UNIONIZATION AND COST OF PRODUCTION: COMPENSATION, PRODUCTIVITY, AND FACTOR-USE EFFECTS

By Randall W. Eberts and Joe A. Stone

Randall W. Eberts is an assistant vice president and economist at the Federal Reserve Bank of Cleveland. Joe A. Stone is the W. E. Miner professor of economics at the University of Oregon. The authors wish to thank Steven Allen, Paul Bauer, Erica Groshen, and Kelly Eakin for suggestions and comments.

Working papers of the Federal Reserve Bank of Cleveland are preliminary materials circulated to stimulate discussion and critical comment. The views stated herein are those of the authors and not necessarily those of the Federal Reserve Bank of Cleveland or of the Board of Governors of the Federal Reserve System.

January 1987
ABSTRACT

This paper demonstrates that unionization can affect cost of production through increases in compensation, through shifts in technologies, and through deviations from the least-cost combination of inputs. The first two effects are familiar, but the third, the factor-use effect, is not. Identification of this effect resolves two important questions. One, when unionization compensation and productivity effects are found to be largely offsetting, why doesn't the labor intensity of production decline, as predicted, and why do employers continue to resist unionization? Two, why do employers complain that work and staffing rules reduce "productivity" when there is little theoretical or empirical reason to believe that this is so? The paper presents estimates of compensation, productivity, and factor-use effects to illustrate the answers to these questions.
I. Introduction

Two paradoxes confound analysis of the effect of unions on the cost of production. One is that productivity increases attributed to unions by Brown and Medoff (1978) and others' are often large enough to offset, or almost offset, cost increases from the union compensation premium, suggesting that the total effect of unionization is small for cost of production but large for employment.\(^2\) To borrow the words of Hirsch and Addison (1986), "This finding is, frankly, implausible...." Wessels (1985) agrees and notes, "...these results appear to conflict with other well-known evidence on unions. If unions did have such a substantial impact on labor productivity, they should reduce employment far more than has been commonly observed." Clark (1980), in fact, found that unionized cement plants are characterized by more labor-intensive production than nonunionized plants, whereas the compensation and productivity changes associated with cement unions lead one to expect the reverse.\(^3\) In addition, the small cost effect of unionization implied by comparisons of the productivity and compensation effects is inconsistent with both the intensity of employer opposition to unionization and direct estimates of the total effect.\(^4\)

The second paradox involves the work rules, employment-security protocols, and staffing requirements prevalent in union contracts. Des-
pite complaints by employers that these restrictive employment rules reduce productivity, there is little systematic evidence that this is so.\textsuperscript{5} Johnson (1986) has shown, in fact, that under a broad range of circumstances, unions and employers will not bargain employment restrictions that result in technical inefficiency because such bargains are Pareto-inferior.\textsuperscript{6} Thus, it is not surprising that most research suggests that this type of technical inefficiency is not a substantial source of increased cost. In fact, the overall effect of unions on productivity is typically positive. Why, then, do employers complain about work rules and employment restrictions? How, for example, can the absence of productivity effects for work rules be reconciled with evidence presented by Piore (1982) that work rules are at least as important as labor compensation in firm location decisions?

We resolve these two paradoxes by identifying three theoretical components of the union effect on cost of production. One is the familiar component related to the union/nonunion labor compensation differential. The other two represent separate components of the union effect on efficiency: (1) changes in "technical efficiency" (shifts in the production technology); and (2) changes in "allocative efficiency" (factor use that deviates from the least-cost combination of inputs).

Although the compensation and productivity effects are familiar and subject to wide investigation, the factor-use effect is not. There is, however, a strong a priori reason to believe that factor-use distortions are significant in unionized firms. A number of recent papers (including McDonald and Solow [1981], Clark [1984], Abowd [1985], Brown and Ashenfelter [1986], Eberts and Stone [1986b], Johnson [1986], and MacCurdy and Pencavel [1986]) demonstrate that Pareto-efficient collective bargaining
contracts normally require constraints on the level of employment. These constraints are necessary because compensation exceeds the marginal revenue product at every point on the contract curve except the competitive one.

Distinguishing the factor-use effect of unionization on cost of production is important not only because it resolves the two paradoxes, but also because it helps to clarify the interpretation of previous work on featherbedding and work rules in union contracts. Most previous studies of the effects of union work rules deal only with technical efficiency and the productivity effect, not with allocative efficiency, and the distinction is not always made clear. Furthermore, recent studies that do examine allocative efficiency (Allen [1986]) yield ambiguous results because they do not formally distinguish the compensation, productivity, and factor-use effects.

Our paper is organized as follows. In the next section, we derive a restricted minimum-cost function that incorporates union effects on cost through restrictions on factor use, as well as through changes in compensation and in the production technology. In Section III, we illustrate the three effects graphically, evaluate the significance of the factor-use effect in resolving the contradictions pointed out by Wessels and others, argue that previous attempts to isolate the factor-use effect are inappropriate, and present estimates of the three effects for public elementary schools. A final section summarizes our major conclusions.

II. Theoretical Framework

We begin with a standard problem of cost minimization, where the
firm seeks to minimize the cost of production subject to the level of input prices and constraints of output and production technology. Our firm seeks to minimize the following cost of production:

\[
C = wL + aA
\]

where \( C \) is the total cost of production, \( L \) is the quantity of labor employed, \( A \) is the quantity of an alternative factor employed, and \( w \) and \( a \) are the respective factor prices. For our purposes, it is useful to specify the following constraints faced by the firm:

\[
z_o = z(L,A,U)
\]

\[
w = w(U)
\]

\[
R = \frac{z_A}{z_L} \frac{w}{a} = R(U), \ R(0) = 1
\]

Equation (2) expresses the output of good \( z \) as a function of the level of labor \( (L) \), the alternative factor \( (A) \), and unionization \( (U) \). Unionization is expressed here as a continuous variable, consistent with the central role of bargaining power in the contract-curve literature. We assume that the production function is concave and twice differentiable.

Equation (3) is the function for the union compensation premium and is assumed to be quasi-concave and differentiable. Equation (4) is an implicit specification of the potential side constraint on employment imposed by union work rules, staffing requirements, and the like, where
$z_A$ and $z_L$ are the marginal products. We assume that the restriction is an increasing function of unionization, so that nonzero values of $U$ require deviations from the least-cost combination of inputs (which is obtained by equating the ratio of marginal products to the ratio of factor prices). In practice, the restriction is likely to be in the form of an inequality constraint, but is expressed as a strict equality here. We also abstract from the process that determined the restriction, taking it as given for the current cost-minimization problem. Other specifications of equation (4) are possible, but this specification is especially tractable and is suggested by the role of bargaining power in the contracting literature, where the excess of compensation over the value of the marginal product reflects bargaining power.

Normally, the minimum-cost function for a firm facing parametric input prices can be expressed simply as the sum of the respective products of each factor price and the constant-output demand for that factor. Union effects on cost of production are introduced through compensation and the production function. If the unionized firm is not strictly free to choose the least-cost combination of inputs, however, the constant-output factor-demand equations are conditional on the factor-use restriction. In this case, unionization affects cost through changes in factor use, as well as through changes in compensation and in the production technology. Thus, equations (1) through (4) yield the following restricted minimum-cost function:

$$C^* = w(U).L^*[z_o,U,w(U),a,R(U)] + a.A^*[z_o,U,w(U),a,R(U)]$$

where $C^*$ is the restricted minimum-cost function and $L^*$ and $A^*$ are the
restricted least-cost **factor-demand** equations. The expressions $w(.)$, $L^*(.)$, $R(.)$, and $A^*(.)$ are implicit functions. Equation (5) differs from the standard minimum-cost function because factor-use decisions are restricted by equation (4). Therefore, we refer to equation (5) as a restricted minimum-cost function.

The compensation, productivity, and factor-use effects are derived by differentiating equation (5) with respect to unionization and by arranging terms. The compensation effect is simply

\[(6) \quad w_u \cdot L^* \]

which is the product of the initial cost-minimizing level of labor services ($L^*$) and the marginal union compensation premium ($w_u$).

The productivity effect (a shift in production technology with the factor-use restriction held constant) is

\[(7) \quad w \cdot L^*_u + a \cdot A^*_u \]

which is the change in the input requirements (evaluated at original factor prices) for a fixed output ($z_0$).

The factor-use effect is

\[(8) \quad R_u \cdot (w \cdot L^*_R + a \cdot A^*_R) \]

which is the effect of the union-induced factor-use restriction on the input requirements for a fixed output (evaluated at the original factor prices). If unionization imposes no factor-use distortions, then $R_u$
is necessarily zero. \( L^*_{R} (A^*_{R}) \) is equal to the deviation of the restricted value of \( L^* (A^*) \) from the unrestricted least-cost value that would otherwise obtain.\(^8\) In reality, of course, work rules and staffing requirements may not impose an exact restriction on labor. It is important to note that a restriction on one factor generally distorts the use of other factors, so that analysis of labor use alone is insufficient.

If one is evaluating the total derivative of the restricted cost function at nonunion values (that is, at \( U \) equal to zero), then the three effects listed above exhaust the nonzero terms and can be separated uniquely. Two complications arise when this is not the case, that is, when \( U \) is not equal to zero. One is that the following term becomes nonzero:

\[
(9) \quad w_U(w_L A_w + a A^*_w)
\]

This expression is the product of the marginal union effect on compensation and the reciprocity relation of cost minimization (see Silberberg [1978]). With no factor-use restriction (which obtains at \( U \) equal to zero), this equation is zero because \( w_L A_w \) is equal to, but opposite in sign from, \( a A^*_w \) at the unrestricted least-cost combination of inputs. Because equation (9) is the result of an interaction between the compensation premium and factor-use distortions, it cannot be uniquely attributed to either.

The second complication also involves interactions. As the values of \( w, L^*, A^*, \) and \( R \) move away from the nonunion values, additional interactions arise between the compensation, productivity, and factor-use
effects. In the compensation effect, equation (6), the value of L* at any point will, in general, depend on existing compensation, productivity, and factor-use effects. In the productivity effect, equation (7), the value of w will depend on the existing compensation premium, and the values of L*, and A*, will, in general, depend on existing compensation, productivity, and factor-use effects. Similar interactions arise in the factor-use effect, equation (8). As with all interactions, attribution of any of the interactive parts of equations (7), (8), or (9) to any one of the component effects is arbitrary. Typically, however, interactions are dominated by the accumulation of the main effects, but do introduce a range of ambiguity. Like standard measures of income and substitution effects in consumption, measures of the compensation, productivity, and factor-use effects are path-dependent.

III. Graphic Illustration and Discussion

The compensation, productivity, and factor-use effects of unionization on cost of production (and their path-dependence) can be illustrated using figure 1, which expresses the relationship between cost of production and labor compensation (w) with output and other factor prices held constant and unionization treated as a binary variable. C^N is the minimum-cost function for nonunion firms, and w^N is the level of non-union compensation. C^R is the restricted minimum-cost function for union firms, which includes both the productivity and factor-use effects, and w^U is the level of union compensation. C^R in figure 1 is defined for a given deviation from equality of the ratio of marginal products and the ratio of factor prices, because unionization is held constant. C^U
Figure 1 Illustration of Compensation, Productivity and Factor-Use Effects
is the unrestricted minimum-cost function for union firms that would obtain in the absence of any restrictions on factor use. This function includes the productivity effect but excludes the factor-use effect. Relative positions of the three functions are arbitrary, and the curvature of each function reflects the corresponding elasticity of substitution.

Taking one path, we can measure the compensation effect as the movement from point A to point B along the nonunion cost function, the productivity effect as the shift from point B to point D, and the factor-use effect as the shift from point D to point F. Taking a different path, however, we could measure the productivity effect as the shift from point A to point C, the compensation effect as the movement from point C to point D along the unrestricted union cost function, and the factor-use effect again as the shift from point D to point F. This path will in fact coincide with the first path if the technology shift from $c^N$ to $c^U$ is neutral, as is suggested by much of the empirical work in this area. In this case, $AC$ is proportionately the same as $BD$, and $CD$ is proportionately the same as $AB$. Other paths are obviously possible.

Resolving the Paradoxes

Figure 1 also illustrates how the factor-use effect resolves the paradox raised by comparisons of the compensation and productivity effects. If one ignores the factor-use effect, significant increases in compensation and roughly offsetting increases in technical efficiency imply only a small change in total cost and sharply lower labor intensities in union firms. In fact, however, direct estimates tend to show
significant cost increases and similar, or even higher, labor intensities. The factor-use effect reconciles this apparent paradox. If we measure the compensation effect as AB and the productivity effect as BD, then the total cost effect is quite small in figure 1, and employment will be sharply lower. Adding the factor-use effect, however, increases costs by DF by preventing the firm from using the least-cost combination of factors. Employment will also be higher than otherwise, obscuring the compensation and productivity effects on employment.

The second paradox is also resolved by the factor-use effect. Even if work rules do not generally reduce technical efficiency, they can still increase cost by preventing the employer from using the least-cost combination of inputs (decreasing allocative efficiency). Employers may complain about work rules and employment restrictions not because they reduce technical efficiency, but because they reduce allocative efficiency.

Estimating the Factor-use Effect

At least one previous paper (Allen [1986]) has recognized the effect of union-induced factor distortions on cost. An estimate of the factor-use effect for office construction is obtained by first estimating a minimum-cost function for nonunion firms and a restricted minimum-cost function for union firms, and then by subtracting the estimated compensation effect for nonunion firms (AB in dollars as a percent of costs) from that for union firms (EF in dollars as a percent of costs), which is assumed to reflect both the compensation and factor-use effects and nothing else. As one can see in figure 1, however, the difference EF minus AB is not a general measure of the factor-use effect. It captures
only that portion of the factor-use effect that arises from differences in the substitution parameters between $c^N$ and $c^R$; the shift in the average level of $c^R$ from $c^U$ is not captured. Moreover, the substitution parameters of $c^R$ depend on the productivity effect as well as the factor-use effect.

The difference $EF - AB$ would measure the factor-use effect if the technology shift from $c^N$ to $c^U$ were neutral and if the factor-use effect were zero at $w^N$ for union firms. Under these sufficient conditions (and using proportionate dollar measures), $EF - CD$ equals $DF$, $AB$ equals $CD$, and therefore $EF - AB$ equals $DF$, which is a measure of the factor-use effect.

In general, an estimate of the factor-use effect requires estimates of all three cost functions: $c^N$, $c^R$, and $c^U$. The nonunion minimum-cost function ($c^N$) and the restricted union minimum-cost function ($c^R$) can be estimated directly. The unrestricted union minimum-cost function ($c^U$), however, is not directly observed and must be constructed from estimates of the union production function via duality.\footnote{Although the data requirements for disentangling the compensation, productivity, and factor-use effects are extensive, they are met for at least one sector -- public elementary schools. Assembling evidence from several different but related studies of public schools, we are able to obtain estimates of the compensation, productivity, and factor-use effects. The compensation effect for teacher unions ($AB$ in figure 1) is calculated by multiplying a mid-range estimate of the union wage premium (16.5\% from Baugh and Stone [1982]) by the teacher's share of total cost in nonunion schools (about 60\% from Eberts and Stone [1986a]), which yields an estimate of 10\% for the compensation effect on total cost.}
The productivity effect on cost (BD in figure 1) is calculated from direct estimates of union and nonunion production technologies for individual students presented by Eberts and Stone (forthcoming). Measured across all students, the average union productivity difference is 3%, which results in a productivity effect on cost of -3%.

The factor-use effect on cost (DF in figure 1) is calculated by subtracting the estimates of the compensation and productivity effects (AB and BD, respectively, in proportionate terms) from an estimate of the overall cost difference (the difference F minus A in proportionate terms). The total effect on cost (15%) is obtained from a reduced-form specification of the effect of unionization on cost presented in Eberts and Stone (1986a). Subtracting the compensation and productivity effects from the overall effect yields a residual estimate of about 8% for the factor-use effects, which is almost as large as the compensation effect.

The significance of factor-use distortion in public schools is also documented by Eberts (1984). With factors other than unionization and compensation held constant, they find that average class size is significantly lower in unionized districts. Furthermore, variations within the union sector in individual contract items dealing with class-size limitations and reduction-in-force procedures are associated with variations in both total cost and the cost share for teachers.

IV. Concluding Remarks

The effect of unionization on cost of production consists of three effects: the compensation effect (arising from the union compensation premium), the productivity effect (arising from technology shifts due
perhaps to greater cooperation between management and workers, reductions in turnover, and increased worker morale), and the factor-use effect (arising from deviations from the least-cost combination of inputs). The compensation and productivity effects are familiar, but the factor-use effect is not. This effect resolves two existing paradoxes: (1) productivity and compensation effects of unionization typically imply a small net effect on cost of production, but employer resistance to unionization and direct evidence on cost suggest a large effect; and (2) employers complain that work rules and employment restrictions reduce "productivity," but there is little systematic evidence that this is so.

The factor-use effect explains both paradoxes. For collective bargaining to avoid increasing costs, it is not enough that the productivity effect offset the compensation effect; it must also offset the factor-use effect. Furthermore, work rules and employment restrictions can increase cost without reducing technical productivity by requiring the employer to deviate from the least-cost combination of inputs.

Recent work in the contracting literature suggests that factor-use restrictions are required to enforce Pareto-efficient contracts because employment is "excessive" at all points on the contract curve except the competitive one. There is reason to believe, therefore, that factor-use distortions are commonplace in unionized firms. Work rules and employment restrictions have been widely examined for their effects on technical efficiency (shifts in the production technology), but efficient contracting, estimates of the effect on technical efficiency, and our analysis here suggest that the primary effect of these restrictions is instead on allocative efficiency (distortions from the least-cost combination of inputs).
Footnotes

1. Other studies include Clark (1980), Allen (1984), and Eberts and Stone (forthcoming).

2. Not all estimates of the productivity effect of unions are significantly positive. Clark (1984), Ehrenberg, Sherman, and Schwarz (1983), and Noam (1983), for example, report either negative or insignificant productivity effects.

3. A similar increase in labor intensity is reported by Eberts (1984) for public schools. Clark (1984) reports little difference in factor intensities in manufacturing, which is still inconsistent with the prediction of a sharp decline in labor intensity.

4. Eberts and Stone (1986a), for example, provide evidence that the overall cost effect for public schools is inconsistent with the sum of the estimated productivity and compensation effects.

5. Most of the studies that do present evidence interpreted as the effect of work rules on productivity do not distinguish between technical and allocative efficiency. An exception is Ichniowski (1984).

6. Technically inefficient work rules or staffing arrangements will be bargained only under two conditions: (1) there must be no technology available that would be "appropriate" for the level of nonunion wages; and (2) the union must place a much greater weight on employment than on wages (Johnson [1986]).

7. The contract curve is bounded by the labor demand and labor supply curves. Thus, at one extreme the contract curve may coincide with labor demand. In this case, no employment restrictions are necessary to
enforce the contract.

8. This statement is correct when the deviations are evaluated at the unrestricted values of \( L^* \) and \( A^* \) (\( R \) equal to 1). Elsewhere (\( R \) greater than 1), \( L^*_R \) and \( A^*_R \) are equal to the change in the deviation of the restricted value from the unrestricted value that would otherwise obtain.

9. Allen scales the percentages relative to union costs at point E in figure 1. The calculations are also based solely on changes in labor costs.

10. Alternatively, the factor-use effect could also be estimated using parametric techniques employed by Lau and Yotopoulos (1971), Toda (1976), Lovell and Sickles (1983), and Atkinson and Halvorsen (1984) to estimate allocative inefficiency.
References


