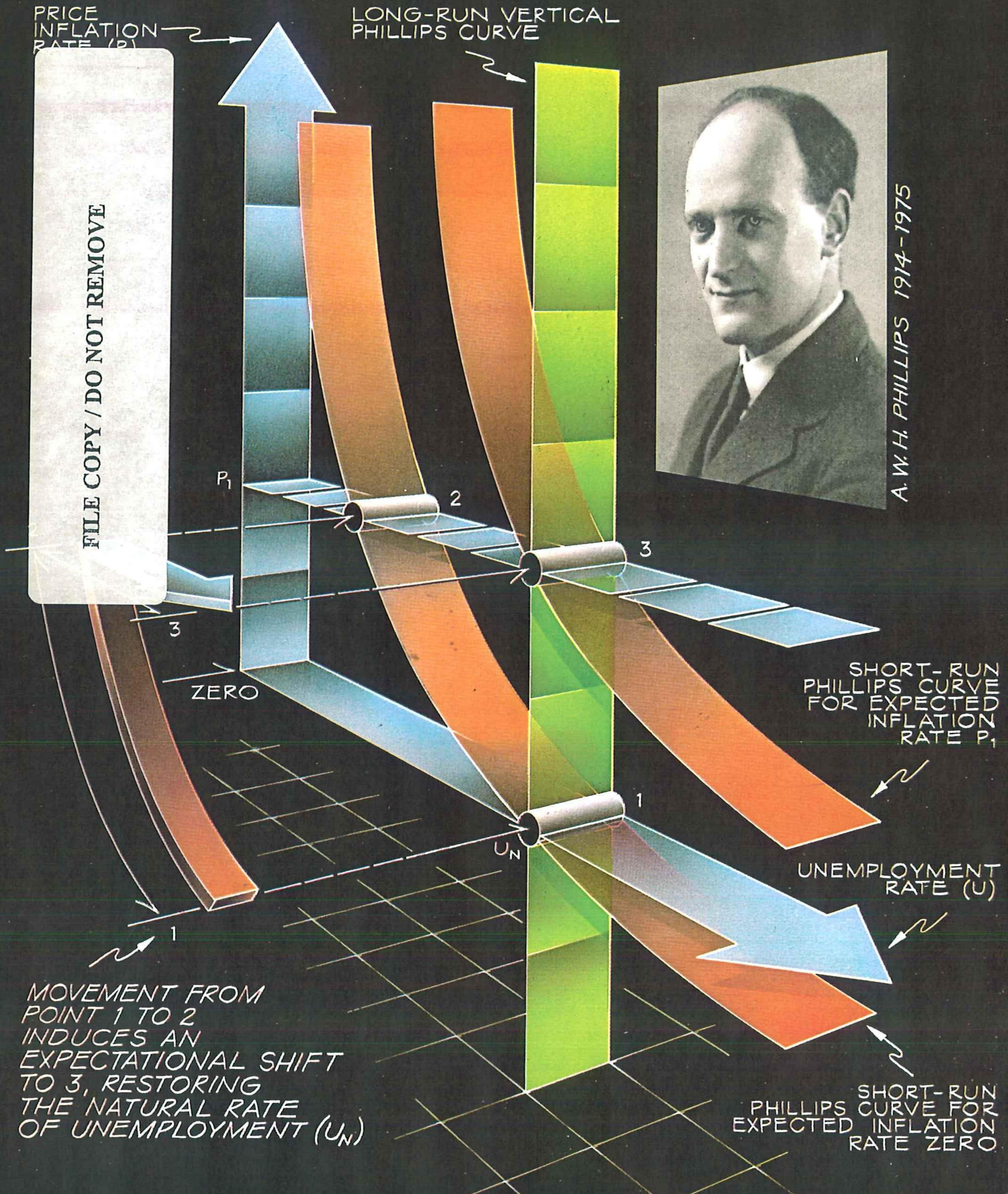
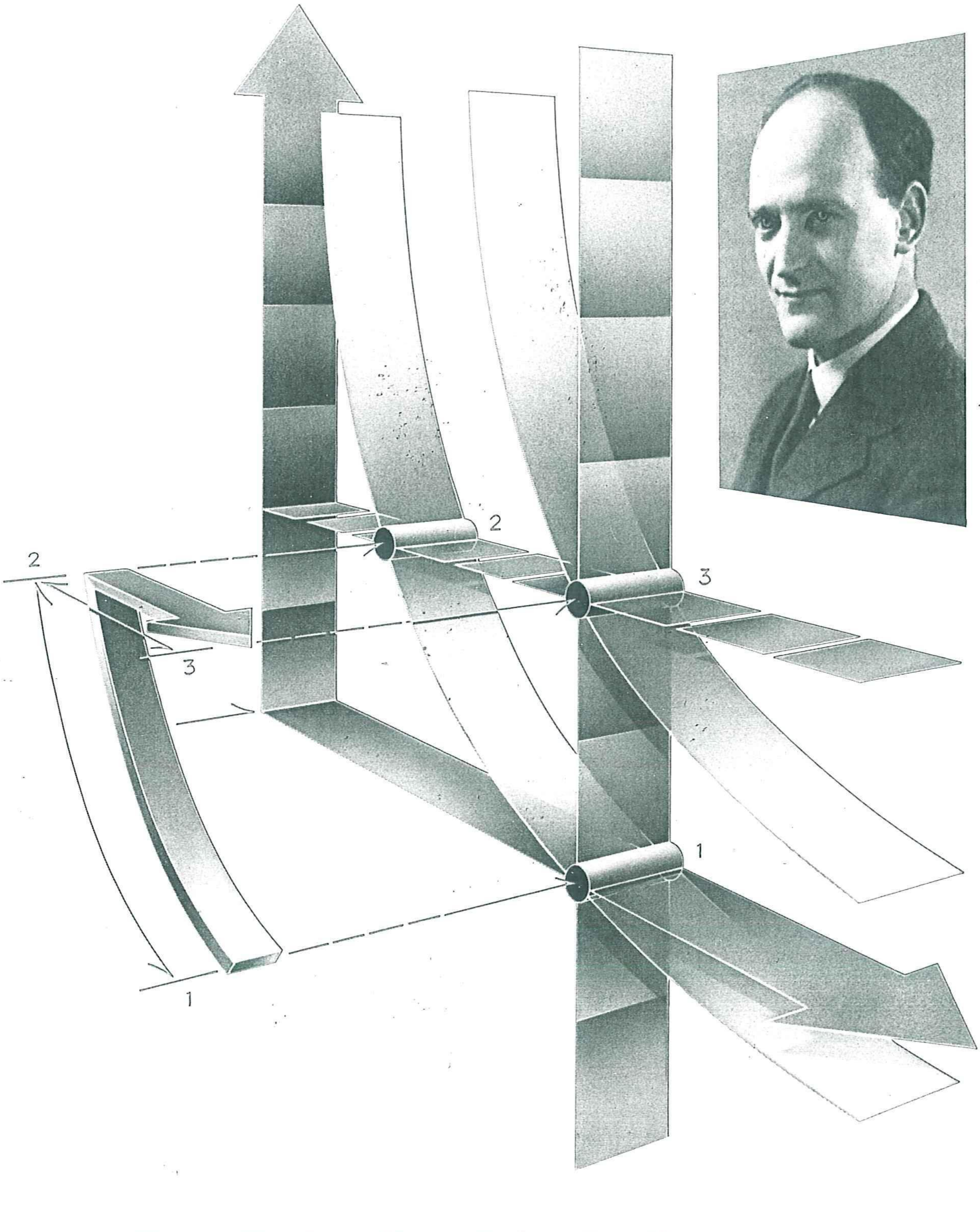


From Trade-offs to Policy Ineffectiveness: A History of the Phillips Curve

Thomas M. Humphrey



From Trade-offs to Policy Ineffectiveness:
A History of the Phillips Curve

Thomas M. Humphrey

October 1986

Contents

Preface

The famous Phillips curve trade-off relationship between inflation and unemployment, whose doctrinal history these essays chronicle, amply illustrates the workings of two well-known "laws" in economics. It illustrates, first, statistician Stephen M. Stigler's Law of Eponymy, according to which no scientific discovery is named for its original discoverer. For, as shown in the initial essay, Phillips was far from the first to describe the relationship bearing his name. On the contrary, in the two hundred years before him at least ten economists, including such celebrated names as David Hume, Henry Thornton, John Stuart Mill, Irving Fisher, and Jan Tinbergen, presented versions of the curve.

The Phillips curve also exemplifies the workings of Goodheart's Law, which, as formulated by the British economist C.A.E. Goodheart, states that any observed statistical regularity will collapse once the authorities try to exploit it for policy purposes. Such a collapse was precisely the fate of the Phillips curve in the 1960s and 1970s. For no sooner had Professor Phillips isolated what he thought was a stable hundred-year empirical relationship between inflation and unemployment than it started breaking down when policymakers sought to exploit it. The resulting failure of the Phillips curve to hold still while the authorities attempted to move the economy along it, trading off higher inflation for lower unemployment until the best attainable combination had been reached, led to disenchantment with and consequently reformulation of the Phillips curve idea. Circumstances forced economists to take account of inflationary expectations viewed as the chief factor causing shifts in the curve.

Of these reformulated expectations-augmented Phillips curves, two in particular have dominated recent thought. First came the Friedman-Phelps error-learning or adaptive-expectations version (according to which trade-offs last only so long as expectations are adjusting to actualities) followed closely by the Lucas-Sargent market-clearing or new classical version (in which expectations adjust instantaneously). This latter version, embodying as it does the natural rate and rational expectations hypotheses, teaches that the Phillips relationship is a purely adventitious phenomenon generated by unforeseeable random shocks and as such cannot be exploited by systematic macroeconomic policies. The fact that John Stuart Mill already had reached this same conclusion as early as 1833 suggests that there have long been at least three competing views of the Phillips curve. One sees it as a stable permanently exploitable trade-off; another as a temporary trade-off that vanishes once expectations catch up with reality; and the third as a purely statistical relationship that cannot be exploited for policy purposes in either the short-run or the long.

Thomas M. Humphrey

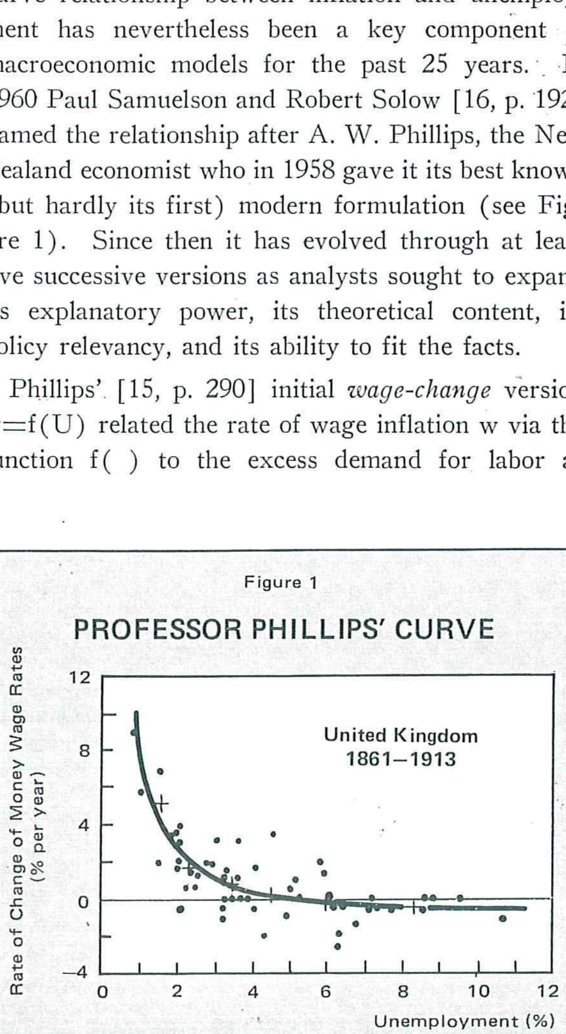
THE EARLY HISTORY OF THE PHILLIPS CURVE

Although critics may dismiss it as a mere empirical correlation masquerading as a tradeoff, the Phillips curve relationship between inflation and unemployment has nevertheless been a key component of macroeconomic models for the past 25 years. In 1960 Paul Samuelson and Robert Solow [16, p. 192] named the relationship after A. W. Phillips, the New Zealand economist who in 1958 gave it its best known (but hardly its first) modern formulation (see Figure 1). Since then it has evolved through at least five successive versions as analysts sought to expand its explanatory power, its theoretical content, its policy relevancy, and its ability to fit the facts.

Phillips' [15, p. 290] initial *wage-change* version $w=f(U)$ related the rate of wage inflation w via the function f( ) to the excess demand for labor as measured by U, the deviation of unemployment from its equilibrium or labor-market clearing rate. Transformed through the assumed markup of prices over wages into the *price-change* equation $p=f(U)$, where p is the rate of price inflation, it was widely interpreted as a stable enduring tradeoff or menu of inflation-unemployment combinations from which the authorities could choose. In its *shift-adjusted* form $p=f(U)+Z$, it incorporated a vector of variables Z, including past price changes, trade union effects, unemployment dispersion, demographic factors and the like, to account for observed shifts in the inflation-unemployment tradeoff or menu of policy choices. In its *expectations-augmented* form $p-p^e=f(U)$, where p^e is the expected rate of inflation, it asserted (1) that the tradeoff is between unemployment and unexpected inflation, (2) that the tradeoff vanishes when expectations are realized, and (3) that unemployment returns to its natural equilibrium rate at this point. Provided expectations adjust to actual inflation with a lag, it also implied the accelerationist notion that unemployment can be pegged permanently below its natural rate only if inflation is continually accelerated so as to always stay a step ahead of expectations. That is, while denying a permanent tradeoff between unemployment and the rate of inflation, it implied that there may be a permanent tradeoff between unemployment and the rate of acceleration of inflation.

Figure 1

PROFESSOR PHILLIPS' CURVE



Phillips obtained his curve by fitting the equation $w = f(U) = -a + bU^{-c}$ to the scatter of annual observations on wage inflation rates (w) and unemployment rates (U) for the United Kingdom, 1861-1913. He found also that observations for the years 1948-1957 lay close to his fitted curve, indicating its apparent long-run empirical stability.

Source: Phillips [15, p. 285]

The preceding versions reflect a non-market-clearing view of the world, expressing as they do the disequilibrium response of wages and prices to a mismatching of demand and supply in the labor market. By contrast, the alternative *New Classical* or *market clearing* version $U=g(p-p^e)$ assumes that the labor market is always in equilibrium and that deviations of unemployment from its natural rate stem solely from inflation misperceptions and vanish when those misperceptions end. When combined with the assumption of rational expectations (according to which actual inflation differs from expected inflation only by a random forecast error) this version says that tradeoffs are solely the result of unpredictable random shocks and cannot be exploited by systematic (predictable) policies.

The foregoing interpretations are well known. Not so well known, however, is the origin and early history of the inflation-unemployment relationship. For the most part, textbooks typically trace the idea to Phillips' famous 1958 *Economica* article without saying anything about what went before. They correctly describe the five versions of the Phillips curve outlined above. But they fail to note that at least three of those versions (including the version presented by Phillips himself) had already been spelled out long before Phillips. The result is to neglect at least ten predecessors whose names deserve to be associated with the Phillips curve. In an effort to redress this oversight and to set the record straight, the paragraphs below document what Phillips' predecessors had to say about the inflation-unemployment relationship.

John Law (1671-1729)

It is probably unrealistic to expect to find a Phillips curve in the writings of John Law, the famous eighteenth century banker and finance minister whose schemes to promote economic development via the creation of a paper currency secured by land ended with the collapse of the Mississippi Bubble in 1720. To be sure, he believed that money stimulates real activity. But he also believed that it does so at constant or even decreasing prices owing to the availability of idle resources and scale economies in production. As a result, there is either no Phillips curve inflation-unemployment relation in his analysis or it works in the wrong direction—falling unemployment being associated with falling, not rising, prices.

David Hume (1711-1776)

The prototypal Phillips curve analysis is to be found in the writings of the eighteenth century Scottish philosopher-economist David Hume. As early as 1752, he presented the essentials of a Phillips curve relationship of the form $U=g(dP/dt)$, where U is the deviation of unemployment from its natural (equilibrium) rate and dP/dt is the change in the price level with respect to time. This relationship derived straight from his assumption that unemployment disturbances stem from price perception errors (the difference between actual and perceived prices) and that such errors persist only when prices are changing. Expressed symbolically, he assumed that

$$U = h(P-P^E) \text{ and}$$

$$P-P^E = k\, dP/dt$$

where P and P^E denote actual and perceived prices and k is a coefficient relating price perception errors to price level changes. Substitution of the latter equation into the former yields Hume's version of the Phillips curve $U=g(dP/dt)$ mentioned above. That version embodied his hypothesis that one must continually raise prices to peg unemployment at arbitrarily low levels since only by doing so can one produce the price perception errors that sustain the tradeoff. In short, Hume's explanation stresses the employment effects of unperceived monetary-induced price changes. He [8, pp. 37-40] says:

> though the high price of commodities be a necessary consequence of the encrease of gold and silver, yet it follows not immediately upon that encrease; but some time is required before the money circulates through the whole state and makes its effect be felt on all ranks of people. At first, no alteration is perceived; by degrees the price rises, first of one commodity, then of another; till the whole at last reaches a just proportion with the new quantity of specie . . . . In my opinion, it is only in this interval or intermediate situation, between the acquisition of money and rise of prices, that the encreasing quantity of gold and silver is favourable to industry . . . . From the whole of this reasoning we may conclude, that it is of no manner of consequence, with regard to the domestic happiness of a state, whether money be in a greater or less quantity. The good policy of the magistrate consists only in keeping it, if possible, still encreasing; because, by that means, he keeps alive a spirit of industry in the nation . . . . There is always an interval before matters be adjusted to their new situation; and this interval is as pernicious to industry, when gold and silver are diminishing, as it is advantageous when these metals are encreasing.

Three points stand out in Hume's analysis [10]. First, the tradeoff is between unemployment and unperceived changes in money and prices; it vanishes once perceptions fully adjust to reality. Second, price perceptions, though slow to adjust, eventually catch up to one-time changes in the level of money and prices. It follows that such changes can at best generate temporary but not permanent tradeoffs.

Third, the only way the tradeoff can be sustained is to generate a continual succession of changes in money and prices. Hume here makes the distinctly non-rational-expectations argument that such changes will, because of the lag in the adjustment of price perceptions, keep prices forever marching a step ahead of perceptions, perpetually frustrating the latter's attempts to catch up. In this way, he claims, the gap between actual and perceived prices will be maintained thus permanently lowering unemployment. Hume notes that this process works symmetrically for price deflation—such deflation, if prolonged, producing an enduring rise in unemployment. It follows at once that a permanent tradeoff U=g(dP/dt) exists between unemployment and the rate of change of money and prices. One must therefore agree with Charles R. Nelson's [14, p. 2] recent judgment that

> Hume was clearly of the opinion that the *level* of activity would be raised *permanently* by a steady increase in the quantity of money, prices, and wages. Hume was therefore a believer in a stable, long-run Phillips curve.

Henry Thornton (1760-1815)

Like Hume, Henry Thornton also described a Phillips curve of the form U=g(dP/dt), where the variables are as defined above [10]. In his classic *An Enquiry into the Nature and Effects of the Paper Credit of Great Britain* (1802) he [19, p. 237] says that a monetary expansion stimulates employment by raising prices:

> . . . additional industry will be one effect of an extraordinary emission of paper, a rise in the cost [i.e., price] of articles will be another. Probably no small part of that industry which is excited by new paper is produced through the enhancement of the cost of commodities.

This same tradeoff, he [19, p. 238] notes, also holds in reverse as monetary and price deflation bring painful rises in unemployment.

> If we assume the augmented paper to be brought back to its ordinary quantity, we must suppose industry to languish for a time through the ill success [of] mercantile transactions.

In his discussion of the Phillips curve, Thornton was careful to distinguish between alternative *levels* of money and prices and continuous *changes* of those variables. Only the latter, he said, can affect real activity and sustain the tradeoff. This is epitomized in his [19, p. 256] remark that "it is the progressive augmentation of bank paper, and not the magnitude of its existing amount, which gives the relief." In other words, money and prices stimulate activity only when they are continually increasing. For, says Thornton [19, p. 238], "While paper is encreasing, and articles continue rising, mercantile speculations appear more than ordinarily profitable." But "as soon . . . as the circulating medium ceases to encrease, the extra profit is at an end," and the stimulus vanishes. Thus a one-time rise in the money stock and level of prices cannot sustain the tradeoff. Instead, a continuous increase or "progressive augmentation" is required. The tradeoff is between output and the rate of change of prices.

As for the tradeoff's source, Thornton attributed it chiefly to a tendency for money wages to consistently lag behind prices. He explicitly stated (1) that inflation stimulates activity, (2) that it does so by reducing real wages and raising real profits, (3) that this output-enhancing redistribution occurs because money wages lag behind prices, and (4) that this wage lag persists as long as inflation is sustained. Like Hume, he did not explain why the lag would persist nor why wages would not eventually catch up with prices once inflationary expectations had fully adjusted to actual inflation. His analysis is largely silent about inflation anticipations; he did not incorporate them into his Phillips curve.

Finally, he disagreed with Hume over the desirability of exploiting the Phillips curve for policy purposes. Hume clearly believed that the policy authorities in the closed world economy should exploit the curve, using monetary gold inflation to stimulate employment. Hume [8, pp. 39-40] says as much in his advice to the policymaker.

> The good policy of the magistrate consists only in keeping [money], if possible still encreasing; because, by that means, he keeps alive a spirit of industry in the nation, and encreases the stock of labor, in which consists all real power and riches.

In contrast, Thornton opposed the exploitation of the Phillips curve for policy purposes. Such exploitation involved inflation, which he saw as an unmitigated evil. All inflationary policy, he [19, p. 239] said, is "attended with a proportionate hardship and injustice." True, output and employment would rise. But such gains, he thought, would be far too small to be worth the costs (uncertainty, injustice, social discontent) of higher inflation. In short, the Phillips curve at the economy's normal level of operations was very steeply sloped, allowing little increase in output per unit rise in inflation. Thus while "paper possesses the faculty of enlarging the quantity of commodities by giving life to some new industry," the unfavorable tradeoff ensures that "the increase of industry will by no means keep pace with the augmentation of paper." Moreover, because the economy normally operates close to its absolute full capacity ceiling, stimulative policy will quickly reach the point where

> it is obvious that the antecedently idle persons to whom we may suppose the [monetary inflation] to give employ, are limited in number; and that, therefore, if the encreased issue is indefinite, it will set to work labourers, of whom a part will be drawn from other, and, perhaps, no less useful occupations.

On these grounds he [19, p. 236] concluded that there exist narrow "bounds to the benefit which is to be derived from an augmentation of paper; and, also, that a liberal, or, at most, a large increase of it, will have all the advantageous effects of the most extravagant emission."

The Attwood-Mill Debate

The Phillips curve concept continued to flourish in the hands of more than one British classical writer after Henry Thornton. That this is so is evident from a glance at the celebrated interchange between Thomas Attwood (1783-1856) and John Stuart Mill (1806-1873) in the 1820s. Attwood, an inflationist proponent of inconvertible paper currency regimes and full employment at any cost, believed in a stable long-run tradeoff relationship of the form $U=g(P)$ where U and P denote unemployment and the price level, both taken relative to their normal (base period) values. Attwood used this relation, in which the inflation variable enters as a price *level* rather than its Hume-Thornton rate of change, to argue (1) that high unemployment stems from low prices, (2) that low unemployment emanates from high prices, and (3) that the government can and should achieve a zero target rate of unemployment with inflationary monetary expansion. For him nothing short of absolute full employment would suffice. Said he [3, p. 467], "so long as any number of industrious honest workmen in the Kingdom are out of employment, supposing such deficiency of employment not to be local but general, I should think it the duty, and certainly the interest, of Government, to continue the depreciation of the currency until full employment is obtained and general prosperity." "Restore the depreciated state of the currency," he [2, p. 66] declared, and "you restore everything that constitutes the commercial prosperity of the nation."

Opposing him was John Stuart Mill who reasoned in terms of the relationship $U=g(P-P^E)$ where U is the discrepancy between unemployment and its natural steady-state level, P is the price level, and P^E is its expected or perceived level. Using this relationship, Mill argued (1) that tradeoffs are temporary, (2) that they stem from unexpected price changes and vanish once perceptions adjust to reality, and (3) that, contrary to Attwood, one cannot peg real activity at arbitrarily low levels simply by pegging a nominal price (or inflation) variable since the two variables are independent of each other in steady-state equilibrium [9].

To be sure, Mill admitted that a temporary inflationary stimulus is possible. It is true, he [13, p. 79] said, that an unexpected inflation, if misperceived as a rise in relative prices, "may create a *false opinion* of an increase of demand; which false opinion leads, as the reality would do, to an increase of production." But it is also true that the real expansion is "followed . . . by a fatal revulsion as soon as the delusion ceases." In other words, once producers correctly perceive price increases as nominal rather than real, economic activity reverts to its steady-state level, but only after undergoing a temporary recession to correct for the excesses of the inflationary boom.

In Mill's view, the steady-state Phillips curve is a vertical line at the economy's natural rate of unemployment. To assert otherwise (as Attwood did), he thought, was to argue that people can be fooled perpetually into believing that nominal gains are real and that commodities can be created from paper money expansion. But according to Mill, one cannot fool all the people all the time. Money illusion, he contended, is not permanent. Attempts to peg real activity are therefore bound to be futile. Inflation cannot permanently stimulate activity. Mill's reply to Attwood dispels the notion that expectations-augmented Phillips curves and the natural rate hypothesis are of recent origin.

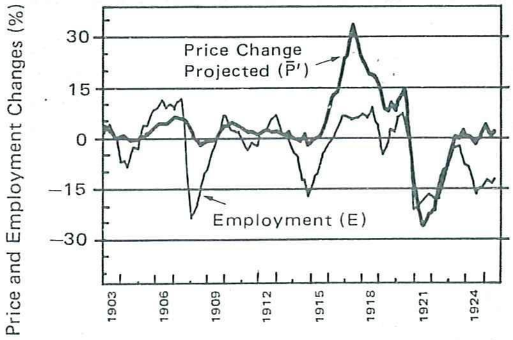
Irving Fisher (1867-1947)

As noted above, Hume and Thornton helped lay the theoretical foundations of the particular Phillips curve relationship $U=g(dP/dt)$. It was Irving Fisher, however, who provided the first statistical evidence of that relationship [7]. In his 1926 *International Labour Review* article, "A Statistical Relationship Between Unemployment and Price Changes," he investigated the correlation between unemployment U and lagged price changes $(dP/dt)_L$, where the subscript L denotes a linear distributed lag (Fisher himself being the inventor of the lag distribution concept) on the price-change variable. Using monthly U. S. data for the period 1915-1925, he obtained correlation coefficients as high as 90 percent between the two variables. Likewise, his time series chart displayed a similar strong correspondence between lagged price changes and employment (see Figure 2). From this evidence he concluded that there was indeed a strong relationship between them. He [7, p. 502] also concluded that the relationship was causal as well as empirical, that causality runs undirectionally from price changes to unemployment, and that there are good theoretical reasons for this being so. His theory of price-to-unemployment causality relies on fixed contracts, the inertia of custom, and other inhibiting factors that prevent costs from adjusting as fast as prices when prices change. Owing to the lag of costs behind prices, changes in the latter affect profits and thereby the level of real activity and employment. Via this linkage, causality, he argued, runs from inflation to unemployment as confirmed by his finding that the former variable leads the latter.

Figure 2

IRVING FISHER'S 1926 CHART

SHOWING CO-MOVEMENTS BETWEEN LAGGED PRICE CHANGES ($\bar{P}'$) AND EMPLOYMENT (E)



$\bar{P}'$ denotes a distributed lag or moving weighted average of past price changes. Fisher interpreted the statistical correlation between the two series as evidence of a causal relation running from price changes to employment.

Source: Fisher [7, p. 502]

Jan Tinbergen

Although he presented no formal econometric equations, Fisher was the first to offer empirical corroboration of the Phillips curve's market clearing version $U=g(dP/dt)$ according to which causality runs from inflation to unemployment. By contrast, Jan Tinbergen [4] in 1936 was the first to estimate the alternative shift-augmented wage-change version $w=f(U)+Z$ in which causality runs from unemployment or some equivalent measure of demand pressure in the labor market to the wage inflation rate and a vector of shift variables enters to affect the wage-unemployment tradeoff. More precisely, his equation was of the form $\Delta W=F(E,\Delta P_{-1})$ where ΔW is the change in money wages, E is employment relative to its normal (i.e., trend) level, and the lagged price-change variable ΔP_{-1} represents catch-up or cost-of-living wage adjustment factors thought capable of shifting the curve. Thus in his "An Economic Policy for 1936" he presents the expression $\Delta W = 0.16\ E + 0.27\ \Delta P_{-1}$ in which the numerical coefficients are estimated from the Netherlands data for the period 1923-1933.

About this equation three things must be said. It was the first econometric Phillips curve equation ever to appear in print. It also was the first to explain the tradeoff in terms of the law of supply and demand according to which the price of any good or service (including labor) varies in proportion to the excess demand for it. In other words, for the first time the Phillips curve was interpreted as a wage-reaction function relating the disequilibrium response of wages to demand pressure in the labor market, this pressure being measured by employment relative to trend. Finally, as mentioned above, Tinbergen's equation was the first to include a price change shift variable to account for observed movements in the wage-employment relationship. In these respects, it foreshadowed 1960s-vintage wage equations that likewise represented the Phillips curve as a demand-pressure wage-response function subject to shifts owing to changes in the cost of living.

Tinbergen returned to the Phillips curve issue once again in his *Business Cycles in the United Kingdom 1870-1914,* published in 1951 fully seven years before Phillips' contribution. There, using W to denote wages and E to denote employment, he [21, p. 50] writes the Phillips curve equation as

$$dW/dt = \dot{W} = f(E) = \lambda E$$

and gives it the excess-demand wage-reaction interpretation. "The theory expressed" in the equation, he says, "may be given the well-known formulation that a high unemployment figure 'exerts a pressure on' the wage rate and that, on the other hand, a small unemployment figure causes wages to go up." He also notes that the equation's empirical fit might be improved if the demand-pressure variable were entered nonlinearly and that this could be accomplished by replacing the employment variable E with the inverse of the unemployment rate U^{-1}. Finally, he suggested adding variables representing cost-of-living changes and the degree of unionization of the labor force to the equation to improve its statistical fit. On all of these innovations he pioneered the practice of fitting econometric Phillips curve equations.

Klein and Goldberger

Lawrence Klein and Arthur Goldberger also estimated econometric inflation-unemployment equations before Phillips. In their famous 1955 study *An Econometric Model of the United States, 1929-1952,* they [11, p. 19] presented a wage-change Phillips curve equation of the form $\Delta W = F(U, \Delta P_{-1})$. More precisely, their equation was

$$\Delta W = 4.11 - 0.74\ U + 0.52\ \Delta P_{-1} + 0.54\ t$$

where U is total unemployment, t is a time trend in years (t=1 in 1929), and the other variables are as defined above.

Like Tinbergen, Klein and Goldberger expressed the wage inflation variable in first difference rather than percentage rate of change form. Besides including a time trend variable, they also entered the unemployment variable linearly rather than nonlinearly into their equation. Except for these minor differences, their equation is virtually the same as the later formulations of Phillips and R. G. Lipsey, who clarified and extended Phillips' work. And like those latter writers, Klein and Goldberger interpreted their equation as a wage-reaction function in which money wages change in response to excess labor demand in an effort to clear the market. According to them [11, p. 18]

> the main reasoning behind this equation is that of the law of supply and demand. Money wage rates move in response to excess supply or excess demand in the labor market. High unemployment represents high excess supply, and low unemployment below customary frictional levels represents excess demand.

Here is the essence of the Phillips-Lipsey interpretation, an interpretation that also runs in terms of the law of supply and demand.

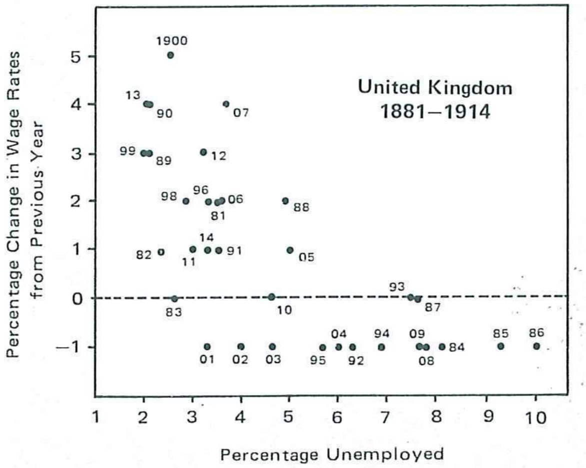
A. J. Brown and Paul Sultan

As documented above, the theoretical, empirical, and econometric foundations of the Phillips curve had been thoroughly established by the mid-1950s, several years in advance of Phillips' own contribution. It remained, however, for someone to present a Phillips-type relationship on a statistical scatter diagram and then to draw the familiar downward-sloping convex tradeoff curve that bears his name. Credit for being the first to accomplish these tasks goes not to Phillips himself but rather to two other economists, A. J. Brown and Paul Sultan.

The former, in his 1955 volume *The Great Inflation 1939-1951,* presented scatter diagrams similar to Phillips' (see Figure 3) that plotted annual wage inflation rates against unemployment rates for the United Kingdom for the periods 1880-1914 and 1920-1951, and for the United States for the period 1921-1948. From these charts Brown [5, pp. 91-101] concluded (1) that the two variables are inversely related, and (2) that the relationship between them is nonlinear since wages change at faster rates at low than at high rates of unemployment. He also used his charts to estimate the critical noninflationary level of unemployment below which wage inflation exceeds productivity growth so that prices rise. He did not, however, fit a curve to his data. Thus, although he presented a Phillips-type graph, he failed to draw the eye-catching curve made famous by

Figure 3

A.J. BROWN'S SCATTER DIAGRAM



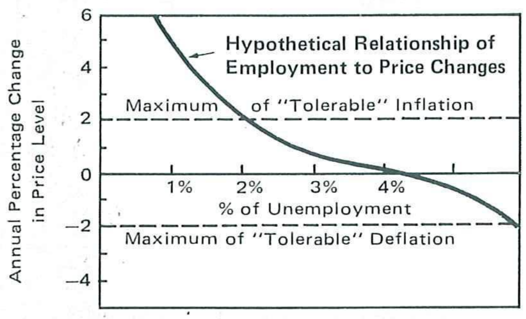
Brown's scatter diagram, presented 3 years before Phillips', shows an inverse nonlinear relation between wage inflation and unemployment.

Source: Brown [5, pp. 99-100]

Figure 4

PAUL SULTAN'S ANTICIPATION OF THE PHILLIPS CURVE

The Hypothetical Relationship of the "Fullness" of Employment to Annual Price Changes



Sultan's hypothetical curve associates 4% unemployment with price stability, 2% unemployment with an assumed maximum tolerable rate of inflation of 2%, and 6% unemployment with a maximum tolerable deflation rate of 2%.

Source: Sultan [17, p. 555]

Phillips. For this reason, one must reject A. P. Thirlwall's [18] contention that the curve should bear Brown's name rather than Phillips'.

Priority for drawing the Phillips curve goes to Paul Sultan, whose contribution predates Phillips' by one year. Thus, in his 1957 textbook *Labor Economics,* Sultan presents the curve in a diagram (see Figure 4) described by him [17, p. 555] as follows:

> the vertical scale measures the annual changes in the price level expressed as a percentage, while the horizontal scale measures the percentage of the work force unemployed. The line relating unemployment to inflation . . . is strictly hypothetical, but it suggests that the tighter the employment situation the greater the hazard of inflation . . . . Assuming that a fairly precise functional relationship exists between inflation and the level of employment, it is possible to determine the "safe" degree of full employment. In our hypothetical case, we are assuming that when unemployment is less than 2 percent of the work force, we face the dangers of inflation. And when unemployment is larger than 6 percent, we face the problem of serious deflation.

Here is the first diagrammatic representation of the price-change Phillips curve as a stable tradeoff relationship $p=f(U)$ between inflation and unemployment. On the basis of this diagram, three writers [1] recently have suggested that the Phillips curve could with equal justification be called the Sultan schedule.

Concluding Comments

Given the evidence presented in the preceding paragraphs, the label "Phillips curve tradeoff" must be judged both misleading and incomplete. For, as documented above, Phillips was far from the first to postulate an inflation-unemployment tradeoff or to draw the curve bearing his name. Even the econometric wage-price equations employed in modern Phillips curve analysis together with their excess demand and alternative market clearing interpretations long predate Phillips. In short, Phillips and his successors inherited (albeit unknowingly) these concepts; they did not invent them. In this sense at least, their work may be said to constitute the continuation rather than the origin of Phillips curve analysis.

Still, it was Phillips' formulation and not those of his predecessors that captured the attention of the economics profession. One must ask why this was so. Certainly it cannot be explained by the novelty of his curve or its empirical derivation; these were hardly innovations at the time he presented them. Nor can it be attributed to any originality in his explanation of his curve. His theory was simply the law of supply and demand according to which the price of any commodity or service (including labor) changes at a rate proportional to the excess demand for it. This explanation of course had been advanced by Tinbergen years before Phillips. Rather his phenomenal success probably stemmed from three factors. First was his striking finding of the apparent near 100-year empirical stability of his curve, a stability not suspected before. Second was the persuasive early expositions of his work provided by such influential economists as Lipsey [12], and Samuelson and Solow [16]. Especially important was the Samuelson-Solow interpretation of Phillips' curve as a menu of policy choices, a menu from which the authorities could select the best (or least undesirable) inflation-unemployment combination and then use their policy instruments to attain it. By providing a ready-made justification for discretionary intervention and activist fine tuning, this interpretation helped make the Phillips curve immensely popular among Keynesian policy advisors. Third was Phillips' presentation of his curve at just the right time to satisfy the Keynesians' search for an explanation of how changes in nominal income divide into price and quantity components. Whatever the reason, his name alone was attached to the tradeoff concept even though at least ten predecessors over a period of roughly 250 years also shared in its formulation.

References

1. Amid-Hozour, E.; D. T. Dick; and R. L. Lucier. "Sultan Schedule and Phillips Curve; an Historical Note." *Economica* 38 (August 1971): 319-20.

2. Attwood, Thomas. *The Remedy; or, Thoughts on the Present Distresses.* Second edition, with additions. London: 1816.

3. ————. *Evidence Before the Select Committee on the Bank of England Charter*, 1831-2, p. 467.

4. Bacon, Robert. "The Phillips Curve: Another Forerunner." *Economica* 40 (August 1973): 314-15.

5. Brown, A. J. *The Great Inflation, 1939-1951.* London: Oxford University Press, 1955.

6. Donner, Arthur and James F. McCollum. "The Phillips Curve: An Historical Note." *Economica* 39 (August 1972): 323-24.

7. Fisher, Irving. "A Statistical Relation between Unemployment and Price Changes." *International Labour Review* 13 (June 1926): 785-92. Reprinted as "I Discovered the Phillips Curve." *Journal of Political Economy* 81 (March/April 1973): 496-502.

8. Hume, David. "Of Money" (1752). Reprinted in his *Writings on Economics.* Edited by Eugene Rotwein. Madison: University of Wisconsin Press, 1955.

9. Humphrey, Thomas M. "Two Views of Monetary Policy: The Attwood-Mill Debate Revisited." *Economic Review*, Federal Reserve Bank of Richmond 63 (September/October 1977): 14-22.

10. ————. "Of Hume, Thornton, the Quantity Theory, and the Phillips Curve." *Economic Review*, Federal Reserve Bank of Richmond 68 (November/December 1982): 13-18.

11. Klein, Lawrence R. and Arthur S. Goldberger. *An Econometric Model of the United States 1929-1952.* Amsterdam: North-Holland Publishing Company, 1955.

12. Lipsey, R. G. "The Relation between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1862-1957: A Further Analysis." *Economica* 27 (February 1960): 1-32.

13. Mill, John Stuart. "The Currency Juggle." *Tait's Edinburg Magazine* (1833). Reprinted in Vol. I of his *Dissertations and Discussions.* Boston: 1865.

14. Nelson, Charles R. "Adjustment Lags Versus Information Lags: A Test of Alternative Explanations of the Phillips Curve Phenomenon." *Journal of Money, Credit and Banking* 13 (February 1981): 1-11.

15. Phillips, A. W. "The Relation between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861-1957." *Economica* 25 (November 1958): 283-99.

16. Samuelson, Paul A., and Robert M. Solow. "Analytical Aspects of Anti-inflation Policy." *American Economic Review* 50 (May 1960): 177-94.

17. Sultan, Paul. *Labor Economics.* New York: Henry Holt and Company, Inc., 1957.

18. Thirlwall, A. P. "The Phillips Curve: An Historical Note." *Economica* 39 (August 1972): 325.

19. Thornton, Henry. *An Enquiry into the Nature and Effects of the Paper Credit of Great Britain* (1802). Edited with an introduction by F. A. von Hayek. New York: Rinehart and Company, Inc., 1939.

20. Tinbergen, Jan. "An Economic Policy for 1936." Reprinted in his *Selected Papers.* Edited by L. H. Klaassen, L. M. Koyck, and H. J. Witteveen. Amsterdam: North-Holland Publishing Company, 1959.

21. ————. *Business Cycles in the United Kingdom, 1870-1914.* Amsterdam: North-Holland Publishing Company, 1951.

THE EVOLUTION AND POLICY IMPLICATIONS OF PHILLIPS CURVE ANALYSIS

At the core of modern macroeconomics is some version or another of the famous Phillips curve relationship between inflation and unemployment. The Phillips curve, both in its original and more recently reformulated expectations-augmented versions, has two main uses. In theoretical models of inflation, it provides the so-called "missing equation" that explains how changes in nominal income divide themselves into price and quantity components. On the policy front, it specifies conditions contributing to the effectiveness (or lack thereof) of expansionary and disinflationary policies. For example, in its expectations-augmented form, it predicts that the power of expansionary measures to stimulate real activity depends critically upon how price anticipations are formed. Similarly, it predicts that disinflationary policy will either work slowly (and painfully) or swiftly (and painlessly) depending upon the speed of adjustment of price expectations. In fact, few macro policy questions are discussed without at least some reference to an analytical framework that might be described in terms of some version of the Phillips curve.

As might be expected from such a widely used tool, Phillips curve analysis has hardly stood still since its beginnings in 1958. Rather it has evolved under the pressure of events and the progress of economic theorizing, incorporating at each stage such new elements as the natural rate hypothesis, the adaptive-expectations mechanism, and most recently, the rational expectations hypothesis. Each new element expanded its explanatory power. Each radically altered its policy implications. As a result, whereas the Phillips curve was once seen as offering a stable enduring trade-off for the policymakers to exploit, it is now widely viewed as offering no trade-off at all. In short, the original Phillips curve notion of the potency of activist fine tuning has given way to the revised Phillips curve notion of policy ineffectiveness. The purpose of this article is to trace the sequence of steps that led to this change. Accordingly, the paragraphs below sketch the evolution of Phillips curve analysis, emphasizing in particular the theoretical innovations incorporated into that analysis at each stage and the policy implications of each innovation.

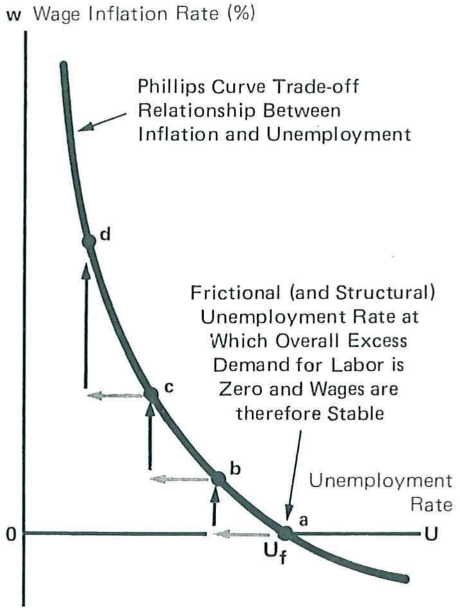
I.

EARLY VERSIONS OF THE PHILLIPS CURVE

The idea of an inflation-unemployment trade-off is hardly new. It was a key component of the monetary doctrines of David Hume (1752) and Henry Thornton (1802). It was identified statistically by Irving Fisher in 1926, although he viewed causation as running from inflation to unemployment rather than vice versa. It was stated in the form of an econometric equation by Jan Tinbergen in 1936 and again by Lawrence Klein and Arthur Goldberger in 1955. Finally, it was graphed on a scatterplot chart by A. J. Brown in 1955 and presented in the form of a diagrammatic curve by Paul Sultan in 1957. Despite these early efforts, however, it was not until 1958 that modern Phillips curve analysis can be said to have begun. That year saw the publication of Professor A. W. Phillips' famous article in which he fitted a statistical equation $w=f(U)$ to annual data on percentage rates of change of money wages (w) and the unemployment rate (U) in the United Kingdom for the period 1861-1913. The result, shown in a chart like Figure 1 with wage inflation measured vertically and unemployment horizontally, was a smooth, downward-sloping convex curve that cut the horizontal axis at a positive level of unemployment.

The curve itself was given a straightforward interpretation: it showed the response of wages to the excess demand for labor as proxied by the inverse of the unemployment rate. Low unemployment spelled high excess demand and thus upward pressure on wages. The greater this excess labor demand the

Figure 1
EARLY PHILLIPS CURVE



At unemployment rate U_f the labor market is in equilibrium and wages are stable. At lower unemployment rates excess demand exists to bid up wages. At higher unemployment rates excess supply exists to bid down wages. The curve's convex shape shows that increasing excess demand for labor runs into diminishing marginal returns in reducing unemployment. Thus successive uniform decreases in unemployment (horizontal gray arrows) require progressively larger increases in excess demand and hence wage inflation rates (vertical black arrows) as we go from point **a** to **b** to **c** to **d** along the curve.

faster the rise in wages. Similarly, high unemployment spelled negative excess demand (i.e., excess labor supply) that put deflationary pressure on wages. Since the rate of change of wages varied directly with excess demand, which in turn varied inversely with unemployment, wage inflation would rise with decreasing unemployment and fall with increasing unemployment as indicated by the negative slope of the curve. Moreover, owing to unavoidable frictions in the operation of the labor market, it followed that some frictional unemployment would exist even when the market was in equilibrium, that is, when excess labor demand was zero and wages were stable. Accordingly, this frictional unemployment was indicated by the point at which the Phillips curve crosses the horizontal axis. According to Phillips, this is also the point to which the economy returns if the authorities ceased to maintain disequilibrium in the labor market by pegging the excess demand for labor. Finally, since increases in excess demand would likely run into diminishing marginal returns in reducing unemployment, it followed that the curve must be convex—this convexity showing that successive uniform decrements in unemployment would require progressively larger increments in excess demand (and thus wage inflation rates) to achieve them.

Popularity of the Phillips Paradigm

Once equipped with the foregoing theoretical foundations, the Phillips curve gained swift acceptance among economists and policymakers alike. It is important to understand why this was so. At least three factors probably contributed to the attractiveness of the Phillips curve. One was the remarkable temporal stability of the relationship, a stability revealed by Phillips' own finding that the same curve estimated for the pre-World War I period 1861-1913 fitted the United Kingdom data for the post-World War II period 1948-1957 equally well or even better. Such apparent stability in a two-variable relationship over such a long period of time is uncommon in empirical economics and served to excite interest in the curve.

A second factor contributing to the success of the Phillips curve was its ability to accommodate a wide variety of inflation theories. The Phillips curve itself explained inflation as resulting from excess demand that bids up wages and prices. It was entirely neutral, however, about the causes of that phenomenon. Now excess demand can of course be generated either by shifts in demand or shifts in supply regardless of the causes of those shifts. Thus a demand-pull theorist could argue that excess-demand-induced inflation stems from excessively expansionary aggregate demand policies while a cost-push theorist could claim that it emanates from trade-union monopoly power and real shocks operating on labor supply. The Phillips curve could accommodate both views. Economists of rival schools could accept the Phillips curve as offering insights into the nature of the inflationary process even while disagreeing on the causes of and appropriate remedies for inflation.

Finally, the Phillips curve appealed to policymakers because it provided a convincing rationale for their apparent failure to achieve full employment with price stability—twin goals that were thought to be mutually compatible before Phillips' analysis. When criticized for failing to achieve both goals simultaneously, the authorities could point to the Phillips curve as showing that such an outcome was impossible and that the best one could hope for was either arbitrarily low unemployment or price stability but not both. Note also that the curve, by offering a menu of alternative inflation-unemployment combinations from which the authorities could choose, provided a ready-made justification for discretionary intervention and activist fine tuning. Policymakers had but to select the best (or least undesirable) combination on the menu and then use their policy instruments to achieve it. For this reason too the curve must have appealed to some policy authorities, not to mention the economic advisors who supplied the cost-benefit analysis underlying their choices.

From Wage-Change Relation to Price-Change Relation

As noted above, the initial Phillips curve depicted a relation between unemployment and wage inflation. Policymakers, however, usually specify inflation targets in terms of rates of change of prices rather than wages. Accordingly, to make the Phillips curve more useful to policymakers, it was therefore necessary to transform it from a wage-change relationship to a price-change relationship. This transformation was achieved by assuming that prices are set by applying a constant mark-up to unit labor cost and so move in step with wages—or, more precisely, move at a rate equal to the differential between the percentage rates of growth of wages and productivity (the latter assumed zero here).[1] The result of this transformation was the price-change Phillips relation

(1) $p = ax(U)$

where p is the rate of price inflation, x(U) is overall excess demand in labor and hence product markets—this excess demand being an inverse function of the unemployment rate—and a is a price-reaction coefficient expressing the response of inflation to excess demand. From this equation the authorities could determine how much unemployment would be associated with any given target rate of inflation. They could also use it to measure the effect of policies undertaken to obtain a more favorable Phillips curve, i.e., policies aimed at lowering the price-response coefficient and the amount of unemployment associated with any given level of excess demand.

Trade-Offs and Attainable Combinations

The foregoing equation specifies the *position* (or distance from origin) and *slope* of the Phillips curve —two features stressed in policy discussions of the early 1960s. As seen by the policymakers of that era, the curve's position fixes the inner boundary, or frontier, of feasible (attainable) combinations of inflation and unemployment rates (see Figure 2). Determined by the structure of labor and product markets, the position of the curve defines the set of all coordinates of inflation and unemployment rates the authorities could achieve via implementation of monetary and fiscal policies. Using these macroeconomic demand-management policies the authorities could put the economy anywhere on the curve. They could not, however, operate to the left of it. The Phillips curve was viewed as a constraint preventing them from achieving still lower levels of both inflation and unemployment. Given the structure of labor and product markets, it would be impossible for monetary and fiscal policy alone to reach inflation-unemployment combinations in the region to the left of the curve.

The *slope* of the curve was interpreted as showing the relevant policy trade-offs (rates of exchange between policy goals) available to the authorities. As explained in early Phillips curve analysis, these trade-offs arise because of the existence of irreconcilable conflicts among policy objectives. When the goals of full employment and price stability are not simultaneously achievable, then attempts to move the economy closer to one will necessarily move it further away from the other. The rate at which one objective must be given up to obtain a little bit more of the other is measured by the slope of the Phillips curve. For example, when the Phillips curve is steeply sloped, it means that a small reduction in unemploy-

[1] Let prices P be the product of a fixed markup K (including normal profit margin and provision for depreciation) applied to unit labor costs C,

(1) $P = KC$.

Unit labor costs by definition are the ratio of hourly wages W to labor productivity or output per labor hour Q,

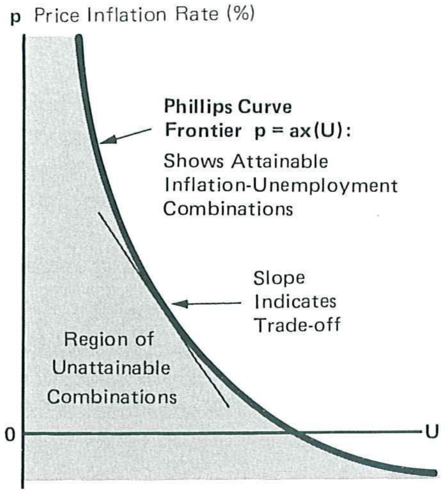
(2) $C = W/Q$.

Substituting (2) into (1), taking logarithms of both sides of the resulting expression, and then differentiating with respect to time yields

(3) $p = w - q$

where the lower case letters denote the percentage rates of change of the price, wage, and productivity variables. Assuming productivity growth q is zero and the rate of wage change w is an inverse function of the unemployment rate yields equation (1) of the text.

Figure 2
TRADE-OFFS AND ATTAINABLE COMBINATIONS



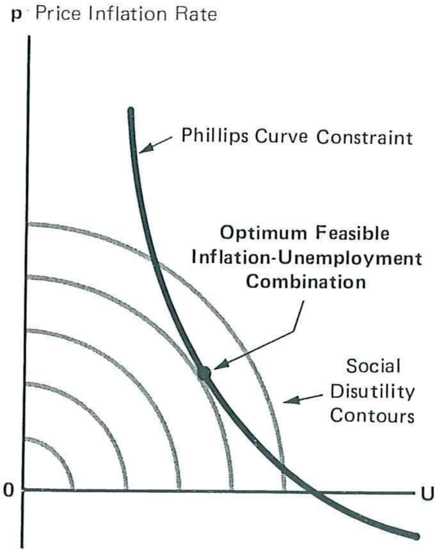
The *position* or *location* of the Phillips curve defines the frontier or set of attainable inflation-unemployment combinations. Using monetary and fiscal policies, the authorities can attain all combinations lying upon the frontier itself but none in the shaded region below it. In this way the curve acts as a constraint on demand-management policy choices. The *slope* of the curve shows the trade-offs or rates of exchange between the two evils of inflation and unemployment.

ment would be purchased at the cost of a large increase in the rate of inflation. Conversely, in relatively flat portions of the curve, considerably lower unemployment could be obtained fairly cheaply, that is at the cost of only slight increases in inflation. Knowledge of these trade-offs would enable the authorities to determine the price-stability sacrifice necessary to buy any given reduction in the unemployment rate.

The Best Selection on the Phillips Frontier

The preceding has described the early view of the Phillips curve as a stable, enduring trade-off permitting the authorities to obtain permanently lower rates of unemployment in exchange for permanently higher rates of inflation or vice versa. Put differently, the curve was interpreted as offering a menu of alternative inflation-unemployment combinations from which the authorities could choose. Given the menu, the authorities' task was to select the particular inflation-unemployment mix resulting in the smallest social cost (see Figure 3). To do this, they would have to assign relative weights to the twin evils of

Figure 3
THE BEST SELECTION ON THE MENU OF CHOICES



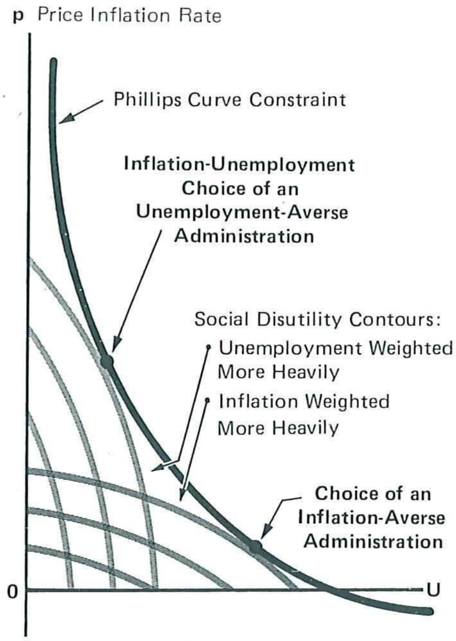
The bowed-out curves are social disutility contours. Each contour shows all the combinations of inflation and unemployment resulting in a given level of social cost or harm. The closer to the origin, the lower the social cost. The slopes of these contours reflect the relative weights that society (or the policy authority) assigns to the evils of inflation and unemployment. The best combination of inflation and unemployment that the policymakers can reach, given the Phillips curve constraint, is the mix appearing on the lowest attainable social disutility contour. Here the additional social benefit from a unit reduction in unemployment will just be worth the extra inflation cost of doing so.

inflation and unemployment in accordance with their views of the comparative harm caused by each. Then, using monetary and fiscal policy, they would move along the Phillips curve, trading off unemployment for inflation (or vice versa) until they reached the point at which the additional benefit from a further reduction in unemployment was just worth the extra inflation cost of doing so. Here would be the optimum, or least undesirable, mix of inflation and unemployment. At this point the economy would be on its lowest attainable social disutility contour (the bowed-out curves radiating outward from the origin of Figure 3) allowed by the Phillips curve constraint. Here the unemployment-inflation combination chosen would be the one that minimized social harm. It was of course understood that if this outcome involved a positive rate of inflation, continuous excess money growth would be required to maintain it. For without such monetary stimulus, excess demand would disappear and the economy would return to the point at which the Phillips curve crosses the horizontal axis.

Different Preferences, Different Outcomes

It was also recognized that policymakers might differ in their assessment of the comparative social cost of inflation vs. unemployment and thus assign different policy weights to each. Policymakers who believed that unemployment was more undesirable than rising prices would assign a much higher relative weight to the former than would policymakers who judged inflation to be the worse evil. Hence, those with a marked aversion to unemployment would prefer a point higher up on the Phillips curve than would those more anxious to avoid inflation, as shown in Figure 4. Whereas one political administration might opt for a high pressure economy on the grounds that the social benefits of low unemployment exceeded the harm done by the inflation necessary to achieve it, another administration might deliberately aim for a low pressure economy because it believed that some economic slack was a relatively painless means of eradicating harmful inflation. Both groups would of course prefer combinations to the southwest of the Phillips constraint, down closer to the figure's origin (the ideal point of zero inflation and zero unemployment). As pointed out before, however, this would be impossible given the structure of the economy, which determines the position or location of the Phillips frontier. In short, the policymakers would be constrained to combinations lying on this boundary, unless they were prepared to alter the economy's structure.

Figure 4
DIFFERENT PREFERENCES, DIFFERENT POLICY CHOICES



Different political administrations may differ in their evaluations of the social harmfulness of inflation relative to that of unemployment. Thus in their policy deliberations they will attach different relative weights to the two evils of inflation and unemployment. These weights will be reflected in the slopes of the social disutility contours (as those contours are interpreted by the policymakers). The relatively flat contours reflect the views of those attaching higher relative weight to the evils of inflation; the steep contours to those assigning higher weight to unemployment. An unemployment-averse administration will choose a point on the Phillips curve involving more inflation and less unemployment than the combination selected by an inflation-averse administration.

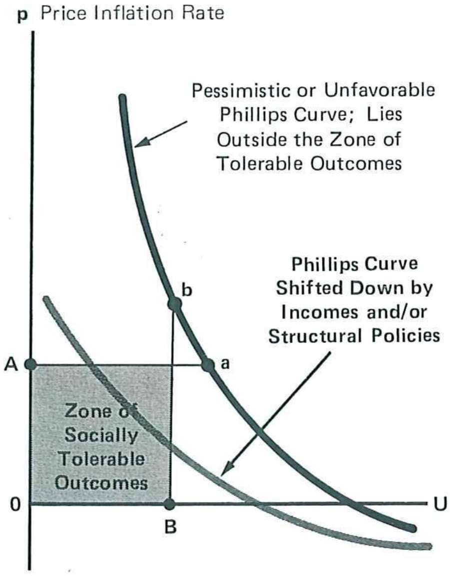
Pessimistic Phillips Curve and the "Cruel Dilemma"

In the early 1960s, there was much discussion of the so-called "cruel-dilemma" problem imposed by an unfavorable Phillips curve. The cruel dilemma refers to certain pessimistic situations where *none* of the available combinations on the menu of policy choices is acceptable to the majority of a country's voters (see Figure 5). For example, suppose there is some maximum rate of inflation, A, that voters are just willing to tolerate without removing the party in power. Likewise, suppose there is some maximum tolerable rate of unemployment, B. As shown in Figure 5, these limits define the zone of acceptable or politically feasible combinations of inflation and unemployment. A Phillips curve that occupies a position anywhere within this zone will satisfy society's demands for reasonable price stability and high employment. But if both limits are exceeded and the curve lies outside the region of satisfactory outcomes, the system's performance will fall short of what was expected of it, and the resulting discontent may severely aggravate political and social tensions.

If, as some analysts alleged, the Phillips curve tended to be located so far to the right in the chart that no portion of it fell within the zone of acceptable combinations, then the policymakers would indeed be confronted with a painful dilemma. At best they could hold only one of the variables, inflation or unemployment, down to acceptable levels. But they could not hold both simultaneously within the limits of toleration. Faced with such a pessimistic Phillips curve, policymakers armed only with traditional demand-management policies would find it impossible to achieve combinations of inflation and unemployment acceptable to society.

Figure 5

PESSIMISTIC PHILLIPS CURVE AND THE "CRUEL DILEMMA"



A = Maximum Tolerable Rate of Inflation
B = Maximum Tolerable Rate of Unemployment

Given the unfavorable Phillips curve, policymakers are confronted with a cruel choice. They can achieve acceptable rates of inflation (point a) or unemployment (point b) but not both. The rationale for *incomes* (wage-price) and ***structural*** (labor market) policies was to shift the Phillips curve down into the zone of tolerable outcomes.

Policies to Shift the Phillips Curve

It was this concern and frustration over the seeming inability of monetary and fiscal policy to resolve the unemployment-inflation dilemma that induced some economists in the early 1960s to urge the adoption of incomes (wage-price) and structural (labor-market) policies. Monetary and fiscal policies alone were thought to be insufficient to resolve the cruel dilemma since the most these policies could do was to enable the economy to occupy alternative positions on the pessimistic Phillips curve. That is, monetary and fiscal policies could move the economy *along* the given curve, but they could not move the curve itself into the zone of tolerable outcomes. What was needed, it was argued, were new policies that would twist or shift the Phillips frontier toward the origin of the diagram.

Of these measures, incomes policies would be directed at the price-response coefficient linking inflation to excess demand. Either by decreeing this coefficient to be zero (as with wage-price freezes), or by replacing it with an officially mandated rate of price increase, or simply by persuading sellers to moderate their wage and price demands, such policies would lower the rate of inflation associated with any given level of unemployment and thus twist down the Phillips curve. The idea was that wage-price controls would hold inflation down while excess demand was being used to boost employment.

Should incomes policies prove unworkable or prohibitively expensive in terms of their resource-misallocation and restriction-of-freedom costs, then the authorities could rely solely on microeconomic structural policies to improve the trade-off. By enhancing the efficiency and performance of labor and product markets, these latter policies could lower the Phillips curve by reducing the amount of unemployment associated with any given level of excess demand. Thus the rationale for such measures as job-training and retraining programs, job-information and job-counseling services, relocation subsidies, anti-discrimination laws and the like was to shift the Phillips frontier down so that the economy could obtain better inflation-unemployment combinations.

II.

INTRODUCTION OF SHIFT VARIABLES

Up until the mid-1960s the Phillips curve received widespread and largely uncritical acceptance. Few questioned the usefulness, let alone the existence, of this construct. In policy discussions as well as economic textbooks, the Phillips curve was treated as a stable, enduring relationship or menu of policy choices. Being stable (and barring the application of incomes and structural policies), the menu never changed.

Empirical studies of the 1900-1958 U. S. data soon revealed, however, that the menu for this country was hardly as stable as its original British counterpart and that the Phillips curve had a tendency to shift over time. Accordingly, the trade-off equation was augmented with additional variables to account for such movements. The inclusion of these shift variables marked the second stage of Phillips curve analysis and meant that the trade-off equation could be written as

(2) $p = ax(U) + z$

where z is a vector of variables—productivity, profits, trade union effects, unemployment dispersion and the like—thought capable of shifting the inflation-unemployment trade-off.

In retrospect, this vector or list was deficient both for what it included and what it left out. Excluded at this stage were variables representing inflation expectations—later shown to be a chief cause of the shifting short-run Phillips curve. Of the variables included, subsequent analysis would reveal that at least three—productivity, profits, and measures of union monopoly power—were redundant because they constituted underlying determinants of the demand for and supply of labor and as such were already captured by the excess demand variable, U. This criticism, however, did not apply to the unemployment dispersion variable, changes in which were independent of excess demand and were indeed capable of causing shifts in the aggregate Phillips curve.

To explain how the dispersion of unemployment across separate micro labor markets could affect the aggregate trade-off, analysts in the early 1960s used diagrams similar to Figure 6. That figure depicts a representative micromarket Phillips curve, the exact replica of which is presumed to exist in each local labor market and aggregation over which yields the macro Phillips curve. According to the figure, if a given national unemployment rate U* were equally distributed across local labor markets such that the same rate prevailed in each, then wages everywhere would inflate at the single rate indicated by the point w* on the curve. But if the same aggregate unemployment were unequally distributed across local markets, then wages in the different markets would inflate at different rates. Because of the curve's convexity (which renders wage inflation more responsive to leftward than to rightward deviations from average unemployment along the curve) the average of these wage inflation rates would exceed the rate of the no-dispersion case. In short, the diagram suggested that, for any given aggregate unemployment rate, the rate of aggregate wage inflation varies directly with the dispersion of unemployment across micromarkets, thus displacing the macro Phillips curve to the right.

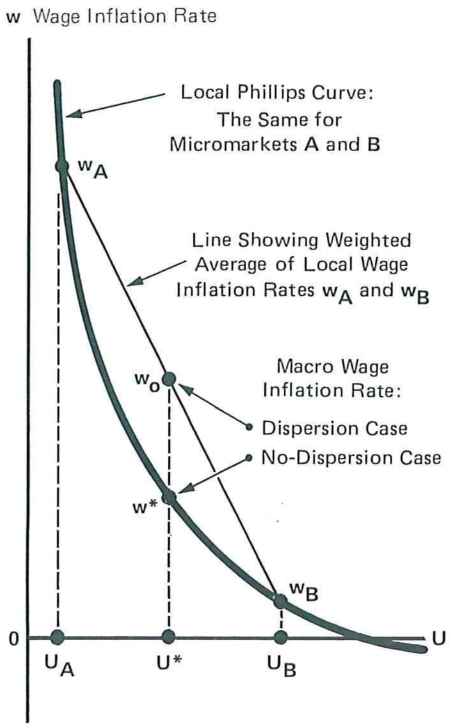
From this analysis, economists in the early 1960s concluded that the greater the dispersion, the greater the outward shift of the aggregate Phillips curve. To prevent such shifts, the authorities were advised to apply structural policies to minimize the dispersion of unemployment across industries, regions, and occupations. Also, they were advised to minimize unemployment's dispersion over time since, with a convex Phillips curve, the average inflation rate would be higher the more unemployment is allowed to fluctuate around its average (mean) rate.

A Serious Misspecification

The preceding has shown how shift variables were first incorporated into the Phillips curve in the early- to mid-1960s. Notably absent at this stage were variables representing price expectations. To be sure, the past rate of price change was sometimes used as a shift variable to represent catch-up or cost-of-living adjustment factors in wage and price demands. Rarely, however, was it interpreted as a proxy for anticipated inflation. Not until the late 1960s were expectational variables fully incorporated into Phillips curve equations. By then, of course,

Figure 6

EFFECTS OF UNEMPLOYMENT DISPERSION



If aggregate unemployment at rate **U*** were evenly distributed across individual labor markets such that the same rate prevailed everywhere, then wages would inflate at the rate **w*** both locally and nationally. But if aggregate unemployment **U*** is unequally distributed such that rate U_A exists in market **A** and U_B in market **B**, then wages will inflate at rate w_A in the former market and w_B in the latter. The average of these local inflation rates at aggregate unemployment rate **U*** is w_o which is higher than inflation rate **w*** of the no-dispersion case.

Conclusion: The greater the dispersion of unemployment, the higher the aggregate inflation rate associated with any given level of aggregate unemployment. Unemployment dispersion shifts the aggregate Phillips curve rightward.

inflationary expectations had become too prominent to ignore and many analysts were perceiving them as the dominant cause of observed shifts in the Phillips curve.

Coinciding with this perception was the belated recognition that the original Phillips curve involved a misspecification that could only be corrected by the incorporation of a price expectations variable in the trade-off. The original Phillips curve was expressed in terms of *nominal* wage changes, $w=f(U)$. Since neoclassical economic theory teaches that *real* rather than nominal wages adjust to clear labor markets, however, it follows that the Phillips curve should have been stated in terms of real wage changes. Better still (since wage bargains are made with an eye to the future), it should have been stated in terms of *expected* real wage changes, i.e., the differential between the rates of change of nominal wages and expected future prices, $w-p^e=f(U)$. In short, the original Phillips curve required a price expectations term to render it correct. Recognition of this fact led to the development of the expectations-augmented Phillips curve described below.

III.

THE EXPECTATIONS-AUGMENTED PHILLIPS CURVE AND THE ADAPTIVE-EXPECTATIONS MECHANISM

The original Phillips curve equation gave way to the expectations-augmented version in the early 1970s. Three innovations ushered in this change. The first was the respecification of the excess demand variable. Originally defined as an inverse function of the unemployment rate, $x(U)$, excess demand was redefined as the discrepancy or gap between the natural and actual rates of unemployment, U_N-U. The natural (or full employment) rate of unemployment itself was defined as the rate that prevails in steady-state equilibrium when expectations are fully realized and incorporated into all wages and prices and inflation is neither accelerating nor decelerating. It is natural in the sense (1) that it represents normal full-employment equilibrium in the labor and hence commodity markets, (2) that it is independent of the steady-state inflation rate, and (3) that it is determined by real structural forces (market frictions and imperfections, job information and labor mobility costs, tax laws, unemployment subsidies, and the like) and as such is not susceptible to manipulation by aggregate demand policies.

The second innovation was the introduction of price anticipations into Phillips curve analysis resulting in the expectations-augmented equation

$$(3)\quad p = a(U_N - U) + p^e$$

where excess demand is now written as the gap between the natural and actual unemployment rates and p^e is the price expectations variable representing the anticipated rate of inflation. This expectations variable entered the equation with a coefficient of unity, reflecting the assumption that price expectations are completely incorporated in actual price changes. The unit expectations coefficient implies the absence of money illusion, i.e., it implies that people are concerned with the expected real purchasing power of the prices they pay and receive (or, alternatively, that they wish to maintain their prices relative to the prices they expect others to be charging) and so take anticipated inflation into account. As will be shown later, the unit expectations coefficient also implies the complete absence of a trade-off between inflation and unemployment in long-run equilibrium when expectations are fully realized. Note also that the expectations variable is the sole shift variable in the equation. All other shift variables have been omitted, reflecting the view, prevalent in the early 1970s, that changing price expectations were the predominant cause of observed shifts in the Phillips curve.

Expectations-Generating Mechanism

The third innovation was the incorporation of an expectations-generating mechanism into Phillips curve analysis to explain how the price expectations variable itself was determined. Generally a simple *adaptive-expectations* or *error-learning mechanism* was used. According to this mechanism, expectations are adjusted (adapted) by some fraction of the forecast error that occurs when inflation turns out to be different than expected. In symbols,

$$(4)\quad \dot{p}^e = b(p - p^e)$$

where the dot over the price expectations variable indicates the rate of change (time derivative) of that variable, $p - p^e$ is the expectations or forecast error (i.e., the difference between actual and expected price inflation), and b is the adjustment fraction. Assuming, for example, an adjustment fraction of ½, equation 4 says that if the actual and expected rates of inflation are 10 percent and 4 percent, respectively—i.e., the expectational error is 6 percent—then the expected rate of inflation will be revised upward by an amount equal to half the error, or 3 percentage points. Such revision will continue until the expectational error is eliminated.

Analysts also demonstrated that equation 4 is equivalent to the proposition that expected inflation is a geometrically declining weighted average of all past rates of inflation with the weights summing to one. This unit sum of weights ensures that any constant rate of inflation eventually will be fully anticipated, as can be seen by writing the error-learning mechanism as

$$(5)\quad p^e = \Sigma v_i p_{-i}$$

where Σ indicates the operation of summing the past rates of inflation, the subscript i denotes past time periods, and v_i denotes the weights attached to past rates of inflation. With a stable inflation rate p unchanging over time and a unit sum of weights, the equation's right-hand side becomes simply p, indicating that when expectations are formulated adaptively via the error-learning scheme, any constant rate of inflation will indeed eventually be fully anticipated. Both versions of the adaptive-expectations mechanism (i.e., equations 4 and 5) were combined with the expectations-augmented Phillips equation to explain the mutual interaction of actual inflation, expected inflation, and excess demand.

The Natural Rate Hypothesis

These three innovations—the redefined excess demand variable, the expectations-augmented Phillips curve, and the error-learning mechanism—formed the basis of the celebrated *natural rate* and *accelerationist* hypotheses that radically altered economists' and policymakers' views of the Phillips curve in the late 1960s and early 1970s. According to the natural rate hypothesis, there exists no permanent trade-off between unemployment and inflation since real economic variables tend to be independent of nominal ones in steady-state equilibrium. To be sure, trade-offs may exist in the short run. For example, surprise inflation, if unperceived by wage earners, may, by raising product prices relative to nominal wages and thus lowering real wages, stimulate employment temporarily. But such trade-offs are inherently transitory phenomena that stem from *unexpected* inflation and that vanish once expectations (and the wages and prices embodying them) fully adjust to inflationary experience. In the long run, when inflationary surprises disappear and expectations are realized such that wages reestablish their preexisting levels relative to product prices, unemployment returns to its natural (equilibrium) rate. This rate is compatible with all fully anticipated steady-state rates of inflation, implying that the long-run Phillips curve is a vertical line at the natural rate of unemployment.

Equation 3 embodies these conclusions. That equation, when rearranged to read $p - p^e = a(U_N - U)$, states that the trade-off is between *unexpected* inflation (the difference between actual and expected inflation, $p - p^e$) and unemployment. That is, only *surprise* price increases could induce deviations of unemployment from its natural rate. The equation also says that the trade-off disappears when inflation is fully anticipated (i.e., when $p - p^e$ equals zero), a result guaranteed for any steady rate of inflation by the error-learning mechanism's unit sum of weights. Moreover, according to the equation, the right-hand side must also be zero at this point, which implies that unemployment is at its natural rate. The natural rate of unemployment is therefore compatible with any constant rate of inflation provided it is fully anticipated (which it eventually must be by virtue of the error-learning weights adding to one). In short, equation 3 asserts that inflation-unemployment trade-offs cannot exist when inflation is fully anticipated. And equation 5 ensures that this latter condition must obtain for all steady inflation rates such that the long-run Phillips curve is a vertical line at the natural rate of unemployment.[2]

The message of the natural rate hypothesis was clear. A higher stable rate of inflation could not buy a permanent drop in joblessness. Movements to the left along a short-run Phillips curve only provoke expectational wage/price adjustments that shift the curve to the right and restore unemployment to its natural rate (see Figure 7). In sum, Phillips curve trade-offs are inherently transitory phenomena. Attempts to exploit them will only succeed in raising the permanent rate of inflation without accomplishing a lasting reduction in the unemployment rate.

[2] Actually, the long-run Phillips curve may become positively sloped in its upper ranges as higher inflation leads to greater inflation variability (volatility, unpredictability) that raises the natural rate of unemployment. Higher and hence more variable and erratic inflation can raise the equilibrium level of unemployment by generating increased uncertainty that inhibits business activity and by introducing noise into market price signals, thus reducing the efficiency of the price system as a coordinating and allocating mechanism.

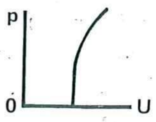
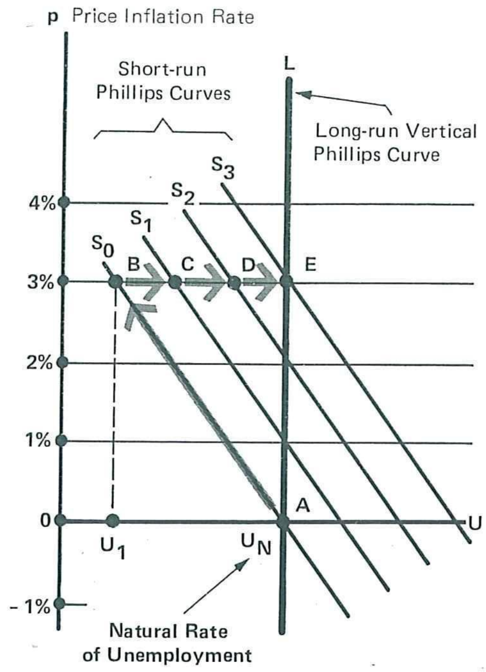


Figure 7

THE NATURAL RATE HYPOTHESIS AND ADJUSTMENT TO STEADY-STATE EQUILIBRIUM



The vertical line **L** through the natural rate of unemployment $\mathbf{U_N}$ is the long-run steady state Phillips curve along which all rates of inflation are fully anticipated. The downward-sloping lines are short-run Phillips curves each corresponding to a different given expected rate of inflation. Attempts to lower unemployment from the natural rate $\mathbf{U_N}$ to $\mathbf{U_1}$ by raising inflation to 3 percent along the short-run trade-off curve $\mathbf{S_0}$ will only induce shifts in the short-run curve to $\mathbf{S_1}$, $\mathbf{S_2}$, $\mathbf{S_3}$ as expectations adjust to the higher rate of inflation. The economy travels the path **ABCDE** to the new steady state equilibrium, point **E**, where unemployment is at its preexisting natural rate but inflation is higher than it was originally.

The Accelerationist Hypothesis

The expectations-augmented Phillips curve, when combined with the error-learning process, also yielded the celebrated *accelerationist* hypothesis that dominated many policy discussions in the inflationary 1970s. This hypothesis, a corollary of the natural rate concept, states that since there exists no long-run trade-off between unemployment and inflation, attempts to peg the former variable below its natural (equilibrium) level must produce ever-increasing inflation. Fueled by progressively faster monetary expansion, such price acceleration would keep actual inflation always running ahead of expected inflation, thereby perpetuating the inflationary surprises that prevent unemployment from returning to its equilibrium level (see Figure 8).

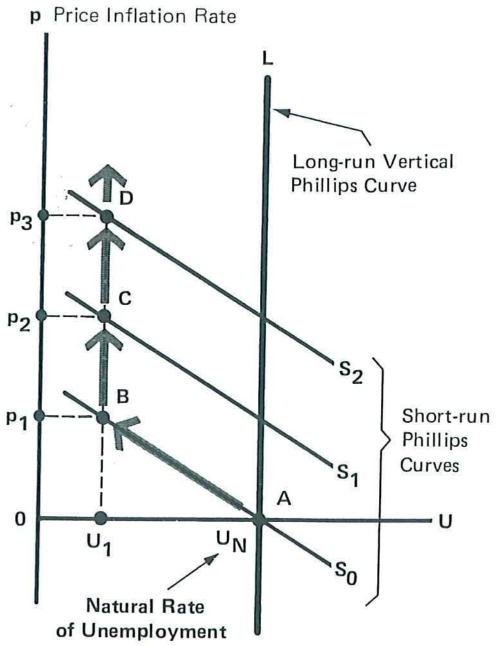
Accelerationists reached these conclusions via the following route. They noted that equation 3 posits that unemployment can differ from its natural level only so long as actual inflation deviates from expected inflation. But that same equation together with equation 4 implies that, by the very nature of the error-learning mechanism, such deviations cannot persist unless inflation is continually accelerated so that it always stays ahead of expected inflation.[3] If inflation is not accelerated, but instead stays constant, then the gap between actual and expected inflation will eventually be closed. Therefore acceleration is required to keep the gap open if unemployment is to be maintained below its natural equilibrium level. In other words, the long-run trade-off implied by the accelerationist hypothesis is between unemployment and the *rate of acceleration* of the inflation rate, in contrast to the conventional trade-off between unemployment and the inflation rate itself as implied by the original Phillips curve.[4]

Policy Implications of the Natural Rate and Accelerationist Hypotheses

At least two policy implications stemmed from the natural rate and accelerationist propositions. First, the authorities could either peg unemployment or stabilize the rate of inflation but not both. If they pegged unemployment, they would lose control of the rate of inflation because the latter accelerates when unemployment is held below its natural level. Alternatively, if they stabilized the inflation rate,

Figure 8

THE ACCELERATIONIST HYPOTHESIS



Since the adjustment of expected to actual inflation works to restore unemployment to its natural equilibrium level U_N at any steady rate of inflation, the authorities must continually raise (accelerate) the inflation rate if they wish to peg unemployment at some arbitrarily low level such as U_1. Such acceleration, by generating a continuous succession of inflation surprises, perpetually frustrates the full adjustment of expectations that would return unemployment to its natural rate. Thus attempts to peg unemployment at U_1 will provoke explosive, ever-accelerating inflation. The economy will travel the path **ABCD** with the rate of inflation rising from zero to p_1 to p_2 to p_3 etc.

[3] Taking the time derivative of equation 3, then assuming that the deviation of U from U_N is pegged at a constant level by the authorities such that its rate of change is zero, and then substituting equation 4 into the resulting expression yields

$$\dot{p} = b(p - p^e)$$

which says that the inflation rate must accelerate to stay ahead of expected inflation.

[4] The proof is simple. Merely substitute equation 3 into the expression presented in the preceding footnote to obtain

$$\dot{p} = ba(U_N - U)$$

which says that the trade-off is between the rate of acceleration of inflation $\dot{p}$ and unemployment U relative to its natural rate.

they would lose control of unemployment since the latter returns to its natural level at any steady rate of inflation. Thus, contrary to the original Phillips hypothesis, they could not peg unemployment at a given constant rate of inflation. They could, however, choose the steady-state inflation rate at which unemployment returns to its natural level.

A second policy implication stemming from the natural rate hypothesis was that the authorities could choose from among alternative transitional adjustment paths to the desired steady-state rate of inflation. Suppose the authorities wished to move from a high inherited inflation rate to a zero or other low target inflation rate. To do so, they must lower inflationary expectations, a major determinant of the inflation rate. But equations 3 and 4 state that the only way to lower expectations is to create slack capacity or excess supply in the economy. Such slack raises unemployment above its natural level and thereby causes the actual rate of inflation to fall below the expected rate so as to induce a downward revision of the latter.[5] The equations also indicate that how fast inflation comes down depends on the amount of slack created.[6] Much slack means fast adjustment and a relatively rapid attainment of the inflation target. Conversely, little slack means sluggish adjustment and a relatively slow attainment of the inflation target. Thus the policy choice is between adjustment paths offering high excess unemployment for a short time or lower excess unemployment for a long time (see Figure 9).[7]

[5] The proof is straightforward. Simply substitute equation 3 into equation 4 to obtain

$$\dot{p}^e = ba(U_N - U).$$

This expression says that expectations will be adjusted downward ($\dot{p}^e$ will be negative) only if unemployment exceeds its natural rate.

[6] Note that the equation developed in footnote 4 states that disinflation will occur at a faster pace the larger the unemployment gap.

[7] Controls advocates proposed a third policy choice: use wage-price controls to hold actual below expected inflation so as to force a swift reduction of the latter. Overlooked was the fact that controls would have little impact on expectations unless the public was convinced that the trend of prices when controls were in force was a reliable indicator of the future price trend after controls were lifted. Convincing the public would be difficult if controls had failed to stop inflation in the past. Aside from this, it is hard to see why controls should have a stronger impact on expectations than a preannounced, demonstrated policy of disinflationary money growth.

IV.

STATISTICAL TESTS OF THE NATURAL RATE HYPOTHESIS

The preceding has examined the third stage of Phillips curve analysis in which the natural rate hypothesis was formed. The fourth stage involved statistical testing of that hypothesis. These tests, conducted in the early- to mid-1970s, led to criticisms of the adaptive-expectations or error-learning model of inflationary expectations and thus helped prepare the way for the introduction of the alternative rational expectations idea into Phillips curve analysis.

The tests themselves were mainly concerned with estimating the numerical value of the coefficient on the price-expectations variable in the expectations-augmented Phillips curve equation. If the coefficient is one, as in equation 3, then the natural rate hypothesis is valid and no long-run inflation-unemployment trade-off exists for the policymakers to exploit. But if the coefficient is less than one, the natural rate hypothesis is refuted and a long-run trade-off exists. Analysts emphasized this fact by writing the expectations-augmented equation as

$$(6)\quad p = a(U_N - U) + \phi p^e$$

where ϕ is the coefficient (with a value of between zero and one) attached to the price expectations variable. In long-run equilibrium, of course, expected inflation equals actual inflation, i.e., $p^e = p$. Setting expected inflation equal to actual inflation as required for long-run equilibrium and solving for the actual rate of inflation yields

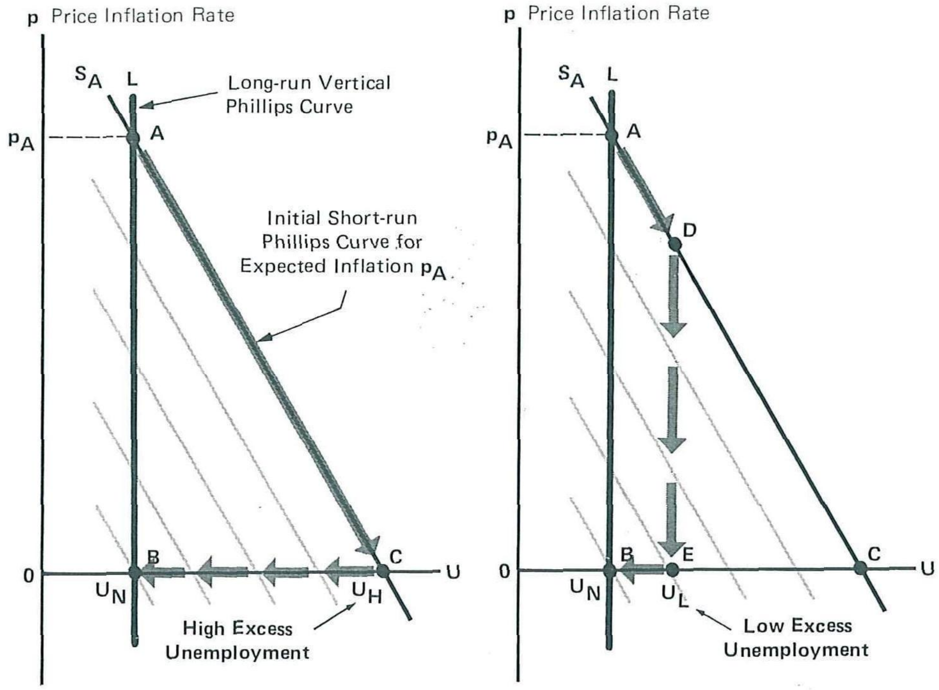
$$(7)\quad p = \frac{a}{1-\phi}(U_N - U).$$

Besides showing that the long-run Phillips curve is steeper than its short-run counterpart (since the slope parameter of the former, $a/(1-\phi)$, exceeds that of the latter, a), equation 7 shows that a long-run trade-off exists only if the expectations coefficient ϕ is less than one. If the coefficient is one, however, the slope term is infinite, which means that there is no relation between inflation and unemployment so that the trade-off vanishes (see Figure 10).

Many of the empirical tests estimated the coefficient to be less than unity and concluded that the natural rate hypothesis was invalid. But this conclusion was sharply challenged by economists who contended that the tests contained statistical bias that

Figure 9

ALTERNATIVE DISINFLATION PATHS



ACB = Fast disinflation path involving high excess unemployment for a short time.

ADEB = Gradualist disinflation path involving low excess unemployment for a long time.

To move from high-inflation point **A** to zero-inflation point **B** the authorities must first travel along short-run Phillips curve S_A, lowering actual relative to expected inflation and thereby inducing the downward revision of expectations that shifts the short-run curve leftward until point **B** is reached. Since the speed of adjustment of expectations depends upon the size of the unemployment gap, it follows that point **B** will be reached faster via the high excess unemployment path **ACB** than via the low excess unemployment path **ADEB**. The choice is between high excess unemployment for a short time or low excess unemployment for a long time.

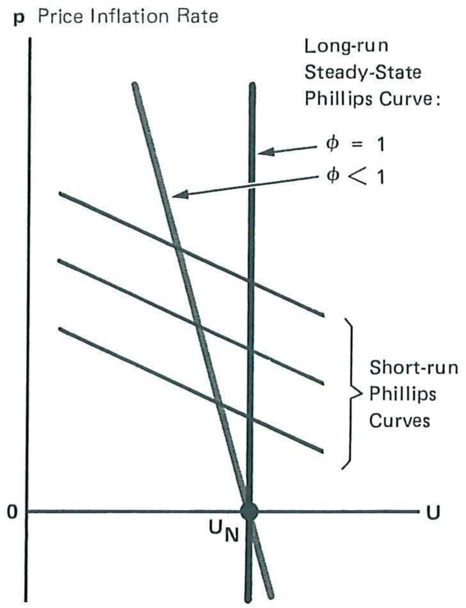
tended to work against the natural rate hypothesis. These critics pointed out that the tests typically used adaptive-expectations schemes as empirical proxies for the unobservable price expectations variable. They further showed that if these proxies were inappropriate measures of inflationary expectations then estimates of the expectations coefficient could well be biased downward. If so, then estimated coefficients of less than one constituted no disproof of the natural rate hypothesis. Rather they constituted evidence of inadequate measures of expectations.

Shortcomings of the Adaptive-Expectations Assumption

In connection with the foregoing, the critics argued that the adaptive-expectations scheme is a grossly inaccurate representation of how people formulate price expectations. They pointed out that it postulates naive expectational behavior, holding as it does that people form anticipations solely from a weighted average of past price experience with weights that are fixed and independent of economic conditions and

Figure 10

THE EXPECTATIONS COEFFICIENT AND THE LONG-RUN STEADY-STATE PHILLIPS CURVE



Statistical tests of the natural rate hypothesis sought to determine the magnitude of the expectations coefficient ϕ in the long-run steady-state Phillips curve equation

$$p = \frac{a}{1-\phi}(U_N - U).$$

A coefficient of one means that no permanent trade-off exists and the steady-state Phillips curve is a vertical line through the natural rate of unemployment. Conversely, a coefficient of less than one signifies the existence of a long-run Phillips curve trade-off with negative slope for the policymakers to exploit. Note that the long-run curves are steeper than the short-run ones, indicating that permanent trade-offs are less favorable than temporary ones.

policy actions. It implies that people look only at past price changes and ignore all other pertinent information—e.g., money growth rate changes, exchange rate movements, announced policy intentions and the like—that could be used to reduce expectational errors. That people would fail to exploit information that would improve expectational accuracy seems implausible, however. In short, the critics contended that adaptive expectations are not wholly rational if other information besides past price changes can improve inflation predictions.

Many economists have since pointed out that it is hard to accept the notion that individuals would continually form price anticipations from *any* scheme that is inconsistent with the way inflation is actually generated in the economy. Being different from the true inflation-generating mechanism, such schemes will produce expectations that are systematically wrong. If so, rational forecasters will cease to use them. For example, suppose inflation were actually accelerating or decelerating. According to equation 5, the adaptive-expectations model would systematically underestimate the inflation rate in the former case and overestimate it in the latter. Using a unit weighted average of past inflation rates to forecast a steadily rising or falling rate would yield a succession of one-way errors. The discrepancy between actual and expected inflation would persist in a perfectly predictable way such that forecasters would be provided free the information needed to correct their mistakes. Perceiving these persistent expectational mistakes, rational individuals would quickly abandon the error-learning model for more accurate expectations-generating schemes. Once again, the adaptive-expectations mechanism is implausible because of its incompatibility with rational behavior.

V.

FROM ADAPTIVE EXPECTATIONS TO RATIONAL EXPECTATIONS

The shortcomings of the adaptive-expectations approach to the modeling of expectations led to the incorporation of the alternative rational expectations approach into Phillips curve analysis. According to the rational expectations hypothesis, individuals will tend to exploit *all* available pertinent information about the inflationary process when making their price forecasts. If true, this means that forecasting errors ultimately could arise only from random (unforeseen) shocks occurring to the economy. At first, of course, price forecasting errors might also arise because individuals initially possess limited or incomplete information about, say, an unprecedented new policy regime, economic structure, or inflation-generating mechanism. But it is unlikely that this condition would persist. For if the public were truly rational, it would quickly learn from these inflationary surprises or prediction errors (data on which it acquires costlessly as a side condition of buying goods) and incorporate the free new information into its forecasting procedures, i.e., the source of forecasting mistakes would be swiftly perceived and systematically eradicated. As knowledge of policy and the inflationary process improved, forecasting models would be continually revised to produce more accurate predictions. Soon all systematic (predictable) elements influencing the rate of inflation would become known and fully understood, and individuals' price expectations would constitute the most accurate (unbiased) forecast consistent with that knowledge.[8] When this happened the economy would converge to its rational expectations equilibrium and people's price expectations would be the same as those implied by the actual inflation-generating mechanism. As incorporated in natural rate Phillips curve models, the rational expectations hypothesis implies that thereafter, except for unavoidable surprises due to purely random shocks, price expectations would always be correct and the economy would always be at its long-run steady-state equilibrium.

Policy Implications of Rational Expectations

The strict (flexible price, instantaneous market clearing) rational expectations approach has radical policy implications. When incorporated into natural rate Phillips curve equations, it implies that systematic policies—i.e., those based on feedback control rules defining the authorities' response to changes in the economy—cannot influence real variables such as output and unemployment even in the short run, since people would have already anticipated what the policies are going to be and acted upon those anticipations. To have an impact on output and employment, the authorities must be able to create a divergence between actual and expected inflation. This follows from the proposition that inflation influences real variables only when it is unanticipated. To lower unemployment in the Phillips curve equation $p - p^e = a(U_N - U)$, the authorities must be able to alter the actual rate of inflation without simultaneously causing an identical change in the expected future rate. This may be impossible if the public can predict policy actions.

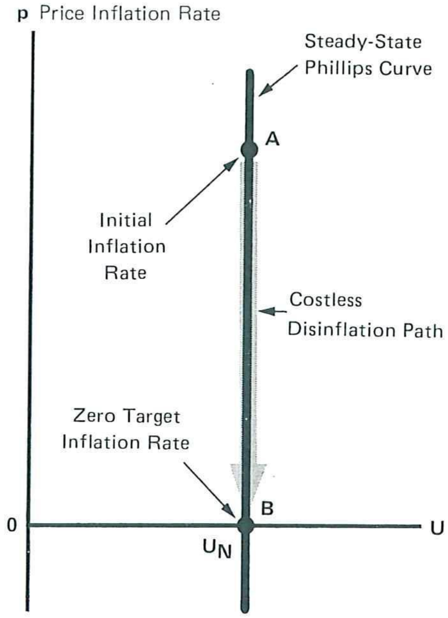
[8] Put differently, rationality implies that current expectational errors are uncorrelated with past errors and with all other known information, such correlations already having been perceived and exploited in the process of improving price forecasts.

Policy actions, to the extent they are systematic, are predictable. Systematic policies are simply feedback rules or response functions relating policy variables to past values of other economic variables. These policy response functions can be estimated and incorporated into forecasters' price predictions. In other words, rational individuals can use past observations on the behavior of the authorities to discover the policy rule. Once they know the rule, they can use current observations on the variables to which the policymakers respond to predict future policy moves. Then, on the basis of these predictions, they can correct for the effect of anticipated policies beforehand by making appropriate adjustments to nominal wages and prices. Consequently, when stabilization actions do occur, they will have no impact on real variables like unemployment since they will have been discounted and neutralized in advance. In short, rules-based policies, being in the information set used by rational forecasters, will be perfectly anticipated and for that reason will have no impact on unemployment. The only conceivable way that policy can have even a short-run influence on real variables is for it to be unexpected, i.e., the policymakers must either act in an unpredictable random fashion or secretly change the policy rule. Apart from such tactics, which are incompatible with most notions of the proper conduct of public policy, there is no way the authorities can influence real variables, i.e., cause them to deviate from their natural equilibrium levels. The authorities can, however, influence a nominal variable, namely the inflation rate, and should concentrate their efforts on doing so if some particular rate (e.g., zero) is desired.

As for disinflation strategy, the rational expectations approach generally calls for a preannounced sharp swift reduction in money growth—provided of course that the government's commitment to ending inflation is sufficiently credible to be believed. Having chosen a zero target rate of inflation and having convinced the public of their determination to achieve it, the policy authorities should be able to do so without creating a costly transitional rise in unemployment. For, given that rational expectations adjust infinitely faster than adaptive expectations to a credible preannounced disinflationary policy (and also that wages and prices adjust to clear markets continuously) the transition to price stability should be relatively quick and painless (see Figure 11).

Figure 11

COSTLESS DISINFLATION UNDER RATIONAL EXPECTATIONS AND POLICY CREDIBILITY



Assuming expectational rationality, wage/price flexibility, and full policy credibility, a preannounced permanent reduction in money growth to a level consistent with stable prices theoretically lowers expected and thus actual inflation to zero with no accompanying transitory rise in unemployment. The economy moves immediately from point **A** to point **B** on the vertical steady-state Phillips curve. Here is the basic prediction of the rational expectations—natural rate model: that fully anticipated policy changes (including credible preannounced ones) affect only inflation but not output and employment.

No Exploitable Trade-Offs

To summarize, the rationality hypothesis, in conjunction with the natural rate hypothesis, denies the existence of exploitable Phillips curve trade-offs in the short run as well as the long. In so doing, it differs from the adaptive-expectations version of natural rate Phillips curve models. Under adaptive-expectations, short-run trade-offs exist because such expectations, being backward looking and slow to respond, do not adjust instantaneously to eliminate forecast errors arising from policy-engineered changes in the inflation rate. With expectations adapting to actual inflation with a lag, monetary policy can generate unexpected inflation and consequently influence real variables in the short run. This cannot happen under rational expectations where both actual and expected inflation adjust identically and instantaneously to anticipated policy changes. In short, under rational expectations, systematic policy cannot induce the expectational errors that generate short-run Phillips curves.[9] Phillips curves may exist, to be sure. But they are purely adventitious phenomena that are entirely the result of unpredictable random shocks and cannot be exploited by policies based upon rules.

In sum, no role remains for systematic countercyclical stabilization policy in Phillips curve models embodying rational expectations and the natural rate hypothesis. The only thing such policy can influence in these models is the rate of inflation which adjusts immediately to expected changes in money growth. Since the models teach that the full effect of rules-based policies is on the inflation rate, it follows that the authorities—provided they believe that the models are at all an accurate representation of the way the world works—should concentrate their efforts on controlling that nominal inflation variable since they cannot systematically influence real variables. These propositions are demonstrated with the aid of the expository model presented in the Appendix on page 21.

VI.

EVALUATION OF RATIONAL EXPECTATIONS

The preceding has shown how the rational expectations assumption combines with the natural rate hypothesis to yield the policy-ineffectiveness conclusion that no Phillips curves exist for policy to exploit

---

[9] Note that the rational expectations hypothesis also rules out the accelerationist notion of a stable trade-off between unemployment and the rate of acceleration of the inflation rate. If expectations are formed consistently with the way inflation is actually generated, the authorities will not be able to fool people by accelerating inflation or by accelerating the rate of acceleration, etc. Indeed, no systematic policy will work if expectations are formed consistently with the way inflation is actually generated in the economy.

even in the short run. Given the importance of the rational expectations component in modern Phillips curve analysis, an evaluation of that component is now in order.

One advantage of the rational expectations hypothesis is that it treats expectations formation as a part of optimizing behavior. By so doing, it brings the theory of price anticipations into accord with the rest of economic analysis. The latter assumes that people behave as rational optimizers in the production and purchase of goods, in the choice of jobs, and in the making of investment decisions. For consistency, it should assume the same regarding expectational behavior.

In this sense, the rational expectations theory is superior to rival explanations, all of which imply that expectations may be consistently wrong. It is the only theory that denies that people make systematic expectation errors. Note that it does not claim that people possess perfect foresight or that their expectations are always accurate. What it does claim is that they perceive and eliminate *regularities* in their forecasting mistakes. In this way they discover the actual inflation generating process and use it in forming price expectations. And with the public's rational expectations of inflation being the same as the mean value of the inflation generating process, those expectations cannot be wrong on *average*. Any errors will be random, not systematic. The same cannot be said for other expectations schemes, however. Not being identical to the expected value of the true inflation generating process, those schemes will produce biased expectations that are systematically wrong.

Biased expectations schemes are difficult to justify theoretically. Systematic mistakes are harder to explain than is rational behavior. True, nobody really knows how expectations are actually formed. But a theory that says that forecasters do not continually make the same mistakes seems intuitively more plausible than theories that imply the opposite. Considering the profits to be made from improved forecasts, it seems inconceivable that systematic expectational errors would persist. Somebody would surely notice the errors, correct them, and profit by the corrections. Together, the profit motive and competition would reduce forecasting errors to randomness.

Criticisms of the Rational Expectations Approach

Despite its logic, the rational expectations hypothesis still has many critics. Some still maintain that expectations are basically nonrational, i.e., that most people are too naive or uninformed to formulate unbiased price expectations. Overlooked is the counterargument that relatively uninformed people often delegate the responsibility for formulating rational forecasts to informed specialists and that professional forecasters, either through their ability to sell superior forecasts or to act in behalf of those without same, will ensure that the economy will behave as if all people were rational. One can also note that the rational expectations hypothesis is merely an implication of the uncontroversial assumption of profit (and utility) maximization and that, in any case, economic analysis can hardly proceed without the rationality assumption. Other critics insist, however, that expectational rationality cannot hold during the transition to new policy regimes or other structural changes in the economy since it requires a long time to understand such changes and learn to adjust to them. Against this is the counterargument that such changes and their effects are often foreseeable from the economic and political events that precede them and that people can quickly learn to predict regime changes just as they learn to predict the workings of a given regime. This is especially so when regime changes have occurred in the past. Having experienced such changes, forecasters will be sensitive to their likely future occurrence.

Most of the criticism, however, is directed not at the rationality assumption per se but rather at another key assumption underlying its policy-ineffectiveness result, namely the assumption of no policymaker information or maneuverability advantage over the private sector. This assumption states that private forecasters possess exactly the same information and the ability to act upon it as do the authorities. Critics hold that this assumption is implausible and that if it is violated then the policy ineffectiveness result ceases to hold. In this case, an exploitable short-run Phillips curve reemerges, allowing some limited scope for systematic monetary policies to reduce unemployment.

For example, suppose the authorities possess more and better information than the public. Having this information advantage, they can predict and hence respond to events seen as purely random by the public. These policy responses will, since they are unforeseen by the public, affect actual but not expected inflation and thereby change unemployment relative to its natural rate in the (inverted) Phillips curve equation $U_N - U = (1/a)(p - p^e)$.

Alternatively, suppose that both the authorities and the public possess identical information but that the latter group is constrained by long-term contractual obligations from exploiting that information. For example, suppose workers and employers make labor contracts that fix nominal wages for a longer period of time than the authorities require to change the money stock. With nominal wages fixed and prices responding to money, the authorities are in a position to lower real wages and thereby stimulate employment with an inflationary monetary policy.

In these ways, contractual and informational constraints are alleged to create output- and employment-stimulating opportunities for systematic stabilization policies. Indeed, critics have tried to demonstrate as much by incorporating such constraints into rational expectations Phillips curve models similar to the one outlined in the Appendix of this article.

Proponents of the rational expectations approach, however, doubt that such constraints can restore the potency of activist policies and generate exploitable Phillips curves. They contend that policymaker information advantages cannot long exist when government statistics are published immediately upon collection, when people have wide access to data through the news media and private data services, and when even secret policy changes can be predicted from preceding observable (and obvious) economic and political pressures. Likewise, they note that fixed contracts permit monetary policy to have real effects only if those effects are so inconsequential as to provide no incentive to renegotiate existing contracts or to change the optimal type of contract that is negotiated. And even then, they note, such monetary changes become ineffective when the contracts expire. More precisely, they question the whole idea of fixed contracts that underlies the sticky wage case for policy activism. They point out that contract duration is not invariant to the type of policy being pursued but rather varies with it and thus provides a weak basis for activist fine-tuning.

Finally, they insist that such policies, even if effective, are inappropriate. In their view, the proper role for policy is not to exploit informational and contractual constraints to systematically influence real activity but rather to neutralize the constraints or to minimize the costs of adhering to them. Thus if people form biased price forecasts, then the policymakers should publish unbiased forecasts. And if the policy authorities have informational advantages over private individuals, they should make that information public rather than attempting to exploit the advantage. That is, if information is costly to collect and process, then the central authority should gather it and make it freely available. Finally, if contractual wages and prices are sticky and costly to adjust, then the authorities should minimize these price adjustment costs by following policies that stabilize the general price level.

In short, advocates of the rational expectations approach argue that feasibility alone constitutes insufficient justification for activist policies. Policies should also be socially *beneficial*. Activist policies hardly satisfy this latter criterion since their effectiveness is based on deceiving people into making expectational errors. The proper role for policy is not to influence real activity via deception but rather to reduce information deficiencies, to eliminate erratic variations of the variables under the policymakers' control, and perhaps also to minimize the costs of adjusting prices.

VII.

CONCLUDING COMMENTS

The preceding paragraphs have traced the evolution of Phillips curve analysis. The chief conclusions can be stated succinctly. The Phillips curve concept has changed radically over the past 25 years as the notion of a stable enduring trade-off has given way to the policy-ineffectiveness view that no such trade-off exists for the policymakers to exploit. Instrumental to this change were the natural rate and rational expectations hypotheses, respectively. The former says that trade-offs arise solely from expectational errors while the latter holds that systematic macroeconomic stabilization policies, by virtue of their very predictability, cannot possibly generate such errors. Taken together, the two hypotheses imply that systematic demand management policies are incapable of influencing real activity, contrary to the predictions of the original Phillips curve analysis.

On the positive side, the two hypotheses do imply that the government can contribute to economic stability by following policies to minimize the expectational errors that cause output and employment to deviate from their normal full-capacity levels. For example, the authorities could stabilize the price level so as to eliminate the surprise inflation that generates confusion between absolute and relative prices and that leads to perception errors. Similarly, they could direct their efforts at minimizing random and erratic variations in the monetary variables under their control. In so doing, not only would they lessen the number of forecasting mistakes that induce deviations from output's natural rate, they would reduce policy uncertainty as well.

Besides the above, the natural rate-rational expectations school also notes that microeconomic structural policies can be used to achieve what macro demand policies cannot, namely a permanent reduction in the unemployment rate. For, by improving the efficiency and performance of labor and product markets, such micro policies can lower the natural rate of unemployment and shift the vertical Phillips curve to the left. A similar argument was advanced in the early 1960s by those who advocated structural policies to shift the Phillips curve. It is on this point, therefore, that one should look for agreement between those who still affirm and those who deny the existence of exploitable inflation-unemployment trade-offs.

APPENDIX

A SIMPLE ILLUSTRATIVE MODEL

The policy ineffectiveness proposition discussed in Section V of the text can be clarified with the aid of a simple illustrative model. The model consists of four components, namely an (inverted) expectations-augmented Phillips curve

(1) $U_N - U = (1/a)(p - p^e)$,

a monetarist inflation-generating mechanism

(2) $p = m + \epsilon$,

a policy reaction function or feedback control rule

(3) $m = c(U_{-1} - U_T) - d(p_{-1} - p_T) + \mu$,

and a definition of rational inflation expectations

(4) $p^e = E[p|I]$.

Here U and U_N are the actual and natural rates of unemployment, p and p^e the actual and expected rates of inflation, m the rate of nominal monetary growth per unit of real money demand (the latter assumed to be a fixed constant except for transitory disturbances), ϵ and μ are random error terms with mean values of zero, E is the expectations operator, I denotes all information available when expectations are formed, and the subscripts T and -1 denote target and previous period values of the attached variables.

Of these four equations, the first expresses a trade-off between unemployment (relative to its natural level) and surprise (unexpected) inflation.[1] Equation 2 expresses the rate of inflation p as the sum of the growth rate of (demand adjusted) money m and a random shock variable ϵ having a mean (expected) value of zero. In essence, this equation says that inflation is generated by excess money growth and transitory disturbances unrelated to money growth. Equation 3 says that the policy authorities set the current rate of monetary growth in an effort to correct last period's deviations of the unemployment and inflation rates from their predetermined target levels, U_T and p_T. Also, since money growth cannot be controlled perfectly by the feedback rule, the slippage is denoted by the random variable μ with a mean of zero that causes money growth to deviate unpredictably from the path intended by the authorities. Note that the disturbance term μ can also represent deliberate monetary surprises engineered by the policy authorities. Finally, the last equation defines anticipated inflation p^e as the mathematical expectation of the actual inflation rate conditional on all information available when the expectation is formed. Included in the set of available information are the inflation-generating mechanism, the policy reaction function, and the values of all past and predetermined variables in the model.

To derive the policy ineffectiveness result, first calculate mathematical expectations of equations 2 and 3. Remembering that the expected values of the random terms in those equations are zero, this step yields the expressions

[1] There exists a current dispute over the proper interpretation of the Phillips curve equation 1. The rational expectations literature interprets it as an aggregate supply function stating that firms produce the normal capacity level of output when actual and expected inflation are equal but produce in excess of that level (thus pushing U below U_N) when fooled by unexpected inflation. This view holds that firms mistake unanticipated general price increases for rises in the particular (relative) prices of their own products. Surprised by inflation, they treat the price increase as special to themselves and so expand output. An alternative interpretation views the equation as a price-setting relation according to which businessmen, desiring to maintain their constant-market-share relative prices, raise their prices at the rate at which they expect other businessmen to be raising theirs and then adjust that rate upward if demand pressure appears. Either interpretation yields the same result: expectational errors cause output and unemployment to deviate from their natural levels. The deviations disappear when the errors vanish.

$$(5) \quad p^e = m^e \text{ and}$$

$$(6) \quad m^e = c(U_{-1}-U_T)-d(p_{-1}-p_T)$$

which state that, under rational expectations and systematic feedback policy rules, the anticipated future rate of inflation equals the expected rate of monetary growth which in turn is given by the deterministic (known) component of the monetary policy rule. The last step is to substitute equations 2, 3, 5, and 6 into equation 1 to obtain the reduced form expression

$$(7) \quad U_N-U = (1/a)(\epsilon+\mu)$$

which states that deviations of unemployment from its natural rate result solely from inflation surprises caused by random shocks.

To see the policy ineffectiveness result, note that only the unsystematic or unexpected random component of monetary policy, $m-m^e=\mu$, enters the the reduced form equation.[2] The systematic component is absent. This means that systematic (rules-based) monetary policies cannot affect the unemployment rate. Only unexpected money growth matters. No Phillips curve trade-offs exist for systematic policy to exploit.[3]

To summarize, the strict (flexible price, continuous market clearing) rational expectations-natural rate model depicted here implies that expectational errors are the only source of departure from steady-state equilibrium, that such errors are random, short-lived, and immune to systematic policy manipulation, and therefore that rules-based policies can have no impact on real variables like unemployment since those policies will be fully foreseen and allowed for in wage/price adjustments. Thus, except for unpredictable random shocks, steady-state equilibrium prevails and systematic monetary changes produce no surprises, no disappointed expectations, no transitory impacts on real economic variables. In short, Phillips curves are totally adventitious phenomena generated by unforeseeable random shocks and as such cannot be exploited by systematic policy even in the short run.

---

2 Note that both the monetary-surprise equation $m-m^e=\mu$ and the price-surprise equation $p-p^e=\epsilon$ embody the famous **orthogonality** property according to which forecast errors $m-m^e$ and $p-p^e$ are independent of (orthogonal to) all information available when the forecast is made. In particular, the forecast errors are independent of the past and predetermined values of all variables and of the systematic components of the policy rule and inflation-generating mechanism. This is as it should be. For if the errors were not independent of the foregoing variables, then information is not being fully exploited and expectations are not rational.

3 Of course **random** policy could affect output. That is, the authorities could influence real activity by manipulating the disturbance term μ in the policy reaction function in a haphazard unpredictable way. Randomness, however, is not a proper basis for public policy.

Federal Reserve Bank of Richmond
Post Office Box 27622, Richmond, Virginia 23261