

Working Paper #9 Revised
Industrial Relations Section
Princeton University

Unionism, Relative Wages, and Labor Quality
in U. S. Manufacturing Industries

By

Orley Ashenfelter and George E. Johnson

Princeton University
The University of Michigan

November 1971

Unionism, Relative Wages, and Labor Quality
in U. S. Manufacturing Industries^{*/}

Orley Ashenfelter and George E. Johnson

H. G. Lewis' estimate^{1/} that unions had raised the average wage rate of union workers by 10-15 percent above that of non-union workers^{2/} in the period 1957-58 has served as a benchmark for further study, but it has not gone unchallenged. Most of the more recent work has sought to take advantage of the more extensive data on labor quality which has since become available in order to correct for quality differences prior to estimating the union-nonunion wage differential. Leonard Weiss, for example, used microeconomic census data and estimated average differentials for 1959 for craftsmen and operatives of about 30 percent.^{3/} Frank Stafford, using microeconomic data from the Survey of Consumer Finances, estimated average differentials for 1965 of 26 percent for operatives, 24 percent for craftsmen, 52 percent for laborers, and 18 percent for clerical and sales workers.^{4/} Using more aggregative industry data, but controlling for quality and other

* The authors are indebted for suggestions and criticism to T. A. Finegan, S. Hymans, H. G. Lewis, J. Pencavel, A. Rees, and F. Stafford.

^{1/} H. Gregg Lewis, Unionism and Relative Wages in the United States, Chicago 1963.

^{2/} Lewis' work has established the precedent of measuring the effect of unionism on wages in relative terms. Thus, if the estimating equation is of the form $\ln W_i = \alpha_0 + \alpha_1 U + \dots$, where W_i is the average wage rate and U_i the extent of unionism in the i^{th} industry, the estimated proportionate wage advantage of union workers is $\exp(\hat{\alpha}_1) - 1$.

^{3/} Leonard W. Weiss, "Concentration and Labor Earnings," American Economic Review, March 1966, pp. 96-117.

^{4/} Frank Stafford, "Concentration and Labor Earnings: Comment," American Economic Review, March 1968, pp. 174-80.

differences, Victor Fuchs has estimated union/nonunion differentials of between 28 and 35 percent, and Adrian Throop has estimated a differential of 30 percent.^{5/} Most recently, Sherwin Rosen has produced a series of estimates of union/nonunion wage differentials that range for the most part between 16 and 25 percent, but apparently reach as high as 35 percent in some cases.^{6/} Despite the differences in data and measurement used in these studies it is difficult to ignore both their tendency to lie above Lewis' estimate of 10-15 percent for essentially the same period and their large absolute size. According to Lewis' results a union/nonunion differential as large as 30 percent has not existed in the U. S. since the early 1930's. In view of such a sizeable differential, and the concomitant rewards to organization, it is puzzling why so much of the work force remains unorganized.

One factor which all these estimates of the effect of unionism on relative wages have in common is their basic dependence on the accuracy of a model which posits that unionism and labor quality are exogenous determinants of wages, i.e., that there is a unicausal relationship from the level of labor quality and the extent of unionism to the level of the wage.

^{5/}Victor R. Fuchs, The Service Economy, National Bureau of Economic Research, New York, 1968, Table 54, and Adrian W. Throop, "The Union-Nonunion Wage Differential and Cost Push Inflation," American Economic Review, March 1968, pp. 79-99.

^{6/}For the former see Rosen's "Trade Union Power, Threat Effects and the Extent of Organization," The Review of Economic Studies, April 1969, pp. 185-196 and "On the Interindustry Wage and Hours Structure," Journal of Political Economy, March/April 1969, pp. 249-273. For the latter see Sherwin Rosen, "Unionism and the Occupational Wage Structure in the United States," International Economic Review, June 1970, pp. 269-286.

However, the interesting analysis by Reder^{7/} and the fuller development by Lewis^{8/} suggest that the extent of unionism should be considered as jointly determined with the wage rate. Further, elementary theoretical considerations suggest that the average level of labor quality in an industry should also be treated as an endogenous variable. Consequently, the twin purposes of this paper are to develop an exploratory model in which the extent of unionism, labor quality, and wages are jointly determined endogenous variables and then to estimate the parameters of this model for a cross-section of manufacturing industries. In addition to these new estimates of the effect of unions on relative wages, development of such a model should be of interest in itself because it is more consistent with the traditional framework of neoclassical price theory and because it makes the estimation of the union/nonunion wage differential a problem not only of measurement but an economic problem as well. That is, it becomes necessary both to measure accurately the extent of unionism, labor quality, and wages and to obtain a priori economic information on how these variables are influenced by other exogenous forces.

The plan of the paper is as follows: in Section I we develop a theoretical framework and roughly evaluate the simultaneous equations bias that results from ignoring the endogeneity of labor quality and unionism; in Section II we provide consistent estimates of the model under

^{7/}M. W. Reder, "Unions and Wages: The Problems of Measurement," Journal of Political Economy, April 1965, pp. 188-96.

^{8/}H. Gregg Lewis, "Competitive and Monopoly Unionism" in Philip D. Bradley (ed.), The Public Stake in Union Power, Charlottesville, 1959, pp. 181-208.

alternative specifications; and in Section III we point out the limitations of the results and offer some concluding remarks.

I. A Conceptual Framework

A. Unionism and Relative Wages

With homogeneous labor and an infinitely elastic long run supply curve of labor to each industry, inter-industry wage differences determined under competitive conditions will reflect only differences in the nonpecuniary attributes of employment between industries.^{9/} Letting W_i^C be the competitive wage rate, and P_i be the proportionate money value of the nonpecuniary attributes of the i^{th} industry, inter-industry labor market equilibrium requires that the total wage rate, $W_i^C(1+P_i)$, be the same for all industries. If this were not so, all workers would desire to move to the industry with the highest total wage rate, and this would be inconsistent with equilibrium.

In practice, of course, labor is not homogeneous, so the equilibrium total wage rate will depend on the quality of the workers in an industry. We shall suppose that a major determinant of quality differences is the average extent of investment in human capital, i.e., schooling.^{10/} If, as an

^{9/} These nonpecuniary attributes include such factors as working conditions in and location of the industry. The assumption of an infinitely elastic labor supply curve at the equilibrium wage rate is equivalent to the assumption that all individuals have similar assessments of the money value of the nonpecuniary attributes of the different industries. For evidence that industry labor supply elasticities are in fact very large see Timothy W. McGuire and Leonard A. Rapping, "The Role of Market Variables and Key Bargains in the Manufacturing Wage Determination Process," Journal of Political Economy, September/October 1968, pp. 1015-36.

^{10/} Differences in the extent of other types of human capital, for example on-the-job training, are probably important for some applications. As we will note below, on the basis of available data they seem to be relatively unimportant for this one.

approximation, we ignore the finite nature of life and assume that the equilibrium total wage rate for each individual, $V \equiv W^c(1+P)$, depends only on the number of years of schooling (E), the rate of return on the j^{th} year of schooling (r_j) is implicitly defined by setting the discounted present value of the incremental earnings due to the j^{th} year equal to the sum of its direct and opportunity costs. Algebraically, r_j is given by

$$(1) \quad \int_0^{\infty} (V_j - V_{j-1}) e^{-r_j t} dt = \frac{V_j - V_{j-1}}{r_j} = C_j + V_{j-1},$$

where V_j is the total wage after j years of schooling and C_j the direct schooling costs of the j^{th} year. Now (1) can easily be solved for the recursion

$$(2) \quad V_j = [1 + r_j(1 + k_j)]V_{j-1},$$

where k_j is the ratio of direct schooling costs to opportunity costs (C_j/V_{j-1}), so that after repeated substitution we have

$$(3) \quad V_E = V_0 \prod_{j=1}^E [1 + r_j(1 + k_j)].$$

Assuming the r_j 's and k_j 's are similar over the relevant range, the log transform of (3) is

$$(4) \quad \ln V \approx \ln V_0 + r(1 + k)E. \frac{11/}{}$$

^{11/} For more discussion of this approach see Gary S. Becker and Barry R. Chiswick, "Education and the Distribution of Earnings," American Economic Review, May 1966, pp. 358-69. Note that we use the approximation $\ln(1+X) \approx X$ for small values of X .

If we now write the log linear labor supply condition as

$$(5) \quad \ln W_i^c + \ln(1 + P_i) = a_0 + a_1 E_i,$$

it should be clear that it will do double duty. First, it explicitly allows for the effect of labor quality on the supply price of labor. Second, since (5) is identical to (4), an estimate of a_1 may be taken as an estimate of the average adjusted rate of return to investment in schooling, $r(1+k)$.^{12/} Since a great deal is already known about rates of return to schooling, this procedure has the obvious advantage that the adequacy of estimates of equation (5) may be tested by an appeal to independent evidence.

The proportionate wage advantage of union over nonunion workers is assumed constant and is defined as $m \equiv \frac{W_i^u - W_i^n}{W_i^n} \equiv \frac{W_i^u}{W_i^n} - 1$, where W^u and W^n are the union and nonunion wage rates. It follows that

$$(6a) \quad \ln W_i^u \equiv \ln W_i^n + \ln(1+m).$$

In formulating (6a) we have implicitly allowed for the possibility that W^n differs from W^c , and this difference is $D_i \equiv \frac{W_i^n - W_i^c}{W_i^c} \equiv \frac{W_i^n}{W_i^c} - 1$,

so that

$$(6b) \quad \ln W_i^n = \ln W_i^c + \ln(1+D_i).^{13/}$$

^{12/} Since it is likely that $k = 0$ over the range of E that we will consider (because the direct cost of schooling will be inconsequential over this range), a_1 can probably be taken as simply the average rate of return.

^{13/} In general, we might expect W^n to differ from W^c for one of two reasons. First, the presence of unionism in one part of an industry may result in increased wages in that part of the industry along with a contraction of employment. Wages in the nonunion sector of the industry may thus be bid below what

Now the average wage rate in the typical industry may be taken as the weighted geometric mean of the union and nonunion wage rates and hence its logarithm is

$$(7) \quad \ln W_i = U_i \ln W_i^u + (1-U_i) \ln W_i^n ,$$

where U_i is the fraction of workers who are union members. Although we leave the empirical specification for Section II, if we posit that the nonpecuniary attributes of each industry may be represented by a single variable A_i , we can substitute (6b) into (6a) and the result plus (5) into (7) to obtain the estimating equation

$$(8) \quad \ln W_i = \alpha_0 + \alpha_1 E_i + \alpha_2 U_i + \alpha_3 A_i + \epsilon_{1i} ,$$

where $\epsilon_{1i} = \epsilon'_{1i} + \ln(1+D_i)$ and ϵ'_{1i} is a disturbance term added to (5), $\alpha_1 = r(1+K)$, and $\alpha_2 = \ln(1+m)$. The received procedure is to estimate the α_j by an ordinary least squares regression, which is unexceptionable so long as $\text{Cov}(U_i, \epsilon_{1i}) = \text{Cov}(E_i, \epsilon_{1i}) = 0$. In general, there are two reasons why this assumption may fail. First, if $\ln(1+D_i)$ is correlated

they otherwise would have been by the workers displaced from the union sector. Under our assumption that long run labor supply curves are very elastic to industries, however, D_i must be the same in every industry in this case. This is true because $W_i^n(1+P_i)$ must now be the same for each industry, else all nonunion workers would desire to move to the industry with the highest total wage rate, which would be inconsistent with labor market equilibrium. It then follows from (6B) that $\ln[W_i^n(1+P_i)] - \ln[W_i^C(1+P_i)] = \ln(1+D_i)$, so that $\ln(1+D_i)$ is the difference between two numbers that are the same in each industry. Second, the threat of unionism may induce nonunion employers to raise wages above W^C . These difficulties are discussed in Lewis, Unionism ..., op. cit., pp. 27-40, and Orley Ashenfelter, "The Effects of Unionization on Wages in the Public Sector: The Case of Fire Fighters," Industrial and Labor Relations Review, January 1971, pp. 191-202.

with either U_i or E_i , then ordinary least squares will be a biased and inconsistent estimator for (8). An instrumental variables estimator would be called for. Second, as we shall point out in the remainder of this section, there are important theoretical reasons for supposing that U and E are not exogenous variables in (8).

B. The Determinants of Union Membership by Industry

Although it does not seem to have been tested empirically, the notion that the degree of unionism in an industry depends on the absolute or relative wage rate in that industry has existed for some time.^{14/} The derivation of the equilibrium relation between the extent of unionism and wages is most fruitfully approached by examining separately the factors influencing both the demand for and supply of union services in a particular industry and then obtaining the reduced form relation. To the typical worker, the benefits of unionism are derived from (i) the potential relative wage advantage due to union membership and (ii) the influence of the union on the non-pecuniary aspects of his work attachment, especially through grievance procedures and seniority systems.^{15/} From this point of view the purchase of unionism

^{14/} John Dunlop, for example, posited a "membership function" which shows "the total amount of labor that will be attached to the labor organization at each wage rate." See his Wage Determination Under Trade Unions, New York, 1944, p. 33.

^{15/} For a lucid analysis of these institutions see Albert Rees, "Some Non-Wage Aspects of Collective Bargaining," in Bradley (ed.), op. cit., pp. 124-42. Since the argument that non-pecuniary aspects of unionism are an important factor in worker benefits is one of the primary hypotheses from which equation (24) below is deduced, the evidence for it requires elaboration. First, there is apparently near unanimity in the industrial relations literature that non-pecuniary factors are a key determinant in the worker's decision to join a

should be treated in part as an investment good and in part as a consumption good. The decision to join a union on the part of the typical potential union member will then depend on the union/nonunion wage differential and on the price of membership, initiation fees and dues, relative to the prices of other goods (M_i) and on his income level, which, holding hours of work constant, can be represented by the industry wage rate (W_i). Aggregating over all workers in the industry, the fraction of workers who demand union services (U_i^D) will depend upon M_i and W_i as well as a taste parameter (τ_i) which reflects the preferences for unionism in the industry and the union relative wage advantage (m). Symbolically,

$$(9) \quad U_i^D = f(M_i, W_i, \tau_i, m)$$

where $f_M < 0$, $f_W > 0$ (on the assumption that unionism is a normal good),

union. See Joel Seidman, et al., The Worker Views His Union, Chicago 1958, especially pp. 74-78. In their study of a large United Steelworkers local these same authors note in "Why Workers Join Unions," Annals of the American Academy of Political and Social Science, March 1951, pp. 75-84, that not a single worker they interviewed mentioned joining the union "to get higher wages." Apparently most workers had joined out of sympathy or conviction, a good bit of which was the result of unhappy work experiences which it was believed union procedures would have eliminated. This is also the view in E. Wight Bakke's "Why Workers Join Unions," in Joseph Shister (ed.), Readings in Labor Economics and Industrial Relations, Chicago, 1956. More recently, a well-known management consultant has concluded after experience with employee attitude surveys in recent years that, "Employees who vote in favor of the union usually feel aggrieved, and believe that the union will correct their personal grievances." See Jules Bank, "Why Companies Lose Union Elections," Factory, April 1968. Second, it should be recognized that this hypothesis has other testable implications. For example, if the presence of unionism adds significant non-pecuniary benefits to a job, then, holding earnings constant, voluntary labor mobility (i.e., quits) in unionized industries should be less than in industries which are not unionized. J. H. Pencavel has recently found in an inter-industry cross-section for 1959 that quits per hundred employees are reduced about .01 percentage points per percentage point of unionism and that this relation is highly significant. See J. H. Pencavel, An Analysis of the Quit Rate in American Manufacturing Industry, Research Report No. 114, Industrial Relations Section, Princeton University, 1970.

and $f_m > 0$.^{16/} At the same time there will be an equation that represents the supply of union services. With a given stock of union entrepreneurial talent in the economy, the supply of union services to an industry - expressed in terms of the proportion of the work force (U_i^S) which can be accommodated by that amount of union services - will depend positively on the price of unionism, negatively on some index of the cost of providing union services to that industry (Z_i),^{17/} and also on a parameter reflecting employers' tastes (or distastes) for unionism (ζ_i). In equation form this is

$$(10) \quad U_i^S = g(M_i, Z_i, \zeta_i) ,$$

where $g_M > 0$ and $g_Z < 0$.

In static equilibrium $U_i = U_i^D = U_i^S$, and the solution values of U_i and M_i depend on the values of W_i , Z_i , τ_i , ζ_i , and m . We are especially interested in the union membership function.

$$(11) \quad U_i = U(W_i, Z_i, \tau_i, \zeta_i, m)$$

^{16/} For estimates of a linear version of (9) produced as a response to these remarks see John H. Pencavel, "The Demand for Union Services: An Exercise," Industrial and Labor Relations Review, January 1971, pp. 180-190. There is also some direct evidence on the relevance of the price of union membership to membership decisions. A management consultant argues the following in "Why They Voted Against the Union," The Bobbin [a management magazine], December 1967: "Another economic factor which has helped management is the cost of union membership. The pocketbook issue - dues, assessments, fines, the cost of the campaign - can be a telling one ... the dues and assessments union members must pay, and over which they have little control, are tantamount to a reduction in wages. This issue has resulted in ... pro-management votes."

^{17/} The precise specification of the variable Z is discussed in more detail in Section II, but evidence that in a stabilized industrial relations system the costs of organization are an important factor in union decisions to organize workers is contained in Jack Barbash, The Practice of Unionism, New York, 1956, pp. 18-25.

and it is easily shown that $\partial U_i / \partial W_i = -g_M f_W / (f_M - g_M) > 0$ and $\partial U_i / \partial Z_i = f_M g_Z / (f_M - g_M) < 0$.^{18/} There are many interesting implications concerning determinants of the taste parameters and their resultant effects on the growth and dispersion of unionism,^{19/} but initially we shall assume that we are dealing with a period of time in which both τ and ζ are invariant with respect to industry. In addition, we should expect that with rising real wages over time both the demand for and supply of unionism functions would shift upward, reflecting the income effect and the opportunity cost of union services, respectively. The conjunction of these shifts and the former assumption about tastes means that this model does not have unambiguous implications for the growth of union membership in the aggregate. Rather this model only suggests that the extent of union membership will tend to be relatively greater in industries with relatively high wages and in which the costs of organizing and servicing union members are relatively low.^{20/}

The linearized stochastic equivalent of (11) is

$$(12) \quad U_i = \beta_0 + \beta_1 \ln W_i + \beta_2 Z_i + \epsilon_{2i} ,$$

^{18/} Differentiate the equilibrium condition totally, $(f_M - g_M) dM_i + f_W dW_i - g_Z dZ_i = 0$, in order to obtain the partial derivatives $\partial M_i / \partial W_i$ and $\partial M_i / \partial Z_i$; then $\partial U_i / \partial W_i = g_M (\partial M_i / \partial W_i)$ and $\partial U_i / \partial Z_i = f_M (\partial M_i / \partial Z_i)$.

^{19/} See both Lewis, "Competitive . . .," *op. cit.*, and Reder, *op. cit.*, for an interesting discussion of some of these.

^{20/} For a model of the growth of trade unionism which can be rationalized in terms of changes over time of the parameters τ and ζ , see Orley C. Ashenfelter and John H. Pencavel, "American Trade Union Growth: 1900-1960," *Quarterly Journal of Economics*, August 1969.

where ϵ_{21} is an error term.^{21/} Suppose now we ignore the presence of (12) and proceed to estimate (8) by ordinary least squares. This will lead to biased and inconsistent estimators for α_1 and α_2 . Intuitively, we would expect the OLS estimator of α_1 to be biased downward and the OLS estimator of α_2 to be biased upward because we will inadvertently have given some of the "credit" for the effect of wages on unionism and for the effect of labor quality on wages to the effect of unionism on wages. More formally, it can readily be shown that

$$(13) \quad \text{plim}_{N \rightarrow \infty} (\hat{\alpha}_2 - \alpha_2) = \frac{1 - \rho_{E,A}^2}{(1 - \alpha_2 \beta_1) \Delta} [\beta_1 \text{Var}(\epsilon_1) + \text{Cov}(\epsilon_1, \epsilon_2)],$$

where Δ is the determinant of a positive definite matrix (hence $\Delta > 0$) and $\rho_{E,A}^2$ is the squared simple correlation coefficient between quality and nonpecuniary attributes. Except for very special cases we clearly require that $\beta_1 = \text{COV}(\epsilon_1, \epsilon_2) = 0$ for the bias to vanish, in which case this model would be diagonally recursive. Further if we ignore the term $\text{COV}(\epsilon_1, \epsilon_2)$, about which we lack information, the a priori conditions on the signs of the coefficients of (8) and (12) ensure that the bias will be positive.^{22/} Similarly, the OLS estimator for β_1 will be biased upward

^{21/} It should be clear that a complete model would contain specific reference to equations (9) and (10). Reliable information on union dues and fees is so scanty, however, that in what follows we restrict our attention to (11), the membership function.

^{22/} This statement is strictly correct only if $(1 - \alpha_2 \beta_1) > 0$, but this must be so for dynamic stability of the system.

if $\alpha_2 > 0$. It can also be shown that the bias of the OLS estimator of α_1 is

$$(14) \quad \text{plim}_{N \rightarrow \infty} (\hat{\alpha}_1 - \alpha_1) = - \frac{b_{QU \cdot A} (1 - \rho_{UA}^2)}{(1 - \alpha_2 \beta_1) \Delta} [\beta_1 \text{Var}(\epsilon_1) + \text{Cov}(\epsilon_1, \epsilon_2)]$$

where $b_{QU \cdot A}$ is the regression coefficient of unionism on labor quality given the nonpecuniary attributes of the industry and ρ_{UA}^2 is the squared simple correlation coefficient between U_i and A_i . Presumably $b_{QU \cdot A} > 0$ due to the positive effect of quality on wages and the positive effect in turn of wages on unionism, so α_1 is most likely biased downward.

C. The Determinants of Labor Quality by Industry

In the preceding discussion we have assumed that E_i , a proxy for the average quality of an industry's labor force, may be taken as fully exogenous. In fact, the firms in an industry must set their hiring standards just as they must set (or bargain with a union over) wage rates. Consequently, we should expect that E_i will be systematically determined by economic variables and that it will be necessary to recognize this when equations (8) and (12) are estimated.

One straightforward and potentially fruitful approach to this question is to inquire as to how a typical firm benefits from employing higher quality workers.^{23/} In order to do this it is necessary to specify the way in which

^{23/}A somewhat similar approach to the one employed here is suggested by M. Blaug, "Approaches to Educational Planning," Economic Journal, June 1967, esp. pp. 281-82.

$$(18) \quad \frac{\partial \pi}{\partial K} = P \frac{\partial X}{\partial K} - r = 0, \text{ and}$$

$$(19) \quad \frac{\partial \pi}{\partial Q} = P \frac{\partial X}{\partial(aL)} \frac{da}{dQ} - \frac{\partial W}{\partial Q} = 0,$$

subject to satisfaction of the second-order conditions which include

$$\frac{d^2 a/dQ^2}{a} - \frac{\partial^2 W/\partial Q^2}{Q} > 0. \text{ Combining (17) and (19) gives}$$

$$(20) \quad \frac{da}{dQ} \frac{1}{a} = \frac{\partial W}{\partial Q} \frac{1}{W},$$

which says that the proportionate increase in labor efficiency with respect to an increase in labor quality must be set equal to the proportionate increase in wages with respect to an increase in labor quality.^{25/} This is illustrated in Figure 1. The second-order condition requires that the slope of the "demand" schedule be less than the slope of the "supply" schedule. If we assume that E is a good proxy for Q , which we shall do in estimation, then our assumptions regarding (8) imply that $\frac{\partial W}{\partial Q} \frac{1}{W} \approx \frac{\partial W}{\partial E} \frac{1}{W} = a_1 \approx r(1+k)$, a constant, so that $\frac{\partial^2 W}{\partial Q^2} \frac{1}{W} = 0$, and the "supply" schedule is perfectly elastic as shown by the dashed line in Figure 1. For this case there must be diminishing returns to labor quality for stability.

Now on this interpretation differences in labor quality across firms and industries will be due to differences in the position of the $\frac{da}{dQ} \frac{1}{a}$

^{25/} Note that (20) implies that for this specification of the production function the labor quality decision is independent of decisions concerning L and K . All the firm must do regarding Q is to minimize cost per efficiency unit of labor, i.e. minimize W/a , which leads to (20). A more general specification assumes that an efficiency index, which is a function of Q , is also associated in (15) with K . We hope to report results with this more general approach in the future.

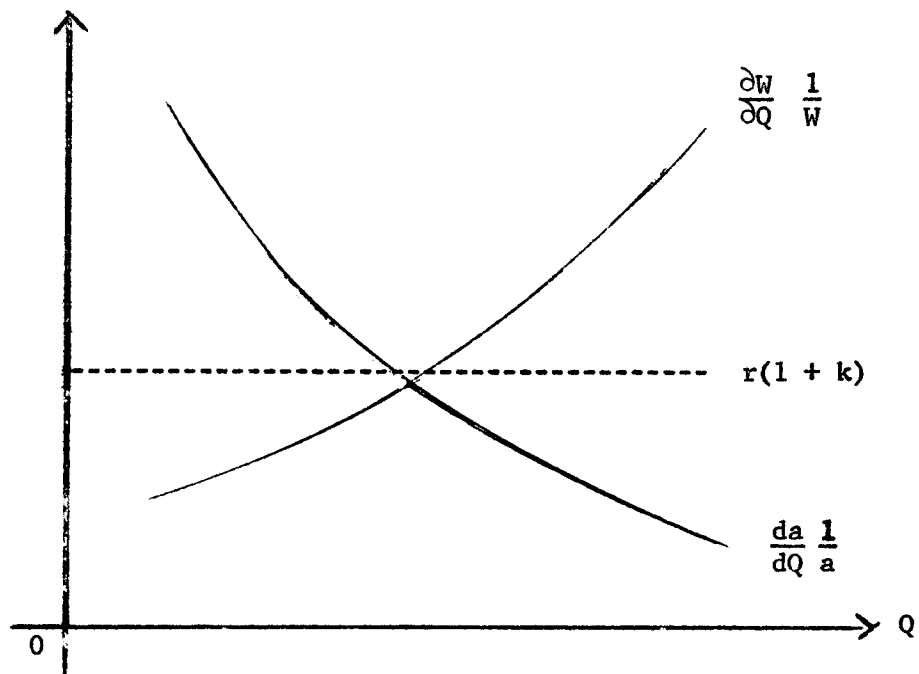


Figure 1

schedule. Suppose that in the typical firm labor quality has a higher marginal effect on labor efficiency the greater are the variables in the (row) vector R. Then the simplest specification of the labor efficiency function which is consistent with the above considerations is

$$a = \exp[b_1 Q + (R\beta)'Q - b_2 Q^2] ,$$

where β is a vector of parameters. The marginal condition (20) then requires that we set $\alpha_1 = b_1 + R\beta' - 2b_2 Q$, so that "the equilibrium" level of Q is given by

$$(21) \quad Q = \frac{b_1 - \alpha_1}{2b_2} + R\beta' \frac{1}{2b_2} .$$

This suggests a third estimating equation

$$(22) \quad E_i = \gamma_0 + R_i \gamma' + \epsilon_{3i} ,$$

where ϵ_{3i} is a disturbance term and γ is implicitly defined by (21). Since (22) contains no endogenous variables on the right hand side, the estimating equations (8), (12), and (22) are block recursive, and it is necessary to consider the last along with the first two only if $\text{COV}(\epsilon_{1i}, \epsilon_{3i}) \neq 0$ or $\text{COV}(\epsilon_{2i}, \epsilon_{3i}) \neq 0$. Although this is primarily an empirical issue, there are a priori reasons for supposing that at least $\text{COV}(\epsilon_{1i}, \epsilon_{3i}) \neq 0$. Intuitively, when unmeasured labor quality causes the deterministic part of (8) to under-predict $\ln W_i$, i.e. a positive fillip to ϵ_1 , we should expect the deterministic part of (22) to over-predict E_i , i.e. a negative fillip to ϵ_3 , because true labor quality will be greater than measured

labor quality. Hence, $\text{COV}(\epsilon_{1i}, \epsilon_{3i})$ is likely to be negative,^{26/} and it will be essential, as well as efficient, to consider all these equations jointly.

II. Empirical Results

Our discussion of empirical results is divided into three parts. First, we develop the specification of equations (8), (12), and (22) in terms of observable variables. Second, we estimate these equations for two-digit manufacturing industries with alternative consistent methods under various specifications about the endogeneity of unionism and labor quality and with the assumption that average hourly earnings are a reasonable approximation to the average wage rate in an industry. Third, we consider certain modifications of the specifications of the wage and unionism equations. Although we will return to this theme in Section III, it is useful to state at the outset that the results presented below are intended only as a preliminary investigation of the empirical relevance of the model presented in Section I. In particular, we do not consider them directly comparable to any other set of results hitherto produced.

^{26/} More formally, write (8) as $\ln W_i = \overline{\ln W_i} + u_i + \epsilon_{1i}'$, where $\overline{\ln W_i}$ is the deterministic component of $\ln W_i$ and ϵ_{1i}' is decomposed into an unmeasured quality difference (u_i) and other forces (ϵ_{1i}'). For simplicity let $Q_i = E_i + u_i$. Then (21) becomes $Q_i = E_i + u_i = \gamma_0 + R\gamma' + \epsilon_{3i}'$ so that (22) is $E_i = \gamma_0 + R\gamma' - u_i + \epsilon_{3i}'$. Then one component of $\text{COV}(\epsilon_{1i}', \epsilon_{3i}') = \text{COV}(\epsilon_{1i}' + u_i, \epsilon_{3i}' - u_i)$ is $-\sigma_u^2 < 0$.

A. Specification of the Model

The labor supply equation developed above posits that the wage in an industry depends upon the degree of unionization, the average educational attainment of the industry's work force (E_i), and an index of nonpecuniary rewards. As a measure of unionization we adopt estimates of the proportion of production workers who are employed in establishments covered by collective bargaining agreements (U_i).^{27/} One important aspect of inter-industry wage determination not discussed thus far is the problem of wage differentials due in part, at least, to discrimination. Quantitatively, the most important form of discrimination is perhaps that against females, and we would expect that the greater the proportion of female workers in an industry (F_i) the lower its average wage.^{28/} More formally, if we specify the estimating equation

$$(23) \quad \ln W_i = \alpha_0 + \alpha_1 E_i + \alpha_2 U_i + \alpha_3 F_i + \epsilon_{1i} ,$$

with $\alpha_1 > 0$ and $\alpha_3 < 0$,^{29/} as an approximation of (8), and if the ceteris paribus

^{27/} See the data appendix for the specific sources of all variables.

^{28/} Another important factor may be systematic male/female differences in the extent of on-the-job training. For more detail on this issue and other reasons for male/female earnings differences, see Henry Sanborn, "Pay Differences Between Men and Women," Industrial and Labor Relations Review, July 1964, pp. 534-50. This subject deserves much further research. In principle the proportion of Negro workers in the industry should also enter equation (23), since discrimination against Negroes presumably affects the average industry wage. Our preliminary results, however, paralleled those of Throop, op. cit., in failing to find statistical significance for such a variable.

^{29/} As has been demonstrated by Lewis, "Competition . . .," op. cit., pp. 186-87, the possibility of $\alpha_2 < 0$ cannot be ruled out a priori. It is conceivable that worker and management tastes for unionism are juxtaposed so that workers are willing to forfeit earnings in order to obtain the nonpecuniary benefits of unionism.

ratio of female to male wages is λ , then by an argument analogous to the derivation of the coefficient α_2 in (8), $\alpha_3 = \ln \lambda$ in (23). At the same time, however, it has been noted that a low value of F_1 may in large part be a proxy for the onerousness of production work which women see to avoid, since male production workers also receive lower wages in industries with a high value of F_1 .^{30/} Thus, F_1 may be a partial proxy for the variable A_1 in equation (8) which measures nonpecuniary returns. If this latter hypothesis were true we would tend to find that $\alpha_3 < \ln \lambda$, where λ is calculated directly from data on the personal earnings of women, because of the lower wages of men in industries with a high F_1 . A test of the null hypothesis $\alpha_3 = \ln \lambda$, which we will carry out below, may then be interpreted as a test of the "no nonpecuniary returns" hypothesis.

The model of trade union membership was developed in terms of the proportion of workers in all bargaining units who join a union, but the data on unionization which we shall use are in terms of the proportion of workers in an industry covered by collective bargaining contracts. This poses a difficulty of interpretation because the membership function was developed on the basis of the individual as a decision maker while the data are more applicable to a situation in which the bargaining unit is treated as a decision maker. In fact, because unions provide collective services which can usually be provided only if a majority of the workers in an

^{30/} See Sumner H. Slichter, "Notes on the Structure of Wages," Review of Economics and Statistics, February 1950, pp. 84-85. Since we treat F_1 as an exogenous variable in what follows, we are implicitly assuming that firms have little control over the conditions of work they offer.

establishment desire them, this difference should not be very important for what follows. Establishments which do not have a collective bargaining agreement should have very few union members, and establishments which do have a collective bargaining agreement should be very fully organized due to the resulting pressure from those workers who desire union services upon those who do not.^{31/}

Equation (12) also contains a variable Z_1 , which represents the potential cost to established unions of the provision of services to potential union members. Put the other way around, we wish to develop a proxy variable for the desirability and ease of organizing and servicing workers in an industry. It has long been maintained that an industry's concentration ratio (CON_1) is such a variable.^{32/} First, the substantial fixed costs

^{31/}This pressure has led to the popularity of the union shop and of dues checkoff schemes as methods for enforcing compulsory union membership. In fact, as of 1959 only about 20 percent of workers covered by collective bargaining agreements were not also covered by a union security provision. See Leo Troy, "Trade Union Membership, 1897-1962," The Review of Economics and Statistics, February 1965, pp. 93-113. Further development of the implications of treating unionism as a collective good is contained in Mancur Olson, The Logic of Collective Action: Public Goods and the Theory of Groups, Cambridge, Harvard University Press, 1965. Evidence to support the view in the text on the attitudes of union members toward the very small minority of non-union members found in unionized plants in manufacturing is contained in Seidman et al., op. cit., esp. chapters 4-5.

^{32/}See Joseph Skister, "The Logic of Union Growth," Journal of Political Economy, October 1953, pp. 413-33; Martin Segal, "The Relation Between Union Wage Impact and Market Structure," Quarterly Journal of Economics, February 1964, pp. 96-114; and Harold M. Levinson, "Unionism, Concentration, and Wage Changes: Toward a Unified Theory," Industrial and Labor Relations Review, January 1967, pp. 198-205. The last reference contains explicit discussion of this issue in the non-manufacturing sector.

associated with the provision of union services suggest that the financial and organizational effort spent per potential union member will be smaller in industries with large establishments. Second, in industries which are characterized by relatively free entry the costs of union services must include a continuous campaign to organize new firms. Third, the financial strength of unions is usually derived from holding on to workers once they are brought into the union, a practice which is more easily performed in the more stable, concentrated industries.^{33/} As a rough estimating equation we thus have

$$(24) \quad U_i = \beta_0 + \beta_1 \ln W_i + \beta_2 \text{CON}_i + \epsilon_{2i} ,$$

where we expect $\beta_1 > 0$ and $\beta_2 > 0$.

Finally, we must specify the variables in the vector R_i of equation (14) which determine E_i , the average quality of the industry's work force. The analysis of Section I-C implies that anything which increases $\frac{1}{a} \frac{da}{dE}$, the marginal proportionate response of efficiency to quality, will result in a higher equilibrium value of E . What these factors are is not entirely clear and their determination will require much further research, but we suggest the following two hypotheses. First, we suppose that the greater the variety of task and individual decision making involved in a particular job category the further to the right will be the position of the $\frac{da}{dE} \frac{1}{a}$ schedule.

^{33/} See the eminently practical discussion and documentation of these last two points in Jack Barbash, The Practice of Unionism, New York, 1956, esp. chapters II and III.

For production workers, therefore, some index of the skill requirements of an industry (SK) should have a positive effect on E .^{34/} If we then take $\frac{\partial W}{\partial E} \frac{1}{W} = \alpha_1$ from equation (23) and use the simplest log-linear polynomial approximation which is consistent with (22),^{35/} i.e.,

$$(25) \quad a = \exp[\lambda_1 E + \lambda_2 E \cdot SK - \lambda_3 E^2], \quad \lambda_j > 0,$$

the equilibrium value of E is

$$(26) \quad E = \frac{1}{2\lambda_3} [(\lambda_1 - \alpha_1) + \lambda_2 SK].$$

Other influences on the position of the $\frac{1}{a} \frac{da}{dE}$ schedule are subsumed in λ_1 . Second, it seems plausible that an ex post proxy for λ_1 for production workers might be the fraction of workers who reside in urban areas (URB). The reason for this is as follows: It is well known that people who live in urban areas are on average more highly educated than people who live in rural areas. This is due to the facts that better educated persons have a higher taste for urban life, schooling is more available in urban areas, and there are pressures for individuals in rural areas to leave school earlier to help on the farm. Consequently, if firms with high values of λ_1 tended to locate in rural areas with the same frequency as in urban areas, the price of highly educated labor would be driven up in rural areas relative to urban areas. The same thing the other way around

^{34/} The particular skill index used here is that developed by Throop, op. cit.; we are indebted to Professor Throop for making his data available to us.

^{35/} To be consistent with (22), we must have in (25) that $\frac{d^2 \ln a}{dE^2} < 0$ to satisfy second-order conditions and that $a = 1$ for $E = 0$, so that $F(aL, K)$ will equal $F(L, K)$ when $E = 0$.

is true for firms with low values of λ_1 . Since locational decisions are not made randomly, one would expect firms which can use high E more profitably to have located in urban areas and firms which cannot do so to have located in rural areas. Hence, URB is taken as an ex post proxy for λ_1 and we have the estimating equation

$$(27) \quad E_i = \gamma_0 + \gamma_1 SK_i + \gamma_2 URB_i + \epsilon_{3i} ,$$

where $\gamma_1, \gamma_2 > 0$.

B. Initial Empirical Results

The initial results of fitting equations (23) and (24) to the 1960 data for 19 manufacturing industries are contained in Table I. The sample is restricted to manufacturing industries both because of the availability of adequate data and because it seems most plausible that the requisite conditions for static equilibrium implicit in equations (23), (24), and (27) are most likely to be satisfied for this sector. In line with the discussion of Section I we present both the ordinary least squares estimates and the estimates obtained by a consistent technique, in this case two stage least squares (2SLS). For each equation we present the results in the order of increasing endogeneity discussed in Section I. Hence, equations (23a) and (24a) are the OLS estimates, (23b) and (24b) are the 2SLS estimates on the

Table 1
 Estimated Regression Coefficients for Equations (23) and (24) for
 2-Digit Manufacturing Industries, Average Hourly Earnings as Wage Rate, 1960^a

Equation Number	Estimation Procedure	Exogenous Variables	"Dependent" Variable	Log of Wage Rate	Labor Quality	Unionism	Proportion Female	Concentration	Constant	R ²	SE ^b
23a	OLS	F _i , U _i , F _i	ln W _i	1.0	.077 (.013)	.382 (.090)	-.473 (.082)		-.172 (.127)	.939	.053
23b	2SLS	E _i , F _i , CON _i	ln W _i	1.0	.090 (.021)	.176 (.267)	-.521 (.108)		-.160 (.145)	.921	.059
23c	2SLS	F _i , CON _i , URR _i , SK _i	ln W _i	1.0	.120 (.031)	.040 (.262)	-.495 (.123)		-.398 (.238)	.874	.076
24a	OLS	V _i , CON _i	U _i	.576 (.122)		1.0		.280 (.142)	.109 (.100)	.667	.098
24b	2SLS	E _i , F _i , CON _i	U _i	.534 (.131)		1.0		.290 (.143)	.139 (.106)	.664	.098
24c	2SLS	F _i , CON _i , URR _i , SK _i	U _i	.481 (.130)		1.0		.304 (.145)	.175 (.107)	.654	.100

^a See data appendix for sources.

^b Standard Error of the Estimate for the regression equation.

assumption that wages and unionism are jointly determined,^{36/} and (23c) and (24c) are the 2SLS estimates on the assumption that wages, unionism, and labor quality are jointly determined. Estimated asymptotic standard errors are enclosed in parentheses beneath the coefficients.

OLS equation (23a) suggests a highly significant unionism coefficient of .46^{37/} Although this estimate is somewhat higher than those obtained by the other investigators which we have noted, it seems likely that this is due primarily to the fact that the present estimate is restricted to production workers in manufacturing industries only.^{38/} If we now turn to 2SLS equation (23b), where wages and unionism are endogenous variables, the unionism coefficient drops to .19 and the regression coefficient on unionism is about two-thirds its standard error. Finally, in 2SLS equation (23c) the unionism coefficient has declined to .04, and the regression coefficient on unionism is a startling one-sixth of its standard error. In line with the decline in the regression coefficient on unionism as we move from (23a) to (23c) is the expected increase in the coefficient on years of schooling. Most investigators have obtained estimates of the average rate of return on

^{36/}Note that on this assumption equation (23) is exactly identified, so that (23b) also gives the limited information - least variance ratio (LI-LVR) estimates and indirect least squares estimates. A standard reference for these procedures is Arthur S. Goldberger, Econometric Theory, New York, 1964, ch. 7. We note in passing that the LI-LVR estimates analogous to (23c) and (24c) are so little different from those obtained with 2SLS that they are not reported.

^{37/}See footnote 2 for the method of going from α_2 to the unionism coefficient.

^{38/}In a recent study, Raimon and Stoikov estimated a unionism coefficient on a sample limited to manufacturing industries, and their estimate was .32. See Robert L. Raimon and Vladimir Stoikov, "The Effect of Blue-Collar Unionism on White-Collar Earnings," Industrial and Labor Relations Review, April 1969, pp. 358-74.

schooling, for the range of E in which we are interested, of between 11 and 18 percent.^{39/} OLS equation (23a) implies an average rate of return on the order of 8 percent, which by conventional standards is significantly below the expected interval for this parameter. 2SLS equation (23c), on the other hand, implies an average rate of return on the order of 12 percent, which is fully consistent with the results obtained by other investigators. The highly significant coefficient on F_1 implies a ratio of female to male wages of about .60. The actual ratio of female to male wages in manufacturing (λ) is about .67 so that $\hat{\alpha}_3 < \ln \lambda$ and the results favor the hypothesis that F_1 is to some extent a proxy for the nonpecuniary benefits associated with an industry, although the hypothesis fails of statistical significance ($t = -.8$) by a wide margin.

The estimated regression coefficients of equation (24) are significantly positive, and as one would expect the coefficient on $\ln W_1$ declines somewhat as we move from OLS equation (24a) to 2SLS equation (24c). The estimated elasticity of unionism with respect to the wage rate at the mean level of unionism is about .72, but it is, of course, not possible to identify any of the parameters of the demand and supply functions for unionism discussed in Section I-B. Finally, the OLS and 2SLS estimates of (27) will not differ since this equation contains no endogenous variables on the right hand side.

^{39/} The most elaborate study of rates of return to education in the U.S. using data for 1959 of which we are aware is Giora Hanoch, "An Economic Analysis of Earnings and Schooling," Journal of Human Resources, Summer 1967, pp. 310-29. But see also W. Lee Hansen, "Total and Private Rates of Return to Investment in Schooling," Journal of Political Economy, April 1963, Table 6, p. 138 for results for 1949 which are not dissimilar for these groups. Both of these studies use different methods and a different source of data to that used here, so that information from them may be considered truly independent.

The estimated equation is:

$$(27a) \quad E_i = 1.21 SK_i + 5.97 URB_i + 1.47 , \\ \quad \quad \quad (.36) \quad \quad (1.47) \quad \quad (1.55)$$

where $R^2 = .705$ and S.E.E. = $.703$.^{40/} Given these results, we tentatively conclude that the model developed in Section I is both consistent with the data and worthy of further consideration.

Although we do not wish to dwell on this point at length, it is worth reiterating that the properties of the 2SLS estimator used on equations (23) and (24) are not generally known for small samples. In fact, it is not

^{40/} The simple correlation coefficient between the 2SLS estimated disturbances of equations (23c) and (27a) is $-.60$. The substantial size of this correlation, of course, is what makes it essential to consider (27) when estimating (23). There are three additional considerations relevant to the model which are worth noting at this point. First, it has sometimes been argued that the variable URB should enter a labor supply equation like (5) because: (i) wages may be higher in urbanized areas to offset differentials in the cost of living or (ii) wages may be lower in urbanized areas because workers prefer to live in them. When URB is added to (23) its coefficient is never significant and its addition does not change any of the other estimated coefficients. This result parallels that of Throop, *op. cit.*, who does not find a significant effect on wages of a variable measuring city size. Second, a possible refutation of our industry location hypothesis, by which we rationalized the inclusion of URB in (27), is for the equilibrium rate of return to schooling to be higher for low E in urban than in rural areas and lower for high E in urban areas. This would suggest that firms do not locate according to availability of labor supply, and it can be tested by adding the cross-product term $E_i \times URB_i$ to (23), where we would expect it to have a negative coefficient. When this is done, the estimated coefficient of $E_i \times URB_i$ has a small positive value which is not significantly different from zero. The estimated rates of return to schooling, for example, are $.11$ for $URB = 0$ and $.12$ for $URB = 1$, while the estimated unionism coefficient is $.02$ under this specification. Finally, it might be argued that the presence of unionism would tend to lower the marginal efficiency of labor quality $\frac{da}{dE} \frac{1}{a}$ because, for example, of work rules introduced under collective bargaining. To test this hypothesis U_i was added to (27), where its estimated coefficient (and standard error) was $.69$ (and 2.21), clearly not significant by conventional standards. The other coefficients in (27) were unchanged by this modification.

altogether clear what differentiates a small from a large sample in the context of any particular stochastic model, such as that contained in (23), (24), and (27). In any event, there are available estimators for this model which are at least asymptotically more efficient than the 2SLS estimator. These three stage least squares (3SLS) estimates of the parameters are contained in Table 2.^{41/} As can readily be seen by comparison of equations (24d) and (27d) with equation (24c) and (27a), the 3SLS estimates of the parameters of the union membership and labor quality equations are little different from the corresponding 2SLS and OLS estimates. In the case of the wages function, however, the coefficient on years of schooling increases again and the regression coefficient of unionism actually becomes negative, although it clearly would not be judged significantly different from zero at conventional test levels, (the point estimate of α_2 implies that unions lower average hourly earnings by 10 percent). In sum, the 3SLS estimates do not alter our basic conclusion concerning the empirical relevance of this model.

C. Further Empirical Results

The preceding empirical results were based on a particular simple specification of the wage and unionism equations. It could be argued that the nature of the results -- in particular those concerning the effect of

^{41/} See Goldberger, op. cit., pp. 346-52, for the basic idea behind three-stage least squares estimation, and A. Zellner and H. Theil, "Three-Stage Least Squares: Simultaneous Estimation of Simultaneous Equations," Econometrica, January 1962, pp. 59-78 for further details and discussion.

Table 2
 Three Stage Least Squares Estimates of Regression Coefficients
 for Equations (23), (24), and (27),^a 1960

Equation Number	"Dependent" Variable	Coefficients of							
		Wage Rate ($\ln W_i$)	Labor Quality (E_i)	Unionism (U_i)	Proportion Female (F_i)	Concentration (CON_i)	Skill Index (SK_i)	Urbanization (URB_i)	Constant
23d	$\ln W_i$	1.0	.128 (.023)	-.087 (.179)	-.555 (.101)				-.379 (.188)
24d	U_i	.542 (.120)		1.0		.261 (.122)			.142 (.097)
27d	E_i		1.0				1.157 (.333)	6.21 (1.33)	1.50 (1.42)

^a Average hourly earnings used as wage rate. See data appendix for sources.

unions on relative wages -- is due to the particular specification. In this section we alter the specification in various ways to see if this is so.

First, some investigators have stressed the role of concentration in determining wage levels.^{42/} This suggests that CON_i should be included in the wage equation, and these results are reported in Table 3. The basic results of Table 1 are unaffected by the inclusion of CON_i . The coefficient on U_i goes from .426 and significant to .038 and insignificant as we move from the single to the three equation model. The equivalent of 23b is, of course, underidentified. The concentration variable is not significant in either of the equations.

Second, the estimates in II-b implicitly assume that all workers are drawn from the same geographical labor market. It is generally felt that labor market conditions, including attitudes toward unionism, are different in the South than in the rest of the country. To test this, we added a variable SOU_i which represents the proportion of workers in each industry who reside in the South. The corresponding estimates of (23) and (24) are shown in Table 4. First, 23g is the OLS wage equation with SOU_i added, and it is seen that wages in the South are estimated to be about 4.5 percent less than elsewhere, but this estimated difference is far from significant. 23h is the wage equation estimated by 2SLS on the assumption that the other equations, (24) and (27), hold. The coefficient on SOU_i is now slightly

^{42/} See, for example, Weiss, op. cit.

Table 3

Estimated Regression Coefficients for Equation (23)

with CON_i Added

Equation Number	Estimation Procedure	Exogenous Variables	Labor Quality	Unionism	Proportion Female	Concentration	Constant	R^2	SEE
23e	OLS	E_i, U_i, F_i, CON_i	.078 (.012)	.426 (.107)	-.451 (.083)	-.079 (.084)	-.188 (.127)	.944	.052
23f	2SLS	$E_i, U_i, F_i, CON_i, URB_i, SK_i$.120 (.033)	.038 (.343)	-.490 (.137)	.010 (.069)	-.396 (.240)	.876	.078

positive but very insignificantly so. The coefficient on unionism declines to a low, insignificant value, as it did in previous cases. The same effect is noted in (23i) in which the original wage equation is estimated, but using SOU_1 as an additional instrumental variable.

It is also possible, of course, that many other variables, including the other exogenous variables in (23), and (27), should enter the unionism equation (24). Since the introduction of these variables into (24) will not change the list of exogenous variables used in the 2SLS estimation of (23), the estimates of this equation will be unaffected so long as (23) remains identified. In this sense, the results of estimation of the wage equation (23) are completely insensitive to specification of the determinants of the extent of unionization. Nevertheless, equation (24) is of some interest in itself and equations (24e) and (24f) in Table 4 are a test of the hypothesis that for cultural, legal, or other reasons the extent of unionization is lower in the South than would be expected.^{43/} As can be seen from the table, the results suggest that U may be some 10 to 15 percentage points lower in the South than elsewhere, although neither of these coefficients is significantly different from zero. It seems unlikely that this would be the case for nonmanufacturing industries.

^{43/} In terms of the discussion in Section I-B, SOU_1 is expected to be an element of τ_1 . This was suggested by Reder, op. cit.

Table 4

Estimated Regression Coefficients for Equations (23)
and (24) with SOU_1 Added

Equation Number	Estimation Procedure	Exogenous Variables	"Dependent" Variable	Coefficients of								R^2	SEE
				$\ln W_i$	Labor Quality	Unionism	Proportion Female	Concentration	South	Constant			
23g	OLS	E_i, U_i, F_i, SOU_i	$\ln W_i$	1.0	.074 (.014)	.375 (.101)	-.471 (.084)		-.046 (.087)	-.123 (.168)	.942	.053	
23h	2SLS	$F_i, SOU_i, CON_i, URB_i, SK_i$	$\ln W_i$	1.0	.121 (.036)	.053 (.261)	-.484 (.128)		.016 (.140)	-.422 (.335)	.877	.077	
23i	2SLS	$F_i, SOU_i, CON_i, URB_i, SK_i$	$\ln W_i$	1.0	.119 (.029)	.053 (.249)	-.486 (.120)			.395 (.226)	.880	.074	
24e	OLS	W_i, CON_i, SOU_i, U_i	U_i	.527 (.148)		1.0		.297 (.147)	-.095 (.157)	.167 (.140)	.674	.100	
24f	2SLS	$F_i, SOU_i, CON_i, URB_i, SK_i$	U_i	.446 (.164)		1.0		.320 (.150)	-.142 (.163)	.235 (.153)	.668	.101	

III. Concluding Remarks

We have proposed a static three-equation model of relative wage determination by industry in which wages, unionism, and labor quality are considered as jointly determined variables. Introducing the endogeneity of unionism and labor quality into the wage determination process provides a model which is more closely grounded in traditional economic analysis and which we tentatively conclude is consistent with the data.

On an empirical level we find that allowing for the simultaneous determination of wages, unionism, and labor quality in estimation tends to produce an estimate of the equilibrium union/nonunion wage differential in the range of 0 to 20 percent, but this is never significantly different from zero. Given the qualitative and quantitative limitations of the data, we are prepared to say only that we are uncertain of the magnitude of the effect of unions on interindustry wage differences. We do conclude, however, that further research into the causes of relative wage differentials should explicitly consider the investigation of the simultaneous determination of wages and other relevant variables. In line with the above comments it is necessary to stress the following limitations inherent in our approach to this problem.

1. Both our theoretical and empirical analyses are concerned with static equilibrium positions. For example, the model proposed above does not deal with deviations from long run behavior due to the relatively recent organization of unions in an industry, and one would suppose that the full adjustment of labor quality to a change in relative wages would take many

years. In line with this point of view, the above empirical analysis does not imply that trade unions are ineffectual with respect to pressure on aggregate money wage changes.^{44/}

2. Our empirical analysis is restricted to manufacturing industries. This is in part because of data availability but also because differences in tastes for unionism by employers and employees -- which we hope can be ignored for the manufacturing sector -- are probably quite important in some of the non-manufacturing industries (like finance and the services). One might very well expect different results for an alternative set of industries.

3. Although the general specification of the model in Section I may be a plausible simplification of the static inter-industry wage determination process, there are a number of unresolved practical difficulties which must be faced in order to prepare the model for estimation. The costs of unionization and the supply of educational quality must be specified.^{45/} Needless

^{44/} Indeed, there is an accumulating body of evidence to the contrary. See, for example, Otto Eckstein and Thomas A. Wilson, "The Determination of Money Wages in American Industry," Quarterly Journal of Economics, August 1962, pp. 379-414; George L. Perry, "The Determinants of Wage Rate Changes and the Inflation-Unemployment Tradeoff for the United States," Review of Economic Studies, October 1964, pp. 287-308; and A. G. Hines, "Trade Unions and Wage Inflation in the United Kingdom 1893-1961," Review of Economic Studies, October 1964, pp. 221-52. For the exposition of a model of individual trade union behavior emphasizing the absolute change in rather than the structure of wages, see Orley Ashenfelter and George E. Johnson, "Bargaining Theory, Trade Unions, and Industrial Strike Activity," American Economic Review, March 1969, pp. 35-49.

^{45/} For example, the concentration ratio is used as a proxy in the empirical analysis for the costs of union organization in manufacturing industries. This variable might very well be unsatisfactory for this purpose in other industry groups. See Levinson, op. cit., on this issue.

to say, there are undoubtedly procedures which might be used in this context which are substantially different from those developed in this paper. Further, it might be argued that the proxy variables chosen are not unambiguous measures which can arbitrarily be excluded from, for example, the wages function; in which case that equation would be under-identified and not susceptible to estimation. The empirical economist cannot long remain sanguine on this problem, and only further a priori economic analysis can mitigate it.

Data Appendix:

For the regressions in Tables 1 and 2, W represents average hourly earnings for 19 two-digit manufacturing industries (two industries, ordinance and miscellaneous manufacturing, are excluded from the sample because of data unavailability) and was obtained from Employment and Earnings for the United States, 1909-67, B.L.S., Bull. 1312-5. The unionization ratio (U) data are those presented by H.M. Douty, "Collective Bargaining Coverage in Factory Employment, 1958," Monthly Labor Review, 83 (April 1960), pp. 345-49. F, the ratio of female to total employment, and URB, the ratio of employed workers residing in urban areas to total employment, are obtained from the "Industrial Characteristics" volume of the 1960 Census of Population, PC(2)-7F. E, the median years of schooling variable, is a weighted average (by F) of median years of schooling by industry for males and females, and these data were obtained from Table 21 of the same Census volume. Four-firm concentration ratios by industry, CON, are for 1954 and are from William G. Bowen, Wage Behavior in the Postwar Period, Industrial Relations Section, Princeton University, 1960, Table D-1. The skill index, SK, was provided by Professor A. Throop and "is calculated for [production workers in] any industry by weighting current median annual sex-occupational earnings throughout the economy by the current relative industry employments of corresponding sex-occupational groups." [See Throop, "The Union-Nonunion ...," p. 81.]