Cowles Foundation Paper 361
Reprinted from: American Economic Review, 62, 1972

Inflation and Unemployment

By JAMES TOBIN*

The world economy today is vastly different from the 1930's, when Seymour Harris, the chairman of this meeting, infected me with his boundless enthusiasm for economics and his steadfast confidence in its capacity for good works. Economics is very different, too. Both the science and its subject have changed, and for the better, since World War II. But there are some notable constants. Unemployment and inflation still preoccupy and perplex economists, statesmen, journalists, housewives, and everyone else. The connection between them is the principal domestic economic burden of presidents and prime ministers, and the major area of controversy and ignorance in macroeconomics. I have chosen to review economic thought on this topic on this occasion, partly because of its inevitable timeliness, partly because of a personal interest reaching back to my first published work in 1941.

I. The Meanings of Full Employment

Today, as thirty and forty years ago, economists debate how much unemployment is voluntary, how much involuntary; how much is a phenomenon of equilibrium, how much a symptom of disequilibrium; how much is compatible with competition, how much is to be blamed on monopolies, labor unions, and restrictive legislation; how much unemployment characterizes "full" employment.

Full employment—imagine macroeconomics deprived of the concept. But what is it? What is the proper employment goal of policies affecting aggregate demand? Zero unemployment in the monthly labor force survey? That outcome is so inconceivable outside of Switzerland that it is useless as a guide to policy. Any other numerical candidate, yes even 4 percent, is patently arbitrary without reference to basic criteria. Unemployment equal to vacancies? Measurement problems aside, this definition has the same straightforward appeal as zero unemployment, which it simply corrects for friction.[1]

A concept of full employment more congenial to economic theory is labor market equilibrium, a volume of employment which is simultaneously the amount employers want to offer and the amount workers want to accept at prevailing wage rates and prices. Forty years ago theorists with confidence in markets could believe that full employment is whatever volume of employment the economy is moving toward, and that its achievement requires of the government nothing more than neutrality, and nothing less.

After Keynes challenged the classical notion of labor market equilibrium and the complacent view of policy to which it led, full employment came to mean maximum aggregate supply, the point at which expansion of aggregate demand could not further increase employment and output.

Full employment was also regarded as the economy's inflation threshold. With a deflationary gap, demand less than full employment supply, prices would be declining or at worst constant. Expansion of aggregate demand short of full employment would cause at most a one-shot

* Presidential address delivered at the eighty-fourth meeting of the American Economic Association, New Orleans, Louisiana, December 28, 1971.

[1] This concept is commonly attributed to W. H. Beveridge, but he was actually more ambitious and required a surplus of vacancies.

increase of prices. For continuing inflation, the textbooks told us, a necessary and sufficient condition was an inflationary gap, real aggregate demand in excess of feasible supply. The model was tailor-made for wartime inflation.

Postwar experience destroyed the identification of full employment with the economy's inflation threshold. The profession, the press, and the public discovered the "new inflation" of the 1950's, inflation without benefit of gap, labelled but scarcely illuminated by the term "cost-push." Subsequently the view of the world suggested by the Phillips curve merged demand-pull and cost-push inflation and blurred the distinction between them. This view contained no concept of full employment. In its place came the tradeoff, along which society supposedly can choose the least undesirable feasible combination of the evils of unemployment and inflation.

Many economists deny the existence of a durable Phillips tradeoff. Their numbers and influence are increasing. Some of them contend that there is only one rate of unemployment compatible with steady inflation, a "natural rate" consistent with any steady rate of change of prices, positive, zero, or negative. The natural rate is another full employment candidate, a policy target at least in the passive sense that monetary and fiscal policy makers are advised to eschew any numerical unemployment goal and to let the economy gravitate to this equilibrium. So we have come full circle. Full employment is once again nothing but the equilibrium reached by labor markets unaided and undistorted by governmental fine tuning.

In discussing these issues, I shall make the following points. First, an observed amount of unemployment is not revealed to be voluntary simply by the fact that money wage rates are constant, or rising, or even accelerating. I shall recall and extend Keynes's definition of involuntary unemployment and his explanation why workers may accept price inflation as a method of reducing real wages while rejecting money wage cuts. The second point is related. Involuntary unemployment is a disequilibrium phenomenon; the behavior, the persistence, of excess supplies of labor depend on how and how fast markets adjust to shocks, and on how large and how frequent the shocks are. Higher prices or faster inflation can diminish involuntary, disequilibrium unemployment, even though voluntary, equilibrium labor supply is entirely free of money illusion.

Third, various criteria of full employment coincide in a theoretical full stationary equilibrium, but diverge in persistent disequilibrium. These are 1) the natural rate of unemployment, the rate compatible with zero or some other constant inflation rate, 2) zero involuntary unemployment, 3) the rate of unemployment needed for optimal job search and placement, and 4) unemployment equal to job vacancies. The first criterion dictates higher unemployment than any of the rest. Instead of commending the natural rate as a target of employment policy, the other three criteria suggest less unemployment and more inflation. Therefore, fourth, there are real gains from additional employment, which must be weighed in the social balance against the costs of inflation. I shall conclude with a few remarks on this choice, and on the possibilities of improving the terms of the tradeoff.

II. Keynesian and Classical Interpretations of Unemployment

To begin with the *General Theory* is not just the ritual piety economists of my generation owe the book that shaped their minds. Keynes's treatment of labor market equilibrium and disequilibrium in his first chapter is remarkably relevant today.

Keynes attacked what he called the classical presumption that persistent unemployment is voluntary unemployment. The presumption he challenged is that in competitive labor markets actual employment and unemployment reveal workers' true preferences between work and alternative uses of time, the presumption that no one is fully or partially unemployed whose real wage per hour exceeds his marginal valuation of an hour of free time. Orthodox economists found the observed stickiness of money wages to be persuasive evidence that unemployment, even in the Great Depression, was voluntary. Keynes found decisive evidence against this inference in the willingness of workers to accept a larger volume of employment at a lower real wage resulting from an increase of prices.

Whenever unemployment could be reduced by expansion of aggregate demand, Keynes regarded it as involuntary. He expected expansion to raise prices and lower real wages, but this expectation is not crucial to his argument. Indeed, if it is possible to raise employment without reduction in the real wage, his case for calling the unemployment involuntary is strengthened.

But why is the money wage so stubborn if more labor is willingly available at the same or lower real wage? Consider first some answers Keynes did not give. He did not appeal to trade union monopolies or minimum wage laws. He was anxious, perhaps over-anxious, to meet his putative classical opponents on their home field, the competitive economy. He did not rely on any failure of workers to perceive what a rise in prices does to real wages. The unemployed take new jobs, the employed hold old ones, with eyes open. Otherwise the new situation would be transient.

Instead, Keynes emphasized the institutional fact that wages are bargained and set in the monetary unit of account. Money wage rates are, to use an unKeynesian term, "administered prices." That is, they are not set and reset in daily auctions but posted and fixed for finite periods of time. This observation led Keynes to his central explanation: Workers, individually and in groups, are more concerned with relative than absolute real wages. They may withdraw labor if their wages fall relatively to wages elsewhere, even though they would not withdraw any if real wages fall uniformly everywhere. Labor markets are decentralized, and there is no way money wages can fall in any one market without impairing the relative status of the workers there. A general rise in prices is a neutral and universal method of reducing real wages, the only method in a decentralized and uncontrolled economy. Inflation would not be needed, we may infer, if by government compulsion, economy-wide bargaining, or social compact, all money wage rates could be scaled down together.

Keynes apparently meant that relative wages are the arguments in labor supply functions. But Alchian (pp. 27–52 in Phelps et al.) and other theorists of search activity have offered a somewhat different interpretation, namely that workers whose money wages are reduced will quit their jobs to seek employment in other markets where they think, perhaps mistakenly, that wages remain high.

Keynes's explanation of money wage stickiness is plausible and realistic. But two related analytical issues have obscured the message. Can there be involuntary unemployment in an equilibrium, a proper, full-fledged neoclassical equilibrium? Does the labor supply behavior described by Keynes betray "money illusion"? Keynes gave a loud yes in answer to the first question, and this seems at first glance to compel an affirmative answer to the second.

An economic theorist can, of course, commit no greater crime than to assume money illusion. Comparative statics is a

nonhistorical exercise, in which different price levels are to be viewed as alternative rather than sequential. Compare two situations that differ only in the scale of exogenous monetary variables; imagine, for example, that all such magnitudes are ten times as high in one situation as in the other. All equilibrium prices, including money wage rates, should differ in the same proportion, while all real magnitudes, including employment, should be the same in the two equilibria. To assume instead that workers' supply decisions vary with the price level is to say that they would behave differently if the unit of account were, and always had been, dimes instead of dollars. Surely Keynes should not be interpreted to attribute to anyone money illusion in this sense. He was not talking about so strict and static an equilibrium.

Axel Leijonhufvud's illuminating and perceptive interpretation of Keynes argues convincingly that, in chapter 1 as throughout the *General Theory*, what Keynes calls equilibrium should be viewed as persistent disequilibrium, and what appears to be comparative statics is really shrewd and incisive, if awkward, dynamic analysis. Involuntary unemployment means that labor markets are not in equilibrium. The resistance of money wage rates to excess supply is a feature of the adjustment process rather than a symptom of irrationality.

The other side of Keynes's story is that in depressions money wage deflation, even if it occurred more speedily, or especially if it occurred more speedily, would be at best a weak equilibrator and quite possibly a source of more unemployment rather than less. In contemporary language, the perverse case would arise if a high and ever-increasing real rate of return on money inhibited real demand faster than the rising purchasing power of monetary stocks stimulated demand. To pursue this Keynesian theme further here would be a digression.

What relevance has this excursion into depression economics for contemporary problems of unemployment and wage inflation? The issues are remarkably similar, even though events and Phillips have shifted attention from levels to time rates of change of wages and prices. Phillips curve doctrine[2] is in an important sense the postwar analogue of Keynesian wage and employment theory, while natural rate doctrine is the contemporary version of the classical position Keynes was opposing.

Phillips curve doctrine implies that lower unemployment can be purchased at the cost of faster inflation. Let us adapt Keynes's test for involuntary unemployment to the dynamic terms of contemporary discussion of inflation, wages, and unemployment. Suppose that the current rate of unemployment continues. Associated with it is a path of real wages, rising at the rate of productivity growth. Consider an alternative future, with unemployment at first declining to a rate one percentage point lower and then remaining constant at the lower rate. Associated with the lower unemployment alternative will be a second path of real wages. Eventually this real wage path will show, at least to first approximation, the same rate of increase as the first one, the rate of productivity growth. But the paths may differ because of the transitional effects of increasing the rate of employment. The growth of real wages will be retarded in the short run if additional employment lowers labor's marginal productivity. In any case, the test question is whether with full information about the two alternatives labor would accept the second one—

[2] Phillips himself is not a prophet of the doctrine associated with his curve. His 1958 article was probably the most influential macro-economic paper of the last quarter century. But Phillips simply presented some striking empirical findings, which others have replicated many times for many economies. He is not responsible for the theories and policy conclusions his findings stimulated.

whether, in other words, the additional employment would be willingly supplied along the second real wage path. If the answer is affirmative, then that one percentage point of unemployment is involuntary.

For Keynes's reasons, a negative answer cannot necessarily be inferred from failure of money wage rates to fall or even decelerate. Actual unemployment and the real wage path associated with it are not necessarily an equilibrium. Rigidities in the path of money wage rates can be explained by workers' preoccupation with relative wages and the absence of any central economy-wide mechanism for altering all money wages together.

According to the natural rate hypothesis, there is just one rate of unemployment compatible with steady wage and price inflation, and this is in the long run compatible with any constant rate of change of prices, positive, zero, or negative. Only at the natural rate of unemployment are workers content with current and prospective real wages, content to have their real wages rise at the rate of growth of productivity. Along the feasible path of real wages they would not wish to accept any larger volume of employment. Lower unemployment, therefore, can arise only from economy-wide excess demand for labor and must generate a gap between real wages desired and real wages earned. The gap evokes increases of money wages designed to raise real wages faster than productivity. But this intention is always frustrated, the gap is never closed, money wages and prices accelerate. By symmetrical argument, unemployment above the natural rate signifies excess supply in labor markets and ever accelerating deflation. Older classical economists regarded constancy of money wage rates as indicative of full employment equilibrium, at which the allocation of time between work and other pursuits is revealed as voluntary and optimal. Their successors make the same claims for the natural rate of unemployment, except that in the equilibrium money wages are not necessarily constant but growing at the rate of productivity gain plus the experienced and expected rate of inflation of prices.

III. Is Zero-Inflation Unemployment Voluntary and Optimal?

There are, then, two conflicting interpretations of the welfare value of employment in excess of the level consistent with price stability. One is that additional employment does not produce enough to compensate workers for the value of other uses of their time. The fact that it generates inflation is taken as prima facie evidence of a welfare loss. The alternative view, which I shall argue, is that the responses of money wages and prices to changes in aggregate demand reflect mechanics of adjustment, institutional constraints, and relative wage patterns and reveal nothing in particular about individual or social valuations of unemployed time vis-à-vis the wages of employment.

On this rostrum four years ago, Milton Friedman identified the noninflationary natural rate of unemployment with "equilibrium in the structure of real wage rates" (p. 8). "The 'natural rate of unemployment,' " he said, ". . . is the level that would be ground out by the Walrasian system of general equilibrium equations, provided that there is embedded in them the actual structural characteristics of the labor and commodity markets, including market imperfections, stochastic variability in demands and supplies, the costs of getting information about job vacancies and labor availabilities, the costs of mobility, and so on." Presumably this Walrasian equilibrium also has the usual optimal properties; at any rate, Friedman advised the monetary authorities not to seek to improve upon it. But in fact we know little about the existence of a

Walrasian equilibrium that allows for all the imperfections and frictions that explain why the natural rate is bigger than zero, and even less about the optimality of such an equilibrium if it exists.

In the new microeconomics of labor markets and inflation, the principal activity whose marginal value sets the reservation price of employment is job search. It is not pure leisure, for in principle persons who choose that option are not reported as unemployed; however, there may be a leisure component in job seeking.

A crucial assumption of the theory is that search is significantly more efficient when the searcher is unemployed, but almost no evidence has been advanced on this point. Members of our own profession are adept at seeking and finding new jobs without first leaving their old ones or abandoning not-in-labor-force status. We do not know how many quits and new hires in manufacturing are similar transfers, but some of them must be; if all reported accessions were hires of unemployed workers, the mean duration of unemployment would be only about half what it is in fact. In surveys of job mobility among blue collar workers in 1946–47 (see Lloyd Reynolds, pp. 214–15, and Herbert Parnes, pp. 158–59), 25 percent of workers who quit had new jobs lined up in advance. Reynolds found that the main obstacle to mobility without unemployment was not lack of information or time, but simply "anti-pirating" collusion by employers.

A considerable amount of search activity by unemployed workers appears to be an unproductive consequence of dissatisfaction and frustration rather than a rational quest for improvement. This was the conclusion of Reynolds' survey twenty-five years ago, p. 215, and it has been re-emphasized for the contemporary scene by Robert Hall, and by Peter Doeringer and Michael Piore for what they term the secondary labor force. Reynolds found that quitting a job to look for a new one while unemployed actually yielded a better job in only a third of the cases. Lining up a new job in advance was a more successful strategy: two-thirds of such changes turned out to be improvements. Today, according to the dual labor market hypothesis, the basic reason for frequent and long spells of unemployment in the secondary labor force is the shortage of good jobs.

In any event, the contention of some natural rate theorists is that employment beyond the natural rate takes time that would be better spent in search activity. Why do workers accept such employment? An answer to this question is a key element in a theory that generally presumes that actual behavior reveals true preferences. The answer given is that workers accept the additional employment only because they are victims of inflation illusion. One form of inflation illusion is over-estimation of the real wages of jobs they now hold, if they are employed, or of jobs they find, if they are unemployed and searching. If they did not under-estimate price inflation, employed workers would more often quit to search, and unemployed workers would search longer.

The force of this argument seems to me diluted by the fact that price inflation illusion affects equally both sides of the job seeker's equation. He over-estimates the real value of an immediate job, but he also over-estimates the real values of jobs he might wait for. It is in the spirit of this theorizing to assume that money interest rates respond to the same correct or incorrect inflationary expectations. As a first approximation, inflation illusion has no substitution effect on the margin between working and waiting.

It does have an income effect, causing workers to exaggerate their real wealth. In which direction the income effect would work is not transparent. Does greater wealth, or the illusion of greater

wealth, make people more choosy about jobs, more inclined to quit and to wait? Or less choosy, more inclined to stay in the job they have or to take the first one that comes along? I should have thought more selective rather than less. But natural rate theory must take the opposite view if it is to explain why under-estimation of price inflation bamboozles workers into holding or taking jobs that they do not really want.

Another form of alleged inflation illusion refers to wages rather than prices. Workers are myopic and do not perceive that wages elsewhere are, or soon will be, rising as fast as the money wage of the job they now hold or have just found. Consequently they under-estimate the advantages of quitting and searching. This explanation is convincing only to the extent that the payoff to search activity is determined by wage differentials. The payoff also depends on the probabilities of getting jobs at quoted wages, therefore on the balance between vacancies and job seekers. Workers know that perfectly well. Quit rates are an index of voluntary search activity. They do not diminish when unemployment is low and wage rates are rapidly rising. They increase, quite understandably. This fact contradicts the inflation illusion story, both versions. I conclude that it is not possible to regard fluctuations of unemployment on either side of the zero-inflation rate as mainly voluntary, albeit mistaken, extensions and contractions of search activity.

The new microeconomics of job search (see Edmund Phelps et al.), is nevertheless a valuable contribution to understanding of frictional unemployment. It provides reasons why some unemployment is voluntary, and why some unemployment is socially efficient.

Does the market produce the *optimal* amount of search unemployment? Is the natural rate optimal? I do not believe the new microeconomics has yet answered these questions.

An omniscient and beneficent economic dictator would not place every new job seeker immediately in any job at hand. Such a policy would create many mismatches, sacrificing efficiency in production or necessitating costly job-to-job shifts later on. The hypothetical planner would prefer to keep a queue of workers unemployed, so that he would have a larger choice of jobs to which to assign them. But he would not make the queue too long, because workers in the queue are not producing anything.

Of course he could shorten the queue of unemployed if he could dispose of more jobs and lengthen the queue of vacancies. With enough jobs of various kinds, he would never lack a vacancy for which any worker who happens to come along has comparative advantage. But because of limited capital stocks and interdependence among skills, jobs cannot be indefinitely multiplied without lowering their marginal productivity. Our wise and benevolent planner would not place people in jobs yielding less than the marginal value of leisure. Given this constraint on the number of jobs, he would always have to keep some workers waiting, and some jobs vacant. But he certainly would be inefficient if he had fewer jobs, filled and vacant, than this constraint. This is the common sense of Beveridge's rule—that vacancies should not be less than unemployment.

Is the natural rate a market solution of the hypothetical planner's operations research problem? According to search theory, an unemployed worker considers the probabilities that he can get a better job by searching longer and balances the expected discounted value of waiting against the loss of earnings. The employed worker makes a similar calculation when he considers quitting, also taking into ac-

count the once and for all costs of movement. These calculations are like those of the planner, but with an important difference. An individual does not internalize all the considerations the planner takes into account. The external effects are the familiar ones of congestion theory. A worker deciding to join a queue or to stay in one considers the probabilities of getting a job, but not the effects of his decision on the probabilities that others face. He lowers those probabilities for people in the queue he joins and raises them for persons waiting for the kind of job he vacates or turns down. Too many persons are unemployed waiting for good jobs, while less desirable ones go begging. However, external effects also occur in the decisions of employers whether to fill a vacancy with the applicant at hand or to wait for someone more qualified. It is not obvious, at least to me, whether the market is biased toward excessive or inadequate search. But it is doubtful that it produces the optimal amount.

Empirically the proposition that in the United States the zero-inflation rate of unemployment reflects voluntary and efficient job-seeking activity strains credulity. If there were a natural rate of unemployment in the United States, what would it be? It is hard to say because virtually all econometric Phillips curves allow for a whole menu of steady inflation rates. But estimates constrained to produce a vertical long-run Phillips curve suggest a natural rate between 5 and 6 percent of the labor force.[3]

So let us consider some of the features of an overall unemployment rate of 5 to 6 percent. First, about 40 percent of accessions in manufacturing are rehires rather than new hires. Temporarily laid off by their employers, these workers had been awaiting recall and were scarcely engaged in voluntary search activity. Their unemployment is as much a deadweight loss as the disguised unemployment of redundant workers on payrolls. This number declines to 25–30 percent when unemployment is 4 percent or below. Likewise, a 5–6 percent unemployment rate means that voluntary quits amount only to about a third of separations, layoffs to two-thirds. The proportions are reversed at low unemployment rates.

Second, the unemployment statistic is not an exhaustive count of those with time and incentive to search. An additional 3 percent of the labor force are involuntarily confined to part-time work, and another 3/4 of 1 percent are out of the labor force because they "could not find job" or "think no work available"—discouraged by market conditions rather than personal incapacities.

Third, with unemployment of 5–6 percent the number of reported vacancies is less than 1/2 of 1 percent. Vacancies appear to be understated relative to unemployment, but they rise to $1\frac{1}{2}$ percent when the unemployment rate is below 4 percent. At 5–6 percent unemployment, the economy is clearly capable of generating many more jobs with marginal productivity high enough so that people prefer them to leisure. The capital stock is no limitation, since 5–6 percent unemployment has been associated with more than 20 percent excess capacity. Moreover, when more jobs are created by expansion of demand, with or without inflation, labor force participation increases; this would hardly occur if the additional jobs were low in quality and productivity. As the parable of the central employment planner indicates, there will be excessive waiting for jobs if the roster of jobs and the menu of vacancies are suboptimal.

In summary, labor markets characterized by 5–6 percent unemployment do not display the symptoms one would ex-

[3] See Lucas and Rapping, pp. 257–305, in Phelps et al.

pect if the unemployment were voluntary search activity. Even if it were voluntary, search activity on such a large scale would surely be socially wasteful. The only reason anyone might regard so high an unemployment rate as an equilibrium and social optimum is that lower rates cause accelerating inflation. But this is almost tautological. The inferences of equilibrium and optimality would be more convincing if they were corroborated by direct evidence.

IV. Why is There Inflation without Aggregate Excess Demand?

Zero-inflation unemployment is not wholly voluntary, not optimal, I might even say not natural. In other words, the economy has an inflationary bias: When labor markets provide as many jobs as there are willing workers, there is inflation, perhaps accelerating inflation. Why?

The Phillips curve has been an empirical finding in search of a theory, like Pirandello characters in search of an author. One rationalization might be termed a theory of stochastic macro-equilibrium: stochastic, because random intersectoral shocks keep individual labor markets in diverse states of disequilibrium; macro-equilibrium, because the perpetual flux of particular markets produces fairly definite aggregate outcomes of unemployment and wages. Stimulated by Phillips's 1958 findings, Richard Lipsey proposed a model of this kind in 1960, and it has since been elaborated by Archibald, pp. 212–23 and Holt, pp. 53–123 and 224–56 in Phelps et. al., and others. I propose now to sketch a theory in the same spirit.

It is an essential feature of the theory that economy-wide relations among employment, wages, and prices are aggregations of diverse outcomes in heterogeneous markets. The myth of macroeconomics is that relations among aggregates are enlarged analogues of relations among corresponding variables for individual households, firms, industries, markets. The myth is a harmless and useful simplification in many contexts, but sometimes it misses the essence of the phenomenon.

Unemployment is, in this model as in Keynes reinterpreted, a disequilibrium phenomenon. Money wages do not adjust rapidly enough to clear all labor markets every day. Excess supplies in labor markets take the form of unemployment, and excess demands the form of unfilled vacancies. At any moment, markets vary widely in excess demand or supply, and the economy as a whole shows both vacancies and unemployment.

The overall balance of vacancies and unemployment is determined by aggregate demand, and is therefore in principle subject to control by overall monetary and fiscal policy. Higher aggregate demand means fewer excess supply markets and more excess demand markets, accordingly less unemployment and more vacancies.

In any particular labor market, the rate of increase of money wages is the sum of two components, an equilibrium component and a disequilibrium component. 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—a monotonic function, positive for positive excess demand, zero for zero excess demand, non-positive for excess supply. I begin with the disequilibrium component.

Of course the disequilibrium components are relevant only if disequilibria persist. Why aren't they eliminated by the very adjustments they set in motion? Workers will move from excess supply markets to excess demand markets, and from low wage to high wage markets. Unless they overshoot, these movements are equilibrating. The theory therefore

requires that new disequilibria are always arising. Aggregate demand may be stable, but beneath its stability is never-ending flux: new products, new processes, new tastes and fashions, new developments of land and natural resources, obsolescent industries and declining areas.

The overlap of vacancies and unemployment—say, the sum of the two for any given difference between them—is a measure of the heterogeneity or dispersion of individual markets. The amount of dispersion depends directly on the size of those shocks of demand and technology that keep markets in perpetual disequilibrium, and inversely on the responsive mobility of labor. The one increases, the other diminishes the frictional component of unemployment, that is, the number of unfilled vacancies coexisting with any given unemployment rate.

A central assumption of the theory is that the functions relating wage change to excess demand or supply are non-linear, specifically that unemployment retards money wages less than vacancies accelerate them. Nonlinearity in the response of wages to excess demand has several important implications. First, it helps to explain the characteristic observed curvature of the Phillips curve. Each successive increment of unemployment has less effect in reducing the rate of inflation. Linear wage response, on the other hand, would mean a linear Phillips relation.

Second, given the overall state of aggregate demand, economy-wide vacancies less unemployment, wage inflation will be greater the larger the variance among markets in excess demand and supply. As a number of recent empirical studies, have confirmed (see George Perry and Charles Schultze), dispersion is inflationary. Of course, the rate of wage inflation will depend not only on the overall dispersion of excess demands and supplies across markets but also on the particular markets where the excess supplies and demands happen to fall. An unlucky random drawing might put the excess demands in highly responsive markets and the excess supplies in especially unresponsive ones.

Third, the nonlinearity is an explanation of inflationary bias, in the following sense. Even when aggregate vacancies are at most equal to unemployment, the average disequilibrium component will be positive. Full employment in the sense of equality of vacancies and unemployment is not compatible with price stability. Zero inflation requires unemployment in excess of vacancies.

Criteria that coincide in full long-run equilibrium—zero inflation and zero aggregate excess demand—diverge in stochastic macro-equilibrium. Full long-run equilibrium in all markets would show no unemployment, no vacancies, no unanticipated inflation. But with unending sectoral flux, zero excess demand spells inflation and zero inflation spells net excess supply, unemployment in excess of vacancies. In these circumstances neither criterion can be justified simply because it is a property of full long-run equilibrium. Both criteria automatically allow for frictional unemployment incident to the required movements of workers between markets; the no-inflation criterion requires enough additional unemployment to wipe out inflationary bias.

I turn now to the equilibrium component, the rate of wage increase in a market with neither excess demand nor excess supply. It is reasonable to suppose that the equilibrium component depends on the trend of wages of comparable labor elsewhere. A "competitive wage," one that reflects relevant trends fully, is what employers will offer if they wish to maintain their share of the volume of employment. This will happen where the rate of growth of marginal revenue product—the com-

pound of productivity increase and price inflation—is the same as the trend in wages. But in some markets the equilibrium wage will be rising faster, and in others slower, than the economy-wide wage trend.

A "natural rate" result follows if actual wage increases feed fully into the equilibrium components of future wage increases. There will be acceleration whenever the non-linear disequilibrium effects are on average positive, and steady inflation, that is stochastically steady inflation, only at unemployment rates high enough to make the disequilibrium effects wash out. Phillips tradeoffs exist in the short run, and the time it takes for them to evaporate depends on the lengths of the lags with which today's actual wage gains become tomorrow's standards.

A rather minor modification may preserve Phillips tradeoffs in the long run. Suppose there is a floor on wage change in excess supply markets, independent of the amount of excess supply and of the past history of wages and prices. Suppose, for example, that wage change is never negative; it is either zero or what the response function says, whichever is algebraically larger. So long as there are markets where this floor is effective, there can be determinate rates of economy-wide wage inflation for various levels of aggregate demand. Markets at the floor do not increase their contributions to aggregate wage inflation when overall demand is raised. Nor is their contribution escalated to actual wage experience. But the frequency of such markets diminishes, it is true, both with overall demand and with inflation. The floor phenomenon can preserve a Phillips tradeoff within limits, but one that becomes ever more fragile and vanishes as greater demand pressure removes markets from contact with the zero floor. The model implies a long-run Phillips curve that is very flat for high unemployment and becomes vertical at a critically low rate of unemployment.

These implications seem plausible and even realistic. It will be objected, however, that any permanent floor independent of general wage and price history and expectation must indicate money illusion. The answer is that the floor need not be permanent in any single market. It could give way to wage reduction when enough unemployment has persisted long enough. But with stochastic intersectoral shifts of demand, markets are always exchanging roles, and there can always be some markets, not always the same ones, at the floor.

This model avoids the empirically questionable implication of the usual natural rate hypothesis that unemployment rates only slightly higher than the critical rate will trigger ever-accelerating deflation. Phillips curves seem to be pretty flat at high rates of unemployment. During the great contraction of 1930–33, wage rates were slow to give way even in the face of massive unemployment and substantial deflation in consumer prices. Finally in 1932 and 1933 money wage rates fell more sharply, in response to prolonged unemployment, layoffs, shutdowns, and to threats and fears of more of the same.

I have gone through this example to make the point that irrationality, in the sense that meaningless differences in money values *permanently* affect individual behavior, is not logically necessary for the existence of a long-run Phillips tradeoff. In full long-run equilibrium in all markets, employment and unemployment would be independent of the levels and rates of change of money wage rates and prices. But this is not an equilibrium that the system ever approaches. The economy is in perpetual sectoral disequilibrium even when it has settled into a stochastic macro-equilibrium.

I suppose that one might maintain that asymmetry in wage adjustment and tem-

porary resistance to money wage decline reflect money illusion in some sense. Such an assertion would have to be based on an extension of the domain of well-defined rational behavior to cover responses to change, adjustment speeds, costs of information, costs of organizing and operating markets, and a host of other problems in dynamic theory. These theoretical extensions are in their infancy, although much work of interest and promise is being done. Meanwhile, I doubt that significant restrictions on disequilibrium adjustment mechanisms can be deduced from first principles.

Why are the wage and salary rates of employed workers so insensitive to the availability of potential replacements? One reason is that the employer makes some explicit or implicit commitments in putting a worker on the payroll in the first place. The employee expects that his wages and terms of employment will steadily improve, certainly never retrogress. He expects that the employer will pay him the rate prevailing for persons of comparable skill, occupation, experience, and seniority. He expects such commitments in return for his own investments in the job; arrangements for residence, transportation, and personal life involve set-up costs which will be wasted if the job turns sour. The market for labor services is not like a market for fresh produce where the entire current supply is auctioned daily. It is more like a rental housing market, in which most existing tenancies are the continuations of long-term relationships governed by contracts or less formal understandings.

Employers and workers alike regard the wages of comparable labor elsewhere as a standard, but what determines those reference wages? There is not even an auction where workers and employers unbound by existing relationships and commitments meet and determine a market-clearing wage. If such markets existed, they would provide competitively determined guides for negotiated and administered wages, just as stock exchange prices are reference points for stock transactions elsewhere. In labor markets the reverse is closer to the truth. Wage rates for existing employees set the standards for new employees, too.

The equilibrium components of wage increases, it has been argued, depend on past wage increases throughout the economy. In those theoretical and econometric models of inflation where labor markets are aggregated into a single market, this relationship is expressed as an autoregressive equation of fixed structure: current wage increase depends on past wage increases. The same description applies when past wage increases enter indirectly, mediated by price inflation and productivity change. The process of mutual interdependence of market wages is a good deal more complex and less mechanical than these aggregated models suggest.

Reference standards for wages differ from market to market. The equilibrium wage increase in each market will be some function of past wages in all markets, and perhaps of past prices too. But the function need not be the same in every market. Wages of workers contiguous in geography industry, and skill will be heavily weighted. Imagine a wage pattern matrix of coefficients describing the dependence of the percentage equilibrium wage increase in each market on the past increases in all other markets. The coefficients in each row are non-negative and sum to one, but their distribution across markets and time lags will differ from row to row.

Consider the properties of such a system in the absence of disequilibrium inputs. First, the system has the "natural rate" property that its steady state is indeterminate. Any rate of wage increase that has been occurring in all markets for a long enough time will continue. Second, from irregular initial conditions the system will

move toward one of these steady states, but which one depends on the specifics of the wage pattern matrix and the initial conditions. Contrary to some pessimistic warnings, there is no arithmetic compulsion that makes the whole system gravitate in the direction of its most inflationary sectors. The ultimate steady state inflation will be at most that of the market with the highest initial inflation rate, and at least that of the market with the lowest initial inflation rate. It need not be equal to the average inflation rate at the beginning, but may be either greater or smaller. Third, the adjustment paths are likely to contain cyclical components, damped or at most of constant amplitude, and during adjustments both individual and average wage movements may diverge substantially in both directions from their ultimate steady state value. Fourth, since wage decisions and negotiations occur infrequently, relative wage adjustments involve a lot of catching up and leap-frogging, and probably take a long time. I have sketched the formal properties of a disaggregated wage pattern system of this kind simply to stress again the vast simplification of the one-market myth.

A system in which only relative magnitudes matter has only a neutral equilibrium, from which it can be permanently displaced by random shocks. Even when a market is in equilibrium, it may outdo the recent wage increases in related markets. A shock of this kind, even though it is not repeated, raises permanently the steady state inflation rate. This is true cost-push —inflation generated neither by previous inflation nor by current excess demand. Shocks, of course, may be negative as well as positive. For example, upward pushes arising from adjustments in relative wage *levels* will be reversed when those adjustments are completed.

To the extent that one man's reference wages are another man's wages, there is something arbitrary and conventional, indeterminate and unstable, in the process of wage setting. In the same current market circumstances, the reference pattern might be 8 percent per year or 3 percent per year or zero, depending on the historical prelude. Market conditions, unemployment and vacancies and their distributions, shape history and alter reference patterns. But accidental circumstances affecting stragetic wage settlements also cast a long shadow.

Price inflation, as previously observed, is a neutral method of making arbitrary money wage paths conform to the realities of productivity growth, neutral in preserving the structure of relative wages. If expansion of aggregate demand brings both more inflation and more employment, there need be no mystery why unemployed workers accept the new jobs, or why employed workers do not vacate theirs. They need not be victims of ignorance or inflation illusion. They genuinely want more work at feasible real wages, and they also want to maintain the relative status they regard as proper and just.

Guideposts could be in principle the functional equivalent of inflation, a neutral method of reconciling wage and productivity paths. The trick is to find a formula for mutual deescalation which does not offend conceptions of relative equity. No one has devised a way of controlling average wage rates without intervening in the competitive struggle over relative wages. Inflation lets this struggle proceed and blindly, impartially, impersonally, and nonpolitically scales down all its outcomes. There are worse methods of resolving group rivalries and social conflict.

V. The Role of Monopoly Power

Probably the most popular explanation of the inflationary bias of the economy is concentration of economic power in large corporations and unions. These powerful

monopolies and oligopolies, it is argued, are immune from competition in setting wages and prices. The unions raise wages above competitive rates, with little regard for the unemployed and under-employed workers knocking at the gates. Perhaps the unions are seeking a bigger share of the revenues of the monopolies and oligopolies with whom they bargain. But they don't really succeed in that objective, because the corporations simply pass the increased labor costs, along with mark-ups, on to their helpless customers. The remedy, it is argued, is either atomization of big business and big labor or strict public control of their prices and wages.

So simple a diagnosis is vitiated by confusion between levels and rates of change. Monopoly power is no doubt responsible for the relatively high prices and wages of some sectors. But can the exercise of monopoly power generate ever-rising price and wages? Monopolists have no reason to hold reserves of unexploited power. But if they did, or if events awarded them new power, their exploitation of it would raise their real prices and wages only temporarily.

Particular episodes of inflation may be associated with accretions of monopoly power, or with changes in the strategies and preferences of those who possess it. Among the reasons that wages and prices rose in the face of mass unemployment after 1933 were *NRA* codes and other early New Deal measures to suppress competition, and the growth of trade union membership and power under the protection of new federal legislation. Recently we have witnessed substantial gains in the powers of organized public employees. Unions elsewhere may not have gained power, but some of them apparently have changed their objectives in favor of wages at the expense of employment.

One reason for the popularity of the monopoly power diagnosis of inflation is the identification of administered prices and wages with concentrations of economic power. When price and wage increases are the outcomes of visible negotiations and decisions, it seems obvious that identifiable firms and unions have the power to affect the course of inflation. But the fact that monopolies, oligopolies, and large unions have discretion does not mean it is invariably to their advantage to use it to raise prices and wages. Nor are administered prices and wages found only in high concentration sectors. Very few prices and wages in a modern economy, even in the more competitive sectors, are determined in Walrasian auction markets.

No doubt there has been a secular increase in the prevalence of administered wages and prices, connected with the relative decline of agriculture and other sectors of self-employment. This development probably has contributed to the inflationary bias of the economy, by enlarging the number of labor markets where the response of money wages to excess supply is slower than their response to excess demand. The decline of agriculture as a sector of flexible prices and wages and as an elastic source of industrial labor is probably an important reason why the Phillips trade off problem is worse now than in the 1920's. Sluggishness of response to excess supply is a feature of administered prices, whatever the market structure, but it may be accentuated by concentration of power per se. For example, powerful unions, not actually forced by competition to moderate their wage demands, may for reasons of internal politics be slow to respond to unemployment in their ranks.

VI. Some Reflections on Policy

If the makers of macro-economic policy could be sure that the zero-inflation rate of unemployment is natural, voluntary, and optimal, their lives would be easy.

Friedman told us that all macro-economic policy needs to do, all it should try to do, is to make nominal national income grow steadily at the natural rate of growth of aggregate supply. This would sooner or later result in price stability. Steady price deflation would be even better, he said, because it would eliminate the socially wasteful incentive to economize money holdings. In either case, unemployment will converge to its natural rate, and wages and prices will settle into steady trends. Under this policy, whatever unemployment the market produces is the correct result. No tradeoff, no choice, no agonizing decisions.

I have argued this evening that a substantial amount of the unemployment compatible with zero inflation is involuntary and nonoptimal. This is, in my opinion, true whether or not the inflations associated with lower rates of unemployment are steady or ever-accelerating. Neither macro-economic policy makers, nor the elected officials and electorates to whom they are responsible, can avoid weighing the costs of unemployment against those of inflation. As Phelps has pointed out, this social choice has an intertemporal dimension. The social costs of involutionary unemployment are mostly obvious and immediate. The social costs of inflation come later.

What are they? Economists' answers have been remarkably vague, even though the prestige of the profession has reinforced the popular view that inflation leads ultimately to catastrophe. Here indeed is a case where abstract economic theory has a powerful hold on public opinion and policy. The prediction that at low unemployment rates inflation will accelerate toward ultimate disaster is a theoretical deduction with little empirical support. In fact the weight of econometric evidence has been against acceleration, let alone disaster. Yet the deduction has been convincing enough to persuade this country to give up billions of dollars of annual output and to impose sweeping legal controls on prices and wages. Seldom has a society made such large immediate and tangible sacrifices to avert an ill defined, uncertain, eventual evil.

According to economic theory, the ultimate social cost of anticipated inflation is the wasteful use of resources to economize holdings of currency and other noninterest-bearing means of payment. I suspect that intelligent laymen would be utterly astounded if they realized that *this* is the great evil economists are talking about. They have imagined a much more devastating cataclysm, with Vesuvius vengefully punishing the sinners below. Extra trips between savings banks and commercial banks? What an anti-climax!

With means of payment—currency plus demand deposits—equal currently to 20 percent of *GNP*, an extra percentage point of anticipated inflation embodied in nominal interest rates produces in principle a social cost of 2/10 of 1 percent of *GNP* per year. This is an outside estimate. An unknown, but substantial, share of the stock of money belongs to holders who are not trying to economize cash balances and are not near any margin where they would be induced to spend resources for this purpose. These include hoarders of large denomination currency, about one-third of the total currency in public hands, for reasons of privacy, tax evasion, or illegal activity. They include tradesmen and consumers whose working balances turn over too rapidly or are too small to justify any effort to invest them in interest-bearing assets. They include corporations who, once they have been induced to undertake the fixed costs of a sharp-pencil money management department, are already minimizing their cash holdings. They include businessmen who are in fact being paid interest on demand deposits,

although it takes the form of preferential access to credit and other bank services. But, in case anyone still regards the waste of resources in unnecessary transactions between money and interest-bearing financial assets as one of the major economic problems of the day, there is a simple and straightforward remedy, the payment of interest on demand deposits and possibly, with ingenuity, on currency too.

The ultimate disaster of inflation would be the breakdown of the monetary payments system, necessitating a currency reform. Such episodes have almost invariably resulted from real economic catastrophes—wars, defeats, revolutions, reparations—not from the mechanisms of wage-price push with which we are concerned. Acceleration is a scare word, conveying the image of a rush into hyper-inflation as relentlessly deterministic and monotonic as the motion of falling bodies. Realistic attention to the disaggregated and stochastic nature of wage and price movements suggests that they will show diverse and irregular fluctuations around trends that are difficult to discern and extrapolate. The central trends, history suggests, can accelerate for a long, long time without generating hyper-inflations destructive of the payments mechanism.

Unanticipated inflation, it is contended, leads to mistaken estimates of relative prices and consequently to misallocations of resources. An example we have already discussed is the alleged misallocation of time by workers who over-estimate their real wages. The same error would lead to a general over-supply by sellers who contract for future deliveries without taking correct account of the increasing prices of the things they must buy in order to fulfill the contract. Unanticipated deflation would cause similar miscalculations and misallocations. Indeed, people can make these same mistakes about relative prices even when the price level is stable. The mistakes are more likely, or the more costly to avoid, the greater the inflationary trend. There are costs in setting and announcing new prices. In an inflationary environment price changes must be made more frequently—a new catalog twice a year instead of one, or some formula for automatic escalation of announced prices. Otherwise, with the interval between announcements unchanged, the average misalignment of relative prices will be larger the faster the inflation. The same problem would arise with rapid deflation.

Unanticipated inflation and deflation—and unanticipated changes in relative prices—are also sources of transfers of wealth. I will not review here the rich and growing empirical literature on this subject. Facile generalizations about the progressivity or equity of inflationary transfers are hazardous; certainly inflation does not merit the cliché that it is "the cruelest tax." Let us not forget that unemployment has distributional effects as well as deadweight losses.

Some moralists take the view that the government has promised to maintain the purchasing power of its currency, but this promise is their inference rather than any pledge written on dollar bills or in the Constitution. Some believe so strongly in this implicit contract that they are willing to suspend actual contracts in the name of anti-inflation.

I have long contended that the government should make low-interest bonds of guaranteed purchasing power available for savers and pension funds who wish to avoid the risks of unforeseen inflation. The common objection to escalated bonds is that they would diminish the built-in stability of the system. The stability in question refers to the effects on aggregate real demand, *ceteris paribus*, of a change in the price level. The Pigou effect tells us that government bondholders whose wealth is diminished by inflation will spend less. This brake on old-fashioned gap

inflation will be thrown away if the bonds are escalated. These considerations are only remotely related to the mechanisms of wage and price inflation we have been discussing. In the 1970's we know that the government can, if it wishes, control aggregate demand—at any rate, its ability to do so is only trivially affected by the presence or absence of Pigou effects on part of the government debt.

In considering the intertemporal tradeoff, we have no license to assume that the natural rate of unemployment is independent of the history of actual unemployment. Students of human capital have been arguing convincingly that earning capacity, indeed transferable earning capacity, depends on experience as well as formal education. Labor markets soggy enough to maintain price stability may increase the number of would-be workers who lack the experience to fit them for jobs that become vacant.

Macro-economic policies, monetary and fiscal, are incapable of realizing society's unemployment and inflation goals simultaneously. This dismal fact has long stimulated a search for third instruments to do the job: guideposts and incomes policies, on the one hand, labor market and manpower policies, on the other. Ten to fifteen years ago great hopes were held for both. The Commission on Money and Credit in 1961, pp. 39–40, hailed manpower policies as the new instrument that would overcome the unemployment-inflation dilemma. Such advice was taken seriously in Washington, and an unprecedented spurt in manpower programs took place in the 1960's. The Council of Economic Advisers set forth wage and price guideposts in 1961–62 in the hope of "talking down" the Phillips curve (pp. 185–90). It is discouraging to find that these efforts did not keep the problem of inflationary bias from becoming worse than ever.

So it is not with great confidence or optimism that one suggests measures to mitigate the tradeoff. But some proposals follow naturally from the analysis, and some are desirable in themselves anyway.

First, guideposts do not wholly deserve the scorn that "toothless jawboning" often attracts. There is an arbitrary, imitative component in wage settlements, and maybe it can be influenced by national standards.

Second, it is important to create jobs for those unemployed and discouraged workers who have extremely low probability of meeting normal job specifications. Their unemployment does little to discipline wage increases, but reinforces their deprivation of human capital and their other disadvantages in job markets. The National Commission on Technology, Automation and Economic Progress pointed out in 1966 the need for public service jobs tailored to disadvantaged workers. They should not be "last resort" or make-work jobs, but regular permanent jobs capable of conveying useful experience and inducing reliable work habits. Assuming that the additional services produced by the employing institutions are of social utility, it may well be preferable to employ disadvantaged workers directly rather than to pump up aggregate demand until they reach the head of the queue.

Third, a number of measures could be taken to make markets more responsive to excess supplies. This is the kernel of truth in the market-power explanation of inflationary bias. In many cases, government regulations themselves support prices and wages against competition. Agricultural prices and construction wages are well-known examples. Some trade unions follow wage policies that take little or no account of the interests of less senior members and of potential members. Since unions operate with federal sanction and protection, perhaps some means can be found to insure that their memberships are open and that their policies are responsive to the unemployed as well as the employed.

As for macro-economic policy, I have

argued that it should aim for unemployment lower than the zero-inflation rate. How much lower? Low enough to equate unemployment and vacancies? We cannot say. In the nature of the case there is no simple formula—conceptual, much less statistical—for full employment. Society cannot escape very difficult political and intertemporal choices. We economists can illuminate these choices as we learn more about labor markets, mobility, and search, and more about the social and distributive costs of both unemployment and inflation. Thirty-five years after Keynes, welfare macroeconomics is still a relevant and challenging subject. I dare to believe it has a bright future.

REFERENCES

W. H. Beveridge, *Full Employment in a Free Society,* New York 1945.

P. Doeringer and M. Piore, *Internal Labor Markets and Manpower Analysis,* Lexington, Mass. 1971.

M. Friedman, "The Role of Monetary Policy," *Amer. Econ Rev.,* Mar. 1968, *58,* 1–17.

R. Hall, "Why is the Unemployment Rate so High at Full Employment?," *Brookings Papers on Economic Activity,* 3, 1970, 369–402.

J. M. Keynes, *The General Theory of Employment, Interest, and Money,* New York 1936.

A. Leijonhufvud, *On Keynesian Economics and the Economics of Keynes.* New York 1968.

R. G. Lipsey, "The Relation Between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1862–1957: A Further Analysis," *Economica,* Feb. 1960, *27,* 1–31.

H. S. Parnes, *Research on Labor Mobility,* Social Science Research Council, Bull. 65, New York 1954.

G. L. Perry, "Changing Labor Markets and Inflation," *Brookings Papers on Economic Activity,* 3, 1970, 411–41.

E. S. Phelps et al., *Micro-economic Foundations* Inflation and Optimal Unemployment Over Time," *Economica,* Aug. 1967, *34,* 254–81.

E. S. Phelps et al., *Micro-economic Foundations of Employment and Inflation Theory,* New York 1970.

A. W. Phillips, "The Relation Between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861–1957," *Economica,* Nov. 1958, *25,* 283–99.

L. G. Reynolds, *The Structure of Labor Markets,* New York 1951.

C. L. Schultze, "Has the Phillips Curve Shifted? Some Additional Evidence," *Brookings Papers on Economic Activity,* 2, 1971, 452–67.

J. Tobin, "A Note on the Money Wage Problem," *Quart. J. Econ.,* May 1941, *55,* 508–16.

Commission on Money and Credit, *Money and Credit: Their Influence on Jobs, Prices, and Growth,* Englewood Cliffs 1961

Economic Report of the President 1962, Washington 1962.

U.S. National Commission on Technology, Automation, and Economic Progress, *Technology and the American Economy,* Washington 1966.