THE EARLY HISTORY OF THE PHILLIPS CURVE

Thomas M. Humphrey

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 relevance, 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 \( f(\cdot) \) 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 \( p - p' = f(U) \), where \( p' \) 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.

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 \( U = g(p-p') \) 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 com-
bined 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 prototypical 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

$$U = h(P - P^e)$$

and

$$P - P^e = k \frac{dP}{dt}$$

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 increasing 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 . . . . 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, in-
justice, 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 increased 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 employ-

ment, 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\textsuperscript{E}) where U is the discrepancy between unemployment and its natural steady-state level, P is the price level, and P\textsuperscript{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 unidirectionally 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 link-age, causality, he argued, runs from inflation to unemployment as confirmed by his finding that the former variable leads the latter.

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 $dW = F(E, dP_{-1})$ where $dW$ is the change in money wages, $E$ is employment relative to its normal (i.e., trend) level, and the lagged price-change variable $dP_{-1}$ 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 $dW = 0.16E + 0.27dP_{-1}$ 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

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

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 $dW=F(U,dP^{-1})$. 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
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


