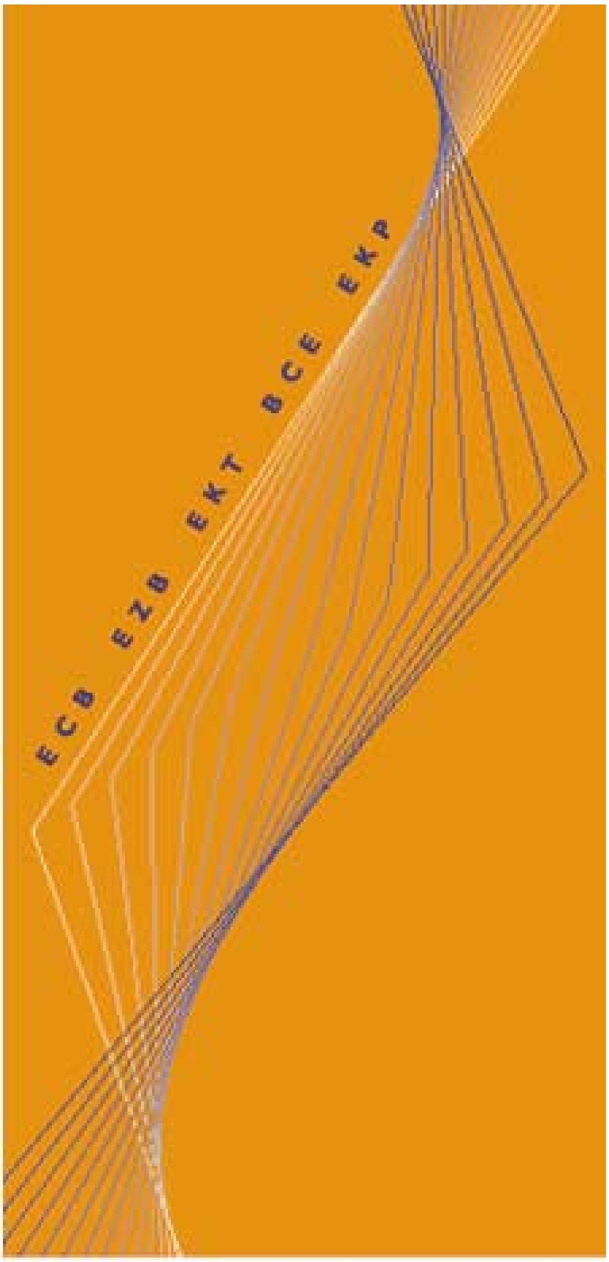
EUROPEAN CENTRAL BANK

WORKING PAPER SERIES



WORKING PAPER NO. 54

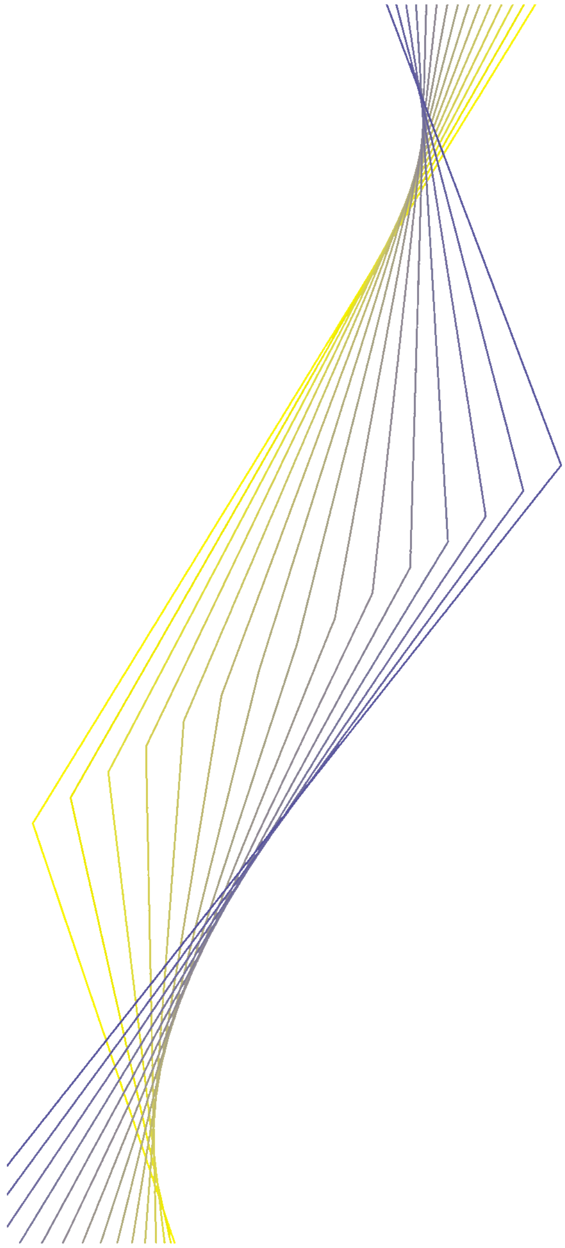
ASSESSMENT CRITERIA FOR OUTPUT GAP ESTIMATES

BY
GONZALO CAMBA-MENDEZ
AND DIEGO RODRIGUEZ-PALENZUELA

April 2001

EUROPEAN CENTRAL BANK

WORKING PAPER SERIES



WORKING PAPER NO. 54

ASSESSMENT CRITERIA FOR OUTPUT GAP ESTIMATES

BY
GONZALO CAMBA-MENDEZ
AND DIEGO RODRIGUEZ-
PALENZUELA*

April 2001

* European Central Bank. The opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the European Central Bank.
We are grateful to Günter Coenen, Gabriel Fagan and Gabriel Pérez Quirós for helpful comments.

© European Central Bank, 2001

Address	**Kaiserstrasse 29**
	D-60311 Frankfurt am Main
	Germany
Postal address	**Postfach 16 03 19**
	D-60066 Frankfurt am Main
	Germany
Telephone	**+49 69 1344 0**
Internet	**http://www.ecb.int**
Fax	**+49 69 1344 6000**
Telex	**411 144 ecb d**

All rights reserved.
Reproduction for educational and non-commercial purposes is permitted provided that the source is acknowledged.
The views expressed in this paper are those of the authors and do not necessarily reflect those of the European Central Bank.

ISSN 1561-0810

Contents

Abstract

This paper assesses the statitical reliability of different measures of the output gap for the Euro-11 area and the US using output, inflation and unemployment systems. In order to assess the reliability of an output gap estimate two criteria are adopted. Firstly, the estimate should have forecasting power over inflation. Secondly, the ex post statitical revisions of the output gap should not differ significantly from previously computed measures. As an additional check on reliability, we find out whether the estimate of the output gap is possitively correlated with standard measures of capacity utilization. We find that under our multivariate specification, unobservable components (UC) type models of the output gap show temporal consistency between sequential and final estimates and are consistent with known cyclical indicators. On the other hand, our UC models for the output gap have limited forecasting power for inflation, since they underperform an arbitrary autoregressive model.

1 Introduction

Recent research on optimal monetary policy has seen a renewed interest in the measurement of the output gap[1]. As argued for instance in Rudebusch (1999) and Ehrmann and Smets (2000), an important result from the recent literature is that the degree of uncertainty (or reliability) surrounding output gap measures has direct and strong implications on optimal monetary policy. In addition, Orphanides (1999) has argued against the advantages of monetary rules based on the estimation of the output gap. The basic idea behind his case is that real time measurements of the output gap are flawed since the monetary authority is unable to perfectly distinguish between cyclical movements in output and changes in its trend component. Infrequent but severe changes in output trend in particular will lead the monetary authority to incorrectly estimate the output gap and therefore to introduce an inflationary (or disinflationary, depending on the type of shift in trend) bias in its reaction function. Orphanides (1999) forcefully argues that an instance of such delay in learning about shifts in potential output on the part of the Fed occurred upon the reduction in US productivity growth since 1965. This delay led the Federal Reserve to underestimate the output gap for a number of years. Misguided monetary policy would then have been responsible for the surge in inflation of the nineteen seventies (the so-called *Great Inflation*).

In this context, the issue of how to produce real-time estimates of potential output and the output gap that are robust to structural changes, proves to be crucial for the design of monetary policy rules. As discussed in Orphanides (1999), Orphanides and van Norden (1999) and Lansing (2000), a minimal requirement for a real-time output gap measure to be robust is that it does not change much from its initial estimate when additional information becomes available over time.

The reliability of output gap estimates obtains when the three types of uncertainty that affect the estimation are small. A first source of uncertainty are *data revisions*, in particular the revisions of published GDP figures. A second source of uncertainty are *statistical revisions*: even if published GDP figures were never revised, the unobservable (and stochastic) nature of potential output and the output gap implies that they need to be estimated. Given an estimate of the output gap at time t, information on GDP in subsequent periods ($t+1$, $t+2$, ...) typically lead to an ex post *statistical* revision[2] of the estimated figure corresponding to period t.

Orphanides (1999) separates the effect of ex post revisions arising from published GDP figures revisions from that arising from statitical revisions through the estimation of the

[1] Optimal rules based on the ouptput gap as an indicator for monetary policy in the recent literature can be found, among others, in Ball (1997), Clarida, Gali, and Gertler (1999) and Svenson and Woodford (2000).

[2] Again, this is so even if the published GDP figures do not need to be revised.

output gap in *quasi real-time*. Within a specific modelling framework, and given the finally published data for the period, say, 1960-1999, the quasi real-time output gap at time t is the result of estimating the output gap with the finally published data for the subperiod 1960-t. The effect of published data revisions is observed by comparing the strictly real-time output gap estimate (which makes use only of the historic data available at time t instead of the finally published data available in 1999) with the qausi real-time estimate.

Interestingly Orphanides (1999) finds that for the US the effect of statitical revisions is about an order of magnitude more important than published data revisions. In addition, Orphanides and van Norden (1999) explore an additional source of uncertainty arising from the lack of agreement of the appropiate model to frame the output gap. Orphanides (1999) estimates a battery of alternative (univariate) models and obtains the quasi real-time and the real-time output gap estimates derived from each model. This evidence allows them to identify modelling strategies which yield relatively more reliable real-time and quasi real time estimates.

It should be noted however that Orphanides criticism of monetary policy based on rules including the output gap is admittedly conditioned on the difficulties in correctly measuring the output gap in real time, and not on such rules *per se*. A relevant counterfactual question that remains open in Orphanides (1999) is whether the inflationary bias in the Fed's monetary policy have still emerged under estimates of the output gap that were based on techniques available today. This question is analysed in Orphanides and van Norden (1999) using univariate methods.

In this paper we take stock on the issue of how to construct sequential (or *quasi real-time*) estimators of the output gap that are robust to statistical revisions (and in particular to shifts in the output trend) using multivariate methods instead, both for the US and the euro-area.

In the recent empirical literature, there are three leading approaches to measuring the output gap. Firstly, there are univariate time series models, of which the Hodrick-Prescott filter is the most prominent example. Secondly, we have the production function approach, relying on measures of capacity utilization, equilibrium unemployment rates and permanent shifts in technology. Thirdly, there are multivariate models built on well known economic relationships, which make use of time series techniques. In the present paper we strictly focus on this third approach.

The two widely used macroeconomic relationships which are typically taken into account in the third approach are Okun's law -which links fluctuations in unemployment with fluctuations in output- and the Philips curve -which links fluctuations in inflation with fluctuations in unemployment. An early reference in this spirit is that of Harvey, Henry, Peters, and Wren-Lewis (1986). Building on Okun's law they lay out an output-employment unobserved components system. Their model relies on a unique underlying force behind cyclical unemployment and cyclical output. Kuttner (1994) adopts a similar strategy, this time applied in a output-inflation unobserved components system. Closely related work in Quah and Vahey (1995) uses as well an output-inflation system in order to obtain a measure of core inflation.

Gerlach and Smets (1999) used an output inflation-system to obtain a measure of the output gap for the Euro wide area. Their paper stresses the relevance of the accuracy in estimating the aggregate EMU output gap for monetary policy-making at the European Central Bank (ECB).

Instead of making use of a two variable system, our approach is closely related to the three variable system proposed by Apel and Jansson (1999). They use an output, inflation and unemployment system to obtain measurements of the output gap, together with the NAIRU. This system incorporates simultaneously the building blocks of Okun's law and the Phillip's curve.

Given that there are different methods in the literature of calculating output gaps, which will yield different estimates, it is essential to have some criteria for judging them. Here we propose two criteria. The first one is forecasting power. The second criterion is the temporal consistency of the initial estimates and their subsequent revisions.

Orphanides and van Norden (1999) consider the measure of the output gap as reliable if it is *consistent* over time, where consistency means that the output gap estimated in real time does not differ significantly from the final output gap estimated with the full sample. The analysis of Orphanides and van Norden (1999) is done in the context of standard univariate methods. Their results for US data indicate that the ex post revisions of the output gap are of the same order of magnitude as the output gap itself. This negative result has potentially important implications on the use of the output gap in monetary policy rules [3].

The focus of our paper is to estimate the output gap, cyclical unemployment and cyclical inflation and to assess the statistical reliability of those measures. We consider that consistency *per se* is not sufficient for reliability of the output gap, since it is always possible to compute consistent measures of the output gap which are void of economic content. A sensible way to account for this is to assess its forecasting power over inflation (the link between inflation and a measure of real activity is illustrated in Calvo (1983) and Taylor

[3] This result from Orphanides and van Norden (1999) refers to what they define as the *Quasi-real* output-gap against the last estimated output gap. Our sequentially estimated output-gap corresponds precisely to their Quasi-real estimate. Orphanides and van Norden (1999) further test for the consistency of the *real time* estimate, which fully takes into account the issue of data revisions.

(1980) in the context of price and wage rigidities).

The paper is organised as follows. Section 2 describes the multivariate models. Section 3 describes the technical details of the strategy followed in assessing the forecasting performance, and in assessing the consistency of the estimates. Section 4 presents the empirical results, and section 5 concludes.

2 The Multivariate Models

The models we implement build on a three variable vector series $\boldsymbol{x}_t = (\pi_t, y_t, U_t)'$; where π_t denotes the inflation rate defined as the first differences of the log of the GDP deflator, y_t denotes the log of Real GDP, and U_t is the unemployment rate in levels. Modelling these three variables together allows us to obtain jointly measures of core inflation, the output gap and the NAIRU. Moreover the system encompasses the well known relationships of Okun's law and the Phillip's curve. In order to model $\boldsymbol{x}_t$ two alternative time series techniques are used: Unobserved Component (UC) models and Vector Autoregressive (VAR) models[4].

2.1 An Unobserved Component (UC) Model

The UC model is specified as follows:

$$\begin{aligned} \pi_t &= \mu_t^\pi + \Phi_\pi(L)\psi_t + u_{1t} \\ y_t &= \mu_t^y + \Phi_y(L)\psi_t + u_{2t} \\ U_t &= \mu_t^U + \alpha\psi_t + u_{3t} \end{aligned} \tag{1}$$

where μ_t^i for $i = \pi, y, U$ are three independent trend components; they refer respectively to core inflation, potential output and the NAIRU; $\Phi_j(L)$ for $j = \pi, y$ are polynomial lag operators; ψ_t is a cyclical component common to the three series, and finally u_{1t}, u_2 and u_3 are iid processes with standard deviations σ_{u1}, σ_{u2} and σ_{u3} respectively. The trends will be modelled as follows:

$$\begin{aligned} \mu_t^i &= \beta_{t-1}^i + \mu_{t-1}^i + \varepsilon_t^i \\ \beta_t^i &= \beta_{t-1}^i + \xi_t^i \end{aligned} \tag{2}$$

where ε_t^i and ξ_t^i are iid processes normally distributed with mean zero and standard deviation $\sigma_{\varepsilon i}$ and $\sigma_{\xi i}$ respectively.

[4]In all cases, although we allow for non-stationarity of first moments of the series in levels, we impose stationarity in second moments. Although this may be at odds with recent evidence for the US where a number of papers have shown a dumpening of the business cycle component of GDP, imposing constant second moments facilitates the comparison of our results with their natural benchmark, which is the results in Orphanides (1999), which does not introduce an adjustment in the second moments for the 1990s.

Three alternatives assumptions will be implemented: model UC1 incorporates to (1) the restriction $\sigma_{\xi i} = 0$: in this model the trend components will be random walks with drift. Model UC2 will be specified exactly as in (1): under model UC2 the trend components will be local linear trends. Finally, model UC3 will impose to (1) the restriction $\sigma_{\varepsilon i} = 0$; this restriction implies smooth trend components in model UC3.

Finally and under all three UC models, the cyclical component will be modelled as an autoregressive process:

$$\Phi_u(L)\psi_t = u_{4t}$$

where u_{4t} is a normally distributed error process with mean zero and standard deviation 1.

Note that while this type of model appears fully atheoretical, it is composed of well known relationships among macroeconomic aggregates. The first equation in (1) links deviations of inflation from core inflation to deviations of unemployment from the NAIRU, and is in effect a Phillips curve. The second equation links the output gap to cyclical unemployment, and builds on the stylised fact known as Okun's law.

The model can be written in State Space form and the Kalman filter is implemented to extract the state component. Given that there are parameters to be estimated, Maximum Likelihood estimation in combination with the Kalman filter must be used [5].

2.2 Reduced VAR Approach: Beveridge Nelson Decomposition

We also implement VAR models to extract the trend and cyclical components. Given the vector series of nonstationary variables $\boldsymbol{x}_t$ defined above, the Wold representation for the first difference is:

$$\Delta \boldsymbol{x}_t = \tilde{\boldsymbol{\mu}} + \boldsymbol{C}(L)\boldsymbol{\varepsilon_t} \tag{3}$$

where $\boldsymbol{\varepsilon_t}$ is and iid process with zero mean and covariance matrix $\boldsymbol{\Omega}$, and where $\boldsymbol{C}(L)$ is a polynomial lag operator defined as $\boldsymbol{C}(L) = \boldsymbol{I} + \boldsymbol{C}_1 L + \boldsymbol{C}_2 L + ...$ with $\sum_{j=0}^{\infty} j\,|\boldsymbol{C}| < \infty$. After inverting the expression above and rearranging, one can write

$$\boldsymbol{x}_t = \boldsymbol{x}_0 + \tilde{\boldsymbol{\mu}}t + \boldsymbol{C}(1)\sum_{j=0}^{t-1} \boldsymbol{\varepsilon}_{t-j} + \boldsymbol{C}^*(L)\boldsymbol{\varepsilon_t}$$

where $\boldsymbol{C}_j^* = -\sum_{i>j} \boldsymbol{C}_i$. This expression is the multivariate version of Beveridge and Nelson (1981) decomposition. The last component, $\boldsymbol{C}^*$ gives the cyclical component. Our vector series $\boldsymbol{x}_t$ does not display any common trends, and therefore $\boldsymbol{C}(1)$ is of full rank. Our previous UC model assumed that there was an unique and common cyclical component

[5] How to do this is well documented in Harvey (1993)

among the elements of the vectors series $\boldsymbol{x}_t$. In the context of the multivariate Beveridge Nelson decomposition this can be enforced by fixing the rank of the matrices in $\boldsymbol{C}^*(L)$ to be equal to one. Under rank one it is possible to rewrite $\boldsymbol{C}^*(L)\boldsymbol{\varepsilon}_t$ as $\tilde{\boldsymbol{C}}^*(L)\zeta_t$, where ζ_t is an *iid* random process resulting from a linear combination of $\boldsymbol{\varepsilon}_t$; and $\tilde{\boldsymbol{C}}^*(L)$ is a polynomial lag operator with matrices of order (3×1). In a VAR framework one can achieve this by means of Reduced Rank regression. Let the vector $\Delta\boldsymbol{x}_t$ be generated by the stationary VAR model of order p,

$$\Delta\boldsymbol{x}_t = \boldsymbol{\mu} + \sum_{i=1}^{p} \boldsymbol{B}_i \Delta\boldsymbol{x}_{t-i} + \boldsymbol{\varepsilon}_t, \tag{4}$$

where $\boldsymbol{B}_i$ is an (3×3) matrix, $i = 1, ..., p$, all roots of $\det(\boldsymbol{I}_3 - \sum_{i=1}^{p} \boldsymbol{B}_i z^i) = 0$ lie outside the unit circle and $\boldsymbol{\varepsilon}_t$ is a noise process defined as above.

A reduced rank VAR model takes the form:

$$\begin{aligned}\Delta\boldsymbol{x}_t &= \boldsymbol{\alpha}\left[\sum_{i=1}^{p} \boldsymbol{\beta}_i' \Delta\boldsymbol{x}_{t-i}\right] + \boldsymbol{\varepsilon}_t \\ &= \boldsymbol{\alpha}\boldsymbol{\beta}'\boldsymbol{z}_t + \boldsymbol{\epsilon}_t \end{aligned} \tag{5}$$

where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}_i$, $i = 1, ..., p$, are $(3, r^*)$ matrices, $r^* \leq 3$, $\boldsymbol{\beta} = (\boldsymbol{\beta}_1', ..., \boldsymbol{\beta}_p')'$ and $\boldsymbol{z}_t = (\Delta\boldsymbol{x}_{t-1}', ..., \Delta\boldsymbol{x}_{t-p}')'$, $t = 1, 2, ..., T$. In (5), it is assumed that the true ranks of the matrices $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}_i$, $i = 1, ..., p$, are identical and equal to r^* which is thus referred to as the rank of system (5). Velu, Reinsel, and Wichern (1986) provide details on estimation of the parameters of system (5), and methods to estimate the rank r^*. We review these methods in a technical appendix at the end of the paper.

Depending on the alternative assumptions on the rank of system (5), we define three VAR-based models that implement the Beveridge-Nelson decomposition. Model $BN1$ is estimated forcing the rank of the system to be equal to 1. Model $BN2$ estimates the rank from the data following Velu, Reinsel, and Wichern (1986). Finally, model $BN3$ imposes the full rank of the system.

2.3 Reduced VAR Approach: Blanchard and Quah Decomposition

The Wold representation of the vector $\Delta\boldsymbol{x}_t$ in (3) does not allow a structural interpretation of the error terms $\boldsymbol{\varepsilon}_t$, since the matrix $\boldsymbol{C}_0$ of contemporaneous effects is restricted to be the identity matrix. Letting $\mathbf{e}_t = \left(e_t^*, e_t^s, e_t^d\right)'$ be a vector of structural shocks, we can represent $\Delta\boldsymbol{x}_t$ as being generated by the structural moving average model[6]:

$$\Delta\boldsymbol{x}_t = \boldsymbol{A}_0\mathbf{e}_t + \boldsymbol{A}_1\mathbf{e}_{t-1} + \boldsymbol{A}_2\mathbf{e}_{t-2} + ...$$

[6]We have taken here the constant term μ as equal to zero for simplicity.

The relation between the structural shocks $\mathbf{e}_t$ and the reduced form shocks $\boldsymbol{\varepsilon}_t$ is given by: $\boldsymbol{\varepsilon}_t = \boldsymbol{A}_o \mathbf{e}_t$. Blanchard and Quah (1989) show how theory-driven restrictions in the matrix of long-run multipliers, $\boldsymbol{A}(1) = \sum_{i=0}^{\infty} \boldsymbol{A}_i$, of the structural model can be used to identify and estimate the matrix of contemporaneous effects $\boldsymbol{A}_0$.

Regarding the identification of the Structural VAR, we take the following approach. First, although it is clear that incorporating additional information (like interest rates) may facilitate the identification of structural shocks, extending the model to include additional variables would be inadequate for our purposes, since it would jeopardise the comparability of output gap estimates across different models. This is because it would be problematic to attribute the greater statistical reliability of an output gap estimate obtained with a given approach to its underlying assumptions or instead to the specific information used. We therefore restrict ourselves to make use only of the vector $\Delta \boldsymbol{x}$ in all implemented models.

Since it is not our purpose here to contribute to the structural VAR literature, we take a pragmatic approach and impose long run restrictions which are in line with existing studies on the dynamics of inflation, output and unemployment in a SVAR framework. Specifically, we introduce the long term relation between variables in $\Delta \boldsymbol{x}_t$ and a vector of structural shocks in the following form:

$$\begin{bmatrix} \Delta\pi \\ \Delta y \\ \Delta U \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_{11}(1) & 0 & 0 \\ \boldsymbol{A}_{21}(1) & \boldsymbol{A}_{22}(1) & 0 \\ \boldsymbol{A}_{31}(1) & \boldsymbol{A}_{32}(1) & \boldsymbol{A}_{33}(1) \end{bmatrix} \begin{bmatrix} e^{*} \\ e^{s} \\ e^{d} \end{bmatrix} \tag{6}$$

This identification is in line with the work of Bullard and Keating (1995) and Blanchard and Quah (1989). Bullard and Keating (1995) study the relation between inflation and output within a structural VAR framework. They represent inflation as determined in the long run by one structural shock (labeled *inflation shock*) and where output is affected in the long run by both structural shocks in the model. The system in Bullard and Keating (1995) is encompassed in our specification by the first two rows and columns in (6). A number of papers have studied output and unemployment systems with the SVAR methodology, most notably the orginal paper by Blanchard and Quah (1989). Blanchard and Quah (1989) estimates two alternative specifications: one in which output is determined in the long run by productivity shocks only and where unemployment is independent in the long run of such shocks, and a second specification where output is similarly determined by productivity shocks but where unemployment is affected both by productivity and a second shock which may be labelled *demand shock*. The latter specification has the adventage that it does not require ad hoc detrending for the unemployment series and the disadvantage (at least for the case of the US) that it implies the existence of hysterisis in unemployment. Our representation implied by (6) encompasses the second specification (with ΔU) in Blanchard and Quah (1989) as a particular case and is therefore grounded on the available literature.

Possibly the most adequate interpretation of the structure in (6) borrows form the search-theoretic models of the monetary exchange literature, in particular the models within this family where money is divisible and prices are endogenous (recent papers in this vein are Molico (1998) and Edward and Zhou (1998))[7]. Trade is fully decentralised in these type of models and characterised by agents' pairwise random matching. Agents hold fiat money in order to reduce transaction costs. Shocks to fundamentals that affect trade technology and transaction costs (like innovations in trade and communications technology) have a long run impact on inflation and output (and, allowing for imperfections in the labour market, also on unemployment). This type of real shock which affects transactions technology corresponds to the first shock (e^{*}) in (6). The second shock in 6, that affects output and unemployment, may be interpreted as a productivity or production technology shock, along the lines in Blanchard and Quah (1989). The third structural shock in (6) affects unemployment in the long run, but not output. Following Blanchard and Quah (1989), we label this shock a *demand* shock.

The Wold representation in (3) is estimated from the reduced rank estimation procedure described above. For the identification scheme in (6), three alternative decompositions, denoted as BQ1, BQ2 and BQ3 will be presented. The numbers correspond, respectively, with: *1*: reduced rank estimation for rank equal to 1; *2*: sequentially estimated rank; and *3*: full rank.

3 Model Assessment

3.1 Measures of Forecasting Performance

Forecasting performance is measured by the RMSFE over the forecasting period chosen. As it is common in the literature we use the results from the forecasting performance of a random walk with drift, the *naive* model as a benchmark and measure the performance of our models relative to this. The Theil statistics, i.e. the relative RMSFE of a particular model with respect to the naive model, are reported. When the Theil statistic is larger than 1, the forecasting performance of this model is worse than that of the *naive* model, and whenever it is smaller its forecasting performance is better. The forecasting performance of the models will also be assessed relative to the performance of an $AR(p)$ model of inflation in first differences. Notice that an $AR(p)$ is a much more powerful and more difficult to beat that a random model, particularly in view of the evidence presented in Stock and Watson (1999) for US data.

[7] For a recent overview to this strand of the literature see section 4 of Rupert, Schindler, Shevchenko, and Wright (2000)).

It is also helpful to use a formal statistical procedure to test the forecasting performance of the models against the benchmark models. This is provided by the test of predictive performance proposed by Diebold and Mariano (1995). The procedure is designed to test the null of equal predictive ability between two models by considering the mean of the differences of squared prediction errors of the two competing models. This mean, suitably normalised, has a standard normal distribution under the null. Details on the test can be found in the technical appendix.

3.2 Measures of Consistency of Output Gap Estimates

The statistical reliability of the output-gap measure will not be assessed exclusively by its forecasting performance. Orphanides and van Norden (1999) compute estimates of the US output gap from alternative univariate models using real-time data. Through a battery of descriptive statistics (no formal tests are provided) they find large disparities between the sequentially estimated measures of the output-gap compared to finally estimated measures. They show that this lack of consistency is mostly due to the unreliability of the models in estimating end of sample values, and to a lower extent to data revisions.

Define the recursively estimated output gap sequence as $\{\hat{g}_t^t\}_{t=1}^T$ and the finally estimated sequence as $\{\hat{g}_t^T\}_{t=1}^T$. The reliability of the output gap measure will be assessed by looking at whether the statistical properties of $\hat{g}_t^t$ and $\hat{g}_t^T$ are the same or not. In particular two formal tests of the following hypothesis will be conducted:

- Equality of the signs in both measures.
- Equality of the variances in both measures.

In order to test for the equivalence of signs of the sequentially estimated measure and the final one we use the predictive accuracy test of directional change developed by Pesaran and Timmermann (1992)[8].

To test for the identity of the standard errors of both measures of the output gap use is made of an F test. Under the null that the size of output-gap computed sequentially is the same as that of the finally estimated output gap, the following result must hold:

$$\frac{\sigma_{t,t}}{\sigma_{t,T}} \rightarrow F(n,n)$$

where $\sigma_{t,t}$ is the estimated standard error for the recursively estimated measure of the output-gap $\hat{g}_t^t$, and $\sigma_{t,T}$ the one corresponding to the finally estimated output-gap $\hat{g}_t^T$. n is the size of the sample over which the series are computed and it is equivalent to the forecasting period.

[8]See Appendix for further details

4 Empirical Results

The study has been conducted for Euro-11 area (henceforth Euro wide) and US seasonally adjusted data[9]. Data for the Euro wide area is aggregated in line with Fagan, Henry, and Mestre (2000). We estimate our models over the period 1970Q2 to 1979Q4 for Euro wide data and from 1960Q2 to 1979Q4 for US data. The remainder of the sample (from 1980Q1 to 1999Q1 for Euro wide data and from 1980Q1 to 1999Q1 for US data) provides us with periods over which we can test both forecasting performance and consistency of the output gap, and by extension consistency of cyclical inflation and cyclical unemployment.

All models are recursively estimated over the forecasting period. Out of sample forecasts from 1 to 6 quarters ahead are computed. For the VAR models, the number of lags is selected using the recursive BIC criterion, and it is recalculated at each stage of the out of sample forecasting process.

4.1 Euro wide Data Results

Forecasting Power. Table 1 displays the results with Euro wide data on the forecasting performance of all six models under study. The Theil statistics reveal that all VAR-based models have an advantage with respect to a naive random walk to forecast inflation and unemployment, but no advantage is shown in the case of output. These results are corroborated by the Diebold-Mariano (D-M) test comparing system forecasts to the univariate random walk.

Still using the *naive* model as a benchmark, UC models perform less well than VAR models, since the Theil statistic is below 1 for inflation and unemployment only when the forecasting horizon is sufficiently long. In the case of the UC3 model, even for inflation forecasting performance is below the random walk.

As for the $AR(p)$ model as benchmark, from table 1 is clear that none of the multivariate models matches the forecasting power of the $AR(p)$ processes (the only exception are models UC1 and UC2 for the case of unemployment). The RMSEs are in most of the cases above one. For inflation and for one quarter ahead forecasts the RMSE is larger than three. The worse forecasting performance of the multivariate methods, relative to the $AR(p)$, is also clear from the respective D-M test. It should be noted however that, for all VAR models and for model UC2, the RMSEs are, for horizons of four quarters or longer, at most 10% above parity.

[9]We are aware of the dangers of mis-specification when conducting a trend cycle decomposition in a series that has been previously adjusted (filtered). Notwithstanding, we have opted for the use of seasonally adjusted data for the following reasons: i) unadjusted series are not avaliable for all Euro area countries, ii) most previous related work has been conducted with seasonally adjusted series.

Among the unobservable components type models, the best performer in terms of forecasting inflation corresponds to the one which specifies the trends as local linear trends, i.e. UC2. If the focus is forecasting unemployment, the best of the UC models corresponds to that which specifies the trends as smooth trends, i.e. UC3.

For one-period-ahead forecasts, all alternative VAR models perform similarly. However, over longer horizons, VAR2 and VAR3 slightly overperform the VAR1 and they are clearly superior to UC models regarding inflation. Recall that VAR1 corresponds to a reduced rank VAR with rank fixed to 1, VAR2 corresponds to a reduced rank VAR with sequentially estimated rank, and VAR3 corresponds to a full rank VAR [10].

Consistency of Recursive versus Final Estimates. Results in table 2 show positive consistency results for all models. The concern raised by Orphanides and van Norden (1999) over the consistency of univariate measures is clearly less relevant in the case of the inflation, output and unemployment systems when applied to Euro-wide data. The smallest correlation between sequentially estimated and finally estimated cyclical inflation, output gap and cyclical unemployment reported in table 2 is 0.674 for UC models and 0.928 for VAR-based models.

The Pesaran and Timmermann (1992) test of predictive accuracy of the direction of change always rejects the null hypothesis of independence between the signs of the sequentially estimated cycles and the final cycles. Visual observation of these cycles in figure 1 for the alternative UC models, figure 2 for the BN models and figure 3 for the BQ models further illustrates this result.

The entry $+/-$ in table 2 denotes the proportion of cases where the signs of the final and recursive estimates of cyclical inflation, output and unemployment coincide. For all models and variables, there is coincidence in sign in more than 90% of the cases for VAR models and in more than 70% of the cases for UC models. For output under model UC2, the signs of the finally and sequentially estimated sequences coincide 82% of the time.

The only drawback on consistency refers to the disparities between the dimensions of the cyclical components of inflation, output and unemployment, as measured by the standard error. Table 2 displays the ratio, F, between the standard error of the sequentially estimated cycle and the last estimated cycle. For the case of output, the null hypothesis of F being one is not rejected for a 10% level of significance for models UC1 and UC2, but this null is rejected for model UC3 at the 10% level. For all VAR models the null of no difference between the standard error of the final and recursive estimates of cyclical output is never

[10]Notice from table 1 that results under models $BN2$ and $BN3$ are numerically identical. This is due to the fact that the endogenously estimated rank of the system under $BN2$ is found to coincide with the imposed rank of 3 under model $BN3$.

rejected at the 10%.

These overall positive results on consistency of quasi real-time estimates of the output gap should be contrasted with those of Orphanides and van Norden (1999) for univariate models with US data. They find that the order of magnitude of the revisions in the output gap measures is equal to the size of the output gap, i.e. F ratios larger than 2 or smaller than 0.5.[11] In terms of consistency, our *worst* results for Euro-wide data corresponds to an F ratio of 1.314 associated to the BQ3 model and to a ratio of 0.7 for the UC3 model. Our *best* results in this respect correspond to an F ratio of 0.81 for output under model UC1 and a ratio of 0.881 for model BN1.

Consistency with known cyclical indicators. In order to explore the extent to which the estimated cyclical components of output should be interpreted as the output gap in the usual sense, we show the correlation between the output gap estimates and the OECD's measure of capacity utilisation. Results for Euro-wide data are in table 3. Such correlations, for the sequentially estimated cycle (SEC) and for the last estimated cycle (LEC) (for two different time periods) are found to be positive for all UC models and negative for all VAR models (BN and BQ). For UC models the correlation between SEC and capacity utiliation is greater than between LEC and capacity utilization. The highest correlation is for the UC1 model (0.674 correlation between capacity utilization and the sequentially estimated cycle).

In summary, we find that all models have a forecasting power that is worse but not far apart (for horizons larger than one quarter) from the $AR(p)$ process. VAR models show particularly high consistency in the comparison of sequential and final estimates, but do not render an estimate of the output gap that is adequate from the point of view of interpretation: BN and BQ output gaps are *negatively* correlated to capacity utilization and are clearly less persistent than standard views on the business cycle. UC models on the other hand share three positive characteristics: they have a forecasting power comparable (if inferior) to that of univariate models, show adequate consistency properties (particularly models UC1 and UC2) and are positively correlated with capacity utilisation (particularly model UC1).

4.2 US Data Results

Forecasting Power. Table 4 presents the forecasting performance of the alternative models. As in the case with Euro-wide data, the UC type models are found to have almost no advantage with respect to a naive random walk to forecast inflation and output. This

[11] See table 1 in Orphanides and van Norden (1999)

remark should be qualified for the UC3 model that slightly overperforms the naive random walk (Theil statistic of 0.945) and for all UC models when the focus is forecasting unemployment. Notwithstanding, the Diebold-Mariano test for this model does not allow rejection of the null of smaller forecasting errors.

When the autoregressive model is used as a benchmark, similar results to the Euro-wide case are also obtained: with some exceptions, the multivariate methods underperform the $AR(p)$ model. For VAR models the RMSEs are close but higher to one (except for the case of unemployment). For UC models and in the case of forecasting inflation, RMSEs are in the average between 20 and 30% above parity for horizons larger than one quarter. For UC models but in the case of output and unemployment, models UC1 and UC2 are in general superior to the $AR(p)$ process.

The performance of the VAR-based forecasts for US data is slightly better than for Euro-wide data. In particular, the D-M test favors the VAR type models against the naive random walk for forecasts 1 and 2 periods ahead but the same does not occur when the comparison is with respect to an $AR(p)$. The best forecasting model among the VAR type model is the VAR1 (VAR of imposed reduced rank equal to 1). The Theil statistic is as low as 0.294 for one step ahead forecasts.

Consistency of Recursive versus Final Estimates. Table 5 shows the consistency results for US data. Again here, the UC type models perform less well than the VAR models. Model UC3 (smooth type trends) shows particularly poor performance: it displays the lowest correlation coefficients, ρ, and F statistics which reject the null of equality of sequentially and last estimated cyclical components at the 5% level of significance for output and unemployment.

Overall consistency of sequential and final estimates for all models except UC3 is found to be adequate from table 5. Correlation coefficients are in general high. For model UC1 the correlation is comparable to that of VAR models (0.952). Again with the exception of model UC3, the null of equality of the standard deviations of final and sequential estimates is never rejected at the 10% level. The Pesaran and Timmermann (1992) test of predictive accuracy rejects the null of independence of the signs even at the 1% level of significance. The proportion of cases with coincidence in sign, in the case of cyclical output estimates, is above 80% for models UC1 and UC2 and above 90% for VAR models.

The degree of the revisions for the UC type models is here again not in line with those reported by Orphanides and van Norden (1999). System estimation yields greater consistency for the sequential estimates of the output gap. The highest F ratio obtained for cyclical output is 1.028 (model BQ1) and the lowest F is 0.608 (UC3 model) followed by 0.793 for model UC2. If we compare these results with those from the univariate BN decomposition reported in Orphanides and van Norden (1999), we see that, excluding model UC3, the consistency of our multivariate decompositions perform clearly better than the univariate case: Orphanides and van Norden (1999) reported a relative standard error of 0.51 against 0.915 in our model UC3 and against 0.994 in our model BN1.

Consistency with known cyclical indicators. Finally, regarding the correlation between cyclical output estimates and capacity utilization (table 6), the results with US data are similar to those with Euro-wide data: correlations are high and positive for UC models, but negative for VAR models. Model UC1 has the highest correlations (0.774 for the sequentially estimated cycle in Period 1). As opposed to the Euro-wide data case, the correlation between SEC and capacity utilization is smaller than the correlation between the last estimated cycle and capacity utilization.

In summary, results are qualitatively similar in the US and Euro-wide cases. Models are overall consistent and have a forecasting power that is not very dissimilar, although inferior, to that of the $AR(p)$ model. Cyclical output measures based on VAR models are more consistent than UC models according to our definition, but are negatively correlated with independent measures of capacity utilization, what casts doubts about their interpretability as a measure of the output gap.

5 Conclusion

In this paper we take stock on the issue of the statistical reliability of output gap measures. The models that we implement include three variables: inflation, output and unemployment and they build upon two well known economic relationships: Okun's law and the Phillips curve.

Our definition of statistical reliability entails two properties. Firstly, forecasting power, i.e. the system should have a forecasting performance comparable to that of a random walk and, preferably, comparable to that of a univariate process with an arbitrary number of lags. Secondly, consistency, i.e. the sequential and the final estimates of the output gap should be similar, in a well-defined sense (Orphanides and van Norden (1999) first raised the issue of consistency in the context of univariate models. They warn against the large size of revisions in sequentially estimated output-gaps).

As an additional check, we explore the consistency of our quasi real-time estimates of the output gap with direct indicators of cyclical fluctuations: we show the correlation of output gap estimates with available series of capacity utilization based on independent information, since it is expected that adequate measures of the output gap should not be uncorrelated or negatively correlated with survey-based measures.

We estimate our models for US and Euro-wide data. Our main finding is that, both for the case of Euro-wide and US data, multivariate unobservable components models that specify the trend component of the series respectively as random walks (model UC1), or as local linear trends (model UC2), yield moderately satisfactory sequential estimates of the output gap.

Such estimates are relatively more adequate since they perform well relative to our alternative specifications, with respect to two criteria. First, they satisfy a minimum criterion for forecasting power, as they overperform a *naive* model of inflation (random walk with drift) and moreover their performance is similar -although inferior- to that of a univariate autoregressive process with an optimal number of lags. Second, they are consistent: subsequent revisions are not statistically significant. Last, and possibly less importantly, they show a positive and relatively high correlation with accepted measures of capacity utilization. VAR based estimates of the output gap on the other hand, although very consistent, do not render an interpretable measure of the output gap.

This evidence suggests that, in our context of system estimation, the concerns raised by Orphanides and van Norden (1999) on the bias in estimating the output gap in quasi real time (recall [12] that the great bulk of the revision arises from ex post statistical revisions rather than from revisions in published data) may be to some extent overdone. It should be noted on the other hand that our qualification to Orphanides and van Norden (1999) is relevant regarding only the component of the bias which is due to statistical shortcomings in separating trend and cycle and that we are mute on the issue of the limitations arising from revisions in published figures.

[12] See the discussion in the Introduction.

References

Apel, M., and P. Jansson (1999): "A Theory-Consistent System Approach for Estimating Potential Output and the NAIRU," *Economics Letters*, 64, 271–275.

Ball, L. (1997): "Efficient Rules for Monetary Policy," NBER Working Paper 5952.

Barlett, M. (1947): "Multivariate Analysis," *Journal of the Royal Statistical Society, Series B*, 9, 176–197.

Beveridge, S., and C. R. Nelson (1981): "A New Approach to the Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Atention to the measurement of the Business Cycle," *Journal of Monetary Economics*, 7, 151–174.

Blanchard, O. J., and D. Quah (1989): "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *American Economic Review*, 79, 655–673.

Bullard, J., and J. W. Keating (1995): "The long-run relationship between inflation and Output in Postwar Economies," *Journal of Monetary Economics*, 36, 477–496.

Calvo, G. A. (1983): "Staggered prices in a Utility Maximazing Framework," *Journal of Monetary Economics*, 12, 383–398.

Clarida, R., J. Gali, and M. Gertler (1999): "The Science of Monetary Policy," NBER Working Paper 6442.

Diebold, F. X., and R. S. Mariano (1995): "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, 13, 253–263.

Edward, G., and R. Zhou (1998): "A Rudimentary Model of Search with Divisible Money and Prices," *Journal of Economic Theory*, 81, 252–271.

Ehrmann, M., and F. Smets (2000): "Uncertain Potential Output: Implications for Monetary Policy," European Central Bank, mimeo, November 2000.

Fagan, G., J. Henry, and R. Mestre (2000): "An Area-Wide model (AWN) for the EU11," European Central Bank, Mimeo.

Gerlach, S., and F. Smets (1999): "Output Gaps and Monetary Policy in the EMU Area," *European Economic Review*, 43, 801–812.

Harvey, A. C. (1993): *Time Series Models*. Harvester and Wheatsheaf.

HARVEY, A. C., B. HENRY, S. PETERS, AND S. WREN-LEWIS (1986): "Stochastic Trends in Dynamic regression models: an application to the employment-output equation," *Economic Journal*, 96, 975–986.

KUTTNER, K. N. (1994): "Estimating potential output as a latent variable," *Journal of Business and Economic Statistics*, 12, 361–368.

LANSING, K. (2000): "Learning About a Shift in Trend Output: Implications for Monetary Policy and Inflation," mimeo, Federal Reserve Bank of San Francisco.

MOLICO, M. (1998): "The Distribution of Money and Prices in Search Equilibrium," mimeo, Department of Economics, University of Pennsylvania.

ORPHANIDES, A. (1999): "The Quest for Prosperity without Inflation," Sveriges Riksbank Working Paper Series, 93.

ORPHANIDES, A., AND S. VAN NORDEN (1999): "The Reliability of Output Gap Estimates in Real Time," Working Paper, Board of Governors of the Federal Reserve System, Washington.

PESARAN, M. H., AND S. POTTER (1997): "A Floor and Ceiling Model of US Output," *Journal of Economic Dynamics and Control*, 21, 661–696.

PESARAN, M. H., AND A. TIMMERMANN (1992): "A Simple Nonparametric Test of Predictive Performance," *Journal of Business and Economic Statistics*, 10(4), 461–465.

QUAH, D., AND S. P. VAHEY (1995): "Measuring Core Inflation," *Economic Journal*, 105, 1130–1144.

RUDEBUSCH, G. (1999): "Is the Fed too Timid? Monetary Policy in an Uncertain World," Federal Reserve Bank of San Francico Working paper, No. 99-05.

RUPERT, P., M. SCHINDLER, A. SHEVCHENKO, AND R. WRIGHT (2000): "The Search-Theoretic Approach to Monetary Economics: A Primer," mimeo, Department of Economics, University of Pennsylvania.

STOCK, J. H., AND M. W. WATSON (1999): "Forecasting Inflation," *Journal of Monetary Economics*, 44, 293–335.

SVENSON, L., AND M. WOODFORD (2000): "Indicator Variables for Optimal Policy," mimeo, Institute for International Economic Studies, Stockholm University.

Taylor, J. (1980): "Aggregate Dynamics and Staggered Contracts," *Journal of Political Economy*, 88, 1–23.

Velu, R. P., G. C. Reinsel, and D. W. Wichern (1986): "Reduced Rank Models for Multiple Time Series," *Biometrika*, 73, 105–118.

Technical Appendix

Reduced Rank Regression

Given the system rank r^*, Velu, Reinsel, and Wichern (1986) considered the reduced rank VAR(p) model (5) and suggested an estimation method for the parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ which may be shown to be quasi-maximum likelihood. Denote the sample second moment matrices by $\boldsymbol{S}_{XX} = T^{-1}\boldsymbol{XY}'$, $\boldsymbol{S}_{XZ} = \boldsymbol{S}'_{ZX}$, $\boldsymbol{S}_{XZ} = T^{-1}\boldsymbol{XZ}'$ and $\boldsymbol{S}_{ZZ} = T^{-1}\boldsymbol{ZZ}'$ where $\boldsymbol{X} = (\Delta\boldsymbol{x}_1, ..., \Delta\boldsymbol{x}_T)$ and $\boldsymbol{Z} = (\boldsymbol{z}_1, ..., \boldsymbol{z}_T)$. Hence, the covariance matrix of the unrestricted LS residuals from (4), $\boldsymbol{S}_{XX.Z} = \boldsymbol{S}_{XX} - \boldsymbol{S}_{XZ}\boldsymbol{S}_{ZZ}^{-1}\boldsymbol{S}_{ZX}$, is the unrestricted quasi-ML estimator of the error process variance matrix $\boldsymbol{\Omega}$. Additionally, let $\{\hat{\lambda}_i^2\}_{i=1}^m$, $\hat{\lambda}_1^2 \geq ... \geq \hat{\lambda}_m^2 \geq 0$, denote the ordered squared eigenvalues of the (m, m) matrix $\boldsymbol{S}_{XX.Z}^{-1/2}\boldsymbol{S}_{XZ}\boldsymbol{S}_{ZZ}^{-1}\boldsymbol{S}_{ZX}\boldsymbol{S}_{XX.Z}^{-1/2}$ with associated eigenvectors $\{\hat{\boldsymbol{\nu}}_i\}_{i=1}^m$ subject to the normalisation $\hat{\boldsymbol{\nu}}_i'\hat{\boldsymbol{\nu}}_j = 1$ if $i = j$ and 0 otherwise. Therefore, the quasi-ML estimators for $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ in (5) with system rank r^* minimize $tr\{\boldsymbol{S}_{XX.Z}^{-1/2}(\boldsymbol{X} - \boldsymbol{\alpha\beta}'\boldsymbol{Z})(\boldsymbol{X} - \boldsymbol{\alpha\beta}'\boldsymbol{Z})'\boldsymbol{S}_{XX.Z}^{-1/2}\}$ and are given by

$$\widehat{\boldsymbol{\alpha}} = \boldsymbol{S}_{XX.Z}^{1/2}\widehat{\boldsymbol{V}}, \quad \widehat{\boldsymbol{\beta}} = \boldsymbol{S}_{ZZ}^{-1}\boldsymbol{S}_{ZX}\boldsymbol{S}_{XX.Z}^{-1/2}\widehat{\boldsymbol{V}}$$

where $\widehat{\boldsymbol{V}} = (\hat{\boldsymbol{\nu}}_1, ..., \hat{\boldsymbol{\nu}}_{r^*})$.

The minimum acceptable rank can be established using the test suggested by Barlett (1947). Under the null hypothesis that the true rank of the system is r, or equivalently under the null of the eigenvalues $\lambda_j^2 = 0$ for $r < j \leq m$

$$T\sum_{j=r+1}^{m} ln\left(1 + \lambda_j^2\right) \sim \chi^2_{(m-r)(m-r)}$$

See Velu, Reinsel, and Wichern (1986) for more details.

Diebold Mariano (1995) test of Predictive Accuracy

Diebold and Mariano (1995) proposed a statistic to test the null hypothesis of equal predictive accuracy of two models. The test statistic is given by

$$S_{DM} = \frac{\bar{d}}{\sqrt{V(\bar{d})}} \xrightarrow{d} N(0,1), \quad V(\bar{d}) = N^{-1}\left(\hat{\gamma}_0 + 2\sum_{i=1}^{n-1}\hat{\gamma}_i\right)$$

where $\bar{d} = \frac{1}{N}\sum_{i=1}^{N}\hat{d}_i$, $\hat{d}_i = \hat{\eta}^2_{\mathrm{UC},i} - \hat{\eta}_i^2$, $i = 1, \ldots, N$, $\hat{\eta}_{\mathrm{UC},i}$ are the prediction errors from the UC model, $\hat{\eta}_i$ are the prediction errors from the naive model, N is the number of prediction errors used, $\hat{\gamma}_i$, $i = 0, 1, \ldots, n-1$ are the estimated autocovariances of the series of prediction error differences and n is the prediction horizon[13].

Pesaran Timmermann (1992) test of Predictive Accuracy

Pesaran and Timmermann (1992) proposed a test of directional change of forecasts which can be used to study the reliability of a sequentially estimated output gap figure versus the final estimate for the output gap. Given the recursively estimated output gap defined as r_t and the final estimate, f_t, define the indicator variables:

$$R_t = \begin{cases} 1 & \text{if } r_t > 0 \\ 0 & \text{otherwise} \end{cases} \quad F_t = \begin{cases} 1 & \text{if } f_t > 0 \\ 0 & \text{otherwise} \end{cases} \quad Z_t = \begin{cases} 1 & \text{if } r_t f_t > 0 \\ 0 & \text{otherwise} \end{cases}$$

[13]Note that for $n = 1$ only $\hat{\gamma}_0$ is used in the variance of $\bar{d}$.

Under the null hypothesis that the sign for both series is the same the following holds:

$$S_n = \frac{\hat{P} - \hat{P}_*}{\left[\hat{V}(\hat{P}) - \hat{V}(\hat{P}_*)\right]^{\frac{1}{2}}} \to N(0,1)$$

where $\hat{P} = n^{-1}\sum_{t=1}^{n} Z_t$, and $\hat{P}_* = \hat{P}_r\hat{P}_f + (1-\hat{P}_r)(1-\hat{P}_f)$ is the estimated probability of Z_t being 1 under the assumption of r_t and f_t being independently distributed; and where $\hat{P}_r$ is the estimated probability of r_t being 1 and $\hat{P}_f$ the estimated probability of f_t being 1. Finally, the estimator for the variances of $\hat{P}$ and $\hat{P}_*$ are defined as $\hat{V}(\hat{P})$ and $\hat{V}(\hat{P}_*)$ and are computed as:

$$\hat{V}(\hat{P}) = n^{-1}\hat{P}_*(1-\hat{P}_*)$$

$$\hat{V}(\hat{P}_*) = n^{-1}(2\hat{P}_f - 1)^2\hat{P}_r(1-\hat{P}_r) + n^{-1}(2\hat{P}_r - 1)^2\hat{P}_f(1-\hat{P}_f) + 4n^{-2}\hat{P}_f\hat{P}_r(1-\hat{P}_r)$$

Table 1

Forecasting Accuracy: Euro-wide Data. [a]

Model	lead	Inflation Theil	Inflation D-M	Inflation RMSE	Inflation D-M	Output Theil	Output D-M	Output RMSE	Output D-M	Unemployment Theil	Unemployment D-M	Unemployment RMSE	Unemployment D-M
UC1	1	1.278	0.988	3.227	1.000	1.130	0.860	1.084	0.866	0.787	0.059	1.112	0.976
	2	1.175	0.997	1.419	1.000	1.000	0.838	1.000	0.837	0.942	0.072	1.016	0.950
	3	1.064	0.790	1.227	1.000	1.000	0.839	1.000	0.840	0.943	0.082	1.012	0.947
	4	1.038	0.675	1.117	0.921	1.000	0.841	1.000	0.844	0.948	0.080	1.022	0.900
	5	0.976	0.418	1.070	0.776	1.001	0.844	1.000	0.848	0.951	0.085	1.027	0.898
	6	0.936	0.325	1.016	0.552	1.001	0.847	1.000	0.852	0.956	0.093	1.026	0.907
BN1	1	0.652	0.000	1.647	0.999	1.142	0.845	1.095	0.998	0.726	0.021	1.026	0.706
	2	0.964	NaN	1.163	1.000	1.000	0.835	1.000	0.835	0.941	0.053	1.015	0.871
	3	0.942	0.162	1.087	NaN	1.000	0.836	1.000	0.836	0.946	0.064	1.016	0.881
	4	0.988	0.379	1.063	0.999	1.001	0.838	1.000	0.837	0.945	0.074	1.018	0.883
	5	0.974	0.370	1.069	0.921	1.001	0.840	1.000	0.839	0.946	0.078	1.022	0.882
	6	0.972	0.369	1.056	0.768	1.001	0.842	1.000	0.840	0.951	0.081	1.021	0.885
UC2	1	1.209	0.933	3.051	1.000	1.129	0.872	1.082	0.879	0.776	0.030	1.096	0.958
	2	1.119	0.920	1.351	0.997	1.000	0.841	1.000	0.839	0.919	0.084	0.991	0.220
	3	1.015	0.575	1.171	1.000	1.000	0.839	1.000	0.841	0.903	0.100	0.970	0.161
	4	1.004	0.519	1.081	0.811	1.001	0.841	1.000	0.843	0.884	0.109	0.953	0.153
	5	0.952	0.352	1.045	0.664	1.001	0.843	1.000	0.845	0.868	0.116	0.938	0.153
	6	0.924	0.305	1.003	0.509	1.001	0.845	1.000	0.847	0.861	0.121	0.924	0.153
BN2	1	0.635	0.000	1.604	0.996	1.178	0.840	1.129	0.945	0.752	0.028	1.063	NaN
	2	0.874	NaN	1.055	1.000	1.000	0.837	1.000	0.836	0.936	0.058	1.010	0.871
	3	0.892	0.000	1.029	NaN	1.000	0.836	1.000	0.836	0.939	0.066	1.008	0.850
	4	0.947	NaN	1.019	NaN	1.000	0.838	1.000	0.837	0.936	0.073	1.009	0.832
	5	0.943	0.162	1.034	NaN	1.001	0.840	1.000	0.839	0.933	0.078	1.008	0.825
	6	0.936	0.156	1.017	0.618	1.001	0.842	1.000	0.840	0.938	0.081	1.007	0.826
UC3	1	1.346	0.875	3.396	0.983	1.012	0.537	0.970	0.329	0.831	0.063	1.174	0.998
	2	1.359	0.889	1.640	0.953	1.000	0.832	1.000	0.832	0.907	0.112	0.978	0.246
	3	1.301	0.839	1.501	0.905	1.000	0.838	1.000	0.851	0.897	0.133	0.963	0.228
	4	1.287	0.814	1.385	0.858	1.000	0.842	1.000	0.854	0.876	0.147	0.943	0.220
	5	1.269	0.792	1.392	0.847	1.000	0.844	1.000	0.852	0.856	0.155	0.925	0.215
	6	1.299	0.805	1.411	0.848	1.001	0.847	1.000	0.854	0.855	0.163	0.918	0.218
BN3	1	0.635	0.000	1.604	0.996	1.178	0.840	1.129	0.945	0.752	0.028	1.063	NaN
	2	0.874	NaN	1.055	1.000	1.000	0.837	1.000	0.836	0.936	0.058	1.010	0.871
	3	0.892	0.000	1.029	NaN	1.000	0.836	1.000	0.836	0.939	0.066	1.008	0.850
	4	0.947	NaN	1.019	NaN	1.000	0.838	1.000	0.837	0.936	0.073	1.009	0.832
	5	0.943	0.162	1.034	NaN	1.001	0.840	1.000	0.839	0.933	0.078	1.008	0.825
	6	0.936	0.156	1.017	0.618	1.001	0.842	1.000	0.840	0.938	0.081	1.007	0.826

[a] The forecasting performance of all models has been tested against that of a naive random walk with drift and that of an AR(p) process. Theil refers to the Theil statistic with respect to the naive random walk and RMSE is the relative root mean square error of the model against that of the AR(p) process; D-M refers to the Diebold and Mariano (1995) test (values displayed are probability values).

Table 2

Consistency of Sequential Cycles Estimates. Euro-wide Data[a]

	Inflation	Output	Unempl.	Inflation	Output	Unempl.	Inflation	Output	Unempl.
		UC1			UC2			UC3	
ρ	0.773	0.799	0.740	0.706	0.818	0.844	0.749	0.718	0.674
F	0.867	0.810	0.473	0.706	0.789	0.703	0.784	0.700	0.877
pv_F	0.732	0.820	0.999	0.934	0.849	0.937	0.854	0.939	0.715
$sign$	6.130	4.586	6.450	5.087	6.678	5.562	5.712	4.819	4.854
pv_s	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$+/-$	0.855	0.763	0.868	0.789	0.882	0.816	0.829	0.776	0.776
		BN1			BN2			BN3	
ρ	0.957	0.951	0.924	0.969	0.978	0.959	0.969	0.978	0.959
F	1.053	0.881	0.600	1.313	1.277	0.743	1.313	1.277	0.743
pv_F	0.411	0.709	0.986	0.119	0.144	0.902	0.119	0.144	0.902
$sign$	7.842	7.842	7.842	8.324	7.293	7.688	8.324	7.293	7.688
pv_s	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$+/-$	0.947	0.947	0.947	0.974	0.908	0.934	0.974	0.908	0.934
		BQ1			BQ2			BQ3	
ρ	0.970	0.928	-	0.980	0.929	-	0.980	0.929	-
F	1.033	0.814	-	1.234	1.314	-	1.234	1.314	-
pv_F	0.444	0.814	-	0.181	0.118	-	0.181	0.118	-
$sign$	8.318	7.048	-	8.109	6.494	-	8.109	6.494	-
pv_s	1.000	1.000	-	1.000	1.000	-	1.000	1.000	-
$+/-$	0.974	0.908	-	0.961	0.868	-	0.961	0.868	-

[a]Results in the table refer to the reliability of the sequentially estimated cycle (SEC) against the last estimated cycle (LEC). ρ gives the correlation coefficient between SEC and LEC. F gives the ratio between the standard deviation of SEC and LEC. This ratio is distributed as an $F(n,n)$ under the null of both standard errors being equal. pv_F gives the level of significance of this F statistic. $sign$ performs the Pesaran and Potter (1997) test of directional accuracy betweem SEC and LEC, and pv_s gives the probability value. Finally $+/-$ gives the proportion of cases where the recursive and final series have the same sign.

Table 3

Correlation between Capacity Utilisation and Output Gap: Euro-wide Data

Model		Period 1[a]	Period 2
UC1	*SEC*	0.674	-
	LEC	0.550	0.508
BN1	*SEC*	-0.644	-
	LEC	-0.705	-0.648
BQ1	*SEC*	-0.588	-
	LEC	-0.541	-0.514
UC2	*SEC*	0.659	-
	LEC	0.391	0.376
BN2	*SEC*	-0.724	-
	LEC	-0.732	-0.649
BQ2	*SEC*	-0.742	-
	LEC	-0.641	-0.603
UC3	*SEC*	0.557	-
	LEC	0.445	0.451
BN3	*SEC*	-0.724	-
	LEC	-0.732	-0.649
BQ3	*SEC*	-0.742	-
	LEC	-0.641	-0.603

[a] Period 1 refers to the period 1980Q1 to 1998Q4 and Period 2 to the period 1976Q1 to 1998Q4. (SEC) denotes the sequentially estimated cycle and (LEC) the last estimated cycle.

Table 4

Forecasting Accuracy: US Data [a]

		Inflation				Output				Unemployment			
Model	lead	Theil	D-M	RMSE	D-M	Theil	D-M	RMSE	D-M	Theil	D-M	RMSE	D-M
UC1	1	1.382	0.806	4.981	0.951	1.040	0.592	1.079	0.806	0.663	0.010	0.992	0.394
	2	1.337	0.833	1.733	0.930	1.001	0.827	0.999	0.165	0.774	0.015	1.073	0.995
	3	1.425	0.876	1.580	0.907	1.000	0.819	0.999	0.163	0.873	0.116	1.088	0.968
	4	1.268	0.799	1.515	0.890	1.000	0.078	0.999	0.162	0.930	0.217	1.131	0.951
	5	1.193	0.739	1.382	0.862	1.000	0.146	0.998	0.160	0.999	0.494	1.151	0.959
	6	1.193	0.764	1.309	0.845	0.999	0.154	0.998	0.159	1.047	0.718	1.173	0.956
BN1	1	0.294	0.000	1.061	0.610	1.097	0.674	1.137	0.999	0.661	0.005	0.989	0.285
	2	0.842	0.055	1.091	0.942	1.002	0.831	1.000	0.186	0.723	0.002	1.001	0.563
	3	1.028	0.606	1.140	0.985	1.001	0.832	1.000	0.173	0.798	0.008	0.995	0.333
	4	0.892	0.196	1.066	0.784	1.001	0.834	1.000	0.872	0.823	0.000	1.002	0.712
	5	0.846	0.121	0.980	0.396	1.001	0.837	1.000	0.987	0.867	NaN	0.999	0.469
	6	0.912	0.184	1.001	0.506	1.001	0.840	1.000	0.865	0.893	NaN	1.000	0.510
UC2	1	1.246	0.728	4.491	0.939	1.040	0.593	1.078	0.842	0.711	0.007	1.064	NaN
	2	1.239	0.788	1.606	0.915	1.002	0.829	1.000	0.166	0.734	0.004	1.017	0.764
	3	1.283	0.843	1.423	0.886	1.001	0.828	0.999	0.164	0.794	0.014	0.989	0.373
	4	1.154	0.732	1.378	0.869	1.001	0.828	0.999	0.164	0.813	0.005	0.989	0.378
	5	1.096	0.654	1.269	0.832	1.001	0.833	0.999	0.161	0.854	0.001	0.985	0.356
	6	1.094	0.677	1.201	0.803	1.000	0.835	0.999	0.160	0.876	0.003	0.981	0.333
BN2	1	0.488	0.000	1.760	0.986	1.090	0.662	1.131	0.998	0.663	0.005	0.993	NaN
	2	0.827	0.034	1.073	0.904	1.002	0.831	1.000	0.190	0.724	0.002	1.003	0.608
	3	0.993	0.473	1.101	0.929	1.001	0.832	1.000	0.174	0.796	0.008	0.992	0.280
	4	0.906	0.205	1.082	0.855	1.001	0.833	1.000	0.863	0.821	0.000	0.999	NaN
	5	0.870	0.123	1.008	0.563	1.001	0.837	1.000	0.924	0.866	NaN	0.998	NaN
	6	0.918	0.177	1.008	0.563	1.001	0.840	1.000	0.861	0.893	NaN	1.000	0.507
UC3	1	0.945	0.263	3.405	1.000	1.160	0.828	1.204	0.966	0.949	0.132	1.421	0.963
	2	0.993	0.478	1.287	0.975	1.003	0.833	1.001	0.843	0.877	0.000	1.215	0.980
	3	1.103	0.725	1.223	0.911	1.002	0.835	1.000	0.847	0.967	0.214	1.206	0.986
	4	1.061	0.629	1.268	0.915	1.002	0.837	1.001	0.844	0.998	0.489	1.214	0.955
	5	1.062	0.619	1.231	0.897	1.002	0.839	1.001	0.843	1.080	0.801	1.245	0.957
	6	1.139	0.759	1.251	0.896	1.002	0.841	1.001	0.844	1.135	0.833	1.271	0.944
BN3	1	0.488	0.000	1.760	0.986	1.090	0.662	1.131	0.998	0.663	0.005	0.993	NaN
	2	0.827	0.034	1.073	0.904	1.002	0.831	1.000	0.190	0.724	0.002	1.003	0.608
	3	0.993	0.473	1.101	0.929	1.001	0.832	1.000	0.174	0.796	0.008	0.992	0.280
	4	0.906	0.205	1.082	0.855	1.001	0.833	1.000	0.863	0.821	0.000	0.999	NaN
	5	0.870	0.123	1.008	0.563	1.001	0.837	1.000	0.924	0.866	NaN	0.998	NaN
	6	0.918	0.177	1.008	0.563	1.001	0.840	1.000	0.861	0.893	NaN	1.000	0.507

[a] The forecasting performance of all models has been tested against that of a naive random walk with drift and that of an AR(p) process. Theil refers to the Theil statistic with respect to the naive random walk and RMSE is the relative root mean square error of the model against that of the AR(p) process; D-M refers to the Diebold and Mariano (1995) test (values displayed are probability values).

Table 5

Consistency of Sequential Cycles Estimates: US Data [a]

	Inflation	Output	Unempl.	Inflation	Output	Unempl.	Inflation	Output	Unempl.
		UC1			UC2			UC3	
ρ	0.631	0.952	0.933	0.635	0.873	0.869	0.420	0.506	0.401
F	1.210	0.915	0.895	1.144	0.793	0.807	0.763	0.608	0.607
pv_F	0.204	0.651	0.685	0.279	0.842	0.824	0.880	0.984	0.984
$sign$	3.744	6.607	6.544	6.463	6.361	5.590	2.731	1.603	1.188
pv_s	1.000	1.000	1.000	1.000	1.000	1.000	0.997	0.946	0.883
$+/-$	0.724	0.882	0.868	0.868	0.855	0.803	0.645	0.579	0.592
		BN1			BN2			BN3	
ρ	0.994	0.994	0.996	0.991	0.993	0.996	0.991	0.993	0.996
F	1.135	0.968	0.985	1.207	0.973	0.979	1.207	0.973	0.979
pv_F	0.291	0.556	0.526	0.207	0.547	0.536	0.207	0.547	0.536
$sign$	8.056	8.056	8.056	8.533	8.286	8.056	8.533	8.286	8.056
pv_s	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$+/-$	0.961	0.961	0.961	0.987	0.974	0.961	0.987	0.974	0.961
		BQ1			BQ2			BQ3	
ρ	0.995	0.985	-	0.990	0.968	-	0.990	0.968	-
F	1.101	1.028	-	1.157	1.023	-	1.157	1.023	-
pv_F	0.338	0.452	-	0.263	0.460	-	0.263	0.460	-
$sign$	8.525	6.359	-	8.045	6.486	-	8.045	6.486	-
pv_s	1.000	1.000	-	1.000	1.000	-	1.000	1.000	-
$+/-$	0.987	0.895	-	0.961	0.921	-	0.961	0.921	-

[a]Results in the table refer to the reliability of the sequentially estimated cycle (SEC) against the last estimated cycle (LEC). ρ gives the correlation coefficient between SEC and LEC. F gives the ratio between the standard deviation of SEC and LEC. This ratio is distributed as an $F(n,n)$ under the null of both standard errors being equal. pv_F gives the level of significance of this F statistic. $sign$ performs the Pesaran and Potter (1997) test of directional accuracy betweem SEC and LEC, and pv_s gives the probability value. Finally $+/-$ gives the proportion of cases where the recursive and final series have the same sign.

Table 6

Correlation between Capacity Utilisation and Output Gap: US Data

Model		Period 1[a]	Period 2
UC1	*SEC*	0.774	-
	LEC	0.871	0.762
BN1	*SEC*	-0.287	-
	LEC	-0.310	-0.267
BQ1	*SEC*	-0.151	-
	LEC	-0.198	-0.122
UC2	*SEC*	0.631	-
	LEC	0.517	0.687
BN2	*SEC*	-0.296	-
	LEC	-0.318	-0.289
BQ2	*SEC*	-0.095	-
	LEC	-0.162	-0.051
UC3	*SEC*	0.759	-
	LEC	0.174	0.309
BN3	*SEC*	-0.296	-
	LEC	-0.318	-0.289
BQ3	*SEC*	-0.095	-
	LEC	-0.162	-0.051

[a]Period 1 refers to the period 1980Q1 to 1998Q4 and Period 2 to the period 1970Q1 to 1998Q4.

Figure 1

Final (solid line) vs Recursive (dashed line) Cycle Estimates. Euro area – UC Models

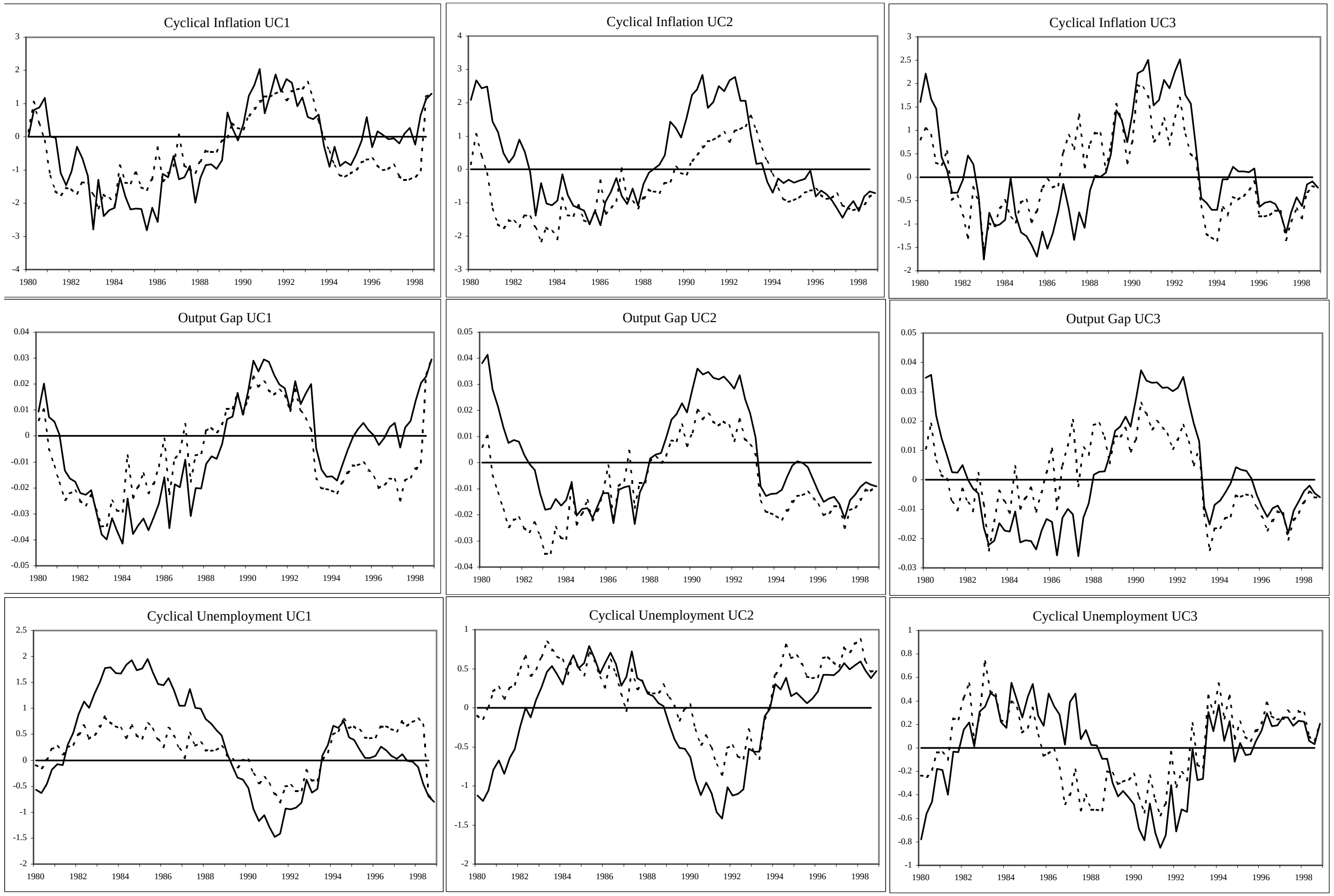


Figure 2

Final (solid line) vs Recursive (dashed line) Cycle Estimates. Euro Area – BN Models

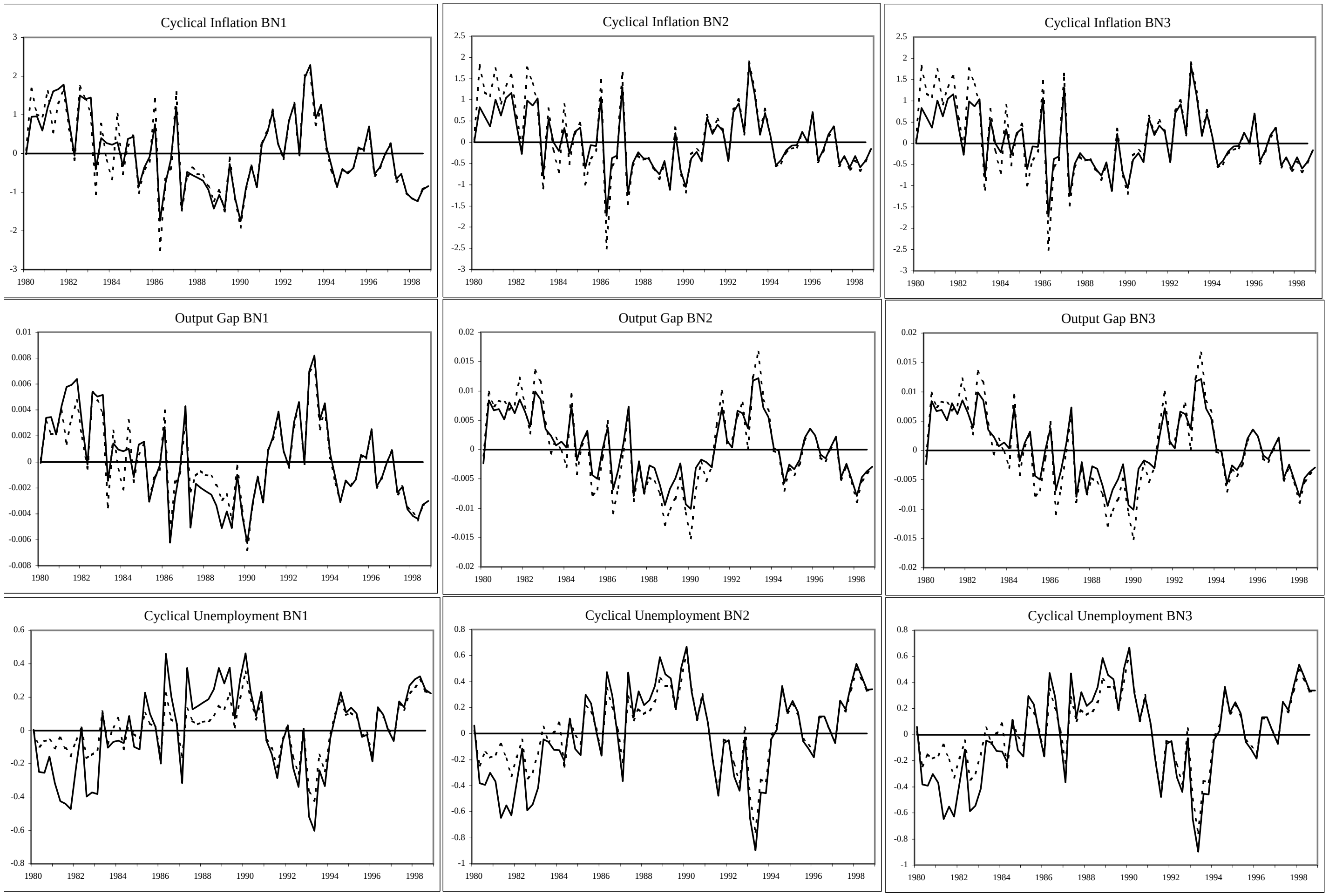


Figure 3

Final (solid line) vs Recursive (dashed line) Cycle Estimates. Euro Area – BQ Models

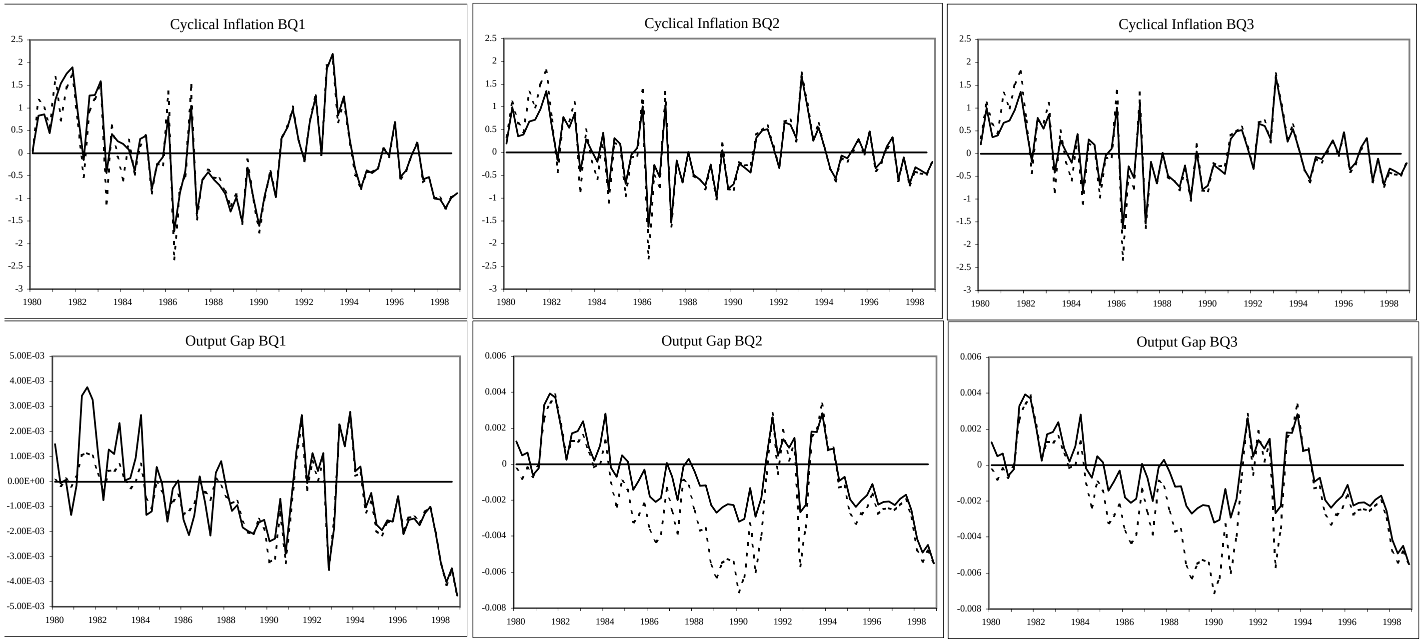


Figure 4

Final (solid line) vs Recursive (dashed line) Cycle Estimates. US Data – UC Models

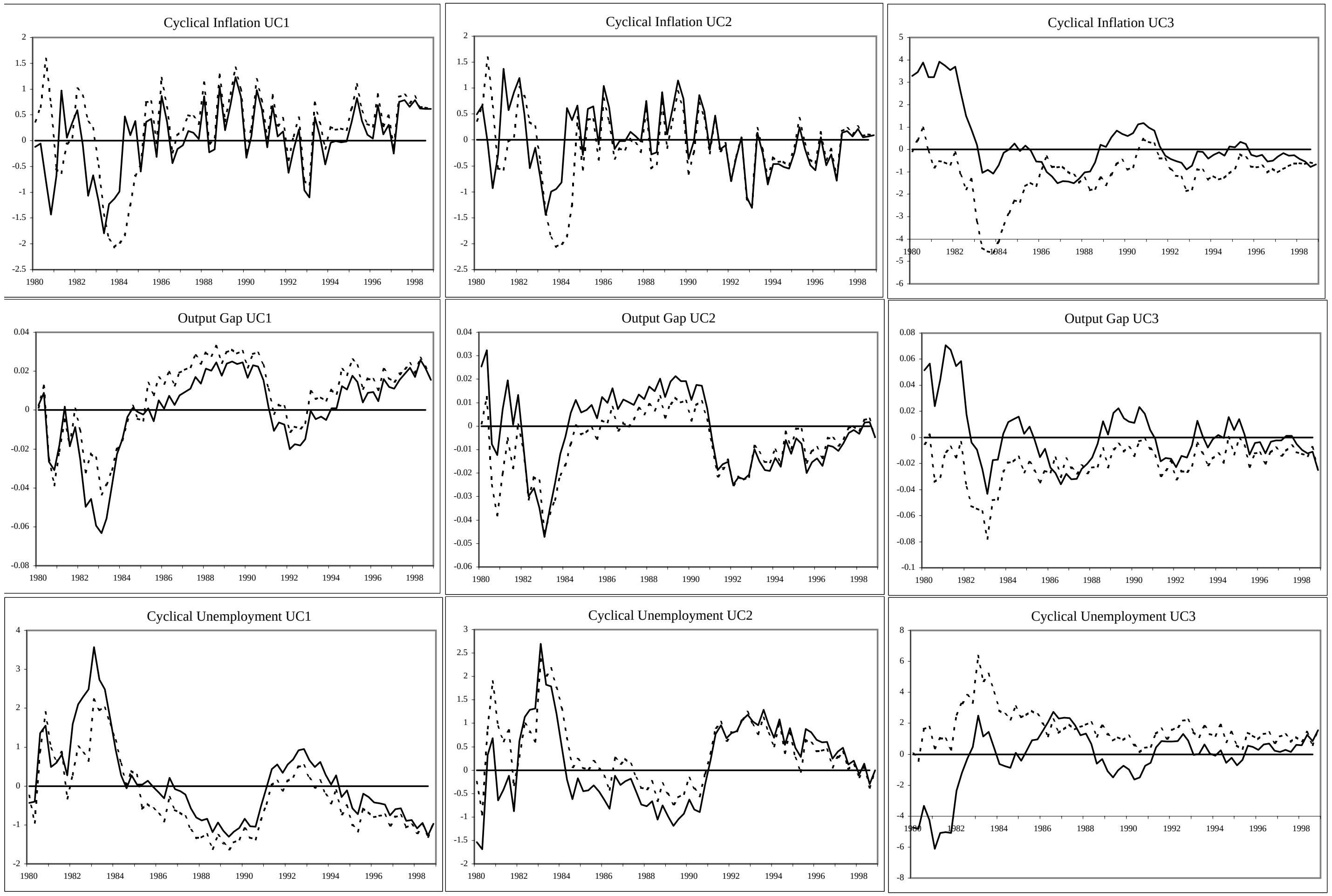


Figure 5

Final (solid line) vs Recursive (dashed line) Cycle Estimates. US Data – BN Models

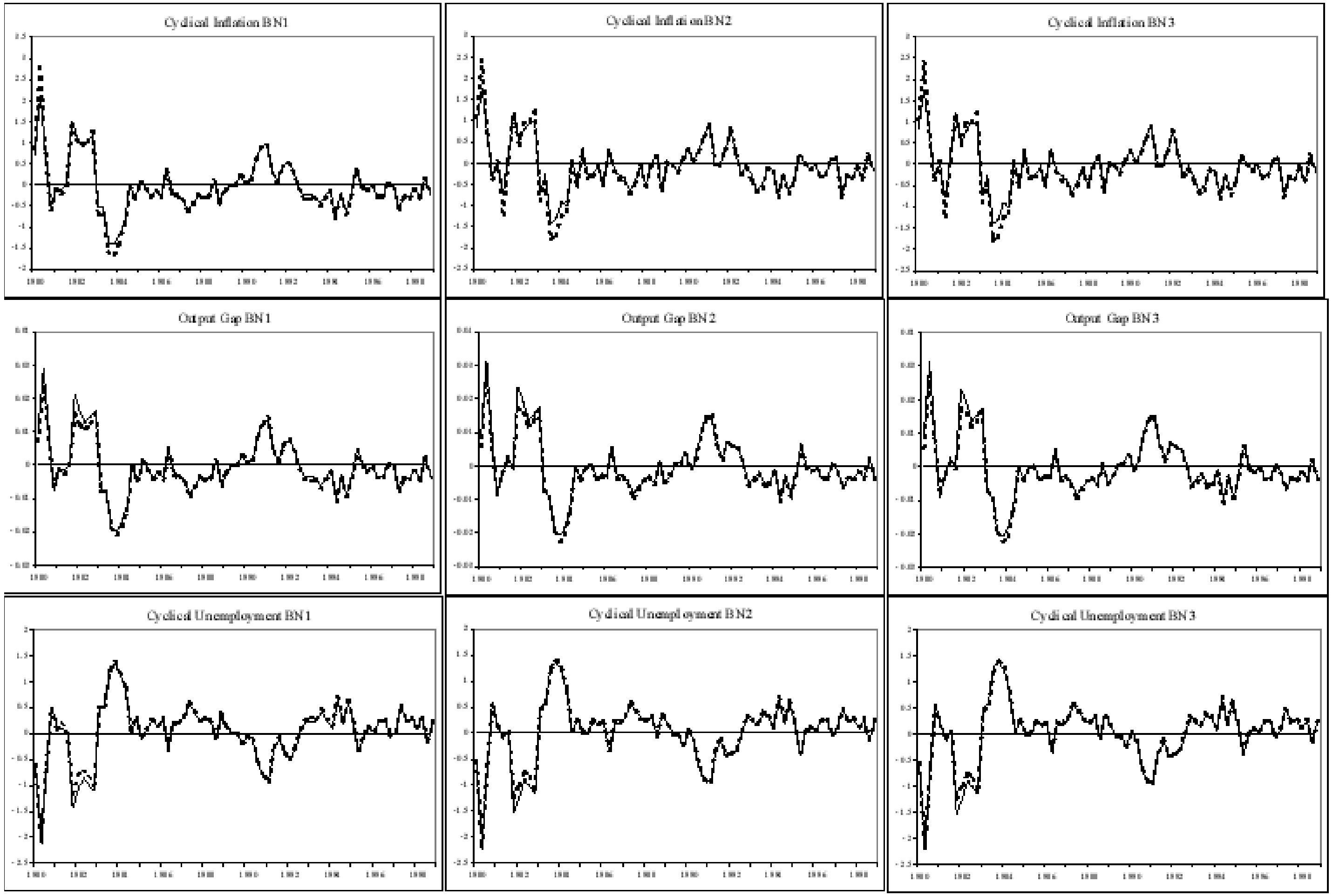
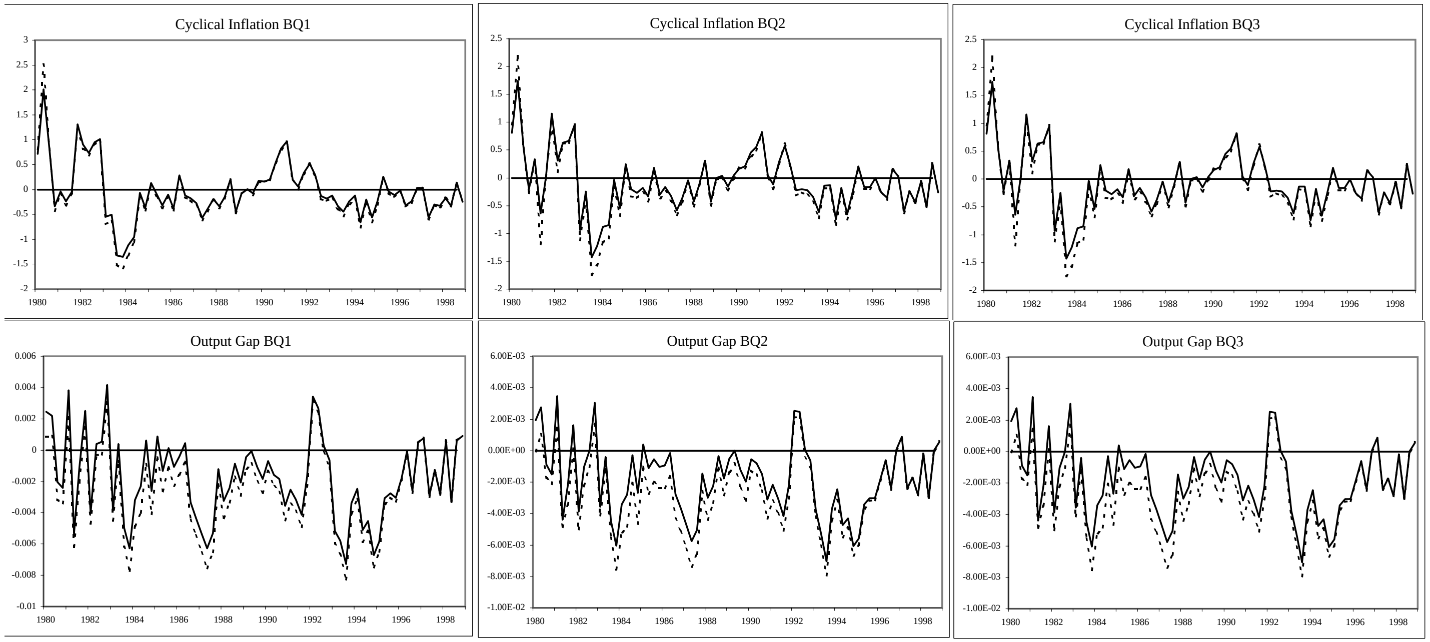


Figure 6

Final (solid line) vs Recursive (dashed line) Cycle Estimates. US Data Area – BQ Models



European Central Bank Working Paper Series

1 "A global hazard index for the world foreign exchange markets" by V. Brousseau and F. Scacciavillani, May 1999.

2 "What does the single monetary policy do? A SVAR benchmark for the European Central Bank" by C. Monticelli and O. Tristani, May 1999.

3 "Fiscal policy effectiveness and neutrality results in a non-Ricardian world" by C. Detken, May 1999.

4 "From the ERM to the euro: new evidence on economic and policy convergence among EU countries" by I. Angeloni and L. Dedola, May 1999.

5 "Core inflation: a review of some conceptual issues" by M. Wynne, May 1999.

6 "The demand for M3 in the euro area" by G. Coenen and J.-L. Vega, September 1999.

7 "A cross-country comparison of market structures in European banking" by O. de Bandt and E. P. Davis, September 1999.

8 "Inflation zone targeting" by A. Orphanides and V. Wieland, October 1999.

9 "Asymptotic confidence bands for the estimated autocovariance and autocorrelation functions of vector autoregressive models" by G. Coenen, January 2000.

10 "On the effectiveness of sterilized foreign exchange intervention" by R. Fatum, February 2000.

11 "Is the yield curve a useful information variable for the Eurosystem?" by J. M. Berk and P. van Bergeijk, February 2000.

12 "Indicator variables for optimal policy" by L. E. O. Svensson and M. Woodford, February 2000.

13 "Monetary policy with uncertain parameters" by U. Söderström, February 2000.

14 "Assessing nominal income rules for monetary policy with model and data uncertainty" by G. D. Rudebusch, February 2000.

15 "The quest for prosperity without inflation" by A. Orphanides, March 2000.

16 "Estimating the implied distribution of the future short term interest rate using the Longstaff-Schwartz model" by P. Hördahl, March 2000.

17 "Alternative measures of the NAIRU in the euro area: estimates and assessment" by S. Fabiani and R. Mestre, March 2000.

18 "House prices and the macroeconomy in Europe: Results from a structural VAR analysis" by M. Iacoviello, April 2000.

19 “The euro and international capital markets” by C. Detken and P. Hartmann, April 2000.

20 “Convergence of fiscal policies in the euro area” by O. De Bandt and F. P. Mongelli, May 2000.

21 “Firm size and monetary policy transmission: evidence from German business survey data” by M. Ehrmann, May 2000.

22 “Regulating access to international large value payment systems” by C. Holthausen and T. Rønde, June 2000.

23 “Escaping Nash inflation” by In-Koo Cho and T. J. Sargent, June 2000.

24 “What horizon for price stability” by F. Smets, July 2000.

25 “Caution and conservatism in the making of monetary policy” by P. Schellekens, July 2000.

26 “Which kind of transparency? On the need for clarity in monetary policy-making” by B. Winkler, August 2000.

27 “This is what the US leading indicators lead” by M. Camacho and G. Perez-Quiros, August 2000.

28 “Learning, uncertainty and central bank activism in an economy with strategic interactions” by M. Ellison and N. Valla, August 2000.

29 “The sources of unemployment fluctuations: an empirical application to the Italian case” by S. Fabiani, A. Locarno, G. Oneto and P. Sestito, September 2000.

30 “A small estimated euro area model with rational expectations and nominal rigidities” by G. Coenen and V. Wieland, September 2000.

31 “The disappearing tax base: Is foreign direct investment eroding corporate income taxes?” by R. Gropp and K. Kostial, September 2000.

32 “Can indeterminacy explain the short-run non-neutrality of money?” by F. De Fiore, September 2000.

33 “The information content of M3 for future inflation” by C. Trecroci and J. L. Vega, October 2000.

34 “Capital market development, corporate governance and the credibility of exchange rate pegs” by O. Castrén and T. Takalo, October 2000.

35 “Systemic risk: A survey” by O. De Bandt and P. Hartmann, November 2000.

36 “Measuring core inflation in the euro area” by C. Morana, November 2000.

37 “Business fixed investment: Evidence of a financial accelerator in Europe” by P. Vermeulen, November 2000.

38 "The optimal inflation tax when taxes are costly to collect" by F. De Fiore, November 2000.

39 "A money demand system for euro area M3" by C. Brand and N. Cassola, November 2000.

40 "Financial structure and the interest rate channel of ECB monetary policy" by B. Mojon, November 2000.

41 "Why adopt transparency? The publication of central bank forecasts" by P. M. Geraats, January 2001.

42 "An area-wide model (AWM) for the euro area" by G. Fagan, J. Henry and R. Mestre, January 2001.

43 "Sources of economic renewal: from the traditional firm to the knowledge firm" by D. R. Palenzuela, February 2001.

44 "The supply and demand for eurosystem deposits – The first 18 months" by U. Bindseil and F. Seitz, February 2001.

45 "Testing the Rank of the Hankel matrix: a statistical approach" by G. Camba-Mendez and G. Kapetanios, February 2001.

46 "A two-factor model of the German term structure of interest rates" by N. Cassola and J. B. Luís, February 2001.

47 "Deposit insurance and moral hazard: does the counterfactual matter?" by R. Gropp and J. Vesala, February 2001.

48 "Financial market integration in Europe: on the effects of EMU on stock markets" by M. Fratzscher, March 2001.

49 "Business cycle and monetary policy analysis in a structural sticky-price model of the euro area" by M. Casares, March 2001.

50 "Employment and productivity growth in service and manufacturing sectors in France, Germany and the US" by T. von Wachter, March 2001.

51 "The functional form of the demand for euro area M1" by L. Stracca, March 2001.

52 "Are the effects of monetary policy in the euro area greater in recessions than in booms?" by G. Peersman and F. Smets.

53 "An evaluation of some measures of core inflation for the euro area" by J.-L. Vega and M. A. Wynne, April 2001.

54 "Assessment criteria for output gap estimates" by G. Camba-Méndez and D. R. Palenzuela, April 2001.