Measuring Excess Demand and its Impact on Inflation

M.PHIL THESIS IN ECONOMICS

Jennifer Castle*
Nuffield College
University of Oxford
Jennifer.castle@nuffield.oxford.ac.uk

Trinity, 2003

* I am most indebted to David Hendry, whose help and guidance throughout my research was invaluable. He was a constant source of inspiration and encouragement. I am very grateful to Sonja Keller and Carlos Santos for their continual advice and support. It was a joy to participate in our research group. I would also like to thank Chris Bowdler and Neil Shephard for their advice and helpful comments. Finally, my thanks extend to the ESRC for their financial support.

Measuring Excess Demand and its Impact on Inflation

Jennifer Castle

Nuffield College
University of Oxford

Trinity, 2003

Abstract

Excess demand pressures play a fundamental role in macroeconomic policy and analysis. The output gap, as a proxy for excess demand, is widely used but is notoriously difficult to measure due to it being a latent variable. This thesis discusses the most prominent measures of the output gap, analysing their advantages and drawbacks. The use of artificially generated data enables explicit evaluation of the performance of univariate 'gap' measures, exposing the degree of disparity between measures. Analysis of the production function approach suggests modelling output in a dynamic setting in order to overcome the problems of systematic and substantial measurement errors in stock variables. The empirical application of the paper assesses the importance of excess demand in explaining inflation. Inflation is modelled in a dynamic single equation framework in which a general to specific modelling strategy is used, encompassing all relevant theories. The output gap, based on a composite measure using principal components analysis, is found to have a substantial impact upon inflation. Furthermore, the paper forecasts inflation over the 1 and 4 quarter horizon using a broad variety of forecasting models. The importance of an accurate, unbiased estimate of excess demand at the forecast origin is emphasized. Inference regarding the business cycle is integral to macroeconomics and this thesis aims to illustrate the importance of the output gap from a variety of angles.

Table of Contents

List of Figures

List of Tables

1 Introduction

The output gap, measuring the difference between actual output and the potential level consistent with full employment of resources in the economy, is one of the most widely used concepts in macroeconomic policy analysis. Excess demand is a significant factor driving inflation and an understanding of the gap, which plays an important role in the conduct of monetary policy, is essential if we are to promote macroeconomic stability. On the fiscal side, the amount of spare capacity in the economy provides an indicator of short term transitory influences that can be used to isolate the impact of cyclical factors on the budget, enabling policy-makers to sustain a stable fiscal policy over the economic cycle. Also, as potential output constitutes the best composite indicator of the supply side of the economy, it plays a key role in long-term policy strategies. A high and stable growth rate of potential output is a precondition for achieving a strong, sustainable non-inflationary growth path. However, whilst the concept of excess demand enjoys a prominence in the literature, it is highly contentious due to it being a latent variable. As there is no clear consensus on how to measure the gap, there is a pressing need for a deeper understanding of the pressures of excess demand and the measurement problems associated with it.

Many users of potential output and the output gap do so blindly without considering the properties of the method used to extract the latent variable. There are a wide variety of methods employed to calculate the output gap, all of which have different properties and many pitfalls. Often linear trends or Hodrick-Prescott filters are used because of their ease, particularly if the researcher is not interested in the output gap per se. Yet inference regarding the gap is sensitive to the measure used. In the absence of prior beliefs based upon economic or statistical reasoning concerning the most appropriate measures of the cycle, conclusions regarding the current state of economic activity can differ widely. A recognition of the limitations of these measures in proxying such a complex variable is requisite. This thesis is designed to give an exposé of the various methods used and their properties in order to provide some guidance as to which measures of excess demand are appropriate. The paper also directly measures the impact of excess demand on inflation by modelling and forecasting UK inflation within a single equation framework, using quarterly data over the past 35 years.

The structure of the thesis is as follows. Chapter 2 provides a review of the most common univariate and multivariate methods of calculating the output gap, discussing their advantages and drawbacks. The section also examines various issues including disaggregation, the equivalence between moving averages and the seasonal adjustment literature, real-time estimation, asymmetries and changes in potential output. Chapter 3 generates real output data as the sum of potential output and the output gap. By generating the DGP, various univariate detrending procedures can be assessed in relation to the known output gap, enabling explicit analysis of the behaviour of these measures when structural breaks occur. Chapter 4 calculates various measures of the output gap for UK data over the period 1965q1-2002q2. Particular attention is given to the production function approach. Firstly, a standard growth accounting model is used and secondly, the production function is estimated in a dynamic framework because of the substantial and systematic measurement errors in stock variables. The section also forms a composite measure of the gap using principal components analysis. The chapter concludes by providing a first pass at a comparison of the methods in terms of the cycles' characteristics, correlations and cointegrating relations between the measures. Chapter 5 of the thesis assesses the impact of the output gap in an inflation equation. Various measures of the output gap are considered in an attempt to quantify the importance of excess demand in determining inflation. Forecasts of inflation for the period 1998q1-2002q2 are evaluated in chapter 6, with the aim of appraising whether a good understanding of excess demand will lead to improved inflation forecasts. Chapter 7 concludes.

The thesis endeavours to assess the importance of excess demand from a variety of angles. Whilst the paper does not offer a solution to the problem of estimating this latent variable, it does expound the problems with current methods and highlights the importance of excess demand, both in terms of its measurement and its transmission mechanism as a major driving force of inflation, thus providing the groundwork for future research on measuring excess demand.

2 Literature Review

There is no unanimous definition of potential output. From a purely statistical viewpoint, potential output is thought to be trend output. From a theoretical perspective, potential output is based on the supply side of the economy and is often defined as the production level at normal utilization of factors of production at the current state of technology. This reflects the idea that potential output is akin to sustainable long-term growth, which implies that the output gap (the transitory component of output) is a consequence of demand shocks. Due to the presence of nominal rigidities, a demand shock will cause output to differ from its supply side level, but as these begin to weaken and prices adjust, the transitory shocks will dissipate and output will revert to its long-run potential. Hence, potential output is the steady state level of output associated with the long-run aggregate supply curve. The definition of potential output differs depending upon the time horizon being examined. In the short-run, physical capital is assumed to be fixed and the gap is determined by how much demand can develop without inducing supply constraints and subsequent inflationary pressures. In the medium-term, investment is assumed to be endogenous implying that a demand expansion may be accommodated. In the long-run, full employment potential output is primarily determined by technological progress and growth of potential labour.

Before proceeding, a note on the terminology used is required. Output in period t is given as y_t, potential output in period t is given as y_t^* and the output gap, y_t^{gap}, is calculated as $(y_t - y_t^*)$. It is assumed that we can use the terms 'potential' and 'trend' output interchangeably but the definition will be clarified by the context in which they are used. Statistical methods of calculating the gap deliver 'trend' output, whereas the production function method produces an estimate of 'potential' output. Also, by custom, the term 'cycle' is used to represent the gap even though it is not periodic.

2.1 Alternative Methods of Estimating the Output Gap

A large literature has developed addressing the question of how trends and cycles should be extracted from data series but there is still no clear consensus. Different methods give rise to variations in the cyclical component in terms of duration, amplitude, autocovariance function, spectrum and whether the cycle is stationary or not.

There are two methods of statistically detrending output, filtering and smoothing. Filtering equates to one-sided estimation. This just relies on backward information and is therefore used for policy-making, but is less accurate than smoothing (two-sided estimation) which uses both backward and forward information. For policy-making, smoothing requires forecast estimates. Note that estimates of the gap from one-sided filters are larger during accelerations because they are purely backward looking. Therefore, trend output is estimated to be lower than it is when the future peak is incorporated. One-sided filters induce a phase shift, which can distort the timing of business cycles. However, 'smoothers' have the problem that future events determine the current path of the estimate, which can lead to pre-recession booms (or vice versa) induced by the smoothing procedure. One measure of robustness is to examine whether estimates of the gap are likely to change with new observations by testing whether one-sided and two-sided estimates differ significantly.

One of the first methods of estimating potential output, and one that is still commonly used, is a linear trend. Potential output is simply a deterministic function of time and the gap is calculated as the residual, implying that all shocks are demand shocks. The trend and cycle will be uncorrelated. There is a voluminous literature on the question of deterministic versus stochastic trends, but the implication of a deterministic trend that potential output is completely predictable with a constant growth rate is theoretically unpalatable.[1] Fitting a deterministic trend to a stochastic series will give erroneous conclusions. Nelson and Kang (1981) observe that assuming a process is trend stationary and hence detrending by regressing on time, when in fact the process is difference stationary, will produce apparent evidence of periodicity in the residuals which is not a property of the underlying process. Also, if used in forecasting, a linear trend has the implication that the long-run forecast error variance converges to a fixed value. In practice, it should grow as the forecast horizon increases.

A segmented linear trend is often used to account for differences in the trend growth rate over time. However, this relies on the ability to identify when breaks occur in the underlying potential of the economy, which is nontrivial. Also, a break is likely to feed through to the underlying potential level of output slowly and will not be accurately captured by cutting the sample at specific dates. For example, a positive technology shock will take time to filter into potential output as research and development, learning, training, time to build and habit formation occurs. Also the aggregation of technology shocks to individual sectors will lead to a smooth diffusion into potential output.

Difference filters are very simple methods of detrending output based on the assumptions that underlying potential output follows a random walk with no drift, the cycle is stationary and $E\left(y_t^* y_t^{gap}\right) = 0$. The stochastic trend is given as $y_t^{*(\Delta)} = y_{t-p}$, where p is the order of the filter.[2] This is a simple and intuitive method, especially as annual

[1] See Chapter 3.1 for a review of the literature examining the question of output persistence. A deterministic trend could be thought of as a limiting case of the general stochastic form, where the variance of the error term is zero.

[2] Often a seasonal filter is used, where $p = 4$.

growth rates may be thought to proxy productive capacity pressures. However, the filter does not match conventional business cycle frequencies. Also, the filter re-weights frequencies by emphasizing higher frequencies and down-weighting lower frequencies, leading to a volatile cyclical component. The filter is not symmetric but introduces a phase shift, given by $\frac{\pi}{2} - 2\omega$ for $p = 4$, where ω is the frequency measured in radians $(-\pi \leq \omega \leq \pi)$. There is also a time shift, given by the phase shift divided by the frequency, and this differs for cycles of varying periodicity. The method can produce negative 4th-order serial correlation in the trend.

Moving averages calculate the algebraic average of a given observation and a specified number of adjacent observations, which can either be one-sided or two-sided. The method is very common as it is convenient and transparent, but it is a naive tool for detrending output. Defining the cyclical component of y_t as:

$$y_t^{gap} = \alpha(L)\, y_t, \tag{1}$$

where $\alpha(L) = \alpha_{-j}L^{-j} + ... + \alpha_{-1}L^{-1} + \alpha_0 + \alpha_1 L + ... + \alpha_k L^k$, implies that we can define the trend component of y_t as:

$$y_t^{*(MA)} = [1 - \alpha(L)]\, y_t = \beta(L)\, y_t. \tag{2}$$

The centered moving average for an even number of periods given by $2m$ is defined by the filter:[3]

$$\beta_n = \begin{cases} \frac{1}{2m}, & n = 0, \pm 1, ..., \pm(m-1) \\ \frac{1}{4m}, & n = \pm m. \end{cases} \tag{3}$$

If β_n contains the properties: i) $\beta_s = \beta_{-s}, \forall s$, ii) $j = k$ and iii) $\beta(1) = 1$, $\alpha(1) =$

[3] The centered MA filter for an odd number of periods, $2m + 1$, is given as:

$$\beta_n = \frac{1}{2m+1},\ n = 0, \pm 1, ..., \pm m.$$

0, which implies symmetry, there will be no phase shift. The effect of the filter will be captured in the 'gain' function.[4] The moving average filter can generate spurious cycles, particularly when y_t is a unit root process, by the gain function exhibiting cyclical behaviour.[5] If a two-sided filter is used the method is not timely as the smoothed series is reduced by k observations. Forecasts are often used to extend the series but these are subject to error. If a one-sided filter is used the mean is treated incorrectly because the series includes a trend. A common form of the 1-sided moving average filter is the exponential smoother. Note that there are direct parallels with the literature on seasonal adjustment, which is discussed in Section 2.3.

One of the most prominent univariate methods of potential output estimation is the Hodrick-Prescott (HP) filter. This is a two-sided symmetric moving average filter. Output can be decomposed into a trend and cycle by optimizing:

$$y_t^{*(HP)} = \arg\min_{y_t^*} \sum_{t=1}^{T} (y_t - y_t^*)^2 + \lambda \sum_{t=3}^{T} \left(\Delta^2 y_t^*\right)^2 . \tag{4}$$

This method obtains a trend that balances the fit to the original series against the degree of smoothness (proxied by the second difference). The level of smoothness depends on the parameter λ. A high λ implies a higher penalization for fit to the original series, yielding a smoother trend.[6] Effectively, the filter captures different priors on the ratio of

[4]This is readily observed in the frequency domain. Applying the Fourier transformation (which enables a series to be represented as the sum of a finite number of sinusoids).to the filter $\alpha(L)$, we can define the frequency response function of $\alpha(L)$ as $\widetilde{\alpha}(\omega)$ where ω is the frequency measured in radians $(-\pi \leq \omega \leq \pi)$.

$$\begin{aligned} \widetilde{\alpha}(\omega) &= \sum_{n=-j} \alpha_n \exp(-in\omega) \\ &= |\widetilde{\alpha}(\omega)| \exp[-iPh(\omega)], \end{aligned}$$

where $i = \sqrt{-1}$. This is the polar form, where $|\widetilde{\alpha}(\omega)|$ is the gain measuring the increase in amplitude of the filtered series over the original series and $Ph(\omega)$ is the phase, which measures the time displacement caused by the filter. This representation completely characterizes $\alpha(L)\, y_t$.

[5]This problem of a moving average generating irregular oscillation, if none exists in the underlying data, is widely documented. It is known as the Slutzky-Yule effect. See Slutzky (1937).

[6]At the extremes, $\lambda = \infty$ results in a trend which is a linear function of time and $\lambda = 0$ produces a trend equal to the original series.

demand to supply side shocks. For quarterly data λ=1600 is used as a rule of thumb. Prescott (1986) interprets the HP filter as a high-pass filter, in which low frequency movements are dampened but high frequencies are untouched, providing an objective way of determining λ.[7] Harvey and Jaeger (1993) argue that it is the solution to a signal extraction problem.[8]

It is important to recognize that whilst the literature terms the HP a filter, it does incorporate all available information (the summations extend to T as opposed to t). Hence, the HP is actually a smoother. There is an end-sample bias stemming from the symmetric property of the HP filter, which requires the output gaps to sum to zero over the estimation period. The trend estimate tends towards the actual estimate at the end of the sample, biasing the gap towards zero. This renders the method of little use when making current policy decisions. As potential output derived from the HP filter is just a moving average of actual output, sustained deviations from actual output, possibly due to substantial nominal rigidities, are not possible. Cogley and Nason (1995) show that the HP filter can lead to spurious cyclical behaviour whereby the cycles are due to the filtering procedure rather than the economic properties of the data.[9] However, the method is robust, easy to use and popular.

The univariate HP filter can be extended by conditioning on information variables, for

[7] For λ=1600, the HP filter can be rationalized as a high-pass filter capturing fluctuations with a period shorter than 32 quarters.

[8] The HP filter is the optimal linear estimator of the trend, μ_t, in the structural time series model given in equations (24), (25) and (26) below. The solution to the signal extraction problem is given by $\lambda = \sigma_\varepsilon^2/\sigma_\zeta^2$. However, this representation assumes that $y_t \sim I(2)$ and that the cyclical component is white noise. It is generally thought that $y_t \sim I(1)$ and the cyclical component contains some persistence.

[9] If the series is trend stationary, the HP filter effectively linearly detrends the data and then smoothes the deviations from trend in an equivalent manner to a high band pass filter. However, if the series is difference stationary, the HP filter is equivalent to a 2-step linear filter that differences the data and then smoothes the differenced data using an asymmetric moving average. This can amplify growth cycles at business cycle frequencies and dampen long-run and short-run fluctuations, leading to spurious cycles. See Nelson and Kang (1981) for a discussion of spurious cycles arising from detrending a RW, which is a direct parallel to applying the HP filter to an I(1) process.

example:

$$y_t^{*(MHP)} = \arg\min_{y_t^*} \sum_{t=1}^{T} (y_t - y_t^*)^2 + \lambda \sum_{t=3}^{T} \left(\Delta^2 y_t^*\right)^2 + \sum_{t=1}^{T} \beta_{\pi,t}\varepsilon_{\pi,t}^2 + \sum_{t=1}^{T} \beta_{u,t}\varepsilon_{u,t}^2 + \sum_{t=1}^{T} \beta_{cu,t}\varepsilon_{cu,t}^2$$

$$\pi_t = \pi_t^e + A(L)(y_t - y_t^*) + \varepsilon_{\pi,t} \tag{5}$$

$$u_t = nairu_t - B(L)(y_t - y_t^*) + \varepsilon_{u,t} \tag{6}$$

$$cu_t = cu_t^* + C(L)(y_t - y_t^*) + \varepsilon_{cu,t} \tag{7}$$

De Brouwer (1998) estimates a multivariate HP filter based on a Phillips curve, equation (5), Okun's Law, equation (6), and a partial indicator of capacity supply, equation (7). Hence, the multivariate filter conditions on structural relationships that contain information about the output gap. This should result in a more precise estimate of the gap. However, the method still holds all the caveats of the univariate filtering method.

The Cubic Spline (CS) is a popular non-parametric method of detrending output that is virtually identical to the HP filter.[10] A CS is a combination of piecewise cubic polynomials that can be fitted to a series of data points. Given knots at x_t (where the piecewise portions join), a CS exactly interpolates the data points. The portions are defined so that at the knots the function and its first two derivatives are continuous.

A CS with knots at x_t, for $t = 1, ..., T$, is defined (letting $x_0 = -\infty, x_{T+1} = \infty$) by:

$$f(x) = a_t + b_t x + c_t x^2 + d_t x^3, \tag{8}$$

[10] See Doornik and Hendry (1996) for an outline of why the natural CS and HP filter result in very similar decompositions.

subject to the restrictions:

$$\begin{aligned}
a_{t-1} + b_{t-1}x_t + c_{t-1}x_t^2 + d_{t-1}x_t^3 &= a_t + b_t x_t + c_t x_t^2 + d_t x_t^3 \\
b_{t-1} + 2c_{t-1}x_t + 3d_{t-1}x_t^2 &= b_t + 2c_t x_t + 3d_t x_t^2 \\
2c_{t-1} + 6d_{t-1}x_t &= 2c_t + 6d_t x_t \\
c_0 = d_0 = c_T = d_T &= 0.
\end{aligned} \tag{9}$$

The first three restrictions ensure the function and its first and second derivatives respectively are continuous at the knots. The final restriction means that CS is a linear function outside the range of the knots, thus avoiding the end-point bias problems of other such smoothers. The CS also has a discontinuous third derivative:

$$f'''(x) = d_t x_t \leq x < x_{t+1}. \tag{10}$$

The CS is the interpolating function that minimizes the sum of squared deviations from a function, f, subject to a roughness penalty given by the integrated squared second difference:

$$y_t^{*(CS)} = \min \sum_{t=1}^{T} \left[y_t - f(x_t)\right]^2 + \alpha \int \left[f''(x)\right]^2 dx. \tag{11}$$

α is the bandwidth, controlling the trade off between minimizing the residual error and minimizing local variation.[11] The solution gives the natural CS with knots belonging to $\{x_1, ..., x_T\}$. However, the method is highly dependent on the decision of the number of knots to impose, k, or the smoothing parameter α.[12]

The Kernel Smoother (KS) is another non-parametric method of detrending output

[11] GiveWin offers 3 ways of specifying the bandwidth. One is to specify an equivalent number of parameters, k_e (which is approximately equivalent to the number of regressors used in a linear regression). A second is to use the default, which corresponds to $\lambda = 1600$ for the HP filter (quarterly data) and the third is to choose the bandwidth by generalized cross validation, computed as:

$$GCV(\alpha) = T\left(\frac{RSS}{T - 1.25K_e + 0.5}\right).$$

Choosing the bandwidth on the basis of GCV tends to undersmooth. See Doornik and Hendry (1996) for further details.

[12] See Green and Silverman (1994) for a more detailed discussion on splines.

based on approximating the probability density function of a random variable. The KS is given as:

$$y_h^{*(KS)} = \left[\sum_{t=1}^{T} K\left(\frac{x - X_t}{h}\right)\right]^{-1} \left[\sum_{t=1}^{T} K\left(\frac{x - X_t}{h}\right) y_t\right], \tag{12}$$

where h is the bandwidth or smoothing parameter and $K(.)$ is the kernal, determining the shape of the weights. This is a continuous, bounded and symmetric real function that integrates to one: $\int_{-\infty}^{\infty} K(u)du = 1$. A variety of kernel functions are possible in general but the most commonly used kernel function is of parabolic shape, usually termed the Epanechnikov kernal:

$$K(u) = \begin{cases} \frac{3}{4}\left(1 - u^2\right), & \text{for } |u| \leq 1 \\ 0, & \text{for } |u| > 1. \end{cases} \tag{13}$$

The optimal bandwidth is given by $h = 0.75\hat{\sigma}_x/T^{0.2}$. This kernal is optimal in the density estimators. A drawback to the KS is that there are problems at the edges. For example, if a variable is trending upwards, the initial values of the KS will be based on a moving average of points that lead the series. These will be higher, causing the initial values of the KS to be higher. As more lags in the data are included the KS adjusts, arising in a classic 'J' shape at the origin.

The Beveridge Nelson (BN) filter is a model-based univariate approach that uses ARIMA methodology to decompose a non-stationary time series into a permanent and transitory component. The trend component is the long-run forecast of output and the cyclical component is output growth in excess of growth in the current state of the economy. Representing Δy_t as:

$$\Delta y_t = \mu + \frac{\left(1 - \theta_1 L - ... - \theta_q L^q\right)}{\left(1 - \varphi_1 L - ... - \varphi_p L^p\right)} \varepsilon_t, \tag{14}$$

where μ is the mean of the process, $\varepsilon_t \sim$NID$(0, \sigma_\varepsilon^2)$ and the roots of $\varphi_p(L) = 1 - \varphi_1 L - ... - \varphi_p L^p$ lie outside the unit circle. We can define:

$$\psi(L) = \frac{\theta_q(L)}{\varphi_p(L)} = \frac{1 - \theta_1 L - ... - \theta_q L^q}{1 - \varphi_1 L - ... - \varphi_p L^p}, \tag{15}$$

which implies that we can decompose Δy_t into a stationary and non-stationary component:

$$\Delta y_t = \underbrace{\mu + \psi(1)\varepsilon_t}_{\Delta y_t^*} + \underbrace{\widetilde{\psi}(L)\varepsilon_t}_{\Delta y_t^{gap}}, \tag{16}$$

where $\widetilde{\psi}(L) = \psi(L) - \psi(1)$. Defining the trend as the value the series would take if it were on the long-run time path in the current time period, the long-run forecast, adjusted for the mean rate of change, is given as:

$$y_t^{*(BN)} = \lim_{h\to\infty} \left(E_t y_{t+h} - \mu h|\Omega_t\right), \tag{17}$$

where $\Omega_t = (y_1, ..., y_t)$. This can alternatively be expressed as the weighted average of current and past values of y_t (see Miller, 1998):

$$y_t^{*(BN)} = \frac{\theta_q(1)}{\varphi_p(1)} \frac{\varphi_p(L)}{\theta_q(L)} y_t. \tag{18}$$

Hence, future information will not modify the trend component, which differs depending upon the stochastic properties of the series. By basing estimates of potential output on forecasts, a priori assumptions on the structure of the economy can be avoided. One problem is that alternative ARIMA models are likely to have different long-run specifications, resulting in differing decompositions.

Two features of the BN filter should be noted. Firstly, the method assumes the trend and cycle components are driven by the same shock as the error terms are perfectly correlated. Secondly, the trend component is modelled as a RW with drift. This implies that the stochastic trend accounts for most of the variation in output and can, in fact, be more volatile than the series itself.[13] The cycle is stationary with weights summing to 0 and it tends to be small and noisy. Whilst the BN decomposition is a 1-sided weighted average

[13] See Watson (1986) for a more detailed discussion regarding the RW trend of the BN decomposition.

filter, Prioretti and Harvey (2000) develop an algorithm for a two-sided BN smoother.

The BN filter can be extended to the multivariate case, as in Evans and Reichlin (1994). Additional information is used to forecast output growth in a VAR framework. The authors find that the cycle component has a larger variance in the multivariate framework compared to the univariate model, although the cycle is still sensitive to the lag length used.

Band Pass (BP) filters eliminate slow moving trend components and high frequency irregular components, retaining the intermediate business cycle elements. The filter passes through components belonging to a prespecified band of frequencies, removing those at higher and lower frequencies. An Ideal Band Pass filter will have a gain function (with upper and lower cut-off frequencies $\overline{\omega}$ and $\underline{\omega}$) given by:

$$F(\omega) = \begin{cases} 0 \text{ if } |\omega| \leq \underline{\omega} \\ 1 \text{ if } \underline{\omega} \leq |\omega| \leq \overline{\omega} \\ 0 \text{ if } |\omega| > \overline{\omega} \end{cases} \tag{19}$$

so that the frequencies belonging to the interval $[\underline{\omega}, \overline{\omega}]$ pass through the filter untouched. Note that an Ideal BP filter requires an infinite order moving average representation. In practice, an approximation to the ideal BP filter is required.

Filtering can be implemented in the time or frequency domain. For filtering in the frequency domain see Hassler et al. (1994). The filtering is implemented by initially smoothing the series using a HP filter (with very high λ) and extending with zeros. Then the spectral measure is derived using a Fast Fourier Transform, which is multiplied by the transfer function (equation (19)) and inverted to extract the cyclical component:[14]

$$y_t^{gap} = \sum_{j=0}^{T-1} S\left(\omega_j\right) F\left(\omega_j\right) e^{it\omega_j}. \tag{20}$$

The method has the drawback of detrending the series prior to the Fourier Transform. Also, results are not invariant to changes in the sample.

For filtering in the time domain, Baxter and King (1999) approximate the Ideal BP filter with a finite order moving average. Defining $b\left(L\right) = \sum_{j=0}^{\infty} b_j$, the 'best' approximation is obtained by solving the constrained optimization problem:

$$\min_{\{\alpha_j\}} \int_{-\pi}^{\pi} \left|\beta\left(\omega\right) - \alpha_k\left(\omega\right)\right|^2 d\omega, \qquad \text{s.t. } \alpha_k\left(0\right) = \phi, \tag{21}$$

where $\beta\left(\omega\right)$ denotes the gain function of the ideal BP filter, $\alpha_k\left(\omega\right)$ is the ideal BP function of the approximating filter and ϕ is the gain of the filter at zero frequency. The resulting optimal solution is:

$$a_j = b_j + \theta, \qquad \text{where } \theta = \frac{\phi - \sum_{j=-k}^{k} b_j}{2k+1}, \tag{22}$$

for $j = 0, ..., k$ where k is the lag length of the symmetric MA representation. Having obtained the weights, which are adjusted by θ, the cycle is given as:

$$y_t^{gap} = \sum_{j=-k}^{k} a_j L^j y_t. \tag{23}$$

Whilst the method of filtering in the time domain is transparent there are severe end-of-sample problems. Increasing k reduces the approximation error of the finite-order MA but k observations are lost at the end-point. In order to overcome this, Harvey and Trimbur (2002) suggest using the difference between two Butterworth (low-pass) filters as an approximation to the BP filter. This is comparable to the Baxter and King (1999) cycle but it provides end-point estimates. Finally, note that the specification of the frequency

[14] A Fast Fourier Transform is a recursive algorithm used to compute the discrete Fourier Transform efficiently.

interval $[\underline{\omega}, \overline{\omega}]$ is subjective. A priori judgement is needed as to what constitutes a business cycle.[15]

Harvey and Jaeger (1993) argue that Structural Time Series (STS) models are the most informative framework in which to determine stylized facts regarding time series data. This is because they are formulated in terms of components that have direct interpretation. The Unobserved Components (UC) model requires i) specification of the statistical model, ii) estimation of the set of hyperparameters and iii) application of the signal extraction algorithm. State Space Form (SSF) provides a framework for finding exact maximum likelihood estimates of a models parameters. Once the model is outlined in SSF, the Kalman Filter can be applied, enabling signal extraction for both stationary and non-stationary components.[16] Both filtering and smoothing can be applied in this framework. Appendix 2 outlines the general multivariate SSF model and the Kalman Filter. The univariate UC model is given as:

$$y_t^{(UC)} = \mu_t + \psi_t + \varepsilon_t \tag{24}$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \tag{25}$$

$$\beta_t = \beta_{t-1} + \zeta_t \tag{26}$$

$$\begin{bmatrix} \psi_t \\ \psi_t^* \end{bmatrix} = \rho \begin{bmatrix} \cos\lambda_c & \sin\lambda_c \\ -\sin\lambda_c & \cos\lambda_c \end{bmatrix} \begin{bmatrix} \psi_{t-1} \\ \psi_{t-1}^* \end{bmatrix} + \begin{bmatrix} \kappa_t \\ \kappa_t^* \end{bmatrix} \tag{27}$$

where $t = 1, ..., T$. λ_c is the frequency in radians, $0 < \lambda_c < \pi$, ρ is the dampening factor, $0 < \rho \leq 1$, and $\varepsilon_t \sim$NID$(0, \sigma_\varepsilon^2)$, $\eta_t \sim$NID$\left(0, \sigma_\eta^2\right)$, $\zeta_t \sim$NID$\left(0, \sigma_\zeta^2\right)$, and $\kappa_t, \kappa_t^* \sim$NID$(0, \sigma_\kappa^2)$. The disturbances of each of the components are assumed to be mutually uncorrelated.

[15] Following Burns and Mitchell (1946), Baxter and King (1999) consider cycles of between 6 and 32 quarters.

[16] The Kalman filter is equivalent to the recursive least squares algorithm. A linear regression model can be represented in SSF and standard OLS is equivalent to the KF for the last observation.

The measurement equation (24) comprises the trend component, μ_t, the cyclical component, ψ_t, and an irregular component, ε_t.[17] The transition equations (25), (26) and (27) define the trend and cycle. The trigonometric specification of the cycle is common, see Harvey (1993). The trend component is an ARIMA(0,2,1) process but recedes to a random walk with drift if $\sigma_\zeta^2 = 0$ and is deterministic if $\sigma_\eta^2 = 0$.[18] Koopman et al. (1999) list the sub-models contained in the general framework, which are obtained by placing restrictions upon the model parameters. Confidence intervals for the gap can be estimated giving a measure of the uncertainty associated with potential output.

Common factor models generalize STS models to the multivariate case by broadening the information set. If the variance matrices are of reduced rank, the model contains 'common' components. For example, Kuttner (1994) uses a bivariate UC model to derive estimates of potential output from the joint behaviour of output and inflation by adding an Expectations Augmented Phillips Curve. Flaig and Plotscher (2001) use a business assessment survey that is assumed to share the same cyclical component as output in order to improve the identification of the gap.

In contrast to the BN decomposition, the UC decomposition assumes that most of the variation in output occurs in the cycle. This is because of the restriction that the trend and cycle innovations are uncorrelated. Morley et al. (2002) find that if this restriction is relaxed, the UC decomposition leads to an identical representation as that of the

[17] A seasonal component can be included in the general form of the model. Also, an AR(1) component can be included, and is a limiting case of the stochastic cycle.

[18] The HP filter is a restricted case of the unobserved components model. The restrictions are given by:

$$\begin{aligned} \sigma_\varepsilon^2 &= 0,\ \rho = 0,\ \sigma_\eta^2 = 0, \\ \psi_t &= \kappa_t \sim \text{NID}\left(0, \sigma_\kappa^2\right),\ \sigma_\zeta^2 = \sigma_\kappa^2/1600. \end{aligned}$$

Hence, σ_κ^2 is the only variance parameter to be estimated. See Appendix 3. The UC model is often preferred to the HP filter as it does not rely on any arbitrary calibration of the variance of the trend term. Also the end-of-sample bias of the HP filter is reduced in the UC framework because the cycle prevents the trend from adjusting to accommodate the end-point problems.

BN decomposition.[19] They also find that the restriction of zero correlation between the trend and cycle for US quarterly GDP (1947-1998) is rejected at the 5% significance level. This lends support to the importance of real shocks in the economy; as real shocks shift the long-run path of the economy, short-term fluctuations lead to adjustments towards a shifting trend. For example, if $y_t = y_t^*$ initially, a positive real shock will cause a negative output gap, implying negative contemporaneous correlation between the trend and cycle. However, a positive transitory shock will have no impact upon the trend.

Kim (1994) shows that Markov-switching models that take account of structural changes in the dependent data can be represented in SSF, with switching in both the measurement and transition equations. The model can be designed to allow for explicit identification and estimation of the trend and cycle components of the underlying series.

A structural VAR, proposed by Blanchard and Quah (1989), is a multivariate method of decomposition based on economic theory. The VAR is given as:

$$y_t = d + \theta(L)\eta_t \tag{28}$$

where d is a vector of deterministic components and η is a vector of structural shocks. As $E(\eta_t \eta_s) = 0, \forall s \neq t$, the variance-covariance matrix can be normalized to the identity matrix. $\theta(L)$ represents the transition mechanism for the shocks, given as $\theta(L) = \sum_{i=0}^{\infty} \theta_i L^i$. Long-run restrictions on output are imposed in order to obtain identification. Blanchard and Quah (1989) take a traditional Keynesian view of fluctuations whereby disturbances with permanent effects are supply side shocks, shifting potential output whereas transitory effects are caused by demand side shocks. However, the distinction between demand and supply shocks, and hence the permanent and transitory components, is nontrivial. Also,

[19] Appendix 2 outlines the equivalence of the BN and UC models.

as the SVAR is a simultaneous equation system with instruments, the quality of the model depends upon the quality of the instruments used.

An alternative multivariate method of estimating potential output is based on the Permanent Income Hypothesis. Cochrane (1994) decomposes output into a trend and cycle by observing consumption, which is a random walk under the PIH. As consumption and income are cointegrated, any change in income must be transitory if consumption remains unchanged, providing the consumption to income ratio is fairly stable. The method relies on the controversial assumptions that PIH holds and that consumption is a RW, thereby constraining potential output to also be a RW.

Another subset of models used to decompose output are the controversial Real Business Cycle models. These are dynamic stochastic general equilibrium models where technology shocks are the only source of disturbance in the economy.[20] There has been an active research programme looking at RBC models but the general consensus is that these models are unable to account for the volatility and persistence of business cycle fluctuations.

Finally, the Production Function (PF) method is a prevalent method of estimating the output gap, particularly within International Institutions and Central Banks. Based on a simple growth accounting framework, the PF method essentially relates inputs to outputs. Potential output is determined by the supply side of the economy; capital, labour and the residual total factor productivity. These are in turn driven by microeconomic foundations such as technology and preferences. The method provides a rigorous framework based in economic theory and enables the impact of disturbances to the inputs to be traced through to potential output, providing a better grounding for policy analysis.[21] Whilst structure

[20] See King et al. (1988) who assume a deterministic trend and King et al. (1991) in which the model is augmented to allow for a stochastic trend.

[21] For example, the impact of policy changes such as changes to unemployment benefits or laws regulating

on the economy is imposed via ad hoc functional forms such as Cobb-Douglas or CES, this enables calculation of the underlying total factor productivity, which is an essential component of the long-run potential output of the economy. There are significant data requirements for the method. In particular, NAIRU (non-accelerating inflation rate of unemployment) estimates are uncertain and data such as that of capital stock are of poor quality. Trend rates of inputs need to be calculated and are often derived using the HP filter and so the problems of univariate detrending methods are not fully avoided.

2.2 Disaggregation

In most applied work there is a problem of aggregation but the issue is rarely addressed explicitly. Whilst there is no clear resolution to the problem, the fundamental issue is in determining at what level of aggregation the analysis should be carried out. This is often restricted by data limitations and conventions but there are strong arguments for working at a disaggregated level.

Espasa, Senra and Albacete (2002) find that breaking down the aggregate price level index into indexes corresponding to groups of markets vastly improves the forecasting performance for European inflation.[22] This is because the component prices are not fully cointegrated. The absence of full rank implies that the trends in the individual price indexes are generated by more than one common factor. Hence, innovations in the aggregate price level will have different long-run effects depending on which common trend they primarily stem from and there will not be full convergence between the indexes.

---

hours worked can be traced though to their effect on potential output.

[22] Espasa et al. (2002) examine the harmonised consumer price index for Europe. They disaggregate the data, both in terms of price indexes corresponding to big groups of markets and countries. The sectoral split is given by non-processed foods, energy, other goods and other services. The country splits examine France, Germany, Italy and Spain. They conclude that forecasts derived from aggregating the forecasts of the individual components outperform forecasts that are calculated by aggregating the components first.

In terms of policy, the feed through effects from a shock to a particular price index will differ depending upon the size and type of innovation and the transmission mechanism. Espasa et al. (2002) distinguish between core and residual inflation on the basis of which price indices are more volatile. It is entirely plausible that policy-makers will react differently to inflationary pressures arising from core inflation as opposed to residual inflation. One example of the differences in inflationary effects is insurance premia. Poor performance by companies drives up premia because of the need to rebuild margins, increasing services price inflation. This differs markedly from goods price inflation. Also, if the manufacturing sector has been less subject to 'new economy' effects than the service sector, the way a production sector output gap will feed into inflation will be more stable than a service sector gap.

There is a trade-off when determining the level of aggregation that should be used. Consumption patterns change over time. If a commodity price index is used as the measure of inflation, an average consumption bundle will be more representative than the disaggregated data because the index uses shares in expenditure as weights. The analysis in this paper examines aggregate inflation due to data limitations, but with the advent of the 'new economy' and the divergence of the goods and services markets, the issue of disaggregation must come to the forefront in applied work.

2.3 Methods of Seasonal Adjustment

The goal of the above filtering and smoothing procedures is to remove unwanted features if the data series, leaving the component that is of interest to the researcher. Macroeconomic data not only contains trends and cycles but also seasonal components. This is often ignored as seasonally adjusted data is used but it is important to recognize that the process of seasonal adjustment will distort the underlying properties of the data.

Seasonality can be described as regular and recurring patterns of variability across years. There are many approaches to modelling seasonality, including dummy variables that parametrically adjust for seasonality, endogenous dynamics such as annual differences, evolving patterns of expenditure and filters that remove components at seasonal frequencies. The last is the most common method of adjusting for seasonality and much of the data published by the ONS is seasonally adjusted using these filters.[23] The problem of seasonal adjustment (SA) is to filter out the seasonal factor without seriously distorting the other elements generating the observed data. Following Hendry (1995a), we shall analyze an approximation to the X-11 procedure which is a linear two-sided filter, see Wallis (1974).[24] Letting $\{x_t^a\}$ be a seasonally adjusted series and $f(L)$ be the two-sided linear filter we can can define the SA series as:

$$x_t^a = f(L)\, x_t, \quad \text{where } f(L) = \sum_{i=-m}^{m} f_i L^i. \tag{29}$$

The fixed, finite weights are given by f_i.[25] We express $f(L)$ as:

$$f(L) = f(1) + f^*(L)\,\Delta = f(1) + f^*(1)\,\Delta + f^{**}(L)\,\Delta^2 ... \tag{30}$$

where the recursion can be repeated to any order. $f^*(L)$ and $f^{**}(L)$ are finite-order, fixed weight, two-sided linear filters with coefficients $\{f_i^*\}$ and $\{f_i^{**}\}$ respectively. The sum of the coefficients in the successive lag polynomials can be obtained from:

$$f^*(1) = -\frac{\partial f(L)}{\partial L}\rfloor_{L=1} \text{ and } f^{**}(1) = -\frac{\partial f^*(L)}{\partial L}\rfloor_{L=1}, \tag{31}$$

which implies $f^*(1) = -\sum_{i=-m}^{m} i f_i$ and $f^{**}(1) = -\sum_{i=-m}^{m-1} i f_i^*$. The properties of the

[23] See Hylleberg (1992) for a discussion of the properties of this and other SA methods.

[24] We ignore features such as graduation of extreme values, constraints on calendar year totals, corrections at the end of sample and multiplicative models of SA.

[25] If some f_i were set to zero, $f(L)$ could be a one-sided filter.

seasonal filter include:

1. Normalization: $f(1)=1$. Therefore, $f(L)=1+f^{*}(L)\Delta$. This ensures that x_t^a and x_t are in the same units, implying the long-run properties of the series will not be affected.

2. Symmetry: $f(L)=f(L^{-1})$. Therefore, $f_i=f_{-i}$, $i=1,...,m$. This is sufficient for $f^{*}(1)=0$, which implies $f(L)=1+f^{**}(L)\Delta^2$.

3. Irrelevance of seasonal dummies: Representing a fixed seasonal pattern by $d(L)s_{1,t}$ where $d(L)$ is a $(p-1)$th-order polynomial (p=periodicity of seasonality) and $s_{i,t}$ is the centred dummy for the ith season, we can define the two-sided linear filter as:

$$f(L)=f^{\blacklozenge}(L)\phi(L), \quad \text{where } \phi(L)=p^{-1}\sum_{i=0}^{p-1}L^i. \tag{32}$$

   As $\phi(L)s_{1,t}=0$:

$$f(L)d(L)s_{1,t}=f^{\blacklozenge}(L)\phi(L)d(L)s_{1,t}=f^{\blacklozenge}(L)d(L)\phi(L)s_{1,t}=0, \tag{33}$$

   removing the seasonal dummies.[26]

This analysis exposes the links between SA smoothers and the centred MA used to detrend output. The centred MA is equivalent to applying a seasonal adjustment onto SA data, imposing more structure. Seasonal adjustment will invariably affect the cyclical properties of the data, distorting business cycle measurement. This is because the dynamic specification will be different. By imposing structured seasonal patterns, constancy

[26] Also note that cointegration should be invariant to seasonal filtering. If $\mathbf{x}_t$ and $\mathbf{x}_t^a$ form a vector of n I(1) time series which satisfy (29) and β is an $n\times r$ cointegrating matrix for $\mathbf{x}_t$, then β is also a cointegrating matrix for $\mathbf{x}_t^a$. This implies that SA will only affect the short-run dynamics of the process, providing the properties outlined are satisfied. $\mathbf{x}_t$ and $\mathbf{x}_t^a$ have the same number of cointegrating vectors. Using the property of symmetry and pre-multiplying by β':

$$\beta'\mathbf{x}_t^a=\beta'\mathbf{x}_t+\beta'\mathbf{f}^{**}(L)\Delta^2\mathbf{x}_t,$$

which implies that β is the cointegrating matrix for both $\mathbf{x}_t$ and $\mathbf{x}_t^a$.

assumptions will be embodied in the filters. If seasonal behaviour is endogenous this can lead to serious errors in the data. Thus, caution is essential when analyzing univariate detrending methods on SA data.

2.4 Real-Time Estimation of the Output Gap

Timely and accurate estimates of the gap are required for policy-making. Estimated reaction functions indicate that Central Banks do respond to the gap and optimal control exercises suggest that it is optimal to do so. However, conventional measures of the gap suffer from errors attributable to end-point problems and revisions to the underlying data. In an attempt to overcome these problems, the gap has been estimated in real-time. Orphanides and van-Norden (1999) show that ex post revisions of the output gap are of the same order of magnitude as the output gap itself, with the bulk of errors attributable to the unreliability of end-of-sample estimates of trend output. The biases are most acute at business cycle turning points, where the costs of policy errors are at their greatest.

High frequency estimation of the output gap may partially reduce the end-point problems associated with estimation of the gap and bring forecasts of the gap closer to those of the real-time estimates. As there is a trade-off between the advantages of high frequency data and the disadvantages of noisy data, estimation across a range of frequencies may provide fruitful results. Data limitations currently prevent estimation of the gap at higher frequencies than quarterly estimation but more timely survey data provides an avenue of research that is likely to be productive.

2.5 Asymmetric Inflation Effects

The question of whether the relationship between inflation and excess demand is non-linear or asymmetric has attracted some attention in recent years. If a given amount of unemployment below the NAIRU is more inflationary than an equivalent excess unemployment is deflationary, parametric trend estimation methods will be biased.[27] If positive output gaps have a stronger effect on inflation than negative gaps, average actual output may well be below potential output and estimation methods would tend to overestimate potential output. By rearranging a simple Phillips curve:

$$\pi_t = \pi_{t-1} + \alpha \left(y_t - y_t^* \right) \tag{34}$$

$$\sum_{t=1}^{T} \left(y_t - y_t^* \right) = \alpha^{-1} \left(\pi_T - \pi_0 \right), \tag{35}$$

we can see that if $\pi_0 = \pi_T$, the sum of the gaps will equal 0 and any demand management policies implemented to reduce the impact of negative gaps will be at the expense of positive gaps. The asymmetry hypothesis emphasizes reducing positive gaps in order to avoid the disproportionately greater recessionary costs that will be realized when the cycle moves into a downturn. If there are substantial asymmetry effects, policy must be designed carefully in order to avoid overshooting when closing the gap as this can be very costly. Demand management policies that are effective in preventing booms could result in a higher average level of output. The literature provides a general conclusion that there is some evidence of asymmetry but the empirical evidence is somewhat inconclusive due to the problems of measuring potential output, the forms of non-linearity and different specifications of the hypothesis. Clements and Sensier (1999) do find the Phillips curve

[27] See Chadha, Masson and Meredith (1992), Laxton Meredith and Rose (1995), and Clark Laxton and Rose (1996).

to be asymmetric for the UK, with only positive gaps affecting inflation. However, in a broader dynamic mark-up model of inflation their finding of asymmetry characterizing the output gap is not reinforced.

Turner (1995) discusses speed limit effects, which are described as a rise in inflation that is attributable to a reduction in the output gap, despite output not rising above its potential level. This asymmetry is based on a Keynesian supply curve that is almost vertical beyond the level of potential output. However, there is limited evidence for such effects. If they do exist, their impact is thought to be minimal.

2.6 The Theoretical Literature

The output gap is essentially a proxy for excess demand in the economy. A positive demand (supply) side shock will impact on the transitory (permanent) component of output, causing excess demand (supply). The initial empirical study of excess demand in relation to inflation was encapsulated in the Phillips curve (1958), which assumed there was a permanent trade-off between unemployment and inflation. When this relationship failed in the early 1970s, the Natural Rate Theory, proposed by Friedman (1968), came to prominence. This theory argues that there is an equilibrating pressure in the long-run and deviations from potential output, which is analogous to the natural rate of unemployment, cannot be permanently sustained.

The Lucas Island Model (1973) aimed to include expectations explicitly in the model. A long-run trade-off between inflation and output would only be sustained if expectations were backward looking and inflation was low and stable, otherwise the economy would always be at the natural rate. New Keynesian Models were developed in an attempt to explain why there may be permanent excess demand effects. These use menu cost models or staggered price models to allow for output persistence. The presence of real and nominal rigidities prevents the immediate price adjustment to return to equilibrium, implying that output can be sustained above potential for long periods.

It is generally accepted that excess demand pressures feed though to inflation in the short-run only. If there is price homogeneity and the output gap has a zero mean, the long-run impact of the gap will be zero. In Chapter 5, a model of inflation based upon a mark-up model is derived, whereby the long-run price level is determined by the supply side but excess demand pressures impact upon inflation in the short-run.

2.7 Changes in Potential Output

The past half century has seen many fluctuations in potential output, but as a latent variable it is difficult to determine exactly when these shifts occurred. There was a decline in potential output in the 1970s from strong growth in the 1960s. Growth recovered in the 1980s but did not reach the rapid growth of the 1960s and the second half of the 1990s has seen an upsurge in output growth but it is still relatively modest.

The causes of a change in potential output stem from numerous sources. These include changes in the rates of capital accumulation, the growth rate of labour inputs and the pace of technological advancement. For example, the slowdown in the 1970s was driven by a fall in total factor productivity (TFP) due to changes in the sectoral share of output, reduced scope for economies of scale, growth in public expenditure, changes in market shares and reduced catch-up effects with the US after the rapid growth of the 1950s and 1960s. Added to lower TFP was a reduction in labour input as average working hours declined without a corresponding increase in employment and increased capital obsolescence following the oil shocks, all of which contributed to lower potential output growth.

The 'new economy' debate has led to a renewed interest in the issue of the sustainable level of output growth.[28] The acceleration of productivity growth over the 1990s, particularly in the US, is largely attributable to the ICT sector.[29] Growth in total labour productivity due to the ICT sector is traced to the production of information technology capital at rapidly falling prices encouraging capital deepening and to the direct contribution of technical progress in information capital to TFP. There is a debate as to TFP growth outside the ICT sector. Oliner and Sichel (2000) find that there is evidence of growth whereas Gordon (2000) attributes any growth outside of this sector to cyclical gains.[30]

These type of policy issues emphasize the need for a method of calculating the growth of the underlying inputs of GDP. Univariate methods are useful in providing estimates of the gap per se but they do not provide any information as to the causes of a change in the gap. This is the reason why the production function method of estimating the gap is so popular with policy-makers. Despite its obvious shortcomings, the production function approach does enable the growth rate of potential output to be dissected into capital inputs, labour and population growth and technical progress.

[28] Coen and Hickman (2002) provide an empirical analysis of the late 1990s productivity growth and projections for the future performance of the US economy using an annual growth model. Cecchetti (2002) assess the implications of the 'new economy' for policy-makers, emphasizing the difficulties associated with estimating potential output when the productivity trend is shifting.

[29] This includes computer hardware, software and communications equipment.

[30] Oliner and Sichel (2000) estimate that one third of the increase in growth in the second half of the 1990s is due to gains outside of the ICT sector, whereas Gordon (2000) claims that once the cyclical component is extracted from the trend, there is no revival of productivity growth outside of the ICT sector.

3 Assessment of Univariate Methods of Estimating the Output Gap

A fundamental criterion for a 'good' output gap measure is that it can accurately identify the 'true' gap under a variety of different circumstances. As the gap is a latent variable, there is no way to determine whether a method is correctly picking up changes in potential output as opposed to the output gap. Instead, we can use artificially generated data to compare methods. This chapter applies univariate detrending procedures to a prespecified data-generation process (DGP). A variety of circumstances shall be examined including i) a change in the growth rate of potential output, ii) an intercept shift in potential output and iii) a large negative output gap.[31] The US productivity increase in the 1990s brought about an increase in the rate of growth of potential output. This should be reflected by an increasing trend as opposed to assuming a larger positive output gap. The Great Depression saw a substantial and prolonged period of negative excess demand. Most univariate estimates of the Great Depression systematically underestimate the magnitude and length of the recession. This is because they attribute a greater proportion of the decline in output to a fall in potential output that, with hindsight, is not warranted.

A shift in the gap will be caused by temporary shocks. These may be characterized as demand side shocks but could include temporary supply side shocks, such as some oil price shocks or exchange rate shocks. A break in potential output will be attributable to a permanent supply side shock such as a change in productivity or the growth rate of the population, although a shift could also be caused by changes in trade union militancy,

[31] Within this analysis, we need to define precisely what is assumed to be a structural break. A vector of parameters, $\theta \in \Theta$ where $\Theta = \{\theta; \forall \text{ admissible } \theta\}$, comprises the structure of a system if invariant and directly characterizes the relations of the economy under analysis. The parameters are invariant to a change in the DGP if they remain constant despite intervention. Thus, we can define a structural change to occur when an element or elements of θ change.

sustained changes to import prices or other long-term effects. Supply side shocks are thought to diffuse slowly into the economy, resulting in a smooth potential output series. Any detrending method would ideally be able to distinguish between these cases in order to accurately estimate the output gap. If using a production function approach, one could make the distinction between permanent and transitory shocks by assuming that transitory shocks are reflected in changes to inventories and utilization rates whereas permanent shocks are incorporated in shifts in employment and the capital stock. This will be dependent on the degree of flexibility in the markets for labour and capital. If there are high adjustment costs this assumption is plausible but in industries with low adjustment costs for both capital and labour, temporary shocks may well be reflected in changes to employment and the capital stock. It is clear that subjective analysis is required to distinguish between shifts in actual and potential output.

3.1 Characterization of UK GDP

This section describes UK real GDP for the period 1965q1-2002q2, and considers its basic properties. Figure 1, panel a, records the log of UK GDP, y, along with the annual growth rate, $\Delta_4 y$, in panel b, the quarterly growth rate, Δy, in panel c and the difference of the quarterly growth rate, $\Delta^2 y$, in panel d.[32] The sharp increase in output growth in 1973q1 to an unprecedented 10.5%pa was caused by a fiscal expansion package and was immediately followed by the oil crisis which brought about negative growth rates. The second oil crisis was followed by the Thatcher period in which there was a permanent intercept shift downwards in output as well as a reduction in volatility. The regime shift takes place over

[32]Graph panels are lettered a, b, c and d, row by row, i.e. $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$. The difference operator is defined as $(1 - L)$ where L represents the lag operator. For example, the lag of variable x at time t is given as $Lx_t = x_{t-1}$. Hence, $\Delta x_t = x_t - x_{t-1}$ and in general $\Delta_i^j = (1 - L^i)^j x_t$.

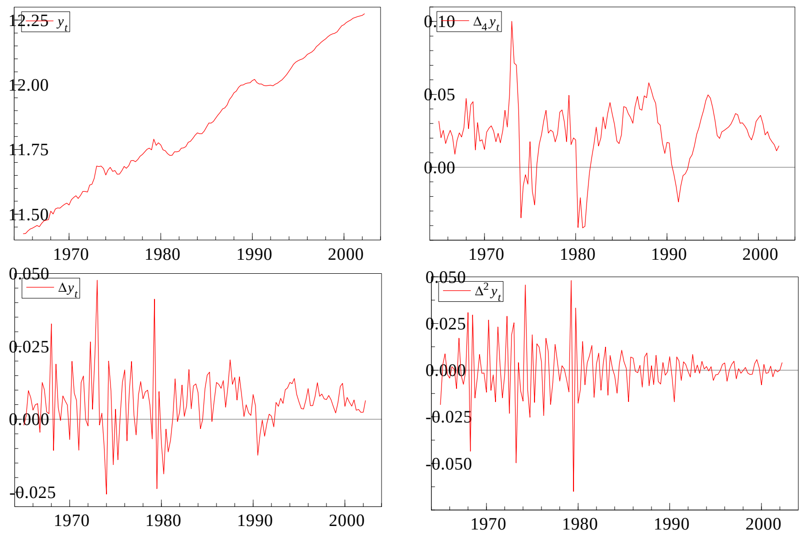

Figure 1: UK GDP, levels and growth rates.

a period but we shall date the break as occurring in the quarter following the imposition of the medium term financial strategy in March 1980. Another break occurs in the early 1990s where there is a period of slowing growth. This is a more prolonged change and it is difficult to determine whether it is a deep recession, reflected in a large negative output gap, or whether this represents a shift in the underlying potential of the economy. There is a decline in the volatility of the growth rate in the latter part of the 1990s, probably driven by the shift towards inflation targeting after exiting the ERM.

The sample ACF and PACF for y_t and Δy_t suggests a unit root in the level of UK GDP.[33] A unit root implies that shocks will have a permanent effect on the dependent variable. There has been a profusion of literature looking at the question of output persistence.

[33] The sample autocorrelation function (ACF) records the correlations between y_t and successive y_{t-j} for $j = 1, ..., J$. If the sample autocovariance is equal to $\widehat{c}_j = \frac{1}{T}\sum_{t=j+1}^{T}(y_t - \overline{y})(y_{t-j} - \overline{y})$ for $j = 0, ..., T-1$, the ACF is given as:

$$\widehat{r}_j = \frac{\widehat{c}_j}{\widehat{c}_0}, \quad \text{for } j = 0, ..., T-1.$$

The partial autocorrelation function (PACF) corrects the autocorrelation function for the effects of previous lags.

Traditionally, aggregate output was thought to be trend stationary. Thus, potential output was given by a deterministic trend and the gap comprised temporary shocks to output. Nelson and Plosser's (1982) seminal paper disagreed with this viewpoint (using annual US real GNP data). They could not reject the hypothesis that output is non-stationary, implying that fluctuations have a permanent component. Campbell and Mankiw (1987) support these findings, although Cochrane (1991) argues that the evidence on unit roots is empirically ambiguous. Perron (1989) includes a structural break in the analysis, in the form of a change in slope at the time of the first oil shock in 1973, and concludes that output may well be trend stationary with a structural break.[34]

Table 1 reports the Dickey-Fuller (DF) test statistics for y_t.[35] The table records the estimated coefficient on the lagged dependent variable, $\widehat{\gamma}$, the ADF(k) statistic, $\widehat{\tau}$, and the number of lags (determined by the last highest significant lag). The results indicate that y does contain a unit root. One problem is that of near observational equivalence where, when the alternative is close to 1, the power is low and a false non-rejection of the null can be a frequent outcome. It is often very difficult to distinguish stable from unit roots.[36]

[34] If output is difference stationary but is treated as trend stationary there is a problem of underdifferencing. Likewise, if the process is trend stationary but is differenced as it is thought to be difference stationary there is a case of overdifferencing. Plosser and Schwert (1978) argue that the problems of over versus under-differencing are not the primary concern. The key is in correctly modelling the error term. By accounting for serial correlation in the errors, the problem of underdifferencing is vastly reduced.

[35] For a variable x, the Augmented Dickey-Fuller (1981) test is outlined as:

$$\Delta x_t = \mu + \beta t + \gamma x_{t-1} + \sum\nolimits_{j=1}^{k} \alpha_j \Delta x_{t-j} + \varepsilon_t,$$

where k is the number of lags on the dependent variable, which are included to eliminate autocorrelation. If there are too few LDVs the size of the test will be adversely affected leading to an over-rejection of the null, whereas too many LDVs will reduce the power of the test. A time trend is included so that the alternative hypothesis is that of trend stationarity. The null hypothesis of a random walk is given by $H_0 : \gamma = 0$ and the test statistic is:

$$\widehat{\tau}_t = \frac{\widehat{\gamma}}{ese\left(\widehat{\gamma}\right)} \sim DF_\tau$$

Critical Values are -3.441 at 5% significance and -4.023 at 1% significance. Asterisks * and ** denote rejection at the 5% and 1% critical values. The null hypotheses of I(2) and I(3) are tested using the same regression where Δx_t and $\Delta^2 x_t$ replace x_t respectively.

[36] Numerous tests have been developed to detect one or more unit roots. These include Pantula,

Null order	$\widehat{\gamma}$	$\widehat{\tau}$	lag
I(1)	-0.08	-2.82	3
I(2)	-0.75	-5.30**	2
I(3)	-3.59	-9.26**	4

Table 1: Dickey-Fuller tests for integration.

Null order	$\widehat{\gamma}$	$\widehat{\tau}$	lags	λ
I(1)	-0.12	-3.11	3	$\frac{89}{150}$
I(2)	-0.78	-5.54*	2	$\frac{89}{150}$

Table 2: Perron test for I(1) with structural break.

It may arise that the finding of non-stationarity is spurious because no account has been taken of possible structural breaks. The Perron (1989) test for a break in the intercept (at a known point in time) is given as:

$$\Delta x_t = \mu + \beta t + \gamma x_{t-1} + \sum_{j=1}^{k} \alpha_j \Delta x_{t-j} + \delta_1 DVTB_t + \delta_2 DVU_t + \varepsilon_t \tag{36}$$

$$DVTB_t = \begin{cases} 1, & \text{if } t = T_B + 1 \\ 0, & \text{otherwise} \end{cases} \qquad DVU_t = \begin{cases} 1, & \text{if } t \geq T_B + 1 \\ 0, & \text{otherwise} \end{cases}$$

T_B is the time when the break occurs and $\lambda = \frac{T_B}{T}$. The dummy variables, $DVTB_t$ and DVU_t, represent the temporary and permanent shift in the intercept respectively. The null hypothesis is I(1) with break and the alternative is trend stationarity with break.[37] The results, given for a break in 1980Q2, are reported in table 2.

Again, we cannot reject the null hypothesis of output containing a unit root with break in 1980Q2 and we shall cautiously conclude that the log of real GDP does contain a unit root.[38] Section 3.2 will model output as the sum of a stochastic but smooth trend and a

Gonzalez-Farias and Fuller (1994) who improve on the power of the DF test using alternative test statistics and Phillips (1987) who allows for serial correlation in the error term. Other tests such as that developed by Kwiatkowski, Phillips, Schmidt and Shin (1992) test for a null of trend stationarity against an alternative of non-stationarity.

[37] The asymptotic critical values are -3.76 at the 5% significance level and -3.47 at the 10% significance level. Asterisk * denotes rejection at the 5% critical value.

[38] Zivot and Andrews (1992) argue that the Perron test biases the results in favour of a rejection of the unit root hypothesis because the break point is treated as known, and develop a testing procedure where there is an unknown break point. Perron (1997) then developed a selection criterion for choosing the

stationary cycle. The trend will be unpredictable but with systematic variation.

3.2 Artificially Generated Data: the DGP

In order to capture the characteristics of the UK business cycle, a few stylized facts based upon a variety of detrending methods applied to actual output over the period 1965q1-2002q2 are listed here. They are assessed in more detail in Chapter 4.4. Output growth over the period is approximately 0.55-0.6% per quarter, resulting in an annual growth rate of 2-2.5%, although this does generalize a period in which growth rates have fluctuated substantially.[39] Estimating a stochastic potential output based on the HP filter ($\lambda = 1600$) results in a trend that fluctuates between a maximum range of $\pm 4\%$ against a deterministic trend. The univariate detrending methods vary with regard to their smoothness. The smoother trends have DF statistics that suggest they are close being I(2) processes.

The gap over this period has a mean of approximately zero due to the symmetry properties of the filters applied. However, if we assume that the gap is a measure of short-run demand side pressures, we would expect the mean to be zero if the end-points coincided with full cycles. The output gap has a standard deviation of between 0.015-0.025 and a range of about $\pm 4\%$ of output. The cycle durations are quite long, with a peak to trough of approximately 14 quarters and a trough to peak of approximately 20 quarters over the entire sample period.[40] Whilst these figures suggest that there is some asymmetry

break point where the t-statistic on the parameter associated with the change shows the greatest evidence of change. However, the Perron test with a known structural break is more powerful than tests where an unknown break point must be estimated.

[39] The maximum quarterly range is almost 8% of output, where the quarterly growth rate reached +4.7% in 1973q1, followed by negative growth of -2.6% in 1974q1. In the last 2 decades growth has been more stable but still has a wide range of approximately 3.5% of output.

[40] Expansions (contractions) are defined as the phase from trough (peak) to peak (trough). A nave dating rule is used, given by the algorithm: a peak (trough) is identified as the highest point during which output is above (below) trend, given that output is above (below) trend for 3 quarters or more. Chapter 4.4 elucidates on this analysis, undertaking a discussion on the algorithms used to identify turning points. The figures quoted in this section are intended only to provide a brief overview of what the gap may look

in the gap, we shall abstract from these issues when generating data. With these general but crude statistics in mind we can generate output data to see which detrending methods perform best.

Assuming that output is persistent, actual GDP is well characterized by a random walk (RW) with drift. Figure 2, panel a records the log of actual GDP, y_t, against the generated data, $\widetilde{y}_t^{(RW)}$ given by:

$$\widetilde{y}_t^{(RW)} = 0.006 + \widetilde{y}_{t-1} + u_t = 0.006t + \sum_{i=0}^{t-1} u_{t-i}, \tag{37}$$

where the assumptions of $u_t \sim$N$(0, 0.001)$ and $y_0 = 0$ are made and the sample size $= 150$. ($\widetilde{y}$ denotes generated as opposed to actual output data).

Whilst the RW representation is plausible, we find a more informative representation of output is obtained by decomposing output into the sum of a cycle and trend in the style of the UC methodology; $\widetilde{y}_t = \widetilde{y}_t^* + \widetilde{y}_t^{gap}$. In this framework potential output is known with certainty, providing the 'true' benchmark against which to compare different detrending procedures. The trend component is formulated as a UC model augmented by a small deterministic component in the level, given as:

$$\widetilde{y}_t^* = \widetilde{y}_{t-1}^* + \beta_{t-1} + 0.0001t + \eta_t$$

$$\beta_t = 0.85\beta_{t-1} + \zeta_t, \tag{38}$$

where $\widetilde{y}_t^*$ is the potential output level and β_t is is the slope. We make the assumptions that $\eta_t \sim$N$(0, 0)$ and $\zeta_t \sim$N$(0, 0.001)$ resulting in a smooth trend with stochastic slope.[41] The small deterministic trend, giving an exogenous growth rate of 0.01%, is included in order

like.

[41] The initial values were set to 0.1 and a sample size of 200 was used, with the first 50 observations being discarded because of the arbitrary initial values.

to better represent the characteristics of trend output. Various coefficients on β_{t-1} were tried in order to derive a smooth trend which is borderline I(2) and has similar properties to the stylized facts. A higher coefficient than 0.85 led to a process that was too close to I(2), whereas a smaller coefficient resulted in a trend that was quite volatile. Treating the generated data as being in logs, the data are scaled in order to compare with the deterministic trend, and are recorded in figure 2, panel b, along with the deterministic trend $\{\widetilde{y}_t^* \, (DT)\}$ calculated as $\widetilde{y}_t^* = 0.006t$. The maximum deviation of the stochastic trend from the deterministic trend is -4.0% and $+3.3\%$, which accords with our stylized facts.

The output gap is characterized by a stationary ARMA(1,1) process with zero mean:

$$\widetilde{y}_t^{gap} = 0.8\widetilde{y}_{t-1}^{gap} + \varepsilon_t + 0.6\varepsilon_{t-1}. \tag{39}$$

The MA component is included to produce a series that is more representative of the actual output gap, with the assumption that $\varepsilon_t \sim \mathrm{N}(0, 0.01)$.[42] The data are scaled to give a maximum positive gap of 4% of output, a maximum negative gap of 3% and a standard deviation of 0.015. $\widetilde{y}_t^{gap}$ is recorded in figure 2, panel c. Summing the trend and cycle leads to the generated output series given in figure 2, panel d, which is plotted against actual output for comparison.

3.3 Univariate Detrending Procedures Applied to the DGP

A direct assessment of various univariate methods can be undertaken given the generated data. A range of descriptive statistics are reported, including the mean, range, standard deviation and correlation between the 'true' and estimated gaps. Durations and amplitudes of the cycles are given, but note that a simple dating rule is applied. The last two columns

[42] A sample size of 200 was used, discarding the first 50 observations and the sample mean was set to 0.

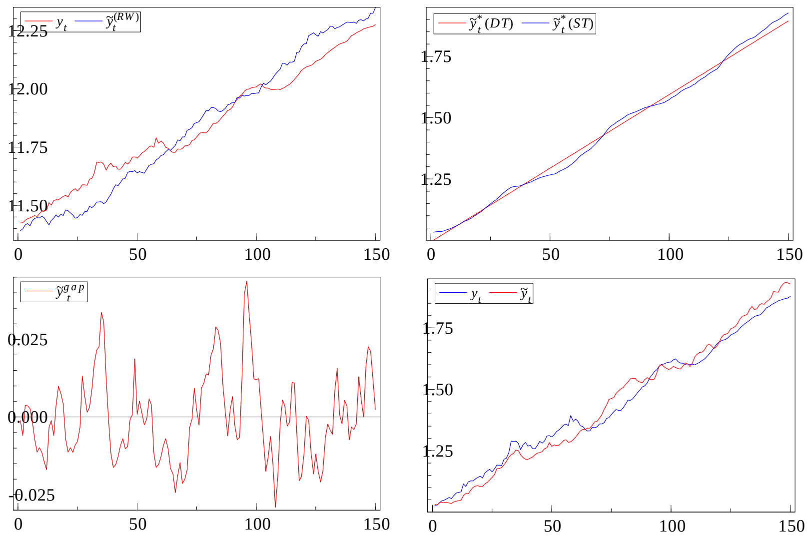

Figure 2: Generating output data.

report the mean absolute error (MAE) and mean squared error (MSE) of the estimated gap measures (reported as percentages). The MSE is given as:

$$MSE = \frac{1}{T}\sum_{t=1}^{T}\left(\widetilde{y}_t^{gap} - \widehat{\widetilde{y}_t}^{gap}\right)^2 \tag{40}$$

This statistic combines both bias and efficiency criteria. The statistical measures that were applied include a linear trend (LIN), linear trend with break (LINXX), Δ_4, HP, CS, KS, centred moving average with a lead and lag of 16 ($MA(16,16)$), filtered MA with a lag of 32 ($MA(32)$) and a UC model based on equations (24), (25), (26) and (27) with the assumption that $\eta_t \sim$N$(0,0)$. All methods apart from Δ_4 and $MA(32)$ are 'smoothers', which incorporate all available information over the sample period.[43] Figure 3 records the generated output gap, $\widetilde{y}_t^{gap}$, against the gaps derived from the applying the detrending methods to the DGP, with LIN and Δ_4 in panel a, HP, CS and KS in panel b, the MAs in panel c and the UC in panel d. Table 3 reports the summary statistics.

[43] Note that a filtered UC gap can also be estimated.

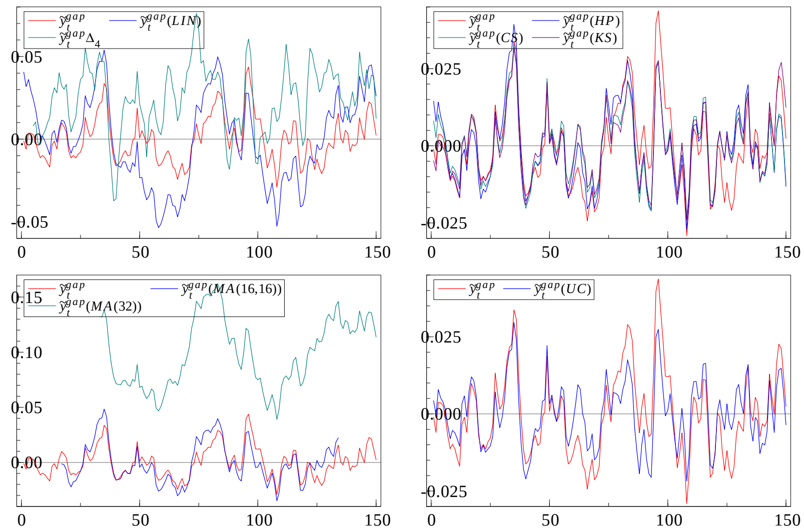

Figure 3: Estimated output gaps from the generated data.

Most methods do pick up the turning points in the generated data, with the notable exception of Δ_4. As LIN attributes all fluctuations in output to the gap as opposed to the trend, it systematically overestimates the magnitude of the gap and the strength of both accelerations and decelerations.[44] The HP, CS and KS perform well in estimating the gap, with MSEs of 0.006%, 0.006% and 0.005% respectively. The conventional $\lambda = 1600$ was found to perform best out of a range of parameters, as would be expected if the generated data accurately reflected true output. A bandwidth of 12 was used for the CS and KS, which is preferred to using GCV as this tends to undersmooth the data. This does suggest that priors regarding the smoothness parameter are important when modelling latent variables. The end-point behaviour of the methods is poor. The magnitude of the generated gap is +0.2% in period 150, but the HP, CS and KS estimate a gap of -1.3%, -1.2% and +1.2% respectively. Hence, the use of such smoothers is highly questionable for

[44] For example, LIN estimates the recession during periods 57-73 to reach a maximum magnitude of 4.7% of output, compared to a true negative gap of 2.4%. This was then followed by a trough to peak range of 9.7% estimated by LIN, whereas the true range was only 5.3%.

	Mean	St. Dev	Range	Corr.	Dur. PtoT	Dur. TtoP	Amp. PtoT	Amp. TtoP	MAE	MSE
$\widetilde{y}^{gap}$	0.00	0.014	0.073	1.00	10.2	10.6	0.043	0.042	0	0
LIN	0.00	0.027	0.107	0.72	15.8	21.0	0.070	0.071	1.728	0.038
Δ_4	0.03	0.020	0.113	0.38	9.0	9.0	0.064	0.063	2.661	0.097
HP	0.00	0.013	0.066	0.84	8.4	9.9	0.040	0.039	0.564	0.006
CS	0.00	0.011	0.056	0.83	8.5	9.5	0.037	0.037	0.623	0.006
KS	0.00	0.011	0.055	0.85	7.4	9.6	0.033	0.038	0.552	0.005
$MA(16,16)$	0.00	0.019	0.083	0.86	17.7	15.0	0.066	0.061	0.634	0.007
$MA(32)$	0.10	0.032	0.123	0.58	24.0	25.5	0.107	0.111	7.792	0.825
UC	0.00	0.011	0.051	0.81	9.4	8.8	0.036	0.035	0.672	0.007

Table 3: Summary statistics for the output gap derived from detrending the DGP.

timely policy analysis that requires accurate end-point estimation.

The centred MA captures the gap characteristics well, with a MSE of 0.007%, although the cycle durations are much longer than the true cycles. In order to prevent observations being lost at the end-point the data would need to be extended with forecasts of output but this increases the uncertainty associated with the gap measure. The duration and amplitude of the cycles for the filtered MA are substantially larger than those of the generated cycle. A MSE of 0.8% exposes the bias and inefficiency of the estimate. Caution should be applied when using MA derived gaps as they can generate spurious cycles. The UC method provides a reasonable approximation to the gap with a MSE of 0.007%, but it tends to underestimate the gap's magnitude.

None of the methods are wholly accurate. The large negative gap in the recession period 122-132 is not picked up at all. Over this period, trend output rose with an average growth rate of 0.9% compared to 0.6% over the sample. At the same time there was a negative output gap of a maximum magnitude of 2.1% of output. This caused actual output to remain on trend and so the detrending procedures do not pick up either the increase in trend or the negative gap. Whilst a quarterly growth rate of 0.9% is substantial and output is unlikely to grow at 3.6%pa for a prolonged period, the senario is entirely

	Mean	St. Dev	Range	Corr.	Dur. PtoT	Dur. TtoP	Amp. PtoT	Amp. TtoP	MAE	MSE
$\widetilde{y}_{\Delta g}^{gap}$	0.00	0.014	0.073	1.00	10.2	10.6	0.043	0.042	0	0
LIN	0.00	0.043	0.163	0.45	45.0	28.5	0.124	0.128	3.411	0.147
$LIN75$	0.00	0.026	0.122	0.66	28.5	23.3	0.060	0.054	1.689	0.041
Δ_4	0.03	0.022	0.113	0.37	9.0	9.0	0.064	0.064	3.027	0.125
HP	0.00	0.013	0.067	0.83	8.8	9.5	0.041	0.040	0.567	0.006
CS	0.00	0.011	0.056	0.81	8.8	9.5	0.037	0.037	0.625	0.006
KS	0.00	0.012	0.056	0.83	7.4	9.6	0.034	0.039	0.572	0.006
$MA(16,16)$	0.00	0.019	0.085	0.83	17.7	15.0	0.068	0.063	0.716	0.009
$MA(32)$	0.12	0.040	0.140	0.48	24.0	25.5	0.104	0.129	9.266	1.190
UC	0.00	0.011	0.050	0.79	8.3	8.8	0.032	0.035	0.687	0.007

Table 4: Summary statistics for the output gap given an increase in the growth rate of potential output.

plausible. If output is initially on trend, a rise in potential output will result in a negative gap if there is negative contemporaneous correlation between the trend and cycle.[45] Thus, the DGP highlights a very real problem with output gap estimation. The ability to distinguish between shocks to potential output or the output gap depends crucially on the degree of correlation between trend and cycle innovations, as exposed by the divergence between BN and UC models.

The latter part of the 1990s has seen an increase in potential output, particularly in the US, due to a marked acceleration of productivity growth. We examine the impact of an increase in the growth rate of potential output to see whether gap measures pick up a change in trend. The break is defined by shifting the average growth rate from 0.6% to 0.8% per quarter in period 75.[46] The generated data, $\{\widetilde{y}_{\Delta g,t}\}$ is recorded in figure 4, panel a, along with the stochastic trend with the break in growth rate, $\left\{\widetilde{y}^{*}_{\Delta g,t}\right\}$ and the stochastic trend without the break in growth rate for comparison, $\{\widetilde{y}^{*}_{t}\}$. Summary statistics are recorded in table 4.

[45] As discussed in Chapter 2.1, Morley et al. (2002) do find evidence of negative contemporaneous correlation between the trend and cycle for US GDP.

[46] The data was obtained by scaling $\widetilde{y}^{*}$ to a deterministic trend as we did in the initial model.

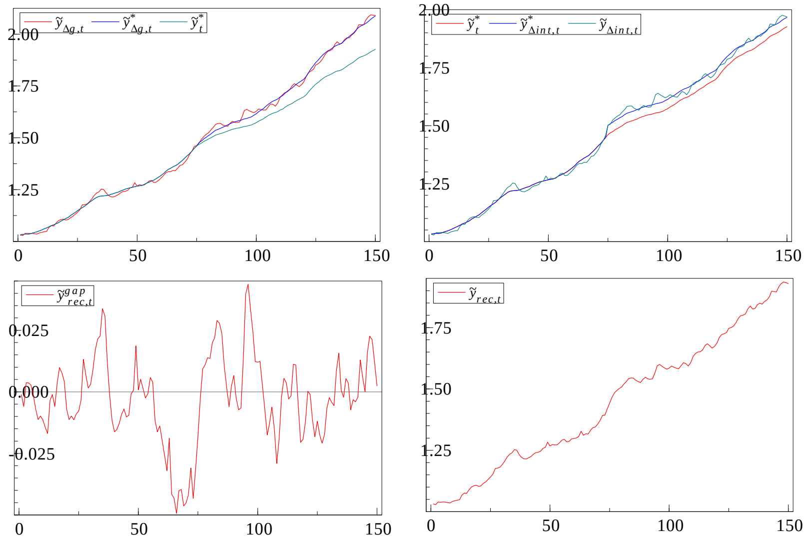

Figure 4: Generated data with an increase in the trend growth rate, an intercept shift in potential output and a deep recession.

The cycle is identical to the initial DGP. LIN performs badly as it does not pick up the increase in trend growth at all. As the method smoothes over the entire sample period, a growth rate above trend is estimated prior to the break and a lower trend is estimated after the break. One would have thought $LIN75$ would perform well as the sample is cut at the known break point and so there is a bias in favour of this method, but a MSE of 0.04% is still large in comparison to other methods. In general, all other measures perform relatively similar to the initial case, with little change in the MSEs. The correlation coefficient falls by 0.02 on average. $MA(32)$ has a very high MSE of 1.2%. The method is not adaptable as it takes 32 quarters following the break to fully capture the change in growth rate.

The second case we examine is an intercept shift in the trend, captured by a 4% increase in output in period 75. This is modelled as a one-off shock with the growth rate remaining at a constant rate of approximately 0.6%. Figure 4, panel b, records the generated output data $\{\widetilde{y}_{\Delta int,t}\}$ along with the stochastic trend $\left\{\widetilde{y}^*_{\Delta int,t}\right\}$ and the trend without the intercept shift, $\{\widetilde{y}^*_t\}$. Summary statistics are recorded in table 5.

The cycle is again identical to the DGP. LIN performs poorly, both before and after the break point. $LIN75$ is comparable to the linear trend in the initial case as expected, although the method actually estimates an upward shift in the trend of 8.5% of output, which is more than double the actual break. The method could not be used in policy applications as the break would not be known at the break origin. There is a fall in the correlation coefficient of 0.04 on average due to the break and the MSEs rise marginally. Regarding the adaptability of the measures to the shift in trend, we can measure the length of time it takes for the estimated trend to equal the actual trend following the break. For the intermittent period the gap will be biased upwards. KS is the most adaptive method, taking 5 quarters for the deviation to be eliminated following the shock. The CS and UC trends also adjust quite quickly, taking 6 quarters for the estimated trend to rise to the actual trend. The HP filter takes 10 quarters, but the least adaptive are the moving averages. The $MA(16,16)$ takes 15 quarters for the estimated trend to equal the actual trend. The mean adjusted $MA(32)$ takes 20 quarters. Also note that the smoothers will estimate a negative gap prior to the break as the trend will incorporate the future intercept rise into their current estimates.[47]

The final case we examine is a deep recession with no change in trend growth. The cycle is augmented by a recession of magnitude 5.9% of output at its maximum, lasting for 20 periods (5 years) dated from period 57-76. Figure 4, panels c and d, record the

[47] This is not as clearly observed in the output gap measures because there is a negative output gap prior to period 75 in the DGP. However, the estimates of this gap are larger than the generated gap by approximately 1% on average in the 2-3 periods prior to the break. Thus, biases resulting from the smoothing properties of the potential output measures do impact substantially on the gap measurement.

	Mean	St. Dev	Range	Corr.	Dur. PtoT	Dur. TtoP	Amp. PtoT	Amp. TtoP	MAE	MSE
$\widetilde{y}^{gap}_{\Delta int}$	0.00	0.014	0.073	1.00	10.2	10.6	0.043	0.042	0	0
LIN	0.00	0.032	0.133	0.70	24.0	26.3	0.115	0.091	2.041	0.058
$LIN75$	0.00	0.023	0.114	0.71	28.5	26.3	0.056	0.048	1.415	0.028
Δ_4	0.03	0.023	0.143	0.36	8.4	6.7	0.058	0.058	2.768	0.111
HP	0.00	0.015	0.068	0.81	8.8	9.5	0.042	0.041	0.645	0.008
CS	0.00	0.012	0.056	0.79	8.5	9.5	0.038	0.038	0.667	0.007
KS	0.00	0.012	0.056	0.84	7.5	9.6	0.034	0.039	0.565	0.006
$MA(16,16)$	0.00	0.022	0.090	0.82	17.7	15.0	0.072	0.065	0.804	0.013
$MA(32)$	0.10	0.038	0.152	0.60	22.8	24.3	0.122	0.126	8.219	0.936
UC	0.00	0.012	0.052	0.78	9.4	8.8	0.037	0.036	0.694	0.007

Table 5: Summary statistics for the output gap given an intercept shift in potential output.

output gap $\{\widetilde{y}^{gap}_{rec,t}\}$ and the generated data $\{\widetilde{y}_{rec,t}\}$ respectively. The motivation behind this case is the Great Depression. Hendry (2000b) finds that fitting a single linear trend from 1860 predicts the Great Depression to last until 1960, whereas a split linear trend estimates that there was no Great Depression at all. Alternatively, the Cubic Spline produces a trend output that declines too rapidly over the recession. Also, as the CS smoothes data, there is a pre-depression boom induced by the early downturn of trend output. Many detrending procedures attribute too much of the fall in output to a decline in the trend, underestimating the magnitude of the gap and predicting the depression to end well before it actually did. The implications of such inferences could be severe. If a government thought the economy was out of a recession when it was actually still trying to recover, inappropriate policy may well be implemented. Summary statistics are reported in table 6.

The split linear trend has break points in periods 62 and 81, which lags the beginning and end of the recession by a year. This is intended to capture the effect of the decline in output being ascribed to a shift in the trend as opposed to a recession, which is only picked up with a delay. The measure performs poorly as it is capturing the wrong effect,

	Mean	St. Dev	Range	Corr.	Dur. PtoT	Dur. TtoP	Amp. PtoT	Amp. TtoP	MAE	MSE
$\widetilde{y}_{rec}^{gap}$	-0.003	0.017	0.093	1.00	10.0	9.2	0.049	0.047	0	0
LIN	0.000	0.030	0.126	0.78	28	23.7	0.114	0.094	1.777	0.039
$LIN(62,81)$	0.000	0.023	0.108	0.64	19	19.8	0.083	0.078	1.480	0.032
Δ_4	0.025	0.021	0.127	0.26	8.6	9.4	0.054	0.065	2.917	0.128
HP	0.000	0.014	0.067	0.81	7.9	9.0	0.035	0.038	0.796	0.011
CS	0.000	0.012	0.056	0.76	8.3	8.6	0.033	0.034	0.855	0.013
KS	0.000	0.012	0.056	0.79	8.3	8.8	0.031	0.036	0.768	0.013
$MA(16,16)$	0.000	0.021	0.092	0.85	15.7	16.5	0.069	0.063	0.699	0.010
$MA(32)$	0.099	0.035	0.135	0.60	24	25.5	0.113	0.117	8.048	0.883
UC	0.000	0.012	0.055	0.77	8.2	8.4	0.032	0.036	0.850	0.013

Table 6: Summary statistics for the output gap given a deep recession.

emphasizing the importance of recognizing whether shocks are attributable to short-run demand side shocks or long-term potential output shocks. The HP, CS, KS and UC trends all drop too rapidly over the recession. The estimated trend is approximately 2% below the actual trend, reducing the estimated magnitude of the negative gap by 2% of output. MSEs double from the baseline case. Only $MA(16,16)$ estimates a comparable sized gap, but it also tends to overpredict the size of the gap when the gaps are not large. Trend output estimated by HP, CS, KS and UC all fall below the actual trend 4-5 quarters prior to the beginning of the recession, inducing an upward bias in the gap. This is a systematic problem with all smoothers.

In conclusion, none of the univariate detrending procedures have accurately estimated the unobserved trend and cycle based on the above DGP given a variety of different scenarios. Admittedly the summary statistics are raw but an ocular judgement is often very informative in comparing the results. However, whilst the gap estimates at any particular point in time are imprecise, the broad profile of the gap is similar across the range of methods examined. The notable exceptions were the two filters, Δ_4 and $MA(32)$. The preference between measures depends upon the trade-off between smoothness and adaptablilty to breaks. Priors regarding the diffusion of shocks and the degree of correlation between the trend and cycle will drive the choice as to which methods are most appropriate.

In the initial case, the MSE for the average of all gap measures excluding the filters (Δ_4 and $MA(32)$) is 0.004%, which represents an improvement on all of the gap measures. If each method measures the true gap with error, then an average would extract the signal relative to the errors. The results suggest that a composite measure of the various gaps may be beneficial (see Chapter 4.3). Also, given the unsatisfactory nature of the univariate decompositions, a multivariate analysis of excess demand pressures using a broader information set is likely to improve the estimation results of these latent variables. The caveat applies that the results are dependent upon the nature of the DGP. If the DGP does not accurately reflect actual output, the above results are not informative in determining which methods 'best' detrend output.

4 Current Measures of the Output Gap

This chapter estimates the UK output gap for the period 1965q1-2002q2 using a broad variety of methods, with the aim of comparing methods based on simple and transparent statistical tests. The estimation of both a static and dynamic production function is outlined initially. A composite measure of the gap, based on principal components analysis, is obtained and finally the various measures are assessed according to a broad range of summary statistics. The methods applied include:

- Linear Trend with break in 1980, [$LIN80$]
- Fourth Difference Filter, [Δ_4]
- Hodrick-Prescott Filter, [HP]
- Cubic Spline, [CS]
- Kernal Smoother, [KS]
- Moving Average with a lag of 32, [$MA(32)$]
- Centred Moving Average with lag and lead of 16, [$MA(16, 16)$]
- Unobserved Components Model with smooth trend,[48] [UC]
- Excess demand for goods and services (Hendry, 2001),[49] [$xd(goods)$]
- Production Function approach, [$PF(Stat)$]
- Dynamic Production Function approach, [$PF(Dyn)$]
- Principal Components, [PC]

[48] See Appendix 3 for estimates of various Unobserved Components models. The smooth trend model with a fixed level and stochastic slope was chosen on the basis of this analysis.

[49] Hendry (2001) estimates excess demand for goods and services as:

$$\begin{aligned} cap_t &= \beta_0 + \beta_1 t + \alpha\left(kpe_t\right), \\ y_t^{gap} &= ype_t - cap_t, \end{aligned}$$

where kpe = capital per worker and ype = output per worker. Under competitive market conditions and constant returns to scale, the factor shares measure their marginal productivities. Note, however, that with imperfectly competitive market conditions, the capital share will include monopoly rents and will thus overstate the marginal productivity of capital services.

4.1 The Production Function Approach

The production function (PF) method is one of the most popular methods of measuring the output gap and is used by the IMF, OECD and most Central Banks.[50] If we assume a Cobb-Douglas technology with constant returns to scale, an elasticity of substitution equal to unity and Hicks-neutral productivity, the PF is given as:

$$y_t = A_t N_t^{\alpha} K_t^{1-\alpha}, \tag{41}$$

where N_t is labour input, K_t is capital input, A_t is total factor productivity (TFP), or the efficiency with which both capital and labour are used to produce output, and α is the elasticity of output with respect to labour $(0 < \alpha < 1)$.[51]

N_t comprises employment, L_t, and the number of paid hours worked per employee, H_t. The normal number of working weeks in a year should be incorporated into the labour input equation but we can assume this has remained relatively constant over the period in question. Employment is broken down into three determinants (lower case represents logs):

$$l_t = wpop_t + pr_t + er_t, \tag{42}$$

where $wpop_t$ is the population of working age, pr_t is the participation rate and er_t is the employment rate.[52] Labour input should be adjusted for labour quality, which is often proxied by educational attainment. Excluding this will imply that changes to labour quality will be picked up in the residual. l_t is recorded in figure 5, panel a, along with trend employment. H_t would be approximated by the difference between average overtime hours and average undertime hours, but as the impact of short-time is negligible hours

[50]It is the recommended approach by the Economic Policy Committee of the EU (2001) and, for the OECD, Giorno et al. (1995) conclude that "the production function approach for estimating potential output provides the best method for estimating output gaps" (p.2).

[51]Under certain conditions, α is the capital and labour participation in income. Appendix 4 provides details as to the calculation of α.

[52]The participation rate is given by: $\frac{\text{No.employed + No.unemployed}}{\text{Population of working age}}$.

can be calculated as:

$$h_t \approx \ln\left(\overline{H}_t\left(1+OH_t\right)\right), \tag{43}$$

where $\overline{H}_t$ is the normal number of hours worked per week, recorded in figure 5, panel b and OH_t is the number of overtime hours worked per week. Muellbauer (1984) finds that data on average hours provides a good approximation to labour utilization.[53] $\overline{H}_t$ declined from 39 hours in 1965 to 32 hours in 2002. The implied fall in output is offset by an increase in efficiency that will be captured in A_t. Average hours are adjusted for a zero mean.

Capital input, K_t, is measured by the net capital stock excluding the dwellings sector, J_t.[54] This is a wealth measure of capital, which weights different types of capital by their asset prices. The ideal measure would be capital services, which measures the flow of productive input from capital. In order to move from this theoretical concept to the available data, an assumption that capital services are proportional to the asset value measure of capital stock must be made.[55] This may well have implications for the order

[53] Muellbauer (1984) proposes a measure of labour utilization based average weekly overtime hours. High utilization rates will arise during periods of high overtime but the corresponding undertime hours will not be observed. Hence, the mean of the truncated upper tail of the distribution of utilization across firms is observed. Given a constant spread, the mean utilization rate can be determined from this truncated upper tail. There is also the problem of an increase in systematic overtime which has accompanied the fall in normal hours. Instead of constructing this series, hours can be approximated by equation (43). Note that short-time is typically less than 10% of overtime.

[54] There is a substantial literature on the treatment of depreciation. The commonly used perpetual inventory method has many problems, see Miller (1983). Perpetual inventory capital stock estimates have serious limitations as measures of capital input because they do not contain any measure of the retirement of capital or transfers to other industries, other than estimates based on historical lengths of life. For example, a capital saving innovation that increases investment and retirements increases the measured capital stock as it does not include the induced retirements. This can lead to serious errors in the measurement of TFP. The capital stock data used in this analysis is obtained from the Bank of England, see Oulton and Srinivasan (2003).

[55] The 'new economy' has had a dramatic impact on the evolution of capital, see Oulton (2001) for a detailed discussion on the impact of ICT growth on output and productivity. The assumption of capital services being proportional to the asset value measure of capital stock may not be feasible if the average life of the stock is changing. This is the case with ICT, which has a much shorter lifespan than traditional capital stock. The substantial ICT capital deepening over the 1990s has led to a divergence in the effects of ICT capital and traditional capital on growth. One way to overcome this problem would be to split the two types of capital in the production function. The IMF are currently looking at this approach.

of integration of capital inputs as we are essentially trying to capture a flow concept by a stock variable. There is some debate as to whether J_t should be adjusted for the degree of capacity utilization, $U_{c,t}$, but we find the utilization variable substantially reduces the procyclicality of the residual:[56]

$$k_t = j_t + U_{c,t}. \tag{44}$$

J_t and $U_{c,t}$ are recorded in figure 5, panels c and d respectively. The CBI produce survey data on the number of respondents reporting that they are operating below normal capacity levels. A capacity utilization variable is constructed on the basis of this survey using the method outlined in Muellbauer (1984), see Appendix 5. Full capacity is assumed to use approximately 91% of the total capital stock available. It should be emphasized that the data applies to manufacturing output. As services have increased dramatically over the period of estimation and the relationship between utilization rates for manufacturing and services is ambiguous, the utilization measure may be a poor approximation. A shortage of data on capacity utilization levels in the service sector prevents a more rigorous, disaggregated measure being derived. It is imperative that the inputs are corrected for utilization of labour and capital. Their exclusion from the PF analysis will result in the residual, representing technological change, picking up many procyclical movements in utilization. The first RBC proponents who admitted this were Burnside, Eichenbaum and Rebelo (1993) who found that the procyclical nature of TFP is vastly reduced when accounting for labour hoarding via an 'effort' variable.

In order to calculate potential output, we need estimates of the latent variables, potential capital, labour and TFP (denoted by superscript *). For capital input, we assume

[56] If capital is thought of as simply being an overhead, a capacity utilization variable may not add much information. However, if respondents to the survey are referring to a much broader measure of capacity than labour inputs alone, a separate capacity measure should be included. Muellbauer (1984) finds that a separate capacity utilization index is dominated by an overtime hours based concept of utilization. However, we find enough variation between the two measures to recommend using both adjustments in the production function approach.

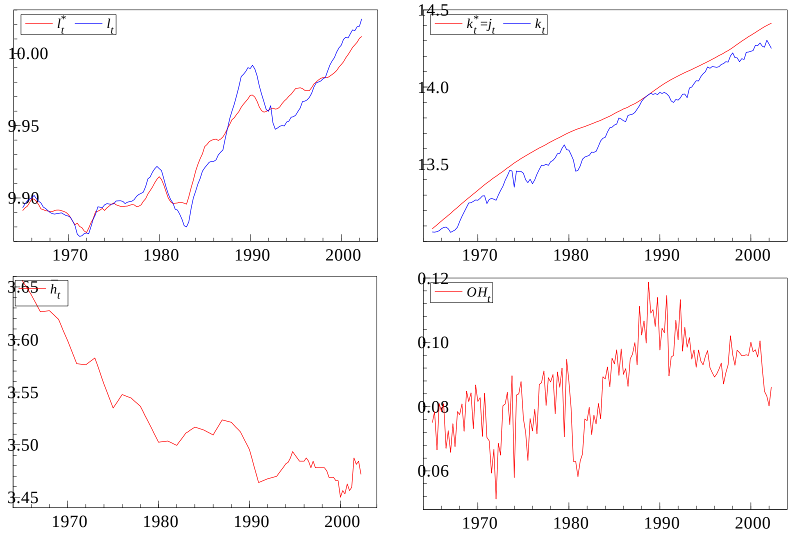

Figure 5: Employment, average hours, capital stock and utilisation of capital.

that capital is always operating at full capacity, hence $U_{c,t} = 0$ and $k^* = j$. As capital stock can be thought of as an indicator of overall capacity, there is no need to smooth the series. Even though net investment per annum is very volatile it is such a small fraction of net capital stock as to have a very limited impact on the stock of capital.

The working population is assumed to be at trend. Most movements in the working population could be thought of as being long-run or permanent changes caused by, for example, a change in pensions provisions, changes in the age of retirement or an increase in the number of women who work. There may be a small cyclical component to the working population, e.g., in the climate of a recession some members may choose to remove themselves from the working population pool by retiring early or choosing not to search for a job, but we shall assume that this effect is negligible. The trend employment rate is derived from the trend unemployment rate, which is used as a proxy for the NAIRU.[57] This

[57]The NAIRU is another latent variable that is notoriously difficult to measure. Staiger, Stock and Watson (1996) investigate the precision of NAIRU estimates based on a variety of models. They conclude that the natural rate is imprecisely measured. For example, for a typical value of the US NAIRU in 1990

is calculated using the UC method of decomposition based on a stochastic level and cycle. This differs from a more structural approach, such as a time varying structural NAIRU which embodies a shifting composition of the labour force as in Coen and Hickman (2002) or a Phillips curve type approach.[58] To estimate the trend participation rate, the total number in employment is smoothed using a HP filter and the level of unemployment is derived from the trend unemployment rate based on a UC decomposition. These are then divided by the actual working population to result in the potential participation rate. Hence, the PF is not immune to the problems of univariate statistical detrending. $\overline{H}$ is assumed to pick up long-run trends only. Any cyclical fluctuations will not be captured in $\overline{H}$ due to labour hoarding. Also, overtime hours are assumed to be 0, therefore $h_t^* = \overline{h}_t$.

The calculation of trend a_t depends on the assumptions made regarding the nature of TFP growth. Theories of technological progress (TP) range from standard neoclassical growth models that regard TP as exogenous to endogenous growth models which assume TP is the result of investment activities. The decision as to which methods are appropriate for detrending a_t depends crucially on whether technical innovations are thought to be random shocks due to a burst of new ideas or whether ideas diffuse gradually as learning is slowly accumulated. This is more of a theoretical question as TFP is a latent variable

of 6.2%, the 95% confidence bands extend from 5.1% to 7.7%.

[58] The method used by the OECD, outlined in Elmeskov (1993) is highly questionable and they are currently looking at methods of improving their estimates for the PF calculations. The method assumes that the change in wage inflation is proportional to the gap between actual unemployment and the NAWRU:

$$\Delta^2 w_t = -\alpha \left(U_t - NAWRU_t \right),$$

where w = ln(wages) and U = unemployment. Also, α is determined by assuming that the NAWRU is constant between two time periods. Hence we can derive the NAWRU as:

$$NAWRU_t = U_t - \frac{\Delta^2 w_t}{-\alpha}, \qquad \alpha = -\frac{\Delta^3 w_t}{\Delta U_t}$$

The NAWRU is highly volatile, as a result of dividing by third differences. The authors then smooth the NAWRU using a HP filter, and so little is gained from using this method. On annual data, the NAWRU is heavily driven by outliers.

but one may expect productivity shocks to take their time feeding through as the learning process, along with R&D, occurs. Also, shocks that are specific to sectors are likely to only impact gradually in the aggregate.[59] Hence, a plausible trend may be quite smooth but will also allow for random productivity shocks. a_t will also pick up efficiency gains in the quality of capital and labour. We use a UC model to detrend a_t based on a smooth trend. This analysis does not address the determinants of TFP growth, which is a flourishing literature in itself.

The resulting output gap is given in figure 6, panel a. The positive mean gap of 0.0002 is negligible. The 1980s recession is estimated to be a lot deeper than the 1990s recession, reaching a magnitude of 3.6% compared to 2.0% of output in the early 1990s. This may be due to the sharp drop in normal hours at the beginning of the 1990s that is unlikely to be offset by increasing productivity due to efficiency gains, causing lower potential output and reducing the size of the negative gap. Panel b records annual actual and trend output growth. There is some divergence at the end of sample. The rise in potential output and the corresponding fall in the gap after 2000 may be driven by the rise in normal hours, which is reversing the previous trend, although the use of the HP filter in detrending pr_t will bias the gap towards 0 at the end-point. Panel c records a_t and the smoothed estimate based on a UC model with fixed level and stochastic slope $\{a_t^* (UC)\}$. The lack of cyclicality in TFP shows that the utilization rates have accounted for business cycle fluctuations. Panel d records the estimated gap against the HP gap measure. Whilst there is some consensus between the two measures, there are periods in which the estimates diverge significantly. Chapter 4.4 compares the gap measures in more

[59] See Caporale (1997), who undertakes a disaggregated analysis of RBC models in order to examine the impact of sector specific versus aggregate shocks.

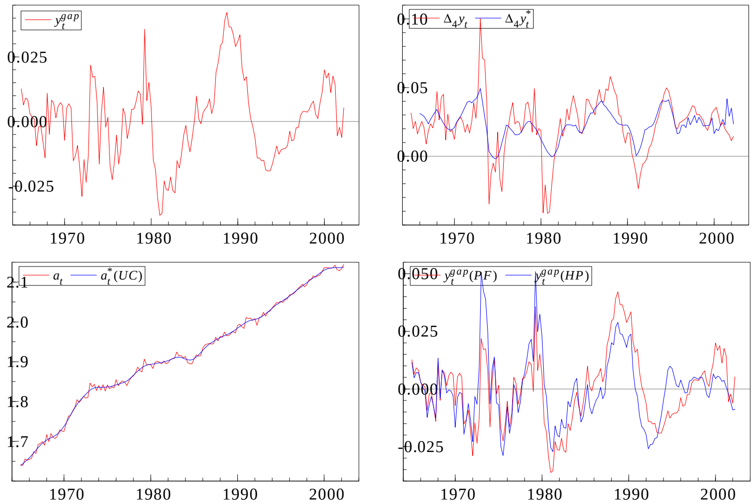

Figure 6: The static production function output gap, output growth and TFP.

detail.

A brief discussion on the order of integration of the inputs is warranted. Output is I(1). As Y comprises C, I, G, X and M, one would expect these components to be I(1) variables as well but if net investment is an I(1) variable, K would be I(2) as it is the cumulation if investment. As the PF models Y as a function of K and N, this would imply Y would also be I(2). The cyclical argument can be followed, suggesting I ~I(2) and K ~I(3) etc. If instead, K were I(1), I would be I(0). This raises the question of whether a consistent accounting solution is obtainable. Assuming K ~I(2), we would require human capital to be I(2) as well in order to give a cointegrating relation for Y ~I(1). The population is I(2) so this argument is plausible. Hence, capital per person will be I(1) as it is calculated as the division of an I(2) variable by and I(2) variable but output per person will be an I(1) variable divided by an I(2) variable. The explanation to this dilemma lies in K and N both containing I(2) components, but to differing extents. Table 7 provides Dickey-Fuller

Null	y	k	$wpop$	l	$k-l$	$y-l$	a
I(1)	-2.813 (0.92)	-3.287 (0.99)	-3.413 (0.98)	-2.502 (0.97)	-3.391 (0.98)	-3.110 (0.89)	-2.848 (0.89)
I(2)	-5.295^{**} (0.25)	-2.295 (0.94)	-3.212 (0.87)	-5.424^{**} (0.65)	-5.043^{**} (0.69)	-5.966^{**} (0.05)	-5.891^{**} (-0.47)
I(3)	-9.259^{**} (-2.59)	-4.347^{**} (0.15)	-11.84^{**} (-0.00)	-11.57^{**} (-0.53)	-11.80^{**} (-0.57)	-9.625 (-2.90)	-10.58^{**} (-3.81)

Table 7: Dickey-Fuller tests for integration.

statistics for the production function variables.[60]

The production function is a static and cointegrating concept. Hence, the standard growth accounting framework should be sufficient. However, the presence of substantial measurement errors in K, N and A imply that a stable relationship may be difficult to identify. Haavelmo (1944) highlights the problem of measurement errors by distinguishing between the latent variables identified in economic theory, their correctly measured empirical counterparts and the actual data available which contains substantial measurement error. For example, in the case of capital, theory tells us that we need a measure of the flow of capital services in the economy, whereas our data is a measure of the capital stock which contains errors due to the assumptions made about depreciation, scrapping, aggregation etc. Some method of allowance for measurement error is required. To do this we analyze the PF in a log-linear dynamic setting, which enables us to find a stable solution for potential output. This approach has the added advantage of setting the PF in the long-run context. Firms do not produce to the PF constraint on a short-run basis. The magnitude and volatility of inventories highlights this fact. In the short-run, firms tend to produce to inventory or order and then sell from these. However, in the long-run the PF constraints will bite, so a dynamic model that allows for adjustments over the short and

[60] The ADF statistic is reported where the number of lags was determined by the highest significant lag. The estimated coefficient on the LDV is reported in parentheses. Note that β is reported (as given in PcGive) which corresponds to $\gamma = \beta - 1$ in footnote 33. Critical Values are -3.441 at 5% significance and -4.023 at 1% significance. Asterisks * and ** denote rejection at the 5% and 1% critical values.

medium term is appropriate.

Measurement errors can be understood in the errors in variables framework. Actual capital stock $\left\{k_t^\bigstar\right\}$ is defined as a stochastic process with a joint sequential density $D_{k^\bigstar}\left(k_t^\bigstar|\mathbf{\Omega}_{t-1},\theta\right)$ for the population parameter, $\theta_p\in\mathbb{R}^l$, given by:

$$k_t^\bigstar|\mathbf{\Omega}_{t-1}\sim \mathrm{N}\left(\lambda k_{t-1}^\bigstar,\mathbf{\Sigma}\right), \tag{45}$$

where $\mathbf{\Omega}_{t-1}=\left(k_1^\bigstar,...,k_{t-1}^\bigstar\right)$. Empirically, $\lambda\approx 0.99$ on quarterly data. Equation (45) can be given as:

$$k_t^\bigstar=\lambda k_{t-1}^\bigstar+\nu_t, \tag{46}$$

where ν_t ~NID$(0,\sigma_\nu^2)$, $E\left[k_t^\bigstar|\Omega_{t-1}\right]=\lambda k_{t-1}^\bigstar$ and $E\left[k_{t-1}^\bigstar\nu_t\right]=0$. The observed capital stock $\{k_t\}$ is contaminated by measurement error, $\{u_t\}$:

$$k_t=k_t^\bigstar+u_t, \tag{47}$$

where u_t ~NID$(0,\sigma_u^2)$. For simplicity we can initially assume that the measurement error $\{u_t\}$ is serially independent and $E\left[k_t^\bigstar u_t\right]=0,\forall t$, which implies that $E\left[\nu_t u_t\right]=0$. The measurement error is systematic and can be represented as an MA(1) process. In practice, $\rho\approx 0.9$ and $\sigma_e^2\approx\sigma_\nu^2$.

$$u_t=\rho u_{t-1}+e_t, \tag{48}$$

where e_t ~NID$(0,\sigma_e^2)$. Therefore:

$$Y_t=AN^\alpha K_t^{\bigstar 1-\alpha}\exp\left(e_t\right). \tag{49}$$

Thus, the PF is static and cointegrating but the presence of systematic and substantial measurement errors requires a dynamic framework. A natural extension to this analysis would be to examine Monte Carlo evidence in order to quantify the impact of the measurement error, see Hendry (1995a).

4.2 A Dynamic Production Function

The dynamic PF model is set in the single equation dynamic framework with a time varying regression intercept that captures unobserved TFP and is augmented by I(0) cyclical factors. The long-run solution, proxying potential output, will be based on the static PF model:

$$y_t^* = \Psi_t + \gamma_1 k_t + \gamma_2 n_t, \tag{50}$$

where Ψ_t is a local level with drift intercept term capturing a_t. We assume that a single equation analysis of Δy_t is valid. This requires that n_t, and k_t are weakly exogenous for y_t. Given an ADL(1,1) model:

$$\begin{aligned} y_t &= \psi a_t + \beta_1 y_{t-1} + \beta_2 k_t + \beta_3 k_{t-1} + \beta_4 n_t + \\ &\quad \beta_5 n_{t-1} + \boldsymbol{\delta}' \left(\text{cyclical factors}\right) + \varepsilon_t, \end{aligned} \tag{51}$$

we can estimate the model in ECM form:

$$\begin{aligned} \Delta y_t &= \psi a_t + \beta_2 \Delta k_t + \beta_4 \Delta n_t + (\beta_1 - 1)\left[y_{t-1} - \kappa_1 k_{t-1} - \kappa_2 n_{t-1}\right] \\ &\quad + \delta_1 \Delta oh_{t-i} + \delta_2 \Delta U_{c,t-i} + \delta_3 \Delta invent_{t-i} + \varepsilon_t, \end{aligned} \tag{52}$$

where $\kappa_1 = \frac{\beta_2+\beta_3}{1-\beta_1}$, $\kappa_2 = \frac{\beta_4+\beta_5}{1-\beta_1}$, and $\varepsilon_t \sim$NID$(0, \sigma_\varepsilon^2)$. The cyclical factors include the change in overtime hours, Δoh, change in capacity utilization, ΔU_c, and change in inventories, $\Delta invent$. The time varying intercept evolves according to the transition equation:

$$a_t = a_{t-1} + \mu + \eta_t, \tag{53}$$

where $\eta_t \sim$NID$(0, \sigma_\eta^2)$. The assumption that σ_ε^2 and σ_η^2 are independently distributed is made. The model is written in SSF and estimated using the Kalman Filter. Equation (52) is generalized to allow for a broader dynamic structure, which is identified using a general to specific modelling strategy.

The use of the time varying trend, modelled as a RW with drift, allows for permanent shifts in TFP. This will robustify the coefficient estimates against the effects of any structural change. The time varying trend will proxy advances in human capital, including knowledge accumulation, experience and educational improvements. Human capital is captured by the process of cohort arrival and departure in the labour force. Those departing from the workforce tend to be due to retirements and have a lot of experience but were educated a long time ago whereas new arrivals have a recent education but a lack of experience. Aggregating across all individuals, given that workers are at different stages in their lifecycles, implies a smooth growth in the effective labour force. Also, the effect of human capital using physical capital stock which embodies TP will be well captured by a RW with positive drift.[61]

The resulting model is given in equation (54), (t-ratios are given in parentheses).

$$\begin{aligned}
\Delta y_t &= \underset{(0.58)}{1.103}a_t + \underset{(1.47)}{0.002}\mu - \underset{(-6.90)}{0.413}y_{t-1} + \underset{(1.99)}{0.098}k_{t-1} + \underset{(2.70)}{0.302}n_{t-1} \\
&\quad + \underset{(2.21)}{0.555}\Delta n_t + \underset{(4.59)}{0.031}\Delta invent_t + \underset{(3.38)}{0.156}\Delta U_{c,t} \\
&\quad + \underset{(4.31)}{0.027}ID68q1 + \underset{(5.43)}{0.035}ID73q1 + \underset{(2.66)}{0.027}BD79q2 \\
LL &= 695.054,\ \widehat{\sigma} = 0.664\%,\ \chi^2_{DH}(2) = 1.347,\ Q_{BL}(11,10) = 12.088. \qquad (54)
\end{aligned}$$

The model represents a good fit given the simplicity of the model, with an equation standard error, $\widehat{\sigma}$, of 0.66%. The goodness of fit, R^2_d, is 0.776 and the model passes all diagnostics. $\chi^2_{DH}(2)$ is a test of normality on the residuals based on the Bowman-Shenton statistic with a correction of Doornik and Hansen (1994). The Box-Ljung statistic,

[61] Note that the estimated a_t from the static PF cannot be used. If there is a unit coefficient on TFP in the dynamic model (which would correspond to the static model), the long-run solution will not be identified. As a_t is calculated as $y_t - f(k_t, n_t)$, y_t will cancel out. If TFP was estimated via a regression model, such as $y_t = \beta_0 + \beta_1 k_t + \beta_2 n_t + a_t$, there would be a problem of generated residuals. Pagan (1989) discusses this issue in which a 2-step procedure that uses the estimated residuals from the 1st step in the second stage will cause the estimated standard errors to be incorrect. The model would need to be augmented by the derivative of the residual with respect to the parameters of the 1st stage model for the standard errors to be asymptotically correct.

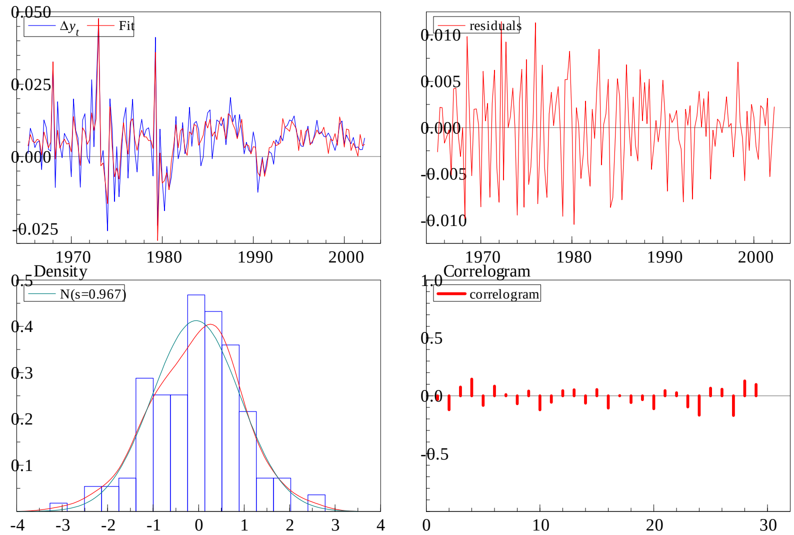

Figure 7: Actual and fitted output growth and the model diagnostics.

$Q_{BL}(11, 10)$, tests the hypothesis that the residuals are uncorrelated up to the 11th order. It is distributed as a $\chi^2(10)$ under the null. Tests for heteroskedasticity, serial correlation at the 1st and 11th lag and the Durbin-Watson test are also satisfactory. Figure 7 records the actual and fitted values along with the diagnostics.

The model has an adjustment coefficient of 0.41, implying that two fifths of the disequilibrium at $t-1$ is removed in the following quarter. The current dated adjustment in k was insignificant (up to 6 lags of Δk were included in the GUM) and so all adjustment to capital takes place in the error correction term. The adjustment term on Δn_t is large. In period t, firms will not only consider whether they were in equilibrium last period but also whether there is a change in labour input in the current period and so current decisions have a direct impact on Δy. Overtime hours are not significant. A convex investment adjustment cost was also included, $\left[\frac{a}{2}I_t^2/K_t\right]$, but was found to be insignificant. As the adjustment to equilibrium for k should be captured in the ECM this is not surprising. Other factors that may impact upon output growth in the short-run such as real interest rates and real exchange rates were included in the GUM but these were also found to be insignificant.

BD79q2 is a blip dummy (i.e. 79q2=1,79q3=-1) and hence integrates to an impulse dummy which does not enter into the long-run solution. With regard to the impulse dummies, $ID68q1$ and $ID73q1$, whilst they are highly significant, their coefficients are reasonably small. A simple plot of output shows the impulses not to be persistent and we conclude that the dummies are capturing one-off shocks or outliers and should not enter the long-run solution as level shifts. They are not included as blip dummies because the counteracting residuals do not occur in the immediate quarter following the positive shock but over the following year and summing the negative residuals over the following 4 quarters removes the majority of the shock. The time varying trend will capture the persistent shocks to output.

Equation (54) suggests that we can impose a restriction of constant returns to scale. Reparameterising the model results in equation (55).

$$\begin{aligned}
\Delta\left(y-n\right)_t &= \underset{(1.75)}{0.975}a_t + \underset{(2.59)}{0.001}\mu - \underset{(-6.11)}{0.375}\left(y-n\right)_{t-1} + \underset{(2.29)}{0.112}\left(k-n\right)_{t-1} + \\
&\quad +\underset{(3.64)}{0.027}\Delta invent_t + \underset{(2.59)}{0.137}\Delta U_{c,t} + \underset{(3.44)}{0.025}ID68q1 \\
&\quad +\underset{(4.11)}{0.030}ID73q1 + \underset{(5.57)}{0.027}BD79q2 \\
LL &= 682.332,\ \widehat{\sigma} = 0.70\%,\ \chi^2_{DH}\left(2\right) = 4.016,\ Q_{BL}\left(11,10\right) = 5.095. \qquad (55)
\end{aligned}$$

The model passes all diagnostics and the equation standard error is only marginally increased to 0.7%. By estimating the model in terms of output and capital per capita, the estimated coefficients on n_{t-1} and Δn_t become insignificant. The parameters are relatively stable when imposing the restriction and the drift of 0.1% is now significant. Figure 8, panel a records the estimated local level, proxying TFP. It clearly shows the productivity

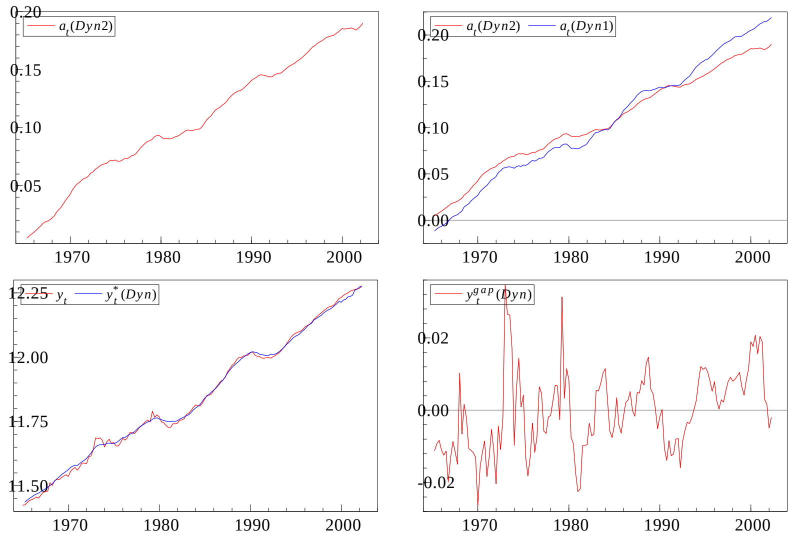

Figure 8: TFP, trend output and the output gap estimated by the dynamic model with time varying intercept.

slowdown in the 1970s, the increase in the second half of the 1980s and the 'new economy' productivity increases of the late 1990s, although this does tail off considerably from 2000. TFP enters the short-run dynamic model with a near unit coefficient resulting in growth of approximately 20% over the period of estimation. Note that the long-run solution determines the total growth in TFP over the period. The q-ratio, determined as the ratio of the variance of the unobserved component to the variance of the model residuals, is 0.14. Panel b records a_t against a_t estimated by equation (54) (denoted $a_t(Dyn1)$). Whilst both models pick up a similar trend, the drift in the unrestricted model is slightly larger, with TFP growing by approximately 26% over the period if the constant returns to scale assumption in not imposed.

The long-run solution is given as:

$$y_t^* = 0.3k_t + 0.7n_t + 2.6a_t. \tag{56}$$

Figure 8, panel c records the long-run solution, which proxies potential output, against actual output. The trend tracks actual output quite closely resulting in a small output gap, recorded in panel d. Whilst the gaps in the 1970s match those of the static PF gap, the late 1980s boom and early 1990s recession are estimated to be much smaller in the dynamic setting. The shocks in the 1970s are quite clearly attributable to short-run shocks and so are not picked up in the long-run trend, whereas the local level component estimates a slowdown in productivity between 1988 and 1992 that is not picked up in the residual based estimation of TFP to the same extent. The static residual based estimation of TFP estimates growth of approximately 50% over the period. With a coefficient of 2.6 in the long-run solution, TFP growth is approximately 53% in the dynamic model, which is comparable. Output growth between 1965 and 2002 is approximately 130% and so TFP growth accounts for about two fifths of output growth over the period.[62] This is difficult to compare with capital and labour as the growth rates depend on how much technical progress is captured by capital and labour. 40% does seem very plausible though, given that there have been large increases in labour participation over the period of estimation.

4.3 Principal Components Analysis

Principal Component (PC) methods are statistical techniques used for data reduction and originated in Hotelling (1933). The method enables the reduction of data by finding linear combinations of the variables that contain most information. Hence, we can compile a measure of the output gap by assembling all measures of the gap into a vector and taking a linear combination with weights determined by maximizing the canonical correlations between variates. The first PC could be defined as a composite measure of excess demand.

[62] Note that taking log changes as percentage growth rates is incorrect when the changes are far from zero. Using this approximation would estimate that output growth was only 85% over the estimation period.

The reasoning behind this analysis is based on the classic signal extraction problem. For each method, the estimated output gap is a combination of the true gap plus some error:

$$\widehat{y}_{i,t}^{gap} = y_t^{gap} + e_{i,t}, \quad e_{i,t} \sim \mathrm{N}(0, \sigma_{e_i}^2). \tag{57}$$

The principal component, which is essentially a weighted average of the individual measures, should extract the signal relative to the errors. The signal to noise ratio will be given by $\sigma_{\widehat{y}_i}^2 / \sigma_{e_i}^2$. On a note of caution, the interpretation of the PC is very difficult, limiting the method's use in practice.

Principal Component methods are concerned fundamentally with the eigenvalue and eigenvector structure of covariance matrices. A criticism of the method is that it is not invariant under linear transformations of the variables because such a transformation will change the eigenstructure of the covariance matrix. Hence, the units of measurement are very important. If the units differ the correlation matrix should be used as opposed to the covariance matrix, but problems of inference are exacerbated when using this.

The population PCs shall be derived, based on Muirhead (1982) and Anderson (1984). Assume a random $m \times 1$ vector $\mathbf{X}$ has a normal distribution, $\mathrm{N}(\boldsymbol{\mu}, \Sigma)$, and let $\lambda_1, \geq \lambda_2 \geq \cdots \geq \lambda_m$ (> 0) be the latent roots of Σ. The $m \times m$ orthogonal matrix of eigenvectors, $H = [\mathbf{h}_1 ... \mathbf{h}_m]$, implies:

$$H' \Sigma H = \Lambda = diag\left(\lambda_1, ..., \lambda_m\right). \tag{58}$$

$\mathbf{U}$ is defined as:

$$\mathbf{U} = H' \mathbf{X} = \left(U_1, ..., U_m\right)', \tag{59}$$

where $cov\left(\mathbf{U}\right) = \Lambda$, and hence, $U_1, ..., U_m$ are uncorrelated and $Var\left(U_i\right) = \lambda_i, i = 1, ..., m$. The components $U_1, ..., U_m$ of $\mathbf{U}$ are the PCs of $\mathbf{X}$, and the first PC is given as $U_1 = \mathbf{h}_1' \mathbf{X}$ with variance λ_1. This is the normalized linear combination of the components of $\mathbf{X}$ with the largest possible variance. The second PC will then account for the maximum of the remaining variance and all the components are derived in this manor. The method serves to combine all variables into a composite variable which reflects the maximum possible proportion of the total variation in the set.

The PCs are determined under the condition that they are orthogonal. If we define an arbitrary linear function as $\boldsymbol{\alpha}'\mathbf{X}$ with $Var(\boldsymbol{\alpha}'\mathbf{X}) = \boldsymbol{\alpha}'\Sigma\boldsymbol{\alpha}$, the condition that $\boldsymbol{\alpha}'\mathbf{X}$ is uncorrelated with the ith PC, U_i, is:

$$0 = Cov(\boldsymbol{\alpha}'\mathbf{X}, \mathbf{h}_i'\mathbf{X}) = \boldsymbol{\alpha}'\Sigma\mathbf{h}_i = \lambda_i\boldsymbol{\alpha}'\mathbf{h}_i, \tag{60}$$

as $\Sigma\mathbf{h}_i = \lambda_i\mathbf{h}_i$, so $\boldsymbol{\alpha}$ must be orthogonal to $\mathbf{h}_i$. Two measures that explain the variability in $\mathbf{X}$ are $\operatorname{tr}\Sigma$ and $\det\Sigma$ where:

$$\operatorname{tr}\Sigma = \operatorname{tr} H'\Sigma H = \operatorname{tr}\Lambda = \sum_{i=1}^{m}\lambda_i, \tag{61}$$

$$\det\Sigma = \det H'\Sigma H = \det\Lambda = \prod_{i=1}^{m}\lambda_i. \tag{62}$$

For the sample PCs, suppose $\mathbf{X}_1, ..., \mathbf{X}_N$ is a random sample of size $N = n + 1$ on $\mathbf{X}$. We can define the sample covariance matrix, S, by:

$$A = nS = \sum_{i=1}^{N}\left(\mathbf{X}_i - \overline{\mathbf{X}}\right)\left(\mathbf{X}_i - \overline{\mathbf{X}}\right)'. \tag{63}$$

The latent roots of S (labelled $l_1, ..., l_m$) are estimates of the latent roots $\lambda_1, \geq \cdots \geq \lambda_m$ of Σ. Defining the matrix of normalized eigenvectors, $Q = [\mathbf{q}_1...\mathbf{q}_m]$ such that:

$$Q'SQ = L = diag(l_1, ..., l_m), \tag{64}$$

we can estimate the eigenvector $\mathbf{h}_i$ by the sample. The sample PCs are given as $\widehat{U}_1, ..., \widehat{U}_m$ of $\widehat{\mathbf{U}} = Q'\mathbf{X}$.[63]

[63] In order to test whether the reduction to the PCs is valid, we can test the null that the latent roots of Σ are equal. Accepting the null implies that all the PCs have the same variance, and so there is no reduction in dimension by deriving the PCs. If the $m-1$ smallest roots are equal and small compared with the largest root, the first PC is explaining much of the variability in the sample and there is a valid

	Eigenvalues	Cumulative %
$PC1$	6.828	68.28
$PC2$	1.366	81.94
$PC3$	0.612	88.06
$PC4$	0.422	92.28
$PC5$	0.304	95.32

Table 8: Estimated eigenvalues for the first five principal components of the output gap.

Whilst the decision as to how many factors to extract is arbitrary (because of the lack of interpretation that can be given to the factors) two methods often used include the Scree test and the Kaiser criterion. The Kaiser (1960) criterion suggests retaining factors with eigenvalues greater than 1 as the factor is only then extracting at least as much information as the original variable. The Scree test (Cattell, 1966) suggests plotting the eigenvalues and seeing when the plot smoothes out horizontally.

10 gap measures were included in the analysis: $LIN80$, Δ_4, HP, CS, KS, $MA(32)^*$, UC, $xd(goods)$, $PF(Stat)$ and $PF(Dyn)$. Both BN and $MA(16,16)$ were excluded from the analysis. The Beveridge Nelson smoother is highly volatile and does not estimate a plausible output gap and the centred MA is not timely. Forecasts were not used to extend the series as this increases the uncertainty of the gap measure. Also note that $MA(32)^*$ was adjusted for a zero mean in order to prevent the PC from being biased upwards. Of the 10 components, 5 were required to achieve the 95% level of significance. Table 8 records

---

reduction in dimension. Sequentially testing the null hypothesis:

$$H_k : \lambda_{k+1} = ... = \lambda_m,$$

for $k = 0, 1, ..., m-2$, is based on a likelihood ratio test whereby only the subset of latent roots appear in the statistic. The test statistic is given by:

$$\Lambda_k = V_k^{N/2}, \quad \text{where } V_k = \frac{\prod_{i=k+1}^{m} l_i}{\left(\frac{1}{m-k}\sum_{i=k+1}^{m} l_i\right)^{m-k}},$$

which is asymptotically distributed under the null as $\chi^2_{(q+2)(q-1)/2}$, where q is the number of subvectors that $\mathbf{X}$ is partitioned into. For proof see Muirhead (1982).

	$PC1$	$PC2$	$PC3$	$PC4$	$PC5$
$LIN80$	15.585	7.258	7.669	24.203	-10.528
Δ_4	11.601	-26.622	11.865	-22.805	8.455
HP	24.032	-0.758	-15.081	12.951	12.050
CS	19.034	6.024	-1.448	9.503	3.083
KS	18.626	8.078	-14.975	-8.752	0.044
$(MA32)^*$	10.015	6.204	15.984	-10.875	0.850
UC	11.473	8.967	0.408	-10.900	-23.039
$xd(goods)$	11.288	-31.569	15.261	21.302	-11.220
$PF(Stat)$	20.978	13.376	5.324	4.524	40.616
$PF(Dyn)$	25.126	-29.948	-58.991	-12.114	-5.463

Table 9: Factor loadings for the first five principal components of the output gap.

the estimated eigenvalues. The first PC is taken as our composite measure of the gap as this accounts for 68% of the variation, but the Kaiser criterion would keep the second PC as well. The factor loadings for the PCs are given in table 9. The first component is recorded in figure 9, panel d. $PC1$ does appear to produce a reasonable estimate of the output gap, as can be seen in the summary statistics.

4.4 Summary Statistics

This section aims to provide a simple first pass at a comparison of the methods. Issues of interest when comparing measures of the gap include what the cycles look like in terms of duration and amplitude, whether the cycles are periodic, whether they are asymmetric and whether the measures co-move. To detect turning points, a simple algorithm was applied. A peak (trough) is identified as the highest point during which output is above (below) trend, given that output is above (below) trend for 3 quarters or more in order to avoid innovations around the trend. This is a very naive dating rule.[64] Whilst it is argued that a policy-maker would not adhere to such a simple rule in practice, the approach does provide a straightforward method for comparing many cycles. Scott (2000a) advocates the

[64]See Harding and Pagan (2000) for details of various business cycle dating rules.

use of simple dating rules, arguing that more complex rules are susceptible to the critique that the stylized facts had been 'dialled in'. Note that this algorithm had to be adjusted for the measures of the gap with a mean that was not close to zero.[65]

Harding and Pagan (2000) argue that turning points should not be determined upon the basis of detrended series, but rather methods should be used that identify peaks and troughs without involving the creation of artificial trend and cycle components. The 'classical' cycle is defined as the period between two turning points in the original series whereas we are examining 'growth' cycles, which are derived by deducting the trend component. They note that inference made on the basis of growth cycles is fraught with perils, emphasizing the degree of disparity in the variety of detrending methods used in the literature.

Plots of the estimated gaps are given in figure 9. Panel a records $LIN80$, HP ($\lambda = 1600$), CS ($\alpha = 8$), and KS ($h = 8$), panel b records the MA gaps, panel c records $xd(goods)$ and finally panel d records the Principal Component of the gap measures.[66] Whilst the broad profile of the gap is similar across measures, there are differences in the magnitude of the gap and the timing of booms and recessions. Most divergence occurs at the end-point, highlighting the considerable policy implications. The KS estimates of the gap to have risen substantially between 2000 and 2002, compared to a sharp drop in the gap estimated by $xd(goods)$. The range of estimates in 2002q2 extends from 3.3% to -3.6% of output and 4 measures estimate a positive gap compared to 6 estimating a negative gap.

[65] The turning points for these measures were estimated by correcting for the mean. This was done by setting the gap approximately equal to zero in 1986Q1, when the Treasury estimates the economy to be on trend. Whilst this is an ad hoc method, it does enable the turning points to be estimated and simple ocular judgements suggest that this algorithm does pick up the main turning points.

[66] Note that figure 6 in Chapter 4.1 records $PF(Stat)$, figure 8 in Chapter 4.2 records $PF(Dyn)$ and

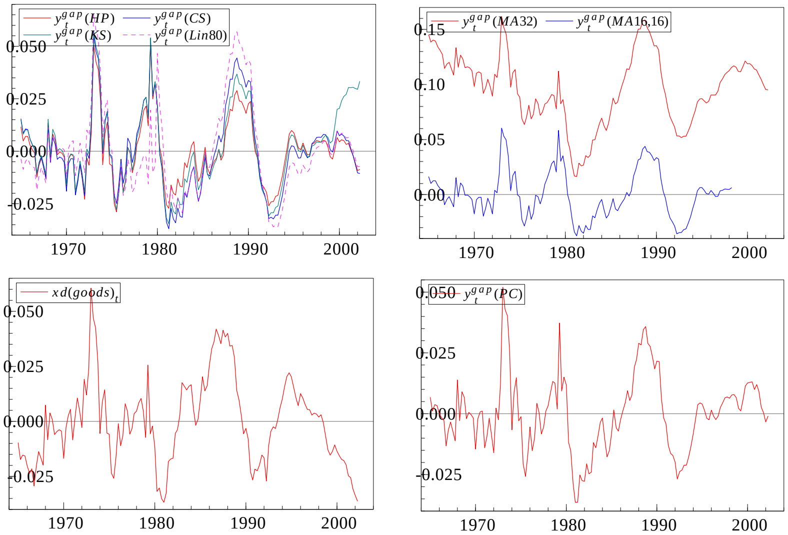

Figure 9: Estimates of the output gap for the UK.

Table 10 reports a variety of summary statistics for the measures.[67] Expansions (contractions) are defined as the phase from trough (peak) to peak (trough). There are a wide range of averages for both duration and amplitude between the gap estimates. However, the duration averages tend to cluster around 3 to 4 years for a contraction and approximately 5 years for an expansion. There is more of a divergence in average amplitudes but most methods tend to produce cycles of approximately 7 percentage points for both contractions and expansions.[68] There is some evidence of asymmetry in cycles, with the duration of expansions approximately a year and a half longer than contractions. This supports the view that the economy gradually builds up pressure throughout an expansion

figure 16, panel c in Appendix 3 records UC.

[67] All statistics reported are multiplied by 100, i.e. percentage of output. Note that Amp/Q = amplitude per quarter.

[68] One could also test for duration dependence to assess whether the cycles are periodic. The Brain-Shapiro (1983) test of duration dependence aims to test whether the longer a series remains in an expansionary (contractionary) phase, the more likely it is to switch to a contractionary (expansionary) phase. However, due to our conclusion of asymmetry there is unlikely to be a periodic cycle and hence the test is not applied. The changing economic environment, from the inflationary 1970s, to the monetarism of the 1980s, commitment to the ERM and a move to inflation targeting in the 1990s, implies that we are unlikely to find evidence of periodic cycles.

	Mean	St.Dev.	Contraction			Expansion		
			Dur.	Amp.	Amp/Q	Dur.	Amp.	Amp/Q
$LIN80$	0.00	2.14	12.33	8.66	0.70	23.75	7.19	0.30
Δ_4	2.29	2.10	7.67	10.24	1.34	21.00	8.57	0.41
HP	0.00	1.49	15.25	6.42	0.42	14.75	6.13	0.42
CS	0.02	1.94	15.75	7.34	0.47	19.25	6.95	0.36
KS	0.24	1.91	17.00	6.57	0.39	16.33	7.66	0.47
$MA(32)$	9.59	3.20	24.50	12.60	0.51	29.00	10.59	0.37
$MA(16,16)$	0.03	2.14	10.67	8.81	0.83	17.67	8.27	0.47
UC	0.05	2.83	12.33	8.74	0.71	24.50	7.73	0.32
$xd(goods)$	0.00	1.94	13.00	7.26	0.56	18.25	6.73	0.37
$PF(Stat)$	0.02	1.58	8.80	4.07	0.46	15.80	4.97	0.31
$PF(Dyn)$	-0.06	1.11	8.40	3.86	0.46	15.50	3.89	0.25
PC	0.00	1.56	11.75	4.28	0.36	17.60	4.50	0.26
$Average$	1.02	1.99	13.12	7.41	0.60	19.45	6.93	0.36

Table 10: Summary statistics for the output gap measures.

	LIN	Δ_4	HP	CS	KS	MA	MA	UC	XD	PF	PF	PC
$LIN80$	1											
Δ_4	0.37	1										
HP	0.79	0.55	1									
CS	0.86	0.49	0.95	1								
KS	0.78	0.47	0.91	0.91	1							
$MA(32)$	0.74	0.50	0.68	0.79	0.74	1						
$MA(16,16)$	0.84	0.49	0.97	0.99	1.00	0.79	1					
UC	0.75	0.37	0.72	0.82	0.84	0.80	0.85	1				
$xd(goods)$	0.47	0.75	0.50	0.48	0.32	0.34	0.50	0.28	1			
$PF(Stat)$	0.77	0.40	0.78	0.86	0.80	0.79	0.83	0.70	0.32	1		
$PF(Dyn)$	0.49	0.61	0.71	0.62	0.70	0.36	0.65	0.52	0.56	0.52	1	
PC	0.87	0.63	0.94	0.96	0.93	0.83	0.96	0.85	0.57	0.86	0.73	1

Table 11: Correlation matrix of output gap measures.

and then, as this bursts and we move into a contraction, the release of pressure is much more rapid. This has serious implications for policy-makers, as addressed in Chapter 2.5.

Another question of interest is whether the measures of the gap co-move with each other. Correlation analysis should tell us whether the different measures give the same signals regarding the economy's position in the cycle. Table 11 reports the correlation matrix for the measures.[69]

[69]Note that the shortened column labels correspond as the transpose of the row labels.

All of the correlations are significant at the 1 percent significance level, implying that the measures almost always co-move.[70] One may naively presume that any measure of the gap may be used as they will all be capturing the same information. However, McDermott and Scott (1999) show that as correlation assesses both amplitude and duration elements, if the amplitude of a swing that is common to both series is large (such as the 1980s boom and recession), this may well dominate the covariance, implying a larger correlation than if duration was assessed alone.[71]

The HP, CS and KS are all highly correlated and pick up similar trends. As these measures are dependent on ad hoc parameter judgements, the high correlation is in a sense 'programmed in' by the choice of parameter. As the PC has a high correlation coefficient with these 3 measures, it is important to ensure that the correlation coefficient for the PC is not biased due to the high correlations between HP, CS and KS. Calculating the PC excluding these measures resulted in correlation coefficients of 0.92, 0.93 and 0.90 between PC and HP, CS and KS respectively, suggesting that these measures are not driving the principal component. This conclusion is supported by the factor loadings, where the weights are fairly well spread.

Whilst the correlation matrix highlights interesting facts regarding the individual mea-

---

[70] The 1 and 5 percent significance levels are given by $2.58\left(1/\sqrt{T}\right)$ and $1.96\left(1/\sqrt{T}\right)$ respectively.

[71] McDermott and Scott (1999) suggest using a concordance statistic as opposed to a correlation statistic, which tests whether different measures signal that the economy is in the same state at the same point in time. The concordance statistic is given as:

$$\begin{aligned} C_{ij} &= T^{-1}\left[\sum\left(S_{i,t}S_{j,t}\right)+\left(1-S_{i,t}\right)\left(1-S_{j,t}\right)\right] \\ \text{where } S_{i(j),t} &= \begin{cases} 1, & \text{if } y_t \text{ is in expansion} \\ 0, & \text{if } y_t \text{ is in contraction} \end{cases} \end{aligned}$$

As this approach is non-parametric, the dating will be almost independent of the sample used. In order to infer significance levels, McDermott and Scott undertook Monte Carlo simulations and computed the response surfaces. Whilst there is some evidence of sensitivity to non-normal innovations, the test has reasonable power properties when the correlation between innovations is high and the power increases dramatically as T is increased.

sures, canonical correlations enable us to examine the correlations between various sets of measures of the gap. Canonical correlation analysis was developed by Hotelling (1936) and is a technique for analyzing the relationship between a linear combination of two sets of variables, each of which can contain several variables such that the correlation between them is maximized. There are direct parallels between Principal Components analysis and canonical correlations. When the variables are regarded as belonging to a single set of variables, PC analysis tends to be used whereas if the variables naturally fall into two sets, canonical correlation analysis can be insightful. The aim is to reduce the correlation structure between two sets of variables **A** and **B** to a simple form by applying linear transformations to the sets. Following Muirhead (1982), partitioning the $(p+q)\times 1$ random vector, **X**, into subvectors **A** and **B**, which are $p\times 1$ and $q\times 1$ respectively, the covariance matrix can be defined as:

$$Cov\begin{pmatrix} A \\ B \end{pmatrix} = \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}, \tag{65}$$

where Σ_{11} is $p\times p$ and Σ_{22} is $q\times q$. Assume $p\leq q$ without loss of generality and let $k = rank\left(\Sigma_{12}\right)$, then there exists a $p\times p$ orthogonal matrix **H** and a $q\times q$ orthogonal matrix **Q** such that:

$$\Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1/2} = \mathbf{H}'\widetilde{P}\mathbf{Q}, \tag{66}$$

where:

$$\widetilde{P} = \begin{bmatrix} \rho_1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \rho_k & \vdots \\ 0 & \cdots & \cdots & 0 \end{bmatrix},$$

and $\rho_1, ..., \rho_k$ are the positive square roots of $\rho_1^2, ..., \rho_k^2$ $(\neq 0)$, which are the latent roots of

$\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$.[72] If we define:

$$
\begin{aligned}
L_1 &= H\Sigma_{11}^{-1/2}, \quad L_2 = Q\Sigma_{22}^{-1/2}, \\
\text{then } L_1\Sigma_{11}L_1' &= I_p, \quad L_2\Sigma_{22}L_2' = I_q \text{ and } L_1\Sigma_{12}L_2' = \widetilde{P}. \qquad (67)
\end{aligned}
$$

Let the first canonical variables be U_1 and V_1, which are linear functions $U_1 = \boldsymbol{\alpha}_1'\mathbf{A}$, $V_1 = \boldsymbol{\beta}_1'\mathbf{B}$. These have the maximum correlation subject to the condition that $Var\left(U_1\right) = Var\left(V_1\right) = 1$. Then if $\mathbf{U} = L_1\mathbf{A}$ and $\mathbf{V} = L_2\mathbf{B}$:

$$
\begin{pmatrix} \mathbf{U} \\ \mathbf{V} \end{pmatrix} = \begin{bmatrix} L_1 & 0 \\ 0 & L_2 \end{bmatrix} \begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix} = L\begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix}, \qquad (68)
$$

$$
Cov\begin{pmatrix} \mathbf{U} \\ \mathbf{V} \end{pmatrix} = L\Sigma L' = \begin{bmatrix} I_p & \widetilde{P} \\ \widetilde{P}' & I_q \end{bmatrix}, \qquad (69)
$$

where $L = diag\left(L_1, L_2\right)$. Thus, we can reduce the covariance matrix, Σ, to a form that only involves the ρ's. Equation (69) is the canonical form of equation (65). If $\mathbf{U}' = \left(U_1, ..., U_p\right)$ and $\mathbf{V}' = \left(V_1, ..., V_q\right)$, U_i and V_i are the ith canonical variables.[73]

The gap measures naturally divide into 2 subsets; univariate statistical methods and multivariate methods.

Set A: $LIN(80)$, Δ_4, HP, CS, KS, $MA(32)$, UC.

Set B: $xd(goods)$, $PF(Stat)$, $PF(Dyn)$.

Table 12 reports the linear combinations for the canonical correlations. The t-tests of significance are based on conditional standard errors. All apart from KS, $MA(32)$ and $PF(Dyn)$ are significant. The eigenvalues for the resulting canonical correlations are given in table 13. The likelihood ratio (LR) is reported, along with the corresponding

[72] For proof see Theorem A9.10, Muirhead (1982).

[73] See Anderson (1984) for a discussion of the properties of canonical correlations.

Canon	Variable	Coeff.	S.E	t-stat (p-val)
U	$LIN80$	12.125	3.284	3.69 (0.00)
	Δ_4	19.015	2.245	8.47 (0.00)
	HP	-26.634	12.164	-2.19 (0.03)
	CS	60.147	10.241	5.87 (0.00)
	KS	-4.914	5.721	-0.86 (0.38)
	$MA(32)$	1.177	2.294	0.51 (0.61)
	UC	-9.643	2.986	-3.23 (0.00)
V	$xd(goods)$	25.367	2.102	12.07 (0.00)
	$PF(Stat)$	44.892	2.501	17.95 (0.00)
	$PF(Dyn)$	2.023	4.060	0.50 (0.62)

Table 12: Canonical correlation estimates.

Can Corr.	Eigenvalue	Cumulative %.	LR	F-test
0.926	5.997	0.632	0.019	$F_{(21,403)} = 56.844$
0.816	1.991	0.842	0.134	$F_{(12,282)} = 40.708$
0.774	1.496	1	0.401	$F_{(5,142)} = 42.482$

Table 13: Estimated eigenvalues for the canonical correlations, and the likelihood ratio tests of significance.

F-statistic.

The canonical correlations are given as $(0.926, 0.816, 0.774)'$. We can use the T^2 statistic proposed by Hotelling (1931) to test for the significance of mean differences in the multivariate case. The generalized T^2 statistic is the multivariate analogue of the square of t and is given as:

$$T^2 = N\left(\overline{x} - \mu\right)' S^{-1}\left(\overline{x} - \mu\right), \tag{70}$$

where $\overline{x}$ is the mean vector of a sample of size N and S is the sample covariance matrix.[74] The Hotelling-Lawley Trace statistic can be converted to the T^2 coefficient by multiplying the trace coefficient by $(N - L)$, where L is the number of groups. Both have the same degrees of freedom and significance level. The Hotelling-Lawley Trace coefficient is 9.484. This gives an F-test coefficient of 62.623, which is compared to a critical value of $F_{(21,416)} \approx 2.38$ at the 1% significance level. Therefore, we can reject the null of significant

[74]If the sample distribution is $\mathrm{N}\left(\mu, \sigma^2\right)$, then $t = \sqrt{N}\frac{\overline{x}-\mu}{s}$ has a t-distribution with $N - 1$ degrees of freedom, where N is the number of observations in the sample.

mean differences and conclude that the two sets of variables are not independent.[75] Univariate statistical detrending methods are picking up common information to multivariate methods that include more information.

4.5 Cointegration of Potential Output Measures

Whilst canonical correlation analysis provides a multivariate framework in which to analyze measures of the gap, it is essentially a static concept. To examine how the measures move over time a cointegrating framework is required. If measures of potential output are pairwise cointegrated with unit coefficients, we can conclude that they have the same common trend and the gaps should be capturing the same information. Thus, a test of whether various gap measures are equivalent can be performed by examining whether the cointegrating vector has full rank. If the potential output measures do not mutually cointegrate, they cannot be cointegrated with the same determinants and the gap measures will be measuring different entities.

We shall use a system cointegration test based on Johansen (1995). Defining a general unrestricted VAR with no exogenous variables as:

$$\mathbf{y}_t^* = \sum\nolimits_{j=1}^{k} \boldsymbol{\pi}_j \mathbf{y}_{t-j}^* + \boldsymbol{\Phi}\mathbf{q}_t + \mathbf{v}_t, \quad \mathbf{v}_t \sim \mathrm{IN}\left[\mathbf{0}, \boldsymbol{\Omega}\right], \tag{71}$$

where $\mathbf{y}_t^*$ is a $(p \times 1)$ vector of potential output measures for $t = 1, ..., T$ and $\mathbf{q}_t$ holds the deterministic variables including a constant and trend. In our analysis we use a VAR(4) so

[75] There are other tests that can be used to test the significance of the canonical correlations. These include Wilks' Lambda, Pillai's Trace and Roy's Greatest Root:

$$\begin{aligned} \text{Wilks}'\ \lambda &= 0.019, \quad F_{(21,403)} = 56.844 \\ \text{Pillai's Trace} &= 2.122, \quad F_{(21,426)} = 49.034 \\ \text{Roy's Greatest Root} &= 5.997, \quad F_{(7,142)} = 121.649 \end{aligned}$$

All are significant, supporting the results of the Hotelling-Lawley Trace test. See Muirhead (1982) for a discussion on testing independence been sets of variables.

$k = 4$. We examine five measures of potential output. These include HP, KS, $MA(32)$, $PF(Stat)$ and $PF(Dyn)$, hence $p = 5$.[76] The model includes an unrestricted constant and a restricted trend. Equation (71) can be rewritten as:

$$\Delta \mathbf{y}_t^* = \boldsymbol{\alpha} \left(\boldsymbol{\beta}_0 \beta_1\right)' \begin{pmatrix} \mathbf{y}_{t-1}^* \\ t \end{pmatrix} + \sum\nolimits_{j=1}^{3} \boldsymbol{\Gamma}_j \Delta \mathbf{y}_{t-j}^* + \boldsymbol{\phi}_0 + \mathbf{v}_t. \tag{72}$$

where, under the null, $\boldsymbol{\alpha} \left(\boldsymbol{\beta}_0 \beta_1\right)' = \boldsymbol{\Pi} = \sum \boldsymbol{\pi}_j - \mathbf{I}$. $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are $(p \times r)$ matrices. The test of cointegration is based the rank, r, of $\boldsymbol{\Pi}$:

$$H(r) : \operatorname{rank}(\boldsymbol{\Pi}) \leq r. \tag{73}$$

The rank of $\boldsymbol{\Pi}$ determines how many linear combinations are I(0). For $0 < r < p$ there will be r cointegrating relations, $\boldsymbol{\beta}' \mathbf{y}_t^*$, which are I(0). Testing the null is done sequentially using nested hypotheses:

$$\underbrace{(rank\Pi \leq 0)}_{H(0)} \subset ... \subset \underbrace{(rank\Pi \leq r)}_{H(r)} \subset ... \subset \underbrace{(rank\Pi \leq p)}_{H(p)} \tag{74}$$

Doornik and Hendry (2001) outline cointegration analysis and the estimation procedure for the cointegrating rank.

The results of the multivariate cointegration test are given in table 14, which reports the log-likelihoods (l), the Johansen eigenvalues (μ) and the trace tests $Tr(r)$ for the hypothesis $H(r)$ of r cointegrating relations.[77] Table 15 reports the estimated cointegrating vector $(\widehat{\beta})$ and the feedback coefficients, $\widehat{\alpha}$ (standard errors are given in parentheses for the case where two cointegrating vectors are imposed).

The results imply that there are two cointegrating vectors between the five estimates of potential output. However, the $\widehat{\alpha}$ matrix strongly suggests that some of the variables

[76] Note that the variables excluded include CS because the trend is very similar to that of the HP filter, $MA(16, 16)$ as it is not timely and UC as it is very close to being a deterministic trend.

[77] * and ** denote significance at the 5% and 1% significance levels respectively.

r	0	1	2	3	4	5
l	5098.373	5150.871	5173.572	5182.855	5188.329	5190.544
μ	–	0.51525	0.26883	0.12018	0.072716	0.030101
$H(r)$	$r = 0$	$r \leq 1$	$r \leq 2$	$r \leq 3$	$r \leq 4$	
$Tr(r)$	184.34^{**}	79.346^{**}	33.944	15.379	4.4317	

Table 14: Cointegration analysis of potential output measures.

	$\widehat{\beta}_1$	$\widehat{\beta}_2$	$\widehat{\beta}_3$	$\widehat{\beta}_4$	$\widehat{\beta}_5$
HP	1.000	-0.936	-3.009	2.414	-11.148
KS	-0.012	1.000	2.217	-3.000	10.261
$MA(32)$	-0.345	-1.776	1.000	-1.816	3.374
$PF(Stat)$	-0.301	0.587	-1.519	1.000	-0.906
$PF(Dyn)$	-0.251	0.647	2.448	2.246	1.000
$trend$	-0.0006	0.0028	-0.0062	-0.0044	-0.016
	$\widehat{\alpha}_1$	$\widehat{\alpha}_2$	$\widehat{\alpha}_3$	$\widehat{\alpha}_4$	$\widehat{\alpha}_5$
HP	-0.0007 (0.00007)	0.00007 (0.00003)	-0.00001	0.00001	-0.000006
KS	-0.002 (0.001)	0.002 (0.001)	-0.0002	-0.001	0.0001
$MA(32)$	0.034 (0.008)	0.019 (0.003)	0.002	0.003	-0.0002
$PF(Stat)$	0.217 (0.061)	0.013 (0.027)	0.030	-0.031	-0.008
$PF(Dyn)$	0.252 (0.057)	0.021 (0.025)	-0.029	-0.024	-0.007

Table 15: Unrestricted estimates of the cointegrating vectors and adjustment coefficients.

are I(2). Given the smoothness of some of the measures this is very plausible and a plot of the roots of the companion matrix suggests that there are two roots that lie outside of the unit circle. DF tests indicate that HP and KS are I(2) processes, and $MA(32)$ may also be I(2). The production function estimates of potential output are estimated to be I(1) processes. A plot of the cointegrating vectors given by the $\widehat{\boldsymbol{\beta}}'\mathbf{y}_t^*$ linear combinations is given in figure 10, panels a-e. The cointegrating vectors do not look stationary. The non-normalized coefficients are recorded against the normalized variables in panels f-j. The first fitted and actual components track each other fairly closely, but there is some deviation in the next four. This analysis does suggest a lack of cointegration for I(1) relations.

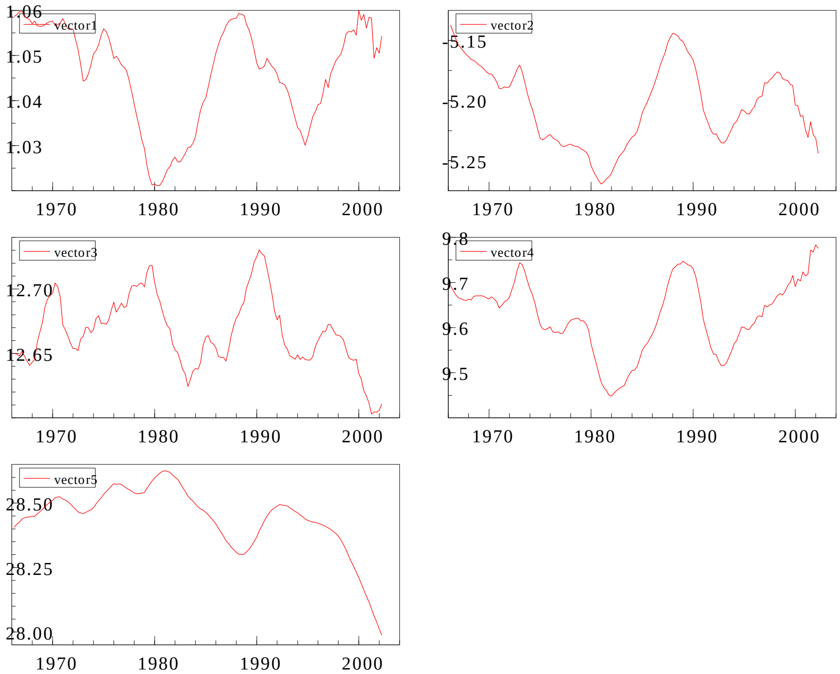

Figure 10: Time series of cointegration vectors.

Tests for I(2) combine the rank test of $\mathbf{\Pi}$ based on equation (72) with an additional reduced rank restriction on the $\mathbf{\Gamma}$ matrix. The second reduced rank condition is given as:

$$\boldsymbol{\alpha}_{\perp}^{\prime}\mathbf{\Gamma}\boldsymbol{\beta}_{\perp}=\boldsymbol{\xi}^{\prime}\boldsymbol{\eta}, \tag{75}$$

where $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ are $(p-r)\times s$ matrices. s is the number of I(1) relations and $p-r-s$ is the number of I(2) relations. Again, testing is done sequentially using nested hypotheses.

$$Q_r : H\left(\text{rank}\left(\mathbf{\Pi}\right)\leq r|\text{rank}\left(\mathbf{\Pi}\right)\leq p\right)$$

$$S_{r,s} : H\left(\text{rank}\left(\mathbf{\Pi}\right)\leq r \text{ and } \leq p-r-s\ I(2)\text{ components}|\text{rank}\left(\mathbf{\Pi}\right)\leq p\right) \tag{76}$$

The I(2) analysis suggests that there are two I(2) relations and one I(1) relation. The test statistic is given as $S_{2,1}=55.438$ with a p-value of 0.42.

The conclusion that we can draw from this analysis is that a lack of full cointegration between different measures of potential output implies the existence of more than one common trend. Given that different measures vary in terms of the order of integration in the series, we would not expect to find that they are related by common trends. Hence, the measures of potential output do not co-move and the resulting output gaps will contain differing properties.

5 Inflationary Pressures

Excess demand is a significant factor driving inflation in the short-run. As the output gap is a proxy for excess demand, a natural extension from the above discussion is to examine the impact of the gap measures on inflation. However, judging the measures of the gap on the basis of inflation alone is problematic as an 'internal' judgement is potentially circular if the gap is defined in terms of inflation. If the gap is just a construct in relation to goods market inflation, the correct measure of the gap would be obtained by 'backing out' an estimate from an inflation model. To then estimate an inflation model based on this type of gap measure would lead to identification problems. If, instead, the gap is a well defined entity determined by the economy's long-run potential growth prospects this problem will not arise. However, the gap is inextricably linked to inflation via the amount of non-inflationary long-term growth that can be sustained. We find that excess demand has a substantial impact upon inflation, strengthening the need for accurate and timely estimates of the gap.

Hendry (2001) argues that there is no single-cause explanation of inflation. Therefore, the model we use is designed to encompass all relevant theories. By adopting a general to specific modelling strategy using PcGets, we can test the relevance of the output gap against all other possible causes of inflation. The use of the single equation framework requires weak exogeneity in the regressors. If this is not the case a VEqCM framework should be used where all variables are modelled explicitly, capturing the variety of channels through which correction to the long-run equilibrium takes place. However, the single equation framework tends to be more robust, particularly for forecasting purposes and so we concentrate on this methodology.

The model of inflation is based on a mark-up model, with excess demand pressures causing short-run cyclical movements in inflation whilst the long-run price level is determined by sectoral price levels including producer prices (ppi), import prices $(import)$, housing rent $(rent)$, wholesale prices (wpi), unit labour costs scaled for the decline in average hours (c^*), oil prices (oil), and national debt (nd). The short-run pressures are captured by the output gap $(xd(pc))$, excess demand for unemployment $(xd(u))$, the growth rate of broad money, $(\Delta m4)$, the short-long real interest rate spread $(rrs - rrl)$, the real effective exchange rate $(reer)$ and asset prices $(assets)$.[78] Some terms are excluded to avoid perfect collinearity and some isomorphic transformations are implemented to limit the parameter space. Also, all t-dated terms in the equilibrium correction model are excluded in an attempt to reduce the possibility of reverse causation bias in the results. If some of the variables were not predetermined a shock may cause a contemporaneous effect on quarterly inflation and other t-dated variables, e.g. an exchange rate shock may impact upon import prices and inflation simultaneously, biasing the results.

[78] Lower case represents logs. See Appendix 1 for a description of the data used.

The GUM, estimated in ECM form, is given as:

$$\begin{aligned}
\Delta p_t &= \beta_0 + \sum_{j=1}^{J-1} \beta_{1,j} \Delta p_{t-j} + \sum_{j=1}^{J-1} \beta_{2,j} \Delta import_{t-j} + \beta_2^* \left(import_{t-J} - p_{t-J}\right) \\
&\quad + \sum_{j=1}^{J-1} \beta_{3,j} \Delta ppi_{t-j} + \beta_3^* \left(ppi_{t-J} - p_{t-J}\right) + \sum_{j=1}^{J-1} \beta_{4,j} \Delta rent_{t-j} \\
&\quad + \beta_4^* \left(rent_{t-J} - p_{t-J}\right) + \sum_{j=1}^{J-1} \beta_{5,j} \Delta wpi_{t-j} + \beta_5^* \left(wpi_{t-J} - p_{t-J}\right) \\
&\quad + \sum_{j=1}^{J-1} \beta_{6,j} \Delta c_{t-j}^* + \beta_6^* \left(c_{t-J}^* - p_{t-J}\right) + \sum_{j=1}^{J-1} \beta_{7,j} \Delta oil_{t-j} + \beta_7^* \left(oil_{t-J} - p_{t-J}\right) \\
&\quad + \sum_{j=1}^{J-1} \beta_{8,j} \Delta nd_{t-j} + \beta_8^* \left(nd_{t-J} - p_{t-J}\right) + f(XD) + \eta D + u_t \\
f(XD) &= \sum_{k=1}^{K} \gamma_{1,k} xd(pc)_{t-k} + \sum_{k=1}^{K} \gamma_{2,k} xd(u)_{t-k} + \sum_{k=1}^{K} \gamma_{3,k} reer_{t-k} \\
&\quad + \sum_{k=1}^{K} \gamma_{4,k} (rrs - rrl)_{t-k} + \sum_{k=1}^{K} \gamma_{5,k} \Delta m4_{t-k} + \sum_{k=1}^{K} \gamma_{8,k} assets_{t-k} \\
u_t &\sim \mathrm{NID}\left(0, \sigma_u^2\right) \qquad (77)
\end{aligned}$$

5.1 The Data

The order of integration of price level data has been discussed extensively in the literature. Hendry (2001) concludes that the price level is I(1) but contains deterministic shifts which give the impression that the series is I(2). DF tests are rarely conclusive due to their low power and results differ across countries and time periods. However, the DF test statistics for the implicit GDP deflator suggest that the price level is I(2) and the inflation rate is I(1).[79] This implies that we have two forms of cointegration. Firstly, the price measures cointegrate to I(1) and secondly, the I(1) cointegrating price measures drive fluctuations in the inflation rate, yielding a polynomially cointegrating relation. This will give a long-run solution for the price level and a long-run solution for the inflation rate based on relative prices. Whilst the model is estimated in I(1) space, I(0) demand side variables drive the short-run fluctuations. Note that many studies examine the consumer price deflator or

[79] ADF test results with constant and trend: $H0 = I(1) : ADF\ \tau = -0.906$, $H0 = I(2) : ADF\ \tau = -2.884$, $H0 = I(3) : ADF\ \tau = -13.82^{**}$.

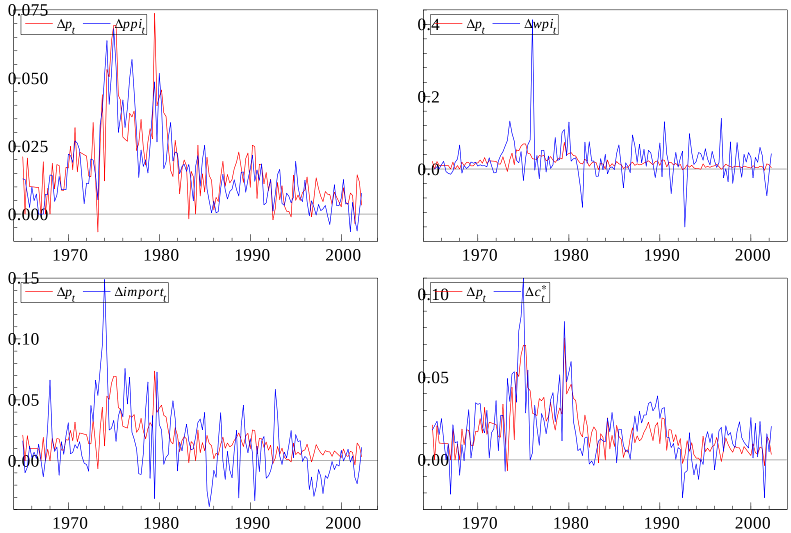

Figure 11: Quarterly growth rates of the producer price index, the wholesale price index, import prices and scaled unit labour costs.

the net national income deflator as opposed to the GDP deflator. Hendry (2001) finds that these series do not mutually cointegrate and so empirical models are specific to the price measure used.

Figure 11 records the quarterly growth rates of *ppi*, *wpi*, *import* and c^* in panels a to d respectively. Δppi follows price inflation fairly closely, but both $\Delta import$ and Δwpi are much more volatile than inflation. Unit labour costs for the whole economy, c, are scaled for the gradual decline in the average number of hours worked per week. If a more disaggregated approach were undertaken, c_t should also control for the effects of self-employment and for the slower evolvement of wage-price linkages in the public sector.[80]

$\Delta rent$ is recorded in figure 12, panel a. Housing market volatility has increased substantially since the late 1980s boom and subsequent recession. The extent of the oil price shocks can be captured by $(oil - p)$ in panel b (scaled for zero mean). Real unit labour

[80] See Batini, Jackson and Nickell (2000).

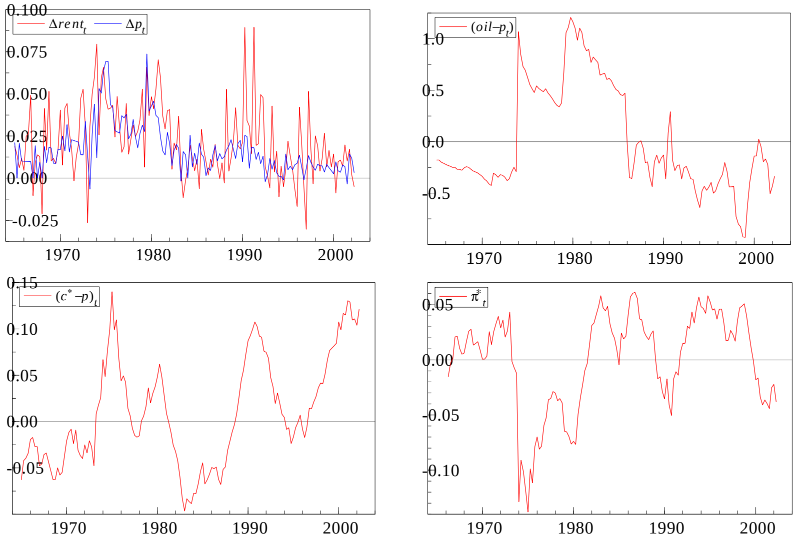

Figure 12: The growth rate of housing rent, oil minus the price level, unit labour costs minus the price level and the markup.

costs $(c^* - p)$ are recorded in panel c and the mark-up, π^*, derived in equation (82) below, is given in panel d.

$(rrs - rrl)$ is included in the GUM as opposed to the interest rates entering independently. The short rate can be thought of as the control variable and the long rate as a proxy for the cost of capital. Hence, the spread captures the inflationary pressures arising from an increase in the cost of capital relative to the borrowing rate. As the interest rates are annual measures they are scaled to represent quarterly interest rates and are adjusted for a sample mean spread of -0.002, recorded in figure **??**, panel a.

Theories of inflation based on purchasing power parity argue that in the long-run exchange rates should adjust to eliminate arbitrage opportunities and hence inflation will be imported via pass-through effects. The $reer$ is derived (setting the sample mean to zero) as:

$$reer_t = p_t - wp_t + 0.02 \tag{78}$$

where wp_t are world prices in sterling. Figure **??**, panel b records $reer$. There are substantial and persistent deviations from PPP over the period, with a range extending from +20% to -30%. Whilst the $reer_t$ is judged to be I(0) over very long data sets, the ADF statistics for the period 1965q1-2002q2 find $reer$ to be I(1).[81]

Monetary theories of inflation stem from Friedman's (1956) seminal work on the 'quantity theory' in which money is treated as exogenous, enabling the money demand equation to be inverted in order to solve for the price level. There is a vast literature looking at money causing inflation, but Hendry (2000a) finds no support for this theory.[82] The growth rate of broad money is included in equation (77), but we do not include an excess demand for money variable. Figure **??** records the velocity of broad money, $v_t = p_t + y_t - m4_t$, in panel c and $\Delta m4$ along with Δp in panel d. The velocity declines sharply over the 1980s when monetarism was operated in the UK via the Medium Term Financial Strategy. The growth rate of broad money tends to exceed price inflation over the 1980s as people transferred their holdings from narrow money to broad money due to the tightening operated.

There is a substantial literature examining the importance of labour market pressures on inflation.[83] We use a measure of excess demand for unemployment based on Hendry (2001). In this model, unemployment rises when the real interest rate exceeds the real growth rate and vice versa. As the unemployment rate, Ur_t, is recorded as in annual units, we derive excess demand for unemployment based on an annual measure of the real interest rate and growth rate and then scale for a quarterly measure.[84] The resulting

[81] Hendry (2001) finds $reer_t$ to be close to its 1872 value in 1991. Also see Rogoff et al. (2001) who examine PPP over 700 years. They find the law of one price holds over the very long-term but that there are substantial and sustained deviations.

[82] See Hendry and Ericsson (1991) and Ericsson et al. (1998) for models of the demand for narrow and broad money respectively in the UK.

[83] See Phillips (1958), Sargan (1980), Nickell (1990) and Layard, Nickell and Jackman (1991) for models of inflation based on the labour market.

[84] A quarterly measure is also derived based upon Ur_t^q and Rl_t^q. Whilst this measure follows the same

model is given as:

$$
\begin{aligned}
\Delta Ur_t &= \underset{(2.58)}{0.001} + \underset{(3.30)}{0.019}\Delta\left(Rl_t - \Delta_4 p_t - \Delta_4 y_t\right) - \underset{(-2.84)}{0.013}Ur_{t-1} + \underset{(23.0)}{0.872}\Delta Ur_{t-1} \\
&\quad \underset{(2.41)}{0.010}\left(Rl_{t-1} - \Delta_4 p_{t-1} - \Delta_4 y_{t-1}\right) - \underset{(-4.16)}{0.006}D71q1 + \underset{(4.93)}{0.007}D71q2 \\
R^2 &= 0.800 \quad \widehat{\sigma} = 0.132\% \quad SC = -6.775 \quad F_{AR}((5,117) = 2.453^* \\
\chi^2_N(2) &= 0.757 \quad F_{ARCH}(4,114) = 1.515 \quad F_{RESET}(1,142) = 0.852 \\
F_H(10,132) &= 1.524 \quad F_{CHOW}(18,125) = 0.466 \quad T = 1965q1 - 2002q2. \qquad (79)
\end{aligned}
$$

The model provides a reasonable fit and passes all diagnostics apart from the AR test at the 5 percent significance.[85] The dummies for the first two periods of 1971 cancel each other out and therefore do not enter into the long-run solution. The long-run solution yields an excess demand for unemployment measure given by:

$$xd(u)_t = Ur_t - 0.05 - 0.55\left(Rl_t - \Delta_4 p_t - \Delta_4 y_t\right). \qquad (80)$$

Figure **??** records excess demand for unemployment in panel b and the annual NAIRU derived in Chapter 4.1 in panel c for comparison. Both measures were tested in the GUM.

As a proxy for excess demand for final goods, the Principal Component of the output gap computed in Chapter 4.3 is used, recorded in panel d (note that this is labelled $xd\,(pc)$). Other measures of the gap are examined in the dominant inflation model derived in Section 5.2.

---

movements as $xd(u)$, the series is much more volatile. This does not accord with the slow moving nature of unemployment and hence the scaled annual measure is preferred.

[85] t-statistics are given in parentheses. R^2 is the squared multiple correlation, $\widehat{\sigma}$ is the residual standard error and SC is the Schwarz Criterion. The diagnostic tests are of the form $F_j\,(k, T-l)$, denoting an F test against an alternative hypothesis given by j. These include k^{th} order serial correlation, F_{AR}, k^{th} order conditional heteroscedasticity, F_{ARCH}, heteroscedasticity, F_H, functional form, F_{RESET}, and parameter constancy over k periods, F_{CHOW}. See Hendry and Doornik (2001) for details of the tests. Normality is tested using the Doornik and Hansen (1994) test and is distributed as a $\chi^2_N\,(2)$.

5.2 The Model

The GUM contains 3 lags of all variables excluding t-dated terms. The model is then reduced to equation (81) by eliminating variables with insignificant t-values. This was conducted in PcGets using a liberal strategy. The liberal strategy minimizes the chances of omitting relevant variables and is therefore less 'tight' than the conservative strategy which minimizes the chances of retaining irrelevant variables.[86] Both strategies are consistent; as $T \to \infty$ the significance level tends to 0.

$$\begin{aligned}
\Delta p_t &= \underset{(4.84)}{0.007} + \underset{(2.98)}{0.185}\Delta p_{t-2} + \underset{(4.58)}{0.092}\left(c^*_{t-1} - p_{t-1}\right) + \underset{(2.34)}{0.082}\Delta m4_{t-7} \\
&\quad + \underset{(6.82)}{0.010}\left(oil_{t-1} - p_{t-1}\right) + \underset{(4.27)}{0.118}\Delta rent_{t-4} + \underset{(3.07)}{0.117}\Delta c^*_{t-3} + \\
&\quad + \underset{(7.94)}{0.318} xd(pc)_{t-1} - \underset{(-5.41)}{0.182} xd(u)_{t-2} - \underset{(-1.91)}{0.169}\left(rrs^q - rrl^q + 0.002\right)_{t-2} \\
&\quad - \underset{(-2.28)}{0.016}\left(reer + 0.02\right)_{t-1} - \underset{(-6.29)}{0.044} D73q2 + \underset{(3.79)}{0.025} D79q3 \\
R^2 &= 0.835 \quad \widehat{\sigma} = 0.625\% \quad SC = -9.795 \quad F_{AR}(5,126) = 0.664 \\
\chi^2_N(2) &= 0.010 \quad F_{ARCH}(4,123) = 1.028 \quad F_{RESET}(1,130) = 1.174 \\
F_H(22,108) &= 1.216 \quad F_{CHOW}(18,113) = 0.824 \quad T = 1966q3 - 2002q2. \qquad (81)
\end{aligned}$$

The model contains elements of most theories of inflation and passes all diagnostics. We can undertake yet another model simplification, following Hendry (2001), by forming a mark-up variable, π^*_t. This is determined by combining c^*, oil and $reer$ in an attempt to capture the mark-up of prices over costs.[87] We make the assumptions of long-run linear

[86] See Hendry and Krolzig (2001) for more details on the strategies of PcGets.

[87] Profit should actually be a function of capital and labour costs, as in the Cobb-Douglas technology used in Chapter 4.1, with weights summing to 1. However, data on capital costs are limited. The long bond rate was tried as a proxy for the cost of capital but the effect is already being captured in the short-long spread. Hence the weight on c^*_t exceeds the Cobb-Douglas weighting of approximately 0.7, because capital costs are not fully captured.

price homogeneity and the adjustment speeds are the same in response to c^*, oil and $reer$.

$$\begin{aligned} \pi_t^* &= 0.016reer_t - 0.092\left(c^* - p\right)_t - 0.010\left(oil - p\right)_t \\ &= p_t - 0.14wp_t - 0.78c_t^* - 0.08oil_t. \end{aligned} \tag{82}$$

Unit labour costs feed through to the GDP deflator with a coefficient of 0.78, which is very similar to Nielsen and Bowdler (2003) who find a coefficient of 0.79 when import prices and unit labour costs enter the long-run solution. Bardsen, Fisher and Nymoen (1998) find a larger coefficient of 0.89 but they exclude import prices and the real exchange rate, which will bias the unit labour cost coefficient upwards. Unit labour costs are dominant in determining the price level and this is consistent with Batini, Jackson and Nickell (2000), who find that the labour share (represented by c) is an important leading indicator of UK inflation. The mark-up is adjusted for a zero mean.[88]

Imposing this restriction yielded $F_{Reduct}\left(2, 131\right) = 4.35^*$ which is marginally significant. However, the restriction does not impact upon the coefficients substantially as they do not change by more than 1 standard error, apart from $xd(u)$ which does not change by more than 2 standard errors, and so the restriction is imposed and the final model is given as:

$$\begin{aligned} \Delta p_t &= \underset{(4.16)}{0.006} + \underset{(3.65)}{0.223}\Delta p_{t-2} + \underset{(4.39)}{0.124}\Delta rent_{t-4} + \underset{(2.94)}{0.111}\Delta c_{t-3}^* + \\ &\quad + \underset{(8.43)}{0.313}xd(pc)_{t-1} - \underset{(-4.91)}{0.128}xd(u)_{t-2} - \underset{(-6.52)}{0.141}\pi_{t-1}^* + \underset{(3.02)}{0.103}\Delta m4_{t-7} \\ &\quad - \underset{(-2.46)}{0.256}\left(rrs^q - rrl^q + 0.002\right)_{t-2} - \underset{(-6.42)}{0.045}D73q2 + \underset{(4.31)}{0.029}D79q3 \\ R^2 &= 0.824 \quad \widehat{\sigma} = 0.641\% \quad SC = -9.800 \quad F_{AR}(5, 128) = 0.566 \\ \chi_N^2(2) &= 0.001 \quad F_{ARCH}(4, 125) = 0.904 \quad F_{RESET}(1, 132) = 3.416 \\ F_H(18, 114) &= 1.257 \quad F_{CHOW}(118, 115) = 0.780 \quad T = 1966q3 - 2002q2. \end{aligned} \tag{83}$$

[88] As the prices are indices there is no natural metric for measuring π_t^*.

The model represents a reasonable fit with a standard error of 0.64%, which is low in view of the turbulence in inflation over the period in question and it passes all diagnostic and constancy tests. The actual and fitted values are recorded in figure 13, along with the scaled residuals, their correlogram and the residual density. Figure **??** records the recursive coefficient estimates and the 1-step residuals with ± 2 standard errors, as well as the 1-step, break-point and forecast Chow tests.[89] The recursive graphics exhibit some evidence of parameter instability, most notably in the spread and $\Delta m4$. As there have been a variety of monetary policy regimes over the period this is not surprising. The 1-step residuals mostly lie within the $\pm$2SE bands, although there does appear to be a slight downward bias over the period. There is some evidence of reduced forecast accuracy post 2000, which can probably be pinpointed as being due to the oil price variable. The large increase in oil prices over 1998 and 1999 have caused an overestimation of quarterly inflation. This is addressed in Chapter 6. There is also a large outlier in 1979 in the 1-step Chow test, again probably due to the oil price shock as this is not reflected in an increase in $\widehat{\sigma}_t$. As the model is relatively stable over time despite many regime changes we can conclude that the implications of the Lucas critique are limited.

The final model contains variables that represent most theories of inflation. The results for quarterly post-war inflation are essentially very close to those obtained by Hendry (2001) for annual inflation over the period 1875-1991, suggesting that the modelling ap-

[89] A 1-step Chow test is given by $\frac{(RSS_t - RSS_{t-1})(t-k-1)}{RSS_{t-1}} \sim_{H0} F(1, t-k-1)$ where the null is given for constant parameters over $t = M, ..., T$. The model is fitted to the sample $M-1$ and the resulting equation is fitted to $M, M+1, ..., T$ observations. Note that normality of Δp_t is needed for this statistic to be distributed as an F distribution.

Break-point Chow tests are sequences of Chow tests as the forecast goes from $N = T - M + 1$ to 1. The statistic is given as $\frac{(RSS_T - RSS_{t-1})(t-k-1)}{RSS_{t-1}(T-t+1)} \sim_{H0} F(T-t+1, t-k-1)$.

The forecast Chow test is a test for constancy over the period 1 to $M-1$ against an alternative which allows for any change over M to T. The test statistic can be given as $\frac{(RSS_t - RSS_{M-1})(M-k-1)}{RSS_{M-1}(t-M+1)} \sim_{H0} F(t-M+1, M-k-1)$ for $t = M, ..., T$.

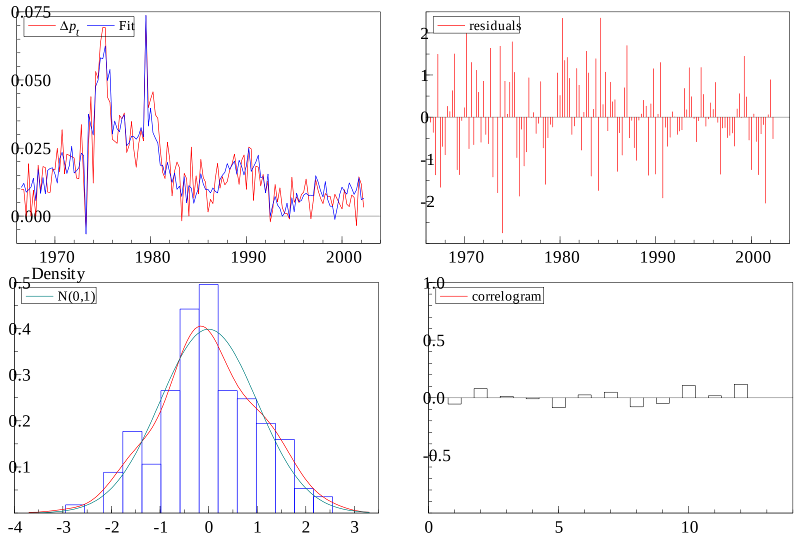

Figure 13: Fitted and actual values of quarterly inflation, the residuals, correlogram and density.

proach used does explain inflation well. There is a small amount of inflation persistence entering through the second lag of quarterly inflation (including the first lag which is positive but insignificant gives an inertia of 25%). The limited evidence for inflation persistence refutes much of the literature, which has suggested that coefficients of the lagged dependent variable are statistically insignificant from 1.[90] Observed inflation persistence in these models may well be due to second round effects in explanatory variables which are not modelled. There is a small but significant constant, suggesting that there is some autonomous inflation of 0.6%.

The short-long spread has a significant impact upon inflation, which is consistent with the long rate being interpreted as a proxy for the cost of capital. $\Delta m4$ enters significantly but with a long lag which is surprising. As a nominal variable it would be expected to feed though to inflation relatively quickly. However, as broad money includes not only assets

[90]See the Fuhrer-Moore (1995) model of inflation stickiness based on relative price rigidities.

used as a medium of exchange but also those used as a temporary store of value, wealth effects may take up to two years to feed through to inflation via potential purchasing power, which would be consistent with a 7 quarter lag.

$\Delta rent$ has a substantial impact of 12%. Rental payment is used as a proxy for the theoretical flow concept of the unit cost of housing. Whilst rents will not control for income effects that arise from the housing market, these will be captured in other demand side variables. The impact occurs with a four quarter lag due to indirect effects. For example, an increase in the cost of housing may reduce labour mobility, increasing wage and price inflation over a longer time horizon. Unit labour costs enter significantly via the mark-up and the growth rate. The mark-up variable is highly significant, with an effect of 14%. Hence, c^* and $reer$ are important determinants of inflation, as well as oil which is highly significant although it has a small coefficient.

Excess demand for unemployment has a significant effect of 13%. Another measure of labour market pressures is the NAIRU. Intuitively, the gap between the level of actual employment and the NAIRU should capture inflationary pressures in the economy. However, replacing $xd(u)$ with $nairu$ led to a significant but smaller impact upon inflation of 6% ($t = -2.24$) dated $t-2$. As with potential output, the NAIRU is a latent variable that is notoriously difficult to measure and hence caution should be applied to these estimates.

The two impulse dummies are highly significant but with relatively small coefficients. As the impulses are not persistent but are instead capturing one-off shocks or outliers, they should not enter the long-run solution as level shifts. The negative residual in 1973q2 is a one-off outlier due to the negative inflation rate recorded in this quarter. The 1979q3 dummy is capturing the increase in VAT after the Thatcher election. This is unlikely to be a step increase as the impact will gradually filter through to inflation. Also, as we are modelling the GDP deflator as opposed to the expenditure deflator, we can assume that the dummy does not impact in the long-run via a level shift.

The output gap has a substantial effect upon inflation of over 30%. Replacing the gap measure by i) a linear trend with break in 1980, ii) a HP filter, iii) a UC measure and iv) excess demand for final goods and services based on Hendry (2001), yields on impact of between 22% and 34%. Capacity utilization only yields an impact of 12%, suggesting that survey based measures of capacity utilization do not accurately reflect the size of the gap. All measures enter significantly with 1 lag, indicating that whilst the timing of the transmission of a shock from the gap onto inflation is captured consistently throughout all measures, the magnitude of the impact is highly dependent upon the measure used.

We can test the impact of various measures of the gap more formally by undertaking model comparisons based on encompassing tests.[91] Encompassing tests based on the F-test assess whether each model from the path search parsimoniously encompasses the union. The test is invariant to the choice of common regressors in models. Defining the first model (M_1) to contain $k_1 + k_2$ regressors, $(\mathbf{x}_{1,t}, \mathbf{x}_{2,t})$, and the second model ($\mathrm{M}_2$) to have $k_2 + k_3$ regressors, $(\mathbf{x}_{2,t}, \mathbf{x}_{3,t})$, where $\mathbf{x}_{2,t}$ are the common regressors, the union model (M_U) comprises the $k = k_1 + k_2 + k_3$ non-redundant set, $(\mathbf{x}_{1,t}, \mathbf{x}_{2,t}, \mathbf{x}_{3,t})$. Let RSS_1, RSS_2, RSS_U and RSS_G denote the residual sums of squares from M_1, M_2, M_U and the GUM respectively. Then, the encompassing test of $\mathrm{M}_1 \varepsilon \mathrm{M}_2$ is equivalent to parsimonious encompassing, $\mathrm{M}_1 \varepsilon_p \mathrm{M}_\mathrm{U}$, where the more simple model is nested within the union model. The PcGets F-test is given as:

$$\eta = \frac{(RSS_1 - RSS_\mathrm{U})/k_3}{RSS_\mathrm{U}/(T-k)} \sim_{H0} F(T-k, k_3).$$

[91] See Hendry and Richard (1982) for a discussion on encompassing.

$M_1 \backslash M_2$	LIN(80)	Δ_4	HP	CS	KS	MA(32)	UC	XD(gs)	PF(S)	PF(D)	PC
LIN(80)	-	0.82	2.66	0.04	1.17	5.06^{*}	6.19^{*}	37.30^{**}	0.65	1.83	0.07
Δ_4	11.77^{**}	-	9.63^{**}	5.62^{*}	1.01	0.05	0.12	56.02^{**}	0.40	3.10	6.46^{*}
HP	4.02^{*}	0.17	-	0.36	8.13^{**}	0.28	1.66	38.82^{**}	0.55	0.06	0.05
CS	5.86^{*}	0.75	4.82^{*}	-	2.45	2.32	4.90^{*}	49.15^{**}	1.28	1.42	1.51
KS	15.24^{**}	3.85	21.38^{**}	10.37^{**}	-	0.19	0.05	59.33^{**}	0.55	2.82	16.25^{**}
MA(32)	23.73^{**}	6.62^{*}	16.79^{**}	14.18^{**}	3.85	-	0.34	78.03^{**}	1.55	7.62^{**}	19.63^{**}
UC	25.39^{**}	7.01^{**}	18.70^{**}	17.33^{**}	4.03^{*}	0.64	-	74.12^{**}	2.09	6.32^{*}	26.23^{**}
xd(good)	0.54	3.98	0.39	3.82	3.47	12.70^{**}	9.68^{**}	-	5.63^{*}	0.20	5.71^{*}
PF(Stat)	17.26^{**}	5.63^{*}	15.66^{**}	11.64^{**}	2.91	0.25	0.49	65.83^{**}	-	6.05^{*}	15.15^{**}
PF(Dyn)	14.14^{**}	4.29^{*}	10.76^{**}	7.54^{**}	1.16	2.18	0.62	52.40^{**}	1.97	-	8.02^{**}
PC	4.36^{*}	0.07	2.97	0.02	6.44^{*}	5.78^{**}	11.44^{**}	49.62^{**}	3.02	0.39	-

Table 16: Encompassing tests for various measures of the output gap in the dominant, congruent inflation model.

Initially, the gap measures were tested one by one, with the results of the encompassing tests reported in table 16. For brevity, only the F-test is reported but the encompassing tests based on Cox (1961), Ericsson (1983) and Sargan (1959) are also checked.[92]

The results suggest that the inflation model using the Principal Component of excess demand does not encompass the models that use $MA(32)$, UC, or $xd(goods)$. It does remain undominated against models that use Δ_4, HP, CS, and the production function estimates. It is difficult to draw conclusions from table 16 regarding the 'best' measure of the output gap to use in an inflation equation because the tests only apply to a comparison between two individual measures. Instead, we can use the general to specific methodology of PcGets to determine whether a particular measure dominates all other measures of the output gap. We do this by inserting all of the gap measures into equation (83) and testing downwards. Obviously a full test would require us to commence the tests with the GUM, equation (77), but we can assume that we have derived the dominant, congruent in-sample

[92] The table reports the F-tests in the format:

$$\begin{bmatrix} M_1 \backslash M_2 & A & B \\ A & - & M_A \varepsilon M_B \\ B & M_B \varepsilon M_A & - \end{bmatrix}.$$

* and ** denote significance at the 5% and 1% levels respectively.

model. Given the use of the composite measure in the initial model, which incorporates all other measures, we would not expect the models to differ significantly depending upon the measure of excess demand used. Two lags of each gap measure were included in order to check the dynamics.

The specific model does find a dominant measure of the gap in explaining inflation. The $xd(goods)$ measure based on Hendry (2001) was selected, with an impact of 0.343% on inflation (s.e.=0.039). This enters significantly with 1 lag. The equation standard error is 0.632%, which is slightly lower than the model containing the Principal Component gap measure. All other coefficients are stable, with none changing by more than 1 standard error. Note that the HP and KS entered with opposing signs (dated $t-2$), exactly cancelling each other out. This analysis exposes the errors in assuming the measures of excess demand are all measuring the same latent variable and are therefore 'substitutes'. The primary concern with the different gap measures is in the varying size of the impact upon inflation. However, in general we can conclude that the impact of the gap is substantial and highly significant.

5.3 A 'Business Cycle' Factor

Using the principal component techniques outlined in Chapter 4.3, we can estimate a composite measure of the business cycle. As inflationary pressures arise via many different channels which can be captured in terms of 'gaps' measuring excess demand or supply in different markets, numerous data series give information regarding the business cycle. If inflation is thought to be driven by a general business cycle factor it may be possible to explain inflation by a linear combination of these gaps which should capture all the business cycle characteristics of the data. If, on the other hand, different gap measures have differing impacts on inflation, information may well be lost by looking at a general business cycle explanation of inflation.

On a note of caution, the maximum variance component need not be a good measure of the business cycle and there is no underlying theory outlining why the leading PCs accurately measure the business cycle. Hence, interpretation of PCs is difficult. One potential product of research into gaps is to construct a dating of cycles similar to that of the National Bureau of Economic Research. The PCs may be informative in dating cycles if they have explanatory power.

The theoretical underpinnings of the principal components analysis lie in the decision as to which variables should be included. Stock and Watson (1998) adopt a very general approach whereby they include 216 variables in the analysis. As the main aim of our analysis is to detect a general structure in the combined variables, a much smaller subset of data is used in order to avoid cluttering with irrelevant variables that may pick up spurious correlations.

Table 17 reports the estimated eigenvalues for the first 7 PCs based on the variables: ppi, wpi, c^*, oil, $rent$, nd, $import$, ur, Rs, Rl, $m4$, $reer$, $assets$, $xd(u)$ and $xd(pc)$. Both levels and first differences were included in order to detect trend and cycle components.[93] The variables included were scaled in order to avoid the series with the greatest amplitude in cycle exerting too much pressure on the PC. The normalized variable is given as $x_i^* = \frac{(x_i - \mu_{x_i})}{\sigma_{x_i}}$. 14 out of the 29 PCs are required to obtain a 95% level of significance. By the Kaiser criterion 7 would be retained and the Scree test suggests that 6 or 7 components should be retained. The factor loadings for the first 7 components are reported in Appendix

[93] The first difference of both housing rent and national debt were excluded from the principal component analysis. As the levels of these variables are so smooth, the differences are very small and this adversely biases the components. Also, p_{t-1} and Δp_{t-1} are excluded to avoid biasing the results.

	Eigenvalues	Cumulative %
$PC1$	8.994	32.12
$PC2$	4.250	47.30
$PC3$	2.877	57.58
$PC4$	2.647	67.03
$PC5$	1.718	73.16
$PC6$	1.374	78.07
$PC7$	1.216	82.42

Table 17: Estimated eigenvalues for the first seven principal components for inflation.

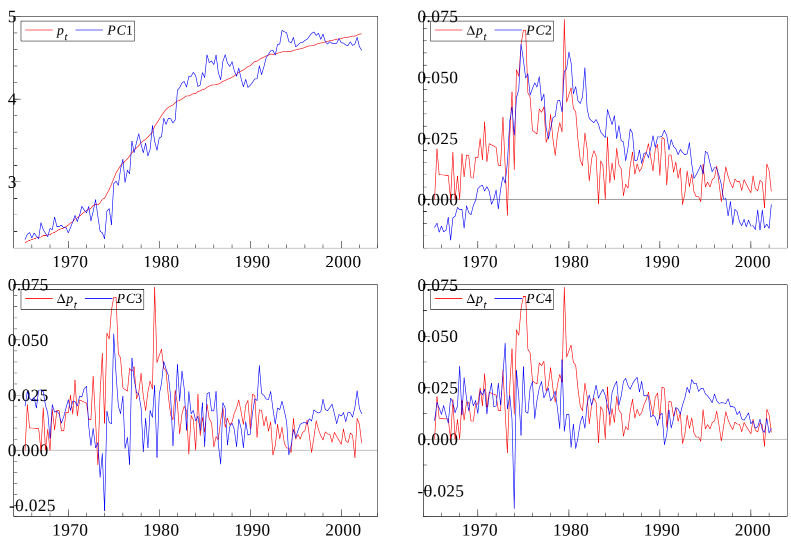

Figure 14: Leading 4 principal components for quarterly inflation.

6.

Figure 14 records the first 4 PCs, scaled by the price level for PC1 and quarterly inflation for PC2, PC3 and PC4. The first component is picking up the trend in the price level, although it is much more volatile. The second component matches inflation reasonably well (correlation = 0.67). The third and fourth components tend to be picking up innovations in the data.

We can estimate a model of inflation based on these PCs. A general to specific modelling strategy was used by including the first 7 PCs with 4 lags of each, along with lags of the dependent variable. The model was the tested down using PcGets to determine the specific model, excluding insignificant variables and imposing the restriction that the first component enters in differences as opposed to levels. The resulting model is given in equation (84).

$$\begin{aligned}
\Delta p_t &= \underset{(8.12)}{0.009} + \underset{(5.59)}{0.322}\Delta p_{t-2} + \underset{(21.0)}{2.489}\Delta PC_{1,t} + \underset{(7.37)}{2.583}\Delta PC_{1,t-1} \\
&\quad + \underset{(11.3)}{1.385} PC_{2,t} - \underset{(-9.14)}{1.236} PC_{2,t-1} + \underset{(4.21)}{0.184} PC_{3,t} - \underset{(-5.05)}{0.361} PC_{4,t} \\
&\quad + \underset{(3.32)}{0.234} PC_{4,t-1} + \underset{(4.91)}{0.343} PC_{4,t-2} - \underset{(-3.77)}{0.278} PC_{5,t} + \underset{(6.62)}{0.432} PC_{5,t-2} \\
&\quad - \underset{(-4.78)}{0.353} PC_{6,t} + \underset{(3.30)}{0.219} PC_{6,t-1} - \underset{(-4.45)}{0.031} D73q2 \\
R^2 &= 0.828 \;\; \widehat{\sigma} = 0.642\% \;\; SC = -9.694 \;\; F_{AR}(5,125) = 1.703 \\
\chi^2_N(2) &= 8.672^* \;\; F_{ARCH}(4,122) = 0.411 \;\; F_{RESET}(1,129) = 1.496 \\
F_H(27,102) &= 1.151 \;\; F_{CHOW}(18,112) = 0.709 \;\; T = 1966q2 - 2002q2. \qquad (84)
\end{aligned}$$

The model provides a good fit with a residual standard error of 0.642%, which is comparable to equation (83). The model does fail normality at the 5% significance level, even when the 1973q2 impulse dummy is included. Further restrictions could not be imposed. The model diagnostics are recorded in figure **??** and the recursive coefficients are stable.[94] Whilst the composite measures do explain inflation well, the inability to interpret the model implies that the model is of limited value to policy-makers. Stock and Watson (1999a) argue that the model's use lies in forecasting.

We can examine the impact of the PCs by adding them into the congruent inflation model, equation (83), to see if they explain inflation by negating the exogenous variables.

[94] Graphs of the recursive coefficients are available upon request. The model passed all constancy tests.

The model is given as:

$$\begin{aligned}
\Delta p_t &= \underset{(12.2)}{0.011} + \underset{(2.76)}{0.071}\Delta rent_{t-4} + \underset{(4.50)}{0.145}\Delta c^*_{t-3} + \underset{(5.59)}{0.236} xd(pc)_{t-1} \\
&\quad -\underset{(-6.17)}{0.159} xd(u)_{t-2} - \underset{(-4.64)}{0.100}\pi^*_{t-1} - \underset{(-4.53)}{0.381}\left(rrs^q - rrl^q + 0.002\right)_{t-2} \\
&\quad -\underset{(-5.63)}{0.034} D73q2 + \underset{(3.12)}{0.019} D79q3 + \underset{(2.33)}{0.462}\Delta PC_{1,t} + \underset{(4.98)}{0.619} PC_{2,t} \\
&\quad -\underset{(-3.27)}{0.358} PC_{2,t-1} + \underset{(5.16)}{0.179} PC_{3,t} + \underset{(2.46)}{0.111} PC_{5,t-1} \\
R^2 &= 0.878 \quad \widehat{\sigma} = 0.540\% \quad SC = -10.063 \quad F_{AR}(5,125) = 0.686 \\
\chi^2_N(2) &= 0.244 \quad F_{ARCH}(4,122) = 0.192 \quad F_{RESET}(1,129) = 1.382 \\
F_H(24,105) &= 0.789 \quad F_{CHOW}(18,113) = 0.843 \quad T = 1966q3 - 2002q2. \qquad (85)
\end{aligned}$$

The model passes all diagnostics and represents an improvement in fit from equation (83), with a residual standard error of 0.54% as opposed to 0.64%. Thus, the PCs are capturing important information, but not to the extent that they can represent inflation alone. The lagged dependent variable is insignificant, suggesting that there is no inflation persistence but that any observed persistence is actually proxying the long-run determinants of inflation captured by the mark-up and the PCs, which contain many input price levels. Also, the growth rate of broad money is negligible once the PCs enter the model. Their presence does impact upon the coefficients, although whilst most do not change by more than 2 standard errors the constant increases by 4 standard errors to 0.011. The impact of excess demand for goods is reduced but the PCs are probably also capturing these pressures. It is very difficult to interpret any of the coefficients because of the lack of interpretability of the components, which are likely to be picking up effects from many causes of inflation. The PCs do not negate the dummies, suggesting that these are modelling effects that are not captured by the economic variables included.

Again, the model should be tested down from the GUM by including all variables and the PCs to derive the congruent inflation model. This is because the relationships between variables may change when the PCs enter the model. To ensure equation (85) is the dominant model, the selection was checked by PcGets and the same model was chosen using 1% significance levels. The evidence does suggest that whilst this data reduction method does capture useful information, it cannot substitute well specified reduced form equations which attempt to model all significant theories of inflation. The single-cause explanation of inflation, in this case represented by what we term general 'business cycle characteristics', is again refuted. The problems of a lack of interpretability and non-robustness to changes in the information set considerably hinder the use of principal component methods.

6 Forecasting Inflation

The move towards an inflation targeting regime in the UK has put pressure on inflation forecasting models to deliver timely, unbiased and efficient forecasts of future inflation. Many current forecasting models fall short of meeting such criteria. The recent inflation forecasts by the Bank of England have contained an upward bias over the last two years, with wide error margins that increase quickly over the forecast horizon.[95] This chapter assesses a variety of forecasting models over the 1998-2002 horizon in an attempt to improve upon current forecasts and to explain why the current forecasts have performed badly.

The role of the output gap is vital in inflation models and its presence in the forecasting models highlights the need for accurate current-dated estimates of the gap. The previous analysis exposes the failure of many gap measures at the end-point and this exacerbates forecast uncertainty. Robust measures of excess demand that are not subject to substantial ex post revisions are essential in models used to forecast inflation.

In a non-stationary and evolving climate, simple naive forecasting devices often outperform causal models. The random walk model has been seen to win forecasting competitions, even when pitted against dominant congruent in-sample models. Clements and Hendry (1999) develop a theory of forecasting that exposes the fundamental source of forecast failure as being location shifts and this provides the reasoning as to why naive devices which track the actual series perform well. In contrast to this theory, Stock and Watson (1999a) develop an entirely disparate approach in the form of factor analysis. They argue that thousands of economic time series contain information regarding future

[95] See Balakrishnan and Lopez-Salido (2002) for a critique of current forecasting models. For an empirical summary, Canova (2001) reports forecasting results for a variety of models designed to forecast 4 quarter ahead inflation for the UK.

inflation. By reducing this entire information set into a small number of estimated factors used to forecast inflation, the forecast estimates should improve.

The forecasting models examined include:

1. Random Walk [RW]
2. Univariate Unobserved Components model [UC]
3. EqCM based on a mark-up model of inflation [Model A]
4. EqCM excluding volatile variables [Model B]
5. Differenced EqCM [Model C]
6. Phillips Curve model [Phillips]
7. Principal Components model [PC]
8. The average forecast [Average]
9. The average forecast excluding the UC model [Ave ex.(UC)][96]

The forecasting theory of the models used is outlined in Section 6.1, followed by forecasts based on models 1-9 along with the corresponding equations in Section 6.2.

6.1 Forecasting Methods

Estimating a model over $t = 1, ..., T$, with a forecast horizon of $t = T + 1, ..., T + H$, the forecast in $T + h$ is given in equation (86), where I_T is the information set at time T, θ_T is the set of estimated model parameters at time T and the forecast is a function h steps ahead, ψ_h. The resulting forecast error in period $T + h$ is given in equation (87).

$$\widehat{y}_{T+h|T} = \psi_h \left(I_T, \widehat{\theta}_T \right) \tag{86}$$

$$e_{T+h|t} = y_{T+h|T} - \widehat{y}_{T+h|T} \tag{87}$$

If $y_t = \mathbf{\Phi}\left(\mathbf{x}_t\right)$, where $\mathbf{x}_T$ is a vector of variables defined in the GUM, we can write the linear model as $y_t = \mathbf{x}_t'\boldsymbol{\beta} + u_t$. The one step estimator is defined by:

$$\widehat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}}{\operatorname{argmin}} \sum\nolimits_{t=1}^{T} \left(y_t - \mathbf{x}_t'\boldsymbol{\beta}\right)\left(y_t - \mathbf{x}_t'\boldsymbol{\beta}\right)', \tag{88}$$

[96] As there is no iterated 1-step forecast for the UC model, the results of the full average are biased in favour of the direct 4-step forecasts. Hence, the average excluding the UC forecast is a more appropriate comparison for the 4-step forecasts.

Equation (89) shows that the forecasts are conditionally unbiased, and the variance of the forecast error is given in equation (90).

$$\begin{aligned} E\left[e_{T+h}|y_T\right] &= E\left[\mathbf{x}'_{T+h}\boldsymbol{\beta} + u_{T+h} - \mathbf{x}'_{T+h}\widehat{\boldsymbol{\beta}}\right] = \mathbf{x}'_{T+h}\left(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}\right) + u_{T+h} = 0, \qquad (89) \\ V\left[e_{T+h}|y_T\right] &= E\left[\left(e_{T+h}|y_T\right)^2\right] = E\left[\left(\mathbf{x}'_{T+h}\left(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}\right)\right)^2 + u^2_{T+h}\right] \\ &= \sigma_u^2 \mathbf{x}'_{T+h}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{x}_{T+h} + \sigma_u^2. \qquad (90) \end{aligned}$$

In order to compare the accuracy of forecasts we shall examine the bias and efficiency (captured by the mean absolute error) of the forecasts derived from each model. Combining these criteria leads to the mean square forecast error (MSFE), which is reported in table 18 as its root (RMSE).

$$MAE = E\left[\left|e_{T+h|T}\right|\right]. \qquad (91)$$

$$MSFE \equiv E\left[e_{T+h|T}e'_{T+h|T}\right] = V\left[e_{T+h|T}\right] + E\left[e_{T+h|T}\right]E\left[e_{T+h|T}\right]. \qquad (92)$$

As forecast accuracy rankings can change as the forecast horizon changes (based on MSFE), multi-step forecasts are also examined. Whilst multi-step forecasts will not be immune from structural breaks, they may capture long memory effects not contained in the 1-step forecasts. In order to forecast more then one step ahead in a single equation framework either an 'iterated' 1-step estimator or a direct h-step estimator can be used. The iterated 1-step forecast is most common, defined as:

$$\widehat{y}_{T+h} = \mathbf{x}'_T\hat{\boldsymbol{\beta}}^h, \qquad (93)$$

$$E\left[\left(y_{T+h} - \widehat{y}_{T+h}\right)|y_T\right] = \left(\boldsymbol{\beta}^h - E\left[\hat{\boldsymbol{\beta}}^h\right]\right)y_T. \qquad (94)$$

Equation (94) gives the average conditional error. It is assumed that the estimators and the latest observations are approximately independent.

The direct h-step estimator is non-recursive in that all information needed to derive a h-step forecast is available at time T. The forecast is obtained by regressing y_T on the regressors lagged h periods. The estimator is given as:

$$\widetilde{\beta}_h = \underset{\boldsymbol{\beta}_h}{\operatorname{argmin}} \sum_{t=h}^{T} \left(y_t - \mathbf{x}'_{t-h}\boldsymbol{\beta}_h \right) \left(y_t - \mathbf{x}'_{t-h}\boldsymbol{\beta}_h \right)'. \tag{95}$$

Hence, in comparison to equations (93) and (94) the forecasts and average conditional errors are given as:

$$\widetilde{y}_{T+h} = \mathbf{x}'_T \widetilde{\boldsymbol{\beta}}_h, \tag{96}$$

$$E\left[\left(y_{T+h} - \widetilde{y}_{T+h} \right) | y_T \right] = \left(\boldsymbol{\beta}^h - E\left[\widetilde{\boldsymbol{\beta}}_h \right] \right) y_T. \tag{97}$$

The relative forecast accuracy of the two multi-step forecasts depends upon the accuracy of the estimators, $\widehat{\beta}^h$ and $\widetilde{\beta}_h$. Chevillon (2002) finds that the iterated 1-step forecasts are preferable when the model is well specified for both stationary and I(1) processes. However, in the case of a mis-specified model for a non-stationary DGP, or if negative residual serial correlation or deterministic shocks are unaccounted for, direct multi-step estimation may lead to more accurate forecasts. The key factor is the size of the drift. As this gets bigger, the benefits of the direct multi-step forecasts outweigh the iterated 1-step procedure.

Hendry and Clements (1999) show that congruent causal models that perform well in-sample often forecast poorly due to their adaptability to structural breaks. The success of the double differenced forecast is understood in this context.[97] As differencing lowers the

[97] For example, examining a mean shift from 0 to 1 at time τ in a simple model:

$$y_t = \mu_1 \left(1 - 1_T^\tau \right) + \mu_2 1_T^\tau + u_t,$$

where 1_τ^τ is an indicator variable defined as $1_{\tau+j}^\tau = 1$ for $t \in [\tau, \tau + j]$ and 0 otherwise. Taking first differences:

$$\begin{aligned} \Delta y_t &= \mu_1 \Delta \left(1 - 1_T^\tau \right) + \mu_2 \Delta 1_T^\tau + \Delta u_t \\ &= \left(\mu_2 - \mu_1 \right) 1_T^\tau + \Delta u_t. \end{aligned}$$

Thus, by differencing, the expected level of y shifts from μ_1 to μ_2 but there is only a non-zero blip of $(\mu_2 - \mu_1)$ at time τ.

We can see that the forecasts are robust to the shifts by looking at the 1-step forecasts, $\widehat{y}_{t+1|t} = \widehat{\Delta y}_{t+1|t} + \widehat{y}_{t|t}$, but $\widehat{y}_{t|t} \equiv y_t$. Therefore, for $t \geq \tau$, $E\left[\widehat{y}_{t+1|t}\right] = \mu_2 = E\left[y_{t+1|t}\right]$ as $E\left[\widehat{\Delta y}_{t+1|t}\right] = 0$. Hence, the forecast is unbiased.

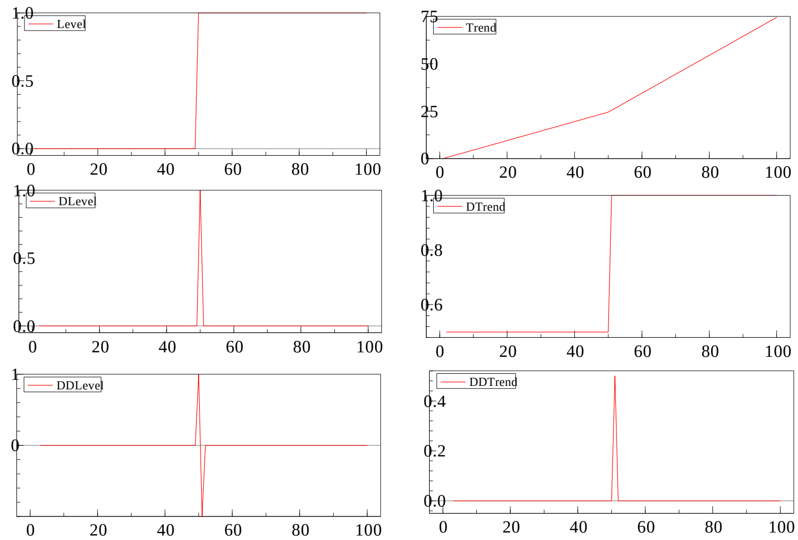

Figure 15: Location shifts and broken trends.

degree of the polynomial in time we can eliminate shifts in trend and location shifts, reducing them to impulses and blips. If we think of 3 degrees of break, from an impulse error to a location shift and then to a break in trend, figure 15 shows the impact of differencing immediately. An impulse error will become a blip when differenced once, a location shift will become an impulse and a trend break will become a level shift. Differencing again will reduce the location shift to a blip and the trend break to an impulse. The process of second differencing effectively removes two unit roots, intercepts and linear trends and so the double differenced model is robust to all of these breaks. After 1 period the forecast will be back on track.

Other methods of robustifying the forecasts include intercept corrections, differencing the equilibrium correction mechanism and the use of composite leading indicators.[98] Intercept corrections are adjustments made at the forecast origin, primarily in an attempt

[98] Emerson and Hendry (1996) find that composite leading indicators do not forecast well in comparison to robust forecasting devices. Also see Camba-Mendez et al. (2001).

to offset structural breaks. If there is deterministic shift prior to the forecast origin, an intercept correction can reduce the bias of the forecasts. However, there is a trade-off in terms of increased forecast error variance and so their application in a stable, unchanged process may lead to a reduction in accuracy as measured by the MSFE. The forecast error variance will depend upon what type of intercept correction strategy is used. There is no evidence of a large structural break in the run up to the forecast horizon 1998q1-2002q2 and hence intercept corrections were not applied.[99]

Another adaptive device that may be used is differencing the EqCM. The reasoning behind the method is that shifts in the mean are the most problematic for forecasting. If there occurs a shift in the equilibrium mean that is unaccounted for, forecasts will be adjusting to the old mean and will therefore be off target for the entire adjustment period. Defining a VAR(1) as $\mathbf{y}_t = \boldsymbol{\tau} + \boldsymbol{\Upsilon}\mathbf{y}_{t-1} + \boldsymbol{\epsilon}_t$, which implies the VEqCM in deviations from means is given as:

$$(\Delta\mathbf{y}_t - \boldsymbol{\gamma}) = \boldsymbol{\alpha}\left(\boldsymbol{\beta}'\mathbf{y}_{t-1} - \boldsymbol{\mu}\right) + \boldsymbol{\epsilon}_t, \tag{98}$$

where the unconditional growth rate of $\mathbf{y}$ is $E\left[\Delta\mathbf{y}_t\right] = \boldsymbol{\gamma}$ and the long-run solution is $E\left[\boldsymbol{\beta}'\mathbf{y}_t\right] = \boldsymbol{\mu}$. Differencing equation (98) leads to:

$$\Delta\mathbf{y}_t = \Delta\mathbf{y}_{t-1} + \boldsymbol{\alpha\beta}'\Delta\mathbf{y}_{t-1} + \Delta\boldsymbol{\epsilon}_t = \left(\mathbf{I} + \boldsymbol{\alpha\beta}'\right)\Delta\mathbf{y}_{t-1} + \boldsymbol{\nu}_t, \tag{99}$$

which is the 1st difference of the initial VAR with the rank restrictions from cointegration imposed. Alternatively, writing equation (99) as:

$$\Delta^2\mathbf{y}_t = \boldsymbol{\alpha\beta}'\Delta\mathbf{y}_{t-1} + \boldsymbol{\nu}_t \tag{100}$$

shows that the double differenced VAR can be augmented by $\boldsymbol{\alpha\beta}'\Delta\mathbf{y}_{t-1}$. As the forecast

[99] We do apply intercept corrections to models A, B, C and the PC model but no substantial gain is obtained in MSFE. However, the increase in uncertainty does not adversely affect the intercept corrected forecasts noticeably either. This is because the corrections imposed are small due to the stability of the period.

differences the mean, a shift in $\boldsymbol{\mu}$ will imply the forecast will fail in the next period but will then correct as $\Delta\boldsymbol{\mu} = 0$ in the following period. Hence, as in the case of the RW, a differenced EqCM will robustify forecasts to deterministic shifts. On a note of caution, unnecessary differencing will lead to increased uncertainty which may increase the MSFE.

In order to apply this procedure to the iterated 1-step forecast, we can define the VAR(1) for the 4-step forecast as:

$$
\begin{aligned}
\mathbf{y}_t &= \sum_{i=0}^{3} \boldsymbol{\Upsilon}^i \boldsymbol{\tau} + \boldsymbol{\Upsilon}^4 \mathbf{y}_{t-1} + \mathbf{u}_t \\
&= \sum_{i=0}^{3} \left(\mathbf{I} + \boldsymbol{\alpha}\boldsymbol{\beta}'\right)^i (\boldsymbol{\gamma} - \boldsymbol{\alpha}\boldsymbol{\mu}) + \left(\mathbf{I} + \boldsymbol{\alpha}\boldsymbol{\beta}'\right)^4 \mathbf{y}_{t-1} + \mathbf{u}_t \\
&= 4\boldsymbol{\gamma} - \left[\mathbf{I} + (\mathbf{I} + \boldsymbol{\alpha}\boldsymbol{\beta}') + (\mathbf{I} + \boldsymbol{\alpha}\boldsymbol{\beta}')^2 + (\mathbf{I} + \boldsymbol{\alpha}\boldsymbol{\beta}')^3\right] \boldsymbol{\alpha}\boldsymbol{\mu} + (\mathbf{I} + \boldsymbol{\alpha}\boldsymbol{\beta}')^4 \mathbf{y}_{t-1} + \mathbf{u}_t \\
&= 4\boldsymbol{\gamma} - \boldsymbol{\alpha}\left[4\mathbf{I} + 6\boldsymbol{\beta}'\boldsymbol{\alpha} + 4(\boldsymbol{\beta}'\boldsymbol{\alpha})^2 + (\boldsymbol{\beta}'\boldsymbol{\alpha})^3\right]\boldsymbol{\mu} \\
&\quad + \boldsymbol{\alpha}\left[4\mathbf{I} + 6(\boldsymbol{\beta}'\boldsymbol{\alpha}) + 4(\boldsymbol{\beta}'\boldsymbol{\alpha})^2 + (\boldsymbol{\beta}'\boldsymbol{\alpha})^3\right]\boldsymbol{\beta}'\mathbf{y}_{t-4} + \mathbf{y}_{t-4} + \mathbf{u}_t \\
&= 4\boldsymbol{\gamma} + \boldsymbol{\alpha}\left[4\mathbf{I} + 6(\boldsymbol{\beta}'\boldsymbol{\alpha}) + 4(\boldsymbol{\beta}'\boldsymbol{\alpha})^2 + (\boldsymbol{\beta}'\boldsymbol{\alpha})^3\right](\boldsymbol{\beta}'\mathbf{y}_{t-4} - \boldsymbol{\mu}) + \mathbf{y}_{t-4} + \mathbf{u}_t, \qquad (101)
\end{aligned}
$$

as:

$$
\begin{aligned}
&\left[\mathbf{I} + (\mathbf{I} + \boldsymbol{\alpha}\boldsymbol{\beta}') + (\mathbf{I} + \boldsymbol{\alpha}\boldsymbol{\beta}')^2 + (\mathbf{I} + \boldsymbol{\alpha}\boldsymbol{\beta}')^3\right]\boldsymbol{\alpha} \\
&= \boldsymbol{\alpha}\left[4\mathbf{I} + 6\boldsymbol{\beta}'\boldsymbol{\alpha} + 4(\boldsymbol{\beta}'\boldsymbol{\alpha})^2 + (\boldsymbol{\beta}'\boldsymbol{\alpha})^3\right]
\end{aligned}
$$

If we define $\boldsymbol{\beta}'\boldsymbol{\alpha} = \boldsymbol{\Lambda}$ as the $r \times r$ matrix of 'roots', which should be negative and relatively small, causing powers to vanish and squares to offset levels, we can derive the multi-step equation for the VEqCM, which is equivalent to equation (98) as:

$$
\begin{aligned}
\Delta_4\mathbf{y}_t &= 4\boldsymbol{\gamma} + \boldsymbol{\alpha}\left(4\mathbf{I} + 6\boldsymbol{\Lambda} + 4\boldsymbol{\Lambda}^2 + \boldsymbol{\Lambda}^3\right)(\boldsymbol{\beta}'\mathbf{y}_{t-4} - \boldsymbol{\mu}) + \mathbf{u}_t \\
&= 4\boldsymbol{\gamma} + \boldsymbol{\alpha}^*(\boldsymbol{\beta}'\mathbf{y}_{t-4} - \boldsymbol{\mu}) + \mathbf{u}_t. \qquad (102)
\end{aligned}
$$

Taking differences of equation (102) gives:

$$
\Delta_4\mathbf{y}_t = (\mathbf{I} + \boldsymbol{\alpha}^*\boldsymbol{\beta}')\,\Delta_4\mathbf{y}_{t-4} + \Delta_4\mathbf{u}_t, \qquad (103)
$$

which is equivalent to equation (99). Hence, the multi-step differenced EqCM forecasts will also be robust to deterministic shifts.

An alternative to theory based forecasting models or causal models are data-driven models. Stock and Watson (1999a) find that Principal Components models can often improve on the forecasting performance of benchmark models such as Phillips curve models or VARs. Bernanke and Boivin (2001) argue that factor models have the advantage of offering a framework for analyzing data that is clearly specified and statistically rigorous, but that remains agnostic about the structure of the economy.

Following Stock and Watson (1998), if y_t is the scalar variable that is being forecasted and $\mathbf{z}_t$ is an N-dimensional matrix of predictors (both written as deviations from means), we can express $(\mathbf{z}_t, y_{t+1})$ in a dynamic factor model representation with r common dynamic factors, f_t.

$$z_{i,t} = \overline{\lambda}_i(L) f_t + \varepsilon_{i,t}, \tag{104}$$

$$y_{t+1} = \overline{\beta}(L) f_t + \epsilon_{t+1}, \tag{105}$$

where the disturbances are given by $\varepsilon_t = (\varepsilon_{1,t}, ..., \varepsilon_{N,t})'$. ϵ_{t+1} is assumed to be a homoskedastic martingale difference sequence with respect to F_t where $F_t = (\mathbf{z}_t, f_t, \epsilon_t, \mathbf{z}_{t-1}, f_{t-1}, \epsilon_{t-1}, ...)$, and so $E(\epsilon_{t+1}|F_t) = 0$ and $E\left(\epsilon_{t+1}^2|F_t\right) = \sigma_\epsilon^2$. $\overline{\lambda}_i(L)$ and $\overline{\beta}(L)$ are finite order lag polynomials. We can rewrite equations (104) and (105) as:

$$\mathbf{z}_t = \Lambda F_t^0 + \varepsilon_t, \tag{106}$$

$$y_{t+1} = \beta' F_t^0 + \epsilon_{t+1}, \tag{107}$$

which is the static form, where $F_t^0 = (f_t, ..., f_{t-q})$ is an $r \times 1$ vector $[r = (q+1)\overline{r}]$, the ith row of $\Lambda = (\overline{\lambda}_{i,0}, ..., \overline{\lambda}_{i,q})$ and $\beta = (\overline{\beta}_0, ..., \overline{\beta}_q)$. The static factor model that we work with arises when F_t^0 and $\varepsilon_{j,t}$ are mutually uncorrelated and i.i.d. and $E(\varepsilon_{i,t}, \varepsilon_{j,t}) = 0, \forall i \neq j$. The factors estimated by Principal Components analysis are consistent as $N \to \infty$ with a fixed T in a static factor model.[100] This is proved by Schneeweiss and Mathes (1995) when $E(\mathbf{z}_t\mathbf{z}_t')$, Λ and $E(\varepsilon_t\varepsilon_t')$ are known.

If I_{t-1} is the full information set in equation (86), the causal model will use a reduced information set, J_{t-1}, where $J_{t-1} \subset I_{t-1}, \forall t$. The PC methodology aims to expand J_{t-1} to be as large a subset of I_{t-1} as possible, with $J_{t-1} = I_{t-1}$ being optimal. However, Hendry (2003) shows that incomplete information about the causal factors is not by itself problematic. The forecast may be less accurate but it will be unbiased. If the process to be predicted is:

$$
\begin{aligned}
y_t &= f_t(I_{t-1}) + \nu_t, \qquad \text{where } E_t[\nu_t|I_{t-1}] = 0, && (108)\\
E_t[y_t|J_{t-1}] &= E_t[f_t(I_{t-1})|J_{t-1}] = g_t(J_{t-1}). && (109)
\end{aligned}
$$

If we define $e_t = y_t - g_t(J_{t-1})$, then $E_t[e_t|J_{t-1}] = 0$, so that e_t is a mean innovation with respect to J_{t-1}. But taking expectations conditional on the full information set implies:

$$
E_t[e_t|I_{t-1}] = f_t(I_{t-1}) - E_t[g_t(J_{t-1})|I_{t-1}] = f_t(I_{t-1}) - g_t(J_{t-1}) \neq 0. \qquad (110)
$$

Thus $V_t[e_t] > V_t[\nu_t]$ so the forecasts are less efficient but are still unbiased.

The Unobserved Components model provides another benchmark forecast. Using the SSF framework outlined in Appendix 2, the forecast of the state vector, $\boldsymbol{\alpha}$, h steps ahead and the MSFE matrix are defined as:

$$
\begin{aligned}
\widehat{\mathbf{a}}_{T+h|T} &= E(\boldsymbol{\alpha}_{T+h}|Y_T), && (111)\\
\sigma^2\widehat{\mathbf{P}}_{T+h|T} &= MSE(\boldsymbol{\alpha}_{T+h}|Y_T). && (112)
\end{aligned}
$$

[100] See Connor and Korajczyk, (1993) for a more detailed discussion.

The forecasts and MSFE matrix are generated recursively by:

$$\widehat{\mathbf{a}}_{T+h|T} = \mathbf{T}\widehat{\mathbf{a}}_{T+h-1|T} + \mathbf{W}^{*}_{T+h}\widetilde{\boldsymbol{\delta}}_{x}, \tag{113}$$

$$\widehat{\sigma}^{2}\widehat{\mathbf{P}}_{T+h|T} = \widehat{\sigma}^{2}\left(\mathbf{T}\widehat{\mathbf{P}}_{T+h-1|T}\mathbf{T}' + \mathbf{H}\mathbf{H}'\right), \tag{114}$$

and hence the resulting forecasts and MSFE matrices are given by:

$$\widehat{y}_{T+h|T} = \mathbf{Z}\widehat{\mathbf{a}}_{T+h|T} + \mathbf{X}^{*}_{T+h}\widetilde{\boldsymbol{\delta}}_{x}, \tag{115}$$

$$\widehat{\sigma}^{2}\widehat{\mathbf{F}}_{T+h|T} = \widehat{\sigma}^{2}\left(\mathbf{Z}P_{T+h|T}\mathbf{Z}' + \mathbf{G}\mathbf{G}'\right). \tag{116}$$

The h-step forecasts are direct forecasts based on $E(y_{T+h}|y_T)$ and the hyperparameters are not re-estimated at each h, i.e. the trend and cycle forecast estimates will be based upon estimation of the hyperparameters at T. Iterated 1-step forecasts cannot be obtained from this framework.

The final forecast statistics reported are for the average forecast. Pooling of forecasts can improve forecasting accuracy immensely. This is because different forecasting models are likely to be affected by different breaks and averaging over them may lead to a more robust forecast. Also, if different forecasts are biased in different directions, the average should provide a more accurate forecast. We find the average does perform well.

6.2 Forecasts

The models were estimated over the period 1966q3-1997q4 and evaluated for forecasting performance over the period 1998q1-2002q2. This analysis looks at the 1-step and 4-step forecasts (both iterated and direct) over 18 quarters. Quarterly inflation in the implicit GDP deflator is examined and the performance is judged on MSFE. The forecast period is relatively stable, with a mean of 0.62% and a maximum range of 1.3%. Table 18 reports

		1-step			Iterated 1-step (h=4)			Direct 4-step	
Model	Bias	MAE	RMSE	Bias	MAE	RMSE	Bias	MAE	RMSE
RW	-0.018	0.432	0.606	-0.049	0.298	0.404	-0.049	0.298	0.404
UC	0.036	0.361	0.477	-	-	-	0.454	0.475	0.490
Model A	-0.441	0.638	0.779	-0.544	0.702	0.847	-0.748	0.763	0.857
Model B	-0.044	0.500	0.614	-0.102	0.446	0.555	0.169	0.480	0.594
Model C	-0.128	0.450	0.566	-0.192	0.435	0.525	-0.532	0.645	0.719
Phillips	-0.017	0.387	0.488	-0.275	0.488	0.556	-0.639	0.706	0.774
PC	-0.016	0.496	0.590	-0.028	0.448	0.543	-0.029	0.394	0.486
Average	-0.104	0.385	0.468	-0.198	0.378	0.459	-0.196	0.346	0.385
Ave (ex.UC)	-0.127	0.400	0.479	-0.198	0.378	0.459	-0.305	0.450	0.492

Table 18: Forecast summary statistics for quarterly inflation, 1998q1-2002q2.

the summary forecast statistics for models 1-9.[101]

A random walk, $\Delta p_t = \Delta p_{t-h} + \varepsilon_t$, can be written as a double differenced model, where $E\left(\Delta^2 p_t\right) = 0$ because the price level does not continuously accelerate. The iterated 1-step and direct 4-step estimators are identical:

$$\Delta p_{T+h} = \widehat{\phi}^h \Delta p_T = \widetilde{\phi}_h \Delta p_T \quad \text{when } \phi = 1. \tag{117}$$

Whilst RW performs well when there are large deterministic shifts, its success in quiet periods is also intuitive. Since the forecast period captures a period of inflation targeting during which the Central Bank has the specific remit of providing low and stable inflation, we have observed relatively constant inflation at approximately 2.5%pa. A RW that tracks inflation by h quarters captures this stability well. Thus, a causal model not only needs to be robust to structural breaks during turbulent periods but it also needs to avoid being contaminated by variables that are volatile during in quiet periods.

Surprisingly, the RMSE of the 4-step forecast is lower than that of the 1-step. This is counter-intuitive as the 4-step forecast error is an accumulation of errors:

$$\Delta p_{T+4} = \Delta p_T \text{ implies that } e_{T+4|T} = \sum_{i=1}^{4} \varepsilon_{T+i}. \tag{118}$$

[101] All statistics are multiplied by 100.

Hence, $MFSE\left(\widehat{\Delta p}_{T+4}\right) \leq MFSE\left(\widehat{\Delta p}_{T+1}\right)$ requires $\sum_{i=1}^{4} \varepsilon_{T+i} \leq \varepsilon_{T+1}$. As inflation fluctuates above and below a constant mean over the sample period, the errors tend to cancel each other out over the 4 quarter horizon, resulting in a lower RMSE than the 1-step. The 4-step forecasts pass through the short-run fluctuations, increasing the efficiency of the forecasts. The model provides a tough benchmark against which to assess other forecasting models.

The second benchmark model is the Unobserved Components model based upon equations (24), (25), (26) and (27). The resulting model can be summarized by the parameters (variance parameters are multiplied by 10^{-5}):

$$\begin{aligned} \sigma_{\eta}^{2} &= 0 \text{ (restricted)},\ \sigma_{\zeta}^{2} = 0.118,\ \sigma_{\kappa}^{2} = 12.02, \\ \rho &= 0.922,\ \frac{2\pi}{\lambda_c} = 18.71,\ \sigma_{\varepsilon}^{2} = 0.00,\ \widehat{\sigma} = 1.36\%, \\ LL &= 521.595,\ Q_{BL}(10,6) = 31.413^{**},\ \chi_{DH}^{2}(2) = 22.299^{**} \end{aligned} \tag{119}$$

Whilst the model is not a good fit based on the diagnostics, the forecasts perform very well. The trend and cycle components are both stable over the forecast period, implying that the forecasts smooth through the quarterly fluctuations giving a similar fit to a moving average with a lag and lead of 2. The forecasts are very close to the average inflation rate over the period.[102] However, the forecasts do not perform as well during volatile periods because of their 'smoothing' properties. Forecasting inflation over 18 quarters following the 1979 oil shock gave rise to systematic overpredictions of inflation for the entire period as the model could not distinguish between a temporary shock and a location shift.

Models A, B and C are the causal model and variants thereof based on equation (77). The models are reported in table 19. In order to forecast over the period 1998q1-2002q2,

[102] The average forecast for the UC model is 0.58%, compared to an actual average over the 18 quarters of 0.62%.

Regressor	Model A	Model B	Model C
$constant$	0.004 (3.73)		0.002 (1.16)
Δp_{t-2}	0.228 (3.22)	0.422 (5.66)	0.383 (4.69)
$xd(pc)_{t-1}$	0.364 (8.40)	0.332 (7.12)	0.256 (4.89)
$xd(u)_{t-2}$	-0.272 (-8.64)	-0.177 (-5.69)	-0.186 (-5.78)
$(c^{*}-p)_{t-1}$	0.061 (3.94)	0.059 (3.22)	$[\Delta_{t-1}]$ 0.078 (1.39)
$(oil-p)_{t-1}$	0.008 (6.33)		$[\Delta_{t-1}]$ 0.004 (0.94)
Δc^{*}_{t-3}	0.157 (3.93)	0.198 (4.39)	0.220 (4.54)
$\Delta rent_{t-4}$	0.129 (4.14)	0.174 (5.14)	0.189 (5.15)
Δppi_{t-2}	0.102 (2.28)	0.087 (1.95)	0.072 (1.871)
$(rrs^{q}-rrl^{q}+\mu)_{t-2}$	-0.124 (-1.99)		-0.157 (-1.97)
$(reer+\mu)_{t-5}$	-0.024 (-2.08)		$[\Delta_{t-5}]$ -0.020 (-0.947)
$\Delta m4_{t-5}$	0.087 (2.83)	0.109 (3.15)	
$D73q2$	-0.040 (-5.82)	-0.034 (-4.94)	-0.053 (-5.95)
$D79q1$	0.027 (4.10)	0.022 (3.37)	0.034 (3.96)
$\widehat{\sigma}$	0.649%	0.766%	0.788%
F_{AR}	(5,108) = 0.812	(5,110) = 1.029	(5,108) = 0.221
F_{ARCH}	(4,105) = 0.725	(4,107) = 4.704^{**}	(4,105) = 2.029
$\chi^2_N(2)$	0.282	2.484	0.228
F_{HET}	(22,90) = 1.209	(18,96) = 1.344	(22,90) = 1.815^{*}
F_{RESET}	(1,112) = 0.500	(1,114) = 0.100	(1,112) = 0.800

Table 19: In-sample forecasting equations.

the dominant model needs to be estimated over the in-sample period (as opposed to the inflation model derived in Chapter 5), or we would bias our results by using information known during the forecast period to derive the best model. Given the parameter stability of our inflation model we would not expect the congruent dominant in-sample model to differ too much. We again use a general to specific methodology.

Whilst Model A provides a good in-sample fit and passes all diagnostics, it has a large RMSE and is easily beaten by the 2 benchmark models over both the 1 and 4-step forecasts. Figure **??** records the model fit in panel a, with the 1-step forecasts, iterated 1-step forecasts and direct 4-step forecasts in panels b-d. The model appears to deliver autocorrelated 1-step and iterated 1-step forecast errors, with a period of underprediction in 1998q3-1999q3 followed by a period of overprediction. The forecast failure in this structural model is likely to be due to breaks in the regressors over the forecast horizon. If the sample is extended to include the forecast horizon, the resulting residuals are smaller than the forecast errors, implying that variables that break need not lead to poor in-sample model specification but do reduce forecasting ability. This emphasizes the fact that the dominant congruent in-sample model will not necessarily produce the best out-of-sample forecast.

The overprediction is suspected to be driven by oil prices. The petroleum spot price more than doubled from 1999q1 to 2000q2, which is picked up by the 1-step and iterated 1-step forecasts. The impact of oil was large in the 1970s, and $(oil - p)$ enters the model highly significantly although with a small coefficient. However, if the economy has become more robust to shifts in the oil price since the 1970s, we would expect the impact upon inflation to be reduced. This can be justified by the reduction in the size of the manufacturing sector, with the service sector not being as susceptible to oil price changes. Hence, the impact of the substantial swings in the oil price in recent periods is likely to be overestimated. The non-recursive direct forecasts do not pick up the oil price shock but the forecasts are systematically higher than inflation by 0.75% on average.

In an attempt to reduce the model's susceptibility to variables that break, Model B excludes oil prices, interest rates (which have the property of breaking when policy-makers adjust the base rate and are therefore prone to location shifts) and exchange rates. The model fit and forecasts are recorded in figure **??** (graphs correspond to panels as in figure **??**). The overprediction of the forecasts from 1999q3-2001q4 is eliminated, vastly improving the RMSE. This does suggest that oil was driving the overprediction in the causal model. Note that the 1-step forecasts still fluctuate substantially. An improvement is obtained by forecasting over the longer 4 period horizon.

Model C attempts to robustify model A by differencing the equilibrium correction terms. Instead of imposing the cointegrating vector as in equation (99), a more general model is estimated by differencing $(c^{*}-p)$, $(oil-p)$ and $(reer+0.02)$, i.e. the estimated coefficients from Model A are not imposed.[103] The resulting model excluded the growth rate of money as this was insignificant. The forecasts are recorded in figure **??**. The results suggest that some improvement can be gained from differencing the EqCM, but as the forecast period is stable (other than for oil) the benefits of the procedure are not substantial. The direct 4-step forecasts do not improve with the differenced EqCM.

Despite the controversy surrounding the Phillips curve, the model has been successful in forecasting inflation over the short-run. However, over the latter part of the 1990s, low and falling inflation has been observed with low and falling unemployment. This has led to numerous papers asking whether the short-run Phillips curve has broken down.[104] The Phillips curve model forecasts are based upon an Expectations Augmented Phillips Curve: $\Delta p_t = \Delta p_{t-1} + \lambda\left(y_t - y_t^{*}\right) + \varepsilon_t$, but is augmented with more complex dynamics:

$$\begin{aligned}
\Delta p_t &= \underset{(2.01)}{0.003} + \underset{(3.80)}{0.331}\Delta p_{t-1} + \underset{(2.78)}{0.248}\Delta p_{t-2} + \underset{(2.74)}{0.238}\Delta p_{t-2} + \underset{(3.97)}{0.237}xd(pc)_{t-5} \\
R^2 &= 0.572 \;\; \widehat{\sigma} = 0.999\% \;\; SC = -9.971 \;\; F_{AR}(5,116) = 0.841 \\
\chi^2_N(2) &= 5.959^{*} \;\; F_{ARCH}(4,113) = 0.826 \;\; F_{RESET}(1,120) = 0.337 \\
F_H(8,112) &= 0.720 \;\; F_{CHOW}(18,121) = 0.234 \;\; T = 1966q3 - 1997q4. \qquad (120)
\end{aligned}$$

Whilst theoretical models assume a contemporaneous relationship between the output

[103] Solving for the static long-run solution of Model A results in $p_t - 0.66c_t^{*} - 0.26wp_t - 0.08oil_t$. The cointegrating vector can then be differenced and estimated imposing the coefficient of $\left(I+\alpha\beta'\right)$ using the constrained simultaneous equations model in PcGive. As the model is not as flexible as Model C in the above analysis, Model C was preferred. Also, as the coefficient on c_t^{*} seems rather low in Model A, the cointegrating vector was not imposed.

[104] Atkeson and Ohanian (2001) argue that, similar to its long-run predecessor, the short-run Phillips curve does not represent a stable empirical relationship that can be exploited for the purpose of constructing reliable inflation forecasts. Fisher, Liu, and Zhou (2002) find that naive inflation forecasts outperform Phillips curve forecasts. Also, Brayton et al. (1999) show that the standard Phillips curve model consistently overpredicted inflation during the late 1990s,

gap and inflation, it takes time for excess demand pressures to feed through, with the most significant effect in the Phillips curve model impacting with a 5 quarter lag. This differs to the full inflation model where excess demand enters with a 1 quarter lag. Also, the impact of gap on inflation is estimated to be smaller at 24% in the Phillips curve model. Note that the unemployment gap was also tried. Whilst the model provides a poor in-sample fit, with a residual standard error of 1%, the 1-step forecasts perform well. However, there is a clear upward bias in the 4-step forecasts, reducing their accuracy considerably. Applying an intercept correction to the iterated 1-step forecast and the direct 4-step forecast gives a RMSE of 0.554% and 0.569% respectively. Hence, the reduced efficiency from applying the correction adversely affects the iterated 1-step forecast but the correction does improve the direct 4-step forecast.

The final model examined is the Principal Components model. The dominant in-sample model is given as:

$$\begin{aligned}
\Delta p_t &= \underset{(4.11)}{0.005} + \underset{(6.48)}{0.436}\Delta p_{t-1} + \underset{(2.55)}{0.160}\Delta p_{t-2} + \underset{(8.35)}{4.442}\Delta PC_{1,t} \\
&\quad + \underset{(9.07)}{1.956}\Delta PC_{2,t} + \underset{(3.36)}{0.147}PC_{3,t} - \underset{(-6.94)}{0.589}PC_{4,t} + \underset{(-7.36)}{0.757}PC_{4,t-1} \\
&\quad - \underset{(-5.64)}{0.525}PC_{5,t} + \underset{(6.83)}{0.617}PC_{5,t-1} - \underset{(-5.59)}{0.422}PC_{6,t} + \underset{(5.09)}{0.387}PC_{6,t-1} \\
R^2 &= 0.809 \quad \widehat{\sigma} = 0.688\% \quad SC = -9.694 \quad F_{AR}(5,109) = 3.061^* \\
\chi^2_N(2) &= 6.118^* \quad F_{ARCH}(4,106) = 2.052 \quad F_{RESET}(1,113) = 3.159 \\
F_H(27,102) &= 1.151 \quad F_{CHOW}(18,114) = 0.720 \quad T = 1966q3 - 1997q4. \qquad (121)
\end{aligned}$$

The model forecasts surprisingly well, beating model A over both the 1 and 4-step horizons. Figure **??** records the model and forecasts. As the factors are not robust to changes in the information set, we cannot conclude that Principal Component models do perform well in forecasting, but rather that this particular specification of variables appears to forecast reasonably accurately for this time horizon. Also, as the period in question is relatively stable we would expect the model to do well and these results concur with Stock and Watson's (1999a) findings. However, in more volatile periods it is difficult to see how the model will be robust to structural breaks. More research is required before we can conclude that the apparent success of the model is not just spurious for the stable forecast period examined.

To conclude, it is very difficult to beat benchmark naive forecasting models, not only during periods when structural breaks are prevalent but also in periods of relative stability. Both the RW and UC models produce accurate forecasts despite their simplicity. The average forecast also performs well, and is the optimal model out of the set examined for the 1-step forecasts. Whilst the inflation model derived in Chapter 6 is a congruent, parsimonious model, the resulting inflation forecasts (for the in-sample model) have a larger MSFE than the RW. The reduced form EqCM can be made more robust by excluding or differencing terms, yielding some improvement. A simple Phillips curve does perform well despite recent criticism and, whilst the Principal Components model produces reasonably accurate forecasts, the methodology requires more research before solid conclusions can be made regarding the model's forecast ability.

This analysis has emphasized that the dominant in-sample model need not be the best forecasting model. What serves as a good forecasting model is one in which the variables are insulated from breaks and where the model can recognize periods of stability. A balance is needed as there is a trade-off in terms of a wider information set increasing the efficiency of estimates for full structural models against the insulation from breaks and hence reduced bias contained in naive devices. For multi-step forecasting in a single equation framework, iterated 1-step forecasts are found to perform better than direct

4-step forecasts in all cases apart from the Principal Components model. This accords with Chevillon (2002) who finds that the iterated 1-step forecasts are more accurate when there is no drift term, which is the case for inflation over the forecast period examined. With regard to the implications for the measurement of excess demand, in order to derive forecasts that can beat naive forecasting devices the output gap needs to be robust at the end-of sample. Forecast uncertainty will be exacerbated by gap estimates that contain end-point bias, reducing a causal model's ability to beat simple devices.

7 Conclusion

This thesis assesses a wide variety of both univariate and multivariate approaches in common use for estimating the output gap. There has been a proliferation of techniques employed for measuring the output gap. The current literature is torn between the agnostic view that estimates of the gap should be based on models where the data 'speaks for itself' and theory driven models in which theoretical priors regarding what excess demand should look like shape the estimates of the output gap. The exposition aims to highlight the differing hypotheses, both statistical and empirical, employed by the various measures and to expose some well known pitfalls of the subsequent gaps.

The use of artificially generated data in Chapter 3 enables explicit evaluation of the performance of univariate output gap measures. Defining the 'best' method of estimating the gap as one that estimates the 'true' cycle accurately in a variety of circumstances, we conclude that none of the univariate methods performed substantially better than any others. Whilst the broad profile of the gap is similar across the range of methods, the magnitude of estimates at a point in time are imprecise, implying that the techniques employed cannot easily distinguish between shocks to the transitory and permanent components of output. This suggests augmenting the techniques by information that will improve signal extraction accuracy.

A production function method of estimating the gap is initially undertaken in a static and cointegrating framework. The residual estimate of TFP accords with our priors regarding this latent variable and the lack of cyclicality suggests that efforts to correct for labour hoarding and capacity utilization are successful. Given the presence of substantial and systematic measurement error in the capital stock, potential output is then modelled as the long-run solution to a dynamic model with a time varying intercept that proxies TFP. The dynamic model attributes more of the fluctuations in output to changes in potential output, resulting in a smaller gap.

Whilst the measures of excess demand do differ, there is some element of consensus amongst the various measures. Hence, a composite measure of excess demand is extracted, based on the reasoning that if all gap estimates measure the true gap with some error, this should extract the signal relative to the errors. The paper provides a comparison of various measures in order to derive some stylized facts regarding the business cycle. The uncertainty in the gap estimates, particularly at the end-point, implies that the gap is best treated as an indicator of the state of the world rather than an exact measure of the precise level of the excess demand. Cointegration analysis exposes a lack of full cointegration, implying that the variety of potential output estimates are driven by more than one common trend.

An empirical model of post-war quarterly inflation is developed, with most extant theories of inflation playing a role in the explanation. The impact of excess demand is found to be substantial, but the magnitude of the impact is not robust to the measure of excess demand used. Moreover, a general 'business cycle' explanation of inflation, based on principal components analysis is refuted.

Chapter 6 evaluates the forecasting performance of a variety of inflation models. Full causal models have difficulty in beating benchmark naive forecasting devices, even for the relatively stable forecast period examined. The average forecast is also found to perform well. Multi-step forecast are assessed, with iterated 1-step forecasts outperforming direct 4-step forecasts due to the lack of drift in the inflation rate over the forecast horizon. The importance of robust estimates of excess demand at the forecast origin is emphasized, exposing the problem of end-of-sample bias associated with many measures. This highlights a key area of future research. The failure of many excess demand measures is paramount at the end-point, which is precisely when accurate, unbiased and timely estimates of excess demand are required, both for forecasting and policy-making. The use of disaggregated data, high frequency data and more timely survey estimates may improve on current measures.

The thesis assesses the importance of excess demand from a variety of angles. Whilst the results of the paper may seem pessimistic at first sight, its importance lies in exposing the difficulty of measuring excess demand and recognizing its fundamental importance in empirical research. Although no solution to the measurement problem is offered, the results hopefully provide a guide as to the most appropriate methods of measuring the gap depending upon their use and paves the way for future research on measures of excess demand.

8 Appendices

Appendix 1: Data Definitions

(All data are seasonally adjusted)

Y_t = Gross Domestic Product at constant 1995 prices, £million. [NS, ABMI]

P_t = Implicit deflator of Gross Domestic Product (expenditure) at market prices, (1995=100). [NS, YBGB]

GVA_t = Gross Value Added at basic prices, constant 1995 prices, £million. [NS, ABMM]

$M4_t$ = Nominal broad money stock, £million. [NS, AUYN]

Rs_t = Three-month treasury bill rate. [DS, UKGBILL3]

Rl_t = Yield on 20-year gilts. [DS, UKGBOND]

$REER_t$ = Real Effective Exchange Rate based on relative Unit Labour Costs, (1995=100). [IFS, REUZF...]

W_t = Total compensation of employees, current price, £million. [NS, DTWM]

Z_t = Total gross operating surplus, current price, £million. (Seasonally adjusted using X-11). [NS, ABNF]

$WPOP_t$ = Population aged 16-59/64, '000s. [NS, YBTF from 1992. Pre-1992, EPG, DEG, EG]

EMP_t = Total number in employment, aged 16+, '000s. [NS, MGRZ from 1992. Pre-1992, EPG, DEG, EG]

$Er_t = EMP_t/WPOP_t$

$U_t = WPOP_t - EMP_t$

$Ur_t = U_t/WPOP_t$

$INACTr_t$ = (Economically Inactive population)/(Population) Both for age 16+, '000s [NS, MGSI/MGSL]

$Pr_t = 1 - INACTr_t$

$L_t = WPOP_t \times Er_t \times Pr_t$

H_t = Average actual weekly hours of work (all workers in main & 2nd job). [NS, YBUV from 1992. Pre-1992, EPG, DEG, EG].

OH_t = (Weekly overtime hours per operative on overtime $\times$ fraction of operatives on overtime)/average hours. [EPG, DEG, EG, LMT]

K_t = Net capital stock for the whole economy excluding dwellings sector, £million. [BoE]

I_t = Total gross fixed capital formation, constant price, £million. [NS, NPQT]

$\Delta INVENT_t$ = Changes in inventories, constant 1995 prices. [NS, CAFU]

$U_{c,t}$ = A capacity utilization index based on the CBI index, % working below capacity. Calculated in appendix 5. [DS, UKCBICAB]

$IMPORT_t$ = Implicit price deflator of imports: (total imports at current prices/total imports at constant prices). [NS, IKBI/IKBL]

PPI_t= Manufacturing output price index, (1995=100) [IFS, 11263...ZF...]

WPI_t = Wholesale price index of materials and fuel purchased by manufacturing industry, (1995=100). [DS, UKPPIMMNF]

C_t = Unit labour cost index for the whole economy, (1995=100). [NS, LNNL]

OIL_t = Petroleum spot price, sterling. [BoE]

ND_t = Public sector net debt, £million. [NS, BKQK]

$ASSET_t$ = FTSE all share index/$(GVA \times P)$. [DS, UKSHRPRCF]

$RENT_t$ = Actual rentals for housing + Imputed rentals for housing, £million. (Seasonally adjusted using X11). [NS, ADFT+ADFU]

Ds = Impulse Dummy equal to unity in period s only

BDs = Blip dummy equal to 1 in period s and -1 in period $s+1$ only.

Sources: [NS] National Statistics database; [IFS] International Financial Statistics Database; [BoE] Bank of England; [DS] Datastream; [EPG] Employment and Productivity Gazette, pre-1971; [DEG] Department of Employment Gazette, 1971-79; [EG] Employment Gazette, 1980-1995; [LMT] Labour Market Trends, 1996-present. Data source codes also in brackets.

Appendix 2: SSF and Model Equivalence of BN and UC

The general multivariate framework for structural time series (STS) models is written in state space form (SSF). Following Koopman et al. (1995), the SSF involves a measurement equation which defines the 'states', $\boldsymbol{\alpha}_t$, (equation (122)) and a transition equation

(equation (123)), which outlines how the states evolve. This is modelled as a VAR(1) process. The assumptions on the error process and the vector of regressors, $\mathbf{b}$, are given in equation (124).

$$\mathbf{y}_t = \mathbf{Z}_t\boldsymbol{\alpha}_t + \mathbf{X}_t\mathbf{b} + \mathbf{G}_t\mathbf{u}_t \tag{122}$$

$$\boldsymbol{\alpha}_{t+1} = \mathbf{T}_t\boldsymbol{\alpha}_t + \mathbf{W}_t\mathbf{b} + \mathbf{H}_t\mathbf{u}_t \tag{123}$$

$$\text{where } \mathbf{u}_t \sim \text{NID}\left(\mathbf{0}, \sigma^2\mathbf{I}\right), \ \mathbf{b} = \mathbf{c} + \mathbf{B}\boldsymbol{\delta}, \ \boldsymbol{\delta} \sim \text{N}\left(\boldsymbol{\mu}, \sigma^2\boldsymbol{\Lambda}\right), \tag{124}$$

for $t = 1, ..., T$. The initial condition is given as $\boldsymbol{\alpha}_1 = \mathbf{W}_0\mathbf{b} + \mathbf{H}_0\mathbf{u}_0$. $\mathbf{Z}_t$ and $\mathbf{T}_t$ are the fixed state system matrices, $\mathbf{X}_t$ and $\mathbf{W}_t$ are the known regression system matrices and the error system matrices are given by $\mathbf{G}_t$ and $\mathbf{H}_t$, which transform the disturbance, $\mathbf{u}_t$, into noise. The unknown values in these matrices are hyperparameters. The role of $\mathbf{b}$ allows for a very general system, enabling many features of time series models to be treated in a uniform manner. Equation (122) enables $\mathbf{b}$ to impact on the observations directly via the $\mathbf{X}_t$ regressors. The states are also affected directly by $\mathbf{b}$ via $\mathbf{W}_t$ in equation (123) and the prior distribution of the initial state is also partly defined by $\mathbf{b}$.

The SSF can be simplified for STS models by making the following assumptions.

1. Set $\mathbf{c} = \mathbf{0}$ such that $\mathbf{b} = \mathbf{B}\boldsymbol{\delta}$. The matrices $\mathbf{B}$ and $\boldsymbol{\delta}$ can be partitioned into regression effects and initial effects.
2. Let $\mathbf{G}_t'\mathbf{H}_t = \mathbf{0}$ such that $\mathbf{G}_t\mathbf{u}_t$ and $\mathbf{H}_t\mathbf{u}_t$ are independently distributed.
3. Assume a time-invariant SSF, i.e. $\mathbf{Z}_t = \mathbf{Z}$, $\mathbf{T}_t = \mathbf{T}$, $\mathbf{G}_t = \mathbf{G}$ and $\mathbf{H}_t = \mathbf{H}$.

Any STS model can be written in SSF but care should be taken when specifying the initial state vector. The Kalman Filter (KF) enables the computation of the 1-step predictions and state vectors, along with the corresponding mean square error, via recursive estimation. The likelihood function is computed using the 1-step ahead prediction error decomposition. The KF that STAMP uses, under the assumption that $\boldsymbol{\delta} = \mathbf{0}$, computes:

$$\mathbf{a}_{t|t-1} = E\left(\boldsymbol{\alpha}_t | Y_{t-1}, \boldsymbol{\delta} = \mathbf{0}\right), \tag{125}$$

$$\sigma^2 \mathbf{P}_{t|t-1} = E\left[\left(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}\right)\left(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}\right)' | Y_{t-1}, \boldsymbol{\delta} = \mathbf{0}\right]. \tag{126}$$

Equation (125) gives the mean of the state and equation (126) estimates the mean square error of the state. The corresponding recursive equations of the KF are given by:

$$\begin{aligned}
\mathbf{v}_t &= \mathbf{y}_t - \mathbf{X}_t\mathbf{b} - \mathbf{Z}\mathbf{a}_{t|t-1}, \\
\mathbf{F}_t &= \mathbf{Z}\mathbf{P}_{t|t-1}\mathbf{Z}' + \mathbf{G}\mathbf{G}', \\
q_t &= q_{t-1} + \mathbf{v}_t'\mathbf{F}_t^{-1}\mathbf{v}_t, \\
\mathbf{K}_t &= \mathbf{T}\mathbf{P}_{t|t-1}\mathbf{Z}'\mathbf{F}_t^{-1}, \\
\mathbf{a}_{t+1|t} &= \mathbf{T}\mathbf{a}_{t|t-1} + \mathbf{W}_t\mathbf{b} + \mathbf{K}_t\mathbf{v}_t, \\
\mathbf{P}_{t+1|t} &= \mathbf{T}\mathbf{P}_{t|t-1}\mathbf{T}' - \mathbf{K}_t\mathbf{F}_t\mathbf{K}_t' + \mathbf{H}\mathbf{H}', \\
\text{where } \mathbf{a}_{1|0} &= \mathbf{W}_0\mathbf{b},\ \mathbf{P}_{1|0} = \mathbf{H}_0\mathbf{H}_0',\ q_0 = 0,\ \text{and } t = 1, ..., T
\end{aligned} \tag{127}$$

$\mathbf{K}_t$ is the Kalman gain, $\mathbf{v}_t$ is the 1-step prediction error and $\sigma^2\mathbf{F}_t$ is the corresponding mean square error. The scaled innovations, $\mathbf{F}_t^{-\frac{1}{2}}\mathbf{v}_t$, are approximately NID with zero mean and variance matrix given by the scale identity matrix in a correctly specified model. The estimate of σ^2 is given by $\widehat{\sigma}^2 = q_T/NT$. $\mathbf{F}_t$ must be a non-singular positive definite matrix, although if this condition does not hold initially the KF can be initialized until the condition is reached. If non-stationary components or fixed regression effects are included in the model, $\boldsymbol{\delta} \neq \mathbf{0}$. In this case an augmented KF is applied. The equations in (127) remain as they are but are augmented by:

$$\begin{aligned}
\mathbf{V}_t &= -\mathbf{X}_t\mathbf{B} - \mathbf{Z}\mathbf{A}_{t|t-1}, \\
\mathbf{A}_{t+1|t} &= \mathbf{T}\mathbf{A}_{t|t-1} + \mathbf{W}_t\mathbf{B} + \mathbf{K}_t\mathbf{V}_t, \\
(\mathbf{s}_t, \mathbf{S}_t) &= (\mathbf{s}_{t-1}, \mathbf{S}_{t-1}) + \mathbf{V}_t'\mathbf{F}_t^{-1}(\mathbf{v}_t, \mathbf{V}_t), \\
\text{where } \mathbf{A}_{1|0} &= \mathbf{W}_0\mathbf{B}. \qquad (128)
\end{aligned}$$

The 1-step ahead predictions and the MSEs are given by:

$$\begin{aligned}
\widehat{\mathbf{a}}_{t|t-1} &= \mathbf{a}_{t|t-1} + \mathbf{A}_{t|t-1}\mathbf{S}_{t-1}^{-1}\mathbf{s}_{t-1}, \\
\sigma^2\widehat{\mathbf{P}}_{t|t-1} &= \sigma^2\left(\mathbf{P}_{t|t-1} + \mathbf{A}_{t|t-1}\mathbf{S}_{t-1}^{-1}\mathbf{A}_{t|t-1}'\right). \qquad (129)
\end{aligned}$$

The estimated σ^2 is given as $\widehat{\sigma}^2 = \frac{1}{NT-d-k}\widehat{q}_T$, where $\widehat{q}_T = q_T - \mathbf{s}_T'\mathbf{S}_T^{-1}\mathbf{s}_T$ and $d+k$ is the number of columns in $\mathbf{B}$. The full sample estimate of $\boldsymbol{\delta}$ is given by $\widehat{\boldsymbol{\delta}} = \mathbf{S}_T^{-1}\mathbf{s}_T$ with a MSE given by $MSE(\widehat{\boldsymbol{\delta}}) = \sigma^2\mathbf{S}_T^{-1}$. The likelihood is obtained via prediction error decomposition. See Koopman et al. (1995) and de Jong (1991) for a more detailed analysis.

Having outlined the general SSF model and the KF, we have the apparatus with which to analyze the differences between the UC and BN decompositions, see Morley et al. (2002). Basing the analysis on the UC model outlined in Chapter 2.1, equation (24) shall be restricted by setting $\varepsilon_t \sim$NID$(0,0)$ so there is no irregular component, equation (26) is restricted by imposing $\zeta_t \sim$NID$(0,0)$ so that the trend reduced to a random walk with drift and the cycle is represented as an ARMA(p,q) process as opposed to the trigonometric specification given in equation (27):

$$\phi_p(L)\,\psi_t = \theta_q(L)\,\epsilon_t \qquad (130)$$

where $\epsilon_t \sim$NID$(0,\sigma_\epsilon^2)$ and $Cov(\eta_t, \epsilon_{t+k}) = \sigma_{\eta\epsilon}$ for $k=0$, 0 otherwise. Harvey and Jaeger

(1993) suggest specifying $p = 2$. Writing the model in SSF implies that the trend and cycle are uncorrelated, $\sigma_{\eta\epsilon} = 0$. The model is estimated using the KF as outlined above. This model has an equivalent univariate ARIMA representation obtained by substituting (25) and (130) into (24) and taking first differences:

$$\begin{aligned} \phi_p(L)(1-L)y_t &= \phi_p(1)\beta + \phi_p(L)\eta_t + \theta_q(L)(1-L)\epsilon_t && (131) \\ &= \beta^* + \theta_{q^*}(L)u_t, && (132) \end{aligned}$$

where $u_t \sim \text{NID}(0, \sigma_u^2)$ and $q^* = \max(p, q+1)$.[105] Equation (132) is obtained using Granger's Lemma. The coefficients of $\theta_{q^*}(L)$ and σ_u^2 are derived by matching the autocovariances of equations (131)and (132). The BN trend can be extracted from the Wold representation of (132).

$$y_t^{*(BN)} = y_{t-1}^{*(BN)} + \varphi(1)u_t = \varphi(1)\sum_{j=1}^{t} u_j \tag{133}$$

where $\varphi(1) = \frac{\theta^*_{q^*}(1)}{\phi_q(1)}$ and $y_0^{*(BN)} = 0$. The variance of the innovation to the BN trend is $\varphi(1)^2\sigma_u^2$. The existence of the BN decomposition guarantees that there will always be at least one UC representation of any ARIMA process. It will not be a unique representation because all the parameters may not be identified. The trend process is always identified but the cycle may not be. Hence, identifying restrictions are required. In general, there will be at least as many non-zero autocovariance relations as parameters if $p \geq q+2$. The conditional expectation of the trend component, given large h and ergodicity of ψ_t, is:

$$E[\mu_t|\Omega_t] = \lim_{h\to\infty} E\left[\mu_t + \psi_{t+h}|\Omega_t\right]. \tag{134}$$

[105] The two representations have the same autocovariance structure, implying the same joint distribution of the data under normality.

Also the expected value of any future innovation in the trend is zero, implying:

$$
\begin{aligned}
E\left[\mu_t|\Omega_t\right] &= \lim_{h\to\infty} E\left[\mu_t + \sum_{j=1}^{h}\eta_{t+j} + \psi_{t+h}|\Omega_t\right] \\
&= \lim_{h\to\infty} E\left[y_{t+h} - \mu h|\Omega_t\right] \qquad (135)
\end{aligned}
$$

which is the BN trend, equation (17). The BN trend is the conditional expectation of the RW component of any UC representation of an I(1) process. The conditional expectations of the trend and cycle can always be computed from the ARIMA reduced form. Two assumptions are required to identify the components:

1. The trend is a RW.
2. The cycles are ergodic.

Hence, if the parameters of the reduced form ARIMA are those implied by the UC model, the KF estimates of the trend and cycle will be identical to the BN estimates. The differences observed in practice between the two models are due to the restrictions imposed on the models. The UC model imposes the restriction that the innovations in the trend and cycle components are uncorrelated, $\sigma_{\eta\epsilon} = 0$, whereas the BN is unrestricted. Relaxing this assumption leads to identical decompositions from both methods.

Appendix 3: Results for the UC Model

The unobserved components model is set out in equations (24), (25), (26) and (27), which comprise the measurement and transition equations. Various restricted versions of the general model can be obtained by placing restrictions on the variance parameters. Parameter estimates are reported in table 20, along with the maximized log likelihood (LL), the Box-Ljung statistic $Q_{BL}(p,q)$, and the Doornik Hansen (1994) normality test, $\chi^2_{DH}(2)$. Variance parameters are multiplied by 10^5. All models fail the normality test,

	Model 1	Model 2	Model 3	Model 4
	Unrestricted	$\sigma_\zeta^2 = 0$	$\sigma_\eta^2 = 0$	$\sigma_\eta^2 = \sigma_\varepsilon^2 = 0$
σ_η^2	7.61	7.61	0(restricted)	0(restricted)
σ_ζ^2	0.00	0(restricted)	0.00	0.00
σ_κ^2	0.00	0.00	6.87	8.50
ρ	1	1	0.96	0.95
$\frac{2\pi}{\lambda_c}$	21.11	21.11	53.06	60.53
σ_ε^2	0.69	0.69	0.76	0(restricted)
LL	685.274	685.274	684.295	683.674
$Q_{BL}(p,q)$	13.508	13.508	13.801	16.069
$\chi^2_{DH}\,(2)$	39.809**	39.809**	41.344**	41.650**

Table 20: Unobserved Components models.

probably due to outliers in the 1970s, but pass the autocorrelation test at the 1% significance level. As the models are simple univariate decompositions we would not expect them to be well specified. This analysis examines smoothed estimates of the output gap but filtered estimates can also be obtained.

Model 1 estimates the unrestricted local linear trend, model 2 estimates a local level and fixed slope, model 3 produces a smooth trend and model 4 results in a deterministic trend. A HP filter can be also be estimated within this framework by imposing the restrictions:

$$\begin{aligned} \sigma_\varepsilon^2 &= 0, \;\; \rho = 0, \;\;\; \sigma_\eta^2 = 0, \\ \psi_t &= \kappa_t \sim NID\left(0, \sigma_\kappa^2\right), \;\; \sigma_\zeta^2 = \sigma_\kappa^2/1600. \end{aligned} \tag{136}$$

Hence, σ_κ^2 is that only variance parameter to be estimated. The different UC models highlight the differences in estimates depending on a priori assumptions regarding the smoothness of the trend. Figure 16 records the gap measures from models 1 and 2 in panel a, model 3 in panel b, model 4 in panel c and the HP gap in panel d.

The unrestricted model estimates a short cycle of approximately 5 years, with a dampening factor of 1. The estimate of the slope is 0. Hence, model 2 results in an identical

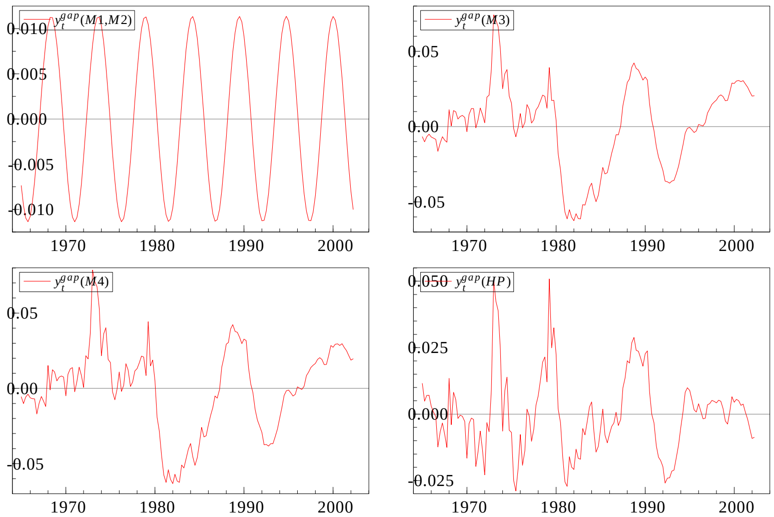

Figure 16: Univariate UC estimates of the output gap.

decomposition. This restriction suggests that Δy_t is stationary, an assumption supported by DF tests. Both models estimate the variance of the cycle to be 0, leading to a deterministic cycle but the sine-cosine waves are not representative of UK business cycles. By imposing a smoothness prior as in model 3, the trend is an integrated random walk. As the slope is estimated to be 0, all of the variation is absorbed into the cycle and irregular. The smooth trend estimates a cycle frequency of 53 quarters, or approximately 13 years which is substantially larger than the stylized facts regarding classical business cycles. Further restricting the model by imposing $\sigma_\varepsilon^2 = 0$ implies that irregular variations are incorporated into the cycle. This lengthens the frequency but does not alter the gap substantially. Finally, by further restricting the model to derive a Hodrick-Prescott trend and cycle, the restrictions on the cycle dampening factor lead to a slightly more volatile cycle, with noticeably different end-point behaviour. The analysis in Chapter 4 uses a gap derived from a smooth trend model as this gap measure accords with our priors regarding the cyclical behaviour of the economy over this period.

Appendix 4: Calculating the Elasticity of Output with respect to Labour

A general form for the production function, which appears to be a good approximation to actual production functions, is the Cobb-Douglas specification. If we assume perfect competition, where the marginal products of labour and capital are equal to the wage rate and profit rate respectively, the elasticity of output with respect to labour, α, is equal to the share of output going to labour. This can be calculated as the share of wages in total income:

$$\alpha = \frac{W}{W+Z}, \tag{137}$$

where W =compensation of employees and Z =gross operating surplus. This should also be augmented by the compensation of those who are self-employed but data shortages restrict this. As the data is quite volatile, a HP filter ($\lambda = 1600$) is used to smooth the data. The wage share is not constant, as can be seen in figure 17. There is a gradual decline from 1975-1985, when the ratio falls from an average of 0.72 to an average of 0.68. However, the series has a small standard deviation of 0.025 and the variation is not substantial. As the share has a mean of 0.702, we can approximate the wage share as being equal to 0.7 over the entire period. This is in line with the Bank of England's Macroeconomic Model. Note that any deviations from this will be picked up in the residual TFP. The smoothed but stochastic series, $\alpha^{(HP)}$, was also tried but the results did not vary substantially, indicating that the constant approximation does not have a large impact upon the results.

Appendix 5: Calculation of a Capacity Utilization Index

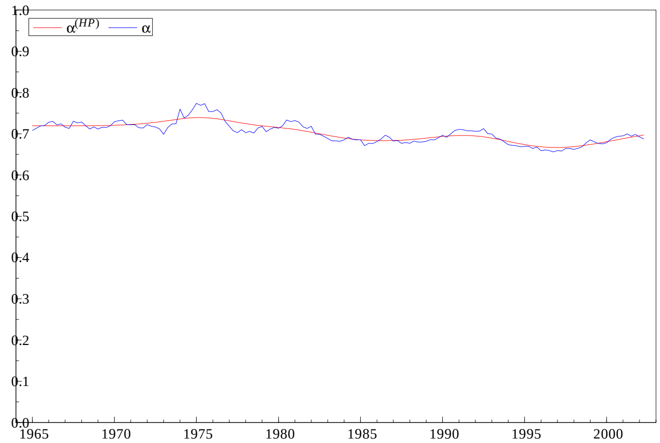

Figure 17: The wage share in national income.

The CBI industrial trends survey reports the response of firms in the manufacturing sector to the question: "Is your present level of output below capacity?" (defined as a satisfactory or full rate of operation). Following Muellbauer (1984), we can define the proportionate deviation of capacity utilization, Uc, by:

$$-Uc = \ln Y(\max) - \ln Y. \tag{138}$$

If different firms have the same view regarding satisfactory levels of operation, we can define z as:

$$Z = \ln Y(\max) - \ln Y(\text{sat}). \tag{139}$$

Assuming a distribution of utilization across firms measured by $\ln Y(\max) - \ln Y$, which shifts through time with a limit fixed at zero, we can calculate the proportion of firms operating below the usual level of capacity, π, which is observed, and link this with the unobserved mean of the distribution, $E(-Uc)$.

If the distribution of capacity utilization is lognormal:

$$\ln(-Uc) \sim \mathrm{N}\left(\mu, \sigma^2\right),$$

then:

$$\pi = 1 - \Phi\left(\frac{\ln Z - \mu}{\sigma}\right).$$

Therefore:

$$E\left[\ln(-Uc)\right] = \mu = \ln Z - \sigma\Phi^{-1}\left(1-\pi\right), \tag{140}$$

where $\Phi\left(\right)$ is the distribution of the standard normal distribution. Due to the normality assumption we can derive $E(-Uc)$ from $E\left[\ln(-Uc)\right]$:

$$\begin{aligned} E(-Uc) &= \exp\frac{1}{2}\sigma^2 \exp\mu \\ &= \beta\exp\left[-\sigma\Phi^{-1}\left(1-\pi\right)\right], \qquad (141) \\ \text{where } \beta &= Z\exp\frac{1}{2}\sigma^2. \end{aligned}$$

In order to calculate this, Amemiya (1981) suggests that if x is a standard normal, the distribution can be well approximated by a logistic distribution:

$$\Phi\left(x\right) \approx \frac{\exp\left(1.6x\right)}{1+\exp\left(1.6x\right)} \sim \text{Log}\left(0, \frac{1.6^2}{\left(\pi^2/3\right)}\right). \tag{142}$$

Using this approximation, we can derive $E(-Uc)$ as:

$$E(-Uc) = \beta\left(\frac{\pi}{1-\pi}\right)^{\frac{\sigma}{1.6}}. \tag{143}$$

Muellbauer (1984) recommends empirical magnitudes of $\sigma = 0.64$ and $z = 0.09$, which suggests that full capacity is approximately 91% of the physical maximum. These values where found by estimating a production function with $\left(\frac{\sigma}{1.6}\right)$ ranging from 0.2 to 0.6.

Appendix 6: Factor Loadings for the PCs of Inflation

	$PC1$	$PC2$	$PC3$	$PC4$	$PC5$	$PC6$	$PC7$
ppi	-0.333	-0.024	0.029	0.064	0.005	0.045	0.006
wpi	-0.326	-0.005	0.028	0.102	-0.009	0.101	-0.043
c^*	-0.330	-0.044	0.024	0.078	0.014	0.058	-0.011
$import$	-0.325	-0.116	0.037	0.044	0.005	0.028	-0.019
nd	-0.327	0.043	0.040	0.067	-0.018	0.071	-0.014
oil	-0.289	-0.212	-0.004	0.081	0.063	0.031	-0.065
$rent$	-0.329	-0.006	0.032	0.082	0.011	0.070	-0.034
Rl	0.090	-0.449	-0.011	-0.037	0.071	-0.021	-0.065
Rs	0.023	-0.384	0.036	0.085	0.077	0.075	0.090
$m4$	-0.327	0.012	0.045	0.082	-0.007	0.069	-0.017
$assets$	-0.004	0.437	0.089	-0.006	-0.123	-0.047	0.086
$reer$	-0.013	0.226	-0.251	0.233	0.164	0.143	0.143
Ur	-0.241	-0.141	0.022	-0.149	0.119	-0.296	0.124
$xd(pc)$	-0.033	-0.056	0.290	-0.399	-0.280	-0.164	0.140
$xd(u)$	-0.201	-0.197	-0.024	-0.300	-0.045	-0.210	0.164
Δppi	0.079	-0.270	-0.179	-0.216	0.060	0.286	0.052
Δwpi	0.019	-0.133	0.0432	-0.043	-0.444	0.033	-0.457
Δc^*	0.132	-0.280	-0.084	0.152	-0.220	0.183	0.157
$\Delta import$	0.132	-0.189	0.314	0.028	0.191	-0.136	-0.269
Δoil	0.005	-0.005	0.024	0.028	-0.008	-0.023	-0.046
ΔRl	0.087	-0.033	0.406	0.163	0.139	0.267	0.159
ΔRs	0.042	-0.029	0.380	-0.016	0.070	0.209	0.298
$\Delta m4$	0.085	-0.235	-0.060	0.212	-0.259	-0.193	0.395
$\Delta assets$	-0.056	0.048	-0.250	-0.137	0.053	-0.499	0.070
$\Delta reer$	-0.047	-0.024	-0.301	0.044	-0.473	0.219	-0.105
ΔUr	0.069	-0.146	-0.291	0.188	0.437	-0.116	-0.157
$\Delta xd(pc)$	-0.029	0.077	0.130	-0.468	0.200	0.197	-0.202
$\Delta xd(u)$	-0.004	-0.038	-0.273	-0.354	0.127	0.305	-0.134

Table 21: Factor loadings for the first seven principal components for inflation.

9 References

Adams, C. and Coe, D. (1990). 'A Systems Approach to Estimating the Natural Rate of Unemployment and Potential Output for the United States.' International Monetary Fund. Staff Papers, 37(2). pp. 232-293.

Amemiya, T. (1981). 'Qualitative Response Models: A Survey.' Journal of Economic Literature, 19. pp. 1483-1536.

Anderson, T.W. (1984). An Introduction to Multivariate Statistical Analysis. Wiley, New York.

Atkeson, A. and Ohanian, L.E. (2001). 'Are Phillips Curves Useful for Forecasting Inflation?' Federal Reserve Bank of Minneapolis Quarterly Review, (Winter). pp. 2-11.

Attfield, C.L.F. and Silverstone, B. (1998). 'Okun's Law, Cointegration and Gap Variables.' Journal of Macroeconomics, 20(3). pp. 625-637.

Bailey, M.N. (1981). 'Productivity and the Services of Capital and Labour.' Brookings Papers on Economic Activity, Issue 1. pp. 1-65.

Balakrishnan, R. and Lopez-Salido, J.D. (2002). 'Understanding UK Inflation: the Role of Openness.' Bank of England Working Paper, No. 164.

Banerjee, A., Dolado, J.J., Galbraith, J.W. and Hendry, D.F. (1993). Cointegration, Error Correction and the Econometric Analysis of Non-Stationary Data. Oxford University Press, Oxford.

Banerjee, A., Hendry, D.F. and Mizon, G.E. (1996). 'The Econometric Analysis of Economic Policy.' Oxford Bulletin of Economics and Statistics, 58(4). pp. 573-600.

Bank of England. (2000). 'Economic Models at the Bank of England - September 2000 Update.'

Bardsen G., Fisher, P.G. and Nymoen, R. (1998). 'Business Cycles: Real Facts or Fallacies?' In Strøm, S. (eds.) Econometrics and Economic Theory in the 20th Century. Econometric Society Monograph 32. Cambridge University Press, Cambridge.

Batini, N., Jackson, B. and Nickell, S. (2000). 'Inflation Dynamics and the Labour Share in the UK.' Bank of England External MPC Unit, Discussion Paper No. 2.

Baxter, M. and King, R.G. (1999). 'Measuring Business Cycles. Approximate Band-Pass Filters for Economic Time Series.' The Review of Economics and Statistics, 81(4). pp. 575-593.

Bell, W. (1984). 'Signal Extraction for Nonstationary Time Series.' The Annals of Statistics, 12. pp. 646-664.

Bernanke, B. S. and Boivin, J. (2001). 'Monetary Policy in a Data-Rich Environment.' National Bureau of Economic Research, Working Paper No. 8379.

Beveridge, S. and Nelson, C.R. (1981). 'A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the Business Cycle.' Journal of Monetary Economics, 7. pp. 151-174.

Blanchard, O.J. and Quah, D. (1989). 'The Dynamic Effect of Aggregate Demand and Supply Disturbances.' American Economic Review, 79(4). pp. 655-673.

Bolt, W. and van Els, P.J.A. (2000). 'Output Gap and Inflation in the EU.' De Nederlandsche Bank. Staff Report No. 44.

Bowdler, C. (2003). 'Essays on Inflation Adjustment in the Open Economy.' D.Phil Thesis, University of Oxford.

Brain, C.W. and Shapiro, S. (1983). 'A Regression Test for Exponentiality: Censored and Incomplete Samples.' Technometrics, 25. pp. 69-76.

Brayton, F., Roberts, J.M. and Williams, J.C. (1999). 'What's Happened to the Phillips Curve?' Federal Reserve Board, Washington. Finance and Economics Discussion Paper No. 49.

Burns, A.M. and Mitchell, W.C. (1946). Measuring Business Cycles. National Bureau of Economic Research, New York.

Burnside, C., Eichenbaum, M. and Rebelo, S. (1993). 'Labor Hoarding and the Business Cycle.' Journal of Political Economy, 101. pp. 245-273.

Camba-Mendez, G., Smith, R. J., Kapetanios, G. and Weale, M. R. (2001). 'An Automatic Leading Indicator of Economic Activity: Forecasting GDP Growth for European Countries.' Econometrics Journal, 4(1). pp. 56-90.

Campbell, J. and Mankiw, N.G. (1987). 'Are Output Fluctuations Transitory?' Quarterly Journal of Economics, 102. pp. 857-880.

Canova, F. (1998). 'Detrending and Business Cycle Facts.' Journal of Monetary Economics, 41(3). pp. 475-512.

Canova, F. (2001). 'G-7 Inflation forecasts.' Available at http://www.york.ac.uk/depts/econ/dept_seminars/canova.pdf

Caporale, G.M. (1997). 'Sectoral Shocks and Business Cycles: A Disaggregated Analysis of Output Fluctuations in the UK.' Applied Economics, 29. pp. 1477-1482.

Cattell, R.B. (1966). 'The Scree Test for the Number of Factors.' Multivariate Behavioural Research, 1. pp. 245–276.

Cecchetti, S.G. (2002). 'The New Economy and the Challenges for Macroeconomic Policy.' National Bureau of Economic Research. Working Paper No. 8935.

Cerra, V. and Saxena, S.C. (2000). 'Alternative Methods of Estimating Potential Output and the Output Gap: An Application to Sweden.' International Monetary Fund. Policy Working Paper. WP/00/59.

Chadha, B., Masson, P. and Meredith, G. (1992). 'Models of Inflation and the Costs of Disinflation.' International Monetary Fund. Staff Papers, 39(2). pp. 395-431.

Chevillon, G. (2002). 'Direct Multi-step Estimation for Forecasting Economic Processes.' Unpublished.

Clark, P., Laxton, D. and Rose, D. (1996). 'Asymmetry in the U.S. Output-Inflation Nexus.' International Monetary Fund. Staff Papers, 43(1). pp. 216-251.

Clements, M.P. and Hendry, D.F. (1999). Forecasting Non-Stationary Economic Time Series. MIT Press, Cambridge, Mass.

Clements, M.P. and Hendry, D.F. (2001). 'Forecasting with Difference-Stationary and Trend-Stationary Models.' The Econometrics Journal, 4(1). pp. 1-19.

Clements, M.P. and Sensier, M. (1999). 'Asymmetric Output-Gap Effects in Aggregate Mark-Up Pricing Models: Evidence for the US, UK and Japan.' Unpublished.

Cochrane, J.H. (1991). 'A Critique of the Application of Unit Root Tests.' Journal of Economic Dynamics and Control, 15(2). pp. 275-284.

Cochrane, J.H. (1994). 'Permanent and Transitory Components of GNP and Stock Prices.' Quarterly Journal of Economics, 61. pp. 241-265.

Coen, R.M. and Hickman, B.G. (2002). 'The Productivity Surge of the Nineties and Future Growth.' Stanford Institute for Economic Policy Research. Discussion Paper No. 01-26.

Cogley, T. and Nason, J.M. (1995). ‘Effects of the Hodrick-Prescott Filter on Trend and Difference Stationary Time Series: Implications for Business Cycle Research.’ Journal of Economic Dynamics and Control, 19(1-2). pp. 253-78.

Connor, G. and Korajczyk, R.A. (1993). ‘A Test for the Number of Factors in an Approximate Factor Model.’ Journal of Finance, 48(4). pp. 1263-1291.

Cox, D.R. (1961). ‘Tests of Separate Families of Hypotheses.’ In Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Vol.1. University of California Press, Berkeley.

De-Brouwer, G. (1998). ‘Estimating Output Gaps.’ Reserve Bank of Australia Research Discussion Paper, No. 9809.

De-Brouwer, G. and Ericsson, N.R. (1998). ‘Modelling Inflation in Australia.’ Journal of Business and Economic Statistics, 16(4). pp. 433-449.

de Jong, P. (1991). ‘The Diffuse Kalman Filter.’ Annals of Statistics, 19. pp. 1073-1083.

De-Masi, P.R. (1997). ‘IMF Estimates of Potential Output: Theory and Practice.’ International Monetary Fund. Working Paper, No. WP/97/177.

Dickey, D.A. and Fuller, W.A. (1979). ‘Distribution of the Estimators for Autoregressive Time Series with a Unit Root.’ Journal of the American Statistical Association, 74. pp. 427-431.

Dickey, D.A. and Fuller, W.A. (1981). ‘Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root.’ Econometrica, 49. pp. 1057-1072.

Dimitz, M.A. (2001). ‘Output Gaps in European Monetary Union: New Insights from Input Augmentation in the Technological Progress.’ Institute for Advanced Studies (IHS), Economics Series, No. 102.

Doornik, J.A. and Hansen, H. (1994). ‘An Omnibus Test for Univariate and Multivariate Normality.’ Nuffield College Discussion Paper. Available at http://www.nuff.ox.ac.uk/users/doornik

Doornik, J.A. and Hendry, D.F. (1996). GiveWin: An Interface to Economic Modelling. Timberlake Consultants LTD, London.

Doornik, J.A. and Hendry, D.F. (2001). PcGive Volume II (Modelling Dynamic Systems Using PcGive), 3rd edition. Timberlake Consultants LTD, London.

Dupasquier, C., Guay, A. and St-Amant, P. (1999). ‘A Survey of Alternative Methodologies for Estimating Potential Output and the Output Gap.’ Journal of Macroeconomics. 21(3). pp. 577-95.

Economic Policy Committee. (2001). ‘Report on Potential Output and the Output Gap.’ European Union. Available at http://europa.eu.int/comm/economy_finance/epc/documents/fina
output_en.pdf

Elmeskov, (1993). ‘High and Persistent Unemployment: Assessment of the Problem and its Causes.’ Organization for Economic Cooperation and Development. Working Paper No. 132.

Emerson, R.A. and Hendry, D.F. (1996). ‘An Evaluation of Forecasting using Leading Indicators.’ Journal of Forecasting, 15. pp. 271-291.

Engsted, T. and Haldrup, N. (1999). ‘Multicointegration in Stock-Flow Models.’ Oxford Bulletin of Economics and Statistics. 61(2). pp. 237-254.

Ericsson, N.R. (1983). ‘Asymptotic Properties of Instrumental Variables Statistics for Testing Non-Nested Hypotheses.’ Review of Economic Studies, 50. pp. 287-303.

Ericsson, N.R., Hendry, D.F. and Prestwich, K.M. (1998). ‘The Demand for Broad Money in the United Kingdom, 1878-1993.’ Scandinavian Journal of Economics, 100(1).

pp. 289-324.

Ericsson, N.R. and Irons, J.S. (1994). Testing Exogeneity. Oxford University Press, Oxford.

Espasa, A., Senra, E. and Albacete, R. (2002). 'Forecasting Monetary Union Inflation: A Disaggregated Approach by Countries and Sectors.' Universidad Carlos III de Madrid. Available at http://www.uc3m.es/uc3m/inst/FL/boletin/

European Commission. (2002). 'Production Function Approach to Calculating Potential Growth and Output Gaps - Estimates for the EU Member States and the US.' European Commission Economic Papers, No. 176.

Evans, G.W. (1989). 'A Measure of the U.S. Output Gap.' Economics Letters, 29(4). pp. 285-89.

Evans, G.W. and Reichlin, L. (1994). 'Information, Forecasts and the Measurement of the Business Cycle.' Journal of Monetary Economics, 33. pp. 233-254.

Fisher, J.D.M., Liu, C.T. and Zhou, R. (2002). 'When Can We Forecast Inflation?' Federal Reserve Bank of Chicago, Economic Perspectives (Q1). pp. 30-42.

Friedman, M. (1956). Studies in the Quantity Theory of Money. University of Chicago Press, Chicago.

Friedman, M. (1968). 'The Role of Monetary Policy.' American Economic Review, 58. pp. 1-17.

Flaig, G. and Plotscher, C. (2001). 'Estimating the Output Gap Using Business Survey Data: A Bivariate Structural Time Series Model for the German Economy.' Ifo-Studien. 47(2). pp. 221-32.

Fuhrer, J.C. and Moore, G.R. (1995). 'Inflation Persistence.' Quarterly Journal of Economics, 110. pp. 127-159.

Gerlach, S. and Smets, F. (1999). 'Output Gaps and Monetary Policy in the EMU Area.' European Economic Review, 43. pp. 801-812.

Gerlach, S. and Yiu, M. (2002). 'Unobservable-Component Estimates of Output Gaps in Five Asian Economies.' Centre for Economic Policy Research. Discussion Paper No. 3393.

Giorno, C., Richardson, P., Roseveare, D. and van den Noord, P. (1995). 'Estimating Potential Output, Output Gaps and Structural Budget Balances.' Organization for Economic Cooperation and Development. Working Paper No. 152.

Gordon, R.J. (2000). 'Does the 'New Economy' Measure up to the Great Inventions of the Past?' Journal of Economic Perspectives, 14(4). pp. 49-74.

Gourieroux, C. and Monfort, A. (1995). Simulation Based Econometric Methods. Universitu catholique de Louvain.

Green, P.J. and Silverman, B.W. (1994). Nonparametric Regression and Generalized Linear Models. A Roughness Penalty Approach. Chapman and Hall, London.

Haavelmo, T. (1944). 'The Probability Approach in Econometrics.' Econometrica, 12(supplement). pp. 1-118.

Hassler, J., Lundvik, P., Persson, T. and Soderlind, P. (1994). 'The Swedish Business Cycle. Stylized Facts over 130 years.' In Bergstrom, V. and Vredin, A. Measuring and Interpreting Business Cycles. Oxford University Press, Oxford.

Harding, D. and Pagan, A. (2000). 'Knowing the Cycle.' In Backhouse, R.E. and Salanti, A. (eds.) Macroeconomics and the Real World, Volume 1. Oxford University Press, Oxford.

Harvey, A. (1989). Forecasting, Structural Time Series Models and the Kalman Filter. Cambridge University Press, Cambridge.

Harvey, A. (1993). Time Series Models, 2nd Edition. Harvester Wheatsheaf, London.

Harvey, A., Henry, S.G.B., Peters, S. and Wren-Lewis, S. (1986). 'Stochastic Trends in Dynamic Regression Models: An Application to the Employment-Output Equation.' Economic Journal, 96. pp. 975-985.

Harvey, A.C. and Jaeger, A. (1993). 'Detrending, Stylized Facts and the Business Cycle.' Journal of Applied Econometrics, 8(3). pp. 231-247.

Harvey, A. and Koopman, S. (2000). 'Signal Extraction and the Formulation of Unobserved Components Models.' Econometrics Journal, 3. pp. 84-107.

Harvey, A. and Streibel, M. (1997). 'Trends, Cycles and Autoregressions.' Economic Journal, 107(440). pp. 192-201.

Harvey, A. and Trimbur, T. (2001). 'General Model Based Filters for Extracting Cycles and Trends in Economic Time Series.' University of Cambridge. Economics Department Working Paper No. 0113.

Hendry, D.F. (1984). 'Monte Carlo Experimentation in Econometrics.' In Griliches, Z. and Intriligator, M.D. (eds.) Handbook of Econometrics. Volume II. North-Holland, Amsterdam.

Hendry, D.F. (1995a). Dynamic Econometrics. Oxford University Press, Oxford.

Hendry, D.F. (1995b). 'Econometrics and Business Cycle Empirics.' Economic Journal, 105(433). pp. 1622-1636.

Hendry, D.F. (2000a). 'Does Money Determine UK Inflation over the Long Run?' In Backhouse, R.E. and Salanti, A. (eds.) Macroeconomics and the Real World, Volume 1. Oxford University Press, Oxford.

Hendry, D.F. (2000b). 'Measuring the 'Gap'.' Mimeo, Nuffield College, Oxford.

Hendry, D.F. (2001). 'Modelling UK Inflation, 1875-1991.' Journal of Applied Econometrics, 16(3). pp. 255-275.

Hendry, D.F. (2003). 'Economic Forecasting.' 2nd Year M.Phil Lecture Notes. University of Oxford.

Hendry, D.F. and Clements, M.P. (2001). 'Economic Forecasting: Some Lessons from Recent Research.' European Central Bank. Working Paper No. 82.

Hendry, D.F. and Doornik, J.A. (2001). PcGive 10 Volume I, 3rd edition. Timberlake Consultants LTD, London.

Hendry, D.F. and Ericsson, N.R. (1991). 'Modelling the Demand for Narrow Money in the United Kingdom and the United States.' European Economic Review, 35. pp. 833-886.

Hendry, D.F. and Krolzig, H.M. (2001). Automatic Econometric Model Selection with PcGets. Timberlake Consultants LTD, London.

Hendry, D.F. and Neale, A.J. (1987). 'Monte Carlo Experimentation Using PC-NAÏVE.' In Fomby, T.B. and Rhodes, G.F. (eds.) Advances in Econometrics, Vol. 6. JAI Press, Greenwich, Connecticut.

Hendry, D.F. and Richard, J.-F. (1982). 'On the Formulation of Empirical Models in Dynamic Econometrics.' Journal of Econometrics, 20. pp. 3-33. Reprinted in Hendry, D.F. (1993). Econometrics, Alchemy or Science? Blackwell Publishers, Oxford.

Hendry, D.F. and Richard, J.-F. (1991). 'Likelihood Evaluation for Dynamic Latent Variable Models.' In Amman, H. Belsley, D. and Pau, L. (eds.) Computational Economics and Econometrics. Kluwer, Dordrecht.

Hendry, D.F. and Richard, J.-F. (1998). 'On the Formulation of Empirical Models in Dynamic Econometrics.' In Whiteley, P.F. (eds.) Economic Policy, Vol. 1. E. Elgar, Cheltenham.

HMT. (2002). 'Trend Growth: Recent Developments and Prospects.' HM Treasury Publications. Available at http://www.hm-treasury.gov.uk/mediastore/otherfiles/ACF521.pdf

Hodrick, R.J. and Prescott, E.C. (1997). 'Postwar US Business Cycles: An Empirical Investigation.' Journal of Money, Credit and Banking, 29(1). pp. 1-16.

Holland, A. and Scott, A. (1998). 'The Determinants of UK Business Cycles.' Economic Journal, 108. pp. 1067-1092.

Hotelling, H. (1931). 'The Generalization of Student's Ratio.' The Annals of Mathematical Statistics, 2(3). pp. 360-378.

Hotelling, H. (1933). 'Analysis of a Complex of Statistical Variables into Principal Components.' Journal of Educational Psychology, 254. pp. 417-441, 498-520.

Hotelling, H. (1936). 'Relations Between Two Sets of Variates.' Biometrika, 28. pp. 321-377.

Hylleberg, S. (eds.) (1992). Modelling Seasonality. Oxford University Press, Oxford.

Johansen, S. (1995). 'Likelihood-Based Inference in Cointegrated Vector Autoregressive Models.' Oxford University Press, Oxford.

Kaiser, H.F. (1960). 'The Application of Electronic Computers to Factor Analysis.' Educational and Psychological Measurement, 20. pp. 141-151.

Kaiser, R. and Maravall, A. (1999). 'Estimation of the Business Cycle: A Modified Hodrick-Prescott Filter.' Banco de Espana Servicio de Estudios, Documentos de Trabajo: 9912.

Kim, C.J. (1994). 'Dynamic Linear Models with Markov Switching.' Journal of Econometrics, 60(1-2). pp. 1-22.

King, R.G., Plosser, C.I. and Rebelo, S.T. (1988). 'Production, Growth and Business Cycles I. The Basic Neoclassical Model.' Journal of Monetary Economics, 21. pp. 195-232.

King, R.G., Plosser, C.I., Stock, J.H. and Watson, M.W. (1991). 'Stochastic Trends and Economic Fluctuations.' American Economic Review, 81(4). pp. 819-840.

Koopman, S.J., Harvey, A.C., Doornik, J.A. and Shephard, N. (1995). STAMP. Timberlake Consultants LTD, London.

Kuttner, K.N. (1994). 'Estimating Potential Output as a Latent Variable.' Journal of Business and Economic Statistics, 12(3). pp. 361-68.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P. and Shin, Y. (1992). 'Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root.' Journal of Econometrics, 54. pp. 159-178.

Kydland, F.E. and Prescott, E.C. (1990). 'Business Cycles: Real Facts and a Monetary Myth.' Federal Reserve Bank of Minneapolis, Quarterly Review. pp. 3-18.

Larsen, J., Neiss, K. and Shorthall, F. (2002). 'Factor Utilization and Productivity Estimates for the United Kingdom.' Bank of England Working Paper No. 162.

Laxton, D., Meredith, G. and Rose, D. (1995). 'Asymmetric Effects of Economic Activity on Inflation: Evidence and Policy Implications.' International Monetary Fund. Staff Papers, 42(2). pp. 344-374.

Layard, R. and Nickell, S.J. (1986). 'The Performance of the British Labour Market.' London School of Economics, Centre for Labour Economics. Discussion Paper No. 249.

Layard, R., Nickell, S.J. and Jackman,R. (1991). Unemployment, Macroeconomic Performance and the Labour Market. Oxford University Press, Oxford.

Lindbeck, A. and Snower, D. (1994). 'How are Product Demand Changes Transmitted to the Labour Market?' Economic Journal, 104. pp. 386-398.

Lucas, R.E. (1973). 'Some International Evidence on Output-Inflation Trade-offs.' American Economic Review, 63. pp. 326-334.

Maddala, G.S. and Kim, I-M. (1998). Unit Roots, Cointegration and Structural Change. Cambridge University Press, Cambridge.

Mankiw, N.G. (2000). 'The Inexorable and Mysterious Trade-off between Inflation and Unemployment.' National Bureau of Economic Research. Working Paper No. 7884.

Massmann, M. and Mitchell, J. (2002). 'Have UK and Eurozone Business Cycles become more Correlated?' National Institute Economic Review, No. 182. pp. 58-71.

Massmann, M., Mitchell, J. and Weale, M. (2003). 'Business Cycles and Turning Points: A Survey of Statistical Techniques.' National Institute Economic Review, No. 183. pp. 90-106.

Maravall, A. (1997). 'Unobserved Components in Economic Time Series.' In Pesaran, H. and Wickens, T.R. (eds.) Handbook of Applied Econometrics. Blackwell Publishers, Oxford.

McDermott, C.J. and Scott, A. (1999). 'Concordance in Business Cycles.' Reserve Bank of New Zealand. Discussion Paper, No. G99/7.

McMorrow, K. and Roeger, W. (2001). 'Potential Output: Measurement Methods, "New" Economy Influences and Senarios for 2001-2010 - A Comparison of the EU15 and the US.' European Commision Economic Papers, No. 150.

Miller, E.M. (1983). 'A Difficulty in Measuring Productivity with a Perpetual Inventory Capital Stock Measure.' Oxford Bulletin of Economics and Statistics, 45(3). pp. 297-306.

Miller, S.M. (1998) 'The Beveridge Nelson Decomposition of Economic Time Series. Another Economical Computational Method.' Journal of Monetary Economics, 21. pp. 141-142.

Morley, J.C., Nelson, C.R. and Zivot, E. (2002). 'Why are Beveridge-Nelson and Unobserved-Component Decompositions of GDP so Different?' Referenced with authors permission.

Muellbauer, J. (1984). 'Aggregate Production Functions and Productivity Measurement: A New Look.' Centre for Economic Policy Research. Discussion Paper No. 33.

Muellbauer, J. (1996a). 'Productivity and Competitiveness.' In Jenkinson, T. (eds.) Readings in Macroeconomics. Oxford University Press, Oxford.

Muellbauer, J. (1996b). 'Income Persistence and Macro-Policy Feedbacks in the US.' Oxford Bulletin of Economics and Statistics, 58(4). pp. 703-733.

Muirhead, R.J. (1982). Aspects of Multivariate Statistical Theory. Wiley, New York.

Nelson, C.R. and Kang, H. (1981). 'Spurious Periodicity in Inappropriately Detrended Time Series.' Econometrica, 49(3). pp. 741-751.

Nelson, C.R. and Plosser, C.I. (1982). 'Trends and Random Walks in Macroeconomic Time Series. Some Evidence and Implications.' Journal of Monetary Economics, 10. pp. 139-162.

Nickell, S.J. (1990). 'Inflation and the UK Labour Market.' Oxford Review of Economic Policy, 6. pp. 26-35.

Nielsen, H. B. and C. Bowdler (2003). 'Inflation Adjustment in the Open Economy: An I(2) Analysis of UK Prices.' Nuffield College Working Paper No. 2003-05.

Oliner, S.D. and Sichel, D.F. (2000). 'The Resurgence of Growth in the Late 1990's: Is Information Technology the Story?' Journal of Economic Perspectives, 14(4). pp. 3-22.

Orphanides, A., Porter, R., Reifschneider, D., Tetlow, R. and Finan, F. (2000). 'Errors in the Measurement of the Output Gap and the Design of Monetary Policy.' Journal of Economics and Business, 52. pp. 117-141.

Orphanides, A. and van-Norden, S. (1999). 'The Reliability of Output Gap Estimates in Real Time.' Board of Governors of the Federal Reserve System. Finance and Economics Discussion Paper No. 99/38.

Oulton, N. (2001). 'ICT and Productivity Growth in the United Kingdom.' Bank of England. Working Paper No. 140.

Oulton, N. and Srinivasan, S. (2003). 'Capital Stocks, Capital Services, and Depreciation: An Integrated Framework.' Bank of England. Working Paper forthcoming.

Pagan, A.R. (1989). 'Econometric Issues in the Analysis of Regressions with Generated Regressors.' International Economic Review, 25(1). pp. 221-247.

Pantula, S.G., Gonzalez-Farias, G. and Fuller, W.A. (1994). 'A Comparison of Unit Root Test Criteria.' Journal of Business and Economic Statistics, 12. pp. 449-459.

Perron, P. (1989). 'The Great Crash, the Oil Shock and the Unit Root Hypothesis.' Econometrica, 57(6). pp. 1361-1401.

Perron, P. (1997). 'Further Evidence on Breaking Trend Functions in Macroeconomic Variables.' Journal of Econometrics, 80. pp. 355-385.

Phillips, A.W. (1958). 'The Relation between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861-1957.' Economica, 25. pp. 283-289.

Phillips, P.B. (1987). 'Time Series Regression with a Unit Root.' Econometrica, 55. pp. 277-302.

Plosser, C.I. and Schwert, G.W. (1978). 'Money, Income and Sunspots: Measuring Economic Relationships and the Effects of Differencing.' Journal of Monetary Economics, 4. pp. 637-660.

Prescott, E.C. (1986). 'Theory Ahead of Business Cycle Measurement.' Federal Reserve Bank of Minneapolis, Quarterly Review, 10(4). pp. 9-22.

Proietti, T. and Harvey, A. (2000). 'A Beveridge-Nelson Smoother.' Economic Letters, 67. pp. 139-146.

Roberts, J.M. (2001). 'Estimates of the Productivity Trend Using Time-Varying Parameter Techniques.' Contributions to Macroeconomics. 1(1), Article 3. Available at http://www.bepress .com/bejm/contributions/vol1/iss1/art3

Rogoff, K., Froot, K. and Kim, M. (2001). 'The Law of One Price Over Seven Hundred Years.' International Monetary Fund. Working Paper No. 01/174.

Sargan, J.D. (1959). 'The Estimation of Relationships with Autocorrelated Residuals by the use of Instrumental Variables.' Journal of the Royal Statistical Society B, 21. pp. 91-105. Reprinted in Sargan, J.D. Contributions to Econometrics, Volume 1. Cambridge University Press, Cambridge.

Sargan, J.D. (1980). 'A Model of Wage-Price Inflation.' Review of Economic Studies, 47. pp. 979-1012.

Schneeweiss, H. and Mathes, H. (1995). 'Factor Analysis and Principal Components.' Journal of Multivariate Analysis, 55. pp. 105-124.

Scott, A. (2000a). 'Stylized Facts from Output Gap Measures.' Reserve Bank of New Zealand. Discussion Paper No. 2000/07.

Scott, A. (2000b). 'A Multivariate Unobserved Components Model of Cyclical Activity.' Reserve Bank of New Zealand. Discussion Paper No. 2000/04.

Shapiro, M.D. (1987). 'Are Cyclical Fluctuations in Productivity Due More to Supply Shocks or Demand Shocks?' American Economic Review Papers and Proceedings, 77(2). pp. 118-124.

Slutzky, E. (1937). 'The Summation of Random Causes as the Source of Cyclic Processes.' Econometrica, 5. pp. 105-146.

Staiger, D., Stock, J.H. and Watson, M.W. (1996). 'How Precise are Estimates of the Natural Rate of Unemployment?' National Bureau of Economic Research. Working Paper No. 5477.

Stock, J.H. and Watson, M.W. (1998). 'Diffusion Indexes.' National Bureau of Economic Research. Working Paper No.6702.

Stock, J.H. and Watson, M.W. (1999a). 'Forecasting Inflation.' Journal of Monetary Economics, 44(2). pp. 293-335.

Stock, J.H. and Watson, M.W. (1999b). 'Business Cycle Fluctuations in US Macroeconomic Time Series.' In Taylor, J.B. and Woodford, M. (eds.) Handbook of Macroeconomics. North-Holland.

Turner, D. (1995). 'Speed Limit and Asymmetric Inflation Effects from the Output Gap in the Major Seven Economies.' OECD Economic Studies, 24. pp. 57-88.

Wallis, K.F. (1974). 'Seasonal Adjustment and Relations Between Variables.' Journal of the American Statistical Association, 69. pp. 18-31.

Watson, M.W. (1986). 'Univariate Detrending Methods with Stochastic Trends.' Journal of Monetary Economics, 18. pp. 49-75.

Wecker, W. and Ansley, C. (1983). 'The Signal Extraction Approach to Nonlinear Regression and Spline Smoothing.' Journal of the American Statistical Association. 78(381). pp. 81-89.

Zivot, E. and Andrews, D.W.K. (1992). 'Further Evidence on the Great Crash, the Oil Price Shock and the Unit Root Hypothesis.' Journal of Business and Economic Statistics, 10. pp. 251-270.