

# Structural vector autoregressive (SVAR) based estimates of the euro area output gap: theoretical considerations and empirical evidence

GIAN LUIGI MAZZI, JAMES MITCHELL  
AND FILIPPO MOAURO

2016 edition





**Structural vector autoregressive  
(SVAR) based estimates of  
the euro area output gap:  
theoretical considerations  
and empirical evidence**

**GIAN LUIGI MAZZI, JAMES MITCHELL  
AND FILIPPO MOAURO**

**2016 edition**

***Europe Direct is a service to help you find answers  
to your questions about the European Union.***

**Freephone number (\*):  
00 800 6 7 8 9 10 11**

(\*) The information given is free, as are most calls (though some operators, phone boxes or hotels may charge you).

More information on the European Union is available on the Internet (<http://europa.eu>).

Luxembourg: Publications Office of the European Union, 2016

ISBN 978-92-79-60293-1

ISSN 2315-0807

doi: 10.2785/764491

Cat. No: KS-TC-16-009-EN-N

© European Union, 2016

Reproduction is authorised provided the source is acknowledged.

For more information, please consult: <http://ec.europa.eu/eurostat/about/our-partners/copyright>

Copyright for the photograph of the cover: ©Shutterstock.

For reproduction or use of this photo, permission must be sought directly from the copyright holder.

The information and views set out in this publication are those of the author(s) and do not necessarily reflect the official opinion of the European Union. Neither the European Union institutions and bodies nor any person acting on their behalf may be held responsible for the use which may be made of the information contained therein.

<b>Abstract .....</b>	<b>4</b>
<b>1. Introduction .....</b>	<b>5</b>
<b>2. SVAR models and BN decomposition .....</b>	<b>6</b>
2.1. The VAR model .....	6
2.2. Cointegration.....	7
2.3. The multivariate BN decomposition.....	8
2.4. The SVAR is also based on the BN decomposition .....	8
2.5. A new interpretation of the trend .....	9
<b>3. Relationship with state-space models .....</b>	<b>10</b>
<b>4. DSGE modelling .....</b>	<b>11</b>
<b>5. Combined density estimates of output gap.....</b>	<b>12</b>
5.1. Constructing the combined estimate of the output gap .....	12
5.2. The combined (density) estimate of the output gap.....	13
<b>6. The empirical comparison .....</b>	<b>14</b>
6.1. Forms of SVAR models.....	14
6.2. Benchmarks: multivariate UC models .....	15
6.3. The euro area output gap estimates .....	16
6.3.1. The shape of the cycle: steepness .....	19
6.3.2. Correlation with inflation .....	20
6.3.3. The permanent component of output .....	21
6.4. The reliability of real-time output gap estimates .....	22
6.5. Comparison of estimated weights .....	24
<b>7. Conclusion .....</b>	<b>26</b>
<b>References .....</b>	<b>27</b>

## Abstract

Users look with particular attention to the output gap either because it gives relevant information on the cyclical behaviour of the economy or for its capacity to anticipate features trends in inflation. Despite its relevance, the output gap is probably one of the unobserved economic variables for which the estimation is more controversial and uncertain. On one hand univariate methods are considered too mechanic and unable to incorporate economic constraints. On the other hand, structural multivariate methods based on production functions, require a big amount of information which is somewhat difficult to be gathered at high frequency. Moreover they often produce quite volatile estimates of the output gap due to the fact that they mainly aim to derive a proper estimate of the potential output. Structural VAR (SVAR) models are used widely in business cycle analysis to estimate the output gap because they combine together a robust statistical framework with the ability of integrating alternative economic constraints.

In this paper we consider several alternative specifications of economic constraints for both co-integrated and non-co-integrated SVAR models.

In the first part of the paper we give an economic interpretation of permanent and transitory components, within Dynamic Stochastic General Equilibrium models (DSGE), generated by alternative long run constraints and we compare them with the multivariate Beveridge-Nelson (BN) decomposition. In particular, we briefly discuss the fact that in the long run both SVAR models and multivariate BN decomposition have a random walk representation for the permanent component while in the short run they could diverge.

In the empirical part we analyse and compare alternative estimates of output gap for the euro area obtained under different constraints specifications referring to the classical UC model as well as to the univariate Hodrick–Prescott (HP) filter decomposition as benchmarks. We also present a detailed analysis of the real time properties of alternative output gap estimates measuring their reliability in terms of correlation with the estimates based on the whole sample. Output gap uncertainty across models is also addressed, with an application to integrate out model uncertainty and arrive at a combined density estimate of the output gap.

A perceived advantage of SVAR models, relative to unobservable components (UC) type models, is that SVAR models can be viewed as one-sided filters; in this sense they overcome the end-point problem associated with UC model that can be seen to involve application of a two-sided filter. We therefore compare the real-time performance of SVAR estimators with those of traditional UC models. Finally we study the capacity of different output gap estimates of forecasting inflation movements. Generally multivariate approaches perform better than the univariate HP filter. Within the multivariate methods, SVAR estimates deliver neither more informative nor less informative indicators of inflation than the other decompositions methods. Nevertheless SVAR models with inflation do better anticipate inflation than the benchmark HP filter, supporting the view that economic theory does help.

**Keywords:** Output gap estimation, Structural VAR, Beveridge-Nelson decomposition, Hodrick and Prescott filter, inflation forecasts.

**Acknowledgements:** Paper first presented at the International Symposium of Forecasters in Nice (June 2008) and to the Joint Statistical Meetings in Washington (August 2009). The authors are grateful to the participants to both conferences for their feedback and comments. The paper has benefitted from the outcomes of the multi-annual PEEIs project financed by Eurostat.

**Authors:** Gian Luigi Mazzi <sup>(1)</sup>, James Mitchell <sup>(2)</sup> and Filippo Moauro <sup>(3)</sup>

<sup>(1)</sup> Gian Luigi Mazzi, Eurostat, Luxembourg, [gianluigi.mazzi@ec.europa.eu](mailto:gianluigi.mazzi@ec.europa.eu).

<sup>(2)</sup> James Mitchell, Warwick Business School, UK, [James.Mitchell@wbs.ac.uk](mailto:James.Mitchell@wbs.ac.uk).

<sup>(3)</sup> Filippo Moauro, Istat, Rome, [email: moauro@istat.it](mailto:moauro@istat.it).



# 1. Introduction

The output gap is probably one of the unobserved economic variables for which the estimation is more controversial and uncertain. Its estimation can follow several approaches: a first method consists into the application of a univariate procedure, even if considered too mechanic and unable to incorporate economic constraints; a second is based on production functions requiring to fit structural multivariate models which make use of a larger amount of information not always easy to be gathered at high frequency. Moreover, they often produce quite volatile estimates of the output gap due to the fact that they mainly aim to derive a proper estimate of the potential output; a third method to estimate the output gap, widely used in business cycle analysis, is based on structural VAR (SVAR) models. This approach has the significant advantage to combine together a robust statistical framework with the ability of integrating alternative economic constraints.

The focus of this paper is on SVAR models, considering several alternative specifications of economic constraints for both co-integrated and non-co-integrated systems. An economic interpretation of permanent and transitory components, within Dynamic Stochastic General Equilibrium models (DSGE) is given in Sections 2 and 3 in contrast with the multivariate Beveridge-Nelson (BN) decomposition. Section 4 is on DSGE modelling, whereas Section 5 discusses output gap uncertainty and combined density estimates.

In the empirical part, Section 6, we analyse and compare alternative estimates of output gap for the euro area obtained under different constraints specifications referring to the classical unobservable components (UC) type model as well as to the univariate HP filter decomposition as benchmarks. Section 7 shortly concludes.

## 2. SVAR models and BN decomposition

This Section is devoted with the use of structural VAR models in business cycle analysis and alternative decomposition methods. Traditionally, within VAR models the output gap or business cycle is defined as those output movements associated with shocks constrained to have no long run effects on output, i.e. "transitory" shocks.

### 2.1. The VAR model

Consider the following VAR model of order  $p$ , [VAR( $p$ )] in the  $m$ -vector process  $\{z_t\}_{t=1}^{\infty}$

$$z_t = a_0 + a_1 t + \phi_1 z_{t-1} + \dots + \phi_p z_{t-p} + u_t. \quad (1)$$

for  $t=1, 2, \dots, T$ , where  $\phi_1, \phi_2, \dots, \phi_p$ , are  $(m \times m)$  matrices  $a_0$  and  $a_1$  are  $m$ -vectors of intercepts and trend coefficients and  $u_t$  is a serially uncorrelated  $m$ -vector of errors with a zero mean and a constant positive definite variance-covariance matrix,  $\Sigma = (\sigma_{ij})$ .

Following Bernanke (1986), Blanchard & Watson (1986) and Blanchard & Quah (1989) (1) is considered as the reduced form of a "structural" VAR model given by

$$A_0 z_t = \zeta + \lambda t + A_1 z_{t-1} + \dots + A_p z_{t-p} + \varepsilon_t \quad (2)$$

where  $A_1, A_2, \dots, A_p$  are  $(m \times m)$  matrices,  $A_0 = (a_{ij})$  is a nonsingular  $(m \times m)$  matrix of contemporaneous coefficients and  $\varepsilon_t$  is a serially uncorrelated  $m$ -vector of structural errors with a zero mean and a constant positive definite variance-covariance matrix,  $\Omega = (\omega_{ij})$ .

There is the following relationship between the remaining reduced form and structural parameters:

$$\begin{aligned} \Phi_i &= A_0^{-1} A_i \quad i = 0, 1, \dots, p, \\ u_t &= A_0^{-1} \varepsilon_t \quad \Sigma = A_0^{-1} \Omega A_0'^{-1}. \end{aligned} \quad (3)$$

Two alternative but equivalent representations of the VAR model that prove useful are the Vector Error Correction Model (VECM) and the Vector Moving Average (VMA) representation:

1. The VECM representation of  $\{z_t\}_{t=1}^{\infty}$ , is obtained by rewriting (1) as

$$\Delta z_t = a_0 + a_1 t + \Pi z_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta z_{t-i} + u_t, \quad (4)$$

where  $\Gamma_i = -\sum_{j=i+1}^p \phi_j$  for  $i=1, \dots, p-1$  and  $\Pi = \sum_{j=1}^p \phi_j - I_m$ .  $a_0$  and  $a_1$  are restricted when  $\Pi$  is rank deficient. The following relations define the mapping between the parameters of the VAR model and that of the VECM formulation  $\phi_1 = I_m - \Pi + \Gamma_1$ ,  $\phi_i = \Gamma_i - \Gamma_{i-1}$ ,  $i = 2, 3, \dots, p-1$ ,  $\phi_p = -\Gamma_{p-1}$ .



2. The infinite VMA representation of  $\{\Delta \mathbf{z}_t\}_{t=1}^{\infty}$  is defined as

$$\Delta \mathbf{z}_t = \mathbf{C}(L)(\mathbf{a}_0 + \mathbf{a}_1 t + \mathbf{u}_t) = \mathbf{b}_0 + \mathbf{b}_1 t + \mathbf{C}(L)\mathbf{u}_t \quad (5)$$

where  $\mathbf{b}_0 = \mathbf{C}(1)\mathbf{a}_0 + \mathbf{C}^*(1)\mathbf{a}_1$ ;  $\mathbf{b}_1 = \mathbf{C}(1)\mathbf{a}_1$ . The matrix lag polynomial  $\mathbf{C}(L)$  is given by

$$\mathbf{C}(L) = \mathbf{I}_m + \sum_{i=0}^{\infty} \mathbf{C}_i L^i = \mathbf{C}(1) + (1-L)\mathbf{C}^*(L), \quad \mathbf{C}^*(L) = \sum_{i=0}^{\infty} \mathbf{C}_i^* L^i \quad (6)$$

$$\mathbf{C}(1) = \sum_{i=0}^{\infty} \mathbf{C}_i, \quad \mathbf{C}^*(1) = \sum_{i=0}^{\infty} \mathbf{C}_i^* \quad (7)$$

The matrices  $\{\mathbf{C}_i\}$  derive from the recursion

$$\mathbf{C}_i = \sum_{j=1}^p \mathbf{C}_{i-j} \boldsymbol{\phi}_j, \quad i > 0, \quad \mathbf{C}_0 = \mathbf{I}_m \text{ and } \mathbf{C}_i = \mathbf{0} \text{ for } i < 0. \quad (8)$$

Similarly for the matrices  $\{\mathbf{C}_i^*\}$ ,  $\mathbf{C}_i^* = \mathbf{C}_i + \mathbf{C}_{i-1}^*$ ,  $i > 0$ ,  $\mathbf{C}_0^* = \mathbf{C}_0 - \mathbf{C}(1)$ . It follows from (1) and (5) that  $\mathbf{C}(L)\boldsymbol{\phi}(L) = \boldsymbol{\phi}(L)\mathbf{C}(L) = (1-L)\mathbf{I}_m$  is implying  $\boldsymbol{\Pi}\mathbf{C}(1) = \mathbf{C}(1)\boldsymbol{\Pi} = \mathbf{0}$  and  $\mathbf{C}(1)\mathbf{a}_1 = \mathbf{0}$ . Therefore  $\mathbf{b}_1 = \mathbf{0}$  and (5) is given by

$$\Delta \mathbf{z}_t = \mathbf{b}_0 + \mathbf{C}(L)\mathbf{u}_t. \quad (9)$$

Since  $\mathbf{u}_t = \mathbf{A}_0^{-1}\boldsymbol{\varepsilon}_t$ , the structural VMA representation corresponding to (9) is  $\Delta \mathbf{z}_t = \mathbf{b}_0 + \mathbf{D}(L)\boldsymbol{\varepsilon}_t$ , where  $\mathbf{D}(L) = \mathbf{C}(L)\mathbf{A}_0^{-1}$ .

## 2.2. Cointegration

The cointegration rank hypothesis is defined by

$$H_0: \text{Rank} [\boldsymbol{\Pi}] = r, \quad r < m \quad (10)$$

Under  $H_0$  of (10),  $\boldsymbol{\Pi}$  maybe expressed as  $\boldsymbol{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'$ , where  $\boldsymbol{\alpha}$  and  $\boldsymbol{\beta}$  are  $(m \times r)$  matrices of full column rank. Let  $\boldsymbol{\alpha}_{\perp}$  and  $\boldsymbol{\beta}_{\perp}$  denote  $(m \times m-r)$  matrices defined as the bases for the orthogonal complements of  $\boldsymbol{\alpha}$  and  $\boldsymbol{\beta}$ , such that  $\boldsymbol{\alpha}'\boldsymbol{\alpha}_{\perp} = \mathbf{0}$  and  $\boldsymbol{\beta}'\boldsymbol{\beta}_{\perp} = \mathbf{0}$ . Furthermore,  $\mathbf{C}(1)$  is a singular matrix which can be written as:

$$\mathbf{C}(1) = \boldsymbol{\beta}_{\perp}(\boldsymbol{\alpha}'_{\perp}\boldsymbol{\Gamma}\boldsymbol{\beta}_{\perp})^{-1}\boldsymbol{\alpha}'_{\perp} \quad (11)$$

## 2.3. The multivariate BN decomposition

Substituting (6) into the VMA representation, (9), and solving for the level,  $\mathbf{z}_t$ , yields the *common trends representation* or the multivariate BN decomposition

$$\mathbf{z}_t = \mathbf{z}_0 + \mathbf{b}_0 t + \mathbf{C}(1)\mathbf{s}_t^u + \mathbf{C}^*(L)(\mathbf{u}_t - \mathbf{u}_0), \quad (12)$$

$$\mathbf{z}_t = \boldsymbol{\mu} + \boldsymbol{\gamma} t + \mathbf{C}(1)\mathbf{s}_t^u + \mathbf{C}^*(L)\mathbf{u}_t, \quad (13)$$

where  $\{\mathbf{z}_t\}$  is initialised by  $\mathbf{z}_0 = \boldsymbol{\mu} + \mathbf{C}^*(L)\mathbf{u}_0$ ,  $\mathbf{s}_t^u$  is the partial sum defined by  $\mathbf{s}_t^u = \sum_{j=1}^t \mathbf{u}_j$ ,  $t = 1, 2, \dots$  and  $\boldsymbol{\gamma} = \mathbf{b}_0$  since  $\mathbf{C}(1)\boldsymbol{\Pi} = \mathbf{0}$ ,  $\mathbf{C}^*(1)\boldsymbol{\Pi} = \mathbf{I}_m$  and  $\mathbf{C}(1)\boldsymbol{\Gamma} - \mathbf{C}^*(1)\boldsymbol{\Pi} = \mathbf{I}_m$ .

Substituting (11) in (13) yields

$$\mathbf{z}_t = \boldsymbol{\mu} + \boldsymbol{\gamma} t + \boldsymbol{\Psi}\boldsymbol{\alpha}'_1 \mathbf{s}_t^u + \mathbf{C}^*(L)\mathbf{u}_t, \quad (14)$$

where  $\boldsymbol{\alpha}'_1 \mathbf{s}_t^u$  are  $(m-r)$  independent “common stochastic trends” and  $\boldsymbol{\Psi} = \boldsymbol{\beta}_1(\boldsymbol{\alpha}'_1 \boldsymbol{\Gamma} \boldsymbol{\beta}_1)^{-1}$  is a factor loading matrix of order  $(m \times m-r)$  measuring the long run response on  $\mathbf{z}_t$  of innovations to the common stochastic trends. The  $(m-r)$  independent common stochastic trends are supplemented by  $r$  linearly independent deterministic trends. Crucially there are fewer common stochastic trends than variables, the difference being the number of cointegrating relations,  $r$ .

$\mathbf{C}(1)\mathbf{s}_t^u = \boldsymbol{\Psi}\boldsymbol{\alpha}'_1 \mathbf{s}_t^u$  represents the *permanent* or *long run* component of  $\mathbf{z}_t$ : the shocks that drive the common stochastic trends are permanent since,  $\boldsymbol{\Psi}\boldsymbol{\alpha}'_1 \mathbf{s}_t^u = \boldsymbol{\Psi}\boldsymbol{\tau}_t$  where  $\boldsymbol{\tau}_t$  is a  $(m-r)$  vector of random walks given by  $\boldsymbol{\tau}_t = \boldsymbol{\tau}_{t-1} + \mathbf{u}_t^*$ , with  $\mathbf{u}_t^* = \boldsymbol{\alpha}'_1 \mathbf{u}_t$  denoting  $(m-r)$  *permanent* shocks. The common stochastic trends are consequently also interpreted as a characterisation of the “long run” in cointegrating VAR models. Note that in (14) the permanent shocks,  $\mathbf{u}_t^*$ , and the transitory shocks,  $\mathbf{u}_t$ , are colinear. But (14) is only one possible permanent/transitory decomposition of the shocks; see King *et al.* [(1987), pp. 40-2] for an alternative decomposition based on a Jordan decomposition of  $\mathbf{C}(1)$ . The lack of uniqueness of any particular decomposition is only resolved by imposing restrictions on  $(\boldsymbol{\Psi}, \boldsymbol{\alpha}'_1)$  motivated by *a priori* theory.

## 2.4. The SVAR is also based on the BN decomposition

When  $\{\mathbf{z}_t\}_{t=1}^\infty$  is a first difference stationary process,  $\boldsymbol{\Pi} = \mathbf{0}$ , Shapiro and Watson (1988) [SW] and Blanchard & Quah (1989) [BQ] pioneered the approach of identifying “structural” shocks by restricting the matrix of structural long run multipliers,  $\mathbf{D}(1) = \mathbf{C}(1)\mathbf{A}_0^{-1}$ , where  $\mathbf{C}(1)$  is a non-singular matrix. These restrictions conventionally take the form of assuming  $\mathbf{D}(1)$  is lower triangular. Motivated by “long run” *a priori* theory (e.g. a long run neutrality hypothesis) a shock is attributed economic meaning on the basis of whether its long run (infinite horizon) impact on particular endogenous variables is permanent or transitory. A structural shock is constrained to have a transitory [permanent] impact on the level of a variable by imposing zero [nonzero] restrictions upon elements of  $\mathbf{D}(1)$ .

Therefore from (13) we can see that what SVAR models are doing is attributing economic meaning to the shocks, i.e. the BN decomposition works off the structural shocks  $\boldsymbol{\varepsilon}_t$  not the reduced form shocks  $\mathbf{u}_t$ :

$$\mathbf{z}_t = \boldsymbol{\mu} + \boldsymbol{\gamma} t + \mathbf{D}(1)\mathbf{s}_t^\varepsilon + \mathbf{D}^*(L)\boldsymbol{\varepsilon}_t. \quad (15)$$

What SVAR models do is add in some additional stationary component to the BN trend by imposing additional identification restrictions; see also Dupasquier, Guay & St-Amant (1999) for an excellent survey. This is seen in the equation above by simply noting that in the SVAR method we define the trend component as the sum of the permanent shocks. Having identified the structural shocks we can see that the SVAR trend will involve summing elements of both  $\mathbf{D}(1)s_t^\varepsilon$  and  $\mathbf{D}^*(L)\varepsilon_t$  since the latter picks up the (short-run) dynamics of the permanent shocks.

That is, the permanent shocks fall both in  $\mathbf{D}(1)s_t^\varepsilon$  and  $\mathbf{D}^*(L)\varepsilon_t$  - hence in SVAR models the trend is not restricted to be a random walk, although it has a random walk component.

This is also the case when there is cointegration, although when there is cointegration a reduced number of shocks drive the common stochastic trends.

In particular, knowledge of the cointegration rank establishes that there are  $(m-r)$  common stochastic trends, but it does not identify the common stochastic trends *individually*. The decomposition of  $\mathbf{C}(1)$  that enables  $\alpha'_\perp s_t^\varepsilon$  to be interpreted as  $(m-r)$  independent common stochastic trends, and  $\Psi$  as the factor loading matrix, is not unique.

Identification of the common stochastic trends is distinct from identification of the cointegrating relations in the sense that even if  $\alpha$  and  $\beta$  are identified there are still an infinite number of candidates for  $\alpha_\perp$  and  $\beta_\perp$ . Therefore  $(m-r)^2$  *a priori* restrictions must be imposed on  $(\alpha'_\perp, \Psi)$  to identify the common stochastic trends. Restrictions can be imposed on either  $\alpha'_\perp$ , or  $\Psi$ , since once  $\Psi$  is identified  $\alpha'_\perp$  can be obtained as  $\alpha'_\perp = (\Psi'\Psi)^{-1}\Psi'\mathbf{C}(1)$ . Equivalently once  $\alpha'_\perp$  is identified we have  $\Psi = \mathbf{C}(1)(\alpha'_\perp\alpha'_\perp)^{-1}\alpha'_\perp$ .

## 2.5. A new interpretation of the trend

Garratt, Robertson & Wright (2006) present a new derivation of the multivariate BN decomposition from the structural cointegrating vector autoregression (VAR) representation. This allows for a clearer economic interpretation of the trend and cyclical components.

As seen above, BN trends are usually derived from the moving average (MA) representation. Garratt, Robertson & Wright (2006) show that, alternatively, they can be derived from the VAR representation. This lets them relate the cycle directly to the underlying stationary processes, which have an economic interpretation, which drive the cycle. Importantly, and in common with DSGE estimates of the output gap, they find that trend or potential output need not be *smooth* as typically implied by statistical and mixed decomposition methods, like UC models. The explicit use of economic theory when decomposing a time-series clearly affects the properties of the trend and cyclical components.

More in particular it can be shown that the trend component, i.e.  $\Delta z_t = \mathbf{C}(1)\mathbf{u}_t$  can be interpreted as conditional cointegrating equilibrium values, i.e. the values to which series tend when the series have converged to their cointegrating equilibria. This lets us decompose the output gap into the contributions of individual cointegrating disequilibria. It means that BN trends can be interpreted directly with respect to the long-run of the SVAR model. This ensures the BN trends are not black-box but have as coherent an economic structure as the cointegrating SVAR model. A last consideration: since BN decompositions assume the permanent component does follow a random walk, if we do not believe that deviations from mean trend output growth are white noise, then we should abandon the BN decomposition. In other words, the BN decomposition implies trend output growth is unpredictable. But if we believe actual output growth equals a constant drift plus a stationary autocorrelated variable then it is sensible.

### 3. Relationship with state-space models

Of course, a VAR model can always be written in state-space form. So, in this sense, there is no distinction between VAR and state-space models. More constructively, recent work (Morley, Nelson & Zivot, 2003) has clarified the relationship between UC and BN decompositions, showing that once the UC model is modified (from its traditional form with uncorrelated innovations - cf. Harvey & Jaeger, 1993) to allow for correlated trend and cycle innovations then the UC model coincides exactly with the BN decomposition. In particular, the BN components (from the reduced form of the UC model) are identical to the real-time estimates from the correlated UC model; cf. Proietti (2002).

It is instructive to distinguish between two state-space approaches for constructing output gap estimates from SVAR models.

UC models with uncorrelated innovations. Proietti (1997) derives expressions for the unobserved components of the multivariate BN decomposition from a SVAR model. This involves writing the cointegrating VAR model (4) in state space form, and then deriving expressions for the trend and cycle based on the Kalman filter. Importantly, the underlying VAR model traditionally is assumed to have finite lag order  $p$ .

UC models with correlated innovations. In contrast, exploiting the Stock & Watson (1988) representation of a cointegrating system, one can write an infinite order VAR model in state-space form and again use the Kalman filter to back out the trend and cyclical estimates.

Morley (2007) proves that this model is identified. In fact, Granger's lemma reveals that the above system has an underlying reduced-form VARMA representation which can be used for the estimation of the output gap when one recalls the debate over the importance of lag length,  $p$ , for estimation of the output gap in SVAR models; cf. Faust & Leeper (1997). In particular, Fernández-Villaverde, Rubio-Ramírez & Sargent (2005) and Ravenna (2007) show that DSGE models imply VARMA not VAR processes.

## 4. DSGE modelling

DSGE models provide a precise and economically coherent definition of, in particular, the output gap. They define potential or trend output as the level of output that an economy attains when the inefficiencies resulting from nominal wage and price rigidities have disappeared; potential output is defined as a flexible price equilibrium when prices and wages have returned to their equilibrium levels. Therefore, the DSGE definition accords with the idea that potential output is the level of output at which inflation tends neither to rise nor fall; see Neiss & Nelson (2005).

But potential output as defined in a DSGE model has quite different properties from many more statistical (or at best mixed) approaches to estimation of the trend; see Neiss & Nelson (2005). In DSGE models potential output is affected by real shocks over the business cycle. This implies that it does not follow a smooth trend, as in many statistical decomposition such as UC models. Government spending, terms of trade shocks, tax changes and taste shocks all affect households' decisions about consumption and so affect labour supply and hence trend output. Trend output, therefore, can vary over the business cycle.

In contrast, more statistical approaches to estimating potential output generally assume that such shocks have no important effects on potential output at business-cycle frequencies. As a consequence, their estimates typically have smaller fluctuations than measures of potential output derived from DSGE models.

Furthermore, as explained by Neiss & Nelson (2005), the business cycle is in fact logically distinct from the output gap in DSGE models. The output gap is explained by nominal rigidities, and when prices are flexible will equal zero even if there is a business cycle driven by real demand and supply shocks causing output to vary cyclically. Indeed Blanchard & Quah (1989) also note that their SVAR trends will not be the same as trends from flex-price equilibria.

## 5. Combined density estimates of output gap

Recent works by Orphanides and van Norden (2002 and 2005) document the unreliability of the output gap measurements implied by a variety of well-known detrending methods, and the associated difficulties of using those measurements to forecast inflation. In real time output gap estimation a considerable uncertainty is generated not only for a given model but also across models. In this latter respect, a more recent work by Garratt, Mitchell & Vahey (2009) develops a methodology to integrate the two sources of uncertainty and arrive at a combined density estimate of the output gap. The weights in this combined measure of the output gap are chosen to maximise the density forecasting performance of the output gap for inflation.

Behind the concept of density forecast there is the recognition that it is important to provide a quantitative indication of the uncertainty associated with a point forecast, along with the balance of risks (skewness) on the upside and downside and the probability of extreme events (fat tails or kurtosis). According to the literature on output gap estimation, the 'better' estimates are those that deliver 'better' inflation forecasts. On this basis it is argued that the appropriate way of evaluating the implied inflation forecasts is not to use some arbitrary statistical loss function, but the appropriate economic loss function; see Granger & Pesaran (2000). Only when the forecast user has a symmetric, quadratic loss function, and the constraints (if relevant) are linear, is it correct to focus on the point forecast alone. In the more general case, the degree of uncertainty matters and it is important to take into account the degree of uncertainty about the point forecast. Uncertainty is expected to attenuate the response or reaction of policy-makers to the point estimate.

### 5.1. Constructing the combined estimate of the output gap

As Orphanides & van Norden (2005) note, monetary policymakers often evaluate output gap measure by their ability to predict inflation. Accordingly, for each measure of the output gap, Garratt, Mitchell & Vahey (2009) estimate a set of vector autoregressive models in inflation and the output gap. The VAR models themselves can be estimated for a range of different lag orders, e.g. lag lengths of 1-4.

We utilize a direct forecast methodology to generate the h-step ahead predictive densities from each VAR. Consider a single equation from a given (constant parameter) VAR model:

$$Y_t = \alpha + \beta X_{t-h} + \sigma e_t \quad (16)$$

where  $t = t_0, \dots, \tau$  refers to the sample (available at vintage  $\tau + 1$  used to fit the model,  $\tau_t \sim iid(0,1)$ , and the parameters are collected in  $\alpha$ ,  $\beta$  and  $\sigma$ . The predictive densities for (with non-informative priors), allowing for small sample issues, are multivariate Student-t; see Zellner (1971), pp. 233-236.



## 5.2. The combined (density) estimate of the output gap

To summarize the density combination approach, for each observation in the policymaker's evaluation period, we use forecast density combination methods to compute the weights on each component based on the recursive log score of the predictive densities for inflation; see Jore, Mitchell & Vahey (2009). Given these weights, denoted as recursive log-score weights (RLSW), we construct an aggregate forecast density for inflation (recursively, at each horizon). We might also use those same weights to construct an h-step ahead ensemble predictive for the output gap from the many component models. Note that each component VAR predicts the latent variable, the output gap, via the second equation in the VAR. The resulting output gap densities reflect the model uncertainty in the many specifications, including the measure of the output gap.

Given  $i = 1, \dots, N$  VAR specifications, the combined densities are defined by the convex combination:

$$p(\pi_{\tau,h}) = \sum_{i=1}^N \omega_{i,\tau,h} g(\pi_{\tau,h} | I_{i,\tau}), \quad \tau = \underline{\tau}, \dots, \bar{\tau} \quad (17)$$

where  $g(\pi_{\tau,h} | I_{i,\tau})$  are the h-step ahead forecast densities from model  $i$ , conditional on the information set  $I_{i,\tau}$ . The non-negative weights,  $\omega_{i,\tau,h}$ , in this finite mixture sum to unity. Furthermore, the weights may change with each recursion in the evaluation period  $\tau = \underline{\tau}, \dots, \bar{\tau}$ .

Since each VAR considered produces a forecast density that is multivariate Student-t (see the discussion in Jore, Mitchell & Vahey, 2009), the combined density defined by equation (17) will be a mixture—accommodating skewness and kurtosis. That is, the combination delivers a more flexible distribution than each of the individual densities from which it was derived. As  $N$  increases, the combined density becomes more and more flexible, with the potential to approximate non-linear specifications. Notice that our focus is on the predictive accuracy of the ensemble density, rather than the (many) individual VAR components.

We construct the weights based on the fit of the individual component forecast densities following Jore, Mitchell & Vahey (2009), where the logarithmic score is used to measure density fit for each model through the evaluation period. The logarithmic scoring rule is intuitively appealing as it gives a high score to a density forecast that assigns a high probability to the realized value.

Specifically the RLSW weights take the form:

$$\omega_{i,\tau,h} = \frac{\exp\left[\sum_{\underline{\tau}}^{\tau-1-h} \ln g(\pi_{\tau,h} | I_{i,\tau})\right]}{\sum_{i=1}^N \exp\left[\sum_{\underline{\tau}}^{\tau-1-h} \ln g(\pi_{\tau,h} | I_{i,\tau})\right]} \quad (18)$$

From a Bayesian perspective, density combination based on recursive logarithmic score weights has many similarities with an approximate predictive likelihood approach. Given our definition of density fit, the model densities are combined using Bayes' rule with equal (prior) weight on each model, which a Bayesian would term non-informative priors.

## 6. The empirical comparison

### 6.1. Forms of SVAR models

1. Bivariate SVAR models in line with Robertson & Wickens (1997) that include output and inflation with the identifying restriction that nominal shocks have no long-run effect on output - this amounts to assuming model is super-neutral. Following Robertson & Wickens (1997) we interpret the two shocks as real and nominal rather than supply and demand. The nominal shocks have no long-run effect on output.
2. Traditional BQ model, in the level of unemployment instead of inflation, which identifies the shocks by assuming that demand shocks do not have a permanent effect on either unemployment or output, but supply shocks may have a long run effect on output but not unemployment. Note both shocks have no effect in the long-run on  $\Delta y_t$  or  $U_t$  since the VAR in these two variables is assumed stationary.
3. Trivariate SVAR models in line with Astley & Yates (1999) and Camba-Mendez & Rodriguez-Palenzuela (2003). We will consider trivariate VAR models of output, inflation and unemployment; see Camba-Mendez & Rodriguez-Palenzuela (2003). Restrictions then can be imposed on the matrix of long-run multipliers so that inflation is determined in the long-run by only one structural shock (the inflation shock), whereas output is determined by two structural shocks (the inflation shock plus a supply shock), and unemployment is determined by three structural shocks (the inflation shock, the supply shock and a demand shock). The output gap is then the cumulative sum of the demand shocks to output. Specifically inflation is determined in the long run by one structural shock (the inflation shock), output by two shocks (the inflation and supply shocks) and unemployment by all three shocks. It is important that the estimated (reduced-form) VAR model, that provides the basis for the identification of the output gap, is stationary. This requires the variables in the VAR to be appropriately transformed prior to estimation. Following Camba-Mendez & Rodriguez-Palenzuela (2003) in their application to the euro area we consider the first differences of inflation, GDP and unemployment. We did, however, experiment with representations in the level of inflation and unemployment.
4. Rather than estimating a stationary (perhaps first-differenced) VAR model a cointegrating VAR model is also considered. We follow King, Plosser, Stock & Watson (1991) and consider a three variable system of GDP, consumption and investment. There are two cointegrating vectors: the ratios of consumption and investment to GDP are stationary. Therefore, there is one permanent shock and two transitory shocks. The output gap is the cumulative sum of the two transitory shocks to output. This model remains of great interest; e.g. see Attfield & Temple (2003). Attfield & Temple (2003) consider the stationarity of the great ratios, noting that these ratios have been subject in the UK and the US to occasional structural (mean) breaks. Once these breaks are accommodated the great ratios are once again (stochastic) stationary.

The long-run restrictions implicit to identification require estimation of the matrix of long run responses (the sum of the lag polynomial in the moving-average representation corresponding to the VAR model). This is based on estimation of the sum of the coefficients in the VAR model. Therefore the reliability of SVAR estimates of the output gap rests on reliable estimation of the AR coefficients; see Faust & Leeper (1997) for details. A key issue then is selection of lag order in the VAR model. Using too small a lag order can lead to significant biases in estimation of the permanent and transitory components; see DeSerres & Guay (1995). DeSerres & Guay (1995) found that use of information criteria leads to too low an estimated lag order. This is consistent with our findings in the sense that the BIC tended to select a lag order of one, but it was only with much larger lag orders that sensible looking output gap estimates were generated. For robustness in the empirical comparison below we consider three different lag lengths,  $p=1$ ,  $p=6$  and  $p=12$ .

## 6.2. Benchmarks: multivariate UC models

*A bivariate UC model of output and inflation.* To give the output gap a more economic interpretation than in univariate unobserved components models, it is related to data on inflation. This is sensible as the causal relationship between the output gap and inflation is central to the Phillips Curve and the conduct of monetary policy.

We consider general models of a form with one equation for the decomposition of the log of the actual output in its trend level and the output gap; a second for the Phillips Curve equation where the inflation is measured at the annual rate.

The Phillips Curve equation provides a link between inflation and aggregate demand (measured here by the output gap). Since inflation is commonly assumed to depend only on nominal factors in the long-run we can impose a long run homogeneity restriction. This can be seen expressing the equation, in terms of level of inflation. From the other hand long run homogeneity requires an expression of the Phillips Curve in terms of first differences of inflation.

The output gap takes a lag polynomial form of the second order with complex roots constrained to be stationary. The trend of output follows a general local linear trend.

*A trivariate UC model of output, inflation and unemployment.* See Apel & Jansson (1999) and Fabiani & Mestre (2001). This is based on the view that the output and unemployment gaps are closely related. Inflation is likely to contain information about the size of both gaps. Restrictions can then be imposed in an unobserved components model that identify not just the Phillips Curve, linking measures of excess demand to inflation, but Okun's Law, that relates the unemployment and output gaps.

The Phillips Curve equation considered is a version of Gordon's triangle Phillips Curve model, relating inflation to movements in the unemployment gap. Expectations are implicit in the inflation dynamics. A further equation is for the Okun's Law relationship, relating cyclical unemployment and output movements.

Both the trend for output and unemployment are assumed to follow a local linear trend model. This representation was used with success by Fabiani & Mestre (2001) in an application to the Euroarea. The system does not allow for full endogeneity between real disequilibria and inflation, that would involve neither real disequilibria causing inflation nor vice-versa. A restrictive path for the transmission of demand shocks is implied as demand shocks lead to inflation via the unemployment gap, and then and from the unemployment gap to the output gap; see Astley & Yates (1999).

*The Cogley-Sargent (STAMP 8) approach.* Inflation and output are modelled jointly with trend inflation following a random walk and trend output an integrated random walk. The irregular terms may be correlated and have a covariance matrix. A seasonal component can be added in too when appropriate.

The assumption that trend inflation follows a random walk has become something of a consensus in macroeconomics; see Cogley & Sargent (2007). While inflation targets might be thought to suggest that trend inflation does not follow a random walk, Cogley & Sargent (2007) argue that inflation remains highly persistent - they argue that target inflation has not stopped drifting, even if its conditional variance has declined post Great Moderation; see Mitchell (2007) for discussion of the changing nature of forecasters' impressions of the conditional mean and conditional variance of inflation in the United States. We therefore follow this tradition and model trend inflation as a random walk.

The stochastic cycles are modelled as *similar* cycles. The trend, cyclical and any seasonal components are assumed orthogonal; but the similar cycles allow for the cycles in the different series to have the same autocorrelation functions. So, as opposed to Basistha & Nelson (2007) who explicitly allow for error dependence (à la MorleyNelsonZivot03), the STAMP/Harvey model handles the dependence by defining the cycles as similar.

*The Morley 2007 correlated UC model (with cointegration).* We can also consider representation of a cointegrated system using the common stochastic trends, implying an underlying VARMA process.

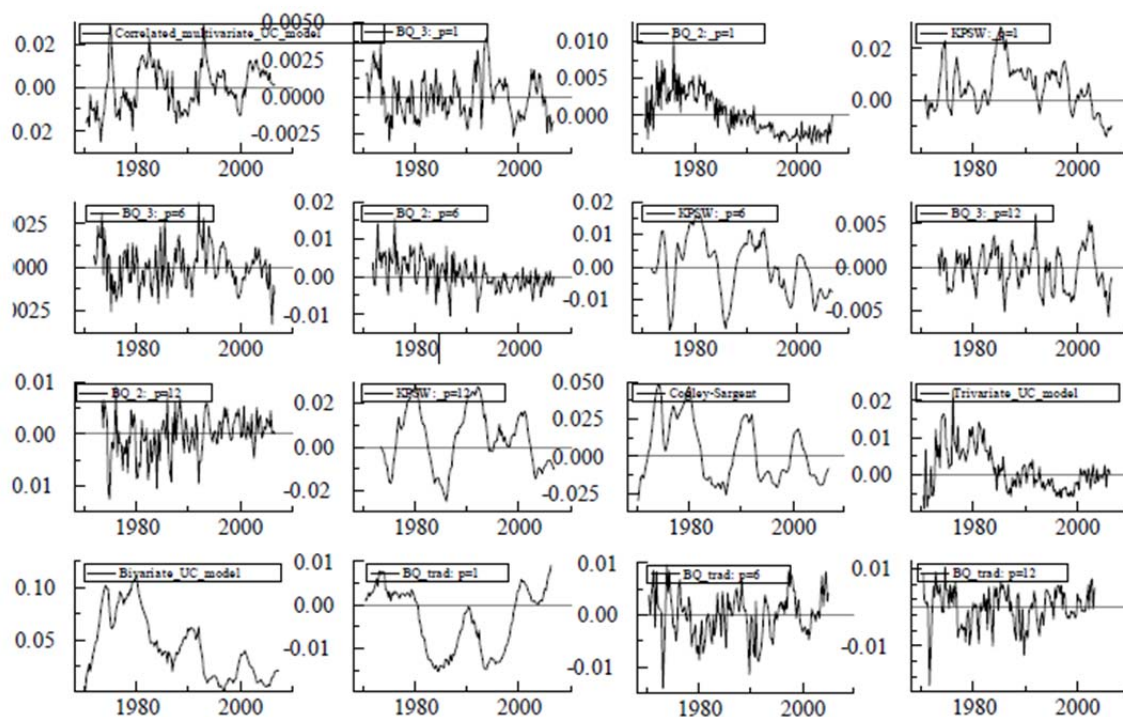
### 6.3. The euro area output gap estimates

Euro area data are taken from the ECB's Area Wide Model (AWM) database released by the ECB in the Autumn of 2007. These AWM data involve aggregating the available national data using the so-called index method that, for example, defines the log of euro area GDP as the weighted sum of the log of country-level GDP. The weights are based on relative GDP shares; see Fagan, Henry & Mestre (2005) for further details. These data span 1970q1-2006q4 and offer a useful tool for business cycle analysis.

Figure 1 plots the output gap estimates implied by the different estimators. Broadly, across the decomposition methods we observe peaks in economic activity in the early 1980s and early 1990s. But inference is sensitive to measurement. While the KPSW, correlated UC model (Morley) and Cogley-Sargent (STAMP) estimates exhibit pronounced cyclical activity, some of the Blanchard-Quah type decompositions (especially with inflation rather than unemployment) are more noisy.

Indeed inspecting the scale on the y-axes reveals that the BQ decompositions can yield output gaps of small magnitude - results are also sensitive to the lag order selected. In general, Figure 1 suggests that more *sensible* looking output gap estimates are obtained for lag orders greater than unity. This supports use of the correlated UC model, in our case the model of Morley (2007), since this imposes no restrictions on the order of the underlying VAR model.

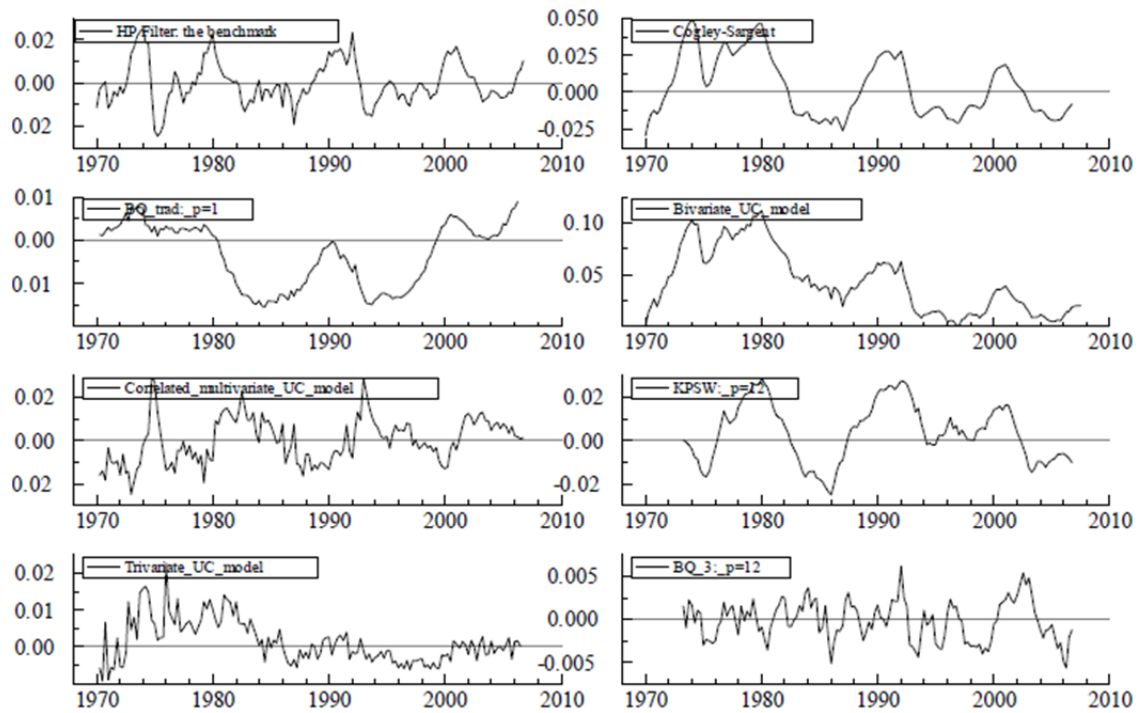
**Figure 1:** Euro area output gap estimates generated using different estimators



Source: Authors' calculations

Figure 2 then plots, as a reference, selected output gap estimates from Figure 1 against the Hodrick–Prescott (HP) filter.

**Figure 2:** The benchmark: HP estimates of the euro area output gap against selected multivariate estimates



Source: Authors' calculations

Tables 1 and 2 distill the information in Figure 1, by reporting the full-sample correlation between the different output gap estimates. Table 1 reveals that the alternative cycles are not closely related, despite some evidence that their peaks coincide. For example, the correlated UC cycle and the Cogley-Sargent (CS) cycles appear from Figure 1 reasonably consistent in terms of when they identified peaks in economic activity, but they are correlated -0.36.

**Table 1:** Correlation of SVAR output gap estimates with each other and alternative decompositions

	CorUC	BQ3:p=1	BQ2:p=1	KPSW:p=1	BQ3:p=6	BQ2:p=6	KPSW:p=6
CorUC	1.00	0.12	-0.06	-0.23	0.02	-0.03	-0.05
BQ3:p=1	0.12	1.00	-0.11	0.07	0.49	-0.09	0.09
BQ2:p=1	-0.06	-0.11	1.00	0.01	-0.17	0.77	0.33
KPSW:p=1	-0.23	0.07	0.01	1.00	0.17	-0.14	-0.17
BQ3:p=6	0.02	0.49	-0.17	0.17	1.00	-0.26	0.08
BQ2:p=6	-0.03	-0.09	0.77	-0.14	-0.26	1.00	0.18
KPSW:p=6	-0.05	0.09	0.33	-0.17	0.08	0.18	1.00
BQ3:p=12	-0.01	0.36	0.01	0.06	0.41	-0.12	0.22
BQ2:p=12	-0.16	0.20	0.06	-0.05	-0.04	0.56	-0.08
KPSW:p=12	-0.27	-0.02	-0.03	-0.39	0.05	-0.01	0.71
CS	-0.36	-0.23	0.51	-0.17	-0.07	0.28	0.62
TriUC	0.00	-0.12	0.90	-0.05	-0.15	0.61	0.44
BiUC	-0.27	-0.24	0.79	-0.01	-0.12	0.43	0.56
BQt:p=1	-0.35	-0.35	0.27	-0.36	-0.22	0.21	0.20
BQt:p=6	-0.02	-0.01	-0.09	0.25	0.15	-0.10	-0.35
BQt:p=12	0.18	0.17	-0.11	0.21	0.01	-0.14	-0.40

Source: Authors' calculations

Notes: BQ3 denotes SVAR based estimates using Blanchard-Quah identification in a VAR with 3 variables, namely output, inflation and unemployment. BQ2 denotes a VAR with just inflation and output. BQt is a (traditional) VAR with output and unemployment. KPSW denotes SVAR based estimates from a cointegrating VAR in output, investment and consumption; BiUC and TriUC are uncorrelated UC models in output, inflation and for tri also unemployment. CS denotes Cogley-Sargent. CorUC is the correlated UC model of Morley (2007) in output and consumption. p denotes lag order in the VAR

**Table 2:** Correlation of SVAR output gap estimates with each other and alternative decompositions (cont.)

	BQ3:p=12	BQ2:p=1 2	KPSW:p=1 2	CS	TriUC	BiUC	BQt:p=1	BQt:p=6	BQt:p=1 2
CorUC	-0.01	-0.16	-0.27	-0.36	0.00	-0.27	-0.35	-0.02	0.18
BQ3:p=1	0.36	0.20	-0.02	-0.23	-0.12	-0.24	-0.35	-0.01	0.17
BQ2:p=1	0.01	0.06	-0.03	0.51	0.90	0.79	0.27	-0.09	-0.11
KPSW:p=1	0.06	-0.05	-0.39	-0.17	-0.05	-0.01	-0.36	0.25	0.21
BQ3:p=6	0.41	-0.04	0.05	-0.07	-0.15	-0.12	-0.22	0.15	0.01
BQ2:p=6	-0.12	0.56	-0.01	0.28	0.61	0.43	0.21	-0.10	-0.14
KPSW:p=6	0.22	-0.08	0.71	0.62	0.44	0.56	0.20	-0.35	-0.40
BQ3:p=12	1.00	-0.16	0.07	0.17	0.13	0.13	0.14	-0.13	0.04
BQ2:p=12	-0.16	1.00	0.00	-0.15	-0.06	-0.29	0.02	0.12	-0.04
KPSW:p=12	0.07	0.00	1.00	0.59	0.07	0.31	0.33	-0.38	-0.46
CS	0.17	-0.15	0.59	1.00	0.65	0.85	0.79	-0.29	-0.26
TriUC	0.13	-0.06	0.07	0.65	1.00	0.82	0.41	-0.10	-0.09
BiUC	0.13	-0.29	0.31	0.85	0.82	1.00	0.55	-0.24	-0.23
BQtrad:p=1	0.14	0.02	0.33	0.79	0.41	0.55	1.00	-0.15	-0.12
BQtrad:p=6	-0.13	0.12	-0.38	-0.29	-0.10	-0.24	-0.15	1.00	0.11
BQtrad:p=12	0.04	-0.04	-0.46	-0.26	-0.09	-0.23	-0.12	0.11	1.00

Source: Authors' calculations

Table 3 then reports the correlation against the HP filter, still the most widely used reference. The highest correlation, of 0.70, is against the Cogley-Sargent output gap estimates. The HP filter also exhibits a strong correlation against the KPSW estimates when a lag order of 12 is selected (0.53), against the bivariate UC model (0.46) and against the traditional BQ estimates when a lag order of 1 is selected (0.49). This again reminds us that it is hard to classify the behaviour of output gap estimates according to whether the estimator is univariate or multivariate; we observe a high correlation of the HP filter against these multivariate estimators.



**Table 3:** A benchmark: correlation of the alternative estimates against the HP filter

	HP
CorUC	-0.45
BQ3:p=1	-0.12
BQ2:p=1	0.06
KPSW:p=1	0.14
BQ3:p=6	0.12
BQ2:p=6	-0.11
KPSW:p=6	0.57
BQ3:p=12	0.31
BQ2:p=12	-0.10
KPSW:p=12	0.53
CS	0.70
TriUC	0.29
BiUC	0.46
BQtrad:p=1	0.49
BQtrad:p=6	-0.23
BQtrad:p=12	-0.13

Source: Authors' calculations

Table 4 then provides some summary descriptive statistics on the properties of the alternative output gap estimates. It reports the mean and standard deviation of the estimates, alongside information on the steepness and average durations (in quarters) of expansions and contractions (described in the sub-section below), as well as reporting the proportion of time the output gap estimate is positive (i.e. the economy is in an expansionary phase of the business cycle).

Table 4 again indicates that inference is sensitive to measurement with the properties of the output gap dependent on which estimator is used.

### 6.3.1. The shape of the cycle: steepness

In particular in Table 4 we follow Harding & Pagan (2001) and Sichel (1993) and report the average duration in quarters of upturns (periods when the cycle,  $c_t$ , is positive) and downturns (periods when the cycle is negative). We also follow Sichel and report a measure of symmetry between upturns and downturns. Specifically we report the sample means of both positive and negative values of the cycle. As Harding & Pagan (2001) explain, these statistics can be interpreted as measures of “steepness”.

This is seen by noting that steepness can be defined algebraically as follows.

Let  $S_t = I(c_t > 0)$ . Then the sample mean of positive values of the cycle is given as

$$\text{Steepness (Expansions)} = \frac{\sum_{t=1}^T S_t c_t}{\sum_{t=1}^T S_t} \quad (19)$$

In turn the sample mean of negative values of the cycle is given as

$$\text{Steepness (Contractions)} = \frac{\sum_{t=1}^T (1-S_t) c_t}{\sum_{t=1}^T (1-S_t)}. \quad (20)$$

**Table 4:** Properties of the output gap estimates

	Mean	SD	PropE	Av Dur E	Av Dur C	SteepE	SteepC
CorUC	0.001	0.01	0.484	6.556	6.3	0.01	-0.008
BQ3:p=1	0	0.001	0.459	3.5	4.125	0.001	-0.001
BQ2:p=1	0	0.003	0.467	6.333	8.125	0.003	-0.002
KPSW:p=1	0.007	0.008	0.828	16.833	4.2	0.009	-0.004
BQ3:p=6	0	0.001	0.492	2.727	2.818	0.001	-0.001
BQ2:p=6	0.001	0.004	0.533	2.708	2.478	0.004	-0.003
KPSW:p=6	0.001	0.009	0.623	12.667	9.2	0.007	-0.009
BQ3:p=12	0	0.002	0.525	4.267	3.867	0.002	-0.002
BQ2:p=12	0	0.004	0.5	2.542	2.542	0.003	-0.004
KPSW:p=12	0.005	0.014	0.672	20.5	13.333	0.013	-0.011
CS	0.006	0.021	0.541	22	28	0.023	-0.014
TriUC	0.002	0.006	0.566	6.273	5.3	0.006	-0.003
BiUC	0.048	0.031	1	.	.	.	.
BQtrad:p=1	-0.005	0.007	0.377	46	76	0.003	-0.009
BQtrad:p=6	0	0.004	0.525	4	3.412	0.003	-0.004
BQtrad:p=12	0	0.005	0.467	2.714	3.095	0.004	-0.003
HP	0.001	0.01	0.459	6.222	8.25	0.009	-0.007

Source: Authors' calculations

Mean and SD are the mean and standard deviation of the output gap estimate. Av Dur E and Av Dur C are the average duration in quarters of expansionary and contractionary periods, namely when the output gap is positive or negative. SteepE and SteepC indicate the steepness of expansionary and contractionary periods

### 6.3.2. Correlation with inflation

We also report in Table 5 the full-sample correlation of the alternative output gap estimates with inflation. We believe that this offers one objective means to rank alternative output gap estimators. It also provides a means of integrating out model uncertainty via (Bayesian) model averaging with respect to each estimator's ability to explain inflation.

Table 5 reveals considerable variation in the ability of the different output gap estimates to explain inflation. The SVAR estimates appear no better and no worse than the other decomposition methods. The SVAR estimates themselves vary considerably in terms of their explanatory power. While the BQ decomposition (with inflation not unemployment), when  $p=1$ , is highly correlated at 0.8, the KPSW SVAR estimate when  $p=1$  is uncorrelated. In turn, the Cogley-Sargent estimate, and the bivariate and trivariate UC models (with uncorrelated innovations) are also highly correlated with inflation - but the correlated UC model yields output gap estimates with little explanatory power for inflation. Interestingly, the HP filter is less well correlated with inflation than these multivariate estimators (with a correlation coefficient of 0.16 only).

**Table 5: Correlation (r) with inflation**

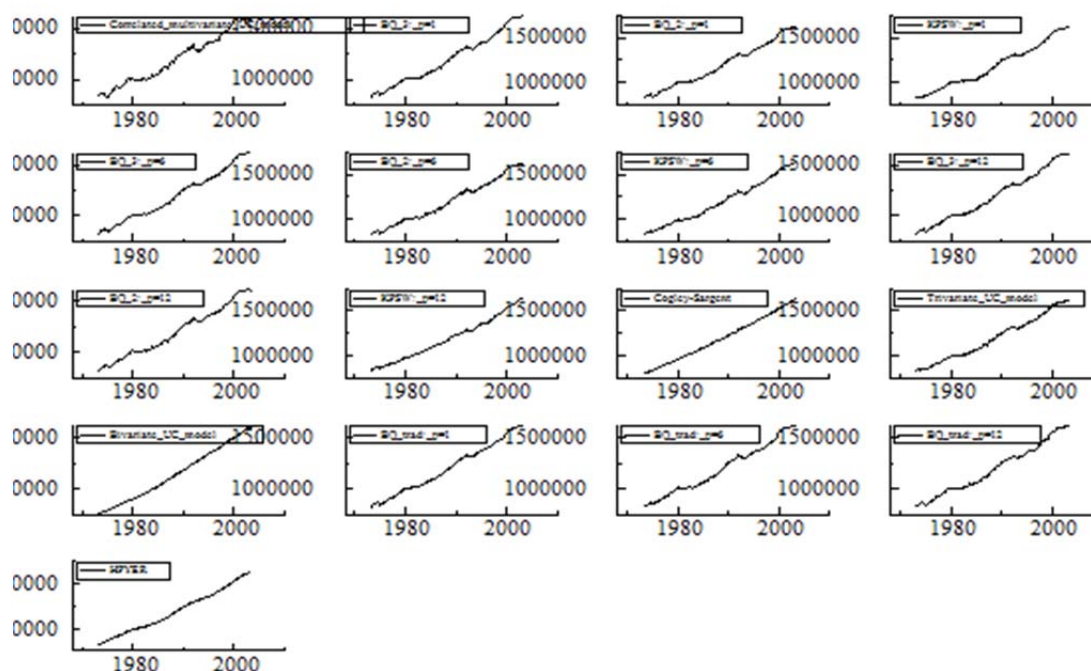
CorUC	0.019958
BQ3:p=1	-0.1813
BQ2:p=1	0.78941
KPSW:p=1	0.058165
BQ3:p=6	-0.14619
BQ2:p=6	0.4519
KPSW:p=6	0.38486
BQ3:p=12	0.037056
BQ2:p=12	-0.32861
KPSW:p=12	-0.01844
CS	0.59766
TriUC	0.81926
BiUC	0.86757
BQtrad:p=1	0.30423
BQtrad:p=6	-0.08418
BQtrad:p=12	-0.07333
HP	0.15572

Source: Authors' calculations

### 6.3.3. The permanent component of output

Figure 3 also provides an indication of the behaviour of the permanent component of output implied by each series. Unsurprisingly these trend upwards, and we therefore do not report the correlation between them as this would be distorted by the nonstationary nature of the series (spurious correlation). However, it is clear from Figure 3 that there is more agreement about the pattern of trend output than the output gap across the decomposition methods. Nevertheless, we can see that the permanent components do differ, especially around the boom and recession of the early 1990s. While some of the estimators attribute this to the trend, others do not. For example, the HP permanent component is not affected much, while many of the structural VAR estimates are. Interestingly, the Cogley-Sargent estimates are also little affected and exhibit a constant rate of growth (almost deterministic growth) over the sample period.

**Figure 3:** Euro area permanent output estimates generated using different estimators



Source: Authors' calculations

## 6.4. The reliability of real-time output gap estimates

To examine the reliability of real-time output gap estimates the following experiment is undertaken. Full sample or *final* estimates of the output gap are derived using data available over the sample-period, 1970q1-2007q4. Real-time output gap estimates are computed recursively from 1980q1. This involves using data from 1970q1-1980q1, to provide an initial estimation period of 20 years to compute the real time estimate for 1980q1. Then data for 1970q1-1980q2 are used to re-estimate the output gap (that involves re-estimation of the parameters of the models used to measure the output gap) and obtain real-time estimates for 1980q2. This recursive exercise, designed to mimic real-time measurement of the output gap, is carried on until data for the period 1970q1-2007q4 are used to estimate the real-time output gap for 2007q4.

Recursive out-of-sample simulations, from 1980q1-2007q4, are used to assess the real-time properties of the SVAR and HP measures of the output gap. We then compare the end-of-sample cyclical estimate, the so-called real-time estimate, with the final estimate. We do not leave any time to allow for the fact that real-time estimates take time to converge to their 'final' values. This should be expected to bias the results in favour of suggesting the real-time estimates are more reliable than in fact they are.

The real-time unreliability of output gap point estimates then is summarised in Table 6 for the different SVAR output gap estimates and the HP filter. The SVAR specifications include the traditional BQ model of output and unemployment, denoted BQ2, the trivariate specification including inflation, BQ3 and the King, Plosser, Stock & Watson (1991) specification, KPSW. In all cases the lag orders 1 and 4 in the SVAR models are considered; for the KPSW form, the lag order 12 is also computed.

Table 6 summarises the reliability of these real-time estimates by reporting their correlation with the final estimate. We also report their correlation against inflation. Table 6 shows that real-time estimates of the output gap vary in terms of their ability to track the final estimates, with correlation coefficients ranging from 0.55-0.88.

However, it is unclear whether we should prefer an output gap estimator with a higher correlation. An objective means to evaluate the quality of an output gap estimate is according to its ability to forecast

inflation; an output gap estimator that offers good forecasting power for inflation may or may not be estimated reliably in real-time.

Table 6 shows that the ability of the different output gap estimates to explain inflation varies considerably, with some estimates having a negative correlation and some a positive correlation, which we might expect as a higher output gap, via the Phillips Curve, should mean higher inflation.

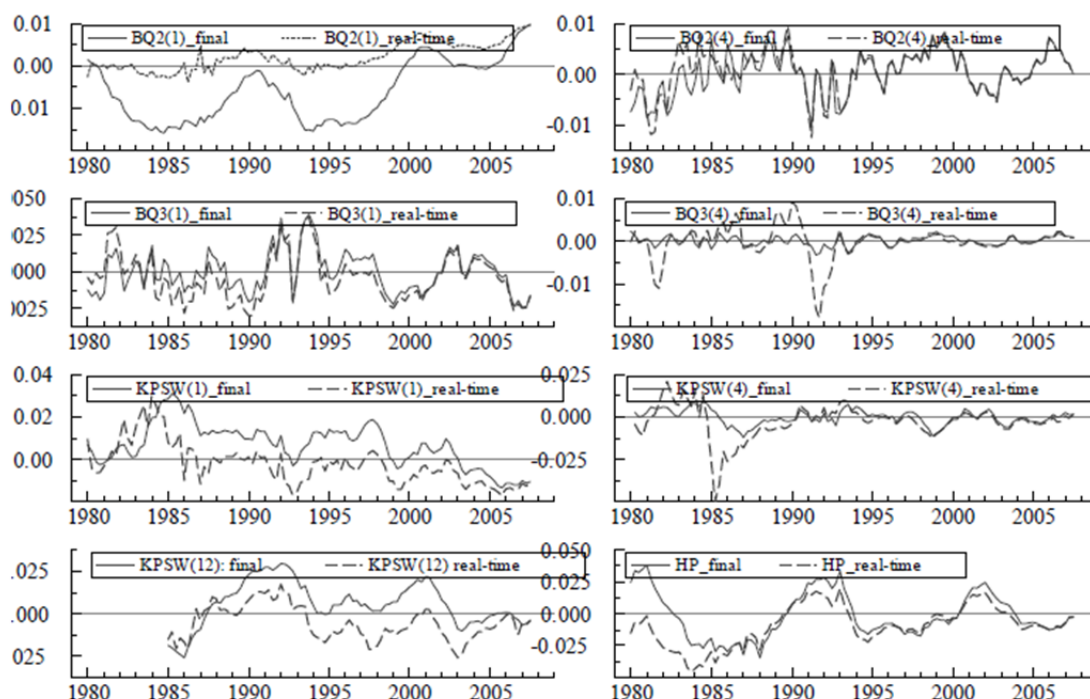
**Table 6:** Correlation of real time output gap estimates against the final estimate and inflation: 1980q1-2007q4

	Final estimate	Inflation
KPSW(1)	0.665	0.382
KPSW(4)	0.554	0.071
KPSW(12)	0.727	0.450
BQ2(1)	0.880	-0.264
BQ2(4)	0.861	-0.092
BQ3(1)	0.888	0.041
BQ3(4)	0.641	-0.065
HP	0.738	-0.070

Source: Authors' calculations

Figure 4 plots the real-time and final estimates of the output gap for the different detrending methods. Figure 4 illustrates that there is uncertainty associated with the real-time estimates (relative to the final estimate) not just for a given estimator of the output gap, but also across estimators, including a given SVAR estimator estimated using a different lag order. The different measures of the output gap can disagree considerably about the cyclical state of output.

**Figure 4:** Real-time and final output gap estimates for SVAR and HP measures. SVAR estimates are distinguished (in brackets) according to the lag order in the underlying VAR model

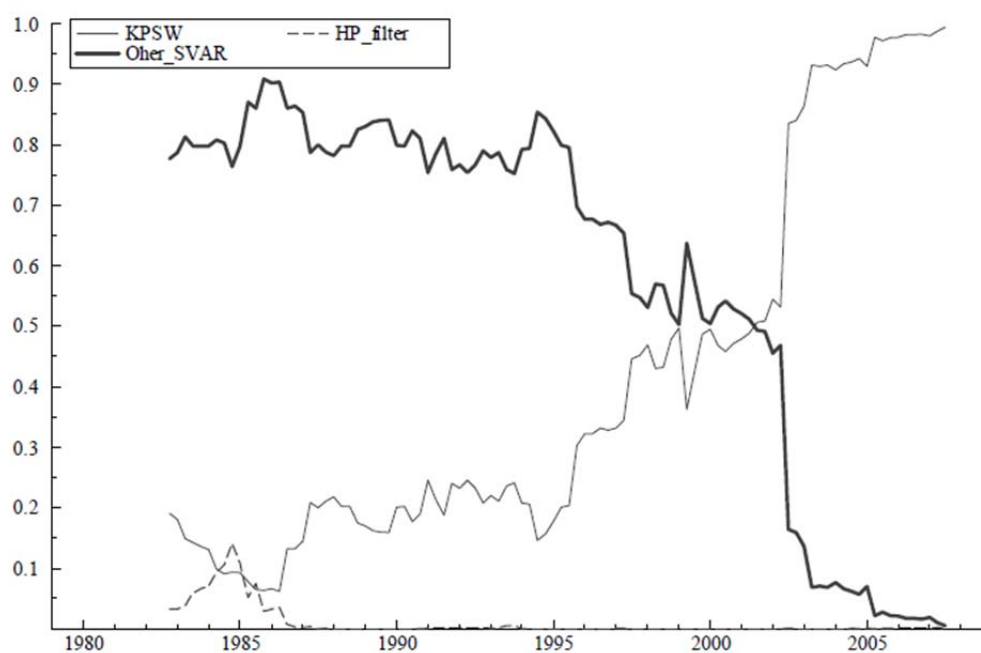


Source: Authors' calculations

## 6.5. Comparison of estimated weights

Figure 5 plots the estimated weights on the different measures of the output gap across time. Recall these are the weights that deliver the best density forecast of inflation (here 1-step ahead, i.e.  $h=1$ ). Figure 5 shows that the weights vary over time. The SVAR measures of the output gap receive a higher weight than the HP filter; in this sense the additional economic content of the SVAR measures of the output gap appears to be paying off. By the end of the evaluation period, 2007q4, the HP filter receives no weight at all: it is best left out of the combination. The KPSW measures of the output gap are found to be the most informative of the SVAR measures, receiving an increasing share of the weight throughout the evaluation period. Importantly, the combined measure of the output gap delivers density forecasts of inflation that have a higher logarithmic score than both an equal weighted combination and any of the individual output gap estimates.

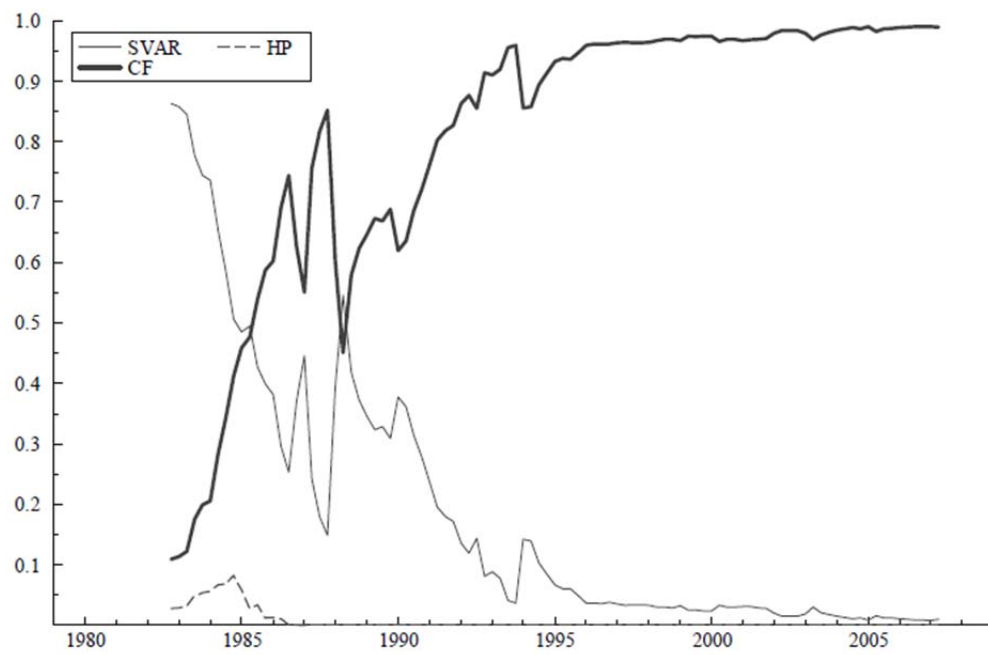
**Figure 5:** Recursive weights on SVAR and HP measures of the output gap



Source: Authors' calculations

Finally, we note that the performance of a combined forecast is always dependent on the set of models over which one combines. If we augment our model space to include the Christiano-Fitzgerald univariate filter we no longer find that the SVAR models dominate; see Figure 6.



**Figure 6:** Recursive weights on SVAR, HP and CF measures of the output gap

Source: Authors' calculations

## 7. Conclusion

We have reviewed the SVAR methodology for estimating the output gap. Then we have carried out an application to the euro area dataset. The SVAR estimates have been compared with several alternative methods used as benchmarks.

It has been shown that in SVAR models potential output has a random walk component but also an additional stationary component, reflecting the short-run dynamics of the permanent shocks. This is widely seen to be more attractive than assuming potential output follows a random walk alone; such an assumption is common in traditional multivariate UC models and multivariate BN decompositions. However in correlated multivariate UC models, as seen in Morley (2007), Basistha & Nelson (2007) and the STAMP 8 manual pp. 119-124, potential output is no longer restricted to consist of only a random walk component.

To provide an economic interpretation to SVAR based output gap estimates we have discussed the relationship with DSGE models and how their linearised solution can, under certain conditions, be approximated by a (restricted) finite order VAR model; see Del Negro & Schorfheide (2004). We presented also the methodology by Garratt, Mitchell & Vahey (2009) to integrate out model uncertainty and arrive at a combined density estimate of the output gap.

In the empirical application we have found that the SVAR estimates of the output gap deliver neither more informative nor less informative indicators of inflation than the other decomposition methods. Indeed there is considerable disparity across the decomposition methods, with some positively correlated with inflation, as we should expect, and some negatively correlated. The multivariate approaches with inflation do deliver better indicators of inflation than the benchmark HP filter, supporting the view that economic theory does, to some degree at least, help. On output gap uncertainty, we found from one side that SVAR measures dominate the HP filter and from another that performance of combined forecasts is always a function of the set of models under consideration.

Finally we stressed how recursive-weight density combination of different output gap estimates delivers improved density forecasts of inflation and provides a means of integrating out uncertainty about the appropriate measure of the output gap.

## References

- Apel, M. & Jansson, P. (1999), 'A theory-consistent approach for estimating potential output and the NAIRU', *Economics Letters* 64, 271–275.
- Astley, M. & Yates, T. (1999), *Inflation and real disequilibria*. Bank of England Working Paper.
- Attfield, C. & Temple, J. (2003), *Measuring trend output - how useful are the great ratios?* Bristol University Working Paper.
- Basistha, A. & Nelson, C. R. (2007), 'New measures of the output gap based on the forward-looking new Keynesian Phillips curve', *Journal of Monetary Economics* 54, 498–511.
- Bernanke, B. (1986), 'Alternative explorations of the money-income correlation', *Carnegie Rochester Conference Series in Public Policy* 25, 49–100.
- Blanchard, O. J. & Watson, M. (1986), *Are business cycles all alike?*, in R. Gordon, ed., 'The American Business Cycle', Chicago, IL: University of Chicago Press.
- Blanchard, O. & Quah, D. (1989), 'The dynamics of aggregate demand and supply disturbances', *American Economic Review* 79, 655–673.
- Camba-Mendez, G. & Rodriguez-Palenzuela, D. (2003), 'Assessment criteria for output gap estimates', *Economic Modelling* 20, 529–562.
- Cogley, T. & Sargent, T. J. (2007), *Inflation gap persistence in the U.S.* Stanford University.
- Del Negro, M. & Schorfheide, F. (2004), 'Priors from general equilibrium models for VARS', *International Economic Review* 45, 643–673.
- DeSerres, A. & Guay, A. (1995), *Selection of the truncation lag in structural VARs (or VECMs) with long run restrictions*. Bank of Canada Working Paper 95-9.
- Dupasquier, C., Guay, A. & St-Amant, P. (1999), 'A survey of alternative methodologies for estimating potential output and the output gap', *Journal of Macroeconomics* 21, 577–595.
- Fabiani, S. & Mestre, R. (2001), *A system approach for measuring the Euro area NAIRU*. European Central Bank Working Paper No. 65.
- Fagan, G., Henry, J. & Mestre, R. (2005), 'An area-wide model (AWM) for the Euro area', *Economic Modelling* 22, 39–59.
- Faust, J. & Leeper, E. M. (1997), 'When do long run identifying restrictions give reliable results?', *Journal of Business and Economic Statistics* 15, 345–353.
- Fernández-Villaverde, J., Rubio-Ramírez, J. F. & Sargent, T. J. (2005), *A,b,c's (and d's)'s for understanding vars*, Levine's Bibliography, UCLA Department of Economics. available at <http://ideas.repec.org/p/cla/levrem/172782000000000096.html>.
- Garratt, A., Mitchell, J. & Vahey, S. (2009), *Measuring output gap uncertainty*. Unpublished manuscript, Birkbeck College, London, April.
- Garratt, A., Robertson, D. & Wright, S. (2006), 'Permanent vs transitory components and economic fundamentals', *Journal of Applied Econometrics* 21, 521–542.
- Granger, C. W. J. & Pesaran, M. H. (2000), 'Economic and statistical measures of forecast accuracy', *Journal of Forecasting* 19, 537–560.
- Harding, D. & Pagan, A. (2001), *Extracting, analysing and using cyclical information*. Mimeo, University of Melbourne.
- Harvey, A. C. & Jaeger, A. (1993), 'Detrending, stylized facts and the business cycle', *Journal of Applied Econometrics* 8, 231–247.
- Jore, A., Mitchell, J. & Vahey, S. (2009), 'Combined forecast densities from vars uncertain instabilities', *Journal of Applied Econometrics*. Forthcoming.

- King, R. G., Plosser, C., Stock, J. & Watson, M. (1991), 'Stochastic trends and economic fluctuations', *American Economic Review* 81, 819–840.
- Mitchell, J. (2007), Understanding revisions to density forecasts: an application to the Survey of Professional Forecasters. Discussion Paper No. 296, NIESR.
- Morley, J. C. (2007), 'The slow adjustment of aggregate consumption to permanent income', *Journal of Money, Credit and Banking* 39, 615–638.
- Morley, J., Nelson, C. R. & Zivot, E. (2003), 'Why are Beveridge-Nelson and unobserved component decompositions of GDP so different?', *Review of Economics and Statistics* 85, 235–243.
- Neiss, K. S. & Nelson, E. (2005), 'Inflation dynamics, marginal cost, and the output gap: evidence from three countries', *Journal of Money, Credit and Banking* 37, 1019–1045.
- Orphanides, A. & van Norden, S. (2005), 'The reliability of inflation forecasts based on output gap estimates in real time', *Journal of Money, Credit and Banking* 37, 583–601.
- Proietti, T. (1997), 'Short-run dynamics in cointegrated systems', *Oxford Bulletin of Economics and Statistics* 59(3), 405–422.
- Proietti, T. (2002), Some reflections on trend-cycle decompositions with correlated components. EUI Working Paper 2002/23.
- Ravenna, F. (2007), 'Vector autoregressions and reduced form representations of DSGE models', *Journal of Monetary Economics* 54, 2048–2064.
- Robertson, D. & Wickens, M. R. (1997), 'Measuring real and nominal macroeconomic shocks and their international transmission under different monetary regimes', *Oxford Bulletin of Economics and Statistics* 59, 5–27.
- Sichel, D. E. (1993), 'Business cycle asymmetry', *Economic Inquiry* 31, 224–236.
- Stock, J. H. & Watson, M. W. (1988), 'Testing for common trends', *Journal of the American Statistical Association* 83(404), 1097–1107.
- Zellner, A. (1971), *An Introduction to Bayesian Inference in Econometrics*, J. Wiley and Sons, Inc., New York.

## HOW TO OBTAIN EU PUBLICATIONS

### Free publications:

- one copy:  
via EU Bookshop (<http://bookshop.europa.eu>);
- more than one copy or posters/maps:  
from the European Union's representations ([http://ec.europa.eu/represent\\_en.htm](http://ec.europa.eu/represent_en.htm));  
from the delegations in non-EU countries ([http://eeas.europa.eu/delegations/index\\_en.htm](http://eeas.europa.eu/delegations/index_en.htm));  
by contacting the Europe Direct service ([http://europa.eu/europedirect/index\\_en.htm](http://europa.eu/europedirect/index_en.htm)) or  
calling 00 800 6 7 8 9 10 11 (freephone number from anywhere in the EU) (\*).

(\*) The information given is free, as are most calls (though some operators, phone boxes or hotels may charge you).

### Priced publications:

- via EU Bookshop (<http://bookshop.europa.eu>).

# Structural vector autoregressive (SVAR based estimates of the euro area output gap: theoretical considerations and empirical evidence

GIAN LUIGI MAZZI, JAMES MITCHELL  
AND FILIPPO MOAURO

This paper analyses alternative constraints in Structural VAR models allowing for the decomposition between permanent and transitory components. Alternative SVAR specifications are then compared in a real-time exercise using the HP filter as a benchmark.

---

**For more information**

**<http://ec.europa.eu/eurostat/>**

