Wage Bargaining and Employment

By IAN M. MCDONALD AND ROBERT M. SOLOW*

One of the perennial problems of business cycle theory has been the search for a convincing empirical description and theoretical explanation of the behavior of wage rates during fluctuations in output and employment. Even the empirical question is hardly settled, although the most recent careful study (see P. T. Geary and John Kennan) confirms the prevailing view that real-wage movements are more or less independent of the business cycle. There are really two subquestions here. The first presumes that nominal wage stickiness is the main route by which nominal disturbances have real macroeconomic effects, and asks why nominal wages should be sticky. The second focuses on real wages, and asks why fluctuations in the demand for labor should so often lead to large changes in employment and small, unsystematic, changes in the real wage.

We address only the second of these subquestions. We do so in the context of explicit bargaining over wages and employment by a trade union and a firm or group of firms, though one could hope that the results might apply loosely even where an informally organized labor pool bargains implicitly with one or more long-time employers. We do not harbor the illusion that trade unions are the only important source of wage stickiness. There are other plausible (and implausible) stories. Some, like this one, rest partially on optimizing decisions; others do not.

The impulse to this study was macroeconomic, but our focus is on a single employer and a single labor pool. Our methods, and therefore our conclusions, are entirely partial equilibrium. If the short-run mobility of labor is slight, and if fluctuations in real aggregate demand affect many sectors synchronously, then perhaps the mechanism we uncover here could be important in the business cycle context. But the work of embedding it in a complete macroeconomic model remains to be done.

We begin with a model in which the union is a simple monopolist, setting the wage rate unilaterally to maximize the expected or total utility of its members, and allowing the employer complete discretion over employment. We then consider a more complex institutional setup in which the union and the firm are supposed to bargain over both wage and employment and reach an outcome efficient for them both. (The monopoly outcome is not efficient, for the traditional reason.) There is, of course, a whole range of efficient bargains. A complete theory must single out one of them, but there is unlikely ever to be universal agreement on the right way to do so. Our approach is simply to try out several simple conventions and several formal solutions to the bargaining problem. We provide a framework within which they are all seen to bear a family resemblance to one another. Moreover, there is a certain assumption which makes all the proposed solutions share an important characteristic: the effects of a downswing or upswing in final demand on the negotiated outcome can be decomposed into two steps which reinforce each other with respect to employment and offset each other with respect to the wage. So it would not be surprising to find large fluctuations in employment and small unsystematic fluctuations in real wages during business cycles.

The key assumption is that product-market conditions are more sensitive to the business cycle than the reservation wage is. This would be the case, for instance, if (a) nonmarket opportunities including unemployment insurance benefits, which are not cyclically vulnerable, play an important role in the determination of the reservation wage, and/or (b) interemployer mobility is so limited that outside market opportunities figure only slightly in workers' calculations.

*University of Melbourne and Massachusetts Institute of Technology, respectively.

to different degrees by aggregate fluctuations. For this reason alone—as a referee has pointed out—one might expect the relevant reservation wage, and even its average across industries, to vary systematically during business cycles. Nevertheless, to the extent that industrial and occupational mobility is limited in the time span relevant to business cycles, we believe our story retains plausibility.

The best developed analytical approach to this problem is the theory of implicit contracts (surveyed by Costas Azariadis). In that literature, a contract is a long-term agreement in the sense that the economic environment will change in an only probabilistically known way during the life of the contract. On the reasonable assumption that the firm is less risk averse than its employees, the typical outcome is that an efficient contract will be wage stabilizing and (unless special features are introduced) employment stabilizing as well. In our approach, by contrast, the wage-employment bargain is struck after economic conditions in the firm's product market and in the surrounding labor market are known. It is a short-term or one-shot contract. Risk enters only in a trivial sense: if the contract calls for a fraction of the union's homogeneous membership to be unemployed, the unlucky ones are chosen at random.

In real life, negotiated contracts are usually long term. But they do not specify employment, which typically is left to the discretion of the employer; in consequence, employment fluctuates a lot. Our reconciliation of the stylized facts, the theory of implicit contracts, and our own theory goes like this: if a series of short contracts would lead as our model suggests to wide variation in employment and fairly stable wages, then the same outcome might reasonably well be achieved by a long-term contract in which a stable wage is specified but the level of employment is chosen at will by the firm. (We do need a general restraint on employers of a sort that could be accomplished by "featherbedding" work rules.)

I. A Simple Monopoly Union

The simplest interesting noncompetitive institutional setup is that of a monopoly union which can set the wage unilaterally. The employer (or employers) then chooses the volume of employment. Most collective bargaining agreements do give the employer discretion over the volume of employment. Why this should be so is an interesting question (see Robert Hall and David Lilien). But it is a rare trade union that literally controls the wage and we take up more complicated bargaining arrangements later. The simple monopoly case has been analyzed before, of course (see, for example, Allan Cartter), and we have only a few novelties to add. We use this analysis mainly as a vehicle to introduce concepts, establish notation, and draw some diagrams.

The firm is characterized by a revenue function $R(L)$ giving sales proceeds as a function of employment. If the firm were a price taker in its product market, $R(L)$ would be simply $pF(L)$ where p is the parametric product price and $F(L)$ is the production function relating employment to output. We assume, as usual, that $R(0)=0$ and $R(L)$ is concave, with marginal revenue eventually becoming very small or even negative. Profit is then $R(L)-wL$.[1] If the firm is a profit maximizer, it is indifferent among (w, L) combinations that leave $R(L)-wL$ constant. These isoprofit curves in the (w, L) plane serve as indifference curves for the firm. The slope of an isoprofit curve through (w, L) is $dw/dL=(R'(L)-w)/L$. For any L, isoprofit curves have positive slope until w reaches $R'(L)$, then negative. For higher L, the switch occurs at a lower w, so the firm's indifference map is as shown in Figure 1. For any L, a smaller w creates a bigger profit, so lower isoprofit curves are better for the firm.

Let the union quote a wage w_1. The firm then seeks the lowest indifference curve that touches the horizontal line at height w_1. That is to say, it solves $R'(L_1)-w_1=0$: marginal

[1] Product price and wage rate are to be thought of as deflated by a general price index.

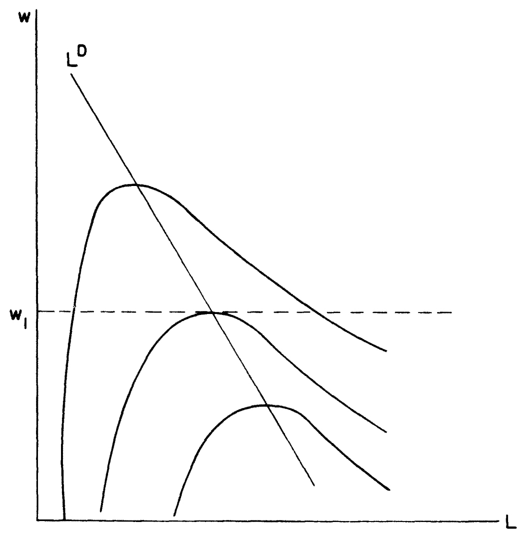

FIGURE 1

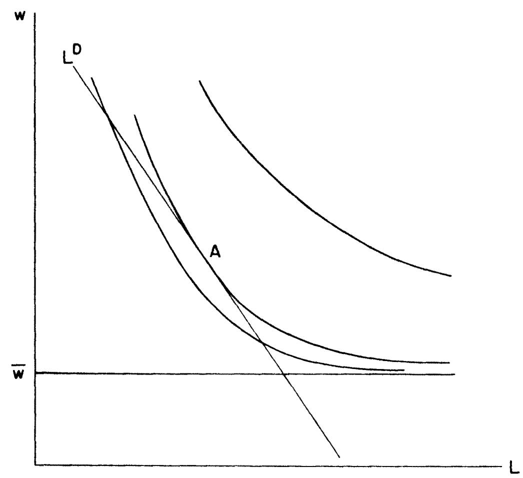

FIGURE 2

revenue product of labor equals the wage. In other words, the firm's demand curve for labor is the locus of maximum points of the indifference curves in the (w, L) plane. It is downward sloping, by virtue of the concavity of $R(L)$.

The union can achieve any point along the firm's demand curve. What is the union's objective? That is an old question in labor economics. We choose a particular answer and use it throughout. Suppose the union has N members, all alike. If L of them are employed, each member has probability L/N of having a job and achieving a level of utility $U(w)-D$ and probability $1-(L/N)$ of not being employed by the firm, where D is the fixed additive disutility of holding a job.[2] If not employed by the firm a worker achieves a level of utility $U(w_u)$, where w_u can be thought of, for short, as an unemployment compensation benefit, but should really include all the other contributions to the standard of living that would not be received if the worker were employed by the bargaining firm. $U(x)$ is the standard sort of concave utility function.

[2]We ignore—by choice—the possibility that workers are free to choose the hours and intensity of work.

The expected utility of a union member is therefore $N^{-1}\{L(U(w)-D)+(N-L)U(w_u)\}$, which can be written as $U(w_u)+N^{-1}L\{U(w)-D-U(w_u)\}$. Since w_u and N are treated as data for the purpose of union wage setting, we can set $D+U(w_u)=\bar{U}$ and summarize by saying that the union wishes to maximize $L(U(w)-\bar{U})$. The logic of this is that $L(U-\bar{U})$ is the membership's aggregate gain from employment, over and above the income w_u that every member starts with. The union's indifference map is derived from $L(U(w)-\bar{U})=\text{constant}$; the indifference curves have the usual downward-sloping convex shape in the (L, w) plane. They have the special property that they are all asymptotic to the horizontal at $w=\bar{w}$, where $\bar{w}$ is derived from $U(\bar{w})=\bar{U}$. This is shown in Figure 2.

The best wage for the union to set is determined in the obvious way by the tangency of an indifference curve with the employer's labor-demand curve as shown in Figure 2. Mathematically, this amounts to finding the maximum of $L(U(w)-\bar{U})$ with respect to L and w, subject to the constraint $R'(L)-w=0$. We can write down the first-order condition directly by equating the slope of the indifference curve through (w, L) (i.e., $-(U-\bar{U})/LU'$) to the slope of the (inverse) demand function (i.e., $R''(L)$). Since $w=R'(L)$ at any eligible point, the first-order condition can be written as

$$(1)\quad -LR''(L)/R'(L) = \left(U(w)-\bar{U}\right)/wU'(w).$$

The left-hand side is the reciprocal of the wage elasticity of the demand for labor, taken positively; the right-hand side is the reciprocal of the elasticity of the gain from employment $(U-\bar{U})$ with respect to the wage. So the condition is that the two elasticities should be equal. (There is a second-order condition that we assume to be satisfied.)

What is the nature of wage behavior implied by this model? A change in demand conditions will affect the union's wage decision via two routes, the elasticity of demand for labor and $\bar{w}$. We consider them in order.

Solve $w=R'(L)$ to give the demand function in direct form, and insert a parameter B (for business cycle), with the convention that an increase in B increases the demand for labor at any wage. Thus the demand for labor is $L=G(w,B)$. As B rises and falls, how is the effect divided between changes in w and changes in L? Consider the first-order condition (1) written as

$$(2)\quad wG_w(w,B)/G(w,B) = wU'(w)/\left(U(w)-\bar{U}\right).$$

The cyclical sensitivity of the wage clearly depends on the extent to which changes in B affect the elasticity of demand for labor at any given wage. For instance, if the demand function shifts isoelastically—that is, the demand for labor falls in a recession, but with its elasticity unchanged at each wage—then we can always write $G(w,B)=BG(w)$, and it is obvious that (2) does not depend on B at all. In that case, the wage will be rigid during business cycles and fluctuations will fall entirely on employment. One can easily imagine cases in which the monopoly wage will move countercyclically, or procyclically for that matter, thus diminishing or magnifying the accompanying fluctuations in employment.

The other way in which the level of aggregate activity can affect the desired wage is through $\bar{w}$, which is composed of several elements. Some of these elements, such as unemployment benefits, the value of leisure, the value of working around the house, net gains from illegal activities, would appear to be affected very little, if at all, by aggregate conditions. (Unemployment benefits are sometimes raised in recession, imparting an upward effect on $\bar{w}$ and thus w.) But the other major element in $\bar{w}$ is the expected value of alternative employment opportunities and this should have a strong procyclical fluctuation through changes in the probability of finding alternative jobs and in their wages. The effect this has on the wage rate will depend on just how important a component of $\bar{w}$ it is. If job mobility is low, and/or if changes in wage rates in other jobs are small, then the effect of changes in alternative job opportunities will be slight.

We conclude this section with a canonical example. Let f be a constant elasticity of demand for labor, however generated. Take $U(w)=w^b/b$, where b is less than one, but may be negative. Then (1) yields $w/\bar{w}=(1-b/f)^{-1/b}$. Thus the monopoly wage depends negatively on the elasticity of labor demand and negatively on the risk-aversion parameter $(1-b)$. Intuitively this is how it should be. For example, if $f=\frac{1}{2}$ and $(1-b)=3$, then the monopoly wage is $(5)^{1/2}$ times the "minimum supply price" $\bar{w}$. If f is as low as $\frac{1}{4}$, $w=3\bar{w}$. If $f=\frac{1}{2}$ and $(1-b)=2, w=3\bar{w}$. If $f=\frac{1}{4}$, $(1-b)=2$ then $w=5\bar{w}$. These low values of f are in accord with econometric results. Notice that if they are combined with positive values of $(1-b)$ less than 1 the outcome is much less "realistic": thus $f=\frac{1}{2}$ and $1-b=2/3$ implies $w=27\bar{w}$.

II. Efficient Bargains

The model of wage determination just described is even more like simple product-market monopoly than it looks. The difference in appearance arises because the monopolist, in this case the union, maximizes a utility function and not profits. It is not surprising, then, that the wage-employment outcome shown at point A in Figure 2 is not efficient. There are wage-employment points at which both parties are better off. This is

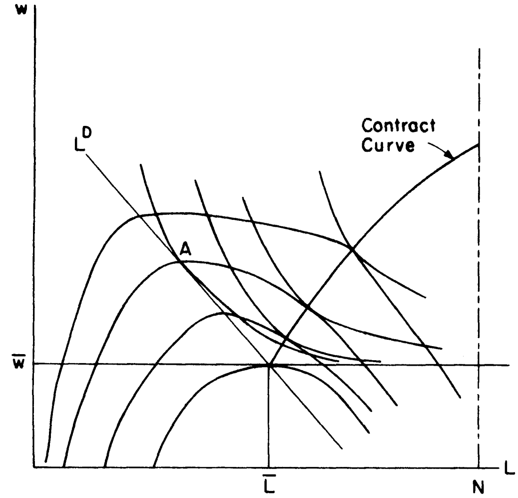

FIGURE 3

easily seen in Figure 3. The constant-profit curve passing through A is by construction horizontal at A. Therefore it cuts the downward-sloping indifference curve through A. The region to the southeast of A, between the isoprofit curve and the indifference curve, is the set of outcomes Pareto superior to A. The monopoly wage is too high and employment too low. Obviously efficient bargains are points of tangency between an isoprofit curve and an indifference curve. We call the locus of such points the contract curve; in this context that is the *mot juste*. An example is shown in Figure 3.

More complicated institutional arrangements are necessary for the achievement of efficient bargains. In particular, the union has to exercise some sort of influence over the level of employment, in contrast to the simple case where the level of employment is set unilaterally by the employer. Since the objective is to increase employment beyond the level given by the labor demand schedule, manning agreements or "featherbedding" are likely to be adopted. If it is impractical to specify the level of employment in the contract, an efficient outcome may be approximately achievable if the contract specifies the number of workers per machine, or some other similar rule, and leaves the overall aggregate to the discretion of the employer.

The contract curve is characterized by equality of the slopes of a union indifference curve and an isoprofit curve. This condition yields the equation

$$(3)\quad (U(w)-U(\bar{w}))/U'(w)=w-R'(L).$$

The first thing to notice is that the contract curve intersects the firm's labor demand curve at $w=\bar{w}$, because the right-hand side of (3) vanishes along the demand curve, and the left-hand side[3] is zero only at $\bar{w}$. The point $(\bar{w}, \bar{L})$ is actually the competitive outcome for this model. If there were no union and $\bar{U}$ were the level of utility attainable elsewhere in the economy,[4] then $\bar{w}$ would be the given supply price of labor to the employer, who would maximize profits at $\bar{L}$.

The slope of the contract curve is, by differentiation of (3),

$$dw/dL=-U'(w)R''(L) \Big/ \big(U''(w)(R'(L)-w)\big).$$

Thus the contract curve is momentarily vertical at $(\bar{w}, \bar{L})$, and positively sloped elsewhere (because, from (3) $w \geqq \bar{w}$ implies $w \geqq R'(L)$).[5] No bargain can be struck with $w<\bar{w}$ so the contract curve does not extend below $\bar{w}$. If we take the total membership of the union as a given number N, then the contract curve rises to the northeast until it reaches the vertical at N, where it ends. The effective part of the contract curve might end earlier if there is an $L<N$ at which the firm's operating profit becomes small enough to induce it to shut down.

Everywhere along the contract curve, except at $(\bar{w}, \bar{L})$, the wage exceeds the marginal revenue product of employment. The firm is

[3] The expression on the left comes up frequently in the theory of contracts. If expected utility is $pU(w)+(1-p)U(\bar{w})$, then the left-hand side of (3) is the marginal rate of substitution between p and w evaluated at $p=1$ (full employment).

[4] That would be the case, for instance, if there were no unemployment compensation, but a large supply of jobs at wage $\bar{w}$.

[5] If the marginal utility of income were constant then the contract curve would be vertical, as in Hall and Lilien.

thus being induced, presumably by an all-or-none offer, to employ more workers than it would like at the agreed-upon wage. This is the insight that led to Wassily Leontief's pioneering paper. An even stronger statement is true: all along the contract curve, except at $(\bar{w}, \bar{L})$, the marginal revenue product of employment falls short of $\bar{w}$. If one thinks of $\bar{w}$ as the true supply price of labor to the employer or industry, then this is a strong reminder that the bargains along the contract curve are efficient only from the point of view of the employer and the fixed membership of the union.

To see how the contract curve is affected by changes in the economic environment, we rewrite the revenue function as $R(L, B)$, and assume that R_B and R_{LB} are both positive: prosperity increases total revenue and the marginal revenue product of labor at any level of employment. Then (3) becomes

$$(3') \quad (U(w)-U(\bar{w}))/U'(w) = w - R_L(L, B).$$

If we now differentiate (3′) with respect to B, holding L constant, we find

$$\partial w/\partial B = R_{LB}(L, B)U'(w)^2 /((U(w)-U(\bar{w}))U''(w)) < 0,$$

and similarly

$$\partial w/\partial \bar{w} = -U'(w)U'(\bar{w}) /((U(w)-U(\bar{w}))U''(w)) > 0.$$

Thus an increase in B (an improvement in the firm's product market) shifts the contract curve to the right. It starts at $\bar{w}$ on the new demand curve; the negotiated level of employment is higher at any wage in an efficient bargain. An increase in $\bar{w}$ (an improvement in the economy-wide alternatives open to workers) has a different effect. If the new value is $\bar{\bar{w}}$, the new contract curve begins at $(\bar{\bar{w}}, \bar{\bar{L}})$, where $\bar{\bar{L}}$ comes from the labor-demand curve. Hence the starting point of the new contract curve is shifted to the NW, and the new contract curve lies everywhere above the old one. If the external labor market improves, efficient bargains have a higher wage at any level of employment. (See Figure 4.)

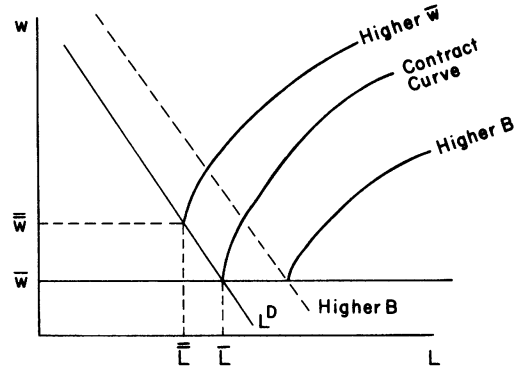

FIGURE 4

Generally speaking, simultaneous improvements in the labor market and the product market produce offsetting effects on the contract curve. Since B and $\bar{w}$ are not easily commensurable, it is hard to know how to model a generalized upswing or downswing in the economy. When we need to do so, we will tentatively assume that the B response outweighs the $\bar{w}$ response. For a "typical" labor pool, therefore, the contract curve shifts to the SE when economic conditions improve and to the NW when they deteriorate.

The contract curve probably has some approximate descriptive value. Even so, there is no generally acceptable solution concept that singles out a point on it as a likely outcome. Before we analyze some of the simpler possibilities, we digress to consider a slightly different institutional setup, as a useful conceptual exercise.

III. The Union as a Commune: A Digression

Everywhere along a positively sloped contract curve, the marginal revenue product of labor is less than $\bar{w}$, the supply price of labor. Efficiency implies starkly excessive employment. Why? The answer appears to lie in the fact that it is *ex post* more attractive to be employed than to be unemployed. To see this, let us change the rules and imagine the union acting as a family or commune, pooling all earnings and redistributing income from its employed to its unemployed members, so that they all have equal utility. Specifically, suppose the union pays out y_e to each of its employed members and y_u to each of its unemployed members. These payments are connected by

$$U(y_u)=U(y_e)-D; \tag{4}$$

the employed worker is compensated for the disutility of work. (For any old-timers who remember the Art Young cartoon: "Me slaving over this hot stove and you working in a nice cool sewer," it is perfectly all right to think of D as negative.) The union can only pay out what its members pay in, so there is a second constraint

$$Ly_e+(N-L)y_u=Lw+(N-L)w_u; \tag{5}$$

the employed contribute the negotiated wage and the others their unemployment benefits. Since everyone is equally well off *ex post*, the aggregate utility function is simply $NU(y_u)$. Here (4) and (5) can be solved for y_e and y_u as functions of w and L, so that the collective utility function can be thought of as a function of the negotiated outcome as before.

Straightforward differentiation of (4) and (5) leads to the slope of the union's indifference curve through (w, L):

$$dw/dL=-L^{-1}[(w-w_u)-(y_e-y_u)].$$

This expression is negative because the union actually redistributes cash income from the employed to the unemployed.

The locus of tangencies between the indifference curves and the isoprofit curves defines a new contract curve. Its equation is

$$R'(L)=y_e-y_u+w_u,$$

where y_e and y_u are functions of w and L. We omit the details but record that this contract curve is *downward* sloping, passes through $(\bar{w}, \bar{L})$, and lies to the right of the demand curve for labor. The picture is thus as in Figure 5.

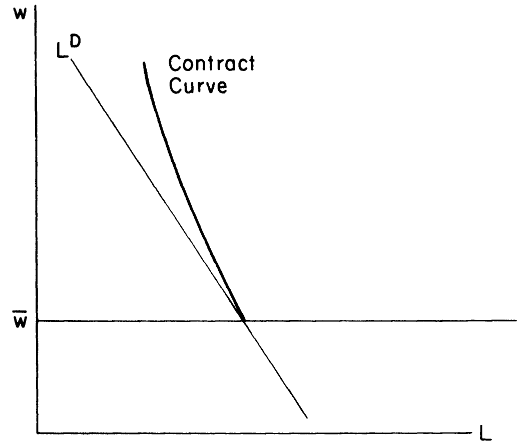

FIGURE 5

It remains true that the wage exceeds the marginal revenue product of employment in efficient bargains. The employer would prefer to reduce employment at the negotiated wage if that were permitted. But employment at a given wage is smaller in the setup of Figure 5 than it is at the same wage in the setup of Figure 3. The reason is that the *ex post* equalization of Figure 5 diminishes the incentive of the individual member to be among the employed. One important consequence is that in Figure 5 the marginal revenue product of employment exceeds the supply price of labor at every efficient bargain.

IV. Some Simple Conventions

Most formal theories of bargaining assume that the negotiated outcome will lie on the contract curve, except perhaps for the occasional conflict—a strike, say—when bargaining breaks down. We have some doubts about the empirical relevance of this assumption. But it is hard to see how one could proceed without it; so we will use it tentatively, with an eye out for its compatibility with common observation. Even so, as we have said, there is no generally accepted rule for selecting a point or other small subset of the contract curve as an especially likely candidate for the actual negotiated outcome. (The book by George de Menil contains an excellent, fairly recent, summary of the state of the theory of wage bargaining. Models of auction markets or sealed bid procedures hardly seem to apply in this context.) In this section, we consider a few very simple conventions, any one of which might seem plausible in a specific context, but none of which has any serious claim to generality. In the next section we will take up a couple of formal bargaining models which do make such a claim. After that, we turn to conventions that might apply to the renegotiation of an original bargain when the environment changes. Throughout, we emphasize the "business cycle" implications of each solution, not its place in the theory of bargaining.

A. *A Dominant Union and "Fair Shares"*

Points to the northeast along the contract curve are successively less profitable for the employer and more favorable to the workers. A powerful union might be able to force the firm to accept zero profits, if we take zero somewhat arbitrarily as the level of profit below which the firm would leave the industry or shut down. That suggests adjoining to the equation of the contract curve (3′) the zero-profit condition $R(L, B)=wL$. Geometrically speaking, this hypothesis singles out the point at which the contract curve intersects the zero-isoprofit curve.

The hypothesis of a zero level of profits can easily be generalized and made less extreme. Suppose that history has led to the notion that there is a "fair" division of net revenue between the workers and the employer. If the normal share of wages is $100k$ percent, we can write

$$wL=k\,R(L,B). \tag{6}$$

The case of zero profits is simply $k=1$. Now (3′) and (6) are the two equations defining the negotiated wage and employment. (Except when $k=1$, (6) does not coincide with a particular isoprofit locus.)

The contract curve (3), as we know, is upward sloping in the (w, L) plane; (6), on the other hand, represents w as a fraction k of the average revenue product of labor and slopes downward by our assumptions on R. This pattern will repeat itself: the negotiated outcome is at the intersection of an upward-sloping *efficiency* locus and a downward-sloping locus that can be interpreted as reflecting *equity* (or power) considerations.

Suppose the economic environment deteriorates in a recession. If the product-market effect dominates the labor-market effect, we can concentrate on a reduction in B. The contract curve shifts to the left, as shown earlier. The locus (6) shifts down. From the crude geometry it is clear that employment must fall, but the negotiated wage can go either way. That is a promising beginning for a wage-stickiness story, so we work out the result exactly.

Differentiation of (3′) and (6) leads in the conventional way to:

$$\begin{pmatrix} z & -R_{LL} \\ L & w-kR_L \end{pmatrix}\begin{pmatrix} dw/dB \\ dL/dB \end{pmatrix}=\begin{pmatrix} R_{LB} \\ kR_B \end{pmatrix}$$

where z stands for $d/dw[w-(U(w)-U(\bar{w}))/U'(w)]=((U(w)-U(\bar{w}))U''(w))/U'(w)^2$. The determinant has sign pattern $\left(\begin{smallmatrix} - & + \\ + & + \end{smallmatrix}\right)$ and is therefore negative. Calculation, and substitution of the value of k from (6) shows that

$$sgn\, dw/dB=-sgn\{R_{LB}(1-LR_L/R)+LR_{LL}R_B/R\} \tag{7}$$

This does not depend explicitly on the utility function, except as it helps determine the point at which R is evaluated. The first term is positive and the second negative, confirming the indeterminacy of the sign of dw/dB.

Two special cases are worth noting. First of all, suppose $R(L, B)$ can be written in the form $BS(L/B)$; this gives rise to the isoelastically shifting labor-demand curve discussed earlier in connection with (2). Then $dw/dB=0$, always. So whenever the elasticity of demand for labor at the going wage is approximately invariant to the business cycle, the wage will be sticky.

The second special case puts $R(L, B)=BS(L)$; this makes the inverse demand curve for labor shift isoelastically in the business cycle so that the demand elasticity at the going level of employment is invariant. Then dw/dB is opposite in sign to $d/dL(LR_L/R)$, which again suggests the lack of any strong directionality in the business cycle.

If it were the case that firms typically become sales constrained in recessions, so that revenue elasticity falls, the model would indicate a countercyclical rise in the wage. That seems extreme; but perhaps one might conclude that efficient bargaining will make employment, more than the wage, bear the brunt of cyclical adjustment.

We do not explore dL/dB in detail, because it is obvious from the geometry that this model makes employment strongly cyclical.

B. *A Dominant Employer*

If the union calls the tune, it is limited in its demands by the possibility that the firm will shut down. If the employer calls the tune, there is (usually) some similar limit to complete freedom of action. It may come from the possibility of a strike or other disruption, or it may come from the need of the employer to preserve a labor pool when there are opportunities for employment elsewhere in the economy. Even a dominant employer will push only so far to the *SW* along the contract curve. We can imagine that there is an indifference curve below which the firm will not wish to push its labor pool. There is a conceptual choice here: we could take such an indifference curve to be given by $L(U(w) - U(\bar{w}))=$constant, that is, by the *gain* to the workers from membership in the firm's labor pool. The significance of this choice is that if $\bar{w}$ falls, the firm can lower its wage at given employment to keep the workers' gain constant. On the other hand if we had fixed the limiting indifference curve by $LU(w)+(N-L)U(\bar{w}) = NU(\bar{w}) + L(U(w) - U(\bar{w}))=$ constant, the firm would have to increase its wage offer to make up for a reduction in $\bar{w}$. That seems rather too paternalistic for real life.

In this excessively paternalistic case, in fact, a simultaneous reduction in $\bar{w}$ and B must always lead to a *higher* negotiated wage along the new contract curve. Under the alternative assumption, as usual, there are forces working in both directions. The reduction in B pushes the negotiated wage upward; the contract curve shifts to the *NW* and so does its intersection with the unchanged limiting indifference curve. A reduction in $\bar{w}$ lowers the limiting indifference curve at given employment (more than one for one, in fact) and thus pushes the outcome to the *SW*. Generalized recession thus reduces the employment side of the bargain unambiguously, but the wage can go either way.

V. Formal Bargaining Theory

Most formal theories of the bargaining process proceed axiomatically. Usually one of the axioms is that the bargained outcome is efficient. But then, instead of arguing that this or that outcome on the contract curve is more "natural" than others, the bargaining theorist proposes desirable properties for a rule that would permit a referee equipped with it to go from one bargaining problem to another in some broad class, and produce a solution to each one by application of the rule. (One of the desirable properties, of course, is that the rule should always choose a point on the contract curve.) The goal of the theorist is to find a set of plausible or acceptable properties and show that there is only one rule with those properties. Raiffa argued early on that such solutions of the bargaining problem might best be thought of as Arbitration Rules; they might not have much descriptive validity in predicting the outcome of raw bargaining, but they provide a defensible handbook for an arbitrator whose job is precisely to settle a stream of bargaining conflicts. From our point of view, Raiffa's interpretation is perfectly acceptable.

Formal theories usually operate not in terms of the contract curve but in the "bargaining set" and its efficient frontier. The bargaining set is related to the contract curve in exactly the way that a production possibility set is related to the contract curve in a production box or a utility possibility set is related to the contract curve in an Edgeworth exchange box. To begin with, we need to construct the bargaining set for our model

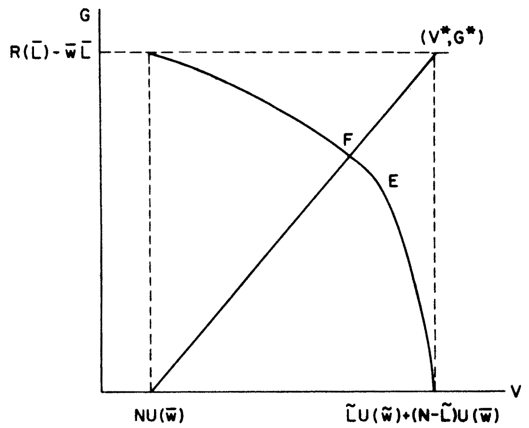

FIGURE 6

of wage bargaining. Each possible outcome (w, L) corresponds to a payoff to each party. In the case of the employer, the payoff is $G=R(L)-wL$; to the union, the payoff is $V=LU(w)+(N-L)U(\bar{w})$. The set of (G, V) swept out as (w, L) ranges over its possibilities is the bargaining set. If no bargain is struck, we take it that $G=0$[6] and $V=NU(\bar{w})$. As (w, L) traverses the contract curve, (G, V) traverses the undominated efficient subset of the bargaining set, by construction.

Axiomatic bargaining theories require that the bargaining set be convex so that, in the usual way, the efficient payoff-possibility curve is a decreasing concave function in the (G, V) plane. They justify an assumption of convexity by the possibility of randomization. That would hardly do in the wage-bargaining context, but fortunately the assumptions we have made on $R(.)$ and $U(.)$ guarantee, as tedious calculation will show, that the frontier of the bargaining set is well behaved. The picture, therefore, is as in Figure 6. (Here $\tilde{w}$ is the highest wage the employer can pay and still break even at a point on the contract curve; see Figure 3.) Selection of a point on the contract curve is equivalent to selection of a point on the efficiency locus.

The best-known formal solution to the bargaining problem is Nash's. It selects the efficient point that maximizes the *product* of the parties' gains over and above the no-contract outcome. In this case, it maximizes $G\cdot(V-NU(\bar{w}))$ or $(R(L)-wL)(U(w)-U(\bar{w}))L$ over the bargaining set. In principle, one might think of maximizing that product subject to the equation of the contract curve; but on reflection, the constraint can be omitted. Unconstrained maximization of the product by choice of (w, L) will certainly try to maximize $R(L)-wL$ for any given value of $L(U(w)-U(\bar{w}))$, so the equation of the contract curve will reappear as one of the first-order conditions of the unconstrained problem. It does. The other first-order condition turns out to imply, rather oddly,

$$w=(R/L+R'(L))/2. \tag{8}$$

At the Nash solution, the wage is equal to the arithmetic mean of the average and marginal revenue products of labor!

The Nash solution is thus defined by (3) and (8). Under our assumptions about $R(.)$, both the average and marginal revenue products are decreasing. So (8) defines a negatively sloped "equity" locus that intersects the contract curve once, at the Nash solution to the bargaining problem. Once again, we can replace R by $R(L, B)$ and ask how variation in B affects the wage coordinate of the solution. Upon calculation, it turns out that the criterion (7) holds here too, and so do the paragraphs of text immediately following (7).

One of the axioms leading to the Nash solution of the bargaining problem requires the rule to be "independent of irrelevant alternatives." Suppose that E is the solution to the bargaining problem pictured in Figure 6; now define a new bargaining problem by deleting part of the bargaining set, any part so long as the point E remains. The axiom requires that E be the solution of the new bargaining problem. Since the deleted outcomes were not chosen by the rule when they were available, they are "irrelevant" and their absence should make no difference to the outcome. This axiom has been much complained about, and justly. Intuitions about "bargaining power" and "fairness" might include the notion that if A *could* win a lot in a bargaining situation, he or she is "entitled"

[6]It would not be hard to allow for fixed costs F, so that $G=-F$ if no agreement is reached.

to more than if he or she could only, in the best of circumstances, win a little. Anyone who shares that intuition does not believe that "irrelevant" alternatives are irrelevant.

Dissatisfaction with this axiom has led to other definitions of the solution to a bargaining problem. Ehud Kalai and Meir Smorodinsky propose replacing the unsatisfactory axiom with another. Start again with Figure 6 and suppose again that E is the solution chosen by the rule. Now alter the bargaining set in the following way: leave the "best possible" outcomes for the vertical and horizontal parties unchanged, but fix things so that for each possible benefit to the horizontal party the largest possible gain to the vertical party is bigger than it was before. Then the axiom of monotonicity requires that the rule assign to this new bargaining set a solution that gives the vertical party more than at E. If the environment becomes more favorable for the vertical party in this strong sense, the vertical party must profit from the change. (Of course the environment can become more favorable for both parties; then they must both gain.)

Kalai and Smorodinsky show that replacing the axiom of irrelevance of independent alternatives with the axiom of monotonicity leads to a unique solution different from Nash's. It is easily described. Let G^* be the best that the vertical party could hope for, $R(\bar{L})-\bar{w}\bar{L}$ in Figure 6. Let V^* be the best the horizontal party could hope for. Find the point (G^*, V^*) and draw a line connecting it with the no-bargain point, whose coordinates in Figure 6 are $(0, NU(\bar{w}))$. The solution is the unique point at which that line intersects the efficiency frontier. It is shown in Figure 6 as F.

Although the geometry is simple, the arithmetic of the Kalai-Smorodinsky solution is not. The equation to be adjoined to that of the contract curve is

$$(9)\quad (R(L)-wL)/(L(U(w)-U(\bar{w}))) = (R(\bar{L})-\bar{w}L)/(\tilde{L}(U(\tilde{w})-U(\bar{w}))),$$

where $(\bar{w}, \bar{L})$ and $(\tilde{w}, \tilde{L})$ are the left- and right-hand end position of the contract curve. Thus $R'(\bar{L})=\bar{w}$ and $R(\tilde{L})=\tilde{w}\tilde{L}$. (We are assuming here that $\tilde{L}\leq N$; otherwise $\tilde{L}$ has to be replaced by N and $\tilde{w}$ by the wage at which the contract curve generates employment of N workers.)

The comparative statics of the solution is messy, mainly because one must keep track of changes in $\bar{w}$, $\tilde{w}$, $\bar{L}$, and $\tilde{L}$, so we hold our comments to a minimum. If we set $R(L, B) = BR(L)$ and calculate dw/dB by total differentiation, the familiar quantity $d/dL[LR'/R]=RR'/L+RR''-(R')^2$ appears; positive (negative) values are associated with $dw/dB<0(>0)$. But this time there is an additional negative term, so that even if $R(L)$ has constant elasticity, $dw/dB<0$. So the Kalai-Smorodinsky solution is more likely than the Nash solution to yield "perverse" countercyclical wage flexibility. There is not much else to be said. Even the effects of changes in $\bar{w}$ are complicated.

VI. Sales Constraints and Incremental Bargaining

This brief section serves three purposes. We make an initial stab at the potentially important case of a fix-price firm which experiences a progressively more binding sales constraint as a recession proceeds. Then we use this sketch as a vehicle to introduce another idea: that bargaining conventions may apply to the sharing of gains and losses when a change in the environment makes an initial situation untenable. The initial situation could even be arbitrary; this is an entry point for historical happenstance. Finally, the same sketch helps to clarify the underlying reason why efficient bargaining holds the potential for countercyclical movement of wage rates.

We can model a sales-constrained firm by assuming that every isoprofit curve simply ends when it reaches the value of L corresponding to maximal sales, as illustrated in Figure 7. This assumption is a bit too strong, because it ignores possible substitution among variable factors with constant output. But we shall use it for illustration.

When the firm experiences a sales constraint, the contract curve cannot extend to the right of its intersection with the vertical

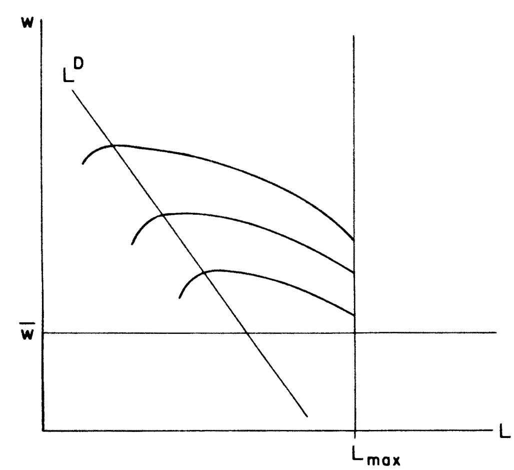

FIGURE 7

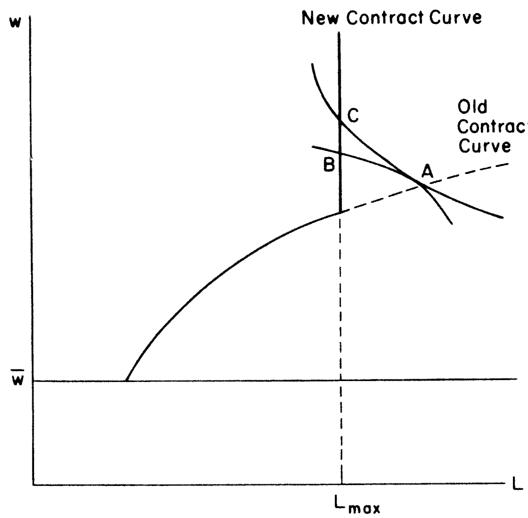

FIGURE 8

at L_{max}, because such points are inaccessible. In fact, the contract curve coincides with the vertical *above* the intersection point, as in Figure 8. (Below the intersection, points on the vertical are dominated by points on the contract curve to the left of L_{max}.)

A simple recession story might go as follows. To begin, the sales constraint is not binding and the wage-employment bargain is somewhere on the contract curve. It remains there in the early stages of the recession as L_{max} moves to the left but is not yet binding. Eventually, the constraint just binds, and then moves still further to the left. The initial bargain is no longer tenable. (We ignore labor hoarding only in order to concentrate on the logic of bargaining.) What happens now?

Given an accepted equity rule (which might also be shifting systematically as the recession proceeds) a new wage-employment bargain might be struck at the intersection of the recession-shifted equity and contract curves. But suppose the initial bargain A had arisen mostly by historical accident. It might not even be efficient. A natural incremental equity rule might be that both parties should gain, or both lose, by the change, but not one gain and the other lose. In the recession case, both must lose. This suggests that the new bargain would have $L=L_{max}$ and a wage somewhere in the interval BC between the isoprofit curve and the indifference curve associated with the initial bargain. This is shown in Figure 8.

Any point on BC has a wage higher than A; the wage would rise as employment falls. This is too sharp to be taken literally. The logic of this result provides, however, a clear insight into the mechanism through which efficient bargaining can generate countercyclical wage variation. At any point on the contract curve, the firm would prefer a lower volume of employment at the bargained wage. We have already suggested that contractual work rules and manning agreements might serve the purpose of enforcing this extra employment on the firm. A binding sales constraint thus benefits the employer by necessitating, or providing the excuse for, a reduction in employment. At the old wage, the firm would be better off and the union worse off. If the incremental equity rule forbids such an outcome, the wage must rise to transfer some of the loss from union to firm. This shows up with great clarity here because the recession is assumed to leave the revenue function unchanged except by imposing a barrier at L_{max}.

VII. Conclusion

We set out to understand why real wages might be sticky, why fluctuations in aggregate demand might have their major effect on employment and little or none on the wage. Our partial-equilibrium bargaining models can hardly be expected to do that. But they do quite generally confirm a tendency for fluctuations in real product demand at the firm or industry level to be accompanied by large correlated fluctuations in employment and small changes in real wage rates that could go in either direction. What is the source of that tendency?

Geometrically speaking, it is because both our efficiency locus and our "equity" or "power" locus shifts to the left in recession and to the right in upswings, provided cyclical changes in product markets dominate those in the effective reservation wage. The shifts operate in the same direction on employment, but in opposite directions on the negotiated wage, so there is a clear possibility that the two will be statistically independent, as empirical investigation suggests.

A deeper, less mechanical, answer might go something like this. Efficient bargaining pushes the firm to hire more workers than it would like at the negotiated wage. The outcome is thus on the falling part of the isoprofit curve. The contract curve slopes upward. Higher employment and higher wages favor the workers; lower employment and lower wages favor the firm. When circumstances enforce a reduction in employment, the employer gains and the workers lose. Equity and bargaining power are likely to seek an adjustment that will transfer some of the employer's gain to the union or shift some of the union's loss back to the firm. The part of this adjustment that falls (efficiently) on wages will involve an increase in the wage. This tendency can, in principle, offset the normal cyclical deterioration of the demand for labor, in part, wholly, or not at all.

Most of the paper represents variations on this theme. If short-term contracting would lead to cyclical fluctuation in employment at a more or less stable wage, then convenience could easily lead the parties to contract for a long-term steady wage, with current employment decisions made by the firm. The union would need protection against excessive (i.e., profitable but "inefficient") reduction in the average level of employment; this could be provided by work rules or manning agreements.

Our main result is in sharp contrast to the outcome of standard models of implicit contracting with symmetric information. There, long-run contracts tend to be employment stabilizing as compared with spot-competitive labor markets. The crucial difference appears to be that implicit-contract models are closed by a utility constraint. We replace this condition by the sort of equity convention that arises naturally in the bargaining context and is less dominated by opportunities available elsewhere in the economy.

REFERENCES

C. Azariadis, "Implicit Contracts and Related Topics: A Survey," unpublished working paper, 1979.

A. M. Cartter, *Theory of Wages and Employment*, Homewood: Irwin, 1959.

George de Menil, *Bargaining: Monopoly Power versus Union Power*, Cambridge, Mass. 1971.

P. T. Geary and J. Kennan, "The Employment-Real Wage Relationship: An International Study," *J. Polit. Econ.*, forthcoming.

R. E. Hall and D. M. Lilien, "Efficient Wage Bargains under Uncertain Supply and Demand," *Amer. Econ. Rev.*, Dec. 1979, *69*, 868–79.

E. Kalai and M. Smorodinsky, "Other Solutions to Nash's Bargaining Problem," *Econometrica*, May 1975, *43*, 513–18.

W. Leontief, "The Pure Theory of the Guaranteed Annual Wage Contract," *J. Polit. Econ.*, Feb. 1946, *54*, 76–79.

H. Raiffa, "Arbitration Schemes for Generalized Two-Person Games," in H. W. Kuhn and A. W. Tucker, *Contributions To The Theory of Games*, II, Princeton 1953, 361–87.