The output gap and inflation – experience at the Bank of England[1]

Paul G. Fisher, Lavan Mahadeva and John D. Whitley

Introduction

Modelling the UK economy at the Bank of England is based on the premise that no single methodological and empirical approach is likely to prove adequate to address the many economic issues which we typically face. Consequently a suite of models has been developed which span the range from highly theoretically representative agent models, based on calibrated data, to VAR-type models where economic theory plays a less well-defined role and statistical considerations are more dominant. This suite of models is described in Fisher and Whitley (1997). As part of this suite we have developed simple reduced-form models which summarise a more structural macroeconomic approach. An example of the reduced-form approach is the use of a measure of the output gap to summarise inflationary pressure as in a Phillips curve. This relationship can then provide a cross-check on forecasts of inflation arising from more detailed models. Its usefulness depends, however, on how well changes in the output gap predict changes in inflation. In this paper we describe our experience with this approach.

The first section of the paper sets out some of the methodological considerations. The second section describes alternative measures of the output gap and then derives a preferred measure together with an estimate of its uncertainty. The third section describes how measures of inflation expectations can be derived and this is followed by econometric evidence which relates the output gap and inflation expectations to actual inflation. Sections 5 and 6 look at issues associated with the Phillips curve: the possibility of sectoral bottlenecks; shifts in the short-run trade-off due to labour market flexibility; the possibility of "speed limit effects" and asymmetry in response to changes in the output gap. The final section sets out the conclusions.

1. Methodological considerations

The output gap is generally used to measure the extent to which the economy is operating at an unsustainable level of resource utilisation – often expressed empirically as the deviation of actual output from trend. In recent years it has usually been used to represent the extent to which the current level of output lies below the equilibrium level: sometimes called potential or full capacity output. Different implicit definitions of equilibrium partly explain different estimates. In ad hoc analyses the output gap is often measured as positive when output is below trend. When examining analytical models it is more consistent to define a positive value as output above trend. In what follows we always use the convention of a positive gap to mean that output is above trend.

In some alternative definitions – not used here – the output gap is defined in terms of how near the economy is to a maximum capacity ceiling; hence it is logically bounded at zero and always defined as positive. An example is full capacity based on peak-to-peak trend fitting of the business cycle.

---

[1] Preliminary results.

The output gap is used for two primary purposes – the analysis of inflationary pressure and cyclical adjustment of other variables, notably the public sector deficit. We concentrate here on the link to inflation.

1.1 An underlying theory of the output gap

In the event of a (positive) output gap caused by a positive demand shock, firms will employ more labour in the short-run for a given capital stock to produce the extra output to meet demand. To induce a greater supply of labour, firms must bid up the real wage rate (this is consistent with a large variety of labour market models). To work existing capital beyond its optimum can also bid up costs per unit of output. On both counts costs rise and, with a constant mark-up of prices over costs, prices also rise. Thereby real wages are reduced and the interaction of attempts to maintain real wages generates further rises in costs and prices (assuming that changes in the money stock accommodate higher demand). Inflation continues until policy reacts to offset the demand shock and reduce inflation if necessary. In the longer run the capital stock is unchanged and hence output should return to its previous level (or trend trajectory).

In the case of a (negative) output gap caused by a permanent and positive supply shock, higher productivity raises the return on capital, and should eventually lead firms to increase output. The extra level of output generates sufficient income to produce an equivalent level of demand. However, in the initial absence of higher demand firms will not expand output immediately. They thus need less labour, causing disinflationary pressure in a cycle opposite to that described for a demand shock. Disinflation only disappears if demand is increased and so, ultimately, policy must initiate, or at least accommodate, such an increase. In the longer run, the higher return on capital should lead to a natural expansion of output and demand through higher investment. In the case of a supply-side shock there is, therefore, a short-run and longer-run definition of potential in which the former is conditional on the existing capital stock. Hence there are corresponding different measures of the output gap. With the focus on inflationary pressure, we concentrate on the short-run definition.

An output gap can arise from either demand-side shocks or supply-side changes. In either case the output gap is consistent with disequilibrium in the labour market – sustained by an (dis)inflationary spiral – and sub-optimal capacity utilisation. Although the immediate consequence of an output gap is likely to be similar whatever the source, the longer-run implications are quite different as a supply-side shock can have a permanent effect on non-inflationary output levels (or even growth). To calculate the output gap one needs to be as clear about what is happening to potential as to actual output. Much of the debate on this topic is concerned with the measurement of potential, whereas the more simplistic calculations assume that the trend in potential output is either fixed or, at least, changes relatively slowly.

Although commonly used in analytical and relatively simple empirical models, the output gap rarely appears directly in structural macroeconometric models used for practical forecasting and simulation analysis. Such models typically focus directly on labour market pressure and, additionally, introduce direct pressure of demand in the goods market. For example, in the Bank's medium-term forecasting model, pressure in the goods market is captured by changes in the capital/output ratio. In the labour market there is a real wage bargaining model with a fixed natural rate of inactivity (variable by assumption). Deviations from the equilibrium inactivity rate and capital/output ratio together give the overall degree of inflationary pressure and hence the overall output gap.

The motivation for looking at the output gap is most simply illustrated by the Phillips curve. It is observed that fluctuations in activity are positively correlated with inflationary pressure. In what follows we illustrate this relationship using a generalisation of the expectations-augmented Phillips curve which can explain shifts in the observed relationship between inflation and activity. Note that in this general model, the *level* of inflation is not determinate unless supplemented by a nominal anchor through monetary policy. As such the Phillips curve only tells us about the dynamics of inflation. A further restriction on interpretation is that the Phillips curve makes no distinction between whether an output gap arrives through a demand or supply shock though the relationship is more likely to be robust for demand shocks.

The terms "output gap" and the "NAIRU" (non-accelerating inflation rate of unemployment) are often used interchangeably but are not identical. The NAIRU relates to equilibrium in the labour market whereas the output gap is a whole economy concept. For example, a shock to technical progress will change potential output, and hence the output gap, but will leave the NAIRU unchanged. In our discussion we concentrate on the output gap. In general we can write the relationship between the output gap and inflation as:

$$\Delta p = \phi \Delta p_{+1}^{e} + (1-\phi)\Delta p_{-1} + \gamma(y - y^*) \qquad (1)$$

where Δp is the rate of inflation, Δp_{+1}^{e} is expected future inflation and y–y^* is the output gap. As the periodic re-setting of prices and wages is not synchronised across the different groups in the economy, the aggregate wage in part responds to expected future inflation and in part to its past value. The parameter ϕ thus measures how sticky inflation is. Expected future inflation is assumed to be given by:

$$\Delta p_{+1}^{e} = \mu \Delta p_{-1} + (1-\mu)\Delta p^* \qquad (2)$$

where μ lies between zero and one and Δp^* is the government's inflation target. This means that individuals expect next period's inflation to be a weighted average of the government's target inflation and past inflation. The parameter μ measures how credible the government's target is: we have assumed that if μ is one, individuals believe that inflation will be unchanged from last period's value irrespective of the government's target. This gives the Phillips curve as:

$$\Delta p = \alpha + \beta \Delta p_{-1} + \gamma(y - y^*) \qquad (3)$$

where $\alpha = \phi(1 - \mu)\Delta p^*$ and $\beta = \phi\mu + (1 - \phi)$.

In general α is non-zero while β lies between zero and unity, and the short-run Phillips curve is non-vertical. Remember that this assumes the existence of a monetary policy rule which delivers an average inflation rate of $\Delta p^* = \alpha/(1 - \beta)$. If the policy target – implicit or explicit – changes, then forward-looking rational agents will take account of this (the Lucas critique applies to this equation) and the parameter α will change. Ultimately, if policy has no credibility, μ and β tend to unity and α to zero such that inflation becomes indeterminate. This implies that a positive (or negative) output gap would cause inflation to be permanently increasing (or decreasing). If, on the other hand, policy is credible, inflation will move towards its new target at a rate depending on the degree of stickiness in wage and price setting.

Conditional on the parameter values and expectations formation, we have alternative policy scenarios for when a positive real demand shock hits the economy (for a negative shock the signs are simply reversed – in fact the first scenario assumes that negative and positive shocks are equally likely so as to maintain an average inflation rate). In the first scenario, we assume that policy is committed to an inflation target and economic agents fully believe in the authorities' commitment ($\mu = 0$) either explicitly or implicitly (e.g. through money growth rules); but inflation is sticky ($\phi<1$). The existence of an output gap puts upward pressure on inflation but the policy rule ensures that the gap is removed through a temporary policy tightening, so as to bring inflation back to target.

In the second scenario we also assume that the rate of inflation is sticky, while economic agents give a zero weight to the determination of the policy authorities to keep to an inflation target: $\mu = 1$, $\alpha = 0$ and $\beta = 1$. This implies that money growth is entirely endogenous such that there is no nominal anchor. In this case, inflation will rise unless the output gap disappears, so that bringing inflation back to the initial level requires an equal and opposite change in the gap – induced perhaps by a policy response.

A third scenario arises if inflation is perfectly flexible (ϕ=0) and an anti-inflationary policy is anticipated (μ=0). Inflation will now return to its initial level of its own accord as soon as the gap is closed.

The stylised facts for the United Kingdom prior to 1992 suggest a degree of price stickiness and less than full credibility. This corresponds to a case closer to the second than the first scenario. Low credibility may have occurred as the result of an insufficiently anti-inflationary policy. Inflation responds to a demand shock, increasing rapidly once agents realise that no policy action is forthcoming. When action is finally taken, policy not only has to compensate for the initial shock but must induce an equal and opposite shock to bring inflation back down again. The greater the degree of price stickiness the longer this process takes.

The monetary arrangements between September 1992 and May 1997 can be seen as approximating the first scenario; establishing credibility to help inflation to return to target, albeit slowly. Achieving the third scenario, with stable inflation and low output costs on the back of an anticipated and credible anti-inflationary policy, will be associated with even lower output costs if wages and prices are less sticky and such changes may themselves be facilitated by the more credible approach.

The value of the output gap approach in monitoring inflationary pressure (and thereby in setting policy) arises from the idea that a change in the output gap usually precedes the change in inflationary pressure it causes. Early action to counteract real shocks can thus minimise fluctuations in inflation. This lagged effect requires a degree of price stickiness whether or not this is accompanied by incomplete credibility.

The presence of lagged adjustment makes the dynamic processes much more complicated and introduces the possibility of "speed limit effects". Inflation may depend as much on the change in the output gap as the level. We discuss the relevance of these effects in Section 6 below.

An open economy affects the analysis in at least three ways:

(i) The direct effect of import prices means that there can be temporary external shocks to domestic retail price inflation. This can be expressed by adding a term to the Phillips curve for the deviation of the real exchange rate from equilibrium.

(ii) The existence of external trade means that capacity pressures may be offset by changes in imports/exports. However, in the absence of a change in relative prices, this tends to be a short-run effect since it is otherwise inconsistent with balance of payments equilibrium.

(iii) The external sector is an additional source of demand shocks.

2. Measuring the output gap

To measure the output gap one needs to estimate the unobservable level of potential output. The methods can be broadly summarised by the following categories:

(i) Smoothing and de-trending methods.

(ii) Econometric estimates (including production function estimates).

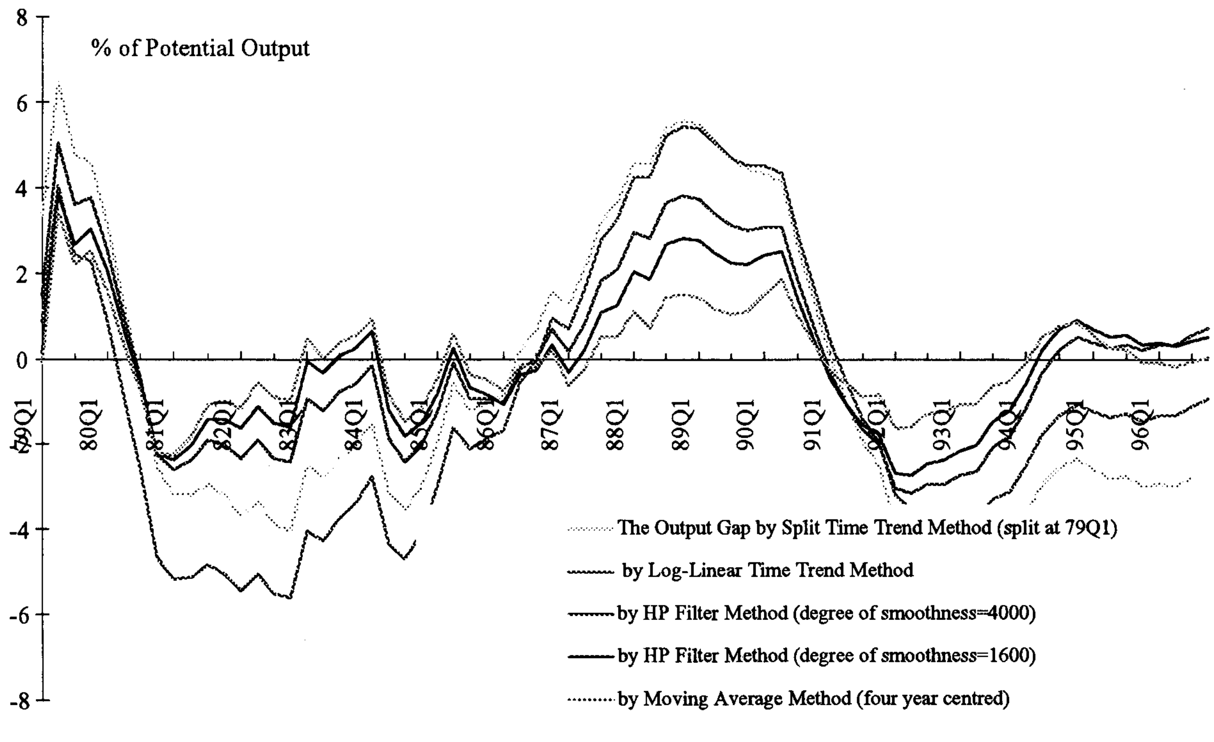
(iii) Survey data on capacity utilisation.

2.1 Smoothing and de-trending methods

The first category of methods are essentially statistical because they assume only that potential output follows some simple trend-like behaviour over time. They therefore focus on isolating this trend as a measure of potential output. Methods range from assuming that potential output grows at a constant exponential rate (the log-linear time trend); through assuming that it grows at a rate which is constant except for a few sudden breaks (the split time trend); to assuming that potential output can change freely but smoothly over time (i.e. the trend can "bend"). An example of a smoothing method is a moving average filter which calculates potential output at any point in time as a weighted average of the current future and past values of output around that point in time and thus evens out the cyclical effects (the filter should be of cycle length). Another popular example of a smoothing method is the Hodrick-Prescott filter (HP filter) which trades off the degree of smoothness in potential output against the extent to which it tracks actual output. This trade-off is determined by a parameter (λ, the degree of smoothness) which can be set to any value between the extremes of zero (where potential output is always equal to actual output and there is no output gap) and infinity (where potential output corresponds to the log-linear time trend). Although all de-trending methods are biased when a structural break has occurred, the Hodrick-Prescott method suffers from the additional disadvantage that it relies on extrapolations of data beyond the end of the period under consideration.

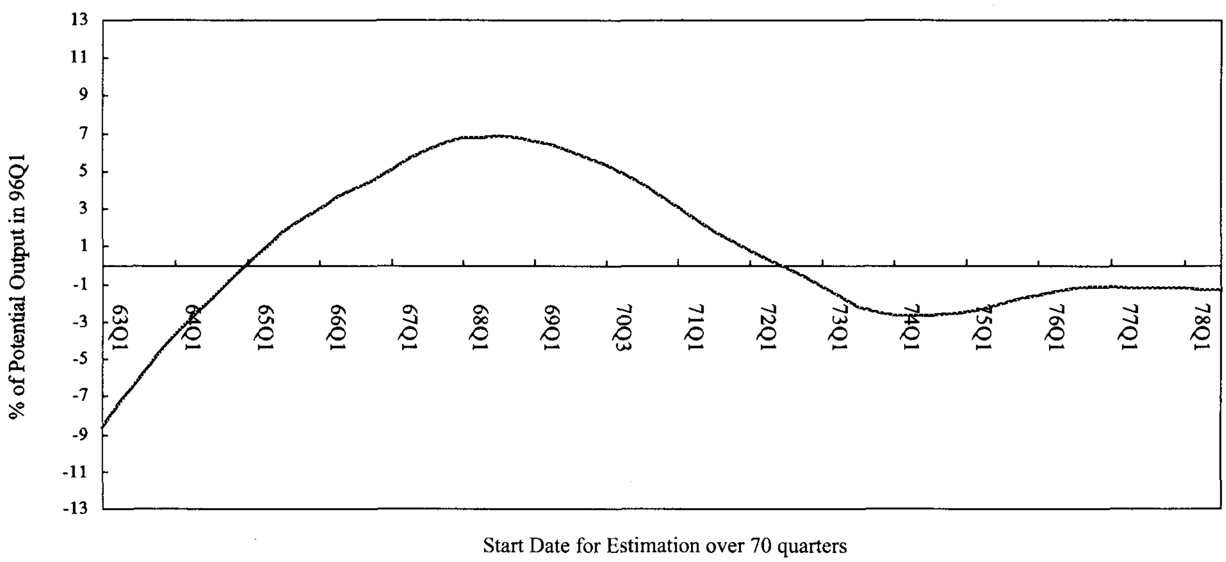
Chart 1

Output gaps by different smoothing and de-trending methods



In Chart 1 we present five measures of the output gap, corresponding to different trend methods. The HP filter is calculated for two degrees of smoothness, one of which (λ = 1,600) is that, typically found in the literature. Although there are similarities in the direction of change of the output gap over time, there is considerable variation in the most useful estimates: the value of the output gap at the end-point. This uncertainty is general as each method tends to produce a diverse range of estimates, depending on the particular assumptions made and the methods are always most uncertain about the current situation. For example, in Chart 2 we present the measure of the 1996Q1 output gap that would be given by using the log-linear time trend method to estimate a constant exponential growth rate for potential output using different time periods. Inclusion of the 1960s gives a significant positive gap in 1996Q1 but if, instead, we restrict our estimation to a sample beginning in the late 1970s, 1996Q1 would be a little less than 2% below trend. Further investigation using samples of the same size, over different periods, confirms the sensitivity of the 1996Q1 end-point estimate. This suggests that much of the disagreement about current estimates of the output gap in the United Kingdom should be cast in terms of economic arguments about what historical period is relevant for deriving the estimate.

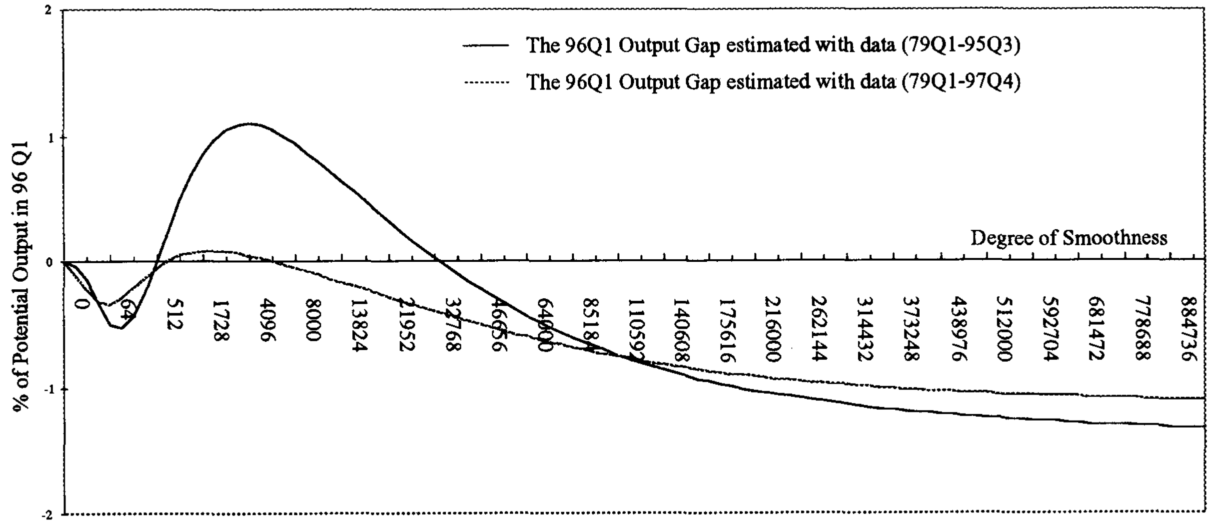
Chart 2
The sensitivity of the 1996Q1 output gap to the start date by the log-linear time trend



Similar problems are encountered with the HP filter method where the measure of the gap is sensitive to the degree of smoothing. Chart 3 presents the projections made with different values of λ with similarly dramatic results. The OECD has calculated the UK output gap using a value of 100 for λ and this would imply a small negative output gap for 1996Q1. A value of 1,600 gives a positive output gap of about 1%. As λ approaches infinity, the HP filter method gives the same result as would the assumption of a log-linear time trend for the sample used.

One plausible explanation as to why these smoothing and detrending methods fail to deliver robust results is that potential output has a random walk component rather than being a simple exponential or linear function of time. When this is the case, we would expect inherent instability in those detrending methods which fail to take account of this (Nelson and Plosser (1992) and Canova (1986)). This instability may manifest itself in the discovery of spurious output gaps (Nelson and Kang (1981)). There are measures of the output gap which begin by assuming that the potential output follows such a stochastic process (for example the Beveridge-Nelson decomposition). However, one problem with these methods is that there are an infinite number of ways in which a univariate time series can be decomposed into a trend and cycle (Quah (1992)). The issue here is that although each measure tends to give more stable estimates, there is a great divergence between the different methods and again, when a time series of GDP is considered in isolation, it is difficult to use economics to discriminate between different members of this class of estimates. The main conclusion is that there are no stylised facts about output gaps.

Chart 3
The sensitivity of the 1996Q1 output gap to the degree of smoothness in the HP filter method



2.2 Econometric methods

A natural response to this is to use economics to help identify differences in the measurement of potential output. Consider the following example using a simple Cobb-Douglas production function. Writing this in log-linear form (with lower case letters for logs):

$$y = \alpha n + (1-\alpha)k + mt + c \qquad (4)$$

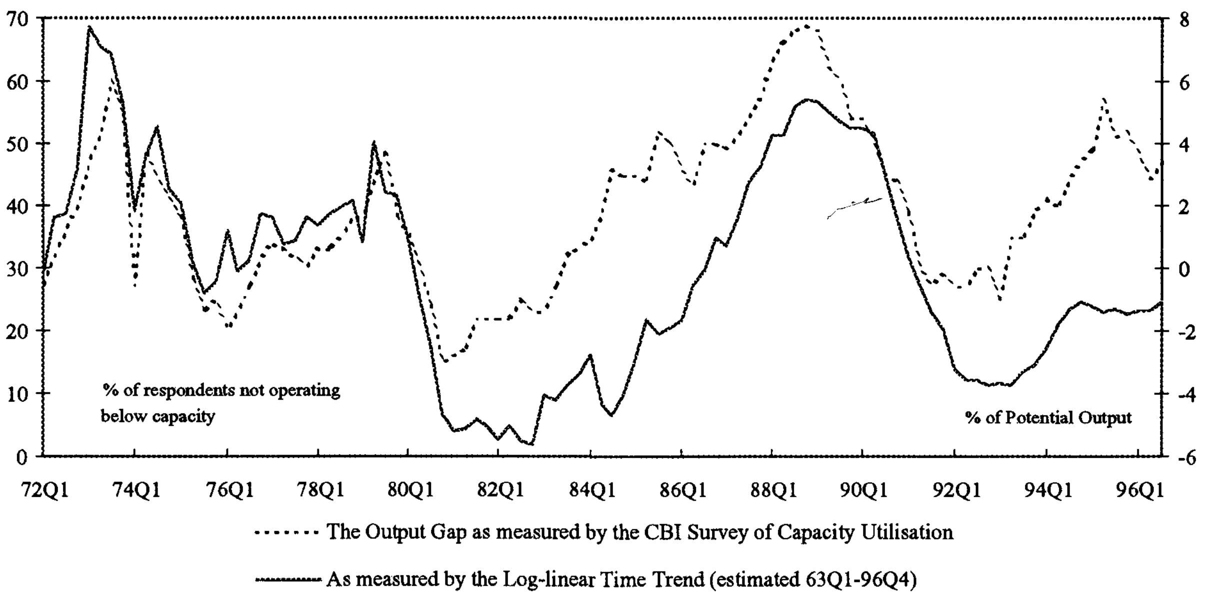
where y is output, n the labour force, k the capital stock and t a time trend with the parameter α giving the labour share. The values of the inputs n and k can be measured at equilibrium utilisation levels. Given an estimate of the labour share of income (α), we can estimate trend technological progress (m) by regressing the residual of log output over the estimates of the full capacity contribution of the factor inputs, on a constant and time trend. In our calculations, we have simply assumed that full capacity labour is given by the working population and capital by cumulated gross investment net of a constant depreciation rate.

Alternatively, econometric estimates of equilibrium output can be derived from some combination of theory and empirical estimation, such as the NAIRU (the non-accelerating inflation rate of unemployment). This involves using wage and price equations to determine directly the equilibrium rates of utilisation of labour and capital and comparing these with actual rates to determine a measure of the output gap. In principle, any uncertainty about the output gap measure can then be more easily linked to theoretical differences. The disadvantage is that these measures tend to be unstable over time. The Layard-Nickell approach (1991) is an example of how the NAIRU can be derived from wage and price behaviour. As noted above, the NAIRU and the output gap are not identical in concept and hence practice. We cannot, therefore, translate empirical estimates of one measure into another.

2.3 Survey evidence

The CBI Industrial Trends Survey asks firms whether they are operating at full capacity and gives a direct, though qualitative, indication of capacity utilisation (see Chart 4). Although this and other surveys tend to cover only the manufacturing sector, it can be generalised to the whole economy if other sectors have a common cycle with manufacturing. Problems can arise if the assumption about normal capacity levels changes over time in the survey responses. Furthermore, it is not clear whether firms take into account both labour and capital in assessing their level of capacity utilisation.

Chart 4
The output gap as measured by the CBI Survey of capacity utilisation in the manufacturing sector



2.4 Probability distribution of estimates of the output gap

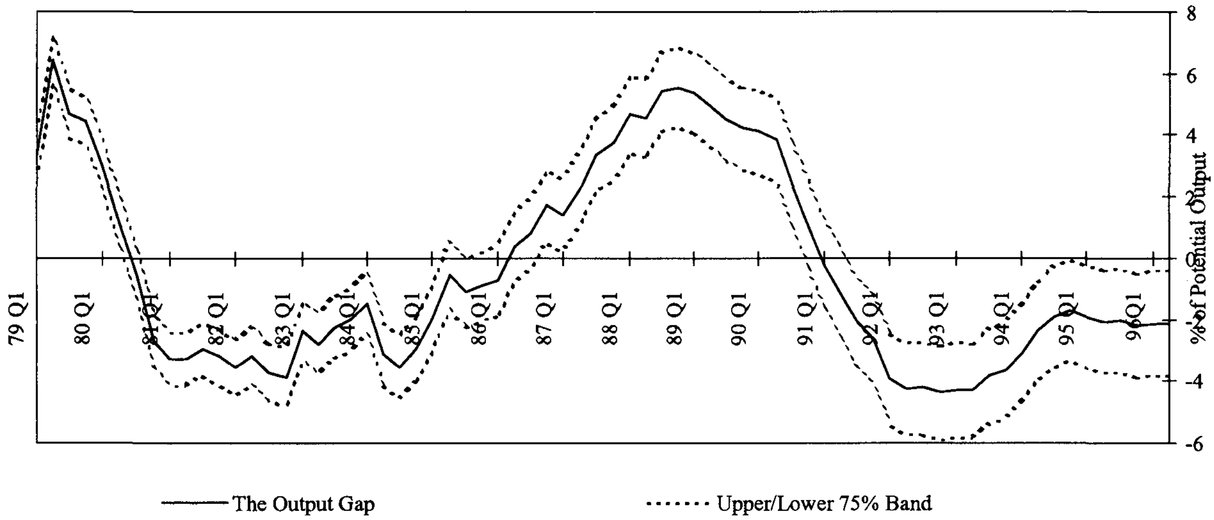
Each method of measuring the output gap typically produces a diverse range of estimates, depending on the particular assumption used. It is not possible to discriminate between these measures in the absence of underlying theory. In consequence, we have greater faith in methods that are more theory-based, such as the production function approach. This incorporates views as to how labour and capital are combined to produce output.

Furthermore, since point estimates of the output gap are unhelpful when there is a wide area of uncertainty, it is important to emphasise the degree of statistical confidence in the cases where we can derive the probability distribution of the estimate of the gap. Where no statistical measure of confidence can be derived there is an argument for not reporting estimates of the mean gap since it implies a misleading degree of precision. In the production function approach we can derive a confidence interval based on the econometric procedure used and in Chart 5 presents the bands of 75% certainty for an estimate using this methodology assuming a shift in technical progress after 1979.[2] The uncertainty in the estimate arises from the uncertainty in estimating the constant and the coefficient on the time trend in the equation described above. The precision of the estimate is also predicated on the assumption that we know the values of full capacity labour and capital, actual GDP,

2 The dating of this shift is confirmed by recursive estimates. Allowing for the shift implies that technological progress improves at an annual rate of 0.6% for the period beginning in the 1970s until 1979 and at a rate of 1.27% thereafter. Potential output as a whole grows at a faster (non-constant) rate because it depends on the rates of growth of equilibrium labour and capital as well as that of technological progress.

the labour share of income and the form of the production function with certainty. Relaxing these assumptions would widen the bands even further. The results show that the mean value of the output gap is -2.1% in 1996Q2 with a 75% confidence interval of -2.9% to -1.3%.

Chart 5
The output gap by the production function method



Our research has shown that the choice of time period is also critical to obtaining estimates from this method. Applying three different assumptions to quarterly real GDP data for the United Kingdom gives estimates of the output gap at the start of 1996 of alternatively 1.3%, -2.1% or -3.4%. The first estimate derives from the assumption that technical progress improves at a constant rate over the last 30 years (1963-93); the second that the whole process of technical progress has changed since 1979; and the third that there has been a change (increase) in the rate of improvement in technical progress after 1979, combined with a step loss in capacity in the 1980-81 recession. The rationale for a shift around 1980-81 is consistent with the effects of the shake-out of the manufacturing sector on aggregate productivity growth. We have adopted the second of these assumptions in deriving our preferred estimate of the probability distribution of the output gaps.

3. Measuring inflation expectations

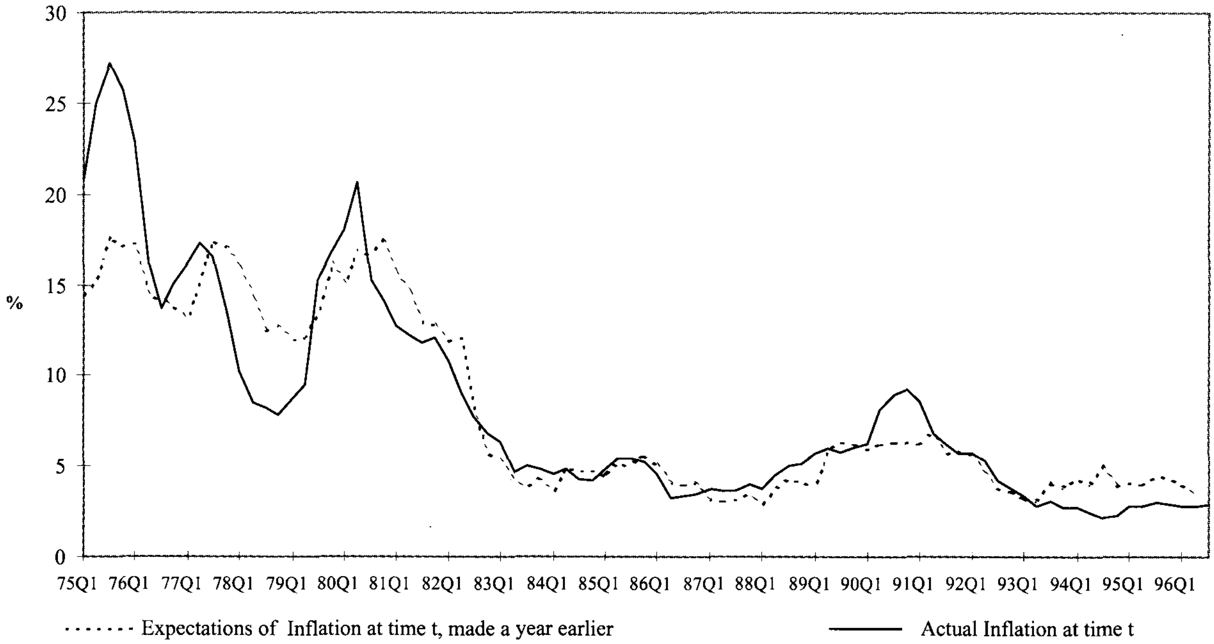
Inflation is affected by past values of inflation, economic agents' expectations of future inflation as well as the output gap. The coefficent on the output gap captures the flexibility of prices: the more flexible are prices, the higher is the value of this coefficient. Substituting out for expectations of future inflation adaptively gives a formulation solely in terms of lagged inflation but such relationships may not be structurally stable. We have derived a direct measure of expected inflation to take account of potential instability from shifts in expectations when estimating the role of the output gap.

Expected inflation is constructed from the Gallup and GFK consumer confidence surveys. These surveys ask a sample of consumers to rank how strong they think inflation will be over the next twelve months. Responses are classified by whether they expect that, in relation to now, there will be a more rapid increase in prices; or that prices will increase at the same rate; or that prices will increase at a slower rate; or that prices will be stable or that prices will fall slightly; or that they don't know. The question on expected inflation has changed twice since the survey has been constructed; hence we construct the quantitative expectations measure over three sub-samples. We quantify the responses by summing the weighted proportions of the sample. The weights are simply determined by a time-series regression of actual inflation on the response proportions (see Pesaran (1988)). This implies that, on average, the weights are the same for actual inflation and expected inflation and the procedure, therefore, imposes unbiasedness on the expected inflation series. Hence tests of rationality are not possible.

The resultant series is shown in Chart 6 alongside actual (RPIX) inflation for the same period. One notable feature arising from comparison of actual and expected inflation rates is the unanticipated shocks due to the large oil price changes in 1974 and 1979. This suggests that the apparent breakdown of the Phillips curves in the 1970s was not entirely due to the lack of an explicit consideration of expectations but partly due to these unanticipated supply-side shocks. The chart also suggests that inflation expectations have behaved adaptively since 1987, lagging behind actual inflation. There does not appear to have been a shift in inflation expectations after 1993 on the introduction of the new monetary arrangements for the United Kingdom, which might have resulted in an increased credibility of economic policy. In particular, inflation expectations rose over 1994 and 1995 whilst actual inflation fell. This result is consistent with measures of inflation expectations derived from a comparison of indexed with conventional bonds. The failure of expectations to adjust to the new monetary regime may reflect the fact that it takes time to establish credibility together with concerns over the impact of sterling depreciation (both after ERM exit in 1992 and in 1995).

Chart 6

Expected and actual rates of inflation (annual rates of change)



4. Empirical estimates of the relationship between the output gap and inflation

We have used our preferred measure of the output gap derived from the production function approach, allowing for a shift in technological progress after 1979, together with our derived measure of inflation expectations to estimate a Phillips curve relationship. It is estimated on quarterly data for 1977Q1–1995Q1. The final form of the equation is derived by starting with a general form with several lags of the output gap and inflation and testing down until all the variables are significant. The final specification is:

$$\Delta_4 p_t = \phi \Delta_4 p^e_{t+4} + (1-\phi)\Delta_4 p_{t-1} + \gamma_1 \Delta\left(y_{t-1} - y^*_{t-1}\right) + \gamma_2\left(y_{t-7} - y^*_{t-7}\right) + \left(\Delta_4 oil_t\right) \tag{5}$$

where $\Delta_4 p_t$ denotes annual RPIX inflation (logs), $\Delta_4 p^e_{t+4}$ expectations of inflation for period t+4 made at the beginning of period t, $y_t - y^*_t$ the output gap at time t, $\Delta\left(y_t - y^*_t\right)$ the change in the output gap from its last quarter's value and $\Delta_4 oil_t$ the rate of oil price inflation.

The underlying form of this equation is explained by New Keynesian theory. Nominal variables are sticky and the output gap is not eliminated in the short run. In addition, the presence of costs in adjusting prices implies that current inflation depends on previous values and expected future values of inflation. Different perceptions of policy, if credible, would be captured in the observed measure of inflation expectations. The oil price inflation term is included to capture errors on expectations made by the agents due to the oil price shocks of the 1970s. The parameters γ_1 and γ_2 measure the flexibility of price adjustment. If prices were more flexible then we would expect these coefficients to be larger; i.e. changes in the demand for output above full capacity lead to faster changes in prices.

The results of estimating this relationship over the period 1977Q1–1995Q1 are shown in Appendix A (Table A.1). The coefficients are all significant at the 5% level. The diagnostic tests reveal no evidence of specification errors apart from slight autocorrelation and non-normality. Autocorrelation in the residuals is increased if the estimation uses data from the early 1970s without the oil price inflation variable. The autocorrelation is thus due to the large temporary supply-side shocks which occurred in the United Kingdom at that time, a period which can be described as the graveyard of the UK Phillips curve. In part, this problem is resolved by including oil price inflation as an explanatory variable in the Phillips curve and in part by excluding data before 1977. Estimates of Phillips curves for the United States, which is a more closed economy and perhaps less subject to supply shocks, tend to find more robust reduced-form relationships (see, for example, Clark et al. (1995a)).

There is a fairly long lag on the output gap in the estimated equation; but, as we discuss later, interpretation of the dynamic response of inflation needs to take account of the endogeneity of lagged and expected inflation. The presence of non-normality in our errors raises the interesting possibility that the effect of the output gap on inflation is asymmetric, with a positive output gap exerting more inflationary pressure than the deflationary pressure exerted by a negative output gap of the same size. Equation errors may in fact be picking up this potential misspecification. We test for this possibility below and discover that allowing for asymmetries makes the residuals more normal.

As a further test of the specification we compare these results on quarterly data to a similar estimation performed on annual data reported in the Appendix (Table A.2). The estimation with annual data has less noise and hence a better fit but is also much less robust because the information content of annual data is too small. Although a comparison between different frequencies is not straightforward, there does seem to be some consistent features of the Phillips curve. In particular, the output gap affects inflation with a similar lag in both quarterly and annual data and, in general, the weights on backward-looking versus forward-looking inflation terms are similar.

The equation appears reasonably well-specified despite the considerable uncertainty as to any point estimate of the output gap. This is because it is the variation in the output gap estimate with respect to inflation that is important in least squares regression. As these different gap estimates move closely together, they presumably move with the true unobserved output gap. Therefore, the variation in the estimated output gaps used to calculate the coefficient in the Phillips curve is close enough to the true variation so as not to induce misspecification. To emphasise this further, we compare the fitted values to the values of actual inflation and expectations of inflation made one year earlier in the table below.

Out of sample forecasts of annual RPIX inflation

t=	95Q2	95Q3	95Q4	96Q1	96Q2	96Q3	96Q4	97Q1	97Q2
Expectations made at t-4	4.07	3.97	4.44	4.38	4.01	3.48	3.67	2.84	2.93
Forecasts	2.63	2.36	2.81	2.66	2.79	-	-	-	-
Actual inflation	2.75	2.93	2.92	2.83	2.79	2.87	-	-	-
Forecast errors	-0.12	-0.57	-0.11	-0.17	0	-	-	-	-

As confirmed by a Chow forecast test at a 1% level of significance, the out-of-sample forecasts are close to the actual values, even though point estimates of the output gap are uncertain (and agents make persistent errors in predicting inflation over this period).

4.1 Interpretation of the estimates

The estimation confirms that the effect of the output gap on inflation is unlikely to be instantaneous and may not just depend on the level of the output gap. The output gap begins to affect inflation after one quarter but this is temporary and subsequent effects appear with a further six-quarter lag. If the output gap is being closed quickly, "speed limit" effects create inflationary pressure more than a year before the level effect feeds through.

We can accept the restriction that the sum of the coefficients on expected and lagged inflation is unity at a 5% level of significance. This restriction means that inflation is dynamically homogenous. The coefficient on lagged inflation is about 0.8 and inflation is characterised as a sticky process.

These features together have an important implication for the inflationary process. If expectations are formed rationally, once we solve out for the expected inflation term, actual inflation is determined by past inflation with a unit coefficient (as well as the discounted path of past, current and future output gaps). This means that when agents are forward-looking, past inflation feeds one-to-one onto current inflation and there is an inherent tendency for the rate of inflation to be explosive. This does not mean that the rate of inflation will necessarily explode, or has no anchor, because the path of future output gaps determines where it will settle. Indeed, as the output gap eventually closes following a shock, the rate of inflation must eventually reach some constant equilibrium value which depends on the whole historical path of output gaps. It is the role of policy to bring inflation back to its long-run target by determining the current and future path of output gaps.

At first glance, it seems that inflation is not particularly sensitive to the level of the output gap with a 1 percentage point (pp) deviation in the gap from one quarter to the next generating a 0.55pp rise in the inflation rate after a quarter due to the "speed limit effect" and then only an additional 0.13pp after a year and a half. However, this interpretation fails to estimate the total effect because it considers only direct effects of the output gap on inflation.

More realistically, past and future inflation are also affected by the output gap and this feeds through to the path of inflation. In particular, the stickiness of inflation can make the long-term inflationary impact of a movement above potential much greater especially when policy is not credible in dampening future inflationary pressure.

To calculate the full effect of the output gap on inflation, we must determine the process by which expectations are formed. To illustrate this, we make two alternative assumptions and determine the inflationary effect of the output gap in each case. First, we assume that expectations are rational. We solve out for expectations of future inflation to find that actual inflation depends on the direct effect of the output gap, as before, as well as on the expected future path of output gaps and past inflation. Assume that the output gap increases by 1pp from an initial value of zero and then dies down gradually without any intervention from the policy maker. This can be interpreted as a simulation of an accommodated temporary increase in the output gap. The rate of inflation jumps immediately by 0.13pp because agents anticipate that the future output gaps are going to widen, and revise upwards their expectations of future inflation. The direct effect of the output gap raises inflation at one quarter, two quarters and seven quarter horizons and, over time, these direct effects reverberate on inflation, gradually disappearing as the gap closes by itself. The feedback of past inflation becomes increasingly dominant and, even after the output gap has closed, continues to drive inflation. Eventually, the rate of inflation settles 2.8pp higher than it was before the increase in the output gap. But this result depends on the speed at which the output gap is closed.

To illustrate the full effect of the output gap for an alternative expectations-generating process, we assume expected future inflation is as described in equation (6):

$$\Delta_4 p_{t+4}^e = \mu \Delta_4 p_{t-1} + (1-\mu)\Delta_4 p^* \qquad (6)$$

where $\Delta_4 p^*$ is the government's annual inflation target and the parameter μ measures how credible the government's target is such that the lower the value of μ the more credible is the target. We repeat the same exercise of increasing the output gap, letting it close by itself over time. Even with a small amount of credibility, the inflation rate eventually returns to its target value (2.5%) but because inflation is so sticky it can take a long time (typically ten years) to be within 0.1pp of its target. The speed of adjustment depends on credibility. With a degree of credibility corresponding to μ = 0.8, the rate of inflation is 3.14pp after 3 years whereas with more credibility (μ = 0.8) the inflation rate is 2.83pp at the same point.

The full effect of the output gap on inflation seems high when we take account of expectations, but it is worth emphasising that this scenario assumes that the policy maker does nothing to the future course of output gaps to bring inflation back to some target value.

5. Key issues

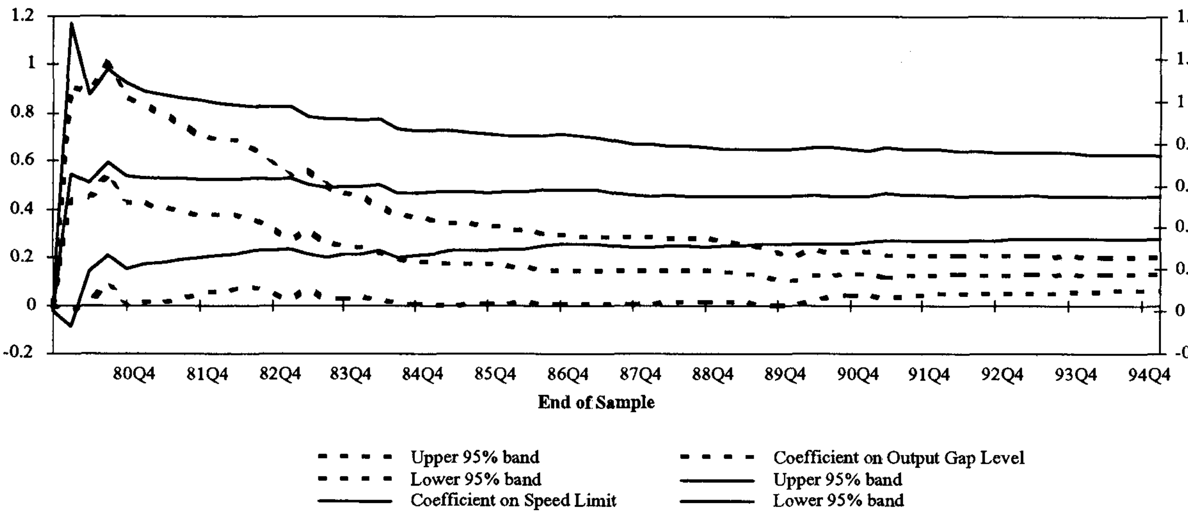
Once a stable relationship with inflation has been estimated, testing features of the framework with which short-term inflationary pressure is generated can be used to shed light on some key issues relevant to forecasting and policy. In this section we present two examples: one which arises from possible shifts in the trade-off between inflation and the output gap and the other which focuses on the bias in estimating inflationary pressure.

5.1 A change in flexibility in the labour markets in the 1990s

In this model, firms set prices as a fixed mark-up on wages and productivity. If workers become more willing to accept cuts in nominal wages as a result of legislation or market conditions, firms would be more free to change prices to absorb any change in demand. More flexible wages would in turn lead to a lower cost in changing prices and raise the parameters representing the trade-off between the output gap and inflation (γ_1 and γ_2).

We tested for a change in labour market flexibility by estimating the parameters γ_1 and γ_2 recursively over the sample. These are shown in Chart 7. Chow tests confirm the null hypothesis that there was no change in labour market flexibility represented by a shift occurring either jointly or separately in these parameters at any quarter from 1986Q1–1994Q4, even at very low levels of significance. It is important to bear in mind that more flexible wages do not necessarily arise from all changes towards more flexible labour markets. For example, lower hiring and firing costs would make it easier for firms to hold wages fixed by altering employment although it would also make it harder for unions to impose their targeted nominal wage.

Chart 7
Recursive estimates of coeeficients capturing flexibility



5.2 Testing for service sector bottlenecks

A further issue is whether sectoral considerations provide information about aggregate inflation, additional to the whole economy output gap. In particular we examine the possibility that the service sector may be more inflationary than other sectors. The aggregate Phillips curve may understate the degree of inflationary pressure because the service sector is more intensive in skilled labour. When the service sector is operating at a greater over-capacity than elsewhere it acts as a bottleneck generating extra inflationary pressure which would not be captured by an aggregate Phillips curve.

We can test this hypothesis with our aggregate Phillips curve estimates because it implies that the service sector output gap would have extra explanatory power in explaining inflation above the aggregate output gap and expected and past inflation terms. There is one complication to this straightforward test as the service sector output gap might affect inflation with a different lag than the rest of the economy. This is more in keeping with the idea of the service sector acting as a bottleneck and creating inflationary pressure before it arises elsewhere and even when there is significant downward pressure on inflation from other parts of the economy.

We deal with this by estimating a Phillips curve with several lags of both the aggregate output gap and the service sector output gap and testing down to see if the service sector output gap terms are significant. Allowing for a general lag structure takes account of the possibility of bottlenecks without imparting bias. The final specification, where none of the variables can be rejected at a 30% level, is given in the Appendix (Table A.3). A joint test on the significance of the service sector output gap terms rejects the hypothesis that they are insignificant and that there is no aggregation bias at even very low levels of significance. Although the diagnostic tests indicate some misspecification in the final estimates, our simple test is very supportive of the view that the service sector does act as a bottleneck. The results also indicate that the effect on inflation of allowing for differences in the service sector output gap can be sizeable. This can be illustrated with a simulation exercise. We compare the effects of a uniform rise in the output gap of 1pp everywhere with a 2pp rise in the service sector output gap accompanied by a lower rise or even a fall in the gap elsewhere. Both simulations imply that the output gap in aggregate increases by 1pp but the effect on inflation is quite different. Assuming that expectations are formed according to equation (6) with the degree of credibility given by $\mu = 0.4$, inflation is 3.5pp after three years when the rise is uniform and 4.7pp when the service sector acts as a bottleneck.

It is not easy to model these bottlenecks in an aggregate Phillips curve because they occur with a different lag and it is difficult to distinguish the extra inflationary impact of the service sector from the uniform tightening that occurs across all sectors. Estimating Phillips curves for each sector should help resolve this bias, but this requires assumptions about the relevant production functions and levels of factor inputs.

6. Speed limit effects and asymmetric effects in the Phillips curve

Short-term monetary policy would be easier if the effects of the output gap on inflation (given expected future and past inflation) were proportional even after a lag. Unfortunately, it may be more plausible to assume that the rate of closure of the gap itself (and whether it is positive or negative) also determines the amount of inflationary pressure a gap generates. In this section we examine how to measure these "speed limit" and asymmetric effects and whether they play a significant role in the UK economy.

6.1 Speed limit effects

It seems plausible that the rate of closure of the output gap can create inflationary pressure above that of the level of the output gap itself. It would be interesting to know how much difference this "speed limit effect" makes to the operation of monetary policy. If we can quantify the policy relevance of the "speed limit effect" then we can compare it across countries or scenarios. We would like to do this by measuring the extent to which the output gap cannot be closed because the "speed limit effect" evokes the threat of rising inflation. However, in a single equation this is not possible when inflation depends on past and expected future inflation.

The difficulty with measuring the "speed limit effect" as the coefficient on the change in the output gap in a single equation is that it depends on how we present the dynamics of the output gap and inflation. This can be explained by considering the following Phillips curve:

$$\Delta p = \phi \Delta p^{e}_{+1} + (1-\phi)\Delta p_{-1} + \gamma_1 \left(y_{-1} - y^{*}_{-1}\right) + \gamma_2 \Delta\left(y - y^{*}\right)$$

The same Phillips curve can be re-written as:

$$\Delta p = \phi \Delta p^{e}_{t+1} + (1-\phi)\Delta p_{-1} + (\gamma_2 - \gamma_1)\Delta(y - y^{*}) + \gamma_1 (y - y^{*})$$

with a different coefficient on the change in the output gap. If we measure "the speed limit effect" as the coefficient on the change in the output gap we would find that the same Phillips Curve gives different measures of the speed limit effect (γ_2 and $\gamma_2 - \gamma_1$) depending only on how it is written.

The idea of a "speed limit effect" arises from a scenario where the output gap cannot be easily closed because of the threat of rising inflation. Thus Turner (1995) defines it with the following sentence: "If output begins by being below potential, then the greatest proportion of this gap that can be closed to keep next period's inflation from increasing is an inverse measure of the speed limit effect." If closing all the gap still keeps inflation constant then there is no "speed limit effect". If closing even a very small amount of the gap means that inflation starts to accelerate, then the 'speed limit effect' is strong.

Once we accept that current inflation is determined by its past and expected future values, this definition can be misleading. If inflation has just risen sharply (e.g. because of a supply shock), and is autocorrelated, it may be difficult, if not impossible, to stop it rising via demand management. Conversely, if inflation expectations are relatively constant then demand pressures on inflation will generally be much less. This measure of the speed limit effect which is based only on its threat to current inflation would report that the speed limit term was more important in the first case than the second even though the same Phillips curve underlies both scenarios. This shows that the intuition behind the "speed limit effect" can only be measured relative to where inflation is now; where it has been recently; and where it is expected to go.

The problem with the definition is that, if inflation is affected by past inflation and expected future inflation, it is more plausible that the policy maker would prefer a more gradual adjustment of inflation to its target value. These considerations make it clear that in order to make this policy experiment valid, we need to know the optimal path of the output gap over time so that we could examine the influence of the speed limit term on this path. In order to determine this optimal path we would have to consider not only the objectives of the policy maker but also other aspects of the transmission mechanism of monetary policy not captured by the Phillips curve equation.

We have explained why a measure of the policy relevance of a speed limit term depends crucially on the objectives and constraints of the policy-maker. This needs a fuller model of the economy and a historical context. We cannot adapt the single equation approach to adequately measure the "speed limit effect".

6.2 Asymmetric output gaps

It may be that the Phillips curve is asymmetric: a positive output gap exerts more inflationary pressure than the deflationary pressure exerted by a negative output gap of the same size. This can have very different implications, compared to a symmetric Phillips curve. For example, it implies that policy makers should err on the side of caution during a recovery because the costs of making mistakes and correcting past mistakes in monetary policy are higher than for a symmetric Phillips curve.

As Clark et al. (1995a) pointed out formally, a Phillips curve such as ours rules this possibility out for two reasons. First, the elasticities of positive and negative output gaps in affecting inflation are equal. Secondly, the OLS estimates of the output gap from past data place the same weight on positive as on negative deviations of actual output from estimated potential output implying that the economy has spent the same amount of time above and below potential. If, on the other hand, the true Phillips curve were asymmetric, output would have to spend more time in recession to make inflation (taking account of past and expected future inflation) stationary and the true output gap would be less than the calculated output gap $\left(y_t - y_t^*\right)$ by a constant α.

Incorporating both these considerations implies that we should estimate a Phillips curve of the following form:

$$\Delta_4 p_t = \phi \Delta_4 E_t p_{t+4} + (1-\phi)\Delta_4 p_{t-1} + \gamma_1 \Delta w_{t-1}\left(y_{t-1} - y_{t-1}^* - \alpha\right) + \gamma_2 w_{t-7}\left(y_{t-7} - y_{t-7}^* - \alpha\right) + \left(\Delta_4 oil_t\right) \quad (7)$$

where the weight w_t at each time t has to be estimated and itself depends on whether the output gap at this time $\left(y_{t-1} - y^*_{t-1} - \alpha\right)$ is positive or negative. The difficulty arises in estimating these weights along with the coefficient α. One simple way of doing this is to follow Turner (op cit) in assuming that the two forms of asymmetry (the weights and the coefficient) are related so that they can be estimated together. The estimated degree of asymmetry is that which best fits the data, and if this implies that the Phillips curve is symmetric, it should be preferred to a range of asymmetric alternatives.

Turner assumes that the weight on a negative output gap is a fixed proportion of the weight on a positive output gap of the same size, and calculates this proportion. We assume that the size as well as the sign of the output gap determine its weight. This exhibits the plausible property that inflation is stubborn and becomes increasingly harder to reduce when it is lower (and expectations of future inflation do not change) but easy to raise when it is high. In Appendix B we illustrate in more depth how this weight can be estimated and what it implies for inflation.

We found that the Phillips curve of best fit is not symmetric. Rather, we estimate that the true output gap is 0.55pp less than the calculated output gap. This means that when the true output gap is 2%, the inflation rate is 0.25pp more than the average rate; whereas when the output gap is -2%, the inflation rate is 0.20pp below the average rate. In addition we find that the asymmetric output gap terms are fairly significant in a regression with past inflation, expected inflation and symmetric output gap terms. The likelihood ratio test indicates that the asymmetric terms are significant at a 20% level. Although this latter test is biased against asymmetry, it supports our overall finding that there is some very mild asymmetry in the UK Phillips curve.

The results of re-estimating the Phillips curve with the asymmetric output gap terms are reported in the Appendix (Table A.4). The equation fits the data better than the symmetric Phillips curve as we should expect but without engendering any great changes in the other parameter estimates. There are still elements of skewness in the residuals but less than with the symmetric Phillips curve. These results do not necessarily contradict Turner who found that the Phillips curve for the United Kingdom was symmetric using annual data. This is because he was testing symmetry against the nearest alternative that a positive output gap exerted at least twice as much inflationary pressure as a negative output gap of the same size whereas our results estimate the degree of asymmetry to be too subtle to discern from his tests.

Conclusions

In this paper we have discussed the use of measures of the output gap and illustrated the diversity of estimates that can be obtained. We have shown how more theory-based estimates can be made and associated with a probability distribution for the output gap. We have further shown that, when coupled with an explicit measure of expected inflation, these output gap measures can be useful in estimating a relationship between the output gap and inflation. By considering the role of expected future inflation separately from the feedback of past inflation, we have a richer understanding of the dynamic effect of the output gap on inflation.

Using the estimated relationship between the output gap and inflation, we have shown that no change in labour market flexibility in the 1990s is apparent. There is some evidence, however, of the service sector acting as a bottleneck, generating more inflationary pressure than the rest of the economy. We have also discovered some mild asymmetry in the relationship between the output gap and inflation which implies that a 2% positive output gap exerts 0.05pp more inflationary pressure than the deflationary pressure exerted by a 2% negative output gap and that OLS estimates of the output gap are on average 0.55pp too high.

The simple idea of a "speed limit effect" asks whether the speed at which the output gap is closed has any bearing on the inflation rate; more strictly it asks what is the fastest rate of closure of the gap that can occur without an increase in the rate of inflation. Our discussion reveals that full appreciation of a "speed limit effect" needs to deal with the endogeneity of inflation expectations and the output gap itself. Inspection of the single equation Phillips curve is insufficient to answer this question. A proper conclusion requires a model of the output gap and the role of policy in influencing this gap, since the response of current inflation depends on past and future output gaps, and not only on the current level. Even without any bottlenecks or asymmetric effects, inflation may rise before any output gap is closed, or continue to change when output is growing at trend.

In conclusion, empirical estimation of expectations-augmented Phillips curves provides a role for the output gap in predicting inflation and the relationship can be robust when not faced with external or supply shocks. The output gap does, therefore, give us a simple method of judging short-term inflationary pressures alongside other modelling approaches. However, we emphasise that the single equation chosen has fairly limited use without some additional information about the determination of the output gap itself. The Phillips curve is not a reduced form in the proper sense of the term; it is just one part of a wider system.

Appendix A

Table A.1

An estimated UK Phillips curve equation on quarterly data[1]

	Annual RPIX inflation **Sample period: 1977Q1–1995Q1**
Output gap$_{t-7}$	0.279** (3.62)
Output gap$_{t-1}$ - output gap$_{t-2}$	0.5455** (5.42)
Expected inflation$_{t+4}$	0.217** (5.08)
Lagged inflation	0.783** (5.08)
Oil price inflation (spot price of Brent Crude)	0.0102** (3.61)
Equation diagnostics $\bar{R}^2$ DW SE Serial correlation (F) Heteroskedasticity (F) Functional form (F) Normality	 0.603 1.43 0.008604 LM(1) = 5.56 3.00 0.005 27.0**
Restriction that coefficients on expected and lagged inflation sum to unity	F=3.69*
Chow Forecast Test for 1995Q2–1996Q2 (F)	0.1

* Indicates statistically significant at 10% level. ** Indicates statistically significant at 5% level.

[1] t-statistic in brackets.

Table A.2

An estimated UK Phillips curve equation on annual data[1]

	Annual RPIX inflation Sample period: 1977–1994
Constant	-0.00639 (1.2)
Output gap_{t-1}	0.1404* (1.8)
Expected $inflation_{t+1}$	0.3696** (7.3)
Lagged inflation	0.6362** (12.0)
Equation diagnostics	
$\overline{R}^2$	0.961
DW	2.33
SE	0.00898
Serial correlation (F)	LM(1) = 1.17
Heteroskedasticity (F)	0.85
Functional form (F)	0.002
Normality	9.0**
Restriction that coefficients on expected and lagged inflation sum to unity	F = 0.01

* Indicates statistically significant at 10% level. ** Indicates statistically significant at 5% level.

[1] t-statistic in brackets.

Table A.3

An estimated UK Phillips curve equation on quarterly data allowing for service sector bottlenecks[1]

	Annual RPIX inflation Sample period: 1977Q1–1995Q1
Aggregate output gap terms at lags $_{t\text{-}1,\ t\text{-}2,\ t\text{-}4,\ t\text{-}5,\ t\text{-}7}$	0.263 (3.84), -0.441 (-3.58), -0.683 (-2.31), 0.146 (2.26), 0.120 (1.60)
Service sector output gap terms at lags $_{t\text{-}1,\ t\text{-}2,\ t\text{-}4,\ t\text{-}5}$	0.228 (0.837), -0.082 (-0.282), 1.03(3.72), -0.841 (-3.39)
Expected $inflation_{t+4}$	0.202 (4.52)
Lagged inflation	0.798 (4.52)
Oil price inflation (spot price of Brent Crude)	0.0103 (4.04)
Equation diagnostics	
R^2	0.695
DW	1.28
SE	0.007648
Serial correlation (F)	LM(1) = 19.42**
Heteroskedasticity (F)	1.375
Functional form (F)	0.513
Normality	12.0**

* Indicates statistically significant at 10% level. ** Indicates statistically significant at 5% level.

[1] t-statistic in brackets (affected by multicolinearity despite estimation with reparameterisation).

Table A.4

An estimated asymmetric UK Phillips curve equation on quarterly data[1]

	Annual RPIX inflation Sample period: 1977Q1–1995Q1
Asymmetric output gap $_{t-7}$	0.111 (4.07)**
Asymmetric output gap $_{t-1}$ - asymmetric output gap $_{t-2}$	0.576 (7.20)**
Expected inflation $_{t+4}$	0.206 (5.63)**
Lagged inflation	0.794 (5.63)**
Oil price inflation (spot price of Brent Crude)	0.0104 (4.50)**
Equation diagnostics	
R^2	0.718
DW	1.83
SE	0.00687
Serial correlation (F)	LM(1) = 0.15
Heteroskedasticity (F)	2.78**
Functional form (F)	0.289
Normality	15.99**

* Indicates statistically significant at 10% level. ** Indicates statistically significant at 5% level.
[1] t-statistic in brackets.

Appendix B: Asymmetric effects

This technical appendix describes our method of estimating asymmetries in the relationship of output gaps to rates of inflation. This improves on past methods because it is less arbitrary and allows the size as well as the sign of the output gap to determine its effect on inflation. For ease of exposition, consider the simplest Phillips curve of the symmetric (linear) form:

$$\Delta p = c_0 + c_1(y - y^*) \tag{7}$$

As we described earlier, this differs from an asymmetric output gap in two ways. First, it imposes the same coefficient on positive and negative output gaps. Secondly, the OLS-estimated output gap ($y - y^*$) places the same elasticity on positive as on negative deviations of actual output from the estimated potential output, implying that the economy has spent the same amount of time above and below potential.

This second point should be considered in more depth. If the true Phillips curve were asymmetric and (quite reasonably) we assume that actual inflation is stationary,[3] then the economy

3 In our more general Phillips curve equation, we are assuming that actual inflation minus the effects of expected and lagged inflation is stationary.

would have spent more time below than above potential and the actual output gap in each period would be more negative than the calculated output gap ($y - y^*$) by a constant α.

Taking account of both of these aspects means that we can define the weighted output gap (*wtog*) as the calculated output gap adjusted for miscalculation and allowing for different weights for positive and negative values:

$$wtog_t = w_t(y - y^* - \alpha) \tag{8}$$

where w_t is the weight on the true output gap at time t and can itself depend on the output gap at time t. This weighted output gap will then have a linear (symmetric) effect on inflation.

Note that the sum of the weighted output gaps (just like the calculated output gaps) over the sample are zero; although the economy spends more time below trend, it affects inflation more when it is positive than when it is negative and the rate of inflation is stationary. This is true, even though the sum of the unweighted output gaps will be negative. Thus:

$$\sum_{t=0}^{n} w_t(y - y^* - \alpha) = 0 \tag{9}$$

Combining implies that, allowing for an asymmetry, we should estimate a Phillips curve of the following form:

$$\Delta p_t = c_0 + c_1 w_t(y - y^* - \alpha) \tag{10}$$

Remember that the weight at time t depends on whether the calculated output gap is greater than or less than α. The difficulty in estimating this Phillips curve is hence that the coefficient α and the weights should be estimated together.

Turner (op cit) suggested a simple test by assuming that the weight on a positive output gap is n times a negative output gap of the same absolute size. If the scale factor, n, is equal to one, then $\alpha = 0$ and the Phillips curve is linear (symmetric) as in equation 7.

For each value of the scale factor, we can calculate the value of α from equation 9 by a grid search. For this calculated α, the asymmetric output gaps are used in the estimation of 10. The fit of this estimation is compared with other estimations for different values of the scale factor. The scale factor which gives the best fit is chosen as best reflecting the DGP. If this optimal scale factor is one, then we have found that the symmetric Phillips curve is preferred to a range of alternatives.

There are two problems with this approach. First, for each scale factor, we have to find the optimal α by trial and error. Apart from being time-consuming, there may be more than one α which satisfies for any given scale factor, n. Secondly, as Turner himself notes, the form of the asymmetry is discrete whereas it may be that the weight an output gap receives in determining inflation depends on its absolute size as well as on its sign such that inflation becomes increasingly harder to reduce the lower it is.

Our alternative specification of the weights which overcomes both these problems is the exponential form:

$$w_t = \exp\rho(y - y^* - \alpha)$$

Here the more positive *and* the larger the output gap, the more weight it receives in determining inflation. The scale factor ρ plays an analogous role to the scale factor n in Turner's method: if it is zero then the Phillips curve is linear. Now, using equation (9), there is a unique, readily-calculable value of α for each ρ given by:

$$\alpha = \frac{\sum_{t=1}^{n}(y - y^*)e\rho(y - y^*)}{\sum_{t=1}^{n}e\rho(y - y^*)}$$

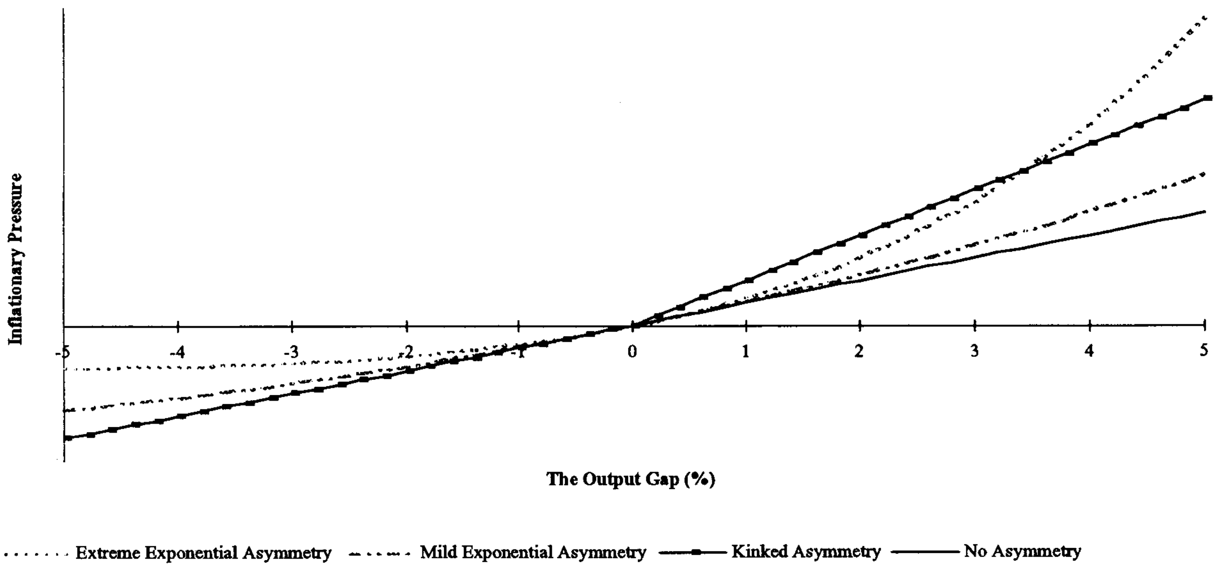
Thus for any value of ρ and given the data we can write down the value of α. This means that estimating the optimal degree of asymmetry is much more straightforward. The optimal degree chosen can then be used to derive policy implications.

Implications for inflation

Chart 8 compares the relative effect on inflation of an output gap according to the kinked weights used by Turner; our method for two degrees of asymmetry and a symmetric system of weights. The inflationary implications of each system differ. According to kinked weights, the elasticity of a given output gap only changes once it crosses full capacity. In contrast, with our weights, to reduce inflation to ever lower levels, ever deeper recessions are necessary.

Other specifications for the weights have been suggested in the literature. Clark et al. (1995b) illustrate the strengths and weaknesses of the different approaches. The system of weights they prefer is the method employed by Chadha, Masson and Meredith (1992) which assumes that the weights are given by a function which has similar properties to our exponential weights: a large negative output gap exerts relatively less deflationary pressure. However, for reasons similar to the kinked weights, this is not as easy to calculate as our system of weights.

Chart 8
Asymmetric output gaps



References

Batchelor, R.A. and A.B. Orr (1988): "Inflation Expectations Revisited". *Economica*, Vol. 33, pp. 317-31, August.

Beveridge, S. and C.R. Nelson (1981): "A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to the Measurement of the Business Cycle". *Journal of Monetary Economics*, No. 7, pp. 151-74.

Canova, F. (1993): "Detrending and Business Cycle Facts". *CEPR Discussion Paper* No. 782, June.

Carlson, J.A. and M. Parkin (1975): "Inflation Expectations". *Economica*, Vol. 42, pp. 123-38, May.

Chadha, B., P. Masson and G. Meredith (1992): "Models of Inflation and the Costs of Disinflation". *IMF Staff Papers*, Vol. 31, pp. 395-431, June.

Clark, P., D. Laxton and D. Rose (1995a): "Asymmetry in the Inflation-Output Nexus: Issues and Evidence". *IMF Working Paper* No. 95/76.

Clark, P., D. Laxton and D. Rose (1995b): "Asymmetric Effects of Economic Activity on Inflation: Evidence and Policy Implications". *IMF Staff Papers*, Vol. 42, No. 2, pp. 344-74, June.

Dicks, M. (1996): "Output Gaps and All That". Lehman Brothers Global Economics, *International Economics*.

Fisher, P.G. and J.D. Whitley (1996): "Macroeconomic Model Projects at the Bank of England". Bank of England (forthcoming).

Giorno, C., P. Richardson, D. Roseveare and P. van den Noord (1995): "Potential Output, Output Gaps and Structural Budget Balances". *OECD Economic Studies* Vol. 1, No. 24, pp. 167-209.

Layard, R., S. Nickell and R. Jackman (1991): "*Unemployment: Macroeconomic Performance and the Labour Market*", OUP.

Nelson, C.R. and H. Kang (1991): "Spurious Periodicity in Appropriately Detrended Time Series". *Econometrica*, Vol. 49, No. 3, pp. 741-51, May.

Nelson, C.R. and C. Plosser (1982): "Trends and Random Walks in Macroeconomic Time Series". *Journal of Monetary Economics*, No. 10, pp. 139-62.

Pesaran, M.H. (1988): "*The Limits to Rational Expectations*", Basil Blackwell.

Quah, D. (1995): "The Relative Importance of Permanent and Transitory Components: Identification and Some Theoretical Bounds". *Econometrica,* Vol. 60, No. 1, pp. 107-18, January.

Roberts, J.M. (1995): "New Keynesian Economics and the Phillips Curve". *Journal of Money, Credit and Banking*, Vol. 27, No. 4, pp. 975-84, November.

Turner, D. (1995): "Speed Limit and Asymmetric Inflation Effects from the Output Gap in the Major Seven Economies". *OECD Economic Studies* Vol. 1, No 24, pp. 58-87.

Comments on: "The output gap and inflation – experience at the Bank of England" by Paul G. Fisher, Lavan Mahadeva and John D. Whitley

by Frank Smets

This paper nicely summarises some of the research done at the Bank of England on the relationship between inflation and the output gap. The paper consists of three sections. The first briefly discusses the concept of a Phillips curve and the various inputs needed for estimation, such as measures of the output gap and inflation expectations. The second presents estimates of a Phillips curve for the United Kingdom and discusses the implications of a widening output gap for future inflation. The third focuses on a number of related issues, in particular whether structural breaks can be detected, whether service sector bottlenecks have additional power, whether speed limit effects are important, and whether there are asymmetries in the link between inflation and the output gap. My comments will be limited to, first, estimation of the output gap and, second, some of the implications thereof for estimation of the Phillips curve.

Clearly, the usefulness of a Phillips curve for policy purposes critically depends on the reliability of the estimated output gap. The authors show quite convincingly that simple statistical methods, such as a log-linear trend or the HP filter, are very sensitive to the choice of the estimation period and the smoothing parameter in the HP filter. They acknowledge that there is a random walk component in output, but reject statistical models that take that explicitly into account on the argument that there are a million ways such a series can be decomposed into a permanent and transitory component. While this is true, it should be recognised that using economically sensible identifying assumptions and additional information can narrow down the set of possible decompositions. Examples of such fruitful approaches are the extended HP filter, the Blanchard-Quah methodology and the Unobservable Components model by Kuttner.

Instead the authors rely on a more structural approach which consists of using labour and capital inputs in a Cobb-Douglas production function and estimating the trend in productivity residuals using a constant and a linear time trend. The production function approach has its own problems related to the considerable data requirements, the possibility of introducing specification error because of, for instance, the assumed form of the production function and the difficulty in calculating confidence bands. More importantly, however, by fitting a linear time trend on the productivity residual, the authors run into the same kind of problems as the simple statistical methods they dismiss. It is a pity they do not report the diagnostics of the linear time trend regression, but I strongly suspect that the Durbin-Watson statistic is very low, indicating that the productivity residuals have a random walk component. If this is the case then not only will the regression on the time trend result in spurious correlation, it will also affect the calculation of the confidence bands. Basically one gets the same type of problems as in regressing output directly on a linear time trend. This is an old critique of the traditional production function approach (see, for instance, Schwert and Plosser (1979)).

The next question is then whether the estimation of the output gap matters for estimation of the Phillips curve. The authors find that it is mainly the change in the output gap that has a significant effect on inflation and, therefore, argue that it does not matter very much which output gap one uses, because they all move in a very similar way. My conjecture here is that this result may be a direct implication of the misspecification of the output gap. If there is still an important random walk component left in the estimated gap which has more to do with supply than demand conditions, then it should not be surprising that differencing helps in getting significant effects, because this implicitly filters out the low frequency or random walk component. To check whether this is what is going on, one should analyse the productivity residuals a bit further. Two pieces of evidence from Gerlach and Smets (1997) suggest that this is a possible explanation. Thus we find that the United Kingdom has the largest variance of random walk shocks among all the G7 countries. Second, once we filter out this quite variable potential output series, we find that the level of the output gap has a quite strong impact on inflation in the United Kingdom. Moreover, we tested for acceleration effects in our model but did not find any evidence of such effects.

References:

Gerlach, S. and F. Smets (1997): "Output gaps and inflation: evidence from the G7 countries". BIS *mimeo.*

Plosser, C. and G. Schwert (1979): "Potential output: its measurement and significance: a dissenting opinion". *Carnegie-Rochester Conference Series.*