Real Time Representations of the Output Gap*

by

Anthony Garratt,† Kevin Lee,†† Emi Mise†† and Kalvinder Shields†††

Abstract

Methods are described for the appropriate use of data obtained and analysed in real time to represent the output gap. The methods employ cointegrating VAR techniques to model real time measures and realisations of output series jointly. The model is used to mitigate the impact of data revisions; to generate appropriate forecasts that can deliver economically-meaningful output trends and that can take into account the end-of-sample problems associated with the use of the Hodrick-Prescott filter in measuring these trends; and to calculate probability forecasts that convey in a clear way the uncertainties associated with the gap measures. The methods are applied to data for the US 1965q4-2004q4 and the improvements over standard methods are illustrated.

Keywords: Output gap measurement, real time data, data revision, HP end-points, probability forecasts.

JEL Classification: E52, E58.

*†Birkbeck College, London, UK, ††University of Leicester, UK, †††University of Melbourne, Australia. Corresponding author: Kevin Lee, Department of Economics, University of Leicester, Leicester, LE1 7RH, UK; web www.le.ac.uk/economics/kcl2. Version dated April 2005.

1 Introduction

The measurement of the output gap, i.e. the difference between the economy's actual output and its potential or trend level, is central to much applied macroeconometric work and particularly the analysis of monetary policy. However, it is widely recognised that the output gap is measured with considerable uncertainty, and this is especially true for the measures considered in real-time decision-making.[1] For example, Orphanides and van Norden (2002) [OvN] show, using US data, that the standard measures of this central concept are extremely unreliable, with ex post revisions of the gap in the US of the same order of magnitude as the estimated gap itself. Much of the unreliability arises because the gap measures are based on output data which is subsequently revised and on measures of the trend output level which are subject to estimation error. OvN decompose the revisions observed in their output gap measures into two parts reflecting these two sources of change. They show that, for their data, the effect of changes in the measurement of the trend exceed the effects of changes in the published data but that both effects are significant.[2]

The OvN analysis highlights the problems involved in real time decision-making by illustrating how their gap measure changes as new information on the actual and trend output levels becomes available with the release of each new vintage of data. However, the OvN decomposition is based on a recursive analysis of each successive vintage of data taken in turn. This ignores the possibility that the sequence of vintages released over time may in itself contain useful information with which to interpret the most recent vintage of data and to anticipate future outcomes (as discussed in Howery, 1978). Hence, for example, there might be systematic patterns in the data revisions that can be used, in

[1]It is also acknowledged that the use of ex post revised data can yield misleading descriptions of historical policy and that the use of real-time data generates different real-time policy recommendations to those obtained on the basis of ex post revised data (see, for example, Rotemberg and Woodford (1999), Brunner (2000), Orphanides *et al.* (2000), Orphanides (2001),and Amato and Swanson (2001)).

[2]These differences are potentially extremely important given the reliance of recent empirical work on the identification of monetary policy shocks and impulse responses on assumptions on the ordering of decisions and the timing of the release of information. See Christiano, Eichenbaum and Evans (1999) or Garratt *et al.* (2005) for reviews.

conjunction with the real time data, both to moderate the direct impact of the revisions obtained in successive vintages of data on the perceived current output level and to look forward to offset their impact on the output trend measure.[3]

In this paper, we exploit the information contained in the sequence of vintages more fully than OvN through a cointegrating VAR model which, under reasonable assumptions on the nature of the output series and measurement errors, explains both the changes in the real time data and its revisions. The model is used to generate forecasts of contemporaneous and future values of output. The forecasts improve the accuracy with which the true level of activity is measured and they can be used to supplement the historically-observed series to obtain improved measures of the underlying trends also. As explained in Mise *et al.* (2005a,b) [denoted MKN], this latter point helps to address the end-of-sample problems associated with the widely-used Hodrick-Prescott (1997) [HP] filter in the measurement of the trend (this being the source of considerable estimation error variance). The model can be estimated recursively, taking into account successive vintages of data. But, because it describes the revision process as well as the underlying output process, the model makes use of all the information available at each point in time, not just the most recent vintage available.

The proposed approach to measuring the output gap has at least three very useful properties. *First*, the output gap is measured relatively precisely because modelling the revision process moderates the effect of changes in published data, while the use of the forecasts mitigates the end-of-sample problem associated with the use of the HP filter. *Second*, by linking the trend measure to forecasts of future output levels, it can be related to the frequently-used Beveridge-Nelson trend and can be readily interpreted in terms of economically-meaningful concepts such as 'potential output'. And *third*, as well as producing point estimates of the output gap that are measured relatively precisely, the underlying model can be used to describe clearly the uncertainties associated with the

[3] OvN also illustrate the uncertainty in the gap measure arising through the choice of detrending techniques. (See also Canova, 1998.) We do not consider this aspect of the uncertainty in this paper, although the methods described for conveying the extent of the uncertainty on the gap could be extended to accommodate alternative detrending techniques.

measure of the gap. This is extremely useful because, while it is important to recognise the unreliability of the output gap measures, the estimated values of the gap at different horizons are nevertheless an essential requirement in many decision-making contexts. The output gap measures can be used appropriately, taking into account the uncertainties surrounding these, when the model is used to supplement the point forecasts with forecasts of the probability of the occurrence of particular events involving the gap.

The remainder of the paper is organised as follows. In Section 2, the proposed method for measuring the output gap is elaborated through a description of the cointegrating VAR model, through a discussion of the HP filter and its limitations when the focus of attention is the end-of-sample, and through a comment on the calculation of probability forecasts relating to the output gap. Section 3 describes the application of the proposed methods to obtain output gap measures for the US and compares these with measures obtained following the procedures of OvN. Section 4 presents some probability forecasts obtained using our modelling framework, and Section 5 concludes.

2 Measuring the Output Gap with Real Time Data

To describe our proposed method of measuring the output gap, we need to introduce some notation and terminology. We write (the logarithm of) the output level at time $t-j$ by y_{t-j}, and denote the measure of output at time $t-j$ that is released in time t by ${}_{t}y_{t-j}$, $j = 0, 1, 2, , ...$ Throughout the paper, the "vintage-t" dataset is defined by $Y_t = \{{}_{t}y_{t-1},\ {}_{t}y_{t-2},\ {}_{t}y_{t-3}, ...\}$ so that it includes the time-t measure of output at time $t-1$ and before. Note that it is assumed that the first release of output data for any period takes place after a one-period delay; this corresponds to practice in both the US and UK.[4] The full information set available at time t, denoted Ω_t, contains the datasets of all vintages dated at t and earlier; i.e. $\Omega_t = \{Y_t,\ Y_{t-1},\ Y_{t-2}, ...\}$. It is worth noting that the time-$(t+1)$ measure of a variable is simply the time-t measure plus the revision; i.e. ${}_{t+1}y_{t-1} = {}_{t}y_{t-1} + ({}_{t+1}y_{t-1} -_{t} y_{t-1})$. Hence, the full information set grows with the

[4]In the US, for example, the first release of output data provides an indication of output levels in the previous quarter. The subsequent releases provide progressively more information, and the July release in each year provides the most definitive statement of output over the previous year.

addition of successive vintages of datasets by including the news on the output level in the previous period (the 'first release' of information on the output level in that period) and the revisions on the output series in previous periods; i.e. $\Omega_{t+1} = \Omega_t \cup \{_{t+1}y_t,\ (_{t+1}y_{t-1} - _t y_{t-1}),\ (_{t+1}y_{t-2} - _t y_{t-2}), ...\}$.[5] Finally, turning to the output trend, we note that there are a variety of methods employed in the literature to obtain measures of the output trend at time t. Some of these make use of data that becomes available both before and after time t, so that care also needs to be exercised in describing the information set on which the trend measure is based. Specifically, writing the trend output level at time $t-j$ by $\widetilde{y}_{t-j}$, we denote the measure of trend output at time $t-j$ that is calculated using method k on the basis of an information set available at time t, say Ω_t, by $\widetilde{y}^k_{t-j}|\Omega_t$.

In OvN, attention is focused on the differences between 'real time' measures of the output gap based on successive vintages of output data and 'final' measures obtained from the last available vintage of data. Hence the comparison is between the real time measure of the gap $x^{ro}_t = {}_{t+1}y_t - \widetilde{y}^o_t|Y_{t+1}$ and the final measure $x^{fo}_t = {}_T y_t - \widetilde{y}^o_t|Y_T$, where $t = 1, ..., T-1$, and the 'o' superscript denotes the HP filter method used by OvN. OvN also consider a 'quasi-real' estimate, $x^{qo}_t = {}_T y_t - \widetilde{y}^o_t|Y_{T,t}$, in which the time-$T$ measure of output at time t is compared to a trend measure obtained on the basis of a subset of Y_T; namely $Y_{T,t} = \{_T y_t,\ _T y_{t-1},\ _T y_{t-2}, ...\}$, $t < T$. Evaluating the differences between the quasi-real measure of the output gap x^{qo}_t and the real time measure x^{ro}_t isolates the changes in the gap measures arising from the revision of the trend measure in the light of subsequent data. OvN find this element to be significant but relatively small, and it is argued that it is the addition of new points to the sample, which causes $\widetilde{y}^o_t|Y_T$ to deviate from $\widetilde{y}^o_t|Y_{T,t}$, that explains much of the difference between the real time and final measures of the output gap.

OvN's three measures of the output gap highlight the different effects of revisions in published data and of differences in the use of information. But their decomposition is

[5] Realistically, the measurement of output at a particular time will be revised for some time, say p periods, after which no 'news' becomes available on this output level and measured output will remain unchanged. Full information therefore consists of the first release of the output level measurement plus the p subsequent revisions.

potentially misleading. For example, MKN show that the HP filter has very poor properties in estimating the trend component of the end-point of a series, with the estimation variance upto 40 times that of the error inherent in the series in some circumstances (see also Baxter and King (1999), or St-Amant and van Norden (1998)). So, even if is decided to focus only on the vintage-t dataset in determining the trend output measure, it is important to take appropriate account of all the information available in that dataset, including its implications for expected future values of the series, to try to mitigate the end-of-sample effect. OvN's chosen method of calculating the HP filter does not do this so that comparison of the quasi real and final measures do not reflect accurately the impact of the effect of changes in the trend measure over time.[6] Moreover, focusing on vintage-t data without reference to the revisions that have taken place in previous periods' data potentially overstates the effects of changes in the published data in time t, since these revisions might have been anticipated. The conclusion, then, is that all information available at time t should be employed in constructing an output gap measure in real time, with particular attention payed to forecasts of future values of the output series. The appropriate modelling framework for accommodating all information is described in the section below, and this is then used to explain how forecasts can be used to eliminate the end-of-sample problems associated with the HP filter.

2.1 A Joint Model of Actual and Revised Output Series

In order to make use of the full information available, the real time measures of output should be modelled alongside the "actual", realised value of output, taking into account the revision process as well as the underlying output process.[7] In most of this section, we assume for illustrative purposes that data is revised just once after its initial release, so that we can model the two processes jointly in a bivariate VAR. However, we note also

[6] As we discuss below, forecasts of future values of the series (obtained in real time), can be used to supplement the vintage-t data and 'extend' the series beyond the end-of-sample. The 'mid-sample' trend estimate obtained for time t in this way is clearly more comparable to OvN's final measure than the trend underlying the quasi-real output gap.

[7] See also Howery (1978) and Diebold and Rudebusch (1991).

that if revisions continue upto q periods after the first release of data, then a VAR of size $q+1$ would be required to model the processes adequately, and we illustrate this more general case too.

Our modelling approach assumes first that actual output is first-difference stationary. This means that, if data on output is released with a one period delay and the actual output is observed with the revision after one further period, $({}_t y_{t-2} - {}_{t-1} y_{t-3})$ is stationary. The approach also assumes that measurement errors (i.e. revisions) are also stationary. The first of these assumptions is supported by considerable empirical evidence,[8] and the latter is eminently reasonable. Under these assumptions, any linear combination of these two series can be modelled in a bivariate VAR.[9] Hence, the output growth measure $({}_t y_{t-1} - {}_{t-1} y_{t-2})$ and the data revision series have the following joint fundamental Wold representation:

$$\begin{bmatrix} {}_t y_{t-1} - {}_{t-1} y_{t-2} \\ {}_t y_{t-2} - {}_{t-1} y_{t-2} \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} + \mathbf{A}(\mathbf{L}) \begin{bmatrix} \epsilon_t \\ \xi_t \end{bmatrix} \tag{2.1}$$

Here, α_1 is mean output growth (measured by 'first-release' data), α_2 is the mean value of the revisions, $\mathbf{A}(\mathbf{L}) = \sum_{j=0}^{\infty} \mathbf{A}_j(L)$, where the $\{\mathbf{A}_j\}$ are 2×2 matrices of parameters, assumed to be absolutely summable, and L is the lag-operator. Also, ϵ_t and ξ_t are mean zero, stationary innovations, with non-singular covariance matrix $\mathbf{\Psi} = \psi_{jk},\ j,k = 1,2$. The model in (2.1) emphasises the point that the chosen measure of output growth at time $t-1$ and the revision of the measure of output at time $t-2$ between $t-1$ and t are both revealed at time t. For notational convenience, in what follows we write $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)'$, where $\alpha_2 = 0$ if there is no bias in the measurement error.

The general model in (2.1) can be expressed in various different ways. For example, assume that $\mathbf{A}^{-1}(\mathbf{L})$ can be approximated by the lag polynomial $\mathbf{A}^{-1}(\mathbf{L}) = \mathbf{B}_0 + \mathbf{B}_1\mathbf{L} + .. + \mathbf{B}_{p-1}\mathbf{L}^{p-1}$, where $\mathbf{B}_0 = \mathbf{I}_2$ without loss of generality. In this case, (2.1) can be

[8] See, for example, Murray and Nelson (2000) and Pappell and Prodan (2004).

[9] For example, output growth measured by the change in the 'first-release' output level, $({}_t y_{t-1} - {}_{t-1} y_{t-2})$, can be written in terms of actual growth and the relevant revisions and so is itself stationary; i.e. $({}_t y_{t-1} - {}_{t-1} y_{t-2}) = ({}_{t+1} y_{t-1} - {}_t y_{t-2}) + ({}_t y_{t-1} - {}_{t+1} y_{t-1}) - ({}_{t-1} y_{t-2} - {}_t y_{t-2})$.

rewritten to obtain the AR representation

$$\begin{bmatrix} {}_ty_{t-1} - {}_{t-1}y_{t-2} \\ {}_ty_{t-2} - {}_{t-1}y_{t-2} \end{bmatrix} = \mathbf{a} - \mathbf{B}_1 \begin{bmatrix} {}_{t-1}y_{t-2} - {}_{t-2}y_{t-3} \\ {}_{t-1}y_{t-3} - {}_{t-2}y_{t-3} \end{bmatrix} - \cdots - \mathbf{B}_{p-1} \begin{bmatrix} {}_{t-p+1}y_{t-p} - {}_{t-p}y_{t-p-1} \\ {}_{t-p+1}y_{t-p-1} - {}_{t-p}y_{t-p-1} \end{bmatrix} + \begin{bmatrix} \epsilon_t \\ \xi_t \end{bmatrix} \tag{2.2}$$

and hence

$$\begin{bmatrix} {}_ty_{t-1} \\ {}_ty_{t-2} \end{bmatrix} = \mathbf{a} + \mathbf{\Phi}_1 \begin{bmatrix} {}_{t-1}y_{t-2} \\ {}_{t-1}y_{t-3} \end{bmatrix} + \mathbf{\Phi}_2 \begin{bmatrix} {}_{t-2}y_{t-3} \\ {}_{t-2}y_{t-4} \end{bmatrix} + \cdots + \mathbf{\Phi}_p \begin{bmatrix} {}_{t-p}y_{t-p-1} \\ {}_{t-p}y_{t-p-2} \end{bmatrix} + \begin{bmatrix} \epsilon_t \\ \xi_t \end{bmatrix} \tag{2.3}$$

where $\mathbf{a} = \mathbf{A}^{-1}(\mathbf{1})\boldsymbol{\alpha}$,

$$\mathbf{\Phi}_j = \mathbf{B}_{j-1} \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} - \mathbf{B}_j \text{ for } j = 1, ..., p-1, \quad \text{and} \quad \mathbf{\Phi}_p = \mathbf{B}_{p-1} \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}.$$

Seen in the context of (2.3), the vector of errors $(\epsilon_t, \xi_t)'$ has a clear interpretation: ϵ_t is the "news on output level in time $t-1$ contained in the first-release data becoming available at time t"; and ξ_t is the "news on the level of output in time $t-2$ contained in the revised data becoming available at time t".

Alternatively, manipulation of (2.3) also provides the VECM representation explaining the changes in the first release measures and the change in output realisations, $[\Delta_t y_{t-1}, \Delta_t y_{t-2}]$ where $\Delta = (1-L)$ is the difference operator. As shown in the Appendix, the VECM representation includes the lagged value of $({}_ty_{t-1} - {}_ty_{t-2})$ as a regressor since these two series are cointegrated, with cointegrating vector $\boldsymbol{\beta}' = [1, -1]$. This property holds because revisions are taken to be stationary in this model, so that first-release and actual output levels are cointegrated by assumption.[10] Note that the model at (2.1), and its equivalent forms, are quite general and have no implications for the nature of the measurement error other than it is stationary. However, the assumption that real time measures are unbiased (in the sense that measurement errors have no systematic content) can be accommodated in the model through the imposition of restrictions. If first-release measures are unbiased, we would have ${}_ty_{t-2} = {}_{t-1}y_{t-2} + \xi_t$ so that, in (2.3), the second row of $\mathbf{\Phi}_1 = \begin{pmatrix} 1 & 0 \end{pmatrix}$, and the second row of $\mathbf{\Phi}_j = \begin{pmatrix} 0 & 0 \end{pmatrix}$, $j = 2, ..., p$.

[10]The VECM representation also has implications for the corresponding MA representation in first differences; see Appendix for details.

Finally here, we note that the above models can be readily extended when the revision process extends beyond just one period. Hence, for example, if quarterly data continues to be revised for upto a year, then the data requires a four-variable VAR to capture the joint determination of the first-release output series and the three successive revisions. Hence, the model that will accommodate the news on output levels contained in the first-release data (ϵ_t) and in all the revised data becoming available at time t on the previous periods (ξ_{1t}, ξ_{2t}, ξ_{3t}) can be written in a form corresponding to (2.2),

$$\begin{bmatrix} {}_t y_{t-1} - {}_{t-1} y_{t-2} \\ {}_t y_{t-2} - {}_{t-1} y_{t-2} \\ {}_t y_{t-3} - {}_{t-1} y_{t-3} \\ {}_t y_{t-4} - {}_{t-1} y_{t-4} \end{bmatrix} = \mathbf{a} - \mathbf{B}_1 \begin{bmatrix} {}_{t-1} y_{t-2} - {}_{t-2} y_{t-3} \\ {}_{t-1} y_{t-3} - {}_{t-2} y_{t-3} \\ {}_{t-1} y_{t-4} - {}_{t-2} y_{t-4} \\ {}_{t-1} y_{t-5} - {}_{t-2} y_{t-5} \end{bmatrix} - \cdots$$

$$-\mathbf{B}_{p-1} \begin{bmatrix} {}_{t-p+1} y_{t-p} - {}_{t-p} y_{t-p-1} \\ {}_{t-p+1} y_{t-p-1} - {}_{t-p} y_{t-p-1} \\ {}_{t-p+1} y_{t-p-2} - {}_{t-p} y_{t-p-2} \\ {}_{t-p+1} y_{t-p-3} - {}_{t-p} y_{t-p-3} \end{bmatrix} + \begin{bmatrix} \epsilon_t \\ \xi_{1t} \\ \xi_{2t} \\ \xi_{3t} \end{bmatrix}. \quad (2.4)$$

This can be rewritten in levels form, in VECM form and in MA form exactly as in (2.3) and the models of the Appendix.

2.2 Measuring Trend Output and the Output Gap

Estimates of the bivariate or multivariate models derived above can be used to generate forecasts of the output series infinitely into the future and, in this section, we argue that these can be usefully applied in the measurement of the output trend using the HP filter. To motivate this procedure, recall that the HP filter is an additive decomposition $y_t = \widetilde{y}_t + x_t$ where $\widetilde{y}_t$ is identified as a growth (trend) component and x_t as a cyclical component. The HP filter is an exponentially weighted moving average filter, and is two-sided symmetric in the sense that it uses both past and future observations with equal importance in order to decompose any one observation in a series. The HP filter has the desirable property that it is optimal, in the expected squared error sense, for data generating processes of the form

$$\begin{aligned}(1-L)^2\widetilde{y}_t &= A(L)\varepsilon_t \quad ; \quad x_t = A(L)u_t \\ A(L) &= \sum_{j=0}^{\infty} a_j L^j \quad ; \quad \sum_{j=0}^{\infty} a_j^2 < \infty\end{aligned} \tag{2.5}$$

where ε_t and u_t are mutually stochastically uncorrelated white noise processes (i.e. $E(\varepsilon_t u_s) = 0 \ \ \forall t, s$), and where their variance ratio is

$$\lambda = \left[\frac{\sigma_u}{\sigma_\varepsilon}\right]^2, \tag{2.6}$$

with λ being the value of the 'smoothness' parameter.[11]. Moreover, although the optimality conditions (2.5) to (2.6) are expressed in terms of unobserved components, MKN show that all ARIMA$(p, 2, q)$ models that can be fitted to the observed series y_t can be expressed in this framework. In particular, this holds true for all possible ARIMA$(p, 1, q)$ models, with $A(L)$ in (2.5) involving a unit moving average root, so that the series and its trend component are $I(1)$. Here, if y_t is an ARIMA$(p, 1, q)$, then $\widetilde{y}_t$ is ARIMA$(p+2, 1, q)$ and x_t is ARMA$(p+2, q+1)$.

However, an important feature of the HP filter is that, when we have a finite series, the optimality properties only hold for the mid-point of the series. As we move towards the end of the series, the HP filter becomes increasingly one-sided, and for the last observation of the series, the filter is completely one-sided. MKN note that, while the filter continues to provide an unbiased estimate of the quantity x_t at the endpoints of a finite series, the estimates are inefficient. They illustrate the extent of the inefficiency by comparing the estimated HP trend measures with the actual trends present in a variety of simulated series obtained using different trend and cycle specifications. In particular, MKN note Burman's (1980) suggestion for addressing the inefficiency issue by augmenting the observed series with optimal linear forecasts and demonstrate, through their simulation exercises, that the application of the HP filter to the augmented series provides an estimate of the end-of-sample observation which is optimal. Indeed, by augmenting a series by its univariate forecast, the standard deviation of the estimation error for the cyclical component is

[11]This parameter is conventionally set to 1600 for quarterly data, following a suggestion by Hodrick and Prescott (1997), made on somewhat arbitrary grounds.

reduced by upto half (relative to the standard application of the HP filter) in their various simulations.[12]

The clear implication of these results is that the output gap should be calculated using a trend obtained by applying the HP filter to the forecast-augmented output series. For the series described in the previous section, the model at (2.1), or its equivalent forms in (2.2) or (2.3), can provide the vehicle for generating these forecasts. Forecasts of the output series ${}_{t+1}y_t$, ${}_{t+2}y_{t+1}$, ${}_{t+3}y_{t+2}$,... could be generated using a univariate model of the vintage-t data, but this will generally be less efficient than that provided by the bivariate model of (2.1) which uses all the information available.[13] We shall denote the end-of-sample trend measure obtained by applying the HP filter to the output series augmented by forecasts from the univariate model obtained using vintage-t data by $\widetilde{y}^u_{t-1}|Y_t$ and the corresponding measure obtained using the bivariate model of (2.1) by $\widetilde{y}^m_{t-1}|\Omega_t$, where the '$u$' and '$m$' superscripts indicates the use of forecast-augmentations, suggested in MKN, based on univariate and multivariate models respectively.

The application of the HP filter to forecast-augmented series not only improves the statistical properties of the derived series, but it also provides an interpretation of the trend that the traditional HP trend does not have and helps reconcile the use of the HP trend with those who prefer to use Beveridge-Nelson (BN) trends. Specifically, the BN trend obtained from a time series analysis of output measures the infinite-horizon effect of shocks on the series. Since this trend measure shows the permanent long-run effect of a shock to output, the BN trend is often interpreted as 'potential' output since this is the level to which the economy will converge in the absence of any further shocks. However, while this measure has an intuitively reasonable interpretation, it has the disadvantage that it does not pay attention to the dynamic path that is taken by the output series

[12]MKN also note that the HP filter is often used in contexts where there is no assumed underlying 'true' trend and cycle measures of the form (2.5) or indeed any other form. They comment that the reliability of a trend measure can be assessed in these circumstances if a measure based on a sample of data $1,..,T$ is revised as little as possible in the light of subsequent observations; this matches the discussion of OvN on the comparison of their 'quasi real' and 'final' trend estimates. MKN confirm through their simulations that these revisions are indeed minimised when the HP filter is applied to the forecast-augmented series.

[13]The exception is when the real time measures are unbiased as discussed in the earlier footnote.

as it moves to its infinite-horizon forecast. This means that the BN trend may be even more volatile than the actual output series itself (especially if it is based on a simple autoregressive model of the output series) and it is often argued that this reduces the usefulness of the trend in the context of measures of the output gap for monetary policy or business cycle analysis. This is clearly not the case for the trend obtained from the forecast-augmented HP filter, $\widetilde{y}^{m}_{t-1}|\Omega_t$, which will display the smoothness typically desired of a trend measure to be used as a measure of the business cycle or output gap. Further the forecast-augmented HP filter series is reconciled with the BN trend since forecasts of future values of $\widetilde{y}^{m}_{t-1}|\Omega_t$ will, by construction, themselves converge towards the output level forecast at the infinite-horizon. The proposed measure therefore has the smoothness properties and, at the long horizon, it can be interpreted as a potential output series.

2.3 Conveying the Uncertainty Surrounding the Output Gap Measures

In practice, decision-makers faced with the complete set of vintages of data upto and including that at time T are concerned with obtaining a measure of the output gap for the end-of-sample period (and possibly into the future). In some cases, attention focuses simply on whether the gap is positive or negative, but in any case it is the time-T (and future) magnitudes that matter in real time decision-making. Here, assuming again that data is released with a one period delay and there is a single revision made, this means decision-makers are interested in forecasts of $x^{fm}_T =_{T+2} y_T - \widetilde{y}^{m}_T|\Omega_{T+H}$. Hence, the relevant output level to be forecast is $_{T+2}y_T$, the time-T output level that will be observed in $T+2$, taking into account the one period delay in the release of data and after any revisions in the data have been fully taken into account. And the relevant trend measure to be forecast is that obtained on the basis of an information set that is available at some forecast horizon well into the future (at $T+H$) so that there are no end-of sample problems for the measure at T.

We can obtain point forecasts of this magnitude relatively easily: the point forecast of $_{T+2}y_T$ is obtained straightforwardly from the bivariate model of (2.2) based on Ω_T; and the forecast of $\widetilde{y}^{m}_T|\Omega_{T+H}$, based on Ω_T, is simply the period-T observation of $\widetilde{y}^{m}_T|\Omega_T$.[14] But

[14]This follows because the measure $\widetilde{y}^{m}_T|\Omega_T$ is itself based on forecast values of the future unrevised and

the point forecast of the gap obviously does not convey the uncertainty associated with the output gap measure, and this is potentially significant here given that forecasts of the revised and unrevised series are used in various different ways in the construction of the measure. So, using the information set Ω_T at time T for example, there will be uncertainty associated with the output gap measure at time $T-2$ because of the uncertainty over the values of output beyond T and, hence, over the measure of the trend (this is due to the end-of-sample problem which is reduced by the forecast augmentation but not eliminated). This uncertainty is compounded in the measure dated at $T-1$ by the forecast revisions that will be made to the first-release data on ${}_Ty_{T-1}$ and then further compounded at T and beyond as the unrevised output series and revisions are subsequently forecasted.

It is important, therefore, that any output gap measure is supplemented with information on the uncertainties associated with the measure. Indeed, as noted above, it is frequently the case that interest focuses not on the size of the output gap but rather on its sign (i.e. whether it is positive or negative). This reflects the fact that decision-makers' objective functions are often concerned with 'booms' and 'recessions' (irrespective of their size) and that these episodes are not valued symmetrically (so that the costs incurred during a recession might outweigh the benefits experienced in boom, say). In these circumstances, the decision-maker requires the entire probability distribution function (pdf) of the estimated output gap measure, rather than its point forecast, or at least an explicit forecast of the probability that the output gap will exceed or fall below zero.[15]

The calculation of probability forecasts and pdf's of this sort is relatively unusual in economics (where uncertainty is typically conveyed, if at all, by the reporting of confidence intervals). But the methods are relatively straightforward to implement and are described in Garratt *et al.* (2003). For example, abstracting from parameter uncertainty for the time being, to calculate the pdf associated with the forecast of $x_T^{fm} = {}_{T+2}y_T - \widetilde{y}_T^m|\Omega_{T+H}$, one

revised series and in the absense of any additional information, the value of the updated series expected to be observed in $T+H$ is unchanged from that measured in period T (cf. the Law of Iterated Expectations)

[15]There is widespread recognition that the design of optimal monetary policy must take into account the various forms of uncertainties faced by the monetary authorities, including those involving imperfect information about the current state of the economy as well as future developments. See, for example, Svensson (2001, 2002).

would use the estimated model of (2.2) to generate R replications of the future vintages of data, denoted $\widehat{Y}_{T+h}^{(r)}$ for $h = 1, ..., H$ and $r = 1, ..., R$. These include values of ${}_{T+2+h}\widehat{y}_{T+h}^{(r)}$, $h = 0, 1, 2, ..., H-2$, on which the trend measure $\widetilde{y}_T^{m(r)} | \Omega_{T+H}^{(r)}$ can be based. The simulated distribution of $\widehat{x}_T^{fm(r)} = {}_{T+2}\widehat{y}_T^{(r)} - \widetilde{y}_T^{m(r)} | \Omega_{T+H}^{(r)}$ obtained in this way provides the pdf of the output gap measure directly, while counting the number of times an event occurs in these simulations provides a forecast of the probability that the event will occur; the fraction of the simulations in which $\widehat{x}_T^{fm(r)} > 0$ provides an estimate of the forecast probability that the time-T output gap is positive, for example. Extending the simulation exercise to accommodate parameter uncertainty is relatively straightforward (see Garratt *et al.* (2003) for more details), so that a complete characterisation of the uncertainty surrounding the output gap measure can be obtained.[16]

3 Output Gaps in the US

The methods described above are applied to the vintages of US output data provided by the Federal Reserve Bank of Philadelphia at http//www.phil.frb.org/econ/forecast/index.html.. This dataset includes 157 vintages of data; the first vintage is dated 1965q4 and the final vintage is dated 2004q4. All vintages of data run from 1947q1 upto one period prior to the release date; i.e. $Y_t = \{{}_t y_{1947q1}, ...,{}_t y_{t-1}\}$, $t = 1965q4, ...2004q4$.

The first exercise undertaken on this data aims to investigate the gains from using the forecast augmented HP filter approach to defining the trend. In the first instance, we follow OvN and consider the successive vintages of data, applying the HP filter, to derive the 'real-time measure' $\widetilde{y}_t^o | Y_{t+1}$, $t = 1965q4, ...2004q3$ as the end-of-sample observation of the trend in each recursion. We compare this with the 'quasi real' measure $\widetilde{y}_t^o | Y_{T,t}$, also derived recursively, and the 'final' measure $\widetilde{y}_t^o | Y_T$. We also derive the corresponding trends

[16] In this exercise, we restrict attention to the forecast-augmented HP filter as a measure of the trend. But alternative measures of the trend exist and, in principle, these could be calculated in each of the simulation exercises also to provide alternative gap measures. If the 'reasonableness' of the alternative approaches to measuring the trend could be captured by appropriate weights, these various sets of simulations could be pooled to provide density functions for the gap measures, and event probability forecasts, that accommodate stochastic uncertainty, parameter uncertainty, and the uncertainties associated with the appropriate measure of trend.

based on data augmented by forecasts. The forecasts are based on eighth-order univariate autoregressions explaining $({}_{t}y_{t-1} - {}_{t}y_{t-2})$; an eighth-order autoregression is applied to ensure there is no serial correlation in the residuals.[17]

Figure 1 shows the output gaps considered by OvN, namely $x_t^{ro} = ({}_{t+1}y_t - \widetilde{y}_t^o | Y_{t+1})$, $x_t^{qo} = ({}_{T}y_t - \widetilde{y}_t^o | Y_{T,t})$ and $x_t^{fo} = ({}_{T}y_t - \widetilde{y}_t^o | Y_T)$, for $t = 1965q3, ..., 2004q3$, and $T = 2004q4$, and illustrates the considerable differences arising out of data revisions and the end-of-sample effects on the underlying trends. Table 1 shows that the correlation between the real time and final measures of OvN is just 0.526, and the two measures agree on whether output growth is above or below trend in only 63% of the sample period. These figures rise to 0.628 and 69% respectively when the comparison is between the quasi real time and final measures (abstracting from effects of data revision) but the figures are clearly still not high. Taking the final measure x_t^{fo} as the best indicator of the true output gap available to OvN, it is the poor performance of the x_t^{ro} and x_t^{qo} measures to reflect the true output gap that is the basis of OvN's conclusion that real time measures of the gap are unreliable.

Table 1 also describes the effect of employing the forecast-augmentation method of calculating the trends on the three gap measures. This has a substantial impact on the variability of the output gap series, cutting the standard deviation and range of values for the real time measure by around 30% and by nearer 40% in the case of the quasi-real measure (this reduction in variability is clearly illustrated in Figure 2). This illustrates that the forecast-augmentation is having a considerable impact on the trend measure as the estimation error variance associated with the application of the HP filter at the end-of-sample is reduced. The effect is to raise the correlation between the final measure x_t^{fu} and the real and quasi real measures to 0.77 and 0.78 respectively, and agreement on the occurrence of booms and recessions rise to 83% and 81% respectively also. The improvement in reliability using the forecast-augmentation method is pronounced and shows that the augmentation should be applied in output gap measures. However, they remain far from perfect, and it is clear that output gap measures obtained in real time

[17] Details of regressions, and diagnostic tests relating to the order of integration of the output and revision series, are not presented for space considerations but are available from the authors on request.

need to be treated with caution and the uncertainties on their measurement appropriately taken into account.

Next, we turn to the multivariate analysis of the output growth and revision processes together, considering whether there are systematic patterns in the data revisions that underlie the successive vintages of data that are released and the extent to which a model of the output growth data is enhanced by modelling the measured output growth and revisions data jointly. To do this, we need to choose the lag length p in the multivariate model in (2.4) and the length of the "revision horizon" (after which revisions are unsystematic and insignificant). The maximum lag length we consider is $p = 4$ and the maximum length of the revision horizon we consider is 3, as in (2.4). It turns out that the data is described adequately if we allow for a revision horizon of two quarters and lags in the VAR of order 2. To demonstrate this, Table 2 provides estimates of (2.4) obtained using the entire data upto and including Y_{2004q4}.[18] The Table shows that a revision horizon of 2 is sufficient to capture systematic elements in the revision process, since none of the variables in the fourth column, explaining time-t revisions of data at $t-4$ are individually or jointly statistically significant. And the Table also provides variable exclusion tests, denoted $\chi^2_{LM}(10)$, showing that the third and fourth lags of the first three variables in our system and all four lags of the fourth can be safely dropped from the regressions without violating the data. Moreover, it is apparent that the joint modelling of the growth series and the revisions is a useful approach. Both the lagged growth series and the lagged revisions contribute significantly to the explanation of the time-t growth $({}_ty_{t-1} -_{t-1} y_{t-2})$, meaning that the univariate model is misspecified, and there are very significant systematic elements in the revisions $({}_ty_{t-2} -_{t-1} y_{t-2})$ and $({}_ty_{t-3} -_{t-1} y_{t-3})$.[19] We are confident, therefore, that the multivariate model of growth and revision is appropriate and will provide more accuracy in the forecasts of future output on which to base the trends and output gap measures.

[18] In fact, we conducted this exercise recursively on the full information set, Ω_t , consisting of all of the vintages of data upto and including Y_t, for $t = 1965q4, .., 2004q4$. Although we only report the results of the $2004q4$ analysis, qualitatively similar results were obtained throughout.

[19] Similar systemmatic elements are found in Swanson et al. (1999).

The regression analysis shows that the first two revisions $({}_{t}y_{t-2} - {}_{t-1}y_{t-2})$ and $({}_{t}y_{t-3} - {}_{t-1}y_{t-3})$ contain systematic elements but that $({}_{t}y_{t-4} - {}_{t-1}y_{t-4})$ does not. In this sense, ${}_{t}y_{t-3}$ is the first measure of the 'true' output level in time $t-3$ and in practice the measure of the output gap based on 'true' data that is available at time T , using the forecast-augmented technique based on our preferred multivariate model, is given by $x_t^{fm} = {}_{T}y_t - \widetilde{y}_t^m|\Omega_T$, $t = 1, ..., T-3$. Table 3 provides summary statistics relating to this series, and the corresponding real time measure obtained applying the procedure recursively over time, x_t^{rm}, for our data upto 2004q1 (i.e. for $T = 2004q4$). These figures show that the advantages of the forecast-augmentation remain, with a correlation between the real time measure and the final measure of 0.80 and agreement on booms and recessions in 83% of the sample.

The Table also presents corresponding statistics based on an 'adjusted' dataset in which revisions relating to data one year earlier (i.e. after two revisions) are assumed to be zero; i.e. ${}_{t}y_{t-4-k} = {}_{t-s}y_{t-4-k}$, $k = 0, 1, 2, ...$for all $s = 1, 2, ...$ Of course, this is a more severe assumption than that these revisions are unsystematic, as shown in the data. But the assumption effectively means that, in order to calculate the output gap at any point in time, one requires just the three most recent vintages of data (not every vintage of data)[20] which will be most useful for the practical implementation of the methods. The correlation coefficients in the Table show that this simplification has a relatively minor impact on the series: correlations between the unadjusted and adjusted series are 0.95 and 0.93 for x_t^{rm} and x_t^{fm} respectively (with agreement on booms and recessions in 91% and 93% of the sample). Given the practical advantages of the adjusted series in terms of their data requirements, these represent our preferred measures of the output gap and these series are presented in Figure 3 (x_t^{fm} representing the best measure of the output gap that we have available to us, given the information set in 2004q4, and x_t^{rm} represents the corresponding measures that would have been produced in real time).[21] This figure illustrates clearly the advantages of applying our methods for estimating the output gap

[20]Since the older vintages are assumed to be truncated versions of the most recent data; i.e. $Y_{T-3} = Y_{T,T-4}$, $Y_{T-4} = Y_{T,T-5}$, and so on.

[21]It is also the adjusted measures on which we concentrate in the section below.

which show a relatively close correspondence between the real time and final measures over the whole period.

4 Representing the Output Gap under Uncertainty

The analysis above shows that the uncertainty surrounding the output gap measure can be reduced through the appropriate use of forecast-augmented data, and that the forecasts are best calculated using a multivariate model that describes the measured output growth series and data revisions jointly. Nonetheless, it shows that the unreliability of the measures highlighted by OvN remains and, as argued earlier, it is therefore important that the uncertainties surrounding the measure are properly represented for decision-making purposes.

Figure 4 illustrates the order of magnitude of the uncertainties involved using the information set available at 2002q2. The end period was chosen to be 2002q2 to leave ten periods, to 2004q4, for the purpose of "out-of-sample" forecast evaluation. As we see from the solid line, analysis undertaken in 2002q2 would show the measured output gap rising above zero in early 1997, peaking in 2000q2 at close to 2%, before falling to -1.3% in 2001q3. These gap measures relate to periods when 'true' output data is available (i.e. data that will not be subsequently revised given our finding that revisions continue for only two periods) and so the uncertainty surrounding these figures arises from the estimation error in the underlying trend measure only. The 95% confidence intervals plotted in the Figure lie approximately ±0.3% around the point forecast in 1997, rising to ±0.7% in 2000q2 and ±1.9% in 2001q3 reflecting the uncertainties associated with the trend measure as we move towards the end of the sample.[22] As it turns out, the measured output gap falls further in 2001q4 before reversing direction in 2002q1 (to -1.7% and -1.1%). Although these measures are informed by data on y_{2001q4} and y_{2002q1} published in 2002q2, it is recognised that these figures will be revised in the coming

[22]These intervals are generated by the simulation methods discussed in Section 2.3 and relate to the stochastic and parameter uncertainty surrounding the measures. Abstracting from parameter uncertainty, by undertaking simulations taking into account stochastic uncertainty only, generates slightly tighter but very similar confidence intervals.

periods, and the gap measures are subject to the additional uncertainties associated with the forecasts of these revisions. The 95% confidence intervals widen to $\pm 2.4\%$ at this time therefore, and then widen still further to around $\pm 3\%$ over the following two years as the gap estimates rely more comprehensively on forecasts of out-of-sample output levels and trend levels. The (considerable) uncertainty associated with the forecast of the gap at this horizon reflects the unconditional variability in the output gap observed over the sample (as illustrated in Figure 3).

Figure 4 shows that the output gap measures based on information at 2002q2 remain negative but, having reached its minimum level in 2001q4, the gap measure tends to zero by 2003q4. As noted previously, the gap tends to zero by construction, but the speed of adjustment reflects the dynamics of the underlying model of Table 2 and the smoothing properties of the HP filter. The figure indicates that, starting from a position of around -1.5%, an investigator would have expected the impact of shocks to output to be fully worked through, and actual and trend output to reach their 'potential' level, within two years. As it turned out, however, the negative output gap persisted for some time. Figure 4 also plots, with the dashed line, the output gap measures obtained on the basis of the information available in 2004q4, $x_t^{fm}|\Omega_{2004q4}$. These show that 2001q4 was not in fact a turning point for the output gap, defining a turning point to occur when two periods of negative growth is followed by two periods of positive growth (or vice versa), and that the recovery began rather later, in 2003q1. Hence, the model performs relatively well, both in terms of the point forecast (the $x_t^{fm}|\Omega_{2004q4}$ all lie well within the confidence intervals obtained on the analysis of the 2002q2 data) and in terms of forecasting recession. But policy prescription based on the point forecasts alone would have misjudged the extent and duration of the recession over the following two year period.

The results of Figure 4, and associated confidence intervals, give a good indication of the *likely* size of the output gap within the sample, at the end of the sample and at various forecast horizons, but the information is not presented in the most useful way from a decision-maker's perspective. This information can be conveyed more usefully and more directly through the corresponding probability distribution functions showing $prob(x_{2002q2+h}^{fm}|\Omega_{2002q2} < c)$ for a range of critical values c at at various estimated horizons,

h. Figure 5 shows such density functions for $h = -3$, 0 and 4, again taking into account stochastic and parameter uncertainty. The functions shift to the right over time, reflecting the rising value of the point forecast, and show, for example, that the probability of an output gap less than zero is 0.93 in 2001q3, 0.72 in 2002q2 and 0.55 in 2003q2. The density functions are relatively steep in 2001q3, showing that there is relatively little uncertainty on the measure at that time, but become progressively flatter at 2002q2 and 2003q2 reflecting the accumulating uncertainty at the end-of-sample and into the forecasting horizons. The density functions of Figure 5 convey information on the output gap in precisely the form that is required by most decision-makers whose objectives are influenced by the expected output gap. Only in the special case where decision-makers face constraints that are linear in the output gap and pursue objectives that are quadratic in the output gap (the 'LQ' case) will attention focus on the point forecasts. More generally, decision-makers will require to maximise complex objective functions and the solution will require the entire density function describing the likely output gap outcomes. The unreliability of the output gap measures noted by OvN and characterised in the previous section does not mean that measures of the output gap cannot be used. Rather, it means that the decision problem needs to be fully-specified, possibly including statements on the risks involved in decisions as well as possibly non-linear objective functions, and evaluated in the light of the various states of nature that might be faced. The density functions of the type presented in Figure 5 provide precisely this information.

One important example of nonlinearities in objectives arises when decision-makers are concerned with the sign or rate of change of the output gap rather than the size. The emphasis in the media on 'booms' and 'recessions' reflects these ideas, indicating that may be a sharp difference in the consequences of a positive output gap compared to a negative one, irrespective of the size of the gap, or that attention should focus on the rate of change of the output gap, irrespective of its size or sign. Certainly there is a view that monetary authorities do not treat positive and negative output gaps symmetrically, reacting more strongly to the inflationary pressures associated with a positive output gap than they do to the recessionary pressures associated with a negative output gap. And there is an argument that policy-makers are concerned with whether conditions are improving or deteriorating, with the gap rising or falling (see Walsh (2003), for example).

If these arguments are important, then there are particular events that are relevant to decision-makers and direct statements on the likely occurrence of these events will be more helpful to decision-makers than the point forecasts of the output gap and associated confidence intervals of Figure 4. Specifically, if interest focuses on the *sign* of the gap, then $prob(x_t^{fm}|\Omega_T > 0)$ is a key statistic for decision-makers, and the evolution of this statistic over time might provide a useful indicator of inflationary pressure. This indicator is provided in Figure 6a, again based on the data available in $2002q2$, and plotting $prob(x_{2002q2+h}^{fm}|\Omega_{2002q2} > 0)$ for $h = -3, -2, -1, 0, 1, 2, ...$ The probability starts at less than 0.1 in $2001q3$, reflecting the fact that the point forecast of the gap starts low at -1.3% and that there is relatively little uncertainty because this figure relates to output data that is not going to be revised. Nevertheless, the probability of a positive output gap is still non-zero as there remains uncertainty about the underlying trend even at this stage. The probability remains low through to $2002q2$, but rises to 0.5 as the output level is forecast to converge on its trend.

Equally, if interest focuses on whether the gap is rising or falling, then decision-makers will focus on the likelihood of a turning point in the data. Taking it as given that the output gap peaked in $2000q2$, and defining a turning point as two consecutive periods of positive growth following two consecutive periods of negative growth (or vice versa), decision-makers would be interested in establishing the likelihood of an upturn in time t given by $prob(A)$ where $A = \left\{\left[x_{t-2}^{fm}|\Omega_T > x_{t-1}^{fm}|\Omega_T\right] \cap [x_{t-1}^{fm}|\Omega_T > x_t^{fm}|\Omega_T] \cap [x_t^{fm}|\Omega_T < x_{t+1}^{fm}|\Omega_T] \cap [x_{t+1}^{fm}|\Omega_T < x_{t+2}^{fm}|\Omega_T 0]\right\}$. This probability, based again on information available in $T = 2002q2$, is presented in Figure 6b. We noted earlier in the discussion of Figure 4 that, based on information available in $2002q2$, the point forecasts of the gap indicated that a turning point had been experienced in $2001q4$ although in fact, based on information available in $2004q4$, the recovery started later in $2003q1$. Figure 6b reflects the uncertainties surrounding the statements on turning points much more precisely and informatively, showing that the probability that an upturn was experienced in $2001q4$ was estimated to be 0.64.[23] Given that the probability exceeds 0.5, the investigator's best

[23] The probability was zero until $2001q2$ and 0.18 in $2001q3$.

guess would be that the upturn had occurred, therefore, but there is considerable uncertainty associated with this view. The probability that an upturn will have happened by 2003q1 (when the upturn did occur) is estimated to be 0.80.

Obviously these probability measures correspond to the point forecasts described above in Figure 4, and which turned out to be incorrect in terms of the size and duration of the recession. But, because they are expressed in probabilistic terms, rather than as point forecasts, they convey much more accurately the strength of conviction with which the forecasts are held and will be much more directly useful for those whose interest is in the occurrence of booms and recessions.

5 Conclusions

The analysis of this paper starts from the point that output gap measures are an essential element of many decisions but that they are measured with considerable uncertainty both because of the imprecision of the output data that they face at the time decisions have to be made and because of the difficulties in establishing a precise measure of trend output. We have argued that these uncertainties can be mitigated by modelling the output process alongside the revision process, making use of forecasts of future output levels, to obtain more precisely estimated measures of the gap for use in real time decision-making. But the uncertainties surrounding the measures, correctly identified as important by OvN and others, remain and are substantial. We have also argued, therefore, that the production of forecasts of probabilities of events involving the gap convey the information on the level of the gap and the uncertainties associated with this measure more precisely than the point forecasts and confidence intervals typically delivered by analysts. The cumulative distribution functions that we have presented, along with the estimated probabilities of positive gap measures and of turning points in the gap, provide a very informative and helpful means of representing the output gap data for use by decision-makers.

6 Appendix

Manipulation of (2.3) in the text provides the VECM representation

$$\begin{bmatrix} \Delta_t y_{t-1} \\ \Delta_t y_{t-2} \end{bmatrix} = \mathbf{a} + \mathbf{\Gamma}_0 \begin{bmatrix} {}_{t-1}y_{t-2} \\ {}_{t-1}y_{t-3} \end{bmatrix} + \sum_{j=1}^{p} \mathbf{\Gamma}_j \begin{bmatrix} \Delta_{t-j+1} y_{t-j} \\ \Delta_{t-j+1} y_{t-j-1} \end{bmatrix} + \begin{bmatrix} \epsilon_t \\ \xi_t \end{bmatrix} \tag{6.7}$$

where

$$\mathbf{\Gamma}_0 = -(\mathbf{I}_2 - \mathbf{\Phi}_1 - \mathbf{\Phi}_2 - \cdots - \mathbf{\Phi}_p) = -\mathbf{\Phi}(\mathbf{1})$$

$$\mathbf{\Gamma}_j = -(\mathbf{\Phi}_{j+1} + \mathbf{\Phi}_{j+2} + \cdots + \mathbf{\Phi}_p) \qquad j = 1, 2, ..., p-1$$

Given the form of the $\mathbf{\Phi}_i$ described in (2.3), it is easily shown that Γ_0 takes the form

$$\mathbf{\Gamma}_0 = \begin{bmatrix} -k_1 & k_1 \\ -k_2 & k_2 \end{bmatrix} = \begin{bmatrix} -k_1 \\ -k_2 \end{bmatrix} \begin{bmatrix} 1 & -1 \end{bmatrix} \tag{6.8}$$

where k_1 and k_2 are functions of the elements of the $\mathbf{B}_j$, $j = 0, 1, .., p-1$. Hence, the model at (2.1) can be written in a VECM form where $\mathbf{\Gamma}_0 = \boldsymbol{\alpha}\boldsymbol{\beta}'$and $\boldsymbol{\alpha}' = [-k_1, -k_2]$ contains the parameters determining the speed of adjustment to equilibrium and $\boldsymbol{\beta}' = [1, -1]$ is the cointegrating vector. The form of the cointegrating vector captures the assumption that revision errors are stationary through the inclusion of the error correction term $\boldsymbol{\beta}' \left[{}_{t-1}y_{t-2}, {}_{t-1}y_{t-3}\right]' = {}_{t-1}y_{t-2} - {}_{t-1}y_{t-3}$.

A final alternative for describing the model is the MA representation obtained through recursive substitution of (2.3):

$$\begin{bmatrix} \Delta_t y_{t-1} \\ \Delta_t y_{t-2} \end{bmatrix} = \mathbf{b} + \mathbf{C}(L) \begin{bmatrix} \epsilon_t \\ \xi_t \end{bmatrix} \tag{6.9}$$

where $\mathbf{b} = \mathbf{C}(1)\mathbf{a}$, $\mathbf{C}(L) = \sum_{j=0}^{\infty} \mathbf{C}_j(L)$, $\mathbf{C}_0 = \mathbf{I}_2$, $\mathbf{C}_1 = \mathbf{\Phi}_1 - \mathbf{I}_2$ and $\mathbf{C}_i = \sum_{j=0}^{p} \mathbf{C}_{i-j}\mathbf{\Phi}_j$, $i > 1$, $\mathbf{C}_i = 0$, $i < 0$. As is well known, following Engle and Granger (1987), the presence of a cointegrating relationship between the ${}_t y_{t-1}$ and ${}_{t-1}y_{t-2}$ imposes restrictions on the parameters of $\mathbf{C}(L)$; namely, $\boldsymbol{\beta}'\mathbf{C}(1) = 0$. Further, given that $\boldsymbol{\beta}' = [1, -1]$, this ensures that $\mathbf{C}(1)$ takes the form

$$\mathbf{C}(1) = \begin{bmatrix} k_3 & k_4 \\ k_3 & k_4 \end{bmatrix} \tag{6.10}$$

for scalars $k_3$and k_4.

References

Amato, J. D., and N. R. Swanson (2001) “The Real-Time Predictive Content of Money for Output”, *Journal of Monetary Economics*, 48, 3-24.

Brunner, A.D. (2000), “On the Derivation of Monetary Policy Shocks: Should we Throw the VAR Out With the Bath Water”, *Journal of Money Credit and Banking*, 32, 254-279.

Baxter and King (1999), ”Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series”, *Review of Economics and Statistics*, 81(4) 55-593.

Burman, J.P. (1980), ”Seasonal Adjustment by Signal Extraction”, *Journal of Royal Statistical Society* A, 143, 321-337.

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

Christiano, L.J., M. Eichenbaum and C.L. Evans (1999),“Monetary Policy Shocks: What Have We Learned and to What End?”, Chapter 2 in J.B Taylor and M. Woodford (eds.), *Handbook of Macroeconomics*, Volume 1A. North-Holland, Elsevier: Amsterdam.

Diebold, F. X., and G. D. Rudebusch (1991) “Forecasting Output with the Composite Leading Index: A Real-Time Analysis”, *Journal of the American Statistical Association*, 86, 603-610.

Garratt, A., K. Lee, M.H. Pesaran, and Y. Shin, Y. (2003) “Forecast Uncertainties in Macroeconometric Modelling: An Application to the UK Economy”, *Journal of American Statistical Association*, 98, 464, 829-838.

Garratt, A., K.C. Lee, M.H. Pesaran and Y. Shin, (2005), *A Structural Cointegrating Macroeconomic Model of the UK*, monograph in preparation, University of Cambridge.

Hodrick, R.J. and E.C. Prescott (1997), "Postwar U.S. Business Cycles: An Empirical Investigation", *Journal of Money, Credit and Banking*, 29, 116-130.

Howery, E.P. (1978) "The Use of Preliminary Data in Econometric Forecasting", *Review of Economic Statistics*, 6, 193-200.

Mise, E., T-H. Kim and P. Newbold (2005a), "On the Sub-Optimality of the Hodrick-Prescott Filter, *Journal of Macroeconmics*, 27(1), 53-67.

Mise, E., T-H. Kim and P. Newbold (2005b), "Correction of the Distortionary end-effect of the Hodrick-Prescott Filter: Application", *mimeo*, downloadable from http://www.le.ac.uk/economics/staff/em92.html.

Murray, C. and C.R. Nelson (2000), "The Uncertain Trend in US GDP", *Journal of Monetary Economics*, 46, 79-95.

Orphanides, A., R.D. Porter, D. Refschneider, R. Tetlow and F. Finan (2000), *Journal of Economics and Business*, 52, 117-141.

Orphanides, A. (2001) "Monetary Policy Rules Based on Real-Time Data", *American Economic Review*, 91, 964-985.

Orphanides, A. and S. van Norden (2002) "The Unreliability of Output-Gap Estimates in Real Time", *Review of Economics and Statistics*, 84(4), 569-583.

Orphanides, A. and S. van Norden (2003) "The Reliability of Inflation Forecasts Based on Output Gap Estimates in Real Time", *CIRANO Working Paper* 2003-01.

Orphanides, A., (1998), Monetary Policy Evaluation with Noisy Information, *Board of Governors of the Federal Reserve System, Working Paper*, October 1998.

Pappell, D.H. and R. Prodan (2004), "The Uncertain Unit Root in US Real GDP: Evidence with Restricted and Unrestricted Change", *Journal of Money Credit and Banking*, 36, 423-427.

Rotemberg, J.J. and M. Woodford (1997), "An Optimisation-Based Econometric Framework for the Evaluation of Monetary Policy", *NBER Macroeconomics Annual*, 297-345.

Rotemberg, J. and M. Woodford (1999), "Interest Rate Rules in an Estimated Sticky Price Model", pages , 57-119 in Taylor J.B. (ed.) *Monetary Policy Rules*, University of Chicago Press: Chicago.

St-Amant, P. and S. van Norden (1998) "Measurement of the Output Gap: A Discussion of Recent Research at the Bank of Canada", *Bank of Canada Technical Report* No. 79.

Svensson, L.E.O.,(2001), "Inflation Targeting: Should It Be Modelled as an Instrument Rule or a Target Rule?", *European Economic Review*, 46, 771-780.

Svensson, L.E.O. (2002), "What is Wrong with Taylor Rules ? Using Judgement in Monetary Policy through Targeting Rules", *Journal of Economic Literature*, 41, 426-477.

Swanson, N.R., E. Ghysels and M. Callan, (1999), "A Multivariate Time Series Analysis of the Data Revision Process for Industrial Production and the Composite Leading Indicator" in *Cointegration, Causality and Forecasting: Festschrift in Honor of Clive W.J. Granger*.

Walsh, C.E. (2003), "Speed Limit Policies: The Output Gap and Optimal Monetary Policy", *American Economic Review*, 93, 265-278.

Table 1: Univariate Output Gap Measures: 1965q3 – 2004q3

	x_t^{ro}	x_t^{qo}	x_t^{fo}	x_t^{ru}	x_t^{qu}	x_t^{fu}
Mean	-0.002	0.007	0.001	-0.005	-0.003	0.001
SD	0.018	0.020	0.016	0.013	0.012	0.016
Min	-0.066	-0.042	-0.047	-0.055	-0.031	-0.047
Max	0.038	0.052	0.038	0.016	0.024	0.038
x_t^{ro}	1.000	0.928	0.526	0.833	0.822	0.520
x_t^{qo}	*0.847*	1.000	0.628	0.813	0.912	0.623
x_t^{fo}	*0.631*	*0.694*	1.000	0.771	0.780	0.998
x_t^{ru}	*0.771*	*0.771*	*0.822*	1.000	0.928	0.769
x_t^{qu}	*0.796*	*0.796*	*0.796*	*0.885*	1.000	0.779
x_t^{fu}	*0.631*	*0.694*	*0.987*	*0.834*	*0.809*	1.000

Notes: Output gaps are denoted by x_t. The 'r', 'q' and 'f' superscripts refer to real-time, quasi-real time and final measures respectively, as described in the text; the 'o' and 'u' superscripts refer, respectively, to trend measures based on methods described in OvN and MKN, using an eighth-order univariate autoregression for forecasts, again described in the text. Summary statistics in the upper panel refer to the mean, standard deviation, minimum and maximum values respectively. Figures in the lower panel refer to correlation coefficients and, in italics, proportion of the sample for which there is agreement that the output gap is positive or negative.

Table 2: Model of Output Growth and Revisions: 1967q1 - 2004q3

Independent Variable	Dependent Variable			
	${}_ty_{t-1} -{}_{t-1} y_{t-2}$	${}_ty_{t-2} -{}_{t-1} y_{t-2}$	${}_ty_{t-3} -{}_{t-1} y_{t-3}$	${}_ty_{t-4} -{}_{t-1} y_{t-4}$
intercept	0.0034	0.0005	0.0002	0.0002
	(0.0009)	(0.0004)	(0.0004)	(0.0004)
${}_{t-1}y_{t-2} -{}_{t-2} y_{t-3}$	0.5749	0.0840	0.0258	0.0134
	(0.0934)	(0.0453)	(0.0378)	(0.0381)
${}_{t-2}y_{t-3} -{}_{t-3} y_{t-4}$	-0.0171	0.0134	0.0404	0.0462
	(0.0950)	(0.0460)	(0.0384)	(0.0388)
${}_{t-1}y_{t-3} -{}_{t-2} y_{t-3}$	-0.9579	-0.5251	-0.3075	-0.2750
	(0.3713)	(0.1800)	(0.1503)	(0.1516)
${}_{t-2}y_{t-4} -{}_{t-3} y_{t-4}$	0.7657	-0.1743	-0.2177	-0.1570
	(0.3808)	(0.1846)	(0.1541)	(0.1555)
${}_{t-1}y_{t-4} -{}_{t-2} y_{t-4}$	0.6577	0.4072	0.2510	0.2395
	(0.4295)	(0.2081)	(0.1738)	(0.1754)
${}_{t-2}y_{t-5} -{}_{t-3} y_{t-5}$	-0.7371	0.2587	0.1888	0.1177
	(0.4329)	(0.2098)	(0.1752)	(0.1768)
R^2	0.2817	0.0799	0.0430	0.0335
$\hat{\sigma}$	0.0081	0.0039	0.0033	0.0033
$\chi^2_{LM}(10)$	{0.04}	{0.62}	{0.44}	{0.39}
F_{SC}	{0.06}	{0.61}	{0.77}	{0.26}
F_{FF}	{0.42}	{0.94}	{0.75}	{0.81}
F_H	{0.80}	{0.13}	{0.25}	{0.44}
F_N	{1.00}	{1.00}	{1.00}	{1.00}

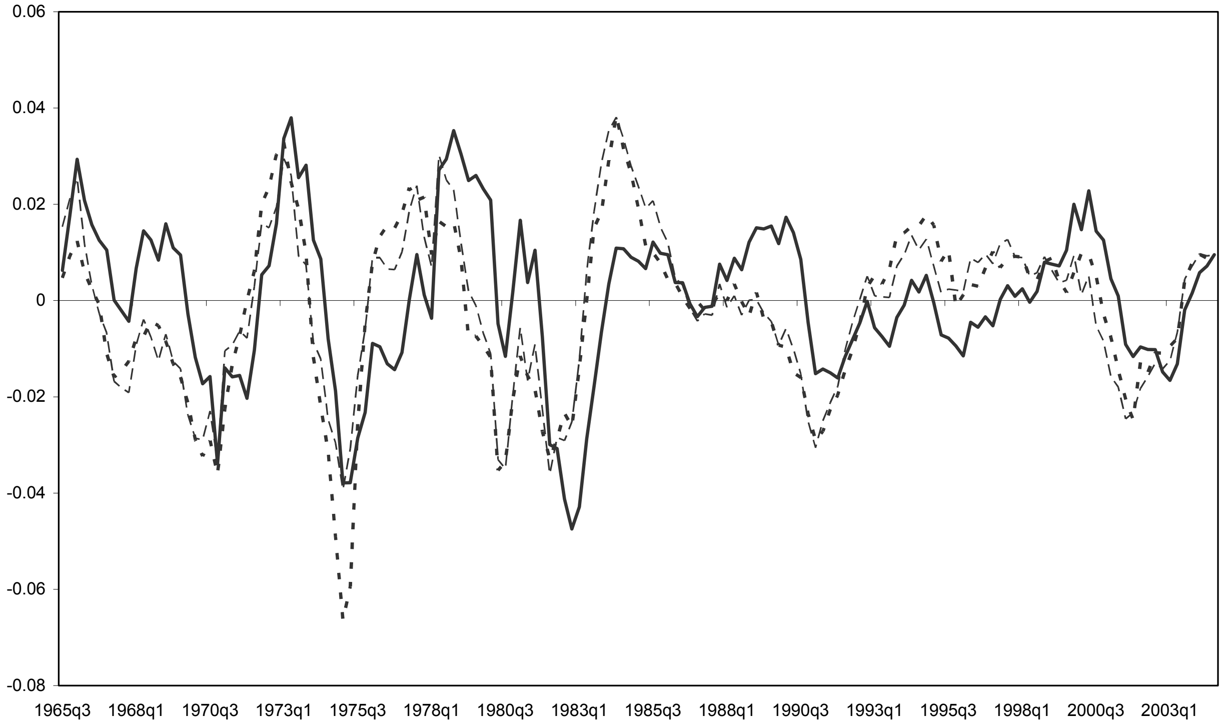
Notes: Standard errors are given in (.). R^2 is the squared multiple correlation coefficient, $\widehat{\sigma}$ the standard error of the regression and χ^2_{LM} a chi-squared test statistic (with 10 d.f.) for the exclusion of the third and fourth lags of the first three dependent variables and all four lags of the fourth from each of the regression equations as described in the text. The remaining diagnostics are p-values, in {.}, for F-test statistics for serial correlation (SC), functional form (FF), normality (N) and heteroscedasticity (H).

Table 3: Multivariate Output Gap Measures: 1970q1 – 2004q1

	x_t^{rm}	x_t^{fm}	adj x_t^{rm}	adj x_t^{fm}
Mean	-0.001	-0.001	-0.001	-0.001
SD	0.017	0.020	0.014	0.016
Min	-0.066	-0.068	-0.066	-0.047
Max	0.060	0.044	0.051	0.038
x_t^{rm}	1.000	0.796	0.950	0.749
x_t^{fm}	*0.832*	1.000	0.760	0.932
adj x_t^{rm}	*0.912*	*0.803*	1.000	0.748
adj x_t^{fm}	*0.854*	*0.934*	*0.839*	1.000

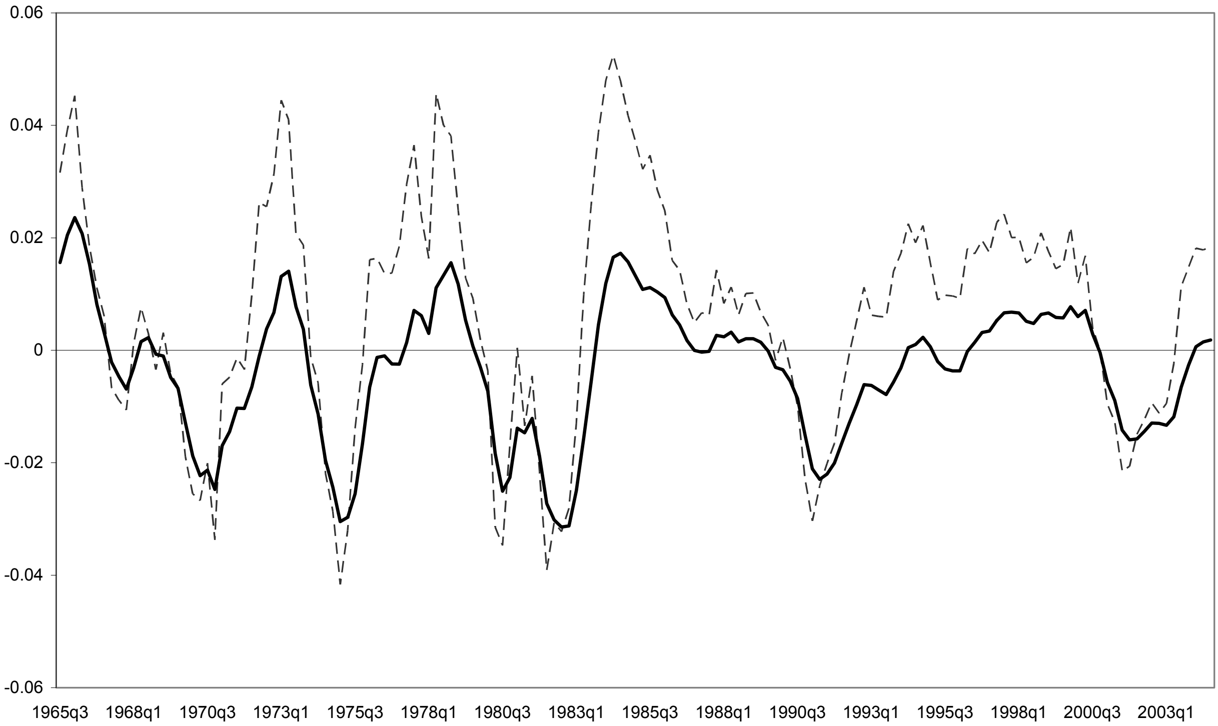
Notes: Output gaps are denoted by x_t. The 'r' and 'f' superscripts refer to real-time and final measures described in the text and the 'm' superscripts refers to trend measures based on the method described in MKN (using a multivariate model), again described in the text. The 'adjusted' figures are based on data assumed unrevised after four quarters. See also notes to Table 1.

Figure 1: Real Time, Quasi Real Time and Final Output Gap Measures following OvN, 1965q3-2004q3



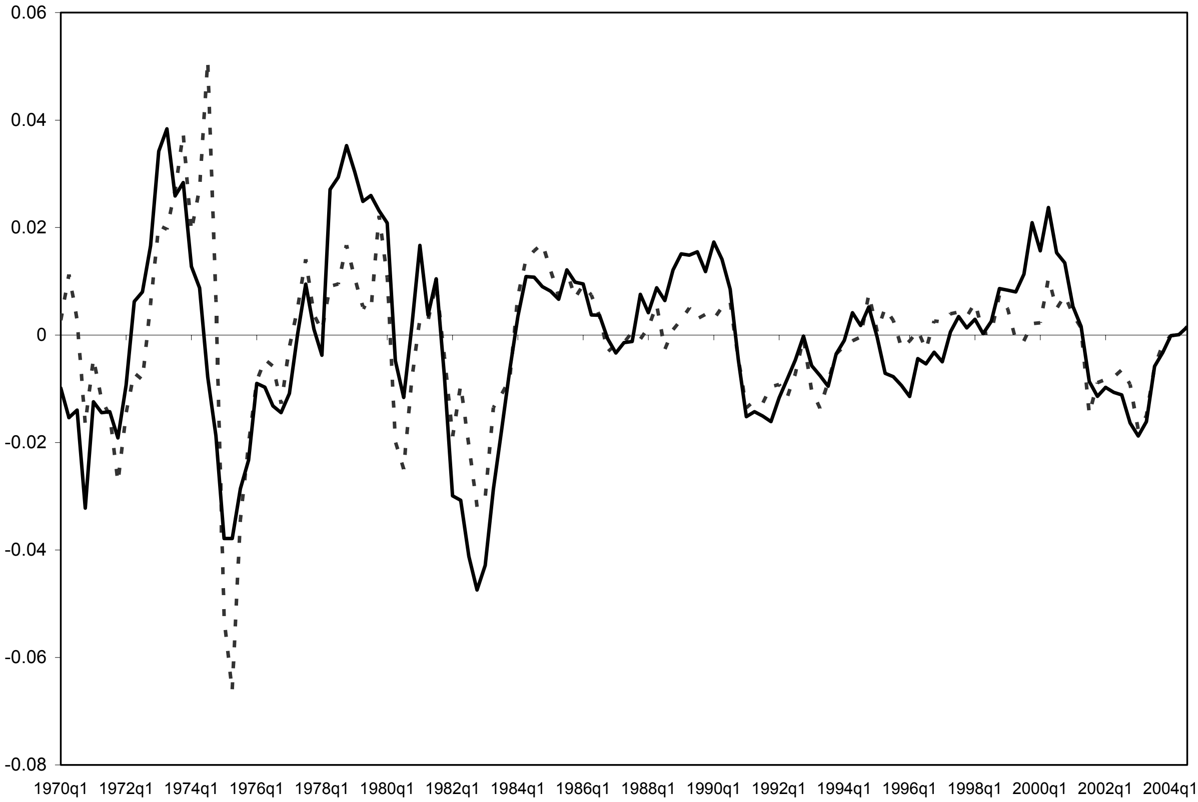
x_t^{ro} - - - x_t^{qo} — — — x_t^{fo} ———

Figure 2: Quasi Real Time Measures with and without Forecast-Augmentation, 1965q3-2004q3



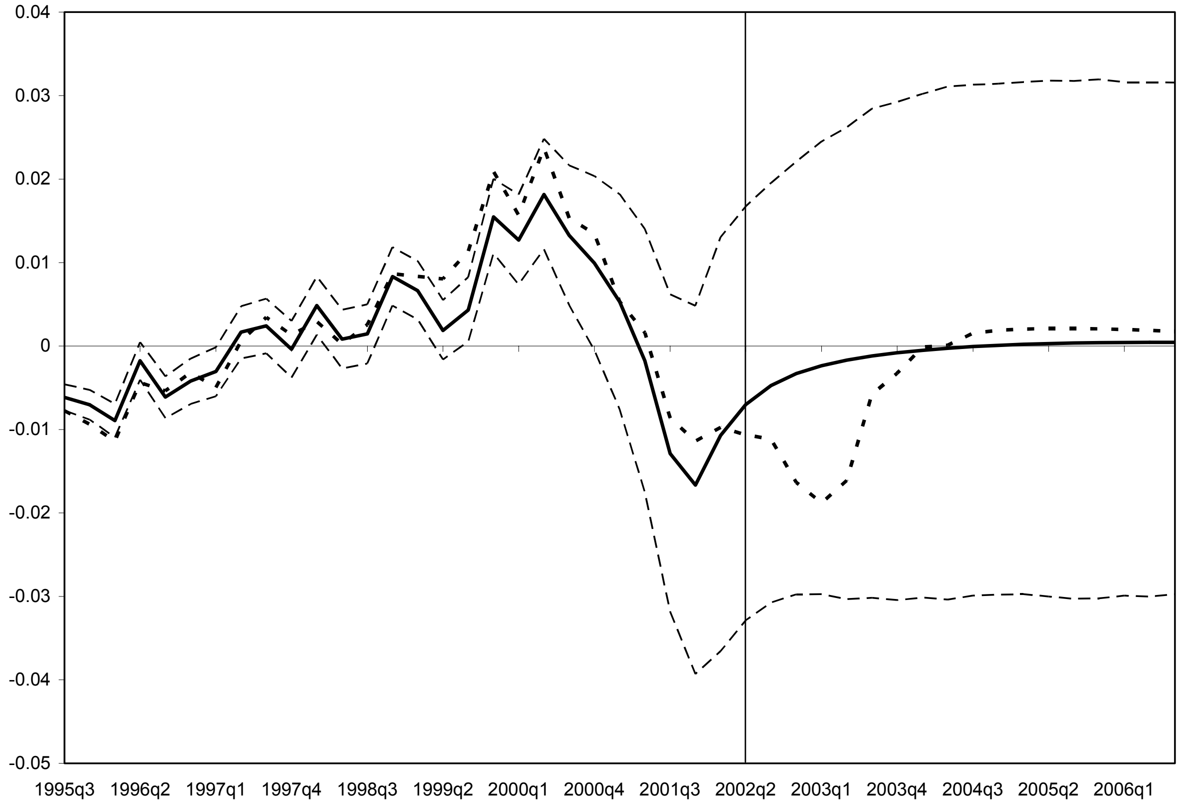
x_t^{qo} – – – x_t^{qu} ——

Figure 3: (Adjusted) Real Time and Final Output Gap Measures using Forecast-Augmentation based on Multivariate Model, 1970q1-2004q3



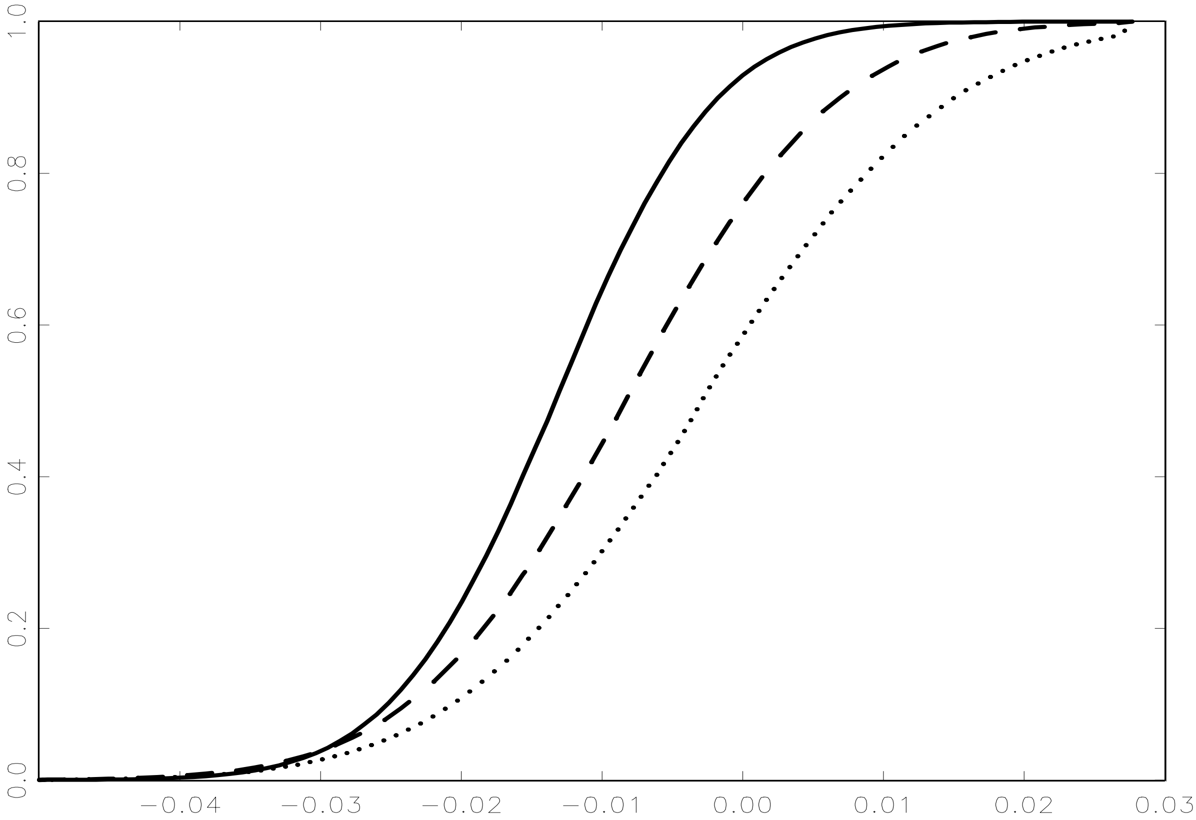
x_t^{rm} - - - x_t^{fm} ———

Figure 4: Output Gap Measures



$x_t^{fm}|\Omega_{2002q2}$ _____ (with 95% confidence intervals $---$) and $x_t^{fm}|\Omega_{2004q4}$.- - - -

Figure 5: Cumulative Density Functions for Forecast Horizons 2001q3 $(h=-3)$**, 2002q2** $(h=0)$ **and 2003q2** $(h=4)$.



2001q3 $(h=-3)$ ——— 2002q2 $(h=0)$ – – – 2003q2 $(h=4)$ · · · · ·

Figure 6a: Probability of a Positive Output Gap, $prob(x_t^{fm}|\Omega_{2002q2} > 0)$

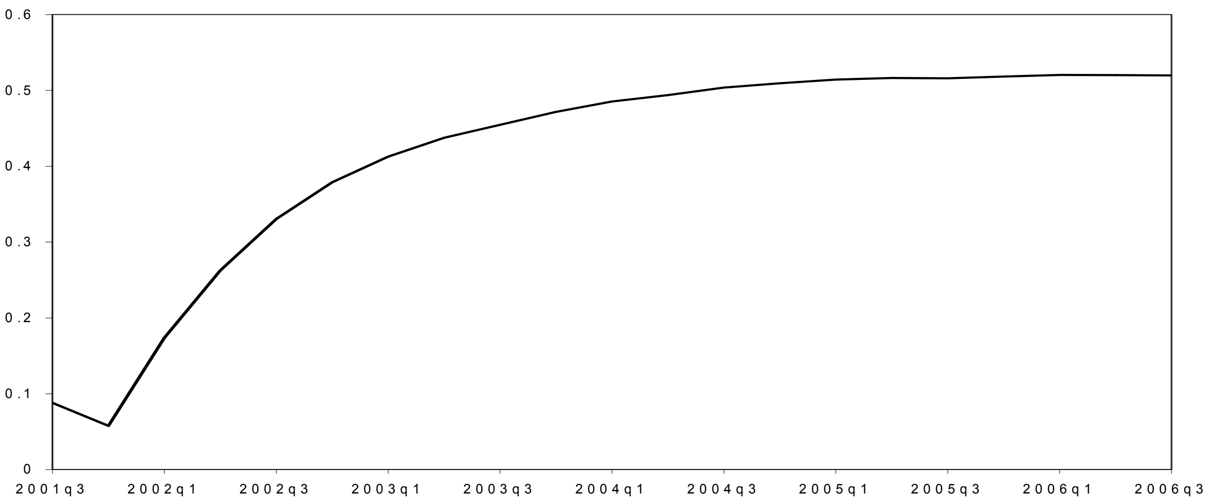


Figure 6b: Cumulative Probability of a Turning Point in the Output Gap, after 2000q2, based on Ω_{2002q2}