35Michael Sattinger suggested this exercise. For a similar, though more limited analysis, see Michael Sattinger, Capital and the Distribution of Labor Earnings (Amsterdam-New York: North-Holland, 1980), pp. 242-44.
Table 3. Effects of Union Membership and Coverage on Earnings Dispersion: Partial Regression Results.a
( asymptotic t-ratios in parentheses)

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>2SLS</th>
<th>3SLS</th>
<th>OLS</th>
<th>2SLS</th>
<th>3SLS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Manufacturing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U-MEM</td>
<td>-.155</td>
<td>-.167</td>
<td>.147</td>
<td>-.099</td>
<td>-.117</td>
<td>-.150</td>
</tr>
<tr>
<td></td>
<td>(-2.14)</td>
<td>(-1.85)</td>
<td>(-1.63)</td>
<td>(-4.15)</td>
<td>(-3.94)</td>
<td>(-5.15)</td>
</tr>
<tr>
<td>U-COV</td>
<td>-.004</td>
<td>-.007</td>
<td>.013</td>
<td>-.043</td>
<td>-.065</td>
<td>-.079</td>
</tr>
<tr>
<td></td>
<td>(-.07)</td>
<td>(.07)</td>
<td>(.14)</td>
<td>(-2.32)</td>
<td>(-2.00)</td>
<td>(-2.43)</td>
</tr>
<tr>
<td><strong>Nonmanufacturing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U-MEM</td>
<td>-.167</td>
<td>-.468</td>
<td>-.716</td>
<td>-.135</td>
<td>-.268</td>
<td>-.383</td>
</tr>
<tr>
<td></td>
<td>(-1.05)</td>
<td>(-1.73)</td>
<td>(-2.76)</td>
<td>(-3.26)</td>
<td>(-3.68)</td>
<td>(-6.42)</td>
</tr>
<tr>
<td>U-COV</td>
<td>-.120</td>
<td>-.405</td>
<td>-.552</td>
<td>-.099</td>
<td>-.180</td>
<td>-.239</td>
</tr>
<tr>
<td></td>
<td>(-1.06)</td>
<td>(-2.38)</td>
<td>(-3.45)</td>
<td>(-3.38)</td>
<td>(-4.06)</td>
<td>(-6.42)</td>
</tr>
</tbody>
</table>

a Dependent variables are $\sigma^2(\ln E)$ and GINI. In manufacturing, n=76; in nonmanufacturing, n=111.

Table 4. Effects of Unionism on Proportion in Earnings Classes, Partial 2SLS Results.a
( asymptotic t-ratios in parentheses)

<table>
<thead>
<tr>
<th></th>
<th>$1\text{ to} $3,000$</th>
<th>$$3,001\text{ to} $5,000$</th>
<th>$$5,001\text{ to} $7,000$</th>
<th>$$7,001\text{ to} $8,000$</th>
<th>$$8,001\text{ to} $10,000$</th>
<th>$$10,001\text{ to} $12,000$</th>
<th>$$12,001\text{ to} $15,000$</th>
<th>$$15,001$ and over</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Manufacturing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean proportion</td>
<td>.10</td>
<td>.12</td>
<td>.20</td>
<td>.11</td>
<td>.19</td>
<td>.12</td>
<td>.08</td>
<td>.08</td>
</tr>
<tr>
<td>in earnings class</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U-MEM</td>
<td>-.080</td>
<td>-.134</td>
<td>-.046</td>
<td>.083</td>
<td>.199</td>
<td>.063</td>
<td>-.002</td>
<td>-.081</td>
</tr>
<tr>
<td></td>
<td>(-3.50)</td>
<td>(-4.05)</td>
<td>(-1.07)</td>
<td>(4.76)</td>
<td>(6.07)</td>
<td>(2.52)</td>
<td>(-.12)</td>
<td>(-3.89)</td>
</tr>
<tr>
<td>U-COV</td>
<td>-.055</td>
<td>-.133</td>
<td>-.062</td>
<td>.072</td>
<td>.190</td>
<td>.057</td>
<td>-.002</td>
<td>-.067</td>
</tr>
<tr>
<td></td>
<td>(-2.17)</td>
<td>(-3.42)</td>
<td>(-1.33)</td>
<td>(3.64)</td>
<td>(4.56)</td>
<td>(2.08)</td>
<td>(-.13)</td>
<td>(-2.93)</td>
</tr>
<tr>
<td><strong>Nonmanufacturing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean proportion</td>
<td>.18</td>
<td>.12</td>
<td>.17</td>
<td>.09</td>
<td>.15</td>
<td>.10</td>
<td>.08</td>
<td>.12</td>
</tr>
<tr>
<td>in earnings class</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U-MEM</td>
<td>.037</td>
<td>-.031</td>
<td>-.043</td>
<td>.004</td>
<td>.106</td>
<td>.081</td>
<td>.024</td>
<td>-.179</td>
</tr>
<tr>
<td></td>
<td>(1.10)</td>
<td>(-.85)</td>
<td>(-1.20)</td>
<td>(.20)</td>
<td>(2.69)</td>
<td>(2.86)</td>
<td>(.92)</td>
<td>(-2.61)</td>
</tr>
<tr>
<td>U-COV</td>
<td>.004</td>
<td>-.048</td>
<td>-.039</td>
<td>.014</td>
<td>.097</td>
<td>.070</td>
<td>.025</td>
<td>-.123</td>
</tr>
<tr>
<td></td>
<td>(.18)</td>
<td>(-2.19)</td>
<td>(-1.80)</td>
<td>(1.04)</td>
<td>(4.07)</td>
<td>(4.18)</td>
<td>(1.56)</td>
<td>(-2.89)</td>
</tr>
</tbody>
</table>

a Dependent variables are proportions in earnings classes. In manufacturing, n=76; in nonmanufacturing, n=111.

some earnings interval, with its value decreasing as we move toward the tails of the distribution.

Table 4 presents the 2SLS union coefficients from such an analysis for manufacturing and nonmanufacturing. The dependent variables, measuring the proportion of workers in each of eight earnings categories, have been regressed on the endogenous unionism variables (U-MEM and U-COV) and the other variables from the inequality model (these coefficients not presented). The results provide further support that unions significantly decrease within-industry earnings dispersion. In manufacturing, unionism significantly increases the proportion of workers in the $7,000 - 12,000 earnings range (1969 dol-
lars), and in nonmanufacturing in the 
$8,000 - $12,000 range. The average median 
income across industries is in the $7,000 - 
$8,000 interval for both sectors.

**Unionism equation.** Table 5 presents the 
3SLS regression results from the unionism 
equations in the manufacturing and non-
manufacturing sectors. While the primary 
interest in this section is the effect of an 
industry's earnings distribution on its level of 
unionism, we shall briefly summarize other 
results, all of which are consistent with 
expectations.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Manufacturing</th>
<th>Nonmanufacturing</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma^2(\ln E)$</td>
<td>.781 (3.49)</td>
<td>- .140 (-2.75)</td>
</tr>
<tr>
<td>$\ln E$</td>
<td>.850 (3.67)</td>
<td>- .019 (-.36)</td>
</tr>
<tr>
<td>BC</td>
<td>.0114 (5.37)</td>
<td>.0016 (2.57)</td>
</tr>
<tr>
<td>FEM</td>
<td>-.0032 (-2.73)</td>
<td>-.0008 (-1.25)</td>
</tr>
<tr>
<td>$\bar{r}$</td>
<td>.013 (1.78)</td>
<td>.001 (.33)</td>
</tr>
<tr>
<td>CON</td>
<td>.027 (.29)</td>
<td>-</td>
</tr>
<tr>
<td>SIZE</td>
<td>.064 (3.78)</td>
<td>-</td>
</tr>
<tr>
<td>AGR</td>
<td>-</td>
<td>-.112 (-1.68)</td>
</tr>
<tr>
<td>TCU</td>
<td>-</td>
<td>.166 (4.84)</td>
</tr>
<tr>
<td>REGION</td>
<td>included</td>
<td>included</td>
</tr>
<tr>
<td>constant</td>
<td>-8.756</td>
<td>.360</td>
</tr>
</tbody>
</table>

*aDependent variable is proportion of workers who are union members, $U_{MEM}$. The variables $\sigma^2(\ln E)$ and $\ln E$ are endogenous. In manufacturing, $n=76$; in nonmanufacturing, $n=111$.

Industries with a higher percent of blue-
collar workers are found to be significantly 
more unionized in both the manufacturing 
and nonmanufacturing sectors. The per-
centage female is negatively related to 
unionism, though this effect is stronger in 
manufacturing. A positive, but not sta-
tistically significant, relationship is found 
between unionism and the average years of 
experience of an industry's labor force. 
After accounting for other determinants of 
unionism, regional location of an indus-
try's work force has no significant effect.

In manufacturing, unionization is found 
to be significantly greater in industries with 
larger firms. Concentration, which is highly 
correlated with $SIZE$, is not significantly 
related to unionization when $SIZE$ is in-
cluded, but is found to have a positive and 
highly significant coefficient when $SIZE$ is 
excluded. As expected, in nonmanu-
facturing, unionism is found to be relatively 
greater in transportation, communication, 
and utilities ($TCU$), and relatively less in 
agriculture ($AGR$).

Industries with higher earnings levels are 
found to be more unionized in the manu-
facturing sector, even after accounting for 
simultaneity. This result is consistent 
with findings in other studies that have 
suggested that the effects of earnings on 
unionism may be statistically stronger than 
the opposite relationship. No significant 
relationship is found between $\ln(\bar{E})$ and $U$ 
in the nonmanufacturing sector.

As discussed previously, we believe 
unionism is systematically related to the 
dispersion of earnings within an industry. 
On the one hand, industries with more 
homogenous workers will have worker 
preferences distributed more closely around 
the median voter, thus increasing the like-
lihood of unionism, *ceteris paribus*. Like-
wise, some benefits from collective bargain-
ing may be higher and the costs of providing 
union services lower within industries with 
more homogeneous workers. On the other 
hand, greater dispersion in earnings in an 
industry is likely to generate increased de-
mand for collective bargaining coverage as 
a means of lowering this dispersion (while 
raising wage levels) via union wage stan-
dardization policies.

Table 6 summarizes the results of the 
unionism-inequality relationship using 
two unionism measures, $U_{MEM}$ and 
$U_{COV}$, and two earnings dispersion 
measures, $\sigma^2(\ln E)$ and $GINI$. Results in
Table 6. Effects of Earnings Dispersion on Union Membership and Coverage: Partial Regression Results.

(Asymptotic t-ratios in parentheses)

<table>
<thead>
<tr>
<th></th>
<th>U-MEM</th>
<th>U-COV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS</td>
<td>2SLS</td>
</tr>
<tr>
<td>Manufacturing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma^2(\ln E)$</td>
<td>.213</td>
<td>.706</td>
</tr>
<tr>
<td>GINI</td>
<td>-.124</td>
<td>-.824</td>
</tr>
<tr>
<td></td>
<td>(-2.22)</td>
<td>(-1.11)</td>
</tr>
<tr>
<td>Nonmanufacturing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma^2(\ln E)$</td>
<td>-.052</td>
<td>-.087</td>
</tr>
<tr>
<td>GINI</td>
<td>-.406</td>
<td>-.876</td>
</tr>
<tr>
<td></td>
<td>(-2.34)</td>
<td>(-2.40)</td>
</tr>
</tbody>
</table>

Dependent variables are U-MEM and U-COV. In manufacturing, n=76; in nonmanufacturing, n=111.

The manufacturing sector are inconclusive and sensitive to the choice of the inequality and union variables. With $\sigma^2(\ln E)$ as the inequality measure, a positive and significant relationship exists using 2SLS or 3SLS, suggesting that a more dispersed earnings structure increases the level of unionization. No clear-cut relationship is found, however, when using the Gini coefficient as the inequality measure.36 In the nonmanufacturing sector, on the other hand, there exists strong evidence of a negative relationship between earnings dispersion and the extent of unionism, regardless of the measures used. Moreover, the estimated strength of this relationship increases as one moves from OLS to 3SLS. Not only does unionism significantly lower within-industry earnings dispersion, but also nonmanufacturing industries with less dispersion have higher levels of union membership and coverage.

While reasons for the difference in results between the manufacturing and nonmanufacturing sectors are not entirely clear, the findings are suggestive. If unionism is not less likely in manufacturing industries where earnings dispersion is larger, then it is likely that demand by workers in these industries for union-induced wage equalization is strong, offsetting fully or in part the greater difficulties involved in organizing and providing services for less homogeneous workers.

The level of unionism in nonmanufacturing industries is found to be significantly larger the lower is earnings dispersion. Greater worker homogeneity, implying less dispersion in preferences around the median worker, facilitates agreement on collective bargaining coverage and may be associated with higher benefits and lower organizing costs. The results also imply that worker demand for union wage standardization policies has played a relatively smaller role in the nonmanufacturing sector. This appears correct even though unionization in fact is found to have lessened earnings dispersion in both sectors, even after accounting for simultaneity.

36 For additional evidence on the relationship in manufacturing, $\sigma^2(\ln E)$ and GINI were included as variables in a logit union membership equation using individual data from the 1973 CPS tapes. GINI was found to be negatively and significantly related to the likelihood of union membership, while the coefficient on $\sigma^2(\ln E)$ was negative but insignificant. These results (available on request) were identical using either actual earnings dispersion or earnings dispersion predicted from a reduced-form equation.
Summary

The major argument of this paper is that not only do unions affect earnings dispersion within industries, but earnings dispersion will probably also affect an industry's level of unionization. A model is developed in which the determination of earnings and earnings dispersion is based on the standard human-capital model, and the equilibrium level of unionism is determined by factors affecting its benefits and costs and the distribution of preferences around the median voter. The resulting three-equation model is estimated by OLS, 2SLS, and 3SLS, using data from 76 manufacturing and 111 nonmanufacturing three-digit industries from 1970, thus providing evidence on the determinants of unionism, earnings, and earnings inequality.

An important finding of this study is that unionism significantly decreases earnings dispersion, even after accounting for their simultaneous determination. Moreover, the estimated equalizing effects of unionism appear to be larger, if anything, using simultaneous-equation techniques. The estimated effect of earnings dispersion on the level of unionization is found to differ between the manufacturing and nonmanufacturing sectors. In manufacturing, no consistent relationship between unionism and earnings dispersion is found, suggesting that worker demand for union-induced wage equalization may be strong in those industries. In nonmanufacturing, earnings inequality significantly decreases the level of unionism. This suggests that worker demand for wage equalization here is relatively weaker, while the homogeneity of worker preferences and their distribution around those of the median worker are important determinants of the level of unionism.

Data Appendix

<table>
<thead>
<tr>
<th>Variable</th>
<th>Manufacturing Mean (Standard Deviation)</th>
<th>Nonmanufacturing Mean (Standard Deviation)</th>
<th>Definition and Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>U-COV</td>
<td>.460 (.190)</td>
<td>.193 (.222)</td>
<td>Proportion of private sector workers who are in establishments where a majority of workers are covered by a union-management agreement. Derived from 1968-72 Expenditures for Employee Compensation by Freeman and Medoff, &quot;New Estimates of Private Sector Unionism,&quot; Table 2, pp. 155-61.</td>
</tr>
<tr>
<td>$\sigma^2(\ln E)$</td>
<td>.532 (.177)</td>
<td>.953 (.377)</td>
<td>Variance of the natural log of earnings. Construction described in text. Upper tail fitted by Pareto distribution. Lorenz curve constructed by cubic spline technique; a &quot;micro&quot; distribution was generated from constructed Lorenz curve. Earnings data from U.S. Bureau of Census, Census of the Population, Table 12.</td>
</tr>
<tr>
<td>GINI</td>
<td>.322 (.050)</td>
<td>.416 (.087)</td>
<td>Gini coefficient of earnings. Construction described in text and above.</td>
</tr>
</tbody>
</table>
### Data Appendix (Continued)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Manufacturing Mean (Standard Deviation)</th>
<th>Nonmanufacturing Mean (Standard Deviation)</th>
<th>Definition and Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S$</td>
<td>10.93 (1.00)</td>
<td>11.73 (1.49)</td>
<td>Mean years of schooling completed by experienced labor force. Constructed from data in U.S. Bureau of Census, <em>Census of the Population</em>, Table 3, organized in eight intervals. Persons in open-ended interval (17 or more years) assigned 18 years. All other assigned midpoint.</td>
</tr>
<tr>
<td>$T$</td>
<td>24.41 (2.15)</td>
<td>24.15 (3.56)</td>
<td>Mean years of experience calculated by Age $- S - 5$. Mean calculated from data in U.S. Bureau of Census, <em>Census of the Population</em>, Table 34, organized in fourteen intervals. Individuals in open-ended interval (70 years and over) assigned 75. All others assigned midpoint. Standard deviation of experience assumed equal to standard deviation of age. This is correct only if schooling is constant across age intervals.</td>
</tr>
<tr>
<td>$\ln(WW)$</td>
<td>3.827 (.036)</td>
<td>3.764 (.086)</td>
<td>Mean of natural log of weeks worked. Constructed from data in U.S. Bureau of Census, <em>Census of the Population</em>, Table 6, organized in five intervals (midpoints used).</td>
</tr>
<tr>
<td>BLACK</td>
<td>8.77 (4.49)</td>
<td>7.29 (4.45)</td>
<td>Percent of employed who are Negro males. Data from U.S. Bureau of Census, <em>Census of the Population</em>, U.S. Summary, Table 236.</td>
</tr>
<tr>
<td>FEM</td>
<td>27.22 (15.87)</td>
<td>33.41 (21.72)</td>
<td>Percent of experienced labor force female. Data from U.S. Bureau of Census, <em>Census of the Population</em>, Table 1.</td>
</tr>
<tr>
<td>BC</td>
<td>70.57 (13.67)</td>
<td>43.58 (27.490)</td>
<td>Percentage of workers who are craftsmen and kindred workers, operatives, nonfarm laborers, service workers, and farm laborers and foremen. Data derived from 1973–75 CPS tapes by Freeman and Medoff, &quot;New Estimates of Private Sector Unionism,&quot; Table 2, pp. 155–61.</td>
</tr>
<tr>
<td>CON</td>
<td>.413 (.162)</td>
<td>—</td>
<td>Share of industry shipments accounted for by the four leading firms in 1972. The data were bridged from 4-digit SIC industries to 3-digit 1970 Census industries and furnished by James Medoff. See Freeman and Medoff &quot;The POW Relationship,&quot; for a fuller description.</td>
</tr>
<tr>
<td>SIZE</td>
<td>1.339 (1.176)</td>
<td>—</td>
<td>The natural log of (the value of industry shipments/number of firms in the industry in 1972). Sources are the same as for the concentration ratios.</td>
</tr>
</tbody>
</table>