BAYESIAN ANALYSIS OF OUTPUT GAP *

C. Planas[(a)], A. Rossi[(a)] and G. Fiorentini[(b)]

July 2004

Abstract

We develop a Bayesian analysis of the bivariate Phillips-curve model proposed by Kuttner (1994) for estimating output gap and potential output. The Bayesian extension is of interest because it enables the gap estimate to incorporate the information available from Phillips curve theory and from business cycle knowledge. For given priors, we implement a Gibbs sampling scheme for drawing model parameters and state vector from their joint posterior distribution. We sample the state conditionally on parameters using the Carter and Kohn (1994) sampler. When sampling parameters given the state, we introduce a Metropolis-Hastings step that removes the conditioning on starting values. We also reparametrize the traditional cyclical AR(2) model in terms of polar coordinates of the polynomial roots in order to elicit a sensible prior on cycle periodicity and amplitude. We then use the Gilks et al. (1995) adaptive rejection Metropolis sampling for drawing periodicity and amplitude from their full conditional distributions. We illustrate our approach with two applications, namely the analysis of output gap in US and EU.

Keywords: Kalman filter and smoothing, MCMC algorithms, priors and parameterizations, unobserved components, Gibbs sampling.

*The ideas expressed here are those of the authors and do not necessarily reflect the positions of the European Commission. Thanks are due to Werner Roeger for useful comments and suggestions. Gabriele Fiorentini acknowledges financial support from MURE.

(a) Joint Research Centre of the European Commission.

(b) University of Florence, Department of Statistics.

E-Mails: *Christophe.Planas@jrc.it*, *Alessandro.Rossi@jrc.it*, *Fiorentini@ds.unifi.it*, .

1 Introduction

In this paper we develop a Bayesian analysis of the Phillips-curve bivariate model put forward by Kuttner (1994) for estimating potential output and output gap. Potential output and output gap are two concepts that are essential to macroeconomic analysis because they are related to different time horizon of the economic dynamics: potential output measures long-term movements associated to the economic growth while output gap captures all short-term fluctuations (see for instance Hall and Taylor, 1991). Accordingly institutions responsible for stabilisation policy monitor the gap and, as the Taylor rule reflects (Taylor, 1993), monetary authorities in charge of inflation control follow its evolution.

Kuttner (1994) original model relates the gap to inflation through a Phillips curve equation. Although somewhat atypical as it involves a regression on an unobserved variable, this model has entertained a certain success. Kichian (1999) for instance applied it to the G7 countries case, Gerlach and Smets (1999) emphasized its appeal for the European Central Bank policy-making process and Apel and Jansson (1999a, 1999b) further extended it to include unemployment. The European Commission uses a similar specification for estimating structural unemployment (see Planas et al., 2003). OECD considers a closely related version for estimating the NAIRU (see OECD, 2000).

In this particular context of cycle estimation embedded into a Phillips curve regression, an important amount of macroeconomic theory and of business cycle knowledge is available. It is thus natural to develop Bayesian tools for incorporating this information into the decomposition. Several advantages result. First, since Bayesian analysis delivers samples from posterior distributions, the finite-sample uncertainty around any quantity of interest can be precisely delineated. This is an important information. Output gap measurements have indeed been strongly criticised by Orphanides and Van Norden (2002) for lacking reliability and knowing uncertainty has become imperative for practitioners. Second, researchers seeking uncertainty reduction through theoretical advances or extra information can use the Bayesian framework in order to evaluate every possible improvement. Third, the Bayesian setting helps in understanding salient features of Phillips curve regressions: for instance the sharpness of the response of inflation to different gap proxies can be compared (see Gali et al., 2001). Fourth, by properly tuning the prior distributions the pile-up problem sometimes faced in classical analysis can be totally avoided. This problem, that consists in obtaining zero variance for the innovations in unobservables even though the true variance is strictly positive (see the simulation study in Stock and Watson, 1998), is generally undesired because the related variable turns out deterministic and hence observed.

Given priors on model parameters, we implement a Gibbs sampling scheme for drawing model parameter and state vectors from their joint posterior distribution (see Casella and George, 1992; also the more general discussion in Geweke, 1999). We sample the state conditionally on parameters on the basis of Carter and Kohn (1994) sampler with initialization handled as in de Jong (1991). Use is made of results in Koopman (1997) for sampling the state in its first time-period. When sampling parameters given the state we introduce a Metropolis-Hastings step (see for instance Chib and Greenberg, 1995) for removing the conditioning on the first observations. We also reparametrize the traditional cyclical AR(2) model in terms of the polar coordinates of the polynomial roots. We find this necessary mainly because specifying non-informative priors for autoregressive parameters has implications for the periodicity that make the AR specification unsuited to Bayesian cyclical analysis. We then resort to the adaptive rejection Metropolis method proposed by Gilks et al. (1995) for sampling the polar coordinates.

Section 2 discusses the model structure, the parametrization adopted and the prior distributions. Section 3 solves the problem of sampling from the joint posterior distribution of the model parameters and unobserved variables. We apply this methodology in Section 4 to the cases of the US and EU-11 economies. Section 5 concludes.

2 Model specification

Let y_t denote the logarithm of real output. Like Watson (1986) and Clark (1987) we assume that it is made up of a trend, p_t, and of a cycle, c_t, according to:

$$\begin{aligned} y_t &= p_t + c_t \\ (1-L)p_t &= \mu_p + a_{pt} \\ \phi(L)c_t &= a_{ct} \end{aligned} \tag{2.1}$$

where L is the lag operator, μ_p is a constant drift and $\phi(L) = 1 - \phi_1 L - \phi_2 L^2$ is an AR(2) polynomial with stationary and complex roots. The permanent and transitory shocks,

a_{pt} and a_{ct}, are independent Gaussian white noises with variances V_p and V_c. The long-term and short-term components, p_t and c_t, are generally interpreted as the potential output and the output gap. Kuttner (1994) complemented (2.1) with an equation that links dynamically output gap to the change in inflation, say π_t, like in:

$$\pi_t = \mu_\pi + \beta c_{t-1} + \lambda \Delta y_{t-1} + \alpha_1 \pi_{t-1} + \alpha_2 \pi_{t-2} + a_{\pi t}, \tag{2.2}$$

where $\Delta = 1 - L$ and $a_{\pi t}$ is a Gaussian white noise with variance V_π. The AR(2) polynomial $\alpha(L) = 1 - \alpha_1 L - \alpha_2 L^2$ is assumed to be stationary. Equation (2.2) introduces a Phillips curve effect by relating the change in inflation to the real output growth and to the latent cycle, both with one lag. Also, a correlation between the shocks in inflation and in output gap is allowed. This adds a contemporaneous link between inflation and output gap through their unpredictable elements. We shall denote $V_{c\pi}$ the covariance between a_{ct} and $a_{\pi t}$. Phillips curve theory expects it to be positive. Kuttner's original model included a moving average polynomial instead of the lagged change in inflation terms that appear in (2.2). We make this slight modification for simplifying the statistical analysis.

The model is completed by specifying the prior probability density function for the model parameters. Thanks to the bulk of business cycles studies it is nowadays generally admitted that typical cycle lengths in G7 countries lie between 2 and 10 years. In order to incorporate this information we parameterize the model for the cycle as in:

$$(1 - 2A\cos(2\pi/\tau)\ L + A^2 L^2) c_t = a_{ct} \tag{2.3}$$

where the parameters A and τ represent the amplitude and periodicity of the cyclical movements, respectively. By construction these parameters describe cycles much more naturally than AR coefficients. Indeed, assuming a normal distribution for (ϕ_1, ϕ_2), we found it very difficult to tune the mean and covariance matrix of the autoregressive parameters in order to reproduce our prior knowledge on the cycle periodicity. And in some cases the implied distribution for the periodicity can be counter-intuitive. Let us consider for instance the non informative setting $(\phi_1, \phi_2) \sim N(0_2, 10 \times Id_2) I_S$, where Id_2 is the 2×2 identity matrix and I_S imposes stationary and complex roots, i.e. $\phi_2 \in (-1, 0)$ and $\phi_1^2 + 4\phi_2 < 0$. Given that $A = \sqrt{-\phi_2}$ and $\tau = 2\pi/acos\{\phi_1/(2A)\}$, the joint distribution of τ and A can be easily derived. Figure 1 shows the cumulative marginal distribution of τ that the almost flat prior on the stationary region of complex roots implies. As can be seen, the median value is 4 and the intervals $[2, 4]$ and $[4, \infty)$ receive equal weight. This happens because the half of the complex region where ϕ_1 is negative yields roots with periodicity in $[2, 4]$, while the other half yields roots with periodicity in $[4, \infty)$. A similar reasoning suggests that the flat prior on the complex area overweights close-to-one amplitudes. Hence, being non-informative on the coefficients of AR polynomials with complex roots amounts to put a strong emphasis on short-term and persistent fluctuations. Putting the prior on the AR coefficients like for instance in Chib (1993), Chib and Greenberg (1994) or Mc Culloch and Tsay (1994), or on the partial autocorrelations like in Barnett et al. (1996) and Billio et al. (1999), can thus be inadequate for cyclical analysis. The trigonometric specification proposed by Harvey (1981, p.182-183) and used in a Bayesian setting in Harvey et al. (2002) appears instead better suited. Here we choose to stick as much as possible to Kuttner's original model by considering the polar coordinates parametrisation (2.3). This parametrisation actually simplifies Harvey's specification by excluding moving average terms. For a similar discussion about the link between prior and parametrizations in the local level model, see Appendix C of Koop and Van Dijk (2000).

We can now set all prior distributions. Let δ and Σ denote the vector and matrix of parameters defined by:

$$\delta = (\mu_\pi, \beta, \lambda, \alpha_1, \alpha_2), \qquad \Sigma = \begin{bmatrix} V_c & V_{c\pi} \\ & V_\pi \end{bmatrix}$$

We shall consider:

$$\begin{aligned} p(A) &= B(a_A, b_A) \\ p(\frac{\tau - \tau_l}{\tau_u - \tau_l}) &= B(a_\tau, b_\tau) \\ p(\delta) &= N_k(\delta_0', M_\delta^{-1}) I_\delta \\ p(\Sigma) &= IW_2(\Delta_0, d_0) \\ p(\mu_p, V_p) &= NIG(\mu_0, M_p^{-1}, s_0, v_0) \end{aligned} \tag{2.4}$$

where $B(.,.)$ is the Beta distribution, τ_l and τ_u are the lower and upper bound of the support for τ, $N_k(\cdot)$ is the k-variate normal distribution, I_δ is an index set imposing constraints on parameters, $IW_2(\cdot)$ is the bivariate inverted-Wishart distribution and $NIG(\cdot)$ is the Normal-inverted Gamma distribution. The hyperparameters τ_l, τ_u, a_A, b_A, a_τ, b_τ, δ_0 (5×1), M_δ (5×5), Δ_0 (2×2), d_0, μ_0, M_p, s_0, and v_0 are supposed given; we shall discuss how we set them in the application. In (2.4) the prior distributions for δ, Σ, μ_p and V_p are naturally conjugate for the full conditionals of interest. Computational convenience is the main reason of this choice. The framework it ensues remains however quite flexible since we can be as non-informative as desired by properly tuning the hyperparameters: for instance, setting M_p, s_0, and v_0 to small quantities leads to $p(\mu_p, V_p) \propto 1/V_p$. The conjugate property is instead lost for the parameters A and τ. We will pay this price for the parameterization (2.3) in terms of computational complexity. Yet (2.3) is quite worth this cost as it is far better suited to cyclical analysis than the traditional AR(2) representation. Equations (2.4) imply a joint prior distribution such as:

$$p(A, \tau, \delta, \Sigma, \mu_p, V_p) = p(A)p(\tau)p(\delta)p(\Sigma)p(\mu_p, V_p)$$

so a block-independence structure holds.

Let θ denote the full set of parameters, i.e. $\theta \equiv (A, \tau, \delta, vech(\Sigma), \mu_p, V_p)$. Our objective is to characterise the joint posterior distribution of potential output, output gap and of the parameters conditionally on the data, i.e. $p(c^T, p^T, \theta | Y^T)$ where $Y^T \equiv \{y^T, \pi^T\}$ and, in general, $x^T \equiv \{x_1, \cdots, x_T\}$. Given model (2.1)-(2.3), no close form expression for this posterior is available but draws from $p(c^T, p^T, \theta | Y^T)$ can be obtained using a Gibbs sampling scheme. The full conditionals of interest are:

- $p(c^T, p^T | \theta, Y^T)$;
- $p(\theta | c^T, p^T, Y^T)$.

We explain in the next section how to sample from these two distributions.

3 Joint posterior distribution of states and model parameters

3.1 Sampling the state given model parameters

We first focus on simulating the unobservable components conditionally on model parameters. It will be useful to cast model (2.1)-(2.3) into a state space format (see for instance Durbin and Koopman, 2001) such that:

$$\begin{aligned} Y_t &= H\xi_t \\ \xi_{t+1} &= D + F\xi_t + w_{t+1}, \end{aligned}$$

where $Y_t = (y_t, \pi_t)'$ is the vector of observations, $\xi_t = (p_t, c_t, c_{t-1}, a_{\pi t})'$ is the state vector, $w_t = (a_{pt}, a_{ct}, 0, a_{\pi t})'$ is a Gaussian error vector with zero mean and singular variance matrix Q. The time-invariant matrices H, D, F and Q can be straightforwardly recovered. Like usual, $\xi_{t|k}$ and $P_{t|k}$ will denote the conditional expectation $E(\xi_t|Y^k)$ and conditional variance $Var(\xi_t|Y^k)$. Samples from $p(c^T, p^T|\theta, Y^T)$ will be obtained through $p(\xi^T|\theta, Y^T)$. We make use of the following identity that gives the basis of the Carter and Kohn (1994) state sampler:

$$p(\xi^T|\theta, Y^T) = p(\xi_T|\theta, Y_T) \prod_{t=2}^{T-1} p(\xi_t|\theta, Y^t, \xi_{t+1}) p(\xi_1|\theta, Y_1, \xi_2)$$

where the last term is isolated because initial conditions need a special treatment. A draw from $p(\xi^T|\theta, Y^T)$ can be obtained as follows:

(i) compute $\xi_{t|t}$, and $P_{t|t}$, $t = 2, \cdots, T$, via the diffuse Kalman filter (de Jong, 1991);

(ii) given $\xi_{T|T}$ and $P_{T|T}$, sample ξ_T from $p(\xi_T|\theta, Y^T) = N(\xi_{T|T}, P_{T|T})$;

(iii) for $t = T-1$ to $t = 2$, sample backward ξ_t from $p(\xi_t|\theta, \xi_{t+1}, Y^t) = N(E[\xi_t|\theta, \xi_{t+1}, Y^t],$ $V[\xi_t|\theta, \xi_{t+1}, Y^t])$.

(iv) sample ξ_1 from $p(\xi_1|\theta, \xi_2, Y_1) = N(E[\xi_1|\theta, \xi_2, Y_1], V[\xi_1|\theta, \xi_2, Y_1])$.

Steps (i) and (ii) only involve classical results. Step (iii) needs the conditional moments $E[\xi_t|\theta, \xi_{t+1}, Y^t]$ and $V[\xi_t|\theta, \xi_{t+1}, Y^t])$. From the joint distribution of ξ_t and ξ_{t+1} conditional on θ and Y^t we get:

$$\begin{aligned} E[\xi_t|\theta, \xi_{t+1}, Y^t] &= \xi_{t|t} + P_{t|t}F'P_{t+1|t}^{-1}(\xi_{t+1} - F\xi_{t|t}) \\ V[\xi_t|\theta, \xi_{t+1}, Y^t]) &= P_{t|t} - P_{t|t}F'P_{t+1|t}^{-1}FP_{t|t}. \end{aligned}$$

Step (iv) is more complicated. For $t = 1$, the formula above involves $\xi_{1|1}$ and $P_{1|1}$ but none of them are available in our model. This occurs because if d is the integration order, the first state that de Jong's algorithm yields is $\xi_{d+1|d}$, i.e. $\xi_{2|1}$ for (2.1)-(2.3). A procedure based on Koopman (1997) for obtaining $\xi_{1|1}$ and $P_{1|1}$ is detailed in Appendix A1. It is true that in our particular context use could be made of the fact that ξ_2 contains c_1 for skipping the sampling of $\xi_{1|1}$, but such a simplification only hold when the trend integration order is 1. The algorithm given in Appendix A1 has the advantage of generality.

Because of the model structure, not all elements of the state need to be simulated. Trivially, given y_t knowledge of c_t determines p_t. We thus end up with two random elements to simulate, c_t and $a_{\pi t}$. In such a low dimension context, using a more efficient simulation smoother as proposed by de Jong and Shepard (1995) and Durbin and Koopman (2002) instead of a state sampler should not give relevant advantages.

3.2 Sampling model parameters given the state

We now turn to the second full conditional distribution, $p(\theta|c^T, p^T, Y^T)$. The structure of model (2.1)-(2.3) together with the block-independence assumption about the parameters priors imply that the posterior density $p(\theta|c^T, p^T, Y^T) \equiv p(\theta|c^T, p^T, \pi^T)$ can be factorised as:

$$p(A, \tau, \delta, \Sigma, V_p, \mu_p|c^T, p^T, \pi^T) = p(A, \tau, \delta, \Sigma|c^T, p^T, \pi^T)p(V_p, \mu_p|p^T)$$

We first consider the conditional $p(V_p, \mu_p|p^T)$. As detailed in Bauwens et al. (1999, p.58), choosing the Normal-inverse-gamma conjugate prior leads to:

$$p(\mu_p, V_p|p^T) = NIG(\mu_{p*}, M_{p*}^{-1}, s_*, \nu_*)$$

where

$$
\begin{aligned}
M_{p*} &= M_p + T - 1 \\
\mu_{p*} &= M_{p*}^{-1}[M_p\mu_0 + (T-1)\widehat{\mu}_p] \\
\nu_* &= \nu_0 + T - 1 \\
s_* &= s_0 + \frac{M_p(T-1)}{T+M_p-1}(\mu_0 - \widehat{\mu}_p)^2 + \sum_{t=2}^{T}(p_t - p_{t-1})^2 - (T-1)\widehat{\mu}_p^2
\end{aligned}
$$

and $\widehat{\mu}_p = \frac{1}{T-1}\sum_{t=2}^{T}(p_t - p_{t-1})$.

Focusing next on $p(A,\tau,\delta,\Sigma|c^T,p^T,\pi^T)$, we consider the full conditionals $p(\Sigma|A,\tau,\delta,c^T,$ $p^T,\pi^T)$ and $p(A,\tau,\delta,|\Sigma,c^T,\ p^T,\pi^T)$. For the first conditional $p(\Sigma|A,\tau,\delta,c^T,p^T,\pi^T)$, we can write:

$$
\begin{aligned}
p(\Sigma|A,\tau,\delta,c^T,p^T,\pi^T) \;\propto\;& p(\pi_1,\pi_2,c_1,c_2|A,\tau,\delta,\Sigma) \\
&\times \prod_{t=3}^{T} p(\pi_t,c_t|A,\tau,\delta,\Sigma,\pi^{t-1},c^{t-1})p(\Sigma)
\end{aligned} \tag{3.1}
$$

Following Bayesian textbooks (see Box and Tiao, 1973, Chap.8; also Geweke, 1995) the product of the last two terms above is such that:

$$
\prod_{t=3}^{T} p(\pi_t,c_t|A,\tau,\delta,\pi^{t-1},c^{t-1},p^T)p(\Sigma) \propto IW_2(\Delta_*,d_*)
$$

with $d_* = d_0 + T - 2$ and $\Delta_* = \Delta_0 + U'U$, U being the $(T-2)\times 2$ matrix of innovations with rows $[a_{\pi t}\ a_{ct}]$. We thus implement a Metropolis-Hastings scheme for getting draws from $p(\Sigma|A,\tau,\delta,c^T,p^T,\pi^T)$ from the proposal distribution $IW_2(\Delta_*,d_*)$. For each candidate $\tilde{\Sigma}$, the acceptance probability is given by:

$$
\min\{1, p(\pi_1,\pi_2,c_1,c_2|A,\tau,\delta,\tilde{\Sigma},p^T)/p(\pi_1,\pi_2,c_1,c_2|A,\tau,\delta,\Sigma,p^T)\}.
$$

Since the innovations $a_{\pi t}$ and a_{ct} are assumed Gaussian, the first two moments of (π_1,π_2,c_1,c_2) are needed for evaluating the acceptance probability. We detail in Appendix A2 a Yule-Walker procedure for computing them.

Next, samples must be obtained from $p(A, \tau, \delta, |\Sigma, c^T, p^T, \pi^T)$. This conditional verifies:

$$\begin{aligned} p(A, \tau, \delta, |\Sigma, c^T, p^T, \pi^T) &\propto p(c^T, \pi^T|A, \tau, \delta, \Sigma, p^T)p(A)p(\tau)p(\delta) \\ &= p(\pi^T|A, \tau, \delta, \Sigma, c^T, p^T)p(\delta) \times \\ &\times p(c^T|A, \tau, \delta, \Sigma, p^T)p(A)p(\tau) \end{aligned}$$

It can easily be seen that $p(c^T|A, \tau, \delta, \Sigma, p^T) = p(c^T|A, \tau, \Sigma)$. Hence the joint density $p(A, \tau, \delta, |\Sigma, c^T, p^T, \pi^T)$ can be split as:

$$p(A, \tau, \delta|\Sigma, c^T, p^T, \pi^T) = p(\delta|A, \tau, \Sigma, \pi^T, c^T, p^T) \times p(A, \tau|\Sigma, c^T) \tag{3.2}$$

Draws from joint distribution $p(A, \tau|\Sigma, c^T)$ above will be obtained using the conditionals:

$$p(A|\tau, \Sigma, c^T) \propto p(c_1, c_2|A, \tau, V_c) \prod_{t=3}^{T} p(c_t|c^{t-1}, A, \tau, V_c)\ p(A) \tag{3.3}$$

and

$$p(\tau|A, \Sigma, c^T) \propto p(c_1, c_2|A, \tau, V_c) \prod_{t=3}^{T} p(c_t|c^{t-1}, A, \tau, V_c)\ p(\tau) \tag{3.4}$$

Sampling directly from (3.3)-(3.4) is not possible but both densities are straightforward to evaluate. Given that log-concavity cannot be insured, we use the adaptive rejection Metropolis scheme proposed by Gilks et al. (1995).

Finally, it remains to sample δ from $p(\delta|A, \tau, \Sigma, c^T, p^T, \pi^T)$. We can write:

$$\begin{aligned} p(\delta|A, \tau, \Sigma, c^T, p^T, \pi^T) &\propto p(\pi_1, \pi_2|A, \tau, \delta, V_\pi, V_{c\pi}, c^T, p^T) \times \\ &\times \prod_{t=3}^{T} p(\pi_t|A, \tau, \delta, V_\pi, V_{c\pi}, \pi^{t-1}, c^T, p^T)p(\delta) \end{aligned}$$

The product of the last two terms in the factorization above can be recognized as proportional to the posterior distribution of δ for given starting values π_1 and π_2 in a regression on the elements of the information set. Let Z denote the $(T-2) \times 5$ matrix of regressors with row $z_t = (1, c_{t-1}, \Delta y_{t-1}, \pi_{t-1}, \pi_{t-2})$. Since A, τ, and c^T altogether make available the innovations a_{ct}, we have $E[\pi_t|\pi^{t-1}, c^T, y^T, A, \tau, \delta, V_\pi, V_{c\pi}] = \delta z_t' + (V_{c\pi}/V_c)a_{ct}$. As the term $V_{c\pi}/V_c a_{ct}$ does not involve δ, the regression of interest is:

$$\pi_t - \frac{V_{c\pi}}{V_c} a_{ct} = \delta z_t' + \epsilon_t$$

with $V(\epsilon_t) = V_\pi - V_{c\pi}^2/V_c$. Given our priors, standard results in Bayesian regression analysis (see for instance Box and Tiao, 1973, Chapter 8; also Geweke, 1995) yield:

$$\prod_{t=3}^{T} p(\pi_t|\pi^{t-1}, c^T, y^T, \delta, V_\pi, V_{c\pi}) p(\delta) \;\propto\; N(\delta_*, M_{\delta*}^{-1}) I_\delta$$

where $M_{\delta*} = M_\delta + Z'Z/(V_\pi - V_{c\pi}^2/V_c)$ and $\delta_* = M_{\delta*}^{-1}[M_\delta \delta_0 + Z'Z\hat{\delta}_{OLS}/(V_\pi - V_{c\pi}^2/V_c)]$, $\hat{\delta}_{OLS}$ being the OLS estimator of δ in the regression of $\pi_t - V_{c\pi}/V_c a_{ct}$ on z_t. This result enables us to get draws from $p(\delta|A, \tau, \Sigma, c^T, p^T, \pi^T)$ through a Metropolis-Hastings scheme with proposal $N(\delta_*, M_{\delta*}^{-1})$. For each candidate say $\tilde{\delta}$, the acceptance probability is given by

$$\min\{1, p(\pi_1, \pi_2|c^T, p^T, \tilde{\delta}, V_\pi, V_{c\pi})/p(\pi_1, \pi_2|c^T, p^T, \delta, V_\pi, V_{c\pi})\}$$

This Metropolis step removes the conditioning on the starting values π_1 and π_2. Notice that equation (2.2) implies that the information set $\{c^T, y^T\}$ is only relevant in its elements $\{c_1, c_2\}$ to the conditional distribution of $\{\pi_1, \pi_2\}$. Use is thus made of Appendix A2 for computing the first two moments of (π_1, π_2) conditional on (c_1, c_2).

This closes the circle of simulations, the full sequence consisting of samples successively drawn from $p(\xi^T|\theta, Y^T)$, $p(\mu_p, V_p|p^T)$, $p(\Sigma|A, \tau, \delta, c^T, \pi^T, p^T)$, $p(A|\tau, V_c, c^T)$, $p(\tau|A, V_c, c^T)$ and $p(\delta|A, \tau, \Sigma, c^T, \pi^T, p^T)$. Markov chains properties discussed in Tierney (1994) insure convergence to the joint posterior $p(\xi^T, \theta|Y^T)$. Notice that in our scheme samples for A, τ and δ are obtained in distinct steps. This separation is a consequence of the loss of the naturally conjugate property that results from the re-parametrization (2.3). This cost is however largely offset by the suitability of (2.3) to the description of cycles. We now illustrate our methodology with an investigation of Phillips-curve-based output gap in US and in EU.

4 Application

4.1 Data

The US data have been downloaded from the US Bureau of Economic Analysis web-site *www.bea.doc.gov*. The EU-11 data comes from a recent update of the data set built by Fagan et al. (2001) for the Euro area. For both US and EU cases the sample is made up of 130 quarterly observations between 1970-3 and 2002-4. Longer series are available for US but imposing the same sample dates facilitates comparisons. Both GDP series are in constant prices. Inflation rate is measured on the basis of the consumer price index; we turned down the possibility of using GDP deflators because the relationship between CPI-based inflation rate and real GDP has been found stronger at cyclical frequencies. This confirms a remark made by Kuttner (1994, p.363). There was some evidence of moderate seasonal movements in the EU price index; we removed them with the seasonal adjustment program TRAMO-SEATS (see Gomez and Maravall, 1998). Finally, a large additive outlier was detected in the EU CPI series at the date 1975-Q1. We thus added a dummy variable to equation (2.2). The Bayesian analysis of the associated coefficient is similar to that of the other δ-parameters in the Phillips curve. All series entering equations (2.1)-(2.3) have been multiplied by 100.

For setting the prior hyperparameters we consider the information available from macroeconomic knowledge and from previously published studies. For the US case, we take into account the results in Kuttner (1994) who characterizes fluctuations in the US economy as occurring with a 5-year recurrence and with a contraction factor around .8. Comparable values are also obtained in the univariate analysis of Harvey et al. (2002) for the so-called first-order cycle with non-informative priors. For EU, we consider the results in Gerlach and Smets (1999) who found EU cycles longer than US ones, namely with a periodicity of 8 years, while the amplitude is similar around .8. We thus center the first moment of the τ and A prior distributions on these values without imposing too much precision. Specifically, our priors are for US $(\tau - 2)/(130 - 2) \sim B(4, 28.00)$ for US, $(\tau - 2)/(130 - 2) \sim B(4, 13.07)$ for EU and $A \sim B(50, 14.10)$ for both US and EU. The support for τ is set to $[2, 130]$ since 2 is the minimum periodicity and 130 is the number of observations available. These priors are displayed in Figure 2; it can be seen that the modes of A and τ are fairly close to the empirical results of the studies we refer to. For the Phillips curve parameters we consider:

$$(\mu_\pi, \beta, \lambda, \alpha_1, \alpha_2)' \sim N(\delta_0, M_\delta^{-1}) \times I_\delta$$

with $\delta_0 = (.0, .01, .01, -.3, -.3)'$ and $M_\delta = diag(10, 50, 50, 10, 10)$ for both US and EU cases. The AR parameters are centered on values that are in broad agreement with those in Kuttner (1994); the precision matrix lets however this prior diffuse enough. We incorporate macroeconomic knowledge by imposing a positive response of inflation to the gap and to the output growth. This is done through the index variable I_δ; we shall check that the data confirm the assumption of positive elasticities. The index variable I_δ is also used for imposing stationarity of the AR(2) polynomial on inflation change. The prior on the coefficient of the dummy variable that accounts for the outlier in EU inflation data has been calibrated around a zero-mean with unit variance.

Finally, we choose the priors for the innovation variances and for the drift parameter on the basis of the data statistical properties. That is we tune the hyperparameters of the priors for Σ, V_p and μ_p in such a way that the prior moments $E(V_c) + E(V_p)$, $E(V_\pi)$ and $E(\mu_p)$ are roughly of the same order than the empirical estimates of $V(\Delta y_t)$, $V(\pi_t)$ and $E(\Delta y_t)$, respectively. For both EU and US we use:

$$\begin{pmatrix} V_c & V_{c\pi} \\ & V_\pi \end{pmatrix} \sim IW\left(\begin{bmatrix} 1.2 & 0 \\ & .6 \end{bmatrix}, 10 \right)$$

and $(\mu_p, V_p) \sim NIG(\mu_0, 10, 1.8, 10)$, $\mu_0 = \sum \Delta y_t / 130 = .76$ for US and $.60$ for EU. We shall comment later on the implied prior for the inverse signal to noise ratio.

Having settled our priors we obtain draws from the joint posterior distribution of gap and model parameters conditional on the observations following the MCMC scheme previously detailed. The chain run 100,000 times after a burn-in phase of 50,000. Simulation are recorded every ten iterations so all statistics presented are based on samples of size 10,000. For both δ and Σ parameters, the acceptance ratio of the Metropolis steps is about .95. This is a high acceptance level that reflects the fact that the proposals are very close to the true conditional distribution. For (A, τ) the acceptance rejection Metropolis procedure almost always accepts the proposal, suggesting that their posterior distribution is only weakly non-log concave.

For a selection of variables of interest, Table 1 reports the sample mean and standard deviation, the autocorrelation at lags 1 and 5, the numerical standard error (NSE) associated with the sample mean, the relative numerical efficiency (RNE) and the p-values of the Geweke (1992) convergence diagnostic (CD). The NSE is computed with a window on autocorrelations with length equal to 4% of the sample size. The RNE is obtained as ratio of the NSEs computed using only the output variance and using the 4% window length. Geweke's CD checks whether the average of the first 20% simulations is significantly different to the average last 50%. The selection of variables we focus on includes the cycle periodicity and amplitude, the inverse signal to noise ratio V_c/V_p, the contemporaneous correlation of the innovations in gap with those in inflation $\rho_{c\pi}$, the elasticity of inflation to gap and to output growth respectively obtained as $\beta/(1-\alpha_1-\alpha_2)$ and $\lambda/(1-\alpha_1-\alpha_2)$ and the gap estimated at three time-periods, namely c_{30}, c_{65} and c_{130}.

Table 1 MCMC efficiency and convergence diagnostics

Model specification:

GDP: $y_t = p_t + c_t, \quad \Delta p_t = \mu_p + a_{pt}, \quad c_t - 2A\cos\{2\pi/\tau\}c_{t-1} + A^2 c_{t-2} = a_{ct}$

Inflation: $\pi_t = \mu_\pi + \beta c_{t-1} + \lambda \Delta y_{t-1} + \alpha_1 \pi_{t-1} + \alpha_2 \pi_{t-2} + a_{\pi t}$

US Phillips curve

Variable	Mean	Sd.	ρ_1	ρ_5	NSE ($\times 10^{-2}$)	RNE	CD
τ	23.84	5.15	.01	.00	5.11	1.01	.28
A	.81	.04	.09	.00	.05	.74	.26
V_c/V_p	.53	.36	.44	.01	.57	.40	.16
$\rho_{c\pi}$	.16	.14	.05	.00	.14	.99	.41
e-gap	.04	.01	.09	.00	.02	.57	.80
e-growth	.03	.02	.00	.00	.01	1.28	.81
c_{30}	.62	.98	.00	.00	1.04	0.89	1.00
c_{65}	.29	.98	.00	.00	.95	1.05	.06
c_{130}	-.56	1.18	.03	.00	1.12	1.11	.70

EU Phillips curve

Variable	Mean	Sd.	ρ_1	ρ_5	NSE ($\times 10^{-2}$)	RNE	CD
τ	39.90	10.96	.01	.00	9.33	1.38	.97
A	.80	.04	.02	.01	.03	1.04	.82
V_c/V_p	.59	.25	.17	.00	.29	.75	.79
$\rho_{c\pi}$	.20	.13	.01	.00	.11	1.54	.08
e-gap	.03	.01	.04	.00	.01	1.08	.77
e-growth	.05	.02	.01	.00	.02	.89	.30
c_{30}	.45	1.01	.02	.00	1.04	.94	.44
c_{65}	-1.44	.94	.00	.00	.86	1.20	.69
c_{130}	-.83	1.08	.01	.00	.98	1.22	.25

Notes: $\rho_{c\pi}$ is the correlation between a_{ct} and $a_{\pi t}$; e-gap and e-growth denote elasticity of change in inflation to output gap and output growth; Sd. stands for standard deviation of the variable; ρ_j represents the lag-j autocorrelation; NSE is the numerical standard error of the mean; RNE is relative numerical efficiency and CD denotes Geweke convergence diagnostic.

The autocorrelations between draws are rather low since after only 5 lags the largest value is about .01. The NSE takes values less than 10^{-1} for the periodicity, about 10^{-2} for the cycle and less than 10^{-2} for all other parameters. Hence the number of draws seems sufficient to estimate the mean of every variable with a fair accuracy. This precision is comparable to that we would have achieved with independent samples, as shows the RNE with values almost always above 70%, apart for the US inverse signal to noise ratio with a RNE at 40%. Higher than one RNE-values are obtained for some variables, in particular for the EU periodicity and cross-correlation between inflation and cycle innovations. The Geweke tests are passed at the 5% level in all cases so the chain seems to have converged. Only for c_{65} the Geweke test rejects convergence at the 10% level but since it is the only case we believe that this does not invalidate the results. We thus turn to analyse the output.

The average periodicity of GDP cycles is about 6 years for US and about 10 years for EU. The US result is in agreement with previous studies such as Kuttner (1994) and Harvey et al. (2002) while for EU such a cycle length may appear a bit large. Figure 2 displays prior and posterior distributions of periodicity and amplitude for both US and EU. The US data give a remarkably clear message about the cycle periodicity. In contrast with EU data the evidence is rather vague; for instance the standard error of τ is twice as that obtained with US data. The posterior distribution of EU periodicity is almost a translation of the prior, the mode being shifted to about 8 years. The results about amplitude are instead very similar, with a mode at .8 and a comparable dispersion. So if both US and EU GDPs embody a periodical unobserved pattern, what is obscuring the periodicity in EU data? One possibility of purely statistical nature is that for a given time span, the longer the cyclical recurrence the less precise the periodicity estimate. We checked this conjecture with a Monte Carlo experiment where we simulate 1,000 times series of length 130 from two AR(2) processes with period, amplitude and innovation variance set close to the posterior means, i.e. $(\tau, A, V_c) = (40, .80, .17)$ and $(\tau, A, V_c) = (24, , .80, .17)$. Figure 3 displays the density of the periodicity derived from the OLS estimates of the AR(2) parameters. Clearly, the longer the period the more diffuse the estimate is. The problem of large dispersion of the periodicity estimate in the EU case seems thus related to an insufficience of data for characterising such cycle periodicity.

Figure 4 displays the prior and posterior distributions of the inverse signal to noise ratio. As can be seen, our prior is diffuse enough. For both US and EU the variance of the gap is about one-half of the variance of the trend, and this is obtained with a fair accuracy. Shocks on potential output are thus dominating the two GDPs. Notice that the posterior distribution puts a zero weight on $V_c = 0$: the pile-up problem mentioned in the introduction is thus totally avoided. Figure 4 also shows the distribution of the cross-correlation between shocks in inflation and in the gap. Macroeconomic theory suggests that this correlation is positive but we did not impose any constraint so as to check whether our results are sensible. In this respect, it is reassuring to see that for both EU and US 90% of the posterior distribution is above the region of positive cross-correlations. Phillips curve theory seems thus to not be contradicted by these data. The average cross-correlation is of the same order in the two cases, .16 for US and .20 for EU.

Figure 5 shows the distribution of elasticity of inflation to the gap and to the output growth. It can be seen that beside imposing a positive response, our priors are rather flat. The weight that the posteriors let in neighbourhood of 0 shows that the data confirm the hypothesis of positive elasticities, maybe less clearly for the elasticity to output growth but still with a high enough confidence level. For both US and EU, the distribution of the response of inflation to shocks in the gap is rater concentrated. In contrast the response to shocks in output growth is much loosier, in particular for EU. The main difference between the two variables lies in the periodicity of their movements: the gap captures some mid-term fluctuations while growth is mainly dominated by short-term movements. Inflation thus reacts more to shocks with a mid-term persistence than to transitory ones. This illustrates the possiblity that a relationship between a dependent variable and an unobserved quantity can depend on the dynamic properties of the unobserved component estimate employed and it can be the case that not all proxies yield similar results (see for instance Gali et al., 2001). So for instance if the unobservable of interest is a detrended variable, it would be recommendable to consider several detrending methods.

Figure 6 shows the posterior mean of the gap all along the sample together with the maximum likelihood (ML) estimates. The ML estimates have been obtained using Program GAP downloadable at *www.jrc.cec.eu.int/uasa/prj-gap.asp*. For both US and EU, the Bayesian and ML estimates result very close to each other. The sequence of turning points suggest a delay in the fluctuations of the EU economy with respect to the US. For instance the four most evident peaks in US data that occur in 2000-II, 1989-II, 1978-IV and 1973-II are followed by peaks in EU data in 2000-III, 1990-IV, 1980-I and 1974-I. Figure 6 also displays the 90% highest posterior density (HPD) interval and the

90% ML confidence bands computed with the Ansley and Kohn (1986) procedure that accounts for parameter uncertainty. Both intervals are centered around the null hypothesis of zero-gap estimate. It can be seen that ML and Bayesian analysis yield a similar accuracy, with a slight advantage to the Bayesian approach that is more pronounced for EU than for US. Figure 7 that shows the densities of concurrent gap estimates further confirms this. This moderate gain in accuracy results from the knowledge that we have incorporated into the analysis. Yet as can be seen on Figure 6 only in very few occasions the gap turns out to be significant: more information is still needed for better characterising it.

5 Conclusion

In this paper we have developed a Bayesian analysis of Kuttner's bivariate Phillips-curve model for estimating output gaps. The analysis involves Gibbs sampling with adaptive rejection Metropolis and Metropolis-Hastings steps for getting draws of parameters conditionally on the state. Our analysis does not condition on first observations. We also reparametrized the cyclical AR(2) specification in terms of the polar coordinates of the AR(2) roots mainly because non-informative priors on AR parameters put too much emphasis on short-term periodicities. Besides being quite natural for describing cycles, the flexibility of the polar coordinates parametrization is important to get meaningful results.

We illustrated our approach with an application to the estimation of cycles in EU and US GDP. In general the results show that the data agree with the Phillips-curve theory. We observed a sharp response of inflation to the gap both in US and EU. In contrast the reaction to output growth is much more vague. The estimate of the periodicity of EU cycle has been found diffuse enough; we could relate this result to the difficulty of inferring about mid-term periodicities using a limited amount of data. Also, the gap posterior mean we obtained presents evidence of a leading behaviour of the US cycle over the EU one. Since only in a few occasions the gap posterior mean turned out significant, the a priori knowledge that we have incorporated into the analysis seems rather modest. More information would be needed for further reducing uncertainty and eventually improving the assessment of the current economic situation. The tools that we have developed in this paper enable analysts to incorporate any extra information into Phillips-curve-based output gap measures.

References

ANSLEY C.F., AND KOHN R. (1986), "Prediction mean squared error for state space models with estimated parameters", *Biometrika*, 73, 467-73.

APEL, M. AND JANSSON, P. (1999a), "A theory-consistent approach for estimating potential output and the NAIRU", *Economics Letters*, 64, 271-75.

APEL M. AND JANSSON, P. (1999b), "System estimate of potential output and the NAIRU", *Empirical Economics*, 24, 373-388.

BARNETT G., KOHN R. AND SHEATER S. (1996), "Bayesian estimation of autoregressive model using Monte chain Monte Carlo", *Journal of Econometrics*, 74, 237-254.

BAUWENS L., LUBRANO M., AND RICHARD J. (1999), *Bayesian Inference in Dynamic Econometric Models*, Oxford University Press.

BILLIO M., MONFORT A. AND ROBERT C. (1999), "Bayesian estimation of switching ARMA models", *Journal of Econometrics*, 93, 229-255.

BOX, G. AND TIAO G. (1973) *Bayesian inference in statistical analysis*, Wiley Classics.

CARTER C., AND KOHN R. (1994), "On Gibbs sampling for state space models", *Biometrika*, 81, 541-553.

CASELLA G. AND GEORGE E. (1992), "Explaining the Gibbs sampler", *The American Statistician*, 46, 3, 167173.

CHIB S. (1993), "Bayesian regression with autoregressive errors. A Gibbs sampling approach", *Journal of Econometrics*, 58, 275-94.

CHIB S., AND GREENBERG E. (1994), "Bayesian inference in regression models with ARMA(p,q) errors", *Journal of Econometrics*, 64, 183-206.

CHIB S., AND GREENBERG E. (1995), "Understanding the Metropolis-Hastings algorithm", *The American Statistician*, 49, 4, 327-335.

CLARK, P. (1987), "The cyclical comovement of U.S. economic activity", *Quarterly Journal of Economics*, 797-814.

DE JONG P. (1991), "The diffuse Kalman filter", *Annals of Statistics*, 2, 1073-83.

DE JONG P., AND SHEPHARD N. (1995), "The simulation smoother for time series models", *Biometrika*, 82, 339-50.

DURBIN, J. AND S. KOOPMAN (2001), *Time Series Analysis by State Space Methods*, Oxford: Oxford University Press.

Durbin, J. and S. Koopman (2002), "A simple and efficient simulation smoother for state space time series analysis", *Biometrika*, 89, 3, 603-616.

Fagan G., J. Henry and R. Mestre (2001), "An area-wide model for the Euro area", European Central Bank Working Paper No. 42.

Gali J., Gertler M. and Lopez-Salido D. (2001), "European inflation dynamics", *European Economic Review*, 45, 1237-1270.

Gerlach, S. and Smets F. (1999), "Output gap and monetary policy in the EMU area", *European Economic Review*, 43, 801-812.

Geweke J. (1992), "Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments", in *Bayesian Statistics*, 4, by J. Bernardo, J. Berger, A. Dawid and A.Smith (eds), Oxford: Oxford University Press.

Geweke J. (1995), "Posterior simulators in Econometrics", Federal Reserve Bank of Minneapolis, Working Paper 555.

Geweke J. (1999), "Using simulation methods for Bayesian econometric models: inference, development, and communication", *Econometric Reviews*, 18, 1, 1-73.

Gómez V. and Maravall A. (1996), "Programs Seats and Tramo: instructions for the user", WP 9628, Bank of Spain.

Hall, R.E. and J.B. Taylor (1991), *Macroeconomics*, New York: Norton.

Harvey A. (1981), *Time Series Models*, 2nd edition, Harvester Wheatheaf.

Harvey A., Trimbur T. and van Dijk H. (2002), "Cyclical components in economic time series: a Bayesian approach", University of Cambridge.

Kichian, M. (1999), "Measuring potential output with a state-Space framework", Working Paper 99-9, Bank of Canada.

Koop G. and van Dijk H. (2000), "Testing for integration using evolving trend and seasonal models: a Bayesian approach", *Journal of Econometrics*, 97, 261-291.

Koopman S. (1997), "Exact initial Kalman filtering and smoothing for nonstationary time series model", *Journal of American Statistical Association*, 92, 440, 1630-1638.

Kuttner K. (1994), "Estimating potential output as a latent variable", *Journal of Business and Economic Statistics*, 12, 3, 361-368.

McCulloch R. and Tsay, R. (1994), "Bayesian analysis of autoregressive time series via the Gibbs sampler", *Journal of Time Series Analysis*, 15, 2, 235-250.

OECD (2000), "The concept, policy use and measurement of structural unemployment", Annex 2 of "Estimating time-varying NAIRU's across 21 OECD countries", Paris.

Orphanides, A. and S. van Norden (2002), "The unreliability of output gap estimates in real-Time", *Review of Economics and Statistics*, 84, 4, 569-583.

Planas C., Roeger W. and A. Rossi (2003), "How much has labour taxation contributed to European structural unemployment?", Economic Paper 183, European Commission, Brussels.

Stock J. and Watson M. (1998), "Median unbiased estimation of coefficient variance in a time-varying parameter model", *Journal of the American Statistical Association*, 349-358.

Taylor J. (1993), "Discretion versus policy rules in practice", *Carnegie-Rochester Conference Series on Public Policy*, 39, 195-214.

Tierney L. (1994), "Markov chains for exploring posterior distributions", *Annals of Statistics*, 22, 4, 1701-1762.

Watson M. (1986), "Univariate detrending methods with stochastic trends", *Journal of Monetary Economics*, 18, 49-75.

Appendix

A1. Derivation of $\xi_{1|1}$ and $P_{1|1}$.

The initial state vector can generally be written as:

$$\xi_1 = a + A\delta + B\eta \quad \eta \sim N(0, I) \quad \delta \sim N(0, kI)$$

where $k \to \infty$, a is a vector of constants and A and B are known matrices. This formulation implies

$$\xi_1 \sim N(a, P)$$

where $P = P_\star + kP_\infty$, $P_\star = BB'$ and $P_\infty = AA'$. Non-zero elements in the matrix P_∞ correspond to the diffuse elements of the state. Let d denote the number of non-stationary elements of the state. The algorithm proposed by de Jong's (1991) yields $\xi_{d+1|d}$ and $P_{d+1|d}$. Koopman (1997) further extended it in order to get $\xi_{1|1}$, $P_{1|1}$, $\cdots$, $\xi_{d|d}$, $P_{d|d}$. We implemented Koopman's algorithm as follows.

Conditional on the first observation, the first two moments of the state at time $t = 1$ are:

$$\begin{aligned} \xi_{1|1} &= a + PH'(HPH')^{-1}(Y_1 - Ha) \\ P_{1|1} &= P - PH'(HPH')^{-1}HP \end{aligned} \tag{5.1}$$

Let S_1 denote the matrix defined by $S_1 = HPH'$. It is easily seen that

$$S_1 = HP_\star H' + kHP_\infty H' = S_{\star,1} + kS_{\infty,1}$$

The dimension of $S_{\star,1}$ and $S_{\infty,1}$ is in general equal to the dimension of Y_t, say N, and $S_{\infty,1}$ is of reduced rank, $r(S_{\infty,1}) = q < N$. Following Lemma 2 in Koopman (1997), we obtain the $N \times N$ matrix $Q = [Q_{1\ (N\times q)} Q_{2\ (N\times(N-q))}]$ that partially diagonalises $S_{\star,1}$ and $S_{\infty,1}$ according to:

$$Q'S_{\infty,1}Q = \begin{bmatrix} I_q & 0_{q\times(N-q)} \\ 0_{(N-q)\times q} & 0_{(N-q)\times(N-q)} \end{bmatrix} \qquad Q'S_{\star,1}Q = \begin{bmatrix} C_{q\times q} & 0_{q\times(N-q)} \\ 0_{(N-q)\times q} & I_{N-q} \end{bmatrix}$$

The matrices Q_1 and Q_2 verify $Q_1' S_{\infty,1} Q_1 = I_q$, $Q_1' S_{\infty,1} Q_2 = 0_{q\times(N-q)}$, $Q_2' S_{\infty,1} Q_2 = 0_{(N-q)\times(N-q)}$, $Q_1' S_{\star,1} Q_2 = 0_{q\times(N-q)}$ and $Q_2' S_{\star,1} Q_2 = I_{N-q}$. Writing the matrix $S_{\star,1}$ as

$$S_{\star,1} = \begin{bmatrix} S_{\star 11\ q\times q} & S_{\star 12\ q\times(N-q)} \\ S_{\star 21\ (N-q)\times q} & S_{\star 22\ (N-q)\times(N-q)} \end{bmatrix}$$

and similarly for $S_{\infty,1}$, it can be seen that the solutions are:

$$Q_2 = \begin{bmatrix} Q_{21\ q\times(N-q)} \\ Q_{22\ (N-q)\times(N-q)} \end{bmatrix} = \begin{bmatrix} 0_{q\times(N-q)} \\ V_2 \Lambda_2^{-1/2} \end{bmatrix}$$

where V_2 and Λ_2 are the matrices of eigenvectors and eigenvalues of $S_{\star 22}$. The matrix Q_1 is instead obtained as:

$$Q_1 = \begin{bmatrix} Q_{11\ q\times q} \\ Q_{12\ (N-q)\times q} \end{bmatrix} = \begin{bmatrix} V_1 \Lambda_1^{-1/2} \\ -(\Lambda_1^{-1/2} V_1' S_{\star 12} V_2 \Lambda_2^{-1/2} (S_{\star 22} V_2 \Lambda_2^{-1/2})^{-1})' \end{bmatrix}$$

where V_1 and Λ_1 are the matrices of eigenvectors and eigenvalues of $S_{\infty,1}$. Then according to Theorem 2 in Koopman (1997), the inverse of S_1 is given by:

$$(S_{\star,1} + k S_{\infty,1})^{-1} = S_\star^- + \frac{1}{k} S_\infty^- - \frac{1}{k^2} S_\infty^- S_{\star,1} S_\infty^- + O(k^{-3})$$

where $S_\star^- = Q_2' Q_2$ and $S_\infty^- = Q_1' Q_1$. Plugging this last result into (5.1) and using the results that in our setting $P_\infty H' S_\star^- = 0$ and $P_\infty H' S_\infty^- H P_\infty = P_\infty$, for $k \to \infty$ we get:

$$\begin{aligned} \xi_{1|1} &= a + (P_\star H' S_\star^- + P_\infty H' S_\infty^-)(Y_1 - Ha) \\ P_{1|1} &= P_\star - P_\star H' S_\star^- H P_\star - P_\infty H' S_\infty^- H P_\star \\ &\quad - P_\star H' S_\infty^- H P_\infty + P_\infty H' S_\infty^- S_{\star,1} S_\infty^- H P_\infty \end{aligned}$$

which allows us to compute $\xi_{1|1}$.

A2. Second moments of (c_1, c_2, π_1, π_2).

Let γ_{ck}, $\gamma_{\pi k}$, $\gamma_{\pi c,k}$ and $\gamma_{\pi\Delta y,k}$ be such that $\gamma_{ck} = E(c_t c_{t-k})$, $\gamma_{\pi k} = E(\pi_t \pi_{t-k})$, $\gamma_{\pi c,k} = E(\pi_t c_{t-k})$ and $\gamma_{\pi\Delta y,k} = E(\pi_t \Delta y_{t-k})$. Equation (2.2) implies:

$$\begin{aligned}
\gamma_{\pi 0} &= \alpha_1\gamma_{\pi 1} + \alpha_2\gamma_{\pi 2} + \lambda\gamma_{\pi\Delta y,1} + \beta\gamma_{\pi c,1} + \mu_\pi E[\pi_t] + V_\pi \\
\gamma_{\pi 1} &= \alpha_1\gamma_{\pi 0} + \alpha_2\gamma_{\pi 1} + \lambda\gamma_{\pi\Delta y,0} + \beta\gamma_{\pi c,0} + \mu_\pi E[\pi_t] \\
\gamma_{\pi 2} &= \alpha_1\gamma_{\pi 1} + \alpha_2\gamma_{\pi 0} + \lambda\gamma_{\pi\Delta y,-1} + \beta\gamma_{\pi c,-1} + \mu_\pi E[\pi_t]
\end{aligned} \tag{5.2}$$

Using both equations (2.1) and (2.2), we have:

$$\begin{aligned}
\gamma_{\pi\Delta y,1} &= \alpha_1\gamma_{\pi\Delta y,0} + \alpha_2\gamma_{\pi\Delta y,-1} + \lambda(V_p + 2(\gamma_{c0} - \gamma_{c1})) + \beta(\gamma_{c0} - \gamma_{c1}) \\
&+ \gamma\mu_p^2 + \mu_\pi\mu_p \\
\gamma_{\pi\Delta y,0} &= \alpha_1\gamma_{\pi\Delta y,-1} + \alpha_2\gamma_{\pi\Delta y,-2} + \lambda(2\gamma_{c1} - \gamma_{c0} - \gamma_{c2}) + \beta(\gamma_{c1} - \gamma_{c0}) \\
&+ \lambda\mu_p^2 + \mu_\pi\mu_p + V_{c\pi} \\
\gamma_{\pi\Delta y,-1} &= \gamma_{\pi c,-1} - \gamma_{\pi c,0} + E[\pi_t]\mu_p \\
\gamma_{\pi\Delta y,-2} &= \gamma_{\pi c,-2} - \gamma_{\pi c,-1} + E[\pi_t]\mu_p
\end{aligned} \tag{5.3}$$

Proceeding similarly for the cross-moments between π_t and c_t yields:

$$\begin{aligned}
\gamma_{\pi c,1} &= \alpha_1\gamma_{\pi c,0} + \alpha_2\gamma_{\pi c,-1} + \lambda(\gamma_{c0} - \gamma_{c1}) + \beta\gamma_{c0} \\
\gamma_{\pi c,0} &= \alpha_1\gamma_{\pi c,-1} + \alpha_2\gamma_{\pi c,-2} + \lambda(\gamma_{c1} - \gamma_{c2}) + \beta\gamma_{c1} + V_{c\pi} \\
\gamma_{\pi c,-1} &= \phi_1\gamma_{\pi c,0} + \phi_2\gamma_{\pi c,1} \\
\gamma_{\pi c,-2} &= \phi_1\gamma_{\pi c,-1} + \phi_2\gamma_{\pi c,0}
\end{aligned} \tag{5.4}$$

The system (5.2)-(5.4) is then solved with the terms γ_{ck}, $k = 0, 1, 2$ directly obtained from (2.1).

Figure 1: Cumulative distribution of periodicity if $(\phi_1, \phi_2)' \sim N(0_2, 10 \times Id_2) I_S$

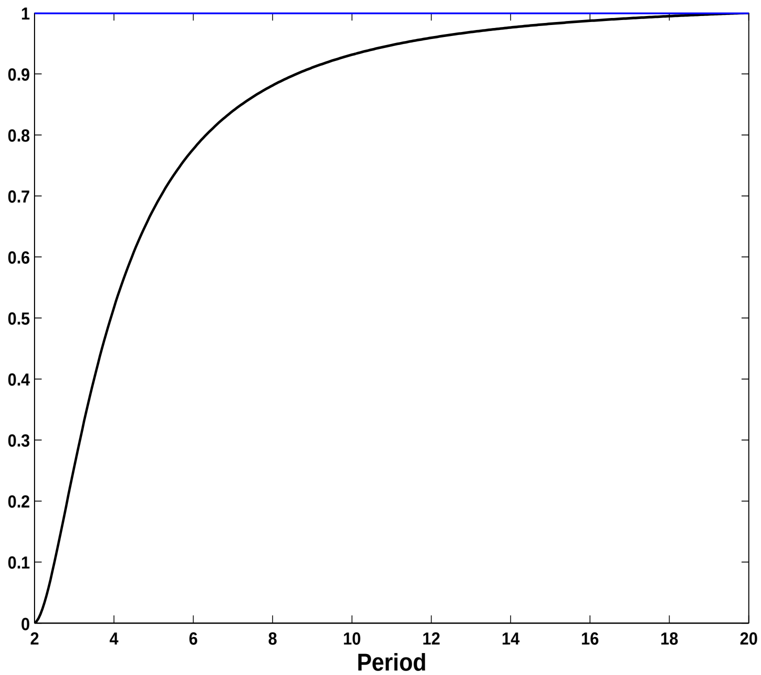
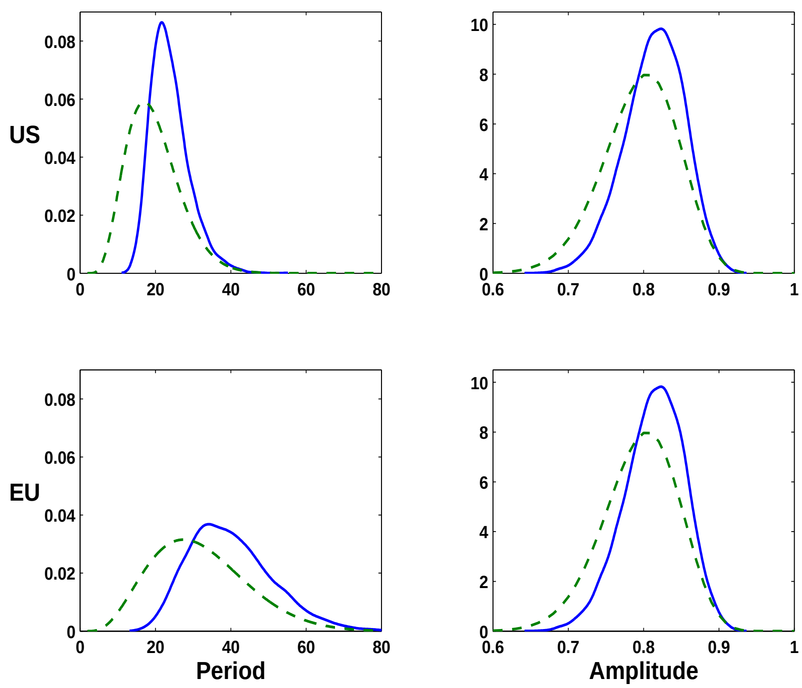
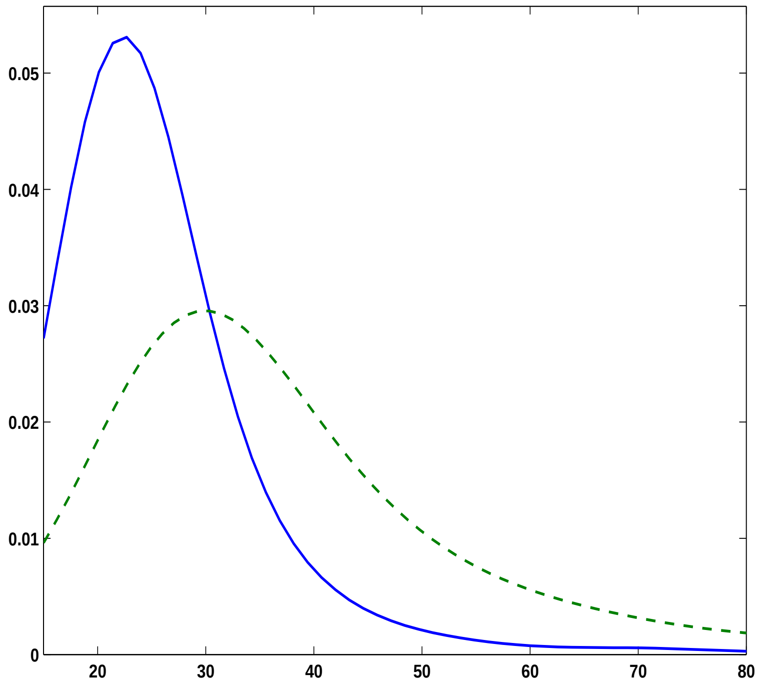


Figure 2: Densities of cycle periodicity and amplitude



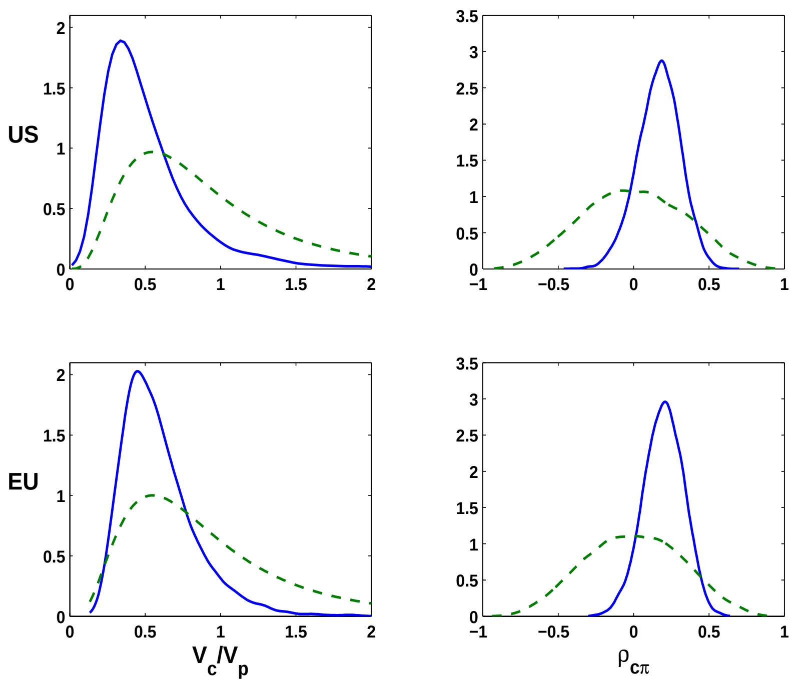
– – Prior —— Posterior

Figure 3: Densities of OLS-estimated periodicity in simulated AR(2) models - 1000 replications, T=130



– – $A = .8$, $\tau = 40$, $V_c = .17$ —— $A = .8$, $\tau = 24$, $V_c = .17$

Figure 4: Densities of inverse signal to noise ratio and of correlation between gap and inflation innovations



– – Prior —— Posterior

Figure 5: Densities of elasticity of inflation to output gap and to output growth

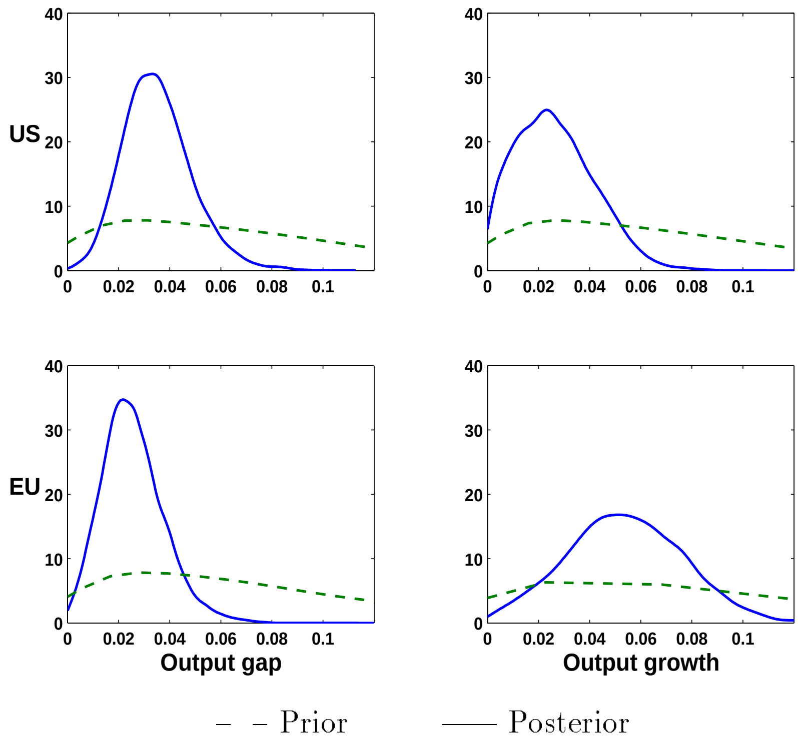


Figure 6: Gap posterior mean (—) with 90% HPD region ML estimate (- -) with 90% confidence bands

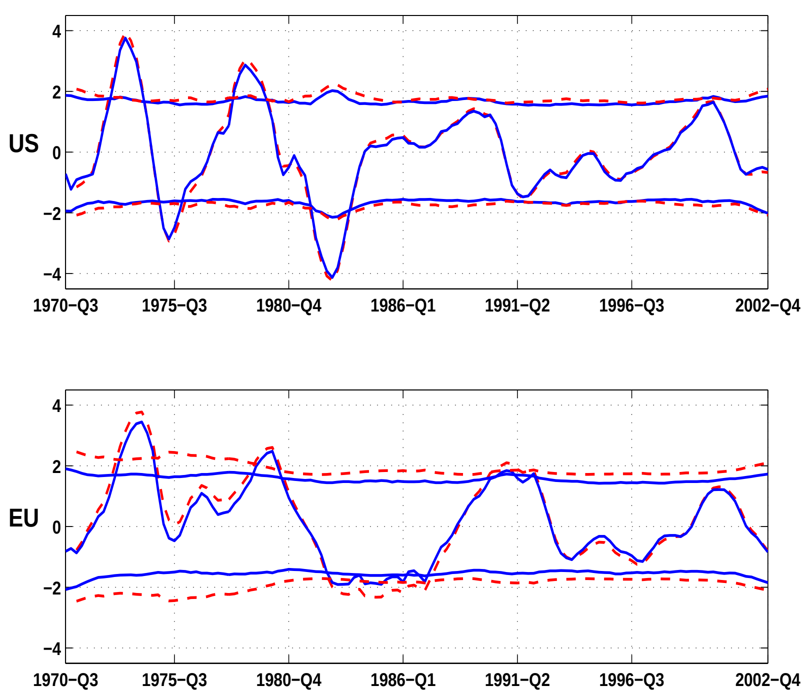


Figure 7: Densities of concurrent gap (c_{130})
Bayesian (—) vs. ML (- -)