ROBERT E. HALL

Massachusetts Institute of Technology

The Process of Inflation in the Labor Market

CHANGES IN AGGREGATE demand are the fundamental source of changes in the price level. In the long run, the supply of resources determines the volume of real output, the quantity of money determines the nominal value of output, and the price level is the ratio of the two. Economists who disagree with Milton Friedman's famous dictum that "inflation is everywhere and always a monetary phenomenon" have doubts mainly about the mechanism linking monetary expansion to inflation. Underlying my discussion of the role of labor markets in the process of inflation is the hypothesis that an increase in aggregate demand raises employment and reduces unemployment. The economy then moves up and to the left along the Phillips curve and wages start to rise more rapidly. Finally, prices rise in the face of increasing costs. If aggregate demand is stabilized at the new, higher, level, the economy comes to rest with a correspondingly higher wage and price level. The inflationary bulge in real aggregate demand disappears as the process reaches its conclusion.

The intellectually weak link in this description of the process of inflation is the Phillips curve. Most economists regard the negative relation between unemployment and wage inflation as a plausible empirical generalization without a firm grounding in economic theory. Recently offered theories of

343

the Phillips curve have been rejected in fairly strong terms by macro and labor economists.[1] Although empirical students of the Phillips curve are generally unsympathetic to the search theorists' notion of unemployment and inflation, they, too, often portray the unemployed as "bidding down the wage level." Labor economists object even to this view on the reasonable grounds that the labor market simply does not function that way. The unemployed never displace the employed by offering to work at lower wages. They compete with each other for openings for which scale wages are already established through collective bargaining or a bureaucratic personnel policy. In turn, the most important consideration in these processes is the wage level in other industries or firms, not the unemployment rate. Economists most familiar with institutions in the labor market are precisely those who are least convinced by the hypothesis that the unemployed bid down the wage level.

Wages do in fact respond to the unemployment rate. The Phillips curve is well established as an econometric relationship.[2] My purpose in this paper is to try to make economic sense out of it in a way that is equally compatible with what is known about institutions in the labor market and with the fundamental economic principle that individual agents always make the best they can out of their situations. In my view the theory of the Phillips curve need not rest on institutional constraints, money illusion, or other failures of this fundamental principle.

The paper follows the theme of my earlier contributions to *Brookings Papers on Economic Activity* in emphasizing the role of turnover and mobility in the operation of labor markets.[3] Its reconciliation of rigid wage scales with the sensitivity of changes in wages to unemployment runs

1. See, for example, James Tobin's critical remarks in his presidential address before the American Economic Association, "Inflation and Unemployment," *American Economic Review*, Vol. 62 (March 1972), pp. 1–18, and Albert Rees' unfavorable comments on a paper by Robert E. Lucas, Jr., and Leonard A. Rapping—"On Equilibrium in Labor Markets," *Journal of Political Economy*, Vol. 78 (March/April 1970), pp. 306–10. The original Lucas-Rapping paper was "Real Wages, Employment, and Inflation," *Journal of Political Economy*, Vol. 77 (September/October 1969), pp. 721–54.

2. As empirical students of the Phillips curve have come to realize the importance of inflationary expectations in shifting the curve, they have lowered their estimates of the slope of the inflation-unemployment relation. Every modern Phillips curve has some downward slope, however.

3. "Why Is the Unemployment Rate So High at Full Employment?" (3:1970), pp. 369–402; "Prospects for Shifting the Phillips Curve through Manpower Policy" (3:1971), pp. 659–701; "Turnover in the Labor Force" (3:1972), pp. 709–56.

as follows: Rigid scales can coexist with rapid changes in wages because workers change jobs frequently. The unemployed are not an inert mass exerting an inexplicable downward force on wages, but a constantly changing group of individuals moving from one job to another. If the labor market is tight, the unemployed tend to find jobs quickly, and, on the average, to get higher-paying jobs than they held previously. In slack markets the opposite is true. Conditions of supply and demand determine the unemployment rate and the rate of wage inflation simultaneously, and the Phillips curve emerges as the locus of alternative combinations of the two.

Plan of the Paper

The first section of the paper develops a two-equation model of the process governing the evolution of the wage level. The first equation deals with the adjustment of the scale wage. It takes the form of a difference equation in the log of the wage with an additive term that varies in response to the unemployment rate. This equation allows direct expression of the accelerationist view that inflation will rise without limit if the unemployment rate is held below an equilibrium level. Under a certain testable restriction on its coefficients, the difference equation has the accelerationist property. The second equation explains the observed average wage as a distributed lag of past values of the scale wage multiplied by a wage-adjustment function. The two equations together predict the course of the observed wage, given values of the unemployment rate which is taken as determined exogenously by aggregate-demand policy.

I have estimated the parameters of the model in two distinct bodies of data. The first pertains to a group of highly unionized manufacturing industries, where scale wages can be observed directly. The second covers the whole private nonagricultural economy, where the scale wage must be inferred. The two sets of estimates agree on the essential points: the process of wage inflation has a pronounced accelerationist tendency, and the strict accelerationist property is consistent with the statistical evidence. The estimated equilibrium level of unemployment is fairly high—5.5 percent of the labor force at its 1974 composition. Following George Perry, I use an unemployment rate that adjusts for the secular shift in the labor force toward groups with high unemployment rates. I then identify three

elements of the adverse combination of unemployment and inflation facing
policymakers today: (1) the increase in the official unemployment rate for
a given degree of inflationary pressure associated with the secular shift in
the composition of the labor force, which has added 0.6 percentage point
to the unemployment rate since 1964; (2) the inheritance of inflationary
momentum from the extremely tight markets of 1965–69, amounting to
about 5 percentage points of inflation; and (3) the further inheritance of
inflationary momentum from 1970–73, amounting to about 2 percentage
points. The third component is the most surprising: in 1970, 1972, and
1973, labor markets were tight enough to *add* to inflation. A reinterpreta-
tion of recent monetary and fiscal policy emerges: it is not that the "old-
time religion" of contraction was tried and found wanting—it was hardly
tried at all. In the concluding section I consider four alternative policies
for the next decade. One ratifies existing inflation for the indefinite future
by holding unemployment at its equilibrium level of 5.5 percent of the labor
force. Two contractionary policies achieve 3.2 percent wage inflation (the
rate consistent with little or no price inflation), one in six years with un-
employment at the 1961 level, and the other in ten years with unemploy-
ment at the 1962 level. Restoration of the wage stability that existed in the
early sixties would be a painful process under either policy. The fourth
policy sets the unemployment rate at its average level for 1961–68 and
brings about a mild acceleration of wage inflation in 1976 and beyond.
None of the four policies yields a combination of inflation and unemploy-
ment over the decade 1975–84 anywhere close to the stated goals of
national policymakers.

Theory of Changes in Money Wages

The integrated theory of the response of the wage level to the supply
and demand for labor developed in the first half of the paper draws exten-
sively on the thinking of labor economists, notably John Dunlop and
Melvin Reder. They have pointed out the rigidity of the scale wage, the
implausibility of a model in which potential workers negotiate with em-
ployers over wages, and the importance of upgrading and downgrading
in response to conditions in the labor market. The theory draws equally
on the work of two macroeconomists, Charles Holt and James Tobin, in

recognizing the importance of mobility in wage change, and of the distinction between the equilibrium and disequilibrium components, a concept that appears in the European literature on wage drift as well. Appendix A presents a systematic discussion of the relation of the theory to its antecedents.

The theory starts from the observation that, unlike the markets for products or financial instruments, the labor market is characterized by extreme heterogeneity. Because many different kinds of workers might be hired for any particular job, it is necessary to make a distinction between the scale wage and the effective wage. The first is the instrument of bureaucratic wage policies and, where workers are organized, the subject of collective bargaining; it can be measured as the wage paid in the typical job. The second is the cost to the firm of adding enough labor to produce one additional unit of output, and can be measured as the wage received by the typical worker. A major point of the theory, documented below, is that the scale and effective wages lead separate lives in the short run. I will argue that the unemployed are unable to bid down the scale wage but that because of the mobility of labor among employers and among jobs, they can bid down the effective wage so that it declines relative to the scale wage in slack labor markets. The essence of the Phillips curve, considered as a relation between the unemployment rate and the rate of change of the effective wage, is that workers move from lower- to higher-paying jobs when the unemployment rate is low and the labor market is tight, and from higher- to lower-paying jobs when the market is slack. These movements cause changes in the effective wage even though the scale wage remains fixed.

JOBS, WORKERS, AND THE SCALE AND EFFECTIVE WAGES

Consider a typical employer with a wide variety of job categories and an equally wide variety of grades of workers. A job category is defined by the type of work performed and not by the specific characteristics of the workers holding it, and a grade of worker is defined by personal characteristics (age, experience, education, training, and the like) and not by the job held. Productivity is the result of an interaction of the characteristics of workers and jobs. A particular grade of worker has a comparative advantage in certain categories of jobs relative to others. Workers of

middle grades will not have the skills to function effectively in more demanding jobs, and would find their own skills underutilized in lower categories. I will suppose that there are M grades of workers, indexed by i, and N categories of jobs, indexed by j. Then I define $c(j)$ as the grade of worker having a comparative advantage in job category j. Many different js have the same value of $c(j)$. I also assume that the grades run in order of productivity from low to high, and that adjacent grades are reasonably close substitutes for one another—in other words, the comparative advantage of grade i over grade $i + 1$ or $i - 1$ in a particular job is slight. Finally, I assume that the wages for workers of grades $1, \ldots, M$ stand in proportions $r_1, \ldots, r_M$. The ratios r_i define a "wage structure" that I assume to be reasonably stable in the short run. In terms of job categories, the relative wage of category j is $r_{c(j)}$. A wage structure among grades of workers implies a wage structure among categories of jobs.

The theory characterizes the wage and hiring policies of the firm in the following way: The firm sets the wage by defining the scale wage, s, and offering wage $sr_{c(j)}$ for job category j. The firm varies wages by adjusting the scale wage in such a way that all wages move in the proportions given by the wage structure. Having established these wages, the firm attempts to fill its needs for labor by a search procedure that takes the most qualified workers available. In tight markets, there may not be any grade $c(j)$ workers interested in working in category j jobs, so the firm must settle for a suitably larger number of workers of lower grades. In slack markets, workers of grades higher than $c(j)$ may be available for category j.

The firm sets the scale wage in advance so as to try to hit the target of hiring grade $c(j)$ workers into category j jobs. After the fact, there may be slippage in one direction or the other. A convenient way to measure this slippage is to define the effective wage, w, as the scale wage that would have hit the target, on the average, in all categories of employment. Let w_i be the wage actually received by workers of grade i. The wage offered for category j jobs should have been $w_{c(j)}$, the wage actually received by workers of grade $c(j)$. Thus w should obey

$$wr_{c(j)} = w_{c(j)}$$

for each job, j. In terms of workers rather than jobs, this implies

$$wr_i = w_i$$

for each grade, i. This cannot be satisfied literally for each grade, but if the

wage structure is reasonably stable, a w that satisfies it in some average sense will come close for each grade. The natural way to average is to use weights corresponding to the employment of each grade of worker, say L_i. Then w should obey

$$w \sum r_i L_i = \sum w_i L_i$$

or

$$w = \frac{B}{L},$$

the ratio of the wage bill, $B = \Sigma w_i L_i$, to an index of labor input, $L = \Sigma r_i L_i$, measured by weighting each grade of labor by its relative wage, r_i.

Note that L is the natural fixed-weight index of total labor input. The effective wage, w, measured in this way is a reasonable approximation of what the scale wage, s, should have been, and thus serves as a guide for future revisions of the scale through collective bargaining or wage policy. I will return to the relation over time between w and s, and to the general range of problems associated with setting the scale wage.

Within this framework, the basic argument of the paper is the following: The scale wage is determined bureaucratically. If it is set correctly, workers with a comparative advantage in category j are hired into it, and the effective wage is the same as the scale wage. If the labor market is unexpectedly tight, workers hired into job j will be of lower grade than $c(j)$, but they will be paid the scale wage for job j. The effective wage will exceed the scale wage as a consequence of the upgrading of the labor force within the job structure. In unexpectedly slack markets the opposite happens and the effective wage falls short of the scale wage. Mobility brings about movements in the effective cost of labor in the short run even though the scale wage is rigid.

The theory has definite, realistic, implications for the composition of unemployment. In slack markets, workers displace one another down the chain of grades because of their new willingness to work at the wages paid in lower grades. The lowest-grade workers have the fewest opportunities to displace others, and so are disproportionately represented among the unemployed. The wage and hiring policies of employers bring about an allocation of unemployment that is economically efficient (given a reduction in total output) but that shifts the distribution of income adversely.

MARGINAL AND AVERAGE EFFECTIVE WAGES

The previous discussion assumed implicitly that a firm hires its entire work force anew each year. Although a large fraction of the labor force does change jobs every year, the majority remain with the same employer from one year to the next.[4] The process just outlined applies to the firm's gross additions to employment ("new hires" in the Bureau of Labor Statistics data), not to its total labor force. The theory thus needs an additional feature, the distinction between the marginal and the average effective wage. The marginal effective wage (called w') is the cost of increasing employment, or the reduction in the firm's total cost if employment falls. It varies in response to conditions in the labor market in the way discussed in the previous section. It corresponds most closely to the labor component of marginal cost that should be relevant in setting profit-maximizing prices. However, the marginal effective wage cannot be observed directly because labor is not infinitely mobile. Only the most recently hired employees receive today's marginal effective wage. Those hired several years ago have a somewhat different marginal effective wage, depending on conditions in the market when they were hired. The average effective wage today, measured as the ratio of total compensation to total labor input, is the weighted average of the present and past marginal effective wages, with weights equal to the fractions of the labor force with various lengths of time on the job.

These definitions of the three distinct notions of the wage (the scale wage, the marginal effective wage, and the average effective wage) underlie the structure of the model developed in detail in the next pages. In brief, the marginal effective wage plays a role closest to that of the unitary concept of the wage employed in previous theories of the Phillips curve. On the one hand, it responds most directly to excess demand or supply in the labor market, and on the other, it is the marginal cost of labor to the firm. The first equation of the model embodies the hypothesis that the scale wage is set each year to the expected level of the marginal effective wage. Later in the paper, I will show that this equation can be rewritten in a form precisely analogous to the conventional Phillips curve, with the important

---

4. One of my earlier papers—"Turnover in the Labor Force"—discusses the sources of turnover and presents a good deal of evidence on its magnitude.

difference that it deals with the marginal effective wage and not an observed wage. The second equation relates the observed average effective wage to the recent history of the marginal effective wage. It has no precise counterpart in previous theories.

Setting the Scale Wage

Consider first an employer hiring in a competitive labor market. His goal each year is to establish a scale wage that is as close as possible to the marginal effective wage. He must therefore predict the marginal effective wage for the coming year, so the issue of expectations enters the theory at this point. In the empirical work presented below, I have used a simple log-linear prediction equation:

$$\log E(w'_t) = \sum_{\tau=1}^{\infty} \beta_\tau \log w'_{t-\tau} ,$$

where $E(w'_t)$ is this year's predicted marginal effective wage.[5] Then the theory holds that s_t is set equal to the projected marginal effective wage, so the basic equation for the scale wage is

$$\log s_t = \sum \beta_\tau \log w'_{t-\tau} .$$

I assume that the β coefficients sum to one.

A central issue in the determination of the scale wage is whether a gap between this year's w'_t and this year's s_t will appear if the economy settles into an inflationary path. Initially, it is precisely such a gap that sets off the process of inflation. Will the gap remain? The answer depends on the shape of the lag distribution, β_τ. Catching up—that is, closing the gap under persistent inflation—requires an extrapolative element in the distribution. Extrapolation occurs when the distribution assigns positive weight to the first differences of $\log w'_{t-\tau}$. One way to see this is to rewrite the equation as

$$\log s_t = \log w'_{t-1} + \sum_{\tau=1}^{\infty} \gamma_\tau \Delta \log w'_{t-\tau} .$$

5. Many readers have pointed out, correctly, that this part of the model is much less fully developed than other parts. Rational economic agents would use all the information available at time t in projecting $E(w'_t)$, not just the history, $w'_{t-\tau}$. The present paper tries to go as far as possible with the autoregressive prediction equation, but does not compare it to more sophisticated alternatives.

This year's marginal effective wage is projected as last year's w' updated by a moving average of past changes in w'. Thus the gap between the actual value of w'_t and s_t is

$$\log w'_t - \log s_t = \Delta \log w'_t - \sum \gamma_\tau \Delta \log w'_{t-\tau}.$$

On a smooth inflationary path, all of the $\Delta \log w'$ terms will have the same value, say m. Then

$$\log w'_t - \log s_t = m - \sum \gamma_\tau m$$
$$= m(1 - \sum \gamma_\tau)$$
$$= m\mu.$$

The parameter μ is one minus the sum of the γ coefficients; the gap will be zero if and only if μ is zero. If μ is positive, inflation will bring about a permanent gap, with consequences for the real economy. In general, μ measures the degree of slippage embodied in the process of setting the scale wage. In terms of the original lag coefficients, it is

$$\mu = \sum \tau \beta_\tau,$$

the first moment, or mean, of the lag distribution. A distribution with a good deal of catching up will have a mean close to zero, which implies that some of the coefficients will be negative. It would not be surprising to find a distribution in which β_1 was substantially larger than one and all of the other lag coefficients were negative.

I will have more to say about the equation for the scale wage in connection with my discussion of the accelerationist hypothesis, which is directly related to the degree of catching up in the process of setting the scale wage under conditions of persistent inflation. Discussion of other aspects of the equation logically follows the next two sections of the paper, on the demand and supply for labor and the determination of the marginal effective wage.

How should the model incorporate the setting of the scale wage by negotiation between the employer and a labor union representing workers? I think this question can be answered without settling the more fundamental issue of how collective bargaining affects the wage level. I assume that the basic determinants of the wage bargain are stable over time—the monopoly position of the firm, the monopoly power of the union, the substitutability of other factors for labor, and the like. Further, both sides in collective bargaining have roughly the same information about future

output, prices, and wages. The bargaining process simply reestablishes the earlier wage bargain, ex ante, making up for any departures on account of recent unforeseen changes in demand. Negotiation sets a scale wage that enters the employer's hiring process in the way described earlier, modified by whatever conditions the union is able to impose to protect its monopoly position. Strong industrial unions in the United States control hiring into jobs above the entry level by requiring that they be offered to union members first; those previously laid off from the jobs have first claim, then those holding the next lowest jobs on the ladder. Unions do not restrict hiring at entry levels, but usually those hired must join the union after a certain length of time. On the other hand, in periods of contraction, senior union members may displace, or "bump," those holding jobs below them on the ladder. Workers at the entry level face the greatest likelihood of layoff.

If the union is successful in raising wages above market levels, the firm is insulated from conditions in the general labor market.[6] Long lines of workers are always available for the jobs that may be filled from the outside. Attention shifts from fluctuations in conditions in the labor market to fluctuations in the demand for the products of organized industries. This seems appropriate, for it is precisely the most highly organized industries, especially those producing durable goods, that are the most sensitive to cyclical variation in economic activity. In fact, it is only a slight overstatement to say that cyclical variation in output is confined to the durables sector, and that the cycle makes itself felt in other sectors only through the labor market.

One outcome of collective bargaining is to reestablish an appropriate scale wage in the face of recent mistakes.[7] If last year's scale wage was too low, the demand for labor will have been unexpectedly high, and the process of upgrading and new hiring at the entry level will have taken place recently. The marginal effective wage for the past year measures the extent of the error. In Appendix B, I argue that the union would have set a scale wage higher than the marginal effective wage turned out to be, had it

6. The firm may still screen applicants for the most qualified available. Restriction of hiring from the outside to entry-level jobs limits the extent to which the firm can undercut the union's effect on the wage by hiring new workers of exceptionally high quality.

7. In each round of bargaining, many other issues are settled, of course; the division of compensation between cash wages and fringe benefits is an important example. These are, however, much less important in the overall process of inflation.

known that demand would have been so high, but that the best it can do after the fact is to set the scale wage to the marginal effective wage. The essence of the argument is that the union acquires responsibility for continuing the employment of the new members it takes on during an unexpected expansion. This obligation prevents it from seeking a wage higher than the marginal effective wage.

To a reasonable approximation, the marginal effective wage should function as a target for setting the scale wage under collective bargaining as well as in an atomistic labor market. The linear extrapolative model ought to serve equally well as a rough summary of the process of setting the scale wage in both cases. The argument of this section is no more than an elaboration of a point made frequently in discussions of the effect of unions on the rate of inflation: unions cause high wages but not rising wages.

EMPLOYERS' STRATEGIES FOR HIRING

A key element in the theory of wage changes is the response of employers to changing conditions in the labor market. Employers look for bargains in the market—high-grade workers willing to work at the quoted wage. In slack markets they find bargains more frequently and the effective wage is depressed. The extent and success of bargain-hunting is a major factor. If employers could interview all unemployed workers in filling each of their jobs each period, bargain-hunting would be perfect, wages would adjust instantaneously, the Phillips curve would be vertical even in the short run, and the unemployment rate would never deviate from a fixed natural or frictional rate. Plainly, the economy does not function this way and employers are nowhere near this successful in locating bargains.

In this section I will discuss an extreme model in which hiring strategies are severely constrained. This model has the advantage of exact compatibility with a model of the behavior of the unemployed that I will present shortly. Suppose that the personnel department of the firm has an inner and an outer office. The inner office knows nothing about conditions in the labor market, and interviews prospective workers according to fixed rules. From the point of view of the worker or of an outside investigator, these rules have a certain randomness. The worker faces a set of probabilities of receiving offers for various categories of jobs, and these probabilities are the same whether the labor market is tight or slack. The outer

office, on the other hand, knows the probability that a given inquiring worker will accept a job, and it knows how many will present themselves each day. In slack markets, the probability of acceptance and the flow of inquirers are both higher than they are in tight markets. The task of the outer office is to admit only the appropriate fraction of the inquirers to the inner office so that the hiring requirements of the firm will be met. In tight markets the outer office keeps long hours, posts a conspicuous help-wanted sign, and so forth. In slack markets, the outer office is open only in the early morning and looks as uninviting as possible. This may be a suboptimal way to run a personnel department, but still it achieves an element of bargain-hunting. The probability that a given job will be filled by a superior worker rises in a slack market because the probability of acceptance by the most desirable workers rises the most. A more astute personnel department would admit more inquirers to the inner office and then make less favorable offers when labor markets slackened. However, considerable evidence suggests that the main difference between tight and slack markets is in the availability of jobs—that is, in the frequency with which an individual looking for work locates a prospective job—and not in the size of the wages paid for the jobs that are offered. Presumably, this is the outcome of efficient recruiting practices; it is simply too expensive to expand the inner office to decide how to adjust job offers to take complete advantage of changes in conditions in the labor market and then to interview substantially larger numbers of prospective workers.

LABOR SUPPLY AND THE BEHAVIOR OF JOB SEEKERS

This section describes a theory of the behavior of job seekers in the environment created by the hiring strategies of employers as discussed previously. Job seekers are all individuals interested in new work, whether employed or not. By definition, all of the unemployed are job seekers, but when the market is tight, many job seekers never become unemployed— that is, they retain their old jobs until they find new ones. There is a large flow of workers through the labor market each month. I will assume that Holt's law[8] applies: the fraction of the labor force taking new jobs each

8. Charles C. Holt, "Job Search, Phillips' Wage Relation, and Union Influence: Theory and Evidence," in Edmund S. Phelps and others, *Microeconomic Foundations of Employment and Inflation Theory* (Norton, 1970).

week is a constant, ϕ, independent of conditions in the labor market. In tight markets a large fraction of job seekers remain employed, planning to quit when they locate a better job, whereas in slack markets a large fraction have been laid off and are seeking new work involuntarily.

In a great many cases job seeking is entirely passive. A worker simply remains in contact with the relevant labor market, waiting for an attractive prospect. Relatively few job seekers spend long hours deliberately searching for work.[9] This theory should deal with job-seeking activities apart from active search. The basic lesson of recent theories of search—that job seekers rationally spend many weeks considering alternative jobs— applies to all their activities in seeking new work. For example, many workers who are laid off face a substantial probability of recall within a few weeks. Probably only a few of them have a good chance of locating a better job elsewhere. Their rational choice is simply to wait for recall. A satisfactory theory should encompass waiting—for an old, or a better, new, job—as well as searching.

The typical job seeker faces the following situation: There is a probability, p, that a job prospect will materialize each week. In terms of the model of the employer's personnel department, p is the probability that at least one employer will admit the worker to the inner office, which varies according to conditions in the labor market. The main innovation in this theory is the working out of the implications of variations in the availability of work, measured by p. Previous theories, whose treatment of the behavior of job seekers is essentially the same as mine, have taken p as constant, thereby missing an important element in the explanation of the Phillips curve.[10] Further, I take p as known to job seekers, and, in general, put much less emphasis on ignorance and mistaken expectations on the part of job seekers. In this theory, it is the unavailability of work that extends the duration of unemployment when demand falls, not incorrect expectations of prevailing wages and especially not incorrect expectations about the relation between prices and wages.

Since the story about the decisions of individual job seekers tells little that is new, it is relegated to Appendix C. The conclusion is that a job

9. This point is documented thoroughly in Robert J. Gordon, "The Welfare Cost of Higher Unemployment," *Brookings Papers on Economic Activity* (1:1973), pp. 188–95, Appendix C. Hereafter this document will be referred to as *BPEA*, followed by the date.

10. Dale T. Mortensen, "A Theory of Wage and Employment Dynamics," in Phelps and others, *Microeconomic Foundations*; and Holt, "Job Search."

seeker of grade i will end up taking a new job at an average wage of w_i', given by

$$w_i' = sr_ig(p).$$

As before, s is the scale wage set by employers, r_i is the relative wage of grade i as given by the wage structure, and $g(p)$ indexes the dependence of the individual's success in finding a high-paying job on conditions in the labor market as measured by p. In tight markets, $g(p)$ is greater than one and the new jobs found by job seekers pay, on the average, above the scale wage for grade i. In slack markets, $g(p)$ is less than one. The dependence of $g(p)$ on p is the explanation of the negative slope of the Phillips curve in the short run.

The relationship between the unobserved weekly probability of a job prospect and the observed unemployment rate requires attention next. The probability, p, determines the related probability, $h(p)$, that a worker will take a prospect in a given week; $h(p)$ rises with p, because the increasing availability of work makes it easier to find a job, but not in strict proportion, because in tighter markets the job seeker can reject prospects he might have taken if p were lower. For the unemployed, $h(p)$ can be observed, if imperfectly, from the data on the duration of unemployment. If all job seekers went through a period of unemployment, the relation between $h(p)$ and the unemployment rate, u, could be derived from the balance between the number of unemployed taking work and the number of workers losing jobs:

$$h(p)uL = \phi L.$$

Here L is the labor force; the left-hand side is the flow into jobs and the right-hand side is the flow out. The two flows are equal when employment is constant, and are very close even when it is changing, because the gross flows in the labor market exceed the net flows by a wide margin even during the sharpest change in employment. Then

$$u = \phi \frac{1}{h(p)} ;$$

this is the familiar proposition from the theory of turnover that the unemployment rate is the product of the frequency and duration of unemployment. This relation between u and p is an oversimplification in one important respect: as Perry has documented, only a fraction of job changers become unemployed at all, and this fraction varies with condi-

tions in the labor market.[11] Suppose the fraction going through unemployment is $f(p)$. Then balance of flows into and out of unemployment requires

$$h(p)uL = \phi f(p)L,$$

and the relation between u and p is

$$u = \phi f(p)\,\frac{1}{h(p)}.$$

Again, the unemployment rate is the product of frequency, $\phi f(p)$, now considered negatively responsive to p, and duration, $1/h(p)$. "Perry's pothole"—his metaphor for the process by which some workers fall into unemployment while changing jobs—only steepens the stable, negatively sloped relation between the unemployment rate and the frequency of prospects.

The last step in this section is to restate the basic conclusion about the response of individual job seekers to conditions in the labor market, now indexed by the observed unemployment rate:

$$w'_i = sr_i g(u).$$

Determination of the Scale Wage

The first equation of the model relates the value of the scale wage to past values of the determinants of the marginal effective wage. The marginal effective wage in the aggregate is

$$w' = \frac{\sum L_i w'_i}{\sum L_i r_i}$$

$$= \frac{\sum L_i sr_i g(u)}{\sum L_i r_i}$$

$$= sg(u).$$

The ratio of w' to s equals the function $g(u)$. When the market is tight, u is low, $g(u)$ exceeds one, and new workers earn above the scale wage.

Substituting this equation for w'_i into the linear extrapolative model for the setting of the scale wage gives

(1) $\log s_t = \sum \beta_\tau [\log s_{t-\tau} + \log g(u_{t-\tau})].$

11. George L. Perry, "Unemployment Flows in the U.S. Labor Market," *BPEA* (2:1972), pp. 245–78.

The scale wage evolves according to a difference equation with an additive term that varies in response to demand as measured by the unemployment rate.

Determination of the Average Effective Wage Rate

The second equation of the theory relates the observed average effective wage to the past history of the marginal effective wage. In deriving this equation I assume that the distribution of the labor force by length of time on the present job is fixed over time—in any year, a fraction ϕ_0 will have started work within the current year, a fraction ϕ_1 last year, and so forth.[12] Promotions and demotions count as new jobs in the definition of the distribution. In tight markets those who started work recently will have been promoted recently or will have quit earlier jobs, while in slack markets they will have been bumped recently or laid off earlier jobs. I also assume that the wages of workers who remain on the job change in proportion to changes in the scale wage. Under these assumptions, w_t, the average effective wage at time t, is the weighted average of the wages received by workers classified by time on the job, with the wage received by those on the job τ years ago given weight ϕ_τ. The assumption about the application of scale increases implies that a worker of grade i who started work τ years ago at wage $w'_{t-\tau}r_i$ now earns

$$\frac{s_t}{s_{t-\tau}} w'_{t-\tau}r_i.$$

His starting wage (the marginal effective wage) was determined by conditions in the labor market when he started:

$$w'_{t-\tau}r_i = s_{t-\tau}g(u_{t-\tau})r_i.$$

In terms of observable variables, his current wage is

$$s_t g(u_{t-\tau})r_i.$$

The average current wage of workers of grade i is the weighted average,

$$w_{i,t} = r_i s_t \sum_{\tau=0}^{\infty} \phi_\tau g(u_{t-\tau}).$$

12. Strictly speaking, Holt's law implies that the distribution should be geometric, with $\phi_\tau = (1 - \phi)^\tau$. However, Holt's law will hold to a close approximation even if the distribution is not geometric.

Finally, the average effective wage for the whole labor force is the weighted average across grades of labor,

$$w_t = \frac{\sum w_{i,t} L_i}{\sum r_i L_i}$$

(2)
$$= \frac{\sum_i r_i L_i s_t \sum_\tau \phi_\tau g(u_{t-\tau})}{\sum r_i L_i}$$

$$= s_t \sum \phi_\tau g(u_{t-\tau}).$$

The Acceleration Theorem

The Acceleration Theorem, discovered and advocated by Friedman and Phelps, holds that the rate of inflation rises without bound if the unemployment rate is held below a critical equilibrium value.[13] At the equilibrium value, any inherited rate of inflation will be sustained indefinitely. In the long run, the economy cannot trade more inflation for less unemployment. The conditions under which the Acceleration Theorem applies to the model of this paper can be derived in the following way. Consider a path with constant inflation at rate m and a constant unemployment rate:

$$w_t = w_0 e^{mt}$$
$$s_t = s_0 e^{mt}$$
$$u_t = \bar{u}.$$

The path must satisfy both equations of the model. First,

$$\log s_t = \sum \beta_\tau [\log s_{t-\tau} + \log g(u_{t-\tau})],$$

or

$$\log s_0 + mt = \sum \beta_\tau [\log s_0 + m(t-\tau) + \log g(\bar{u})].$$

This gives

$$\log g(\bar{u}) = m \sum \tau \beta_\tau$$
$$= m\mu.$$

Thus

$$g(\bar{u}) = e^{m\mu}.$$

13. Milton Friedman, "The Role of Monetary Policy," *American Economic Review*, Vol. 58 (March 1968), pp. 1–17; and Edmund S. Phelps, "Money Wage Dynamics and Labor Market Equilibrium," in Phelps and others, *Microeconomic Foundations*, pp. 124–66.

Recall that μ is the slippage of the scale wage behind the marginal effective wage in the face of sustained inflation. If there is no slippage, then $\mu = 0$, $g(\bar{u}) = 1$, and the constant unemployment rate, $\bar{u}$, is the same for all rates of inflation, m. The equation $g(u^*) = 1$ defines the unique equilibrium value of the unemployment rate, u^*. At this rate, the marginal effective wage equals the scale wage, implying that expectations about conditions in the labor market are exactly fulfilled. Following Phelps, I think this condition deserves the use of the term "equilibrium." I avoid Friedman's term "natural unemployment rate" because it seems to suggest that u^* is a desirable level of unemployment.

In the rest of the paper I will refer to the condition $\mu = 0$ as the *accelerationist hypothesis*. It alone establishes the Acceleration Theorem for the model. All that the accelerationist hypothesis requires is that those responsible for setting the scale wage extrapolate any past history of smooth inflation into the future.

The equation for the average effective wage can be satisfied for any constant unemployment rate, not just the equilibrium rate, as I will now show, starting with

$$(2) \qquad w_t = s_t \sum \phi_\tau g(u_{t-\tau}).$$

On the inflationary path,

$$w_0 \, e^{mt} = s_0 \, e^{mt} \sum \phi_\tau g(\bar{u}),$$

or

$$w_0 = s_0 g(\bar{u}).$$

The level of the average effective wage lies above or below the level of the scale wage depending on whether $g(\bar{u})$ is above or below one. If $\bar{u} = u^*$, then $w_0 = s_0$ and so $w_t = s_t$: at the equilibrium unemployment rate, all workers receive exactly the scale wage along any inflationary path, provided that workers remaining on the job receive the full benefit of increases in the scale wage. When the labor market is in disequilibrium— say, with the unemployment rate below its equilibrium value—w_0 exceeds s_0, and the average wage is above the scale wage in every year. Nothing in this equation rules out the possibility of permanent disequilibrium with a constant rate of inflation. The Acceleration Theorem depends only on the properties of the equation for the scale wage.

This is a good point to sum up the theory developed in this paper and to compare it to the conventional pair of equations for expected inflation

and the Phillips curve. The comparison is clearest if the Phillips curve is considered in terms of the marginal effective wage. Recall that

$$s_t = E(w'_t);$$

the scale wage is set to the expected value of the current w'. The actual marginal effective wage differs from s_t according to the adjustment function:

$$w'_t = s_t g(u_t)$$
$$= E(w'_t) g(u_t).$$

Taking logs and subtracting $\log w'_{t-1}$ from both sides, I get

$$\log w'_t - \log w'_{t-1} = \log g(u_t) + \log E(w'_t) - \log w'_{t-1},$$

or

$$\Delta \log w'_t = \log g(u_t) + E(\Delta \log w'_t).$$

The rate of change in w'_t, measured as the first difference of its log, is the sum of the log of the wage-adjustment function—and the expected change in the log of w'_t—$\log E(w'_t) - \log w'_{t-1}$.[14] The standard Phillips curve has precisely the same form,

$$\Delta \log w_t = f(u_t) + E(\Delta \log w_t),$$

but deals with the wage level itself rather than the marginal effective wage.

In the theory of this paper, the expected change in the marginal effective wage is

$$E(\Delta \log w'_t) = \log E(w'_t) - \log w'_{t-1}$$
$$= \sum \beta_\tau \log w'_{t-\tau} - \log w'_{t-1}$$
$$= \log w'_{t-1} + \sum \gamma_\tau \Delta \log w'_{t-\tau} - \log w'_{t-1}$$
$$= \sum \gamma_\tau \Delta \log w'_{t-\tau}.$$

---

14. *Not* the change in the expected level, $\log E(w'_t) - \log E(w'_{t-1})$. Instead of subtracting $\log w'_{t-1}$ from both sides, I could have taken first differences to get

$$\log w'_t - \log w'_{t-1} = \log g(u_t) - \log g(u_{t-1}) + \log E(w'_t) - \log E(w'_{t-1}).$$

This is a valid implication of the theory, but it obscures the comparison with the conventional Phillips curve by making it appear that the change in w'_t depends not on the level but on the change in $\log g(u_t)$. The $-\log g(u_{t-1})$ is actually canceled by the log $g(u_{t-1})$ in the expression for $-\log E(w'_{t-1})$. The rate of change of w' depends on the *level* of u_t.

The expected change is a weighted average of past actual changes, precisely as in the standard theory. In both theories, the accelerationist hypothesis has the form $\Sigma \gamma_\tau = 1$, or $\Sigma \tau \beta_\tau = 0$.

There are two important differences between the two theories. First, the theory of this paper formulates the adjustment process and the expectations equation in terms of the marginal effective wage rather than a wage that is observed directly. Second, the level of the expected wage (and hence of the scale wage) plays an important role in this theory, while it has no role at all in the standard theory.

Empirical Evidence on Five Manufacturing Employers

I have calculated series for union wage scales for the following five employers:[15]

Employer	*Union*
Members of the Clothing Manufacturers Association of the United States of America	Clothing Workers of America, Amalgamated
General Motors Corporation	United Auto Workers
General Electric Company	International Union of Electrical and Radio and Machine Workers
B. F. Goodrich Company	United Rubber Workers
American Viscose Division, FMC Corporation	Textile Workers Union of America

As a measure of conditions in the labor markets of the five industries, I have used the proportion of blue-collar workers unemployed in the total U.S. economy. This measure is far from ideal, especially where unions are strong. Construction of measures more directly related to the individual industries is clearly one of the next steps in this area of research, but it is beyond the scope of this paper.

The general form of the equation describing the evolution of the scale wage over time is

$$\log s_t = \sum \beta_\tau [\log s_{t-\tau} + \log g(u_{t-\tau})].$$

15. For sources and descriptions of the data used in this section, see Appendix D.

The empirical work in this paper uses the following simple econometric specification for the wage-adjustment function, $g(u)$:

$$g(u) = \psi_0 - \psi_1 \log u.$$

Provided ψ_1 is positive, this function slopes downward and is concave from above, conforming to most prior views about the curvature of the Phillips curve. I have avoided imposing any strong specification on the lag distribution, β_r, beyond truncating it after three years. The equation is then

$$\log s_t = \beta_1 [\log s_{t-1} + \log (\psi_0 - \psi_1 \log u_{t-1})]$$
$$+ \beta_2 [\log s_{t-2} + \log (\psi_0 - \psi_1 \log u_{t-2})]$$
$$+ (1 - \beta_1 - \beta_2) [\log s_{t-3} + \log (\psi_0 - \psi_1 \log u_{t-3})].$$

The first moment of the lag distribution,

$$\mu = \beta_1 + 2\beta_2 + 3(1 - \beta_1 - \beta_2),$$

which plays a crucial role in testing the accelerationist hypothesis, can be examined directly by substituting its definition into the equation to eliminate β_2:

(1a) $\log s_t = \beta_1 [\log s_{t-1} + \log (\psi_0 - \psi_1 \log u_{t-1})]$
$$+ (3 - 2\beta_1 - \mu) [\log s_{t-2} + \log (\psi_0 - \psi_1 \log u_{t-2})]$$
$$+ (\beta_1 + \mu - 2) [\log s_{t-3} + \log (\psi_0 - \psi_1 \log u_{t-3})].$$

This equation is mildly nonlinear in its parameters, but I found no difficulty in estimating them by nonlinear least squares.

Results obtained from estimating separate equations for the five employers did not suggest that the parameters of the equation differed systematically among them, so I imposed the hypothesis that the same lag distribution and wage-adjustment function prevailed for all of them. This brought sixty observations to bear on the estimation of the four parameters.[16] The results appear in Table 1. The first lag coefficient, $\beta_1 = 1.53$,

---

16. As a first step equation (1a) was estimated separately for each of the five employers. The results were suggestive but hard to interpret on account of the size of the standard errors. Inevitably, estimates of four parameters derived from only twelve observations will have a good deal of dispersion. The first lag coefficient, β_1, ranged from 0.95 for American Viscose to 1.75 for B. F. Goodrich; both had standard errors of over 0.40. The first moment, μ, ranged from -0.27 for B. F. Goodrich to 0.89 for American Viscose, with standard errors above 0.60. In no case was μ more than 1.5 standard errors above or below zero, but the standard errors were so large that the evidence in

Table 1. Parameters of the Scale Wage Equation for Five Manufacturing Employers, Sample Period 1961–72

	Parameter[a]			
Form of equation	*First lag coefficient* β_1	*First moment of lag distribution* μ	*Constant of wage-adjust-ment function* ψ_0	*Slope of wage-adjust-ment function* ψ_1
μ estimated	1.53	0.20	1.051	0.024
	(0.12)	(0.12)	(0.013)	(0.007)
μ constrained to be zero	1.60	0.00	1.037	0.020
	(0.11)	...	(0.009)	(0.005)

Standard error, by employer

Clothing manufacturers (men's)[b]	0.023
General Motors Corporation	0.016
General Electric Company	0.019
B. F. Goodrich Company	0.010
American Viscose Division, FMC Corporation	0.033

Sources: Derived from equation (1a) discussed in the text, using basic data described in Appendix D.
a. The parameter estimates are based on sixty observations. The numbers in parentheses are standard errors.
b. Covers clothing manufacturers affiliated with the Clothing Manufacturers Association of the United States of America.

is substantially larger than one, and in fact the other two lag coefficients are negative. From equation (1a) the implicit values of β_2 and β_3 are

$$\beta_2 = 3 - 2\beta_1 - \mu$$
$$= -0.26;$$
$$\beta_3 = \beta_1 + \mu - 2$$
$$= -0.27.$$

In level form, the scale-wage equation is

$$\log s_t = 1.53 \log w'_{t-1} - 0.26 \log w'_{t-2} - 0.27 \log w'_{t-3}.$$

favor of the accelerationist hypothesis was not at all powerful. The slope of the wage-adjustment function, ψ_1, ranged from 0.008 for General Motors to 0.023 for American Viscose, with standard errors of 0.019 and 0.040. Not even the response to the unemployment rate was statistically unambiguous. When the data were combined, the method of estimation took account of the correlation among the residuals for the same year. For details, see E. Berndt, B. Hall, R. Hall, and J. Hausman, "Estimation and Inference in Nonlinear Structural Models," *Annals of Economic and Social Measurement* (forthcoming).

The equation embodies a good deal of catching up—the negative coefficients on w'_{t-2} and w'_{t-3} imply that s_t will not fall behind w'_t even when w'_t is rising rapidly. The catching-up property may be revealed more plainly in the following rewriting of the same equation:

$$\log s_t = \log w'_{t-1} + 0.53 \, \Delta \log w'_{t-1} + 0.27 \, \log w'_{t-2}.$$

This year's scale wage is last year's marginal effective wage updated by 53 percent of last year's change in the marginal effective wage plus 27 percent of the change the year before that. The two coefficients sum to only 80 percent, so there is a slippage of 20 percent of any upward trend in w'_t. This slippage is precisely what is measured by the first moment of the lag distribution, μ, which is 0.20. Its standard error, 0.12, is sufficiently small to cast statistical doubt on the strict accelerationist hypothesis of full catching up, that is, $\mu = 0$. However, the results suggest that the accelerationist hypothesis is not far from the truth. The scale wage catches up to within 20 percent of its target, and even more catching up is statistically entirely plausible. Imposing the hypothesis $\mu = 0$ further reduces the standard errors of the other parameters, as the second set of results in Table 1 shows.

The wage-adjustment function in the constrained equation is

$$g(u) = 1.037 - 0.020 \log u.$$

The equilibrium unemployment rate, u^*, defined by $g(u^*) = 1$, is

$$u^* = \exp \left[(\psi_0 - 1) / \psi_1 \right] = \exp (1.85) = 6.4 \text{ percent.}$$

When 6.4 percent of blue-collar workers are unemployed in the economy at large, the labor market of the five employers considered here is in equilibrium in the sense that the scale wage and the marginal effective wage are equal and the employers are hiring workers of the expected grade into their new jobs. This level of unemployment corresponds to an overall unemployment rate of 5.1 percent, a little higher than most other estimates but slightly lower than my estimate in the next section for the aggregate economy.

The statistical evidence that union wage scales respond to conditions in the labor market is unambiguous: the slope parameter, ψ_1, could not reasonably be positive solely because of random sampling error. The ratio of the marginal effective wage and the scale wage responds to the unemployment rate in the following way:

Unemployment rate		Ratio of marginal effective wage to scale wage
Overall	*Blue-collar workers*	
3.1	3.2	1.014
4.1	4.8	1.006
5.1	6.4	1.000
6.1	8.0	0.995
7.1	9.6	0.992

The second equation of the theory relates the observed average effective wage to the recent history of the marginal effective wage:

$$w_t = s_t \sum \phi_\tau w'_{t-\tau}/s_{t-\tau}.$$

Again, I assume

$$w'_t = s_t(\psi_0 - \psi_1 \log u_t).$$

The equation to be estimated is

$$w_t = s_t \sum \phi_\tau(\psi_0 - \psi_1 \log u_{t-\tau}).$$

The average effective wage, w_t, is the wage received by the typical worker. If all workers were the same, or if the composition of the labor force never changed, it could be measured as average earnings per man-hour. The first source of data used here is precisely that: the series on straight-time average earnings calculated by the Bureau of Labor Statistics, at the appropriate standard industrial classification level for the firm. This is the narrowest possible concept of the average effective wage; it departs from the scale wage only because of changes in the composition of the jobs offered by the firm. Nevertheless, the results of this section show that there are important fluctuations in the BLS series relative to the scale wage, related systematically to conditions in the labor market. The second source of data makes extensive adjustments for changes in the composition of the labor force to get the average earnings of the standard manhour. It is calculated as the ratio of total compensation to an index of labor input prepared by Frank Gollop.[17] Gollop's index weights the various age-sex-occupation categories by relative wage rates and by relative annual hours of work. The wage derived from his measure of labor input is much closer in concept to the average effective wage of

17. "Modelling Technical Change and Market Imperfections: An Econometric Analysis of U.S. Manufacturing, 1947–1971" (Ph.D. dissertation, Harvard University, 1974), Chap. 1 and Appendixes A–E.

the theory, but it suffers relative to the BLS series in that it covers a much
wider group of workers than does the scale wage: Gollop's wage covers
both production and nonproduction workers, and is calculated at the
two-digit level, while the BLS series covers only production workers and
is available at a much more disaggregated level. Further, Gollop's series
includes the imputed value of fringe benefits, which are excluded from
the BLS and scale wages.

In some of the industries, there is a slight upward or downward trend in
w_t relative to s_t, the result of long-term shifts in the composition of em-
ployment. Further, Gollop's wage series is an index, set arbitrarily to
1.000 in 1958, and the straight-time series from the BLS is not on precisely
the same basis as the series for the scale wage. In view of these consider-
ations, I have added a multiplicative constant and a trend to the equation
for the average effective wage:

$$w_t = ke^{\lambda t}s_t \sum \phi_r(\psi_0 - \psi_1 \log u_{t-r}).$$

The coefficients of the lag distribution sum to one. I further constrained
the distribution to be a trapezoid—that is, an Almon specification with
degree one. The lag is assumed to cover three years. The equation finally
estimated was

(2a) $w_t = ke^{\lambda t}s_t[\phi_0 (\psi_0 - \psi_1 \log u_t) + \frac{1}{3} (\psi_0 - \psi_1 \log u_{t-1})$

$$+ (\tfrac{2}{3} - \phi_0) (\psi_1 - \psi_0 \log u_{t-2})].$$

As in the case of equation (1a), I first estimated the parameters of
equation (2a) for the five industries separately. Again, it appeared that
joint estimation, under the constraint that the parameters were the same
in all industries, gave the most usable results (I did let each industry have
its own k and λ). The results of the estimation appear in Table 2. With
Gollop's wage series, the first lag coefficient, ϕ_0, is surprisingly large—
most of the response of the average effective wage to the marginal effective
wage takes place inside of a year. The standard error of ϕ_0 is sufficiently
large, however, that sampling variation alone may explain its unexpected
size. The lag distribution estimated from the BLS data is quite reasonable,
suggesting that about half of the labor force changes jobs, receives pro-
motions, or otherwise benefits from wage changes apart from scale changes,
within the first year after the market tightens. Another third benefits by
the second year, and the rest within three years. The Gollop wage embodies
substantial cyclical corrections for the movement in and out of the labor
force of low-wage workers, so it is not altogether surprising that the slope

Table 2. Parameters of the Equation for the Average Effective Wage for Five Manufacturing Industries, by Alternative Measures, Sample Period 1961–72

Wage measure and industry		*Parameter*[a]	
	First lag coeDcient ϕ_0	*Constant of wage-adjust-ment function* ψ_0	*Slope of wage-adjust-ment function* ψ_1
Wage measure			
Gollop	0.83	1.097	0.058
	(0.18)	(0.031)	(0.010)
Bureau of Labor	0.54	1.061	0.023
Statistics	(0.16)	(0.014)	(0.005)

		Standard error, by industry	
		Gollop wage	*BLS wage*
	Clothing	0.020	0.027
	Automobile	0.036	0.029
	Electrical equipment	0.029	0.026
	Rubber	0.024	0.016
	Rayon	0.027	0.027

Sources: Derived from equation (2a) discussed in the text, using basic data described in Appendix D.
a. The numbers in parentheses are standard errors.

of the wage-adjustment function, ψ_1, is quite large. In the case of Gollop's wage, sampling variation alone cannot explain the discrepancy between the estimate of ψ_1 and the same parameter estimated in Table 1: the estimate in Table 2 is almost three times that in Table 1. With the BLS data, on the other hand, the estimate of ψ_1 in Table 2 is in close agreement with that in Table 1. I am not altogether sure how to explain the superior performance of the conceptually inferior BLS wage in equation (2a). Mismatches in the data may be part of the story: the unemployment rate is much too aggrega-tive for the scale wage, but comes closer to having the same coverage as Gollop's wage.

Taken together, the results for the two equations of the model for the five employers and industries give an internally consistent empirical view of the process of inflation that conforms in all major respects to the predictions of the theory. The only important exception is the disappointing performance of the quality-corrected wage in the equation for the average effective wage. Much more work will be required to create satisfactory, mutually consistent data on effective wages. Finally, while the empirical results generally support the theory, it would not be fair to claim that they

cast doubt on competing theories. I plan further work in testing some
of the crucial points of disagreement among the alternative theories
discussed in this paper.

An Aggregate Wage Equation

Although data on wage scales exist for the large fraction of the labor
force employed by firms with more than a hundred or so employees, no
comprehensive data exist on the scale wage for the whole U.S. economy.
It is necessary, therefore, to estimate a single equation obtained by sub-
stituting one of the two equations of the model into the other to eliminate
the unobserved scale wage. Recall that in equation (2)

$$w_t = s_t \sum \phi_\tau g(u_{t-\tau}),$$

which gives

$$s_t = \frac{w_t}{\sum \phi_\tau g(u_{t-\tau})}.$$

With estimates of the lag distribution ϕ and the function $g(u)$ giving the
response of the marginal effective wage to conditions in the labor market,
it is possible to calculate the implied series for the scale wage from the
observed values of the average effective wage. The result, say $\hat{s}_t$, can be
substituted into the equation for the scale wage to give the one-equation
condensed model,

$$\log \hat{s}_t = \sum \beta_\tau [\log \hat{s}_{t-\tau} + \log g(u_{t-\tau})].$$

Carrying out the substitution for s_t yields the equation finally estimated:

$$(3) \quad \log w_t = \log \left[\sum \phi_k (\psi_0' - \psi_1' \log u_{t-k}) \right]$$
$$+ \sum_\tau \beta_\tau \{ \log w_{t-\tau} - \log \left[\sum_k \phi_k (\psi_0' - \psi_1' \log u_{t-\tau-k}) \right]$$
$$+ \log (\psi_0 - \psi_1 \log u_{t-\tau}) \}.$$

The parameters of the wage-adjustment function that appears in the ef-
fective-wage equation have primes to distinguish them from those in the
scale-wage equation.

For the aggregate average effective wage rate I have used a series calcu-
lated by Peter Chinloy on the ratio of total compensation in the private non-
farm U.S. economy to an index of total labor input.[18] The index of labor

18. "Issues in the Measurement of Labor Input" (Ph.D. dissertation, Harvard
University, 1974), Chap. 1.

input is constructed along the same lines as Gollop's indexes for manufacturing industries, and the resulting wage index accords fairly well with the notion of the average effective wage in the theory. As a measure of conditions in the aggregate labor market I have used a fixed-weight index of unemployment by age-sex groups, u_F. The weights reflect the contributions of the various groups to labor input: each group's weight is the product of its share in the total labor force, its relative hours of work, and its relative wage. Details appear in Appendix D. This measure of unemployment derives from Perry's work,[19] but goes somewhat further in adjusting for the shifting composition of the labor force. Perry's unemployment rate counts unemployed units of labor input in place of unemployed individuals, but shifts upward when all unemployment rates remain unchanged and the composition of the labor force shifts toward groups with high unemployment rates. Mine remains constant in the face of a shift in composition. The official unemployment rate rose about 0.03 or 0.04 percentage point per year relative to Perry's index from the early fifties to the late sixties. For my unemployment rate the rise is almost 0.06 point per year. In 1974, an official unemployment rate of 6 percent indicates the same degree of tightness as 5.4 percent did in 1964.

Table 3 reports the values of Chinloy's wage index, the fixed-weight unemployment rate, u_F, and the implicit scale wage, $\hat{s}_t$. In calculating $\hat{s}_t$, I have held the lag distribution fixed with $\phi_0 = 0.50$, $\phi_1 = 0.33$, and $\phi_2 = 0.17$, values that are consistent with the estimates for manufacturing presented earlier. Further, I held ψ_0' arbitrarily at 1.144; this is just a normalization and any other value would have given the same results. I estimated ψ_1' by finding the value that minimized the sum of squared residuals of equation (3). The best value was 0.099. The results for the estimation of equation (3) conditional on this value appear in Table 4.

The first set of estimates yields a value for μ, the first moment of the lag distribution of the scale wage behind the marginal effective wage, that is slightly but not significantly positive. The results are compatible with the accelerationist hypothesis, although some slippage of the scale wage behind the marginal effective wage under persistent inflation is also consistent with the findings.

The other results are sharpened a bit by imposing the accelerationist hypothesis, $\mu = 0$, as the second set of estimates in Table 4 shows. The

19. George L. Perry, "Changing Labor Markets and Inflation," *BPEA* (3:1970), pp. 411–41.

**Table 3. Wage Index, Fixed-Weight Unemployment Rate,
and Implicit Scale Wage, 1953–72**

Year	Wage index w_t	Fixed-weight unemploy- ment rate u_F	Implicit scale wage $\hat{s}_t$
1953	0.829	2.6	0.792
1954	0.855	5.1	0.843
1955	0.866	4.0	0.861
1956	0.908	3.7	0.902
1957	0.965	3.8	0.954
1958	1.000	6.3	1.013
1959	1.038	4.9	1.056
1960	1.083	4.9	1.103
1961	1.102	5.9	1.128
1962	1.123	4.9	1.145
1963	1.152	4.8	1.170
1964	1.210	4.2	1.216
1965	1.243	3.6	1.234
1966	1.320	2.9	1.287
1967	1.371	2.8	1.323
1968	1.467	2.6	1.403
1969	1.574	2.5	1.497
1970	1.717	3.7	1.661
1971	1.810	4.5	1.789
1972	1.938	4.1	1.932

Sources: Wage index, Peter Chinloy, "Issues in the Measurement of Labor Input" (Ph.D. dissertation, Harvard University, 1974); also see Appendix D below; unemployment rate, see Appendix D; implicit scale wage, calculated from

$$\hat{s}_t = \frac{w_t}{\sum \phi_\tau (\psi_0' - \psi_1' \log u_{t-\tau})},$$

with lag distribution $\phi_0 = 0.5$, $\phi_1 = 0.33$, $\phi_2 = 0.17$, $\psi_0' = 1.144$, $\psi_1' = 0.099$; w_t is the average effective wage, and u_t is the unemployment rate.

lag distribution of the scale wage, s_t, behind the marginal effective wage, w_t', implied by the estimate of β_1 is

$$\log s_t = 1.37 \log w_{t-1}' + 0.26 \log w_{t-2}' - 0.63 \log w_{t-3}'.$$

In the extrapolative form,

$$\log s_t = \log w_{t-1}' + 0.37 \,\Delta \log w_{t-1}' + 0.63 \,\Delta \log w_{t-2}'.$$

The slope of the wage-adjustment function, determined by the parameter ψ_1, is indicated by the following calculations:

Unemployment rate (percent)

Fixed-weight u_F	Corresponding official rate, 1974 composition	Ratio of marginal effective wage to scale wage
2.3	3.7	1.029
3.3	4.9	1.012
4.3	5.8	1.000
5.3	6.8	0.990
6.3	7.9	0.981

Again, there is statistically unambiguous evidence that effective wages depart from scale wages in a way that is systematically related to conditions in the labor market. Some examples from past years will help to draw out the implications of the results. At the equilibrium rate of unemployment, a u_F of 4.3 percent, the marginal effective wage equals the scale wage. Both 1970 and 1971 were years of near equilibrium. In the somewhat tighter labor market of 1965, with u_F at 3.6 percent, it is estimated that the typical worker took a new job at about 1 percent above the scale wage, while in the superheated economy of 1968, with u_F at 2.6 percent, the gap was more than 2 percent. The model implies that in the mild recession of 1954, with u_F at 5.1 percent, new jobs paid about 0.7 percent below the scale, and in the deep recession of 1958, with u_F at 6.3 percent, new jobs paid almost 2 percent less than the scale.

A second way to illustrate the response of wages to conditions in the labor market is the following: Suppose both the scale and effective wage have been stable over time at, say, 100, and that the unemployment rate

Table 4. Parameters of the Aggregate Wage Equation

Form of equation	First lag coefficient β_1	First moment of lag distribution μ	Constant of wage-adjustment function ψ_0	Slope of wage-adjustment function ψ_1	Standard error of the regression
μ estimated	1.31 (0.18)	0.13 (0.20)	1.074 (0.015)	0.040 (0.009)	0.014 ...
μ constrained to be zero	1.37 (0.16)	0.00 ...	1.068 (0.011)	0.047 (0.008)	0.013 ...

Source: Equation (3), discussed in the text, using the value for ψ_1' (0.099) that minimized the sum of the squared residuals, where ψ_1' is as defined in Table 3.
  a. The numbers in parentheses are standard errors.

has been at the equilibrium rate. Then in year zero an expansionary
policy reduces the unemployment rate by one point for a full year. From
then on policy holds the unemployment rate at the equilibrium rate. This
sets off the following inflationary spiral:

Year	Fixed-weight unemployment rate, u_F	Scale wage	Marginal effective wage	Average effective wage Level	Average effective wage Change
− 2	4.3%	100.0	100.0	100.0	...
− 1	4.3	100.0	100.0	100.0	0.0%
0	3.3	100.0	101.2	100.6	0.6
1	4.3	101.7	101.7	102.1	1.5
2	4.3	102.7	102.7	102.9	0.8
3	4.3	103.3	103.3	103.3	0.4
4	4.3	104.2	104.2	104.2	0.9
5	4.3	104.9	104.9	104.9	0.7

After a somewhat irregular pattern in the first few years, wage inflation
settles down to a rate of about 0.8 percent per year for the indefinite
future. The policy that achieves the steady unemployment rate of 4.3 per-
cent is one of continual expansion of aggregate demand; an attempt to
stabilize the wage level would inevitably require that the unemployment
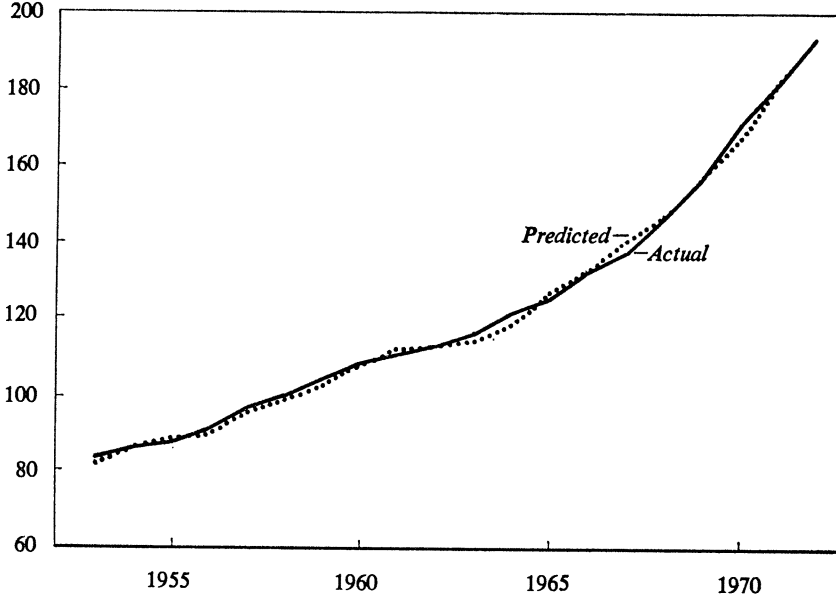rate rise above the equilibrium rate.

It is instructive to examine the response of wages to a policy that holds
the unemployment rate below 4.3 percent for many years, starting from
the stable conditions in the previous example:

Year	Fixed-weight unemployment rate, u_F	Scale wage	Marginal effective wage	Average effective wage Level	Average effective wage Change
− 2	4.3%	100.0	100.0	100.0	...
− 1	4.3	100.0	100.0	100.0	0.0%
0	3.3	100.0	101.2	100.6	0.6
1	3.3	101.7	103.0	102.7	2.1
2	3.3	104.4	105.7	105.7	2.9
3	3.3	107.9	109.2	109.2	3.3
4	3.3	112.4	113.7	113.7	4.1
5	3.3	117.9	119.3	119.3	4.9
6	3.3	124.6	126.2	126.2	5.8
7	3.3	132.7	134.4	134.4	6.5

Finally, it is useful to look at the story equation (3) tells about the
evolution of the wage level over the past twenty years. Figure 1 shows

Figure 1. Actual and Predicted Wage Index, Private Nonfarm Economy, 1953–72

Index (1958 = 100)

Source: Derived from equation (3), discussed in the text.

the actual and predicted values for the wage index over the period of fit, 1953–72. The important errors occur mainly in recession years. In 1961, the wage was overstated by nearly 3 percent of its level; the equation predicted a change of 4.6 percent over 1960 when the actual increase was only 1.8 percent. In 1967, a smaller error in the same direction occurred: the predicted wage is 2.2 percent too high. The predicted change over 1966 was 6.1 percent, against an actual change of about 3.9 percent. On the other hand, the prediction for 1970 is slightly too low and for the sharpest recession by far, 1958, it is almost exactly correct. It is apparent that a single unemployment rate is not a complete measure of conditions in the labor market, especially in recession years. Some of the failures of the equation also can be attributed to the use of annual data. There is no question, though, that the equation does a good job in tracking the enormous acceleration in the rate of inflation in the period 1965–69.

Wage Inflation Today

The aggregate equation projected a wage increase of 6.9 percent for 1973 and a somewhat larger increase, 8.7 percent, for 1974 (based on the assumption that u_F will be 3.9 percent for 1974). Most of this inflation is a hangover from the expansionary binge of 1965–69. The contribution of those years can be measured in the following way: First, project the evolution of wages from 1964 to 1974 using the initial conditions of 1963 and the actual path of unemployment. Then repeat the projection with the fixed-weight unemployment rate set to the equilibrium rate of 4.3 percent for 1964–69. The difference between the two projected rates of inflation for 1974 is today's inheritance of inflationary expectations operating through the mechanism setting the scale wage; it amounts to 5.8 percent per year. In the Phillips curve diagram, this should be visualized as a vertical shift of the whole curve by 5.8 percentage points. Further, aggregate policy since 1969 has been expansionary enough to add significantly to today's inflation by shifting the Phillips curve even further upward. In 1973, the labor market was tighter than it was in 1965, and only in 1971 did the fixed-weight unemployment rate go as high as its equilibrium value. The cumulated effect of tight markets in 1970, 1972, and 1973 shifted the curve upward by another 2.1 percentage points, as measured by the procedure just described.

The worsening of the unemployment-inflation position of the economy since 1964 can be decomposed into three parts on the basis of these calculations: First, according to recent projections, 1974 has 0.3 percentage point *more* inflationary pressure in the labor market than did 1964, as measured by the fixed-weight unemployment rate; but these projections put the 1974 official unemployment rate 0.2 point *higher* than it was in 1964. Second, something over 5 percentage points of the current inflation represents the inheritance of the acceleration that took place from 1964 to 1969. Third, the remainder of about 2 percentage points reflects the further acceleration of 1970–74.

Today's high rate of wage inflation is the result of a decade of continuously tight labor markets, even though the shift in the composition of the labor force has masked the tightness in recent years. The impression has become widespread, even among economists, that contractionary policies can no longer keep inflation under control. For example, Walter Heller

has written: "Why should the economic game plan that failed so miserably in 1968–71 work in 1974–75? Tightening the fiscal and then the monetary screws generated 6% unemployment . . . [in] 1970, yet failed to subdue inflation."[20] The results presented earlier suggest that this is a misinterpretation of recent experience. Conditions in 1970 and later were almost always expansionary, in the sense that labor markets were tighter than their equilibrium levels, especially in 1973. Contractionary policy was tried only briefly in 1971. I think it is unlikely that it will be tried again in the next few years. There is very little room between the equilibrium rate of unemployment with today's labor force (estimated at about 5.5 percent on the official rate, with a standard error of 0.29) and what I judge to be the congressional intolerance for rates much above 6 percent. In the next section I discuss several alternative policies, including two that are genuinely contractionary. Both have a distinct air of unreality about them. Continuation, and perhaps worsening, of inflationary pressure in the labor market appears the probable future course of the economy.

Future Wage Inflation under Alternative Policies

A policy for aggregate demand that maintained unemployment at the equilibrium rate of about 5.5 percent of the total labor force would ratify the current level of wage inflation of around 8 percent for the indefinite future. The implications of such a policy for the average effective wage are shown in the first column of Table 5. Stable, secular inflation of this sort, implying price inflation of about 5 percent a year, is totally without precedent. Its accommodation would require a number of reforms of economic institutions, including the elimination or adjustment of controls on interest rates. The labor market itself is already capable of adjusting to a policy of stable inflation: the catching-up property of the mechanism by which the scale wage is set guarantees that wages do not fall behind when inflation is fully anticipated; formal cost-of-living escalators would not be necessary.

Some economic policymakers within the federal government are reluctant to ratify 8 percent wage inflation. Increasingly, their unhappy experience with price and wage controls has convinced them that the "old-time religion" of contraction in the aggregate economy is the only way out.

20. *Wall Street Journal,* August 12, 1974, p. 8.

Table 5. Projected Rates of Wage Inflation under Alternative Policies, 1974–84 and Beyond

Percent

Year	Held constant at 4.3 percent[a]	Held at 5.8 percent[b]	Held at 4.9 percent[c]	Held constant at 4.0 percent[d]
	Fixed-weight unemployment rate			
1974	8.7	8.7	8.7	8.7
1975	7.8	7.1	7.4	7.9
1976	7.6	5.2	6.5	8.2
1977	7.6	4.4	6.2	8.4
1978	7.8	4.1	6.2	8.7
1979	7.7	3.0	5.7	8.9
1980	7.8	3.0	5.4	9.1
1981	7.7	3.7	4.0	9.3
1982	7.8	3.1	4.6	9.5
1983	7.7	1.7	4.2	9.7
1984	7.8	2.6	3.9	9.9
.				
.				
.				
2000	7.7	2.6	2.9	13.3

Source: Derived by author.
a. 4.3 percent is the equilibrium rate.
b. Held at equilibrium rate plus 1.5 percent whenever last year's wage inflation exceeds 3.2 percent.
c. Same as note b, but unemployment raised by only 0.6 percentage point (to the 1962 level).
d. The average level of unemployment for 1961–68.

There is no doubt the prescription is effective. The second column of Table 5 shows the results of a policy of strong old time-religion that constricts the economy to the point where the fixed-weight unemployment rate is 5.8 percent for as long as it takes to drive the rate of wage inflation down to 3.2 percent per year, approximately the condition required for complete price-level stability. Only in 1958 and 1961 has the unemployment rate reached this level in the postwar economy. Six years of extreme slack in the labor market—the first five and the eighth after imposition of the policy—are required to bring wage inflation down to the target.

The Phillips curve is a curve and not a straight line, so it pays to take the old-time religion in smaller doses over a longer period. In the third column of Table 5, I show the results of a policy that holds the unemployment rate only 0.6 percentage point above the equilibrium rate. This rate (4.9 percent) prevailed in 1962. Ten straight years of this degree of

slack brings the rate of wage inflation down close to the target of 3.2 percent.

Finally, it is worth looking at the implications of a policy that maintains permanently tight labor markets. The last column of Table 5 plots the gradually accelerating inflation that would accompany a fixed-weight unemployment rate of 4.0 percent, the average for the Kennedy-Johnson years, 1961–68. The cost is a rate of wage inflation that is about 1.3 percentage points higher in 1980, 2.1 points higher in 1984, and 5.6 points higher in 2000 than when the unemployment rate is held at 4.3 percent (first column). The process of accelerating inflation takes many years to reach the stage of a wage explosion.

POSTSCRIPT ON THE ROLE OF PRICES IN
THE PROCESS OF WAGE INFLATION

Many readers of earlier drafts of this paper expressed dissatisfaction about the absence of a role for prices or expectations of price changes in my treatment of wage inflation. The paper says nothing about the major concern today that increases in the prices of food and oil will feed back into the wage process, shifting the Phillips curve even further upward. I think it is premature to deal with the issue of feedback on an empirical level; even if there is substantial feedback, it will not appear fully in the observed wage level for another year or so. At a theoretical level, there is little agreement about the mechanism linking prices to wages. I will mention a few hypotheses here, deal with some of them theoretically, and indicate some directions of further research.

First, and easiest to dispose of, is the popular view that labor can set the wage at any level it wants and will always use this power to restore losses in the real wage. This is no more than another manifestation of the fallacy that market power causes inflation. If labor has the power to raise wages, why wait until a price increase to exercise it? No matter how much concentration exists on either side of the bargain, the wage is set by its objective determinants, and only if the price increase is related to one of these determinants will it affect the wage.[21]

Robert J. Gordon avoids the fallacy of unused market power by taking an index of product prices as the price relevant for the determination of

21. James Tobin makes exactly this point in "Monetary Policy in 1974 and Beyond," *BPEA* (1:1974), pp. 219–32.

wages.[22] Within the theoretical framework of this paper, Gordon's view can be interpreted as suggesting an alternative way to measure the excess demand for labor. When the product price exceeds marginal cost as determined by the wage and the costs of other inputs, excess demand for labor and the other inputs appears. A variable that measured the gap between price and marginal cost could substitute for the unemployment rate in my equations. In some cases, the alternative would very likely be superior; for example, in the conditions of late 1974, coal miners will surely be able to drive wages up in response to the dramatic increases in the price of coal, and their unemployment would not be a satisfactory indicator of their ability to do this. I have misgivings about the usefulness of unemployment as an indicator of pressure on wages, and have already begun work on alternatives that are more directly related to the demand for labor. I want to emphasize, though, that measuring excess demand in the price dimension rather than the employment dimension is an alternative, not a necessary additional, feature of the theory. There is nothing theoretically deficient about a model in which unemployment is the measure of excess demand.

The Phillips curves fitted by Gordon and others who believe in a significant feedback from prices to wages do not contain the gap between the level of prices and the level of costs. Rather, they use the rate of change of prices in place of the rate of change of wages in the second term of the Phillips curve, which deals with expectations. I do not believe that this is a correct implication of the view that price less marginal cost is a better measure of excess demand than is unemployment. Further, I know of no convincing rationale for the presence of expected price inflation rather than expected wage inflation in the second term of the Phillips curve. The only plausible justification is that prices are better predictors of future wage inflation than is the recent history of wages themselves. I find this view unconvincing. It is significant that Gordon and Michael Wachter (in his paper in this issue) find that the price index that performs best in the Phillips curve is the price of value added, which is most like a wage because it excludes imports.

22. Gordon has used implicit price deflators rather than consumer price indexes in the favored equations of most of his many papers in *BPEA* on the Phillips curve. As his discussion of this paper shows, he favors the price of the product that labor is producing rather than the price of products labor consumes.

I think the theoretical case for a feedback from prices to wages other than through the excess demand for labor remains to be worked out. I am prepared to defend my omission of such a feedback from the view of the process of wage inflation advocated in this paper. There is no point in a dogmatic position, however, since the economy is in the midst of an accidental experiment with divergent movements of prices and wages. In another year a good deal of new evidence will be available on this question.

APPENDIX A

Survey of Related Work on
Wages and Unemployment

FOR THE PURPOSES of this discussion I will distinguish four groups of economists with divergent views on the theory of wage changes and un-employment: labor economists, search theorists, turnover theorists, and students of wage drift.

Labor economists have always been uncomfortable with the notion of a Phillips curve in which the unemployment rate has a major role in determining changes in wages. Their unease began long before A. W. Phillips' paper of 1958. The following remarkable passage appears in John T. Dunlop's classic *Wage Determination under Trade Unions*, published in 1944:

A cherished view among economists has been that wage rates advance in any market when unemployment has been reduced below a "critical" level and are reduced when unemployment exceeds another "critical" level. Ordinarily the proposition is stated in terms of the wage structure for a total system. The industrial sequence of wage variation suggested in this chapter would render the first formulation of the proposition invalid. Wages fell last (and probably least) in the sector of the economy in which unemployment was clearly relatively greatest and rose first where it was also relatively greatest.[1]

Dunlop's alternative view, widely held among labor economists today, is expressed clearly in his book: The labor market is not a bourse—it is not

1. (Macmillan), pp. 147–48. See also A. W. Phillips, "The Relation Between Un-employment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861–1957," *Economica*, N.S., Vol. 25 (November 1958), pp. 283–99

a market where buyers and sellers deal symmetrically with each other to negotiate prices. Rather, employers quote fixed wages to potential employees on a "take it or leave it" basis. Negotiation over wage levels between individual workers and employers almost never occurs. The quoted or scale wage is set bureaucratically and by collective bargaining, and the major consideration in setting it is the past history of wages in related industries. Dunlop discusses at length the hypothesis that a whole series of wage increases can be set off by an increase in a single key wage. This view remains influential today; many advocates of wage controls believe that the wages of only a few key industries need be controlled to restrain inflation in the entire labor market. Dunlop also demonstrates in some detail that the effective wage differs from the scale wage, a topic that has since received almost no attention among American economists. He attributes the gap between the two wages to overtime premiums and to variations in the productivity of workers paid on the basis of piecework, and not to upgrading within the firm. He takes the strong position that the scale wage is the true cost of labor and that variations in the effective wage around it are merely statistical artifacts.

The theory of this paper starts from the labor economist's observation that the labor market is not a bourse. In particular, the unemployed do not bid the wage down through personal negotiation with employers. I depart from Dunlop's view, however, that the scale wage is the cost of labor input. When firms choose to pay overtime rather than expand employment, the alternative of expanding employment must be even more expensive. Similarly, cyclical upgrading increases the true cost of labor. Except for piecework (which is much less common now than thirty years ago), it seems to me that the effective wage—specifically, the marginal effective wage—is the appropriate measure of the cost of labor. This is a critical point, for it means that variations in the quality of labor in given jobs substitute for direct negotiation of wages as a mechanism of wage adjustment. I think that labor economists seriously understate the responsiveness of the wage to conditions in the labor market. As a consequence, I believe they overstate the importance of the direct transmission of wage increases from one industry to another. Wage increases in related industries are, of course, highly correlated, but this can be explained by their similar responses to fluctuations in the labor market they share.

One labor economist, Melvin Reder, has stated clearly the proposition that quality variations function in the short run to make effective wages

responsive to the supply and demand for labor—the position I adopted here. He has written:

> Wage rates paid for particular jobs are not analogous to factor prices. The skill and other characteristics of workers who apply for given jobs vary with the state of the labor market. . . .
>
> Quality variations in labor markets arise through upgrading and downgrading of members of the labor force relative to the jobs they are to fill. . . .
>
> Shifting workers, reclassifying jobs, etc., are more or less continuous processes, and therefore not subject to the time-lags attendant upon changing contractual prices, i.e., union wage rates.[2]

Reder is concerned mainly with the implications of quality variations for the cyclical behavior of occupational wage differentials, and does not consider their role in a theory of the evolution of the average wage.

Search theorists represent the polar extreme to the views of labor economists. In the most carefully and fully articulated exposition of the search theory, that of Mortensen, employers set wages for jobs in immediate response to shifts in the supply function for labor facing them.[3] Workers quit to look for new work when they feel that wages available elsewhere are enough better to justify the investment in a period of search. Quits are the only source of turnover in Mortensen's theory. They are the result of faulty expectations on the part of employers and workers. The unreality of this aspect of the theory has been pointed out forcefully by Tobin.[4] Quits are never an important source of unemployment, and they are least important when unemployment is highest. A second defect of Mortensen's theory has received less attention: He assumes that the weekly probability of a job offer is a constant independent of conditions in the labor market. Employers make larger numbers of offers at lower wages when the unemployment rate is high. The unemployed individual sees a slackening in the market not as a decline in the availability of work but as a reduction in the wages paid by the jobs available. Again, this is unrealistic. All the evidence suggests the opposite: quoted scale wages respond more slowly to conditions in the market and the immediate change seen by job seekers

2. M. W. Reder, "The Theory of Occupational Wage Differentials," *American Economic Review*, Vol. 45 (December 1955), pp. 834, 835.

3. "A Theory of Wage and Employment Dynamics," in Phelps and others, *Microeconomic Foundations*.

4. "Inflation and Unemployment," p. 7. See also Roger Brinner, "The Accelerationist Model: A Correct Conclusion for the Wrong Reason" (Harvard University, March 1974; processed).

when the market slackens is fewer prospects per week. I emphasize once more that employers do not pay a large cost for their failure to adjust wages instantly; they get almost the same effect from changes in the quality of their labor forces as they would from variations in the scale wage.[5]

Turnover theorists, notably Tobin and Holt, discuss the operation of the labor market in terms closest to those presented here. In Section IV of "Inflation and Unemployment," Tobin sketches a theory of "stochastic macro-equilibrium" that is generally harmonious with the views of this paper.[6] Tobin begins with a distinction between the equilibrium and disequilibrium components of changes in wages: "The first is the rate at which the wage would increase were the market in equilibrium, with neither vacancies nor unemployment. The other component is a function of excess demand and supply. . . ."[7] These correspond roughly to the scale wage and the difference between the effective and scale wages in my exposition. According to Tobin, macro-equilibrium occurs at a positive unemployment rate because of "shocks of demand and technology that keep [individual] markets in perpetual disequilibrium" (p. 10). Thus turnover has a central role in Tobin's thinking. In each market, an unspecified process restores equilibrium by raising wages in the face of excess demand and depressing them in the face of excess supply. The equilibrium component of the wage evolves in response to wages elsewhere in the economy and possibly to its own past values. At the conclusion, Tobin comes close to endorsing Dunlop's view of the influence of key wages: ". . . accidental

5. In a related paper—"Job Search, the Duration of Unemployment, and the Phillips Curve," *American Economic Review*, Vol. 60 (December 1970), pp. 847–62—Mortensen considers a model in which both quality and the wage are instruments of recruiting policy. However, his measure of wages is essentially the same as my scale wage, and he does not consider the implications of changes in the composition of employment for the effective wage or unit labor cost. In his second model, the availability of work does vary in response to the unemployment rate, but workers remain ignorant of the variation.

6. Parts of Tobin's theory were anticipated by Richard Lipsey in "The Relation Between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1862–1957: A Further Analysis," *Economica*, N.S., Vol. 27 (February 1960), pp. 1–31; and G. C. Archibald in "Wage-Price Dynamics, Inflation, and Unemployment," in American Economic Association, *Papers and Proceedings of the Eighty-first Annual Meeting, 1968 (American Economic Review*, Vol. 59, May 1969), pp. 124–34. Although Tobin does not seem to have written on the subject before 1971, his theory became an important part of the oral macroeconomic tradition at the Cowles Foundation many years earlier.

7. "Inflation and Unemployment," p. 9.

circumstances affecting strategic wage settlements also cast a long shadow" (p. 13).

Tobin's theory differs from the model of this paper in two respects. First, he leaves unanswered the basic question of how wages change in response to disequilibrium in individual markets. Holt has addressed exactly this question from the point of view of turnover theory, but I will indicate shortly the defects I find in his theory. Second, the model of the evolution of the scale or equilibrium component of the wage developed here starts from rather a different point. In Tobin's economy, employers and unions are in considerable doubt as to the appropriate level of wages; they need "reference standards" from past history and other industries. Tobin even remarks: "Wage rates for existing employees set the standards for new employees, too" (p. 12). In this model, just the opposite holds. The marginal effective wage, the wage paid to new employees, is precisely the indicator of the appropriate wage level for existing employees needed by both employers and unions. No reference to other industries is needed. The theories wind up at similar conclusions, however. They both imply that the scale or equilibrium component of the wage evolves according to a difference equation that is shifted upward or downward by the disequilibrium wage-adjustment function.

Charles Holt's contributions to turnover theory are complementary to Tobin's.[8] Holt is most concerned with the actual process by which wages change. He takes the flow through the labor market as a given constant (an empirical regularity I have called Holt's law above), and studies the implications of the search rules followed by job seekers for the rate of change of the average wage level. In his model job seekers set an "aspiration wage" when they begin looking for work. As their stretch of unemployment lengthens, they become less confident of finding work at the aspiration level and begin lowering it. At low unemployment rates, job seekers find work rapidly, when they are still accepting only high-paying jobs. In slack markets, the long duration of unemployment forces aspirations down far enough so that the jobs eventually taken pay below the prevailing wage. A central feature of the model of this paper is taken

---

8. "Job Search, Phillips' Wage Relation, and Union Influence"; and James Tobin, "How Can the Phillips Curve be Moved to Reduce Both Inflation and Unemployment?" in Phelps and others, *Microeconomic Foundations*.

directly from Holt: wages rise in tight markets because workers changing jobs take new jobs with higher wages, and drop in slack markets because new jobs have lower wages. However, the influence underlying this process in my theory is rather different. In my model job seekers set their cutoff or aspiration wage in response to conditions in the labor market. Even in the first week of search, whether they take a particular job depends on the availability of other jobs, as measured by the probability of locating other prospects. By contrast, in Holt's model workers arbitrarily set their initial aspiration wage to their past wage plus a constant increment. Their response to conditions in the market occurs only slowly as they remain unemployed. The behavior assumed by Holt conflicts with the evidence that workers have fairly good information about the availability of work within their own labor markets. The equation in my model is compatible with Holt's theory of wage change as well as mine, however.

Tobin, Holt, and other turnover theorists refer extensively to the concept of job vacancies in their expositions. On the other hand, I have constructed a complete model without a single reference to the concept. I am concerned that the symmetry of unemployed workers and unfilled jobs derives from the false belief that both represent unused resources. An unemployed worker is certainly an unused resource but a vacancy is not. The proper analog of an unemployed worker is an unemployed machine or other complementary input. An employer maintains a stock of jobs in the process of being filled precisely to *avoid* idle capital.

In countries that have attempted to control inflation through incomes policies, economists have paid close attention to the gap between scale and effective wages, known in this context as "wage drift." The tendency for actual wage increases to exceed negotiated or controlled increases has frustrated incomes policies in a number of countries. A substantial literature on the relation between wage drift and the excess demand for labor has emerged, starting with the major contribution of Hansen and Rehn.[9] Subsequent work on this problem by continental and British economists has been too extensive to cite or discuss individually here, but generally follows the pattern established by Hansen and Rehn.[10] Wage

9. Bent Hansen and Gösta Rehn, "On Wage-Drift: A Problem of Money-Wage Dynamics," in *25 Economic Essays in Honour of Erik Lindahl* (Stockholm: Ekonomisk Tidskrift, 1956).

10. A few American economists have written on the subject, but always with reference to European economies, as far as I know. Lloyd Ulman and Robert J. Flanagan deal

drift arises from three sources: changes in the productivity of pieceworkers, variations in overtime, and variations in actual straight-time rates from scale rates. Many authors mention upgrading within the firm as part of the third source, and Hansen and Rehn suggest as an aside that movements of workers among firms has the same effect. However, turnover and the mobility of labor do not have a large role in any of the discussions of wage drift. Most studies take the point of view of the empirical literature on the Phillips curve: there is a gross empirical relation between unemployment and wage changes arising from an unexplained influence of excess demand on wages.
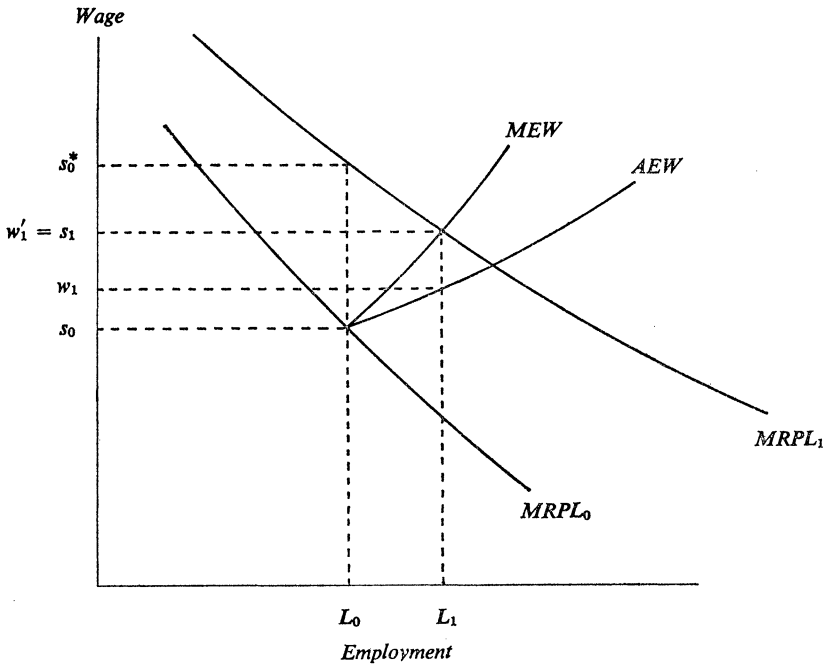
APPENDIX B

Evolution of the Wage under Collective Bargaining

THE CONTENT of this appendix appears in Figure B-1. Ex ante, collective bargaining has established the scale wage, s_0. This is a point on the marginal revenue product schedule, called $MRPL_0$. The corresponding level of employment is L_0. However, demand turns out to be higher than expected, and $MRPL_1$ is the firm's actual demand function for labor. Collective bargaining has established a set of rules by which labor input can be increased, primarily through promotion of existing workers and new hiring at the entry level. The result is seen by the firm as the marginal effective wage schedule, MEW, the supply function to the firm in the short run. Ex post, the firm employs L_1 at a marginal effective wage of w_1'. If collective bargaining had taken place after the shift in demand, then the scale wage would have been s_0^*, and employment would have remained at L_0. After the unexpected shift in demand and consequent unexpected increase in employment, the union has responsibility for an increased membership. The scale wage giving employment to its new membership,

Figure B-1. Impact of Unexpected Shift in Demand for Labor in an Industry with Union Collective Bargaining

Source: See accompanying text in Appendix B, where the symbols are defined.

with demand given by $MRPL_1$, is s_1, which is the same as w_1'. By setting this wage, the union collects the intramarginal part of the new wage bill, $(s_1 - w_1)(L_1 - L_0)$, which accrues to the firm in the first instance. In the figure, AEW signifies the average effective wage.

APPENDIX C

Strategies of Job Seekers

LET $F(w^*)$ be the probability that the wage quoted for a particular prospect is less than a cutoff wage, w^*. I assume that the probability distribution for the prospects available to a worker of grade j is $F(w^*/sr_j)$; all prospective wages rise in proportion to the scale wage, and the distribution

is higher for higher grades in proportion to the wage structure as measured by r_j. The strategy consists in setting a cutoff wage, w^*, and continuing to look for work until a prospect appears that pays a wage at least as high as w^*. Then the weekly probability of taking a job, h, is the product of the probability of locating a prospect and the probability of taking the prospect:

$$h(w^*) = p\left[1 - F\left(\frac{w^*}{sr_j}\right)\right].$$

Suppose, as in the text, that there is a constant probability, ϕ, of losing or leaving a job each week. Then the worker expects to be employed only a fraction of the time,

$$m(w^*) = \frac{h(w^*)}{h(w^*) + \phi}.$$

Now the expected wage upon accepting a job is

$$w'(w^*) = \frac{1}{1 - F\left(\frac{w^*}{sr_j}\right)} \int_{w^*}^{\infty} \omega f\left(\frac{\omega}{sr_j}\right) d\omega,$$

which exceeds w^*. In this expression, f is the probability density function associated with the cumulative distribution function, F;

$$\frac{1}{1 - F\left(\frac{w^*}{sr_j}\right)}$$

is used to scale for the probability of the worker's accepting the job. The expected weekly return is the product of the fraction of the time employed and the average wage when employed, $m(w^*)w'(w^*)$. I assume that the job seeker sets w^* to maximize this return. It turns out that the optimal strategy is to take any job that pays at least as much as the expected return to the combination of waiting and working. Thus w^* is defined implicitly by

$$w^* = m(w^*)w'(w^*).$$

How does the strategy, measured by w^*, respond to changes in conditions in the labor market? It is not hard to show that

$$\frac{\partial w^*}{\partial p} = w' \frac{\partial m}{\partial h} \frac{\partial h}{\partial p},$$

which is positive since all three of its factors are positive. That is, with a

constant distribution of the wages of job prospects, the tighter is the market (the higher is p), the more selective is each job seeker in deciding whether to take a prospect. For the worker making the optimal choice of w^*, w', the average new wage, becomes a function of s, r_i, and p:

$$w' = sr_i g(p).$$

The function $g(p)$ will exceed one when the market is tight and will be less than one when it is slack.

APPENDIX D

Data Sources

THE FOLLOWING paragraphs detail the sources for the data used in the study of five manufacturing employers and in the aggregate wage equation.

Five Manufacturing Employers

SCALE WAGES

Chronologies of wage changes determined by collective bargaining are published by the Bureau of National Affairs in *Collective Bargaining: Negotiations and Contracts* (updated periodically). The chronologies give the month and year of each change, including deferred increases negotiated in earlier years and increases under cost-of-living escalators. Starting from a benchmark in 1958, I applied the reported changes cumulatively to obtain annual averages of the scale wage. The benchmark was the straight-time average hourly earnings in the industry of the employer as calculated by the U.S. Bureau of Labor Statistics (see below).

BLS WAGE

Employment and Earnings reports gross average hourly earnings, average overtime hours, and average hours. The Bureau of National Affairs

calculates average straight-time hours from these data on the assumption that all overtime is paid at time and a half. The data appear in *Collective Bargaining*. Employers and industries are matched as follows:

Employer	*Industry and SIC code*
Clothing manufacturers (men's)	Men's and boys' suits and coats (231)
General Motors Corporation	Motor vehicles and equipment (371)
General Electric Company	Electrical equipment and supplies (36)
B. F. Goodrich Company	Tires and inner tubes (301)
American Viscose Division, FMC Corporation	Weaving mills, synthetics (222)

WAGE DATA FROM FRANK GOLLOP

Gollop constructed his indexes of labor input by calculating estimates of employment and annual hours of work cross-classified by sex, age, education, and occupation.[1] He uses an elaborate interpolation procedure to form these estimates from detailed data from the Census of Population, less detailed annual data from the Current Population Survey, and in-dustry aggregates from establishment data, all collected by the Bureau of Labor Statistics. He then weights these measures of labor input of the various demographic and occupational groups by their relative rates of compensation to form indexes of total labor input to the industries. Rather than using fixed weights throughout the period, he forms a chain or Divisia index by linking together estimates of annual changes in labor input. For each annual change he holds the weights constant, so his index is not very different from a fixed-weight index, and seems an appropriate one for my purposes. The series for effective wages by industries is then calculated as the ratio of total compensation (including supplements) to the index of labor input. Employers and industries are matched as follows:

Employer	*Industry*
Clothing manufacturers (men's)	Apparel and other fabricated textile products
General Motors	Motor vehicles and motor vehicle equipment
General Electric	Electrical machinery
B. F. Goodrich	Rubber and miscellaneous plastic products
American Viscose	Textile mill products

1. Gollop, "Modelling Technical Change," Chap. 1 and Appendixes A–E.

UNEMPLOYMENT RATE

Employment and Earnings reports the unemployment rate for workers who are employed in blue-collar jobs or whose most recent occupations were blue-collar (craftsmen, operatives, and laborers). A regression of the blue-collar rate, u_B, on the official rate, u_O, gives

$$u_B = -1.7 + 1.6\,u_O.$$

Aggregate Wage Equation

WAGE DATA FROM PETER CHINLOY

Chinloy applies a method similar to Gollop's in calculating a Divisia index of labor input to the total private domestic U.S. economy.[2] A particularly useful contribution is his detailed reconciliation of the disparities in the two major sources for data on employment, the household survey and the establishment survey of the BLS. Again, the effective wage is the ratio of total compensation from the national income accounts to the index of total labor input.

UNEMPLOYMENT INDEX WITH FIXED WEIGHTS

The index has the form

$$u_F = \sum v_i u_i.$$

The weight, v_i, reflects the contribution of demographic group i to the unemployment of labor input measured in efficiency units in the base year, 1964. Each weight was calculated by multiplying George Perry's wage-hour weight[3] by the fraction of the labor force in the group in 1964. The weights, normalized to sum to one, are as follows:

2. Chinloy, "Issues in the Measurement of Labor Input," Chap. 1.
3. "Changing Labor Markets and Inflation," pp. 439–41.

Age group	Male	Female
16–17	0.006	0.004
18–19	0.006	0.007
20–24	0.054	0.025
25–34	0.186	0.035
35–44	0.210	0.048
45–54	0.187	0.048
55–64	0.126	0.029
65 and over	0.023	0.006

The official aggregate unemployment rate, u_O, has a noticeable upward trend relative to u_F, owing to the shift in the composition of the labor force toward groups with high unemployment rates. The relation between the two is described fairly well by the following regression:

$$u_O = 0.76 + 1.033\, u_F + 0.059t.$$

The trend variable, t, has the value -1 in 1963, 0 in 1964, 1 in 1965, and so forth. In ten years, u_O shifts upward relative to u_F by 0.6 percentage point.

Comments and Discussion

Robert J. Gordon: Hall's paper makes a major contribution with its theoretical analysis of the role of turnover and quality adjustments in the process of wage setting. But it is seriously deficient when viewed as an attempt at comprehensive analysis of the inflation process. The major problem is the failure to assign any role to prices in the setting of wages; indeed, scarcely anywhere in this entire paper on the "inflation process" does the word "prices" appear. I have become convinced that monetary and fiscal policies influence wages not directly by raising or reducing the rate of unemployment, but indirectly through a chain of effects, first, on excess demand, then on prices in the product market, and finally on the demand for labor. My own contributions to *Brookings Papers* have estimated the effect on wages of product prices (as distinguished from consumer prices); and the experience of 1973, when rising farm prices raised consumer prices relative to nonfarm product prices and wages, can be cited as a prime example of the independent "channel of influence" through which product prices make themselves felt. If we are interested in *anything* about wages in the context of 1974 food and energy inflation, it is the channels by which an exogenous shock in prices relative to wages feeds through to the behavior of nonfarm wages.

Hall has responded in part to this criticism in his final section, "Postscript on the Role of Prices." His basic argument is that the product price adds no information beyond that already contained in the unemployment variable. Consider a standard labor market diagram with the nominal wage rate on the vertical axis and manhours on the horizontal axis, and with a negatively sloped labor demand curve that depends on the product price, and a positively sloped labor supply curve that depends on the ex-

394

pected consumer price level. An increase in the product price relative to
workers' expectations about the price level—which occurs because these
expectations take time to adjust—will shift the demand curve upward to a
new intersection northeast of the original equilibrium position. In this
simple case Hall is correct in asserting that the product price adds no
information beyond that provided by the unemployment variable. The
product price simply measures a vertical shift in labor demand, while the
unemployment rate, usually introduced as a proxy for the net excess de-
mand for labor, measures the same shift horizontally. When workers'
expectations fully adjust to the new, higher, price level, the labor supply
curve shifts up by the amount of the original shift in the demand curve, and
the unemployment rate returns to its initial level. Rising wages and in-
creasing unemployment are observed simultaneously, an anomaly explained
in the conventional framework by lags in the formation of price expecta-
tions, and in Hall's framework by lags in the adjustment of the scale wage
to the effective wage.

The redundancy of the product price variable disappears, however, in an
individual labor market where labor input is supplied inelastically (that is,
a market characterized by a vertical supply curve). For instance, an exog-
enous increase in the relative price of coal, such as has occurred during
1974, would raise the equilibrium nominal wage of coal miners without
simultaneously changing their unemployment rate, *if* coal miners were
initially fully employed and if union restrictions or geographical isolation
prevented the recruitment of new coal miners from other industries or
regions. An aggregate wage equation in which the explanatory variables
include both unemployment and the relation between the current product
price and expected consumer prices could explain the bulge in the aggregate
wage index caused by higher coal wages, but Hall's equation would fail,
since it includes only unemployment and lagged wages.[1]

This example is not an empirical oddity, but may be potentially impor-
tant to the extent that the prime-age segment of the adult-male labor force
is characterized by a vertical labor supply curve. An external event occur-

1. I implicitly assume that nominal income, output, prices, and unemployment in the
noncoal sector are unchanged. An alternative assumption more in tune with recent events
would have policymakers hold nominal income constant in the whole economy, in which
case the higher price of coal would then require a reduction in noncoal real output and
an increase in noncoal unemployment. Hall's equation would be particularly inadequate
in this case, since the aggregate wage index and the aggregate unemployment rate would
be observed to increase together.

ring in a prosperous period—say, a devaluation—might raise the wage rate of adult males without any appreciable effect on their unemployment rate if initially their unemployment rate were near its irreducible "frictional" minimum.

Recent events have reminded us of another route by which prices may raise wage rates without a simultaneous reduction in unemployment. An increase in the relative price of farm products or foreign oil can raise the consumer price index without directly affecting the price of nonfarm products.[2] In graphic terms, in the nonfarm sector, the labor supply curve shifts up along the fixed labor demand curve as expectations adjust to the higher consumer prices, and employment falls while the wage rate increases. Once again Hall's approach provides no explanation of the simultaneous increase in wage rates and unemployment.

In the United States the price deflator for consumer goods includes two important categories—farm products and imports—that are excluded from the index for nonfarm product prices, while the latter includes three categories—exports, structures, and producers' durable equipment—that are excluded from the former. Numerous economic circumstances can cause the two indexes to diverge, and the private nonfarm wage rate will in general depend on both. In my own work for *Brookings Papers*, the wage rate depends on both price indexes, and the distinction between them adds explanatory power in a period like 1956, when an increase in the prices of investment goods relative to consumer goods contributed to an unusually rapid acceleration of wage increases.[3] In this light, the slow growth of wage rates relative to consumer prices in 1973–74 can be seen partly as the result of the reduction in the ratio of private nonfarm prices to consumer prices, which in turn was due to higher farm and oil prices; the more immediate effect of devaluation on import than on export prices; and finally the price controls in the nonfarm sector.

Although Hall's paper makes an important contribution to the theory of labor markets, it is not as general as one might wish. Hall adds varia-

2. Here I ignore substitution between domestic and foreign oil, which would tend to drive up the domestic nonfarm price index.

3. Between 1955:4 and 1956:4 the price deflator for producers' durable equipment increased by 7.9 percent, as compared with an increase in the deflator for consumption of only 3.0 percent. A very large increase in social security contributions also raised wages during that period. Hall's equation, which omits both of these factors, is forced to attribute all of the wage increase to lower unemployment, leading to a short-run Phillips curve that is too steep.

tions in labor quality to wage changes as a second instrument of employer policy. Unlike Mortensen,[4] who allows simultaneous adjustment of worker quality and wage rates, Hall restricts the firm to quality adjustments in the short run, allowing adjustments of the scale wage only in subsequent periods. While this dynamic specification may be appropriate for the unionized sector of the economy with its long-term contracts, in a very substantial portion of the labor market wages may change frequently and at variable intervals. Wachter's paper in this issue provides ample evidence of the longer adjustment lags in the union sector, relative to nonunion industries. For the sizable nonunion segment of the labor market, I prefer Mortensen's approach:

If in the "short run" there is a general increase in demand for the products of firms, then each firm lowers the skill requirements attached to its vacancies and attempts to raise its wage offer relative to all others in the market in order to attract labor at a more rapid rate. The consequence of this joint strategy is that wage inflation is systematically associated with general lowering of minimum skill requirements.[5]

For the union sector and for some nonunion jobs in large organizations, however, Hall does advance beyond Mortensen in emphasizing that the bureaucratically set scale wage introduces a lag between changes in labor market conditions and in measured wage rates.

The Hall and Mortensen approaches share the disadvantage of operating symmetrically during periods of excess labor demand and supply. In Hall's model a decline in product demand leads initially to the appearance of more able applicants in the personnel office, and then to reductions in the scale wage. In many firms high-quality applicants would not be hired during a recession as long as experienced laid-off workers had not been recalled. Hall nevertheless advances the analysis by pointing to the possibility of downward flexibility of labor costs during a recession despite an inflexible scale wage. Even if a firm is not actively hiring, it may enhance the quality of its work force by laying off those workers who are least productive relative to their wage—that is, those on whom the firm makes the smallest profit.

Beyond his distinction between the scale and effective wages, Hall shows how the marginal and average effective wages can diverge. This is a novel

4. Dale T. Mortensen, "Job Search, the Duration of Unemployment, and the Phillips Curve," *American Economic Review*, Vol. 60 (December 1970), pp. 847–62.

5. Ibid., p. 855.

and interesting idea: new workers may be paid more than old ones in the sense that they provide labor services of lower quality for the same wage, and yet old workers may not quit if their wage rate contains a premium for the specific human capital that their employer has invested in them. Unfortunately, an empirical test of this source of lags in the response of average wage rates to changes in product demand cannot be made without personnel records. On a theoretical basis, the same lag in the data can be explained in the context of the formation of price expectations.

While the elimination of price expectations from the wage equation might be justified as one step toward eliminating unobservable variables, Hall's scale wage is just as unobservable in the aggregate economy and must be estimated by extrapolating observable wages, just as the expected rate of inflation has been estimated from data on the response of interest rates to past inflation. In fact, one can argue that expected inflation is more readily observable than the scale wage, given the substantial and growing body of U.S. survey data on the question.

The primary policy conclusions of Hall's empirical estimates are (1) that the natural rate of unemployment is 5.5 percent, so that the unemployment rate must be held above that level if the rate of wage change is to decelerate, and (2) that the slope of the short-run Phillips curve relationship is fairly steep, with an increase of 1 percentage point in the unemployment rate sufficient to reduce the marginal effective wage by 1.0 percent. The 5.5 percent estimate of the natural rate is somewhat higher than that in my own work, but not by enough to merit comment, especially since the data are not capable of drawing fine lines between hypotheses. But the steep slope of Hall's Phillips curve may mislead policymakers into believing that a short, sharp recession that raises the unemployment rate to, say, 7.0 percent will cause an immediate and marked reduction in wage inflation, as occurred in Hall's data in 1958 and 1961.

There are four possible sources of error in Hall's slope coefficient. First, his omission of tax and price variables leads his equation to attribute an acceleration in wage inflation entirely to reduced unemployment in episodes like 1956 and 1966 when taxes and product prices rose during periods of falling unemployment. Second, using compensation instead of earnings corrected for changes in interindustry employment mix means treating a procyclical increase in high-wage employment in durable goods industries as an increase in the wage rate, thus raising the estimated coefficient on the unemployment variable. Third, measurement error tends to intensify the

uptrend of the compensation series relative to wage rates during 1956–57, 1964, and 1966.[6] Finally, sensitivity tests with my data sources indicate that the coefficient on unemployment in the wage equation became substantially smaller in the late 1960s compared with the late 1950s, and thus a 1961-type recession may take substantially longer to tame inflation than Hall's results suggest.[7] This shift in structure may have occurred in response to the greater availability of unemployment benefits, welfare payments, and food stamps which reduce the burden of unemployment, tending both to lengthen it and relieve the pressure on workers to accept lower wages.

James Tobin: As always, Bob Hall has given us a stimulating and thought-provoking paper. It is full of insights on labor markets and wage setting, which in some degree reconcile the often divergent views of labor economists and macro theorists. I believe the paper opens a number of promising lines of inquiry.

It has been a dismal experience over the years to come to the Brookings panel meetings and see the empirical long-run Phillips curve shift steadily upward and become steeper and steeper. Now Hall has outdone his predecessors, including himself, and tells us that the natural rate of unemployment—he calls it the equilibrium rate—exists and is about 5.5 percent. I find one consolation in this estimate. No one can identify a 5.5 percent natural rate with optimal unemployment or voluntary search unemployment or frictional unemployment matched by vacancies. If the lowest unemployment rate consistent with nonaccelerating inflation is so high, it must be because the wage-setting institutions of this economy have an inflationary bias. I do not think Hall himself attaches any normative significance to his equilibrium rate.

This paper gives no prescription for diminishing the equilibrium rate, nor are the appropriate measures apparent from the structure of the model. According to Hall, the basic problem is that the unemployment rate must rise above 5.5 percent before employers can find workers at the prevailing wage scale who are qualified for the job categories for which replacement

6. See Figure 1 in my "Inflation in Recession and Recovery," *Brookings Papers on Economic Activity* (1:1971), p. 117.

7. Walter Y. Oi, "On Measuring the Impact of Controls: A Critical Appraisal" (paper presented to the Rochester Conference on Wage-Price Controls, November 9, 1973; processed), Table 1.

hiring or new recruiting is necessary. In tighter labor markets employers are continuously placing new and old workers in jobs in which their productivity falls short of the scale wage.

Hall's basic insight is his distinction between the scale wage, s, and the marginal effective wage, w'. The wage scale is designed to put workers in jobs for which they are qualified—that is, for an allocation of worker types to job types that maximizes output per effective (or quality-weighted) labor input. When the labor market is tight, as measured by some unemployment rate u, w' exceeds s because new workers are placed in jobs beyond their qualifications. When the labor market is slack, w' is smaller than s because new workers are available who are overqualified for the jobs to which they are assigned. The equation is $w' = sg(u)$, where $g(u)$ is the wage-adjustment function, and the natural rate of unemployment is u^* where $g(u^*) = 1$.

According to Hall, employers—or employers and unions in collective bargains—attempt to set the scale wage each period to equal their estimate of the marginal effective wage for that period. Their belief is that $\ln w'$ depends on its own lagged values. Thus s depends on lagged values of w', which means it depends on lagged values of $sg(u)$. The weight of any past year's scale wage in today's setting of that wage therefore depends on how tight the labor market was that year.

THE FORMAL MODEL

Hall's model of wage determination is a new version of the accelerationist, or natural-rate, hypothesis. In order to understand it, I have found it helpful to look explicitly at its formal properties.

The basic equation is a difference equation in the marginal effective wage, w':

$$(1) \qquad \ln w'(t) = \sum_{\tau=1}^{n} \beta_\tau \ln w'(t - \tau) + \ln g(u(t)).$$

(The numbering of equations here does not correspond to Hall's.) The summation on the right is the prediction of $\ln w'(t)$ from past history. Hall denotes it $\ln Ew'(t)$, and sets the scale wage $s(t)$ equal to it. (I will simply use a hat to denote a predicted value—for example, $\hat{w}$.) He assumes throughout that

$$\sum_{\tau=1}^{n} \beta_\tau = 1.$$

When $g(u(t))$ exceeds 1, the actual wage $w'(t)$ is pushed above the scale wage—that is, above the predicted value $\hat{w}'(t)$. The reverse occurs when $g(u(t))$ is less than 1.

For notational economy, let $z(t) = \ln w'(t)$ and $f(t) = \ln g(u(t))$. Let us also, following Hall's empirical specification, let $n = 3$. Equation (1) then becomes

(2) $\quad z(t) - \beta_1 z(t-1) - \beta_2 z(t-2) - (1 - \beta_1 - \beta_2)z(t-3) = f(t).$

Following Hall, let

$$\mu = \sum_{i=1}^{n} \beta_r \tau = \beta_1 + 2\beta_2 + 3\beta_3.$$

Then the two parameters of the difference equation are β_1 and μ:

(3) $\quad z(t) - \beta_1 z(t-1) - (3 - \mu - 2\beta_1)z(t-2)$
$$- (\beta_1 - 2 + \mu)z(t-3) = f(t).$$

I consider several cases:

(a) $\mu = 0$, $f(t) = 0$ for all t. When $\mu = 0$, according to Hall, there is only one unemployment rate compatible with steady nonaccelerating inflation or deflation, the one for which $g = 1$ and $f = 0$. At this natural rate of unemployment, any steady rate of inflation is possible.

(b) $\mu = 0$, $f(t) = f \neq 0$ for all t. According to Hall, there will be accelerating inflation when f is a constant positive number, accelerating deflation when f is a constant negative number.

(c) $\mu \neq 0$, $f(t) = f$ for all t. In these cases, according to Hall, there is for any given value of μ a relationship $m(f)$ which tells the steady inflation rate consistent with each value of f.

Given case (a), in which $\mu = 0$, $f(t) = 0$, the general solution of (3) is

(4) $\quad\quad\quad\quad z(t) = c_1 + c_2 t + c_3(\beta_1 - 2)^t.$

Recall that, in this case, $\beta_1 - 2 = \beta_3$.

The cs are related to the initial conditions as follows:

(5) $\quad\quad\quad \begin{cases} c_1 \quad\quad\; + c_3 \quad\quad\; = z(0) \\ c_1 - c_2 + c_3\beta_3^{-1} = z(-1) \\ c_1 - 2c_2 + c_3\beta_3^{-2} = z(-2). \end{cases}$

In particular, if the initial values of z show steady inflation—that is, $z(-1) = z(0) - m$, $z(-2) = z(0) - 2m$—then $c_3 = 0$, $c_2 = m$, $c_1 = z(0)$, and $z(t) = z(0) + mt$. This is Hall's "natural rate" proposition for this case. Equation (4) also shows that the steady inflation path is stable pro-

vided the absolute value of β_3 is smaller than 1—that is, $1 < \beta_1 < 3$. The model per se does not exclude instability, but Hall's empirical estimates are well within the interval for stability. Of course, in a stochastic version of (3) and (4), the value of c_2, the component that represents the steady inflation rate, m, would vary, and random shocks could also keep the third component of (4) from vanishing.

For case (b), in which $\mu = 0, f(t) = f \neq 0$, the solution is

$$(6) \qquad z(t) = c_1 + c_2 t + c_3(\beta_1 - 2)^t + \frac{ft^2}{2(3 - \beta_1)}.$$

The final term of (6) is the "acceleration" component. It has the same sign as f if β_1 is less than 3, and it will dominate the path if β_1 is also larger than 1. If β_1 exceeds 3, the third term will dominate, though in which direction it carries the rate of inflation depends on the initial conditions rather than on the sign of f. If β_1 is less than 1, the third term gives an explosive sawtooth oscillation, which will eventually dominate. Thus Hall's accelerationist proposition—that $f \neq 0$ leads to accelerating inflation when $\mu = 0$—is consistent with his empirical estimates of β_1.

Fluctuation of u around the equilibrium rate u^* for which $f = 0$, will activate the final term of (6) even if f averages zero. Moreover, because of the nonlinearity of $f(u)$, unemployment averaging u^* means that f is on average positive.

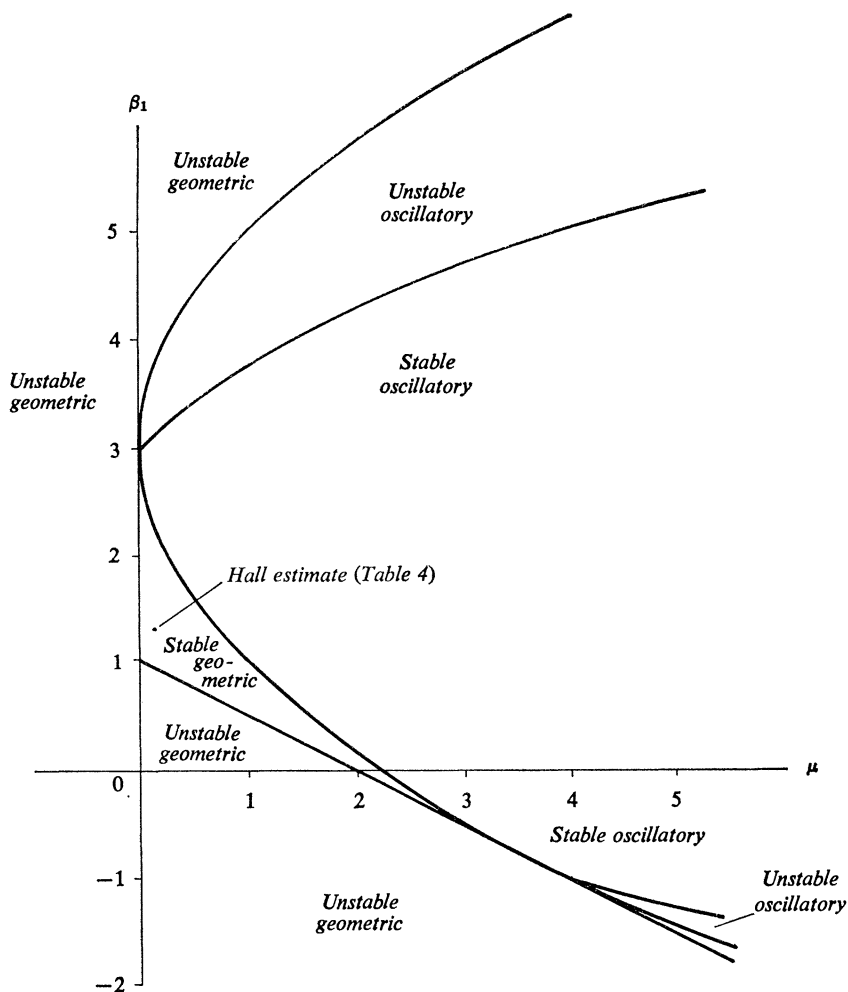For case (c), in which $\mu \neq 0, f(t) = f$, the solution is

$$(7) \qquad z(t) = c_1 + c_2 x_2^t + c_3 x_3^t + (f/\mu)t.$$

Here

$$(8) \qquad x_2 = \frac{\beta_1 - 1}{2} + \frac{\sqrt{(3 - \beta_1)^2 - 4\mu}}{2}$$

$$x_3 = \frac{\beta_1 - 1}{2} - \frac{\sqrt{(3 - \beta_1)^2 - 4\mu}}{2}.$$

In the diagram below the various possibilities of (7) in relation to β_1 and μ are graphed. Here stability means that both x_2 and x_3 are less than 1 in absolute value (or that their modulus is if they are conjugate complex roots). If the stability conditions are satisfied, then steady inflation at the rate f/μ eventually dominates the path. As the diagram indicates, there is *always* instability at negative values of μ; and for positive μ extreme values of β_1 are associated with instability. Hall's empirical estimates fall in the area where the model is stable and nonoscillatory.

The relation $m = f/\mu$ (for $\mu \neq 0$) is a long-run Phillips curve. For $\mu > 0$ it implies a higher steady-state inflation for a lower unemployment rate. For $\mu < 0$ the long-run Phillips curve slopes the wrong way, but it is in any case irrelevant since the path is unstable.

COMPARISON WITH STANDARD THEORY

According to standard theory, a sufficient condition for acceleration is $\Sigma\beta_r = 1$. In Hall's model the additional condition $\mu = \Sigma\beta_r\tau = 0$ is necessary. His empirical results do not reject the hypothesis that this condition

is met. But neither do they exclude positive values for μ. In any case, the paper leaves in mystery the source of the additional condition. For example, the popular model of adaptive expectations implies a positive μ, since all the β_τ are positive, in geometrically decaying series. Yet it has been generally accepted that this model is accelerationist, provided that $\Sigma \beta_\tau = 1$ and that the coefficient of expected wage (or price) inflation is equal to unity in the wage equation.

To illuminate the mystery, I shall compare a standard model and Hall's model. It is more convenient, and I think less artificial, to use continuous rather than discrete time. The symbols are the same as in the preceding section; $\hat{z}$ represents the anticipated value of z, and the notation $x'(t)$ means the derivative of x with respect to t.

My version of the standard model is as follows:

$$(9) \qquad \hat{z}(t) = \int_{-\infty}^{t} \beta(t - x) z(x)dx = \int_{0}^{\infty} \beta(\tau)z(t - \tau)d\tau,$$

where

$$\int_{0}^{\infty} \beta(\tau)d\tau = 1.$$

In the adaptive expectations case,

$$\beta(\tau) = \gamma e^{-\gamma \tau}$$
$$\mu = 1/\gamma.$$

Differentiating (9) with respect to t gives

$$(10) \qquad \hat{z}'(t) = \beta(0) z(t) + \int_{-\infty}^{t} \beta'(t - x)z(x)dx = \beta(0)z(t)$$
$$+ \int_{0}^{\infty} \beta'(\tau)z(t - \tau)d\tau.$$

Note that

$$\int_{0}^{\infty} \beta'(\tau)d\tau = -\beta(0).$$

Differentiating (10) with respect to t gives

$$(11) \qquad \hat{z}''(t) = \beta(0) z'(t) + \int_{0}^{\infty} \beta'(\tau) z'(t - \tau)d\tau.$$

The short-run Phillips relation is

$$(12) \qquad\qquad z'(t) = \alpha \hat{z}'(t) + f(t),$$

where the natural-rate proposition is that $\alpha = 1$. Combining (10), (11), and (12) gives

$$(13) \qquad \frac{\hat{z}''(t)}{\beta(0)} = f(t) + \alpha z'(t) + \int_0^\infty \frac{\beta'(\tau)}{\beta(0)} z'(t - \tau)d\tau.$$

In case of adaptive expectations,

$$\frac{\beta'(\tau)}{\beta(0)} = -\beta'(\tau),$$

and the last term in (13) is $-\hat{z}'(t)$. Thus (13) says, when $\alpha = 1$, that acceleration $\hat{z}''$ will have the same sign as f. For general $\beta(\tau)$, $\hat{z}''$ will also depend on the past history of z'. But there is no long-run Phillips curve, as can be seen by setting all past $z' = m$. The integral in (13) will equal $-m$. Current $\hat{z}'$ will equal m, and there will be zero acceleration only if $f = 0$.

Hall's model differs by replacing (12) with an equation in levels rather than rates of change. With $\alpha = 1$, his short-run "Phillips" curve is

$$(14) \qquad\qquad z(t) = \hat{z}(t) + f(t),$$

or

$$(15) \qquad\qquad z'(t) = \hat{z}'(t) + f'(t).$$

Consequently, in (13), Hall would have $f'(t)$ instead of $f(t)$. This means that zero acceleration is possible when f is constant though not zero, as indicated for case (c) of the previous section.[8]

However, Hall also appears to assume $\beta(0) = 0$. This means that the sum of the $\beta'(\tau)$ is zero. In (11), therefore, there will be zero acceleration of $\hat{z}$ if z' has been forever constant. This is why it is necessary to investi-

---

8. Hall acknowledges equation (15), but he says his short-run Phillips curve is orthodox nonetheless. In proof, he subtracts $z(t - 1)$ from both sides of his discrete-time version of (14). This, of course, cannot change its content, which is that each period the scale wage is set to an effective wage predicted on the assumption that f will be zero. It seems quite arbitrary to assume that the gap between actual and expected wages is closed fully in precisely the time unit of the model. The continuous-time counterpart of Hall's model might be $z' = \eta(\hat{z} - z) + f$. But one would certainly wish to add a $\hat{z}'$ term to that expression, as Hall has not done. He is, I believe, mistaken in his assertion that the standard Phillips curve makes the actual wage change depend, apart from $f(u)$, on the difference between expected and current wage—my $\hat{z} - z$, his $E(\Delta\log w)$—rather than on the change in the expected wage (my $\hat{z}'$, his $\Delta\log \hat{w}$).

gate, as the previous section does, the conditions under which z' can be constant, and why it turns out to be possible even if $f \neq 0$.

Natural-rate theorists will object with some logic to Hall's formulation, which permits his scale wage and effective wage to diverge by a constant percentage forever if $f \neq 0$ and $\mu \neq 0$. Every period employers and workers try to predict the effective wage, and every period they underpredict (or overpredict) by the same margin. Yet they never learn. Clearly, Hall's model needs a catch-up term, so that the scale wage is advanced faster when it is lagging behind the effective wage.

OTHER COMMENTS

Behind Hall's formal model stands a vision of labor markets that invites some comments.

First, Hall's scale wage is the central parameter of a structure of wages corresponding to job classifications, a structure assumed to remain constant over time and the business cycle. Workers come in quality grades, but their wages depend on their job assignments rather than on their productivity. At the equilibrium unemployment rate, an employed worker is assigned to the job for which he has a comparative advantage. The wage scale is geared to this situation. In other states of the labor market, up- or downgrading of workers alters productivity and drives a wedge between effective labor cost and the scale wage.

Job classifications have been determined, capital equipment installed, and relative wage scales set for a particular mix of labor grades. If we are to believe Hall's results, it is the employment mix corresponding to 5.5 percent unemployment. The mix corresponding to 4 percent would be different, particularly in having more low-skill and inexperienced labor, and the comparative-advantage calculations would come out differently. Why do we have one equilibrium rate rather than the other? Why doesn't the availability of unemployed labor alter the structure of wages and jobs and induce a lower equilibrium unemployment rate?

Second, why is the new scale wage set to equal the marginal effective wage? The employer, Hall argues, will not wish his scale wage to diverge from those of other employers, because then he will not attract the most efficient quality mix of workers. This argument seems to assume that the existing mix of employees is proper; if not, marginal recruitment of a proper mix will not rectify the situation. Here as elsewhere, it would help

if Hall specified the production function of the firm in jobs and skill grades, so that the effects of using workers in jobs above or below their grades are clearly derived. Hall also argues that collective bargaining will make no difference, since the union too has an interest in equality of w' and s. This claim seems to be based on the union's responsibility for the jobs of everyone hired, and it is not always justified.

Third, turnover is essential to Hall's theory. Only because there are always some employers and workers in the market, by choice or otherwise, do effective wages change and pull scale wages along. Intersectoral shifts in product demand and labor supply are the sources of constant turnover and mobility. This phenomenon suggests a disaggregated model in which differences among markets, industries, and firms play an essential role. Instead, Hall's model for the economy as a whole is no different from his model for an individual firm. The fiction of the representative firm permits him to ignore aggregation difficulties.

Fourth, Hall's wage adjustment function $g(u)$ takes the form $\psi_0 - \psi_1 \log u$. However, the parameters of $g(u)$ are allowed to take values in the effective wage equation different from those used in the scale wage equation. One of the pairs is just arbitrarily assumed. The difference is contrary to the model, and $g(u)$ should be constrained to be the same wherever it appears in Hall's equation (3).

My last two comments refer to the introductory material. In these days when new mutants of inflation appear before economists have figured out the old types, we should be modest in our claims for our theories. Hall is concerned with the wage-price-wage dynamic in a closed system. Specifically, he is dealing with the part of the U.S. economy Galbraith calls the "planning" system, as distinct from the "market" system. This is where wages, and prices too, are bureaucratically—though not necessarily monopolistically—administered. In this economy competition works indirectly; wages and prices move when excess demand or supply at existing wages or prices induces business firms to change them. Recently we have learned that even the U.S. industrial economy is not immune from external inflationary shocks, whether from overseas or the domestic agricultural sector. In agriculture and other sectors where self-employment is the pattern, wages are not administered and they are not rigid in the face of fluctuations in product demand or prices.

Finally, while I applaud the integrative intent of Hall's opening paragraph, I must register a mild dissent. Inflation certainly is a monetary

phenomenon, if only by its definition as a decline in the purchasing power of the monetary unit of account. But the determinative sovereignty of any particular quantity denominated in the unit of account is another matter. What we can agree is that in the long run the value of a dollar is determined by the magnitudes of all exogenous variables, stocks or flows, denominated in dollars. But this does not commit us to a belief that M_1 or M_2 or M_{19} is sufficient unto itself to determine the price level, or even that it is exogenous.

General Discussion

Considerable discussion centered on Hall's estimate of the equilibrium unemployment rate. Hall thought part of the reason for the high estimate was his use of the quality-corrected wage index. George Perry noted that he had developed and used a similar wage index which displayed a secular rise relative to the official wage index, because the latter was held down artificially by the growing relative importance of lower-wage workers. In a model linking prices to wages via productivity, such a drift would be offset by a comparable drift in quality-corrected productivity relative to the official measure, so long as wages were proportional to workers' productivity. But Hall's model never went through the wage-productivity-price cycle, a fact that need not bias Hall's estimates, though it does make them harder to relate to other studies. In any case, Perry noted, Hall had estimated the equilibrium unemployment rate using his five industries as 5.1 percent. Taken together with his estimate of 5.5 percent from aggregate data, his results are not very different from those of studies that take account of the changing composition of unemployment.

Hall stressed that he was not advocating the equilibrium unemployment rate, but only describing it as the rate that does not cause inflation to accelerate. He did not see any obvious policy prescription arising from his estimate of the equilibrium rate; in particular, he denied any usefulness for public employment, which he characterized as just another way to spend government money and as a movement along, rather than a shift of, the Phillips curve.

James Tobin saw the assumption of fixed relative wages within the firm as unrealistic. Hall pictures an employer, who has a multidimensional labor input problem, as having only one parameter—the scale wage—to

control. Hall replied that this simplification in the theory did not seem crucial to him and that a more elaborate model would become intractable without adding much to his picture of wage determination for the economy as a whole.

Arthur Okun questioned Hall's specification that employers set the scale wage equal to the expected marginal effective wage. Such a practice would be optimal only in a world of certainty equivalence. It would not be optimal in an uncertain world if the loss from setting the scale wage above the marginal effective wage is not equal to the loss from setting it below. There are all sorts of asymmetries in the employer's wage problem: paying too little means getting low quality in the new workers he is hiring; paying more than he needs to means raising the wage of all his workers. He can easily raise a wage through a bonus, but would find it difficult to dock an employee after a wage has been set. Similarly, it is easier to promote an employee than to demote one. On balance, the asymmetries seem to point to setting the scale wage below the marginal effective wage.

Christopher Sims argued that the Hall model should be estimated as a system of at least three simultaneous equations, allowing for the endogeneity of the right-hand variables. Hall agreed, but thought that the parameters would not be very sensitive to specification error.

Sims also offered some insights from the theory of rational expectations on how the adjustments specified in the model might be unrealistic. First, people are likely to change any linear, fixed-weight forecasting scheme in response to the way the variables have been behaving recently. Second, he saw no reason to assume that the weights of the forecasting scheme (the βs) sum to one.