A Kernel Regression of Phillips' Data

by
Nancy J. Wulwick and Y.P. Mack*

Working Paper No. 40

November 1990

Submitted to
The Jerome Levy Economics Institute
Bard College

* Nancy J. Wulwick is at the Jerome Levy Economics Institute of Bard College. Y.P. Mack is in the Statistics Division of the University of California at Davis. The authors acknowledge R.L. Basmann's helpful comments.

Abstract

Economists have assumed that the Phillips curve, which shows a positive (negative) relation between inflation and the output ratio (unemployment rate), may be mapped off the aggregate demand-aggregate supply apparatus. The paper shows that the Phillips curve requires that unlikely restrictions be put on the form of the aggregate supply and aggregate demand curves. In this case, it is inappropriate to treat data on inflation and capacity utililization as the basis for estimating an underlying formal model. The paper therefore uses a nonparametric, data-driven method to describe the data. This method, of kernel regression, shows the inflation-unemployment association in Phillips's sample to be negative on a global scale, yet irregular within particular ranges of unemployment.

I. Introduction

Given the importance of the identification problem in econometrics, it is surprising that few economists have attempted to identify the model underlying the Phillips curve. This curve shows a positive relation between the proportionate rate of change of the price level and the output ratio, or (assuming markup pricing and Okun's law) a negative relation between the proportionate rate of change of the money wage rate and the unemployment rate. The curve, presented by Phillips in the mid 1950s, has played a prominent role in the development of macroeconomic theory and policy analysis.[1]

In the 1960s, when Keynesian economics was in high repute, economists thought that the Phillips curve represented a long run relation between inflation and capacity utilitization. On this basis, economists advised that the management of aggregate demand involved a trade-off between inflation and unemployment. However, orthodox economists saw that such policy advice conflicted with traditional economic teaching, which said that the market tended to eliminate excess supplies of labor. In the 1970s, when the US economy experienced rapid inflation <u>and</u> high unemployment, the long run Phillips curve fell into disrepute. It became the standard view that the economy in the long run maintained an output level consistent with a "natural" rate of unemployment. The Phillips

---

[1]. Wulwick (1987) analyzes the historical development of the Phillips curve.

curve was only a short run phenomenon that arose during a period of adjustment following a shock to aggregate demand. Because a rate of unemployment below the "natural" rate could be maintained only at the cost of increasing inflation, the Phillips curve could not offer an exploitable trade-off. Econometric studies have failed to settle the controversies between economists over the existence and the duration of the Phillips curve phenomenon.

The Phillips curve has remained an important notion in policy-making and teaching at the undergraduate and graduate levels. It always has been a "selling point" of the Phillips curve that it can be explained in terms of aggregate supply and aggregate demand. This paper argues that the Phillips curve requires that *ad hoc* and unlikely restrictions be put on the form of the aggregate supply and aggregate demand curves. If this is the case, then it is inappropriate to treat data on inflation and capacity utililization as the basis for estimating an underlying theoretical model.

Economists require formal, econometric models only in order to estimate relations explained by economic theory. From this perspective, it appears to be unnecessary to impose a formal model on the the data for inflation and capacity utilization. Instead the paper uses kernel regression, which is a data-driven technique, to observe Phillips's data. The inflation-unemployment association appears to be negative on a global scale, yet irregular within particular ranges of unemployment.

Most economists have used only parametric methods of regression analysis.[2] Economists estimate an algebraic model, after having combined all the observations of the dependent and independent variables. The t-test then measures the statistical significance of the estimates, on the assumption that the residual deviations from the model belong to a normal distribution Kernel regression offers a different approach. It computes the value of an *unspecified* function at each of its argument points. The technique does not require any assumptions about the distribution from which the sampled data was drawn. Indeed, kernel regression merely is a data-driven method of curve-fitting. Unlike the curve-fitting in econometrics, kernel regression has a firm basis in mathematical statistics.

Kernel regression was introduced by Rosenblatt in the United States in the latter 1950s. His proposal seems to have been motivated by the smoothing of time-series data in the frequency domain.[3] At this time, Phillips, who was a time series expert, was in London working on his curve. In some respects, Phillips's (1958) article on the curve developed a statistical method that is similar to kernel regression.[4] The next section compares Phillips's method, with which economists are familiar, to kernel

---

[2]. On the use of nonparametric regression in economics, see Ullah, ed. (1989) and Stock (1989).

[3]. Rosenblatt (1956).

[4]. Wulwick (1989) discusses Phillips's statistical method.

Figure 1

Phillips's Scattergraph, 1861-1913.

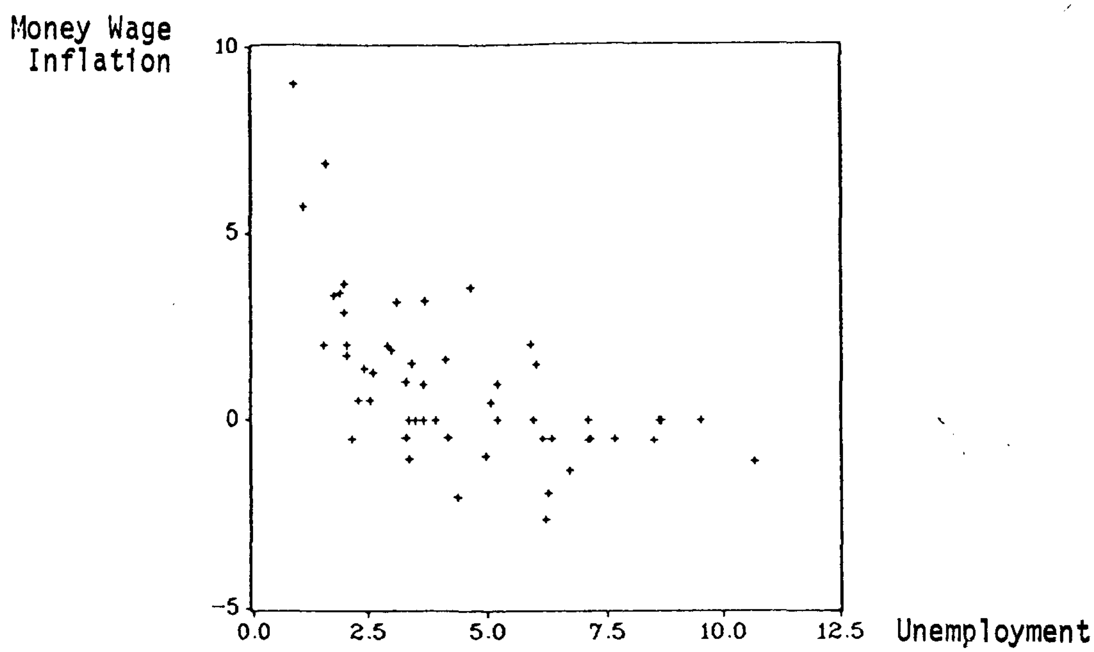
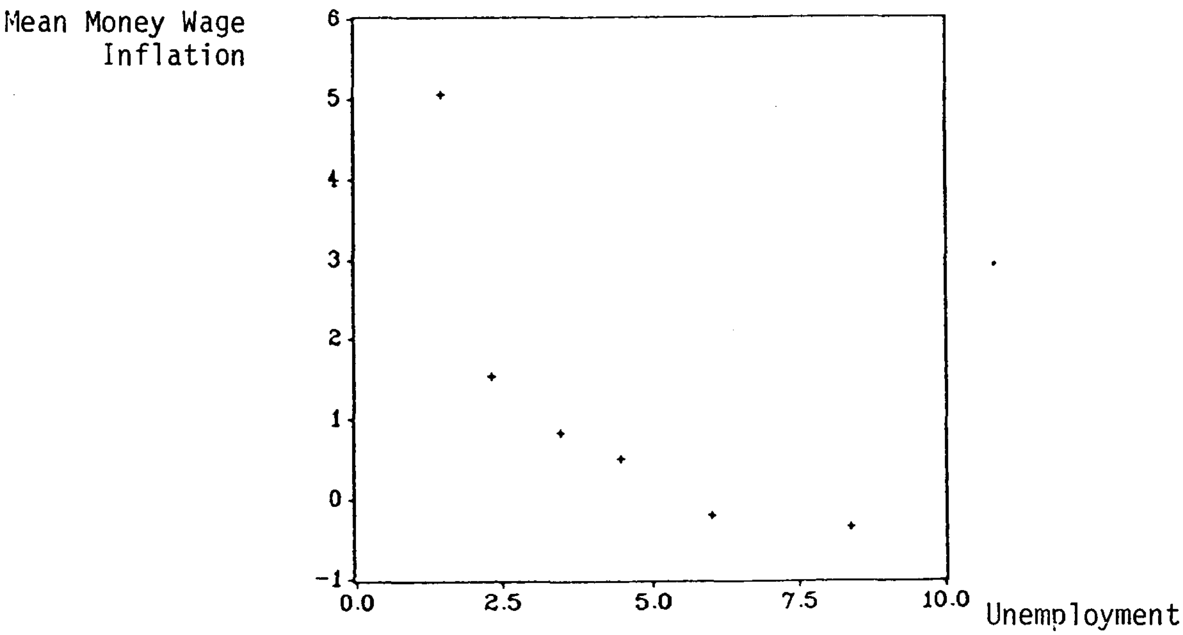


Figure 2

Phillips's Graph of Averages



Mean Coordinates: 1.52, 5.06; 2.35, 1.55; 3.48, 0.85; 4.49, 0.52;
6, -0.2; 8.4, -0.35

regression and carries out a kernel regression on Phillips's data. We then present the economic argument for using kernel regression to describe the Phillips curve.

II. Phillips's Curve and Kernel Regression

A. Phillips's Method

Phillips presented a scattergraph of annual money wage inflation and the annual unemployment rate for Great Britain 1861-1913, as shown in Figure 1. (The data appear in Appendix A.) His previous theoretical work suggested to him that the inflation-unemployment association was negative and markedly nonlinear (Phillips (1954)). Thus Phillips specified a hyperbolic shape to model the scattergraph. Since nonlinear least squares was unavailable in the late 1950s, Phillips exploited what was then a conventional method of curve-fitting. He divided the x-axis of the scattergraph into six intervals and represented (i.) the unemployment variable by the arithmetic mean of the unemployment observations within each interval and (ii.) money wage inflation by the average value of observations of inflation within each interval. The procedure, as Figure 2 shows, yielded a graph of averages relating mean inflation to unemployment which formed a regular, hyperbolic shape.

The regularity of the graph of averages depended on Phillips's choice of intervals, which included six, ten, twelve, five, eleven and nine observations, with varying bandwidths at the unemployment rates 0-2, 2-3, 3-4, 4-5, 5-7 and 7-11. A small change in the number or the width of Phillips's bands produces an irregular graph of averages, with a positive inflation-unemployment relation at some ranges of unemployment.[5]

Like Phillips's curve-fitting method, kernel regression divides the horizontal axis into bandwidths and computes the local average observation of the dependent variables. However, kernel regression provides criteria for the optimal choice of bandwidths.[6]

B. The Kernel Method: A "Moving" Histogram

For constructing a probability density function $p(x)$, the kernel method is analogous to the histogram. To construct a histogram for the given data $\{X_i,\ i=1 \ldots n\}$, the x-axis is partitioned into k ($k = 1 \ldots v$, $v \leq n$) intervals of width b and midpoint m_k. For ease of discussion, we assume the bandwidths are uniform in size. Then,

(1.) $b = (m_k - b/2) + (m_k + b/2)$.

---

5. Wulwick (1989), Figures 4-5.

6. Silverman (1988) presents a introduction to kernel regression.

The number N_k of data points falling inside an interval is counted and then the histogram density, $p_h(x)$, is constructed according to the relative frequency of observations X_i, normalized by the width b of the interval. Thus,

(2.) $p_h(x) = N_k/bn, \qquad k = 1 \ldots v, \quad v \leq n.$

Alternatively, we can construct the histogram density by using a kernel, K, a weighting function which is used in mathematical transforms. The histogram density implicitly assumes a weighting function, K, so that any observation b/2 distance from the midpoint m_k receives a weight of 1 and zero otherwise, that is,

(3.) $$K((m_k-X_i)/b) = \begin{cases} 1, & \text{if } |(m_k-X_i)/b| \leq \tfrac{1}{2}. \\ 0, & \text{otherwise.} \end{cases}$$

The probability density function can be written as

(4.) $p_h(x) = 1/nb\ \Sigma_i^n\ K[(m_k-X_i)/b].$

The histogram density fixes the intervals along the x-axis. In contrast, the kernel method allows the intervals to "move" with the observation X_i and overlap. An interval is formed centered at each observation X_i with variable bandwidths b_k and midpoint x_k. The kernel method then defines the density function as[7]

(5.) $p_k(x) = 1/nb_k\ \Sigma_i^n\ K(z), \qquad z = (x_k-X_i)/b_k,$

$i=1\ldots n,\ k=1\ldots v,\ v<n,$

where

$1, \qquad \text{if } |z| \leq \tfrac{1}{2}.$

---

[7]. If v=n, the histogram will be a series of spikes.

$K(z) =$

$0, \qquad$ otherwise.

The kernel function, K(z), in equation 5 makes each data point within an interval contribute equally to the local average. Other kernel functions employ different weighting schemes. Typically, kernels are symmetric probability densities. The bandwidth, b, controls the amount of local smoothing and satisfies the condition that the width shrink to zero as the same size n grows without bounds -- that is, nb -->0 as n -->∞.

C. Defining Kernel Regression

Consider the problem of analyzing the association between two attributes of the population by using regression analysis. Given the bivariate stochastic variables, X, Y, let the population regression of Y on X be defined by the conditional mean of Y given the value of X,

(6.) $r(x) = E(Y|X=x)$.

The conditional mean is defined by

$$E(Y|X=x) = \int y f(y|x) dy.$$

If

$$p(x) \neq 0,$$

then

$$f(y|x) = g(x,y)/p(x),$$

where $g(x,y)$ is the joint probability density function and $p(x)$ is the marginal probability density function. Thus, the population regression may be defined by

$$(7.) \qquad r(x) = \int y g(x,y)/p(x).$$

Kernel regression, r_k, is analogous to the population regression, $r(x)$, defined by equation 7. The regression based on the data $\{X_i, Y_i, i=1...n\}$ with kernel k and bandwidth b, proceeds as follows: The marginal probability density function $p(x)$ is constructed according to equation (5) and the joint probability density function, $g(x)$ according to

$$(8.) \qquad g_k(x) = 1/nb\ \Sigma_i^n Y_i K(z), \qquad z = (x_k - X_i)/b.$$

The definition of kernel regression then is

$$(9.) \qquad r_k(x) = \Sigma Y_i K(z)/\Sigma K(z).$$

This means that kernel regression, $r_k(x)$, gives a graph of averages that has coordinates with the first term defined by the midpoint of each bandwidth and the second term by the weighted average observation of Y in each bandwidth, where the weights, $w(x)$, are defined by the kernel function. We can write this formally as,

$$(10.) \qquad r_k(x) = \Sigma_i^n Y_i w(x),$$

where the weights are

$$w(x) = K(z)/\Sigma K(z).$$

The sum of the weights equals one,

$$\Sigma_i^n w(x) = 1.$$

When the kernel $K(z)$ is the uniform kernel given by equation (3), r_k is simply the <u>ordinary</u> average of the Y observations with corresponding X values which fall within the distance b/2 of x_k.

There are pros and cons for using kernel regression as opposed to the least squares algorithm to model data. When economists use least squares, they typically assume that the joint density of X and Y is normal. Only in this case will least squares give the maximum likelihood estimate. Economists also assume that the residuals from the least squares estimate are distributed normality, which is a condition of the statistical test of significance. These assumptions of normality appear to be inappropriate in many economic problems.[8] In contrast to least squares, nonparametric regression does not involve assumptions about the distribution of the data. This is the main <u>statistical</u> reason for using the nonparametric approach in economics.[9] However, the use of kernel regression in economics suffers from several technical drawbacks. Economists often study multivariate data. Although statistical theory explains that kernel regression can treat multivariate data, in practice, the statistical studies have been limited to bivariate data to facilitate computer programming. In addition, a "large" sample is necessary to carry out kernel regression, which estimates a data distribution without an estimating model. Statisticians usually lack the information to determine objectively how large is "large" in any particular

---

[8]. For example, Mandlebrot found that stock market prices follow a Cauchy distribution.

[9]. The residuals from Phillips's own nonlinear equation and estimates show only a moderate amount of departure from normality.

case and rely on experience to make their judgements. Certainly, the "large" samples available in biology and engineering, fields that have used kernel regression, are unusual in economics. Taking into account the technical pros and cons of the method, from an economic perspective, kernel regression offers a data-driven approach which is particularly appropriate when economists wish to observe data, rather than use data to estimate a theoretical model.[10]

D. An Approximate Kernel Regression and Phillips's Method Compared.

To give some idea about how the kernel approach would compare with Phillips's procedure, let us compute the kernel regression estimate at the averaged values of unemployment, using the bandwidths, b_k, suggested by Phillips.

We shall use a quadratic kernel, defined as follows[11]

$$(11.)\qquad K(z) = \begin{cases} 0.75(1-z^2), & \text{if } |z| \leq 1, \qquad z = (x_k - X_i)/b_k, \\ 0, & \text{otherwise.} \end{cases}$$

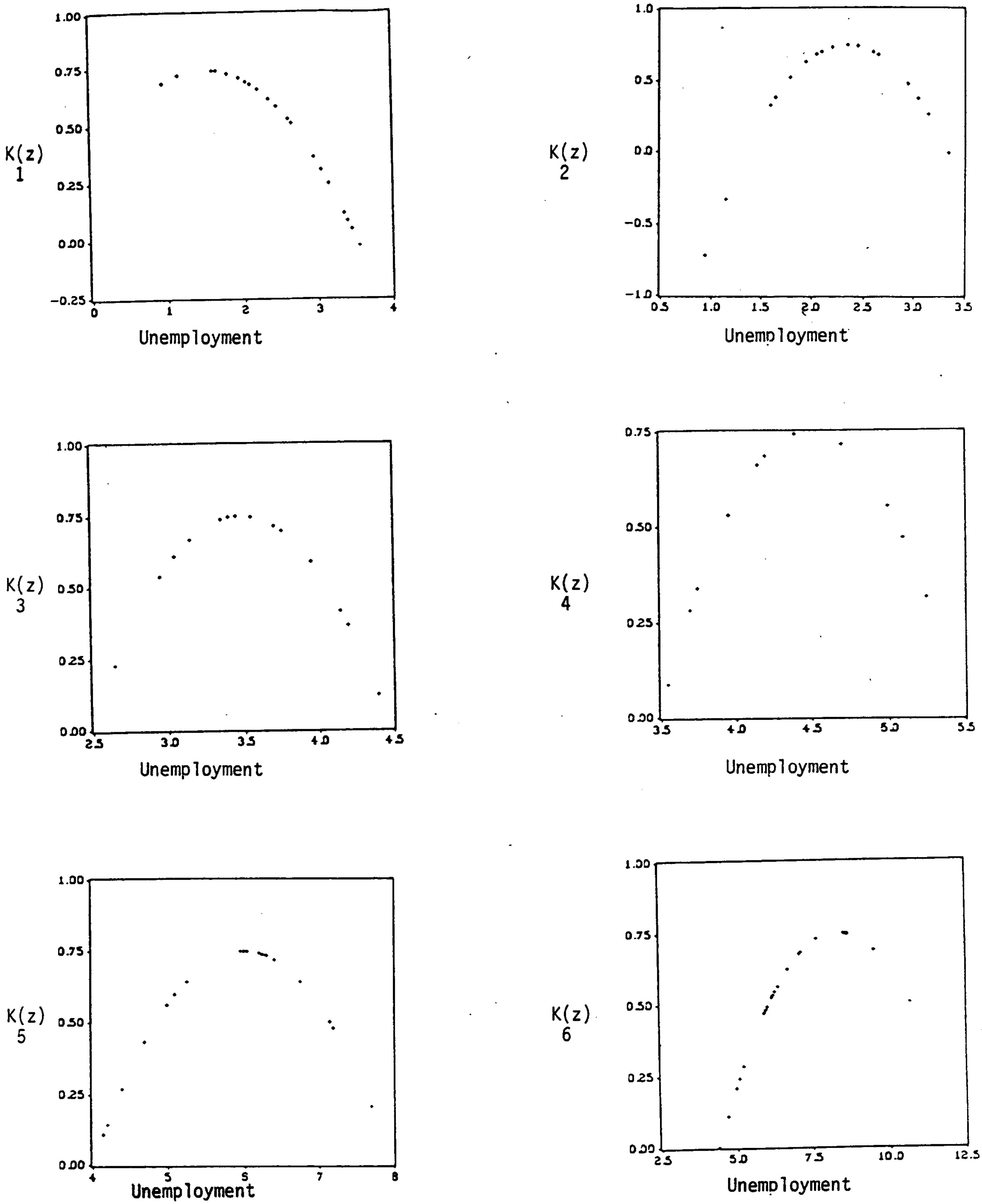
---

[10]. Stock (1989) used kernel regression to evaluate the effect of policy intervention. Kernel regression eliminates the risk of making invalid inferences because of misspecifying a parametric model.

[11]. This kernel is referred to as Epanechnikov's kernel, after the statistician who first applied this mathematical function to kernel regression.

Figure 3

The Quadratic Kernel Functions using Phillip's Six Bandwidths



The quadratic kernel has a unimodal shape that gives more weight to those points with first coordinates, X_i, lying closer to x_k. Given this kernel, the kernel regression, r_k, is given by

$$(12.) \quad r_k(x) = \Sigma Y_i K(z)/\Sigma K(z),$$
$$= \Sigma_i^n Y_i 0.75(1-z_i^2)/\Sigma_i^n 0.75(1-z_i^2),$$
$$= g(x_k) \quad / \quad p(x_k).$$

where and b_k are variable bandwidths. Equation 12 means that the graph of averages will have coordinates with the first term defined by the midpoint of each bandwidth and the second term defined by the weighted average of the Y observations in each bandwidth, where the weights are defined by the quadratic kernel.[12]

Given this kernel, we defined (i.) b_k as six Phillips's bandwidths, (ii.) x_k as the midpoint of Phillips's six bandwidths on the horizontal axis, and (iii.) X_i as the 53 observations of unemployment, sorted by increasing size. Then we computed six series, each of 53 measures of z,

$$z \equiv (x_k - X_i)/b_k, \qquad i = 53,\ k = 6.$$

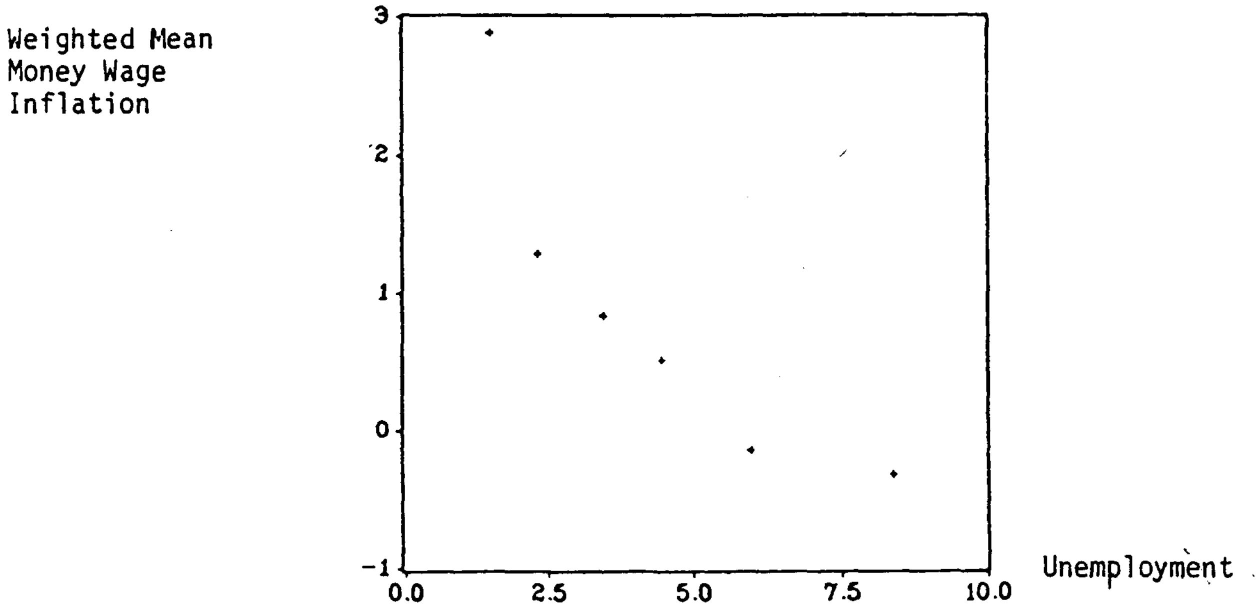
To obtain the kernel weights, we checked which of the 53 values of z in each of the six series equalled or were less than one in absolute terms, according to equation 11. For values of z in that range, we computed the values of the quadratic kernel for each of the six fixed values, x_k, according to equation 11. Figure 3 shows the assignment of weights by the kernel function. It is clear that the six kernels overlap and have a symmetric, unimodal

---

[12]. In contrast, Phillips (1958) did not use weights.

Figure 4

An Approximate Kernel Regression Using Phillips's Bandwidths

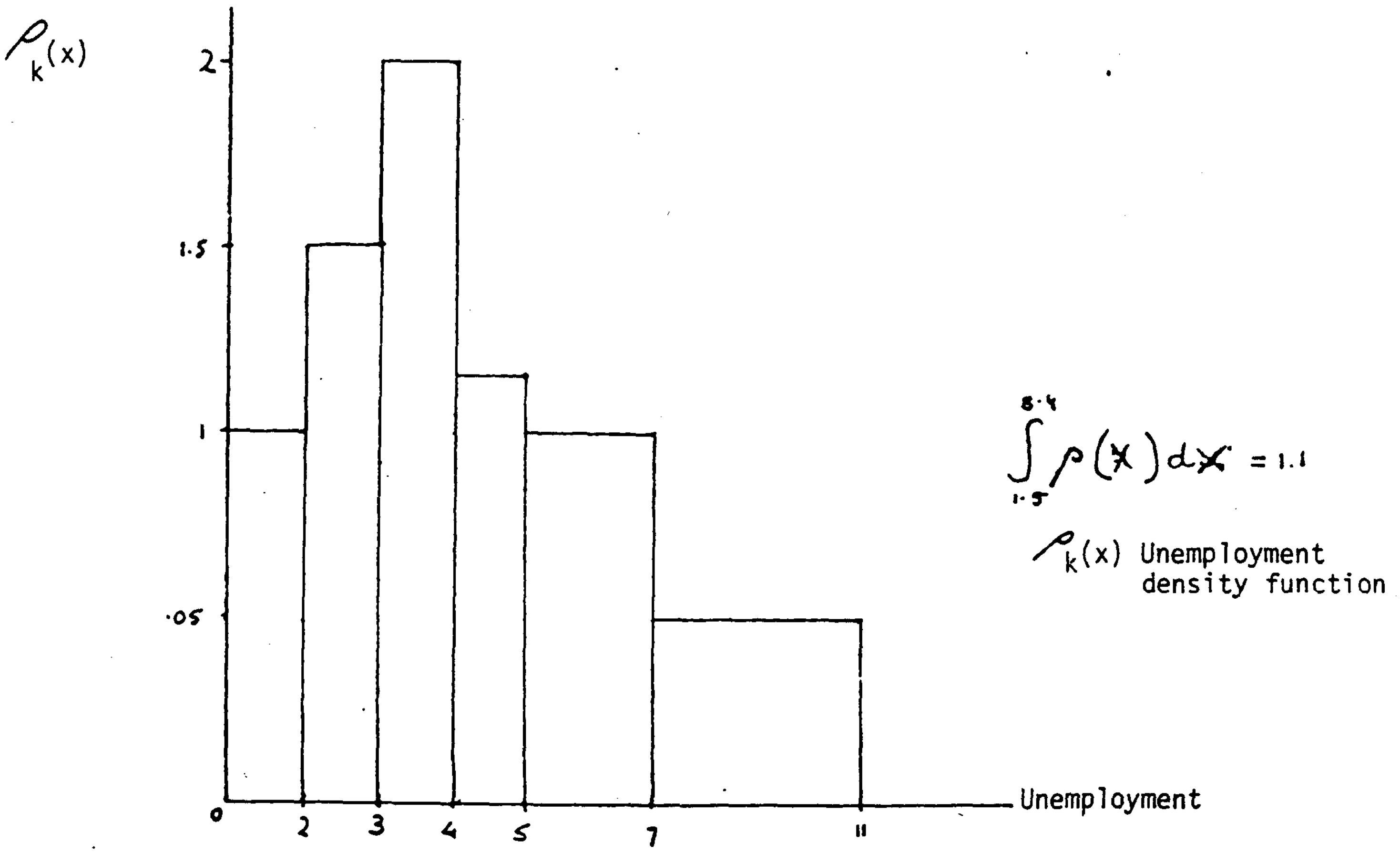


Coordinates: 1.52, 2.87; 2.35, 1.29; 3.48, 0.84; 4.49, 0.52; 6, -0.14; 84, -0.31.

Figure 5

Kernel Estimate of Marginal Probability Density of Unemployment,

Phillips Bandwidths.



shape.

Given the kernel weights, we constructed according to equation 12 the joint density, $g_k(x)$ and the marginal density, $p_k(x)$ and the weighted average of the observations of money wage inflation at the midpoint of each bandwidth, x_k. Figure 4 presents the graph of

averages using the kernel weighting function. Figure 5 constructs the marginal density, $p_k(x)$, of unemployment[13].

Money wage inflation in Figure 4 assumes a lower value when the unemployment rate is very small than in Phillips's graph of averages, shown in Figure 2. This result arose because kernel regression weights each observation of inflation in a way that takes into account the lower inflation values when unemployment is very small in the surrounding region near the left boundary. Otherwise, the two graphs of averages in Figures 2 and 4, which are based on identical bandwidths, show a similar association between mean inflation and mean unemployment.

---

[13]. A Comparison of the Probability of U: The Histogram and the Kernel Estimator.

Bandwidth	Histogram	Kernel Estimator
0-2	0.055	0.11
2-3	0.19	0.16
3-4	0.23	0.19
4-5	0.09	0.11
5-7	0.105	0.10
7-11	0.04	0.05

The marginal probability density function is given by $p_k(x)$ in equation 12.

The illustration of kernel regression using Phillips's bandwidths yields a graph of averages with a space between each point. (Phillips measured his graph of averages, having transformed the data into loglinear terms, by means of ordinary least squares regression.) In contrast, a real exercise of kernel regression actually plots the regression line, given a suitable choice of bandwidths. The kernel regression exercise performed below will replace Phillips's expertise in choosing bandwidths by an expert system.

E. A Kernel Regression of Phillips's Data

As mentioned before, the theory of kernel regression is large-sample based. In an <u>economic</u> context, the issue of asymptotic approximation is irrelevant if, as we shall show, there is no theoretical Phillips curve. In a statistical context, however, kernel regression requires satisfaction of the asymptotic conditions.[14]

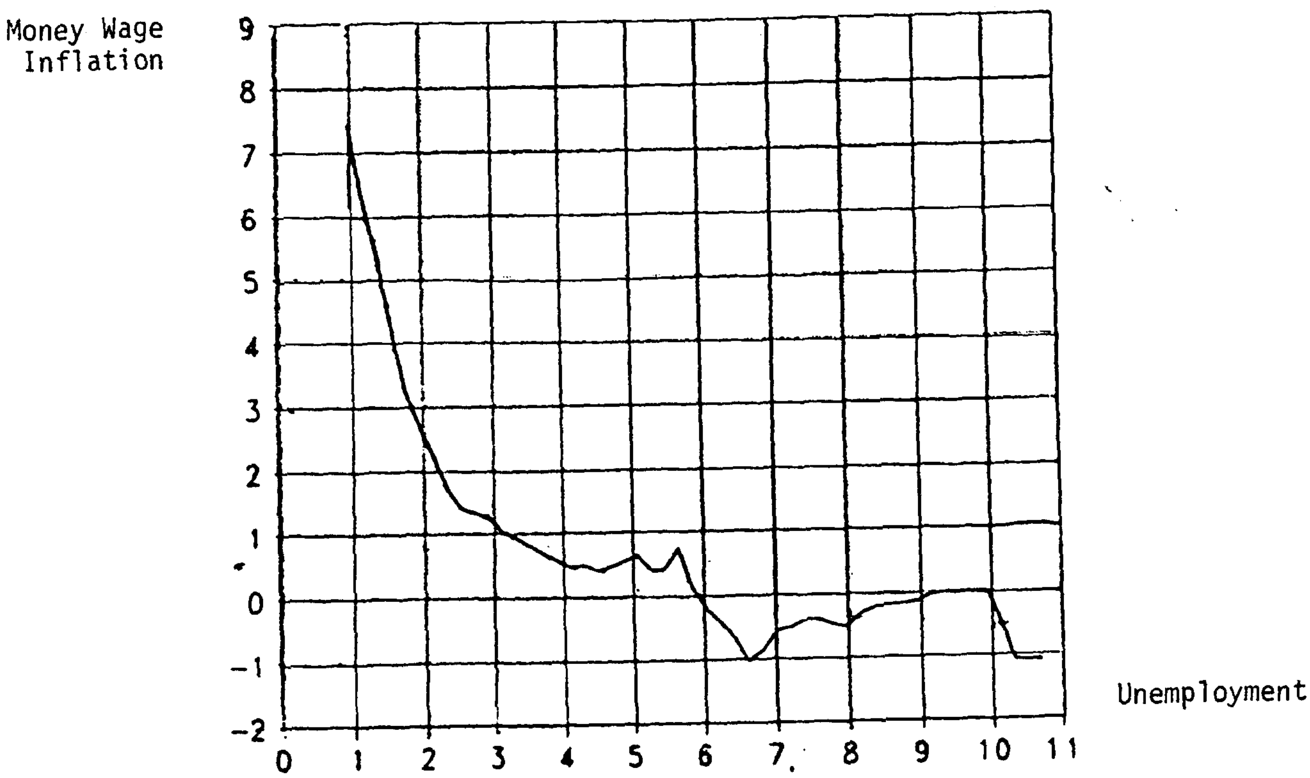
We determined a global optimal bandwidths by means of cross-validation (CV) by least squares. This method minimizes the sum of the squares of the difference between the observed values, Y_i, and the weighted average values of Y given by the kernel regression. A leave-out one regression estimate $r_{-k}(x)$ is used to avoid a trivial solution to the optimization problem. In

---

[14]. Mack and Silverman (1982).

Figure 6

Kernel Regression with Bandwidth = 0.64, Phillip's Data



mathematical terms, the cross-validated bandwidth is that value such that

(13.) $CV(b) = [\Sigma_i^n(Y_i - r_{-k}(x_k)]^2$

is minimized. The estimate of the optimal cross-validated bandwidth on Phillips's data is 0.64. Figure 6, which is based on this global bandwidth and the quadratic kernel, shows a money-wage-unemployment association that is negative on a global scale, but positive within some ranges of unemployment. The wiggles in the regression line at unemployment rates exceeding 4 percent reflect the variability of the inflation-unemployment association in the scatter diagram.[15]

The kernel regressions shown in Figure 6 is peculiar to Phillips's sample. The marked wiggles in the curve suggest the presence of omitted variables that invite historical investigation.

Econometricians also note the presence of omitted variables when they see outlying residual deviations from their formal specifications. But because econometricians seek regularities existing across samples, they will remove the outliers if this improves the results of the statistical tests. The economic rationale for using kernel regression on the data for inflation and the unemployment rate (or output ratio) is that economists do not have a theoretical model of a Phillips curve to estimate or test. In this case, it is unnecessary to replicate the Phillips curve or specify a formal Phillips curve as part of a deductive system. The

---

[15]. See Figure 1.

next section of the paper argues that the standard theoretical model of the Phillips curve is incoherent.

III. The Phillips Curve: Theoretical Structure or Statistical Pattern?

Economic theorists have produced various explanations of the existence of the Phillips curve. Orthodox economists have generated the Phillips curve as an equilibrium relation that arises from the misinformation about relative prices on the part of rational, optimizing suppliers of labor and goods (Friedman (1968); Lucas and Rapping (1969); Lucas (1973)). Keynesian economists have explained the Phillips curve as a price adjustment mechanism which clears markets of excess demand (Phillips (1954); Gordon (1985)).[16] In either case, the Phillips curve can be mapped off a model of supply and demand. We shall analyze geometrical mappings of the Phillips curve from the aggregate supply and aggregate demand curves.[17]

Gordon (1990), who has revived the Keynesian Phillips curve in the 1980s, mapped this curve off the aggregate supply curve.

---

[16]. Gordon (1985), pp.268, 290. Phillips (1954) presented a curve that related the rate of change of the price level (rather than inflation) to the output ratio.

[17]. Note that these two curves are supply and demand curves in a special sense. Figure 7 shows that output, along the horizontal axis is measured in index terms, where an index of 100 stands for the "natural" level of output. Points on the graph with first coordinates that exceed 100 refer to the supply of excess output or the demand for excess output.

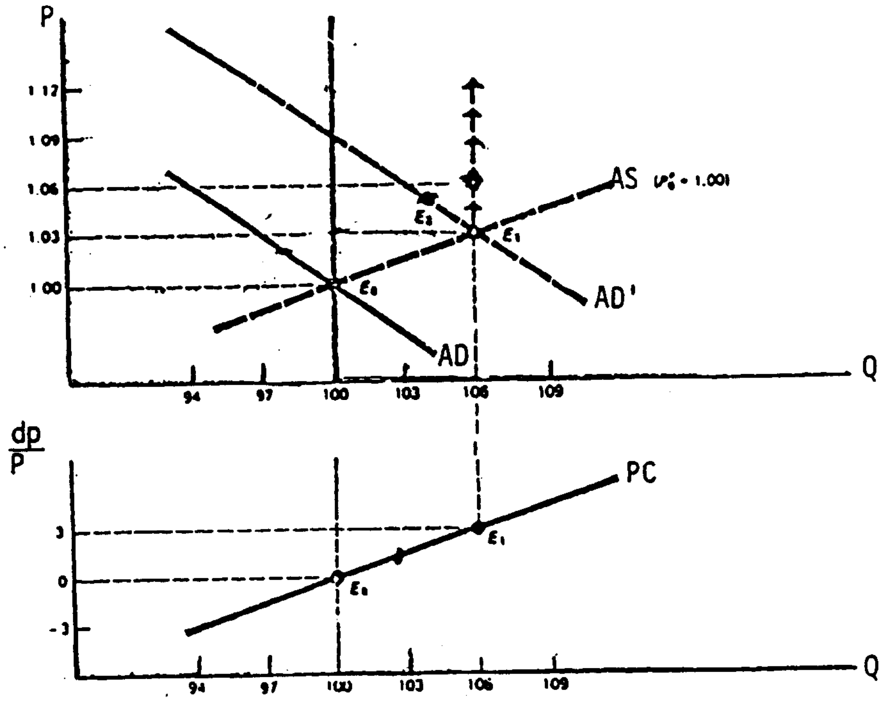


Figure 7a

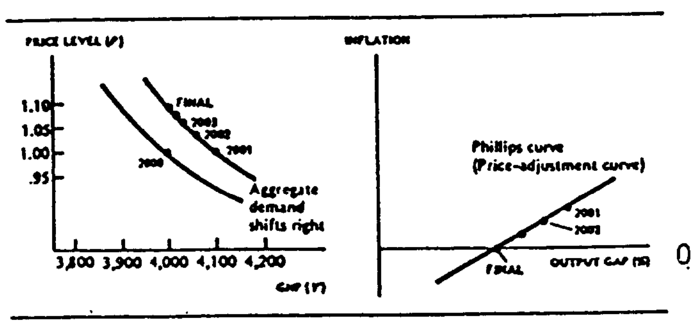


Figure 7b

Figure 7c

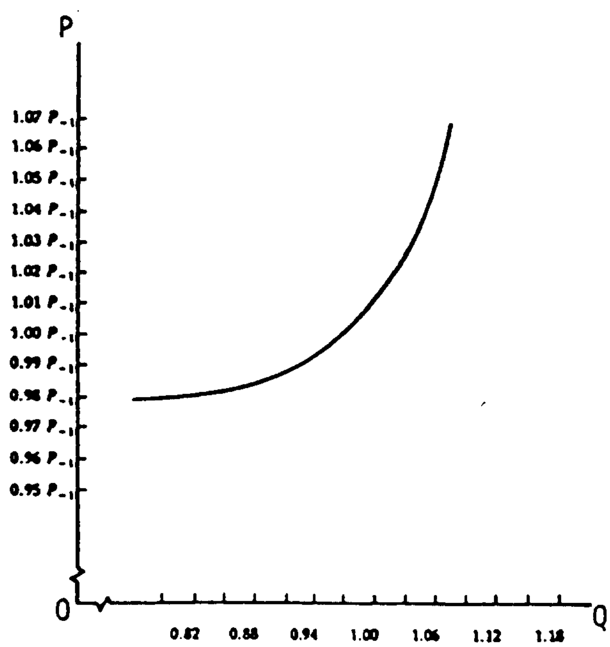


Figure 7a shows a shock to aggregate demand increasing the equilibrium output ratio and the price level over their "natural" levels at E_o. For every possible shock to demand off the "natural" equilibrium level, E_o, Gordon mapped a Phillips curve off the aggregate supply curve, as if the Phillips curve is a logarithmic transformation of the aggregate supply curve.[18] The mapping involves the special assumption that inflation, I, is defined as the change in the second coordinate as one moves up along the aggregate supply curve from the natural level -- that is, as

$$I = (p_a - p_n)/p_n,$$

rather than by the proportionate rate of change of the price level per unit of time.[19]

Hall, the policy adviser at the Hoover Institute, and Taylor, a member of President Bush's Council of Economic Advisers, mapped an aggregate demand curve off the short run Phillips curve (Hall and Taylor (1986)). Figure 7b shows that a shock to aggregate demand creates excess demand in year 2001. Given the output ratio, Q, in 2001, the Phillips curve shows the rate of inflation in that

---

18. Assume that

(i.) $p = bQ$.

Then,

(ii.) $\log p = \log b + \log Q$.

Differentiate (ii.) in respect to time and, as an approximation, write the result in finite terms,

(iii.) $(p_a - p_n)/p_n = (Q_a - Q_n)/Q_n$,

where n stands for the natural rate and a for the actual rate.

19. Gordon shows that unless aggregate demand continues to increase along the shifting aggregate supply curve (the shifts arise because agents fully anticipate inflation), the Phillips curve will correspond to movements along the aggregate demand curve, similar to the analysis of Hall and Taylor, below.

year, which allows one to calculate the higher price level, p, of year 2002. The inflation during 2001 in turn reduces the level of excess demand. The inflation continues at a decreasing rate until the market clears.

Melvin and and Darby, the chief economic adviser at the US Treasury, mapped the short run aggregate supply curve off the short run Phillips curve (Darby and Melvin (1986)). Figure 7c shows that the Phillips curve negatively relates the deviation between the actual and expected inflation rate with the unemployment rate, which is related to the output ratio, Q, by Okun's law. If the expected inflation rate is constant, given last period's price level, p_{-1}, one can compute the current p level from the actual rate of inflation. One then has the Q,p coordinates to construct an aggregate supply curve.

Each of these three explanations presuppose that the price level changes following a change in the output ratio are <u>proportionate changes</u>. The mappings thus presuppose a one-to-one relation between the Phillips curve and aggregate supply (or demand). In fact, the aggregate supply and aggregate demand curves must satisfy certain first and second order conditions to correspond to a Phillips curve. We shall derive these conditions below.

Recall that the definition of the output ratio is the relative difference between the actual and the long run equilibrium levels of output, Y, that is,

(14.) $Q = (Y_a - Y_n)/Y_n.$

The proportionate rate of price inflation, I, normally refers to the change in the price level, p, over time, that is,

$$(15.)\qquad I = (p_t - p_{t-1})/p_{t-1},$$

$$= dp(Q)\ /\ p(Q).$$

Since the actual price level varies with the output ratio, inflation can be defined as

$$(16.)\qquad I = \frac{1}{p(Q)}\ \frac{dp(Q)}{dQ}.$$

The Phillips curve means that the change in inflation in respect to the output ratio is positive, that is,

$$(17.)\qquad \frac{dI}{dQ} > 0.$$

Given equation 17, dI/dQ can be written as

$$\frac{dI}{dQ} = \frac{d}{dQ}\left[\frac{1}{p(Q)}\ \frac{dp(Q)}{dQ}\right] > 0.$$

For the purposes of exposition, let us call $dp/dQ=p'$ and $d^2p/dQ^2=p''$. Expanding the terms on the right-hand side of equation 17 yields,

$$(18.)\qquad \frac{dI}{dQ} = pp'' - p'^2 > 0.$$

Equation 18 is the mathematical condition that will be satisfied by any aggregate demand or aggregate supply curve that is consistent with the Phillips curve.

Appendix B considers four functions, a supply and a demand curve that are each linear and a supply and a demand curve that each take the form of a parabola. It is easy to show that the condition stated in equation 18 is not satisfied by three of those functions, the linear demand curve and the supply curves that are linear or a parabola. Orthodox economics has dictated that the supply curve is nonlinear, on the *a priori* assumption of diminishing marginal returns, an assumption that allows income distribution and the level of output to be determined simultaneously, given the supply of labor and capital. However, economics lacks any reason whatsoever for the aggregate demand only to be nonlinear and the aggregate supply curve not to take on the form of a parabola.

The compatibility of the Phillips curve with the conventional aggregate supply-demand analysis has promoted the continued teaching of the Phillips curve.[20] However, the Phillips curve requires that *ad hoc* and unlikely restrictions be put on the form of the supply or demand curve. Because these curves form the building blocks of the standard macroeconomic model, the theoretical coherence of the Phillips curve itself is brought into question. In this case, it is inappropriate to treat data on inflation and capacity utililization as the basis for estimates of a given theoretical model.

---

[20]. Haynes and Stone (1985) identify the Phillips curve as a dynamic aggregate supply phenomenon.

IV. Conclusion

This paper has raises the methodological issue of how alternative methods of treating given data produce different ways of seeing the data, with the choice of method depending on the assumptions about the phenomenon under study.

Kernel regression and parametric regression provide alternative ways of analyzing data. Least squares regression imposes a simple mathematical form on the data, on the assumption that the data are samples of attributes of a normal distribution, and allows analysis of the residual deviations from the regression, on the assumption that the residuals are drawn from a normal distribution. Kernel regression, a data-driven, smoothing technique, shows the major irregularities as well as the overall pattern of the data in the sample without any assumptions about the distribution of the population.

Economists rarely have known the distributions of their data and not all known distributions (particularly of price data) were distributed normally. The residuals from econometric models often have not appeared to follow a normal distribution. Thus there are statistical reasons for economists not to rely on the standard parametric regression techniques, which assume normality. Economists might obtain a "better" picture of events using other tools of observation.

Economists require parametric regression techniques to estimate models of an economic theory. However, standard economics has lacked a basic, formal model of the Phillips curve. The measures of the Phillips curve given by a parametric regression do not estimate a theoretical model, if there is no theoretical model to be estimated across samples. In this case, it is unnecessary to impose a parametric model on the data. Kernel regression, a data-driven technique, offers a picture of the general movement of inflation and unemployment as a basis for exploration.

Appendix A.

1. Phillips's Data

	Unemployment (X_i)	Money Wage Inflation (Y_i)
1861	3.7	0
1862	6.05	1.4706
1863	4.7	3.5714
1864	1.95	3.4247
1865	1.8	3.3333
1866	2.65	1.2821
1867	6.3	-1.9481
1868	6.75	-1.3333
1869	5.95	2
1870	3.75	3.2051
1871	1.65	6.875
1872	.95	8.9888
1873	1.15	5.7292
1874	1.6	2
1875	2.2	-0.5
1876	3.4	-1.0101
1877	4.4	-2.0408
1878	6.25	-2.6316
1879	10.7	-1.0753
1880	5.25	0
1881	3.55	0
1882	2.35	0.5376
1883	2.6	0.5319
1884	7.15	-0.5319
1885	8.55	-0.5376
1886	9.55	0
1887	7.15	0
1888	4.15	1.6129
1889	2.05	3.6458
1890	2.11	2
1891	3.4	0
1892	6.2	-0.5
1893	7.7	-0.5051
1894	7.2	-0.5051
1895	6	0
1896	3.35	1.0101
1897	3.45	1.5
1898	2.95	1.9608
1899	2.05	2.8846
1900	2.45	1.3889
1901	3.35	-0.4673
1902	4.2	-0.4673
1903	5	-0.9434
1904	6.4	-0.4762

1905	5.25	0.9524
1906	3.7	0.9346
1907	3.95	0
1908	8.65	0
1909	8.7	0
1910	5.1	0.4673
1911	3.05	1.8519
1912	3.15	3.1532
1913	2.1	1.7391

2. Phillips's Bandwidths

	Mean (x_k)	Width (b_k)
0-2	1.5167	2
2-3	2.35	1
3-4	3.4	1
4-5	4.49	1
5-7	6	2
7-11	8.4	4

Appendix B.

The aggregate supply and demand curves consistent with the Phillips curve will satisfy the condition,

(i.) $pp'' - p'2 > 0$.

Suppose that the supply curve takes the form

(ii.) $p = aQ + b$

and the demand curve,

$p = -dQ + m$.

Then,

(iii). $p' = a$

for the supply curve and

$p' = -d$

for the demand curve. For both curves,

(iv.). $p'' = 0$.

Substitute (iii.) and (iv.) into equation (i.). The results,

(v.). $[(aQ + b)0 - a2] > 0$

for the supply curve and

$[(-dQ + m)0 - d^{2]} > 0$

are impossible.

Suppose that

(vi.) $p = bQ^a$.

Then

(vii.) $p' = abQ^{a-1}$

and

(viii.) $p'' = a(a-1)bQ^{a-2}$

Substitute equations (vii.) and (viii.) into (i.) and simplify, which gives

(ix.) $-ab^2Q^{2(a-1)} > 0$.

Since Q and b have positive values, given (ix.) then

(x.). $-a > 0$.

This can only hold if a<0, which means that the curve is downward sloping.

References

Darby, M. and Melvin, M. (1986), Intermediate Macroeconomics, Glenview: Scott Foresman.

Friedman, M. (1968), "The Role of Monetary Policy", American Economic Review, V.68, pp.1-17.

Gordon, R. (1985), "Understanding Inflation in the 1980s", Brookings Papers on Economic Activity, V.1, pp.263-299.

___________(1990), Macroeconomics, Boston: Little, Brown.

Hall, R. and Taylor, J. (1986), Macroeconomics, New York: Norton.

Haynes, S. E. and Stone, J. A. (1985), "A Neglected Method of Separating Demand and Supply in Time Series Regression", Journal of Business and Economic Statistics, V. 3, No. 3.

Lucas R. (1973) "Some International Evidence on Output-Inflation Tradeoffs", American Economic Review, V.63, pp.326-334.

___________ and Rapping, L. (1969), "Real Wages, Employment and Inflation", Journal of Political Economy, V. 77, pp.721-754.

Mack, Y. P. and Silverman, B. W. (1982), "Weak and Strong Uniform Consistency of Kernel Regression Estimates", Zeit. Wahr. verw. Geb., Vol. 61, pp.405-415.

Phillips, A.W. (1954), "Stabilisation Policy in a Closed Economy", Economic Journal, V.67, pp.290-323.

_____________ (1958), "The Relation between Unemployment and The Rate of Change of Money Wage Rates in the United Kingdom, 1861-1957", Economica, pp.283-199.

Rosenblatt, M. (1956), "Remarks on some Nonparametric Estimates of Density Function", Annals of Mathematical Statistics, V. 27, pp.832-837.

Silverman, B. W. (1988), Density Estimation, London: Chapman and Hall.

Stock, James H. (1989), "Nonparametric Policy Analysis", Journal of the American Statistical Association, V.84, pp.567-575.

Ullah, A., editor (1988), Semiparametric and Nonparametric Econometrics, New York: Springer-Verlag.

Wulwick, N. J. (1987), "The Phillips Curve: Which? Whose? to Do What? How?", Southern Economic Journal, V.53, pp.834-857.

_______________(1989), "Phillips's Approximate Regression", Oxford Economic Papers, V.40, pp.17-188.