ROBERT E. HALL

Massachusetts Institute of Technology

The Rigidity of Wages and the Persistence of Unemployment

IN THE MODERN American economy, fluctuations in unemployment usually persist from one year to the next. Past decades have seen two lengthy deviations from equilibrium in the labor market: 1958 through 1963, when the market was slack for six straight years, and 1964 through 1970, when it was tight for seven straight years. Virtually all forecasters agree today that the sharp recession of 1974–75 will mark the beginning of another extended period of slack, with the official unemployment rate above 6 percent perhaps until the end of the decade. These forecasts are fully consistent with the behavior of the economy after the only comparable postwar recession, in 1957–58. Explanation of the persistence of unemployment has been a major focus of macroeconomic theory since the Great Depression. Though his predecessors had begun to think seriously about the challenge to classical economic theory raised by successive years of high unemployment, John Maynard Keynes was the great pioneer in creating a theory that came to grips with the facts of persistent unemployment, and he remains the dominant figure today.

This paper presents a detailed critique of modern explanations of the behavior of unemployment, both within the Keynesian tradition and outside it. The critique begins with a close look at the modern theory of disequilibrium, which claims to provide a microeconomic foundation for the theory of unemployment and wage adjustment. This theory attributes both the

301

rigidity of wages and the persistence of unemployment to the slow diffusion of information within the labor market. The evidence presented here, however, suggests that unemployment in the modern American economy persists far longer than is remotely plausible if this factor is the main cause of the lag in market clearing. This hypothesis is strongly confirmed by a study of the data under the assumption that three months is long enough for both workers and employers to find out what is happening in the labor market.

The paper then turns to a group of theories that are more traditionally Keynesian in flavor. First is the belief that labor unions are responsible for the rigidity of wages. Keynes' theory of effective demand showed convincingly that rigid wages can permit persistent unemployment, and Keynes himself, as well as his successors, appeared to hold labor unions primarily responsible for the nonclassical behavior of wages. Closely related is the theory of wage determination based on "idiosyncratic exchange." This theory emphasizes the bilateral monopoly between individual workers and employers, which biases against the adjustment of wages in the short run and makes equity an important element in wage determination. These two theories share an essential defect: however successful they are in explaining wage rigidity in a particular sector, they do not seem capable of explaining economy-wide rigidity. As long as a competitive residual sector with a flexible wage exists, workers who are unable to find jobs in the rigid-wage sector will find it in the residual sector and its wage will be depressed as a result. Overall, the labor market will clear even if one sector has rigid wages. If unions and other sources of rigidity in particular sectors have a role in persistent unemployment, it is more subtle than the traditional Keynesian view supposes.

Empirical evidence on the behavior of wages, presented next, suggests that the modern American economy contains an important sector with rigid wages. Though unionism and idiosyncratic exchange are present in this sector, its defining characteristic is different. The sector with inflexible wages consists of government, regulated industries, and nonprofit institutions. These industries make up what I will call the "nonentrepreneurial" sector, in which employers do not face the usual incentive to minimize labor cost. The empirical results suggest that the nonentrepreneurial sector is much the most important source of wage rigidity in today's economy.

After isolating the major source of wage rigidity in the contemporary U.S. economy, the paper goes on to present a theory of the transmission, or spillover, of this rigidity to the competitive sector. The modern theory of

the supply side of unemployment is shown to imply that competitive wages will not fall far enough to clear the labor market at low frictional levels of unemployment when aggregate demand falls. Rather, the duration of unemployment will rise as the unemployed search longer for the high-paying jobs in the rigid-wage sector. During a contraction, the payoff to job search rises because the wage premium for these good jobs rises. Conversely, search is less useful to the unemployed when the labor market is unusually tight, as competitive wages rise close to rigid wages. The existence of the rigid-wage sector lessens the ability of the competitive sector to push the wage down to its market-clearing level during a contraction. The paper does not contain direct evidence of this mechanism, but does show that the observed relation between unemployment and relative wages in the two sectors accords roughly with the predictions of the theory.

All of the theories discussed here are Keynesian in the sense that they explain why the mechanism of effective demand can operate without the immediate corrective force of market-clearing wage adjustments. They agree that the basic cause of an increase in unemployment is contractionary aggregate policy or other shocks in the economy, and similarly that expansionary policy can eliminate excessive unemployment. With the possible exception of some versions of the theory of disequilibrium, none of these theories can fairly be accused of making persistent unemployment a voluntary phenomenon arising from the supply side of the labor market. Rather, all of the theories that appear to have any explanatory power concur that unemployment is the result of inadequate demand for labor. The interaction of demand and supply is described in greater detail in the next section.

Theories of Unemployment

A simple two-equation structural model of the economy captures the theoretical relationships that are important for the issues addressed by this paper and can encompass the alternative theories considered here.

(1)   effective demand: $u_t = f(w_t, x_t)$;

(2)   wage adjustment: $w_t - w_{t-1} = g(u_t) + \hat{w}_t - w_{t-1}$.

Here u_t is the unemployment rate and the function f embodies the effective-demand mechanism for the determination of real gross national product

and Okun's law for translating it into the unemployment rate. Effective aggregate demand depends negatively on the nominal wage, w_t, through the Keynes and Pigou effects, and depends on various policy variables, x_t (monetarists and Keynesians disagree about the presence of fiscal variables in x_t, but that does not matter here). In the wage-adjustment equation, $g(u_t)$ is the disequilibrium component and $\hat{w}_t - w_{t-1}$ is the expected or equilibrium component.[1] Monetarists usually write the equation in a precisely equivalent form,

$$(3) \qquad u_t = g^{-1}(w_t - \hat{w}_t),$$

and call it "aggregate supply."[2]

A convenient algebraic specification of the system is

(4)   effective demand: $u_t = \psi_0 - \psi_1(m_t - w_t) + \eta_t$;

(5)   wage adjustment: $w_t - w_{t-1} = -\phi_1(u_t - u^*) + \hat{w}_t - w_{t-1} + \epsilon_t$,

where u^* is the equilibrium value. Here I have switched to taking w as the log of the nominal wage, m as the log of the money supply, and u as the log of the unemployment rate. I assume, for simplicity, that other policy variables, represented by x in the previous form of the equation, do not enter in determining effective demand. The equations also contain random shifts, η and ϵ. In a structural model, the fundamental explanation of the behavior of the endogenous variables resides in the reduced form, which gives the impact of the predetermined variables on the endogenous variables after taking account of the interdependence of the endogenous variables in the equations. Statements dealing with causal relations between endogenous variables—such as "high wages cause excessive unemployment"—are meaningless or contradictory in a structural model. High wages are associated with high unemployment in the aggregate-demand equation, but with low unemployment in the wage-adjustment equation. The reduced-form equation for unemployment sorts out the interaction of the two variables in both equations:

$$(6) \qquad u_t = u^* + \frac{\psi_0 - \psi_1(m_t - \hat{w}_t - \epsilon_t) + \eta_t}{1 + \psi_1\phi_1}.$$

1. More generally, $\hat{w}_t - w_{t-1}$ embodies the inertia that is widely believed to be present in wage determination. The evidence suggests that expectations are only a part of the explanation of inertia, but the use of the term is so widespread that I will adopt it here.

2. Thomas J. Sargent studied exactly this two-equation system in monetarist form in "Rational Expectations, the Real Rate of Interest, and the Natural Rate of Unemployment," *BPEA*, 2:1973, pp. 429–72.

The major issues of the theory of unemployment that are considered in this paper concern the values of the parameters in this equation, the specification of the expected wage, and the behavior of the random shifts. First is the role of the nominal wage in determining real aggregate demand, controlled by the parameter ψ_1. In the crude Keynesian model with the liquidity trap and no Pigou effect, $\psi_1 = 0$, wage adjustment is irrelevant, and unemployment remains at a level determined by the effective-demand process:

(7)   crude Keynesian reduced form: $u_t = u^* + \psi_0 + \eta_t$.

Empirical evidence reviewed later in the paper reveals clearly that the nominal wage exerts a strong force on real aggregate demand through the Keynes and Pigou effects, so the crude Keynesian model deserves no further consideration.

   The second major issue is the responsiveness of the nominal wage to excess supply or demand in the labor market, measured by the parameter ϕ_1. In a classical model with virtually instantaneous clearing of the labor market, ϕ_1 is extremely large and unemployment never deviates from its equilibrium value, u^*:

(8)   classical reduced form: $u_t = u^*$.

The classical model is unable to explain *any* of the fluctuations in unemployment and also deserves no further consideration. All modern theories rest crucially on the hypothesis of temporary inflexibility of wages to explain the behavior of unemployment: *the Phillips curve is an essential part of all contemporary macroeconomic models*. One of the major tasks of this paper is to appraise the current state of theories of the Phillips curve.

   The third major issue is the specification of the expectation term, $\hat{w}_t$. Modern Keynesians generally concede the importance of this variable in determining unemployment, but view it as responding sluggishly to the past history of inflation. Modern Keynesian theory, which combines this view with the beliefs that money matters (ψ_1 is large) and that wages adjust slowly (ϕ_1 is small), permeates the thinking of most economists and policy-makers, including many who would not call themselves Keynesians. The consensus on the Keynesian diagnosis of the present state of the U.S. economy is particularly strong: expectations of high rates of inflation and other sources of inflationary momentum are deeply embedded in the economy today as a result of more than a decade of accelerating inflation. Monetary authorities are determined to bring inflation under control by limiting

monetary growth. As a result, the term $m_t - \hat{w}_t$, which appears in the reduced form with a strong negative coefficient, will keep unemployment above its equilibrium for many years in succession, until inflationary expectations subside and thus raise $m_t - \hat{w}_t$ to its full-employment level. In the modern Keynesian theory, monetary authorities are capable of sustained contractionary policies; its attraction is its apparent ability to explain the persistence of unemployment as a result of such policies on the part of the monetary authorities.[3]

Theorists of rational expectations, especially Robert Lucas, have pointed out a fundamental logical flaw in the Keynesian position.[4] Inflationary expectations are formed by intelligent people, who ought to take advantage of all information available about the economy when they make their plans. As shown below, rational expectations implies the following reduced form for the unemployment rate:

(9)   rational-expectations reduced form: $u_t = u^* + \dfrac{\epsilon_t - v_t}{\phi_1}$.

The random variable, v_t, is the difference between the actual and expected wage levels. According to the theory of rational expectations, v_t is uncorrelated with any variable whose value is known at the time expectations are formed, including its own past values. The theory implies that the reduced form for unemployment contains only random variables. Because it is nonclassical in recognizing the sluggishness of wage adjustment, it is capable of explaining the existence of unemployment in any particular year. However, the theory explains *persistent* movements in the unemployment rate only by relabeling them as shifts in the Phillips curve; v_t is not permitted to move persistently at all, so the entire burden falls on persistent movements of ϵ_t, the random shift in the wage-adjustment equation. The relative importance of the two random variables is examined in a later section.

The present state of the two major macroeconomic theories may be summarized as follows: Keynesian theory is favored by practical economists and policymakers. It seems to give a reasonable account of the important

3. A decade ago, Keynesians relied heavily on persistently contractionary fiscal policy (fiscal drag) to explain persistent unemployment. Since then, Keynesian models have been refined to reduce the impact of fiscal policy through "crowding out."

4. See Robert E. Lucas, Jr., "Econometric Testing of the Natural Rate Hypothesis," in Otto Eckstein, ed., *The Econometrics of Price Determination*, A Conference sponsored by the Board of Governors of the Federal Reserve System and Social Science Research Council (Board of Governors, 1972), pp. 50–59.

facts about unemployment and to support the view that monetary policy has a substantial and lasting impact on the unemployment rate. On the other hand, the basic hypothesis of rational expectations has considerable logical force and has appealed to an important segment of theorists. The task facing students of unemployment and wage adjustment is to create a logically consistent theory capable of squaring the Keynesian facts of sluggish wage adjustment and persistent unemployment with the persuasive criticisms of the rational-expectations school based on the behavior of intelligent economic agents. This paper devotes most of its attention to that task.

THE ROLE OF NOMINAL WAGES IN REAL AGGREGATE DEMAND

In an economy with a predetermined nominal money supply, the persistence of unemployment derives from the rigidity of *nominal* wages— that is, in times of high unemployment, a reduction in money wages would restore full employment. This proposition was the subject of heated controversy in the past, and even today some macroeconomists are skeptical of the stimulative effects of wage cuts (typically, they also deny that wage cuts would ever occur in the first place). The bulk of empirical evidence, however, demonstrates a connection between the money wage and aggregate real output whose strength may be inadequately appreciated by modern Keynesian economists. Because the central argument of this paper rests on the hypothesis that the sluggishness of wages accounts for the persistence of unemployment, the evidence on the connection of wages to real output and unemployment is worth study at this point.

The first step on the logical path from wages to unemployment is from wages to prices. Research on the determination of prices has converged toward the simple view that prices are proportional to costs. Wages are by far the largest component of costs, although recent experience has highlighted the importance of prices of raw materials and agricultural inputs as well. Modern price equations have a measure of wages (usually adjusted for the trend of productivity but not for its cyclical variation) as the dominant right-hand variable.[5] Price and wage inflation take place simultaneously. Except for a smooth trend and, recently, the effects of variations in other real factor prices, there has been essentially no variation in real

5. See James Tobin, "The Wage-Price Mechanism: Overview of the Conference," in Eckstein, ed., *Econometrics of Price Determination*, pp. 5–15.

wages.[6] Rigidity of wages brings about a corresponding rigidity of prices.

The next step is the connection between the price level and aggregate output. In the simple quantity theory the two are exactly inversely proportional; the nominal quantity of money multiplied by velocity is the constant of proportionality. If the money stock remains constant, a reduction of 1 percent in the price level raises real GNP by 1 percent. At the other end of the spectrum is the extreme Keynesian model, in which the demand for money is highly interest elastic and aggregate demand is unresponsive to interest rates. A reduction in prices causes interest rates to fall but has little effect on real GNP. The structure of contemporary macroeconomic models admits both of these extremes as special cases. The interest elasticities of the demand for money and of aggregate expenditure are estimated empirically, and the behavior of the models lies part way between those of the quantity and the extreme Keynesian theories.

Among the large macroeconomic models, the MIT-Pennsylvania-Social Science Research Council (MPS) model has investigated these issues most carefully. In the MPS model, the interest elasticity of the demand for money is fairly low—0.06 in the short run, 0.08 in the long run.[7] Changes in prices cause large changes in interest rates in the same direction. Falling interest rates stimulate expenditure through two principal channels: the increase in demand for investment goods, housing, and consumer durables; and the increase in consumption, stimulated by the increase in wealth in the stock market. (There is also a small stimulus to consumption because money is a part of wealth—the Pigou effect.) Missing from the MPS model is a third channel whose importance is increasingly recognized.[8] When in-

6. The view that excessive unemployment could be attributed to a level of *real* wages that is too high was widespread before Keynes. The *General Theory* seems to have disposed of the view; see John Maynard Keynes, *The General Theory of Employment, Interest and Money* (Harcourt, Brace, 1936). The roles of nominal and real wages in the theory of disequilibrium are discussed clearly and thoroughly by Robert Barro and Herschel Grossman, "A General Disequilibrium Model of Income and Employment," *American Economic Review*, vol. 61 (March 1971), pp. 82–93. They conclude that the fundamental cause of disequilibrium is an inappropriate level of prices and nominal wages, and that the real wage determines only the market in which the disequilibrium appears.

7. See "Equations in the MIT-Penn-SSRC Econometric Model of the United States" (Massachusetts Institute of Technology, Department of Economics, January 1975; processed), sec. 15.

8. See, for example, Ronald McKinnon, "The Limited Role of Fiscal Policy in an Open Economy" (paper presented at the Conference on the Monetary Mechanism in Open Economies, Helsinki, Finland, August 1975; processed).

terest rates fall in the United States, foreigners will increase their borrowing. Since capital flows and trade flows must always sum exactly to zero, every dollar loaned to foreigners must be accompanied by a dollar increase in U.S. net exports. The demand for net exports is interest elastic as a consequence of this process. The omission of the third channel biases the MPS model toward understating the expansionary effect of wage reductions and thus toward understating the cost of wage rigidity in terms of persistent unemployment.

In the MPS model, a one-time reduction of 1 percent in all prices and wages raises real GNP by about 0.1 percent within one quarter.[9] The effect grows rapidly, reaching 0.7 percent after one year and achieving its peak of about 1.5 percent in eight quarters. Except for the lags, the MPS model is more like the quantity theory than like the extreme Keynesian theory. Within the model, persistent unemployment would not exist if wages adjusted rapidly. The model is Keynesian in embodying sluggish wage adjustment, but distinctly non-Keynesian in implying a quick return to full employment after a reduction in nominal wages. Addition of an equation for net export demand that was properly sensitive to interest rates would make the MPS model even less Keynesian in the second respect.

Perhaps the major qualification to the argument that wage reductions could stimulate employment is that deflation brings about expectations of further deflation, and these expectations inhibit the expansion of demand.[10] Symmetrically, inflation should stimulate demand through the same process and lessen the impact of wage increases in reducing excess demand. Buyers will postpone purchasing durable goods if they expect them to be cheaper in the future and will hoard them if they expect their prices to rise. Within modern Keynesian theory, business investment is the component of demand that is viewed as most sensitive to price expectations. The MPS

9. These effects were inferred indirectly from solutions of the current version of the model. Because the model has a great many exogenous variables specified in nominal terms, it is not possible to compute the impact of wage and price reductions directly. Instead, the solutions increased the money supply. The homogeneity of the model makes an increase in the money supply equivalent to a reduction in wages and prices, under the assumption that the wage and price cut does not feed back through the expectations equations. The model has essentially no lag between prices and wages.

10. This issue is discussed by James Tobin, "Keynesian Models of Recession and Depression," *American Economic Review*, vol. 65 (May 1975), pp. 195–202. As Tobin points out, an unfavorable expectation effect is not even certain as a theoretical matter. Here I do not disagree with Tobin's analysis, but merely present evidence that the effect is very small.

investment equation embodies the view that a downward shift of 1 percentage point in the expected rate of inflation has the same discouraging effect on investment as an increase of 1 percentage point in the nominal interest rate. However, expectations respond to the history of inflation or deflation with such a long lag that the immediate sharp response of the nominal interest rate to a change in the price level is offset only to a small extent by a shift in expectations. In other words, in the MPS model a wage-price reduction (or monetary expansion) has a pronounced downward effect on real interest rates. In turn, demand depends on real interest rates and responds to a 1 percent reduction in wages and prices only slightly less than it does to a 1 percent increase in the money supply. Keynesian concerns about the adverse effects of deflation operating through expectations are not sustained by the MPS model.

This brief study suggests that the connection between unemployment and wages is the place to look for an explanation of persistent unemployment. The economy would behave in a thoroughly classical manner if only the short-run Phillips curve were much steeper. The study equally confirms the view that monetary expansion is a swift and sure corrective for excessive unemployment. Even with rapid wage adjustment, only a concern about inflation would cause policymakers to choose laissez-faire over monetary expansion as a response to an unemployment shock.

Microeconomic Disequilibrium Theory

Dissatisfaction with the lack of theoretical explanations of the sluggishness of wage adjustments has led economists to study the behavior of wages and unemployment in an atomistic labor market. Many have argued that, in a market with many small employers and independent workers, wages do not react instantly to excess supply or demand because of the time it takes for information to spread.[11] Employers do not take immediate advantage of a slackening in demand by reducing wage offers because they become aware only gradually that workers have become more readily available. Similarly, workers continue to hold out for the kind of job they would have been able to get under normal conditions. The initial impact of a reduction in demand is an increase in unemployment. Later, wages fall,

11. See Robert E. Hall, "The Process of Inflation in the Labor Market," *BPEA*, *2:1974*, app. A, for further discussion and references.

stimulating demand and reducing unemployment to its normal level. While information is diffusing through the labor market, nonclassical quantity adjustments take place, but in the longer run the market behaves classically.[12] This modern theory of disequilibrium is quantitatively at odds with the facts of persistent unemployment. When unemployment has been excessive for several years running, how can employers remain unaware of the bargains in the labor market in the form of high-quality workers willing to take lower wages? How can workers remain unaware that jobs are hard to find so that reservation wages should be cut? When it is universally predicted that labor markets will be extremely tight, why don't employers recruit effectively by raising wage offers? Why don't workers avoid taking new jobs prematurely by raising their reservation wages? The strong persistence of unemployment implies that the diffusion of information within the labor market can take five years or more, which seems altogether implausible. The theory of disequilibrium is well formulated and internally consistent, but its relevance to the actual behavior of the United States economy is open to serious question.

Persistent Unemployment in the Theory of Rational Expectations

The criticism of modern theories of the Phillips curve just sketched is sharpest within the framework of rational expectations. As mentioned above, within a two-equation model similar to the one discussed at the beginning of this paper, Sargent discussed a Phillips curve embodying rational expectations in which unemployment can persist, though the main focus of the paper is on other issues and the persistence of unemployment is not addressed specifically.[13] The role of the aggregate-demand equation in the rational-expectations model is simply to contribute to the surprise in the disturbance. Thus, the starting point for my purposes is the wage-adjustment equation:

12. Axel Leijonhufvud, *On Keynesian Economics and the Economics of Keynes: A Study in Monetary Theory* (London: Oxford University Press, 1968), pt. 2, argues that modern disequilibrium theory is really what Keynes had in mind. Keynes carefully avoided blaming unemployment on labor unions (see, for example, *General Theory*, p. 9). He also subscribed to the modern view of the Phillips curve that wage adjustment takes time and that strict rigidity of wages is a useful assumption only in the short run (see ibid., pp. 247–51).
13. "Rational Expectations."

(5) $w_t - w_{t-1} = -\phi_1(u_t - u^*) + \hat{w}_t - w_{t-1} + \epsilon_t.$

The natural, or equilibrium, rate occurs when $w_t = \hat{w}_t$, or

(10) $u_t^* = u^* + \dfrac{\epsilon_t}{\phi_1}.$

The random disturbance shifts the equilibrium rate. Now, the criterion for rationality of $\hat{w}_t$ is that it be an unbiased predictor of the actual wage, given all the information available just before period t, say b_t:

(11) $\hat{w}_t = E(w_t | b_t).$

This implies that w_t and $\hat{w}_t$ differ by a random forecast error, ν_t, that obeys $E(\nu_t | b_t) = 0$; b_t contains no information about ν_t, for if it did, the information could be used to improve $\hat{w}_t$. The error ν_t is uncorrelated with everything known before t, including particularly its own lagged values. Purely random serially uncorrelated errors of expectations are the central implication of the theory of rational expectations. Inserting the forecast error in the Phillips curve gives

(12) $\nu_t = -\phi_1(u_t - u^*) + \epsilon_t.$

As an equation explaining unemployment, this is

(9) $u_t = u^* + \dfrac{\epsilon_t - \nu_t}{\phi_1}$

(13) $= u_t^* - \dfrac{\nu_t}{\phi_1}.$

The unemployment rate differs from its equilibrium value by a serially uncorrelated purely random disturbance, $-\nu_t/\phi_1$. If ϵ_t were also serially uncorrelated, then u_t itself would fluctuate in a purely random serially uncorrelated fashion around its mean, u^*. Lucas has pointed out this implication of rational expectations,[14] but the data show the implication to be totally false. Unemployment is highly serially correlated.

Sargent's modification of Lucas' work provides a model that describes the facts about unemployment more satisfactorily. He suggests that the random shift of the Phillips curve, ϵ_t, may be serially correlated, implying that the persistence of unemployment arises from persistence of the shifts. But this is no more than a relabeling of the same phenomenon: *the theory proposed by Lucas and Sargent contains no explanation whatever of the*

14. "Econometric Testing of the Natural Rate Hypothesis."

persistence of unemployment but merely identifies it with an unexplained shift in the equation that describes unemployment. The central contribution of their theory is its demonstration that in an economy of rational agents, the part of unemployment that is explained by the tradeoff between inflation and unemployment is totally random and unpredictable. The predictable component of unemployment (as well as part of the unpredictable component associated with the unpredictable part of ϵ_t) must be explained by forces not considered in their theory.

From the stochastic behavior of the unemployment rate one can calculate bounds on the fraction of its total variation that is attributable to inflationary or deflationary surprises and on the remaining fraction that is attributable to shifts of the Phillips curve.[15] The first step is to rewrite the model in a form that makes explicit the distinction between the two sources of fluctuations in unemployment. The problem is that part of ν_t is the instantaneous response of the economy to the unexpected fluctuation in ϵ_t—shifts of the Phillips curve are one possible source of inflationary surprises. The unexpected wage fluctuation can be separated into two uncorrelated components: ν'_t, the component caused by unexpected monetary policy and other unexpected developments not arising from the Phillips curve, and a residual that is a multiple of the unexpected component of ϵ_t. Then the unemployment equation can be written

$$(14) \qquad u_t = u^* + \frac{\epsilon'_t - \nu'_t}{\phi_1}.$$

Here ϵ'_t is the shift of the Phillips curve net of the compensating response in ν_t, and ν'_t is ν_t net of the same compensating response. I will assume that the net inflationary surprise, ν'_t, is uncorrelated with the net shift, ϵ'_t. This is reasonable if the major sources of unexpected inflationary shocks are monetary or fiscal policy, surprises in world markets, or unexpected changes in demand. None of these should shift the Phillips curve directly. The computation of the bound on the fraction of the variance contributed by ν'_t starts from the observation that a second-order autoregressive equation gives quite an accurate prediction of the unemployment rate. The variance of ν'_t cannot be any larger than the variance around this prediction equation and in fact must be substantially smaller; the prediction error contains lagged values of the serially uncorrelated ν'_t as well as the unpredictable

15. I am indebted to Christopher Sims for pointing out an error in an earlier version of this section.

part of ϵ_t'. The predicted deviation of the unemployment rate from its mean, u^*, is

(15) $$\lambda_1(u_{t-1} - u^*) + \lambda_2(u_{t-2} - u^*).$$

The forecast error, z_t, is then

(16) $$z_t = u_t - u^* - \lambda_1(u_{t-1} - u^*) - \lambda_2(u_{t-2} - u^*).$$

The variance of the forecast error is

(17) $$\sigma_z^2 = (1 + \lambda_1^2 + \lambda_2^2)\sigma_{v'}^2/\phi_1^2 + V(\epsilon_t' - \lambda_1\epsilon_{t-1}' - \lambda_2\epsilon_{t-2}').$$

The term surrounded by the V is the unpredictable part of the random shift, ϵ_t'. Its variance must be positive, so

$$\sigma_z^2 > (1 + \lambda^2 + \lambda_2^2)\sigma_{v'}^2/\phi_1^2,$$

or

$$\sigma_{v'}^2/\phi_1^2 < \frac{\sigma_z^2}{1 + \lambda_1^2 + \lambda_2^2}.$$

For quarterly data for the United States from 1954 through 1974, σ_z^2 is 0.090, with $\lambda_1 = 1.57$ and $\lambda_2 = -0.65$.[16] These results imply that $\sigma_{v'}^2/\phi_1^2$ cannot exceed 0.023. By contrast, the variance of unemployment is 1.35. *Only a trivial fraction, less than 1.7 percent, of the variation of unemployment is attributable to the unemployment-inflation tradeoff in the rational-expectations model.* The remaining 98.3 percent or more is attributable to unexplained shifts of the Phillips curve.

Another Theory of the Phillips Curve

Some other theories of the Phillips curve do not embody explicitly the sharp assumptions of rational expectations, yet claim to explain the response of wages to unemployment as the outcome of the meeting of rational buyers and sellers in the labor market. The model presented in "The Process of Inflation" (hereafter, the PI model) falls into this category. It now appears that the criticism just directed at the rational-expectations model applies to other models as well: they are unable to explain the persistence

16. That is, the standard error of the regression of the unemployment rate on its level lagged once and lagged twice is 0.30. The regression coefficients are 1.57 and −0.65. Very similar results are reported by Sargent, "Rational Expectations," p. 451. A regression is just a convenient way to get reasonable values for λ_1 and λ_2. The resulting bound on $\sigma_{v'}^2/\phi_1^2$ does not require that ϵ_t' actually obey a second-order autoregressive process; it holds for any process.

of unemployment. They treat it as a surprise to participants in the labor market, yet most deviations of the unemployment rate from equilibrium are highly predictable. The point where this defect enters the PI model is particularly easy to identify, but I believe the difficulty exists in all models.

In the PI model, the crucial step in wage determination is the setting of the scale wage by employers. Their goal is a scale wage just high enough to attract the desired grade of worker into each job category. Firms never *expect* upgrading or downgrading of the labor force, since they prefer that these movements not occur and can offset any expected one by an adjustment of the scale wage. Consequently, all movements are totally unpredictable and uncorrelated with past movements or any other information available at the time wages are set. If up- or downgrading has a stable relationship with the unemployment rate, as hypothesized in the PI model, this feature of wage determination implies that firms never expect the unemployment rate to deviate from its equilibrium level. Either firms remain ignorant year after year of widely reported conditions in the labor market, or the relation between the unemployment rate and up- and downgrading of the labor force shifts over time. To maintain its claim to be a theory based on rational behavior, the PI model must make the second assumption. But this makes it exactly like the rational-expectations model, attributing an overwhelming fraction of the variation in unemployment to unexplained changes in the equilibrium unemployment rate (the decomposition of the variance of the unemployment rate in the previous section applies to the PI model as well).

Keynesian Theories of Wage Rigidity

In its modern interpretation, Keynes' *General Theory* is a working out of the macroeconomic implications of the short-run inflexibility of wages. Its major innovation, the theory of effective demand, makes economic sense only when wages are fixed. Keynes accepted wage rigidity and even argued its desirability. Falling wages in a contraction might only further destabilize the economy, he claimed. Chapter 19 of the *General Theory* argues at length against the view that falling nominal wages could cure a depression.[17] The modern evidence reviewed earlier suggests that Keynes was quite mis-

17. The argument rests on two main reasons: (1) wages would fall more than prices, and the marginal propensity to spend profits is less than that for wages; (2) deflation may cause expectations of further deflation.

taken in this respect. The durability of Keynesian economics arises from the truth of the hypothesis of wage rigidity and the usefulness of the theory of effective demand it implies. Keynes' error on the purely hypothetical question of what would happen if wages *were* flexible hardly diminishes the importance of the *General Theory*, because wages are in fact not at all flexible.

To the reader who takes the economic theory of supply and demand seriously, it is curious that the *General Theory* lacks any fundamental explanation for the failure of the wage to vary in order to clear the labor market. The lack has a good historical explanation. The *General Theory* was in large part a reaction against influential books by A. C. Pigou and Lionel Robbins, who attributed unemployment to rigid wages and advocated measures to increase competition in the labor market to alleviate the depression.[18] Keynes' major goal was to demonstrate the inadequacy of this prescription. A lasting contribution of the *General Theory* was its demonstration of the fallacy of the simple view that reduction of the nominal wage would stimulate the demand for labor directly by lowering the real wage. He emphasized that changes in nominal wages have essentially no effect on real wages (a view strongly supported by modern research), and that the stimulus from wage cuts acts more indirectly. Under flexible wages, Keynes wrote,

the wage-unit might have to fall without limit until it reached a point where the effect of the abundance of money in terms of the wage-unit on the rate of interest was sufficient to restore a level of full employment. At no other point could there be a resting-place.[19]

This is an exact description of the central mechanism restoring full employment in, for example, the MPS model.

From his classical predecessors, Keynes inherited the notion that wages were inflexible because they are set by monopolists in the labor market—that is, by labor unions. Since this view is widespread today, it is worth a careful examination. Without doubt, wages are more sluggish when they are negotiated than when they are set in a competitive market. The question is whether concentration in the labor market can account for an important fraction of the persistent movements of unemployment. Can the slow response of unions explain the six straight years of excessive unemployment, 1958 through 1963? Can it explain the seven straight years of low unem-

---

18. Respectively, *The Theory of Unemployment* (London: MacMillan, 1933), and *The Great Depression* (MacMillan, 1934).

19. *General Theory*, p. 253.

ployment from 1964 through 1970? The second question is a particular obstacle for superficially plausible explanations.

The view that labor unions are the major source of the pathological behavior of wages and unemployment has three major flaws:

1. The logic of monopoly does not make wages rigid. Every monopolist faces a negative tradeoff between price and quantity. Even labor unions with full power to dictate wages are restrained by the downward-sloping demand function for their members; if they were not, they would set an infinite wage.

2. Collective bargaining does not in fact fix the wage. Elaborate institutional provisions are made for adaptation to changes in demand not foreseen at the time of negotiation. Effective wages vary in response to demand through upward and downward movements of workers among job categories, variations in overtime, and other mechanisms.

3. Collective-bargaining agreements cover only a fraction of the labor force, perhaps 25 percent. In the rest of the economy, employers are largely unfettered. When demand falls, wages should fall in the nonunion sector until all the excess supply of labor is absorbed and unemployment is restored to its normal level. No matter what they do to union wages, unions cannot cause unemployment unless they control *all* wages.

The first of these flaws involves the deep and unsettled issue of the motivation of labor unions and will not be discussed any further here. The second was developed and documented at considerable length in "The Process of Inflation" and also will receive no further attention here. The third flaw is equally important. Its resolution will occupy most of the latter parts of the paper.

In a standard two-sector model in which the wage is fixed in one sector and flexible in the other, unemployment is impossible. Wages will fall in the competitive sector until unemployment disappears, simply because the unemployed would rather be working at some positive wage than be unemployed at zero wage. Labor unions can cause wage differentials but not unemployment as long as there is a competitive sector able to pay a positive wage. The wage rigidity identified earlier in the paper as the basic source of persistent unemployment must be universal. The presence of persistent unemployment in an economy in which firms employing 75 percent of the work force are unfettered by collective bargaining can be explained only by wage rigidities in the nonunion sector.

Idiosyncratic Exchange in the Labor Market

The theory of idiosyncratic exchange is a very recent attempt to supply a rationale for the apparent rigidity of wages in an atomistic labor market.[20] In the spirit of the microeconomic theory of disequilibrium, this theory attempts to explain wage rigidity as the outcome of the rational activities of buyers and sellers in the labor market, each of whom controls an infinitesimal share of the total demand or supply. However, the theory does not depend on lack of information on the part of either buyers or sellers, and so is not subject to the theoretical or empirical criticisms of the previous section of this paper.

The central hypothesis of the theory of idiosyncratic exchange holds that an important part of the product of a worker is the return to specific human capital. A worker produces more in his present job than he would as an inexperienced employee of another firm. Under competition, there is a "zone of indeterminacy" (in Okun's words) within which the wage can vary: a worker will quit if paid less than the wage of inexperienced workers elsewhere, but the current employer should be willing to pay up to the worker's marginal product, if necessary. To avoid costly bargaining with individual workers over the division of the return to specific capital, institutions have evolved for treating workers equitably and in a way that is well understood from the beginning of their employment. Unexpected wage changes would be a violation of these rules. Wages continue on a smooth trajectory, with labor getting a larger share of the return to specific capital when the market is slack and employers the larger share when demand is strong. Unlike many casual rationalizations of wage rigidity, this theory explains both the failure of wages to rise rapidly in tight markets and their downward rigidity.

The proponents of the theory of idiosyncratic exchange do not claim to explain wage rigidity throughout the economy. For many sectors—trade,

20. The term is due to Oliver E. Williamson, Michael L. Wachter, and Jeffrey E. Harris, "Understanding the Employment Relation: The Analysis of Idiosyncratic Exchange," *Bell Journal of Economics*, vol. 6 (Spring 1975), pp. 250–78, which draws extensively on Peter B. Doeringer and Michael J. Piore, *Internal Labor Markets and Manpower Analysis* (Heath, 1971). Arthur Okun applies this body of thought to the problem of wage rigidity in his paper in this issue, "Inflation: Its Mechanics and Welfare Costs." Another related attempt to create a theory of rigidity is John Hicks, *The Crisis in Keynesian Economics* (Basic Books, 1974), pt. 3, "Wages and Inflation."

services, construction are important examples—and many occupations—
some professionals, clerical workers, drivers, craft workers, laborers—the
character of the employment relation is rather different from that described
by the theory. As in the case of unions, the theory that wage rigidity results
from idiosyncratic exchange requires a complementary theory of trans-
mission to the competitive sector.

Empirical Evidence on Wage Rigidity by Sector

The importance of the various considerations that account for rigid
wages in particular sectors is a matter for empirical determination. To
study the cyclical behavior of relative wages, I have fitted the following
simple model to wage data for fourteen sectors:

$$(18) \qquad w_{it} - \bar{w}_t = \alpha_i + \lambda_i t + \beta_i (u_t - \bar{u}) + \epsilon_{it},$$

where

$\quad w_{it}$ = the log of the annual earnings of full-time equivalent workers
$\qquad\quad$ in sector i and year t

$\quad \bar{w}_t$ = the average level of the log of wages in year t

$\quad \alpha_i$ = the permanent differential of sector i from the common level

$\quad \lambda_i$ = the trend relative to the average

$\quad \beta_i$ = the cyclical response of the differential

$u_t - \bar{u}$ = the departure of the log of the aggregate fixed-weight unem-
$\qquad\quad$ ployment rate[21] from its average level

$\quad \epsilon_{it}$ = a random disturbance.

Estimates for the model appear in table 1.

Industries with high values of β have unresponsive wages. The largest β
is for the federal government, with medicine and education, transportation,
and communication (almost entirely the telephone industry) not far behind.
The remaining regulated industry, utilities, has a smaller but still unam-
biguously positive β. In five industries—state and local government, min-
ing, durables and nondurables manufacturing, and construction, the hy-
pothesis that β is zero cannot be rejected statistically. The cyclical behavior
of wages in these industries roughly parallels the economy-wide average.
The remaining four sectors, retail trade, wholesale trade, finance, insur-

21. See "Process of Inflation," app. D, p. 392.

Table 1. Regression Results for Cyclical Behavior of Wage Differentials, by Major Industry, 1962–73[a]

Industry	Constant differential α	Cyclical differential β	Time trend λ	Standard error
Nonentrepreneurial				
Federal government	0.35	0.076	0.011	0.013
	(0.01)	(0.017)	(0.001)	
Medicine and education	−0.26	0.055	0.013	0.006
	(0.01)	(0.008)	(0.001)	
Transportation	0.31	0.044	0.006	0.011
	(0.01)	(0.015)	(0.001)	
Communication	0.12	0.044	0.000	0.021
	(0.02)	(0.028)	(0.002)	
Utilities	0.24	0.018	0.000	0.003
	(0.00)*	(0.005)	(0.000)	
State and local government	0.05	0.004	0.006	0.009
	(0.01)	(0.012)	(0.001)	
Entrepreneurial				
Mining	0.22	0.012	0.005	0.007
	(0.01)	(0.010)	(0.001)	
Nondurables manufacturing	−0.05	−0.005	−0.005	0.002
	(0.00)*	(0.003)	(0.000)†	
Durables manufacturing	0.12	−0.007	−0.007	0.006
	(0.00)*	(0.008)	(0.000)†	
Construction	0.21	−0.018	0.004	0.019
	(0.01)	(0.025)	(0.002)	
Retail trade	−0.30	−0.018	−0.011	0.004
	(0.00)*	(0.005)	(0.000)†	
Wholesale trade	0.21	−0.021	−0.004	0.003
	(0.00)*	(0.004)	(0.000)†	
Finance, insurance, and real estate	0.06	−0.029	−0.001	0.003
	(0.01)	(0.011)	(0.001)	
Services	−0.19	−0.038	0.000	0.010
	(0.01)	(0.013)	(0.001)	

Sources: Derived from text equation (18). The basic wage data are from the national income and product accounts, table 6.5, "average annual earnings per full-time employee by industry," appearing in *Survey of Current Business,* vol. 54 (July 1974), p. 37, and the various summary volumes of the national income and product accounts published by the U.S. Department of Commerce.

a. The numbers in parentheses are standard errors.

* Less than 0.005.

† Less than 0.0005.

ance, and real estate, and services (other than medical-educational), have wages that are clearly more responsive than the average.

The view of cyclical variations in wage differentials that emerges from table 1 is quite different from the stereotypes of most theoretical discus-

sions. Capital-intensive unionized industries are hardly the major source of rigid wages. On the other hand, the results are not totally incompatible with the two theories of sectoral rigidity discussed earlier. Several of the rigid-wage industries are highly unionized, while none of the clearly flexible industries is substantially unionized. The one industry with an unambiguously high component of specific human capital, communication, does in fact have inflexible wages. However, the most notable feature of the results is suggested by the ranking of industries by their wage responsiveness: nonbusinesses are unresponsive and businesses are responsive. The nonentrepreneurial sector—government, nonprofit institutions, and regulated industries—is precisely the rigid-wage sector. My main purpose here is simply to point out the role of this sector in wage determination, not to explain its behavior from more fundamental considerations. Doubtless, a variety of explanations are important, including the degree of unionization of the sector and the nature of the relations between workers and employers. My own belief is that the insulation of this sector from arbitrage is paramount in the explanation. In the entrepreneurial sector, a business that does not take full advantage of the opportunity to reduce wage costs in a recession, or to keep its workers in a boom by raising wages speedily, will be displaced by a more profitable firm that has a more timely and aggressive wage policy. But nobody is permitted to displace governments or the telephone company, whose wage policies are thus insulated from the basic force of competition.

The Transmission of Wage Rigidity

Previous sections have discussed two theoretical reasons for wage rigidity in specific sectors of the economy: collective bargaining in the unionized sector, and institutions for dividing the return to specific capital in the idiosyncratic-exchange sector. Quantitatively, rigid wages are most important in the nonentrepreneurial sector, which employs 34 percent of the labor force. As emphasized earlier, the mere existence of any of these sectors does not explain the rigidity of the overall wage level or the persistence of unemployment. Workers who cannot find jobs in any of the three overlapping rigid-wage sectors should take jobs that require little specific capital with nonunion entrepreneurial firms. There are many such jobs and there would be even more if the unemployed were to bid the wage

Competitive wage S

Total employment
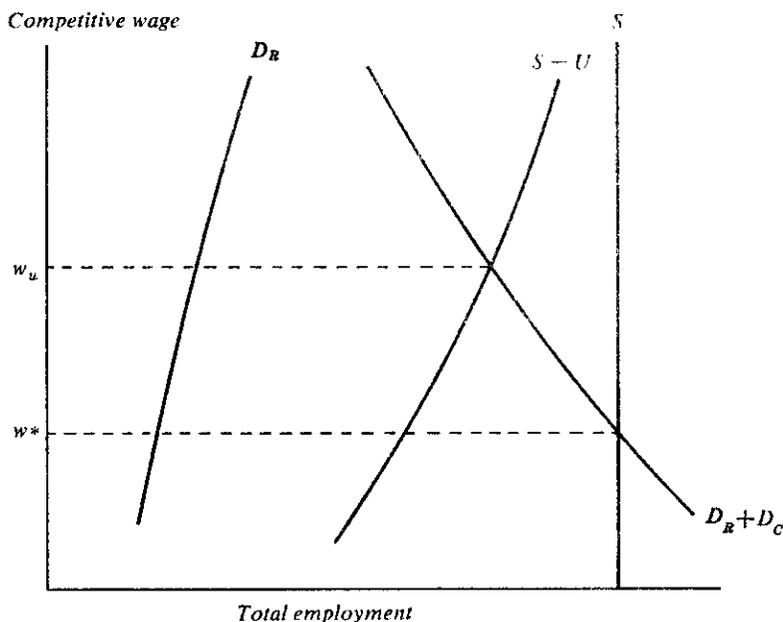
down. Why does the competitive sector fail to absorb the unemployed? This section seeks an answer within a theory of wage rigidity in the residual or competitive sector induced by the presence of the rigid-wage sector.

The essential argument that wage rigidity is transmitted from one sector to another appears in the accompanying diagram. Here the wage in the rigid-wage sector is considered a predetermined constant, while supply and demand determine the competitive wage conditional on the value of the rigid wage. The demand for labor in the rigid-wage sector (D_R) slopes upward to the right because the vertical axis gives the competitive wage level. If the labor market cleared in the ordinary way, then the competitive wage would be w^*, the point where the demand schedule for both the rigid-wage and competitive sectors ($D_R + D_C$) crosses the supply schedule for labor (S). Under ordinary market clearing, a higher rigid wage would imply a lower competitive wage, since the total demand schedule would lie to the left. The competitive wage compensates for the inappropriate level of the rigid wage. This is the basic argument, sketched earlier in the paper, that a rigid wage in one sector does not make the entire economy function as if the average wage level were rigid.

Now, in fact, the labor market does not clear in the normal way. Unemployment is always present. If unemployment were simply a constant frac-

tion of the labor force for frictional reasons, then the previous analysis would hold with no important modification. The possibility considered in the diagram, however, is that the amount of unemployment (U) may depend on the level of the competitive wage. This makes the net supply of labor $(S - U)$ positively wage elastic, and makes the competitive wage (w_u) less responsive to changes in the demand for labor. In particular, higher levels of the rigid wage are no longer fully compensated by lower levels of the competitive wage; rather, the higher the rigid wage, the higher is the unemployment rate. This is the basic theory of the role of a rigid-wage sector in causing overall wage rigidity and excessive unemployment that is pursued in the remainder of this paper.

The proposition illustrated in the diagram clearly rests on a theory of unemployment on the supply side. Before developing that theory in detail, I must emphasize that a theory of unemployment on the supply side is not a substitute or rival for the view that unemployment is a function of the level of demand for labor. The unemployment rate is determined by the interaction of supply and demand forces. The diagram illustrates a model within the framework developed in the introductory section of the paper: the demand schedule is the effective-demand function considered as a function of the wage rate, and the unemployment schedule is the wage-adjustment function, considered now as a relation between the unemployment rate and the current level of the wage. In fact, except for the distinction between the two sectors, the diagram illustrates the standard contemporary macroeconomic model with an aggregate-demand function and a Phillips curve.

A substantial theoretical literature supports the hypothesis that unemployment rises when wage differentials widen.[22] The theory rests on the view that unemployment is the result of conscious decisions by job seekers about the use of their time. According to this search theory, instead of

22. Much of it starts from John R. Harris and Michael P. Todaro, "Migration, Unemployment and Development: A Two-Sector Analysis," *American Economic Review*, vol. 60 (March 1970), pp. 126–43; see especially Stephen A. Ross and Michael L. Wachter, "Wage Determination, Inflation, and the Industrial Structure," *American Economic Review*, vol. 63 (September 1973), pp. 675–92. A recent contribution with many other references is B. Curtis Eaton and Philip A. Neher, "Unemployment, Underemployment, and Optimal Job Search," *Journal of Political Economy*, vol. 83 (April 1975), pp. 355–75. Empirical support for the hypothesis that wage differentials respond to the unemployment rate is found in Michael Wachter's study of the cyclical behavior of wages in the manufacturing sector. See "Cyclical Variations in the Interindustry Wage Structure," *American Economic Review*, vol. 60 (March 1970), pp. 75–84. These results for the manufacturing sector are quantitatively similar to mine for the entire economy.

taking the first job that becomes available, the unemployed decide whether a better job might develop in the future, for which it might pay to wait.[23] When wage differentials widen, the return to waiting increases on the average, because the difference between the wage for a good job and the wage of the average job increases. In a recession, the unemployed do not take jobs immediately in the competitive sector at lower wages. Rather, they are aware that jobs still exist in the rigid-wage sector and some of them will decide to remain unemployed until they find a job there. The wider the differential, the higher is the payoff to waiting.

THE SUPPLY OF UNEMPLOYMENT

Rational behavior on the part of the unemployed involves sampling from the universe of available jobs and deciding after examining each prospect whether to take it or reject it in the hope that a better one will become available. Unemployment is a productive activity to the individual, so it does only slight violence to the economist's vocabulary to discuss the supply of unemployment.[24] Earlier, in "The Process of Inflation," I set out the details of a representative model of the behavior of the unemployed. There, the dispersion of prospective wages facing the unemployed was taken as a constant. The only important modification pursued here (and in more detail in the appendix) is the recognition that one of the determinants of the value of waiting to the unemployed is the probability of locating a job in the rigid-wage sector, and that the desirability of such a job depends on the relative wages paid by the two sectors.

Suppose workers face a normal distribution of wage prospects in the competitive sector with mean μ_1 and standard deviation σ_1, and similarly, a normal distribution in the rigid-wage sector with parameters μ_2 and σ_2. Workers taking jobs in the two sectors receive wages of w_1 and w_2 on the average. An unemployed worker has a daily probability, p, of locating a prospect, of which a fraction, α, comes from the rigid-wage sector while

23. Many economists have misgivings about this view. See, for example, Robert Solow's comments on this paper. Some of the misgivings may arise from the tendency of proponents of the search theory to portray it as an alternative to demand-oriented theories of unemployment.

24. It would be more accurate to discuss the theory of unemployment as one aspect of the complete theory of the allocation of individual time that underlies the net labor supply schedule in the diagram. Though individual workers supply unemployment, employers do not demand unemployment. Rather, the market clears when employers demand exactly the supply of labor net of the supply of unemployment.

$1 - \alpha$ comes from the competitive sector. Both p and α are determined endogenously in the model. In slack markets, all jobs are hard to find (p is low) and jobs in the rigid-wage sector are particularly hard to find (α is low as well).

Table 2 presents the unemployment rates consistent with various relative wages for a hypothetical group in the labor force. These should be considered as alternative points on the unemployment–competitive-wage relation in the diagram. When the typical job prospect in the competitive sector pays only $4.00 per hour, against $5.00 in the other sector, the supply of unemployment is high—11.3 percent of the labor force. The return to holding out for a good job is so high that many workers do so, even though the cost is an extended spell of joblessness. In spite of the high unemployment rate, competitive employers are not swamped with applicants ready to take jobs below the prevailing wage. The average wage paid to a worker taking a job in the competitive sector is $4.14, well above the average prospect. Competitive employers would suffer if they tried to depress the wage further, because even fewer of their offers would be accepted. In other respects, the labor market shows the usual signs of slack conditions—prospects are difficult to locate, arriving on the average about once every eight weeks. Only about one in ten of the prospects located by an individual comes from the rigid, high-wage sector. Two-thirds of the unemployed wind up in the competitive sector. At the other end of the supply schedule, where the two sectors offer identical wages, the unemployment rate is only 1.6 percent. Prospects arrive much more frequently, close to 1.5 per week. Because the rigid-wage sector is no longer a magnet for the unemployed, the mix of prospects available to an individual is the same as the mix of jobs being filled. At all points on the supply schedule, the unemployed are making the best of the situation, given the opportunities available in the two sectors. None of the wage-unemployment combinations depends on imperfect knowledge of actual conditions in the labor market on the part of the unemployed. If the differential between the wages of the two sectors persists, the supply of unemployment will persist as well.

EMPIRICAL APPLICATION

In terms of the empirical results on relative wages presented in table 1, for wage differentials to widen in contractions and narrow in expansions,

Table 2. Measures of Labor-Market Conditions for Alternative Levels of the Competitive Wage

Mean of prospective wage distribution in competitive sector, μ_1 (dollars)	Average wage		Unemployment rate, u (percent)	Prospective jobs	
	Competitive, w_1 (dollars)	Rigid, w_2 (dollars)		Fraction in rigid sector, α (percent)	Weekly frequency of a prospect, p
4.00	4.14	5.00	11.3	9.8	0.136
4.20	4.34	5.00	8.8	9.8	0.179
4.40	4.54	5.00	6.5	9.8	0.249
4.60	4.74	5.00	4.3	9.8	0.385
4.80	4.94	4.97	2.0	11.0	0.857
5.00	5.14	5.14	1.6	34.0	1.450

Source: Derived by author. For explanation, see text and appendix.

the high-β industries must also be high-α industries. Table 1 suggests that this is the case, but the evidence is not conclusive. The pattern of differentials facing a particular worker is somewhat different from the pattern of the αs, because the αs index the variation in occupational mix across industries as well as pure industry differentials. Table 3 attempts to adjust the gross differentials of table 1, shown in column (1), for the effects of variations in quality across industries. The quality index in column (2) of table 3 was derived by estimating relative occupational earnings from 1970 census data and then weighting these by the composition of employment in each industry. The net industry differentials in column (3) suggest that part, but by no means all, of the gross industry differentials are attributable to variations in occupational mix. The high pay of workers in the federal government and regulated industries is largely the result of policies of paying more than the going wage for each occupation and to a smaller extent the result of hiring a disproportionate share of workers in high-paying occupations. By contrast, high wages among unregulated businesses—mining, construction, and wholesale trade—reflect quality differentials to a greater extent. The only conspicuous failure of the positive association of α and β is in the medical-educational sector. This reflects the low pay of nurses, technicians, and private-school teachers relative to their counterparts in other sectors. The basic difficulty of this kind of quality adjustment is its inability to deal with "equalizing differences" across industries: presumably the medical-educational sector is able to hold

**Table 3. Wage Differentials before and after Adjustment
for Occupational Mix, by Major Industry, 1969**

Percent

Industry	Gross differential (1)	Quality differential (2)	Net or pure differential (3)
Nonentrepreneurial			
Federal government	35	5	30
Medicine and education	−26	−1	−25
Transportation	31	2	29
Communication	12	2	10
Utilities	24	8	16
State and local government	5	5	0
Entrepreneurial			
Mining	22	19	3
Nondurables manufacturing	−5	−9	4
Durables manufacturing	12	4	8
Construction	21	13	8
Retail trade	−30	−14	−16
Wholesale trade	21	14	7
Finance, insurance, and real estate	6	5	1
Services	−19	−4	−15

Sources: Occupational wage differentials are regression estimates from data in U.S. Bureau of the Census, *Census of Population, 1970: Occupation by Industry*, Final Report PC(2)-7C (1972), table 4; industry quality is computed from industry occupational composition in ibid., table 1. Column (1) is from table 1 above, first column.

its labor force in spite of its extremely low wage because of the non-pecuniary returns to working in the sector.

With some reservations about the success of the quality adjustments, I have aggregated the fourteen sectors of table 1 into two sectors—the entrepreneurial, or competitive, sector, and the nonentrepreneurial, or rigid-wage, sector—to get the following two-sector relation:

(19) $$w_2 - w_1 = 0.089 + 0.057\,(\log u - \log \bar{u}).$$

For the various unemployment rates in table 4, the percent wage differential between the two sectors implied by this estimated relation appears in column (1). In extremely tight markets (comparable to 1969), the differential is 8.9 percent. At the postwar average (fixed-weight unemployment index at 4.5 percent), the differential is 12.3 percent, and in very slack markets (comparable to 1975), the differential rises to 15.2 percent. None of these differentials is large, and the cyclical response is not sharp. However, the supply of unemployment is sufficiently sensitive to the differential,

Table 4. Wage Differential at Alternative Unemployment Rates, and Unemployment Rates Predicted by the Search Theory

Percent

Unemployment rate[a]	Estimated relation[b]		Adjusted relation[c]	
	Differential between entrepreneurial and non-entrepreneurial sectors (1)	Predicted unemployment rate (2)	Differential between entrepreneurial and non-entrepreneurial sectors (3)	Predicted unemployment rate (4)
2.5	8.9	6.3	2.2	2.7
3.5	10.8	7.4	4.1	3.7
4.5	12.3	8.3	5.6	4.5
5.5	13.4	8.9	6.7	5.1
6.5	14.3	9.5	7.6	5.6
7.5	15.2	10.0	8.5	6.1

Sources: The estimated relation is derived from text equation (19). For the adjusted relation, the constant in equation (19) was reduced so that the predicted wage differential at 4.5 percent unemployment gave a 4.5 percent prediction from the unemployment-supply model.

a. Fixed-weight unemployment rate.

b. $w_2 - w_1 = 0.089 + 0.057 (\log u - \log \bar{u})$, where the symbols are as defined for text equation (19).

c. $w_2 - w_1 = 0.022 + 0.057 (\log u - \log \bar{u})$, which is the equation in note b adjusted as described above.

according to the model, that these differentials are *too* large to be consistent with the unemployment rates in the table. Column (2) gives the unemployment rates predicted by the model; all are about 3 to 4 percentage points too high. Probably the main source of disagreement is the inadequate adjustment for quality variations across industries, though many of the unrealistic simplifications of the model may also contribute to it. It is reasonable to suppose that the quality differentials are stable from year to year, so that the cyclical movements of the differential are measured more accurately than is the level. As a rough adjustment for the error in the level, I have reduced the constant in the two-sector equation far enough so that the predicted wage differential at 4.5 percent unemployment generates a prediction of 4.5 percent from the unemployment-supply model. The adjusted wage differential for the various unemployment rates appears in column (3) of table 4, and the predictions of the supply model in column (4). The predicted unemployment is slightly high at low unemployment rates and somewhat low at high rates. Over the range of unemployment rates, the response of the predicted rate is slightly more than two-thirds of the actual change in unemployment. The evidence suggests that the

observed relation between the wage differential and the unemployment rate is roughly similar to the relation predicted by the model of unemployment supply. There is a quantitative justification for the elasticity of the net labor supply schedule in the diagram with respect to the competitive wage.

The facts about relative wages and the unemployment rate seem consistent with the basic view expressed in this paper: There is a sector with rigid wages, employing an important minority of the labor force. When demand falls, the competitive wage does not fall enough to provide employment for the entire labor force. Rather, part of the labor force joins the queues for good jobs in the rigid-wage sector rather than accept lower-paying work in the competitive sector. When demand is strong, the competitive sector is able to bid labor away from the other sector. This result is compatible with supply behavior because the increase in labor demand also closes the gap between the two wages and reduces the incentive to wait for a job in the rigid-wage sector.

Aggregate Economic Policy with a Predetermined Wage in the Nonentrepreneurial Sector

The unemployment supply equation is

$$(20) \qquad w_2 - w_1 = \delta_0 + \delta_1 u,$$

with $\delta_1 = 0.057$. If w_2 is considered a predetermined variable in the economy, then this equation plays the same role as the wage-adjustment equation in the simple two-equation model of the introduction. The effective-demand equation can be written in terms of the weighted average of the two wages:

$$(21) \qquad u = \psi_0 - \psi_1[m - (1 - \theta) w_1 - \theta w_2];$$

θ is 0.34, the fraction of the labor force receiving the predetermined wage.

The reduced form of the system is

$$(22) \qquad u_t = \frac{\psi_0 - \psi_1(1 - \theta) \delta_0 - \psi_1(m_t - w_{2t})}{1 + \psi_1(1 - \theta) \delta_1} ;$$

$$(23) \qquad w_{1t} = w_{2t} - \delta_0 - \delta_1 \frac{\psi_0 - \psi_1(1 - \theta) \delta_0 - \psi_1(m_t - w_{2t})}{1 + \psi_1(1 - \theta) \delta_1}.$$

Each value of m_t corresponds to a different point on the wage-unemployment tradeoff. The slope of the tradeoff between w_{1t} and u_t is just δ_1, so

the slope between w_t (the weighted average of w_{1t} and w_{2t}) and u_t is $(1 - \theta)\,\delta_1$. If $\theta = 0.34$ and $\delta_1 = 0.057$, as suggested earlier, then the slope is 0.8 percentage point of wage change for each percentage point of unemployment. Monetary authorities face a hard decision because the trade-off is so flat: expansion sharply reduces unemployment today, but at the cost of future inflation after w_{2t} begins to respond. Contraction pulls the wage down disappointingly little considering the increased unemployment it brings about. In this model, rigidity of 34 percent of the wages is sufficient to explain the sluggish behavior of all wages, because the optimizing job search of the unemployed forces competitive employers to moderate wage adjustments. The pessimistic conclusions of modern Keynesians are given a more solid theoretical foundation by this view.

If policy could somehow influence the level of the nonentrepreneurial wage, w_{2t}, the conclusion would be less pessimistic. The reduced-form equation for the unemployment rate shows that a reduction of 1 percentage point in w_2 offers exactly the same stimulus as an increase of 1 percentage point in the money supply: each reduces the unemployment rate by about 0.25 percentage point, assuming $\psi_1 = 5.5$. Moreover, the expansion is accompanied by a fall in wages and thus in prices. Instead of facing a tradeoff between inflation and unemployment, authorities who could reduce w_2 could achieve an unambiguous improvement in both dimensions at the same time. The effect of a change in w_2 on the average wage level is

$$(24) \qquad \frac{\partial w}{\partial w_2} = (1 - \theta)\left(1 - \frac{\delta_1\psi_1}{1 + \psi_1(1 - \theta)\,\delta_1}\right) + \theta$$
$$= 0.83.$$

In contrast, the effect of a change in the money supply on the average wage level is

$$(25) \qquad \frac{\partial w}{\partial m} = (1 - \theta)\frac{\delta_1\psi_1}{1 + \psi_1(1 - \theta)\,\delta_1}$$
$$= 0.17.$$

A reduction of 1 percentage point in w_2 yields almost five times as much deflation as a like reduction in the money supply. But the two policies have opposite effects on real output. Lowering w_2 stimulates real output, so wages fall by more than the decline in nominal output. Contractionary monetary policy, on the other hand, reduces real output and depresses wages by less than the fall in nominal output.

Concluding Remarks

The rigidity of wages in the face of a predetermined money supply and fiscal policy is the source of the persistence of unemployment. Disequilibrium and imperfect information are only a small part of the story. It seems to be a fact that wages are rigid in one sector of the labor market, though debate about the importance of competing explanations for this rigidity will doubtless continue. In addition to documenting the location and extent of the rigidity, this paper has contributed an explanation of the spillover of rigidity to the residual competitive sector of the market. Anti-inflationary policies are largely thwarted by their lack of effect on the rigid wage or the flexible wage, so they depress real output instead of moderating inflation. If the government could manipulate the rigid wage directly, real output could be expanded at the same time that inflation was brought under control. Alternatively, if currently rigid wages could be made more responsive to conditions in the labor market, monetary policy would have a larger anti-inflationary effect and a smaller contractionary effect on real output. The scope for federal action in either direction is narrow. Strengthening the provisions for linking federal to private wages and deregulation in the transportation and utilities sectors offer some hope. In general, however, the diagnosis of this paper offers relatively poor prospects for movement toward an economy in which the wage-price mechanism is free to do its job in protecting real output from severe contractions. Federal policymakers can probably look forward to a continuation of their power to affect real output and of their inability to have much effect on wages and prices.

APPENDIX

The Supply of Unemployment Under Variable Wage Differentials

THIS APPENDIX develops a search model that takes explicit account of the importance of sectoral wage differentials on the value of time spent looking

for work. The fundamental precept of any search model is that the unemployed are waiting for jobs to become available and that this behavior is rational from the point of view of the individual worker. The time may be spent in active search, but more often the unemployed wait passively for information to reach them through well-established channels.[25] It is sometimes best to remain unemployed even after locating a job prospect because a better prospect may become available later. Besides, as Martin Feldstein has pointed out, many workers who are unemployed as a result of layoffs have good chances of returning to their original jobs.[26]

Descriptions of search models have emphasized the diversity of job prospects that may be available to an unemployed worker. Virtually all models view the unemployed as sampling from a probability distribution of prospective wages. Under certain conditions, the optimal search rule is to establish a reservation wage, w^*, and to accept the first job paying more than w^*.[27] A representative model appears in "The Process of Inflation," appendix C. The first step here is to specify a probabilistic characterization of the prospects available in a labor market with two kinds of employers. Suppose workers face a normal distribution of wage prospects in the competitive sector with mean μ_1 and standard deviation σ_1, and similarly face a normal distribution in the high-wage sector with parameters μ_2 and σ_2.[28] Prospects arrive randomly at a weekly rate, p, and the probability that a given prospect comes from the high-wage sector is α. Both p and α are *endogenous in the model. The overall distribution of prospects is not normal*, but a "mixture" of two normals.

The unemployed worker establishes a reservation wage, w^*, and accepts the first job in either sector that pays at least that much. Workers taking

25. For a comprehensive review of evidence on this point, see Robert J. Gordon, "The Welfare Cost of Higher Unemployment," *BPEA, 1:1973*, app. C, pp. 188–95. Gordon concludes that the typical unemployed worker spends 19 percent of his normal working time in job-seeking activities. Less than half of this time is spent in active search; the rest is spent reading want ads and in other passive activities.

26. "Temporary Layoffs in the Theory of Unemployment" (Harvard Institute of Economic Research, July 1975; processed).

27. See Meir G. Kohn and Steven Shavell, "The Theory of Search," *Journal of Economic Theory*, vol. 9 (October 1974), pp. 93–123. Michael Rothschild has investigated cases in which optimal search rules are more complicated, but concludes that qualitative behavior is largely the same under them; see "Searching for the Lowest Price When the Distribution of Prices Is Unknown," *Journal of Political Economy*, vol. 82 (July–August 1974), pp. 689–711.

28. The normal distribution seems a reasonable approximation. It is realistic in assigning very low probability to extremely good wage prospects.

jobs in the two sectors receive wages of w_1 and w_2 on the average (these are the conditional expectations of the two distributions and were called "effective wages" in "The Process of Inflation"). The measured wage differential is $w_2 - w_1$, and is always less than the difference in the means of the two distributions (see table 2). The average wage received by a newly employed worker is

$$w = (1 - \theta) w_1 + \theta w_2,$$

where θ is the fraction of jobs in the rigid-wage sector. A search strategy is optimal when its expected payoff equals the reservation wage.[29] Suppose unemployment compensation pays workers a fraction, z, of their usual wage during periods of search; this "replacement ratio" is apparently about 60 percent for eligible workers and about half the unemployed are eligible, so I will take z to be 30 percent. Then the reservation wage is

$$w^* = [1 - (1 - z) u] w,$$

where u is the unemployment rate.

The probability that a job prospect from the competitive sector will be accepted, q_1, is

$$q_1 = 1 - \Phi \left(\frac{w^* - \mu_1}{\sigma_1} \right),$$

where Φ is the cumulative normal distribution, μ_1 is the mean of the wage distribution for the noncontractual sector, and σ_1 is the standard deviation for that sector; for the competitive sector this probability is

$$q_2 = 1 - \Phi \left(\frac{w^* - \mu_2}{\sigma_2} \right),$$

where μ_2 and σ_2 are the wage mean and standard deviation for the contractual sector. The fraction of job prospects from the high-wage sector is α, so the probability that a prospect that is taken came from the high-wage sector is

$$\theta = \frac{\alpha q_2}{(1 - \alpha) q_1 + \alpha q_2}.$$

The weekly probability that an unemployed worker will take a job, h, is the product of the probability that a prospect will become available and the probability that it will be accepted:

29. See Kohn and Shavell, "Theory of Search."

$$h = p \,[(1 - \alpha)\,q_1 + \alpha\, q_2].$$

The unemployment rate is

$$u = \frac{\phi}{h + \phi},$$

where ϕ is the weekly probability of becoming unemployed.

To complete the supply model, it is necessary to specify the relation between the competitive wage and the frequency of prospects, as determined by the recruiting procedures of competitive employers. I will assume that employers make wage offers to qualified workers by drawing from a normal distribution with mean μ_1 and standard deviation σ_1. Presumably, the randomness of wage offers derives from underlying uncertainty about the quality of workers, but the model will not attempt to deal with that explicitly. Firms choose μ_1 with full knowledge of workers' search rules. If μ_1 is far below the reservation wage of workers, a large number of offers will have to be made before one is accepted. If it is too high, firms will pay needlessly high wages. The expected cost of hiring and paying one worker is

$$E(c) = c_0 \frac{1}{q_1\!\left(\dfrac{w^* - \mu_1}{\sigma_1}\right)} + \frac{40}{\phi}\, E(w \,|\, w \geqq w^*).$$

Here, $q_1 = 1 - \Phi((w^* - \mu_1)/\sigma_1)$ is the probability that a given offer will be accepted; *its inverse is the expected number of offers required to recruit one worker.* The coefficient c_0 is the dollar cost of one offer; $E(w \,|\, w \geqq w^*)$ is the expected hourly wage, taking account of the unwillingness of anyone to work below the reservation wage, w^*; and $40/\phi$ is the total number of hours the worker is expected to remain with the firm, working 40 hours per week. If σ is \$0.10 and c_0 is \$15.00, then cost is minimized when μ is set a little less than one standard deviation below the reservation wage, in which case the expected number of offers to fill one position is five and q is 0.2.

The model contains five parameters: the turnover rate, ϕ, taken as 0.005; the two standard deviations of wage prospects, σ_1 and σ_2, both taken as \$0.10; the fraction of jobs in the rigid-wage sector, θ, taken as 0.34; and the mean of prospects in the rigid-wage sector, μ_2, taken as \$5.00. It contains eleven endogenous variables, μ_1, w^*, q_1, q_2, h, p, α, u, w_1, w_2, and w, related by ten equations. Altogether, it determines a relationship between the unemployment rate and the competitive wage that can be

interpreted as the supply schedule of unemployment. The relationship can be calculated in the following way: Select an arbitrary μ_1. From the requirement that the probability of acceptance in the competitive sector is 0.2, compute w^* (the relation is $w^* = \mu_1 + 0.8\sigma_1$). Compute w_1 and w_2 as the conditional expectations given this reservation wage, and from them the expected wage, w. Finally, compute the unemployment rate from the equation for the optimal reservation wage, as follows:

$$u = \frac{1 - w^*/w}{1 - z}.$$

The other equations can be solved for the remaining variables.