THE RIMINI CENTRE FOR ECONOMIC ANALYSIS

WP 21-08

Gabriele Fiorentini
University of Florence, Italy
and
The Rimini Centre for Economic Analysis, Italy

Chrispophe Planas
Joint Research Centre of the European Commission

Alessandro Rossi
Joint Research Centre of the European Commission

"THE MARGINAL LIKELIHOOD OF STRUCTURAL TIME SERIES MODELS, WITH APPLICATION TO THE EURO AREA AND US NAIRU"

The Rimini Centre for Economic Analysis
Legal address: Via Angherà, 22 – Head office: Via Patara, 3 - 47900 Rimini (RN) – Italy
www.rcfea.org - secretary@rcfea.org

The marginal likelihood of Structural Time Series Models, with application to the euro area and US NAIRU *

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

December 2007

Abstract

We propose a simple procedure for evaluating the marginal likelihood in univariate Structural Time Series (STS) models. For this we exploit the statistical properties of STS models and the results in Dickey (1968) to obtain the likelihood function marginally to the variance parameters. This strategy applies under normal-inverted gamma-2 prior distributions for the structural shocks and associated variances. For trend plus noise models such as the local level and the local linear trend, it yields the marginal likelihood by simple or double integration over the (0,1)-support. For trend plus cycle models, we show that marginalizing out the variance parameters greatly improves the accuracy of the Laplace method. We apply this methodology to the analysis of US and euro area NAIRU.

KEYWORDS: Marginal likelihood, Markov Chain Monte Carlo, unobserved components, bridge sampling, Laplace method, NAIRU.

*The ideas expressed here are those of the authors and do not necessarily reflect the positions of the European Commission.

(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 propose a simple procedure for evaluating the marginal likelihood in univariate Structural Time Series (STS) models. For this we exploit the statistical properties of STS models and the results in Dickey (1968) to obtain the likelihood function marginally to the variance parameters. Our strategy applies under normal-inverted gamma-2 prior distributions for the structural shocks and associated variances, an assumption that is quite common in the time series literature (see Fruhwirth-Schnatter, 1994; Chib and Greenberg, 1994). For trend plus noise models such as the local level and the local linear trend, it yields the marginal likelihood by simple or double integration over the (0,1)-support, without any MCMC sampling. For trend plus cycle models, we show that marginalizing out the variance parameters greatly improves the accuracy of the Laplace method.

Since the first studies in the 1970's (see Pagan, 1975), STS models have become quite widespread in empirical macroeconomics. Researchers typically resort to structural components for describing potential output (Clark, 1987), technological growth (Hansen, 1997), reservation wage (Planas, Roeger and Rossi, 2007), permanent income (Hall and Mishkin, 1982) and trend inflation (Cogley and Sargent, 2005; Stock and Watson, 2007). Not always however the prior information is sufficient for isolating a particular model, and in some cases discriminating between different specifications can be a difficult task. For instance, inferring about a stationary against an integrated process for the trend slope is not immediate. Moreover in the STS framework classical hypothesis testing does not apply straightforwardly because the null hypothesis often lies on the boundary of the parameter space, like for testing for a deterministic component (see Harvey, 2001). Also the null and alternative hypothesis may not be nested. Through the marginal likelihood, the Bayesian framework offer a conceptually simple answer to the model selection problem (see Kass and Raftery, 1995), with the important advantage of involving exact finite sample distributions instead of asymptotic assumptions. The drawback, however, is that evaluating the marginal likelihood is cumbersome: the number of parameters to be integrated out is usually relevant and the likelihood function is typically highly concentrated with respect to the prior distribution (see Fruhwirth-Schnatter, 2005).

The approach we propose here takes advantage of the properties of STS models to simplify the marginal likelihood computation. For trend plus noise decompositions, the result is a simple tool for the selection of trend models, as we shall see in Section 2.

For trend plus cycle models, we suggest in Section 3 to apply the Laplace method on the posterior density defined marginally to variance parameters. We show through a simulation study in Section 4 that our procedure greatly improves over the traditional Laplace marginal likelihood estimates (Tierney and Kadane, 1986), and that it is comparable with bridge sampling (Meng and Wong, 1996) although it does not involve any importance sampling. Finally, in Section 5 we apply this methodology to the analysis of the NAIRU in US and in the euro area. We focus on the NAIRU as the European Commission uses it for estimating the potential growth of Member State economies (see Denis, Grenouilleau, Mc Morrow and Roeger, 2006). There have been some debate about the US NAIRU characteristics (see for instance Staiger, Stock and Watson, 1997; Stiglitz, 1997; Ball and Mankiw, 2002). Our methodology enables us to discriminate between twenty seven models for the US and the euro area NAIRU, and to end up with recommendation for practitioners. Section 6 concludes.

2 The marginal likelihood of trend plus noise STS models with IG-variance parameters

2.1 Background

The results we present in this Section are based on the following Theorem:

Theorem 2 (Dickey, 1968, p.1623) Let $\tau_1,\ldots,\tau_K$ *have independent standard* q_k*-dimensional multivariate t distributions with* $\nu_1,\ldots,\nu_K$ *degrees of freedom (centers 0, matrices* $\nu_k^{-1} I_{q_k}$*). Then the random r-vector* δ*,*

$$\delta = \sum B_k \tau_k$$

has the representation

$$\delta = (\sum u_k^{-1} \nu_k B_k B_k')^{1/2} z$$

where the u_k *are independently chi-squared distributed with* ν_k *degrees of freedom, and* z *is an independent r-dimensional standard normal vector. Consequently,* δ *has the further representation*

$$\delta = (\sum v_k^{-1} (\nu_k/\nu) B_k B_k')^{1/2} \tau$$

where, with $\nu = \sum \nu_k$*, the* $v_k = u_k / \sum u_k$*,* $v_1,\ldots,v_K$ *are jointly Dirichlet distributed:* $v_k > 0$*,* $\sum v_k = 1$*, with density* $\Gamma(\nu/2) \prod_k v_k^{\nu_k/2-1} / \Gamma(\nu_k/2)$ *in* $v_1,\ldots,v_{K-1}$*, and* τ *has an r-dimensional standard multivariate t distribution with* ν *degrees of freedom. If the matrix* $\sum B_k B_k'$ *is non singular, the distribution of* δ *is non degenerate with the density function*

$$f(\delta) = \Gamma(\frac{\nu + r}{2}) \pi^{-r/2} / \prod_k \Gamma(\nu_k/2) \times$$
$$\int_\sigma \prod_k v_k^{\nu_k/2-1} \mid \sum v_k^{-1} \nu_k B_k B_k') \mid^{-1/2} [1 + \delta' (\sum v_k^{-1} \nu_k B_k B_k')^{-1} \ \delta]^{-\frac{\nu+r}{2}} dv_1 \cdots dv_{K-1}$$

the range σ *of the* v_k *as above.* ∎

Dickey's Theorem 2 expresses the density of a linear combination of independently distributed multivariate t vectors as an integral of dimension one less than the number of summands. Assuming standardized t-distributions for the τ_k vectors, the density of δ is obtained as a function of the degrees of freedom ν_k and of the products $\nu_k B_k B_k'$. As we turn to see, this result greatly simplifies the computation of the marginal likelihood of STS models with inverted gamma-2 (IG) priors on the variance parameters.

2.2 First-order random walk trends

We first consider the case of a time series y_t that is made up of a random walk p_t plus a noise c_t like in:

$$\begin{array}{rcll} y_t &=& p_t + c_t & \\ \Delta p_t &=& a_{pt} & a_{pt}|V_p \sim N(0, V_p) \\ c_t &=& a_{ct} & a_{ct}|V_c \sim N(0, V_c) \end{array} \tag{2.1}$$

where $\Delta \equiv 1 - L$ and L is the lag operator. Model (2.1) is also known as the local level model (see Durbin and Koopman, 2001, Chap.2). Given their respective variance, the structural shocks a_{ct} and a_{pt} are independent and Normally distributed. The variance parameters V_c and V_p are assumed to be random variables with IG prior distribution:

$$V_\ell \sim IG(s_{\ell 0}, \nu_{\ell 0}) \qquad \ell = c, p. \tag{2.2}$$

For easing exposition, we shall denote $x_k^T \equiv (x_k, \cdots, x_T)'$ and in short $x \equiv x_1^T$. IG-priors for variance parameters have been intensively used in time series analysis (see for instance Chib, 1993; Chib and Greenberg, 1994). This assumption lets the corresponding shocks marginally distributed according to the Student density:

$$f(a_\ell) = t(0, s_{\ell 0}, I_T, \nu_{\ell 0}) \qquad \ell = c, p$$

where I_T is the $T \times T$ identity matrix. The structural shocks can be expressed as $a_\ell = (s_{\ell 0}/\nu_{\ell 0})^{1/2}\tau_\ell$, where τ_ℓ is the random variable with standard t-distribution:

$$f(\tau_\ell) = t(0, 1, I_T/\nu_{\ell 0}, \nu_{\ell 0}) \qquad \ell = c, p \tag{2.3}$$

Let us define D_1 the $T-1 \times T$ first-difference matrix, i.e. $D_1(i,i) = -1$, $D_1(i, i+1) = 1$, and 0 elsewhere. The stationary transformation of the observed process y verifies:

$$\begin{array}{rcl} D_1 y &=& (s_{p0}/\nu_{p0})^{1/2}\tau_p + (s_{c0}/\nu_{c0})^{1/2} D_1 \tau_c \\ &=& B_p \tau_p + B_c \tau_c \end{array}$$

The local level model is thus a particular case of Theorem 2 with $B_p = (s_{p0}/\nu_{p0})^{1/2} I_{T-1}$ and $B_c = (s_{c0}/\nu_{c0})^{1/2} D_1$. We get:

Lemma 1 The marginal likelihood of the local level model (2.1) with IG-prior (2.2) on the variance parameters is such that:

$$\begin{aligned} f_D(y) &= \pi^{-\frac{T-1}{2}}\Gamma(\frac{\nu}{2})\Gamma(\frac{\nu_{p0}}{2})^{-1}\Gamma(\frac{\nu_{c0}}{2})^{-1} \\ &\times \int_0^1 u^{\nu_{p0}/2-1}(1-u)^{\nu_{c0}/2-1} \mid \frac{s_{p0}}{u}M_p + \frac{s_{c0}}{1-u}M_c \mid^{-1/2} \\ &\times [1+(D_1y-\mu_y)'(\frac{s_{p0}}{u}M_p + \frac{s_{c0}}{1-u}M_c)^{-1}(D_1y-\mu_y)]^{-\frac{\nu}{2}}du \end{aligned} \tag{2.4}$$

with $\nu = \nu_{p0} + \nu_{c0} + T - 1$, $M_p = I_{T-1}$, $M_c = D_1D_1'$, *and* $\mu_y = 0$.$\blacksquare$

Lemma 1 reduces the problem of evaluating the marginal likelihood of T observations in the local level model to a scalar integration over the support $(0, 1)$, the bounds being excluded. Notice that for such a model with only variance parameters, evaluating the marginal likelihood does not require any MCMC simulation. Some numerical tools can however help. In particular, the diagonalization $M_c = P_c\Lambda_cP_c'$, where P_c and Λ_c denote the eigenvectors and eigenvalues matrices of dimension $T - 1$, yields:

$$(\frac{s_{p0}}{u}M_p + \frac{s_{c0}}{1-u}M_c)^{-1} = P_c\{\frac{s_{p0}}{u}I_{T-1} + \frac{s_{c0}}{1-u}\Lambda_c\}^{-1}P_c' \tag{2.5}$$

since for the random walk $M_p = I_{T-1}$. Expression (2.5) is advantageous as the central term is a diagonal matrix and as the eigenvectors do not depend on u. Since $\mid P_c \mid= 1$, it also simplifies the computation of the determinant.

In empirical macroeconomics, a constant slope is often added to the trend (see for instance Stock and Watson, 1988):

$$\Delta p_t = \mu_p + a_{pt} \tag{2.6}$$

Considering the standard assumption that μ_p, V_p are jointly NIG-distributed like in

$$f(\mu_p|V_p) = N(\mu_{p0}, V_pv_{\mu0}), \tag{2.7}$$

the distribution of the trend growth marginally to the parameters μ_p and V_p becomes:

$$f(D_1p) = t(\mu_{p0}, s_{p0}, (I_{T-1} + 1_{T-1}v_{\mu0})^{-1}, \nu_{p0})$$

where 1_k is the $k \times k$ matrix of ones (see Bauwens et al., 1999, p.300, 304). In terms of the standardized t-variables defined in (2.3), we have now $\tau_p = (s_{p0}/\nu_{p0})^{-1/2}(I_{T-1} + 1_{T-1}v_{\mu 0})^{-1/2}a_p$. The stationary transformation of the observed series verifies:

$$D_1 y = \mu_{p0} + (s_{p0}/\nu_{p0})^{1/2}(I_{T-1} + 1_{T-1}v_{\mu 0})^{1/2}\tau_p + (s_{c0}/\nu_{c0})^{1/2}D_1\tau_c$$

We can thus state:

Lemma 2 The marginal likelihood of the random walk with drift plus noise model with NIG prior distributions (2.2) and (2.7) is like in Lemma 1, equation (2.4), with $M_p = I_{T-1} + 1_{T-1}v_{\mu 0}$, $M_c = D_1 D_1'$, *and* $\mu_y = \mu_{p0}$.∎

Because M_p has lost the identity structure, the simple diagonalisation (2.5) cannot be used anymore for speeding up the integration. One must instead resort to the simultaneous diagonalization such that $Q'M_pQ = I_{T-1}$ and $Q'M_cQ = \Lambda_c$, where Λ_c is a diagonal matrix (see Magnus and Neudecker, 1988, p.22). This yields:

$$(\frac{s_{p0}}{u}M_p + \frac{s_{c0}}{1-u}M_c)^{-1} = Q\{\frac{s_{p0}}{u}I_{T-1} + \frac{s_{c0}}{1-u}\Lambda_c\}^{-1}Q'$$

The matrix Q is obtained as $Q = P_p\Lambda_p^{-1/2}P_c$, where P_p and Λ_p are the eigenvectors and eigenvalues matrices related to M_p, and P_c is the eigenvector matrix of the product $(P_p\Lambda_p^{-1/2})'M_cP_p\Lambda_p^{-1/2}$.

Before turning to trend models with stochastic slope, we briefly discuss a consequence of the IG-variance priors:

Corollary 1 Marginally to the variance parameters, the posterior distribution of the increments D_1p *is the poly-t 2-0 density:*

$$f(D_1p|y) \propto t(0, s_{p0}, I_{T-1}, \nu_{p0}) \times t(D_1y, s_{c0}, \{D_1D_1'\}^{-1}, \nu_{c0})$$

for the local level model (2.1) and

$$f(D_1p|y) \propto t(\mu_{p0}, \{s_{p0}, I_{T-1} + 1_{T-1}v_{\mu 0}\}^{-1}, \nu_{p0}) \times t(D_1y, s_{c0}, \{D_1D_1'\}^{-1}, \nu_{c0})$$

for the random walk with drift plus noise model (2.6)-(2.7). ∎

In the two equations above, the first t-kernel corresponds to the prior distribution of $D_1p = \mu_y + a_p$ while the second term is the prior distribution of D_1c evaluated at

$D_1 c = D_1 y - D_1 p$. Trivially, in the two-component model $f(D_1 c|y) = f(D_1 p|y)$. Given Corollary 1, the complete posterior distribution of the latent variables c and p can be easily retrieved. By convolution, it is proportional to the product of the marginal prior distributions of c and p evaluated at $p + c = y$, i.e. $f(c|y) \propto f(c) \times f(D_1 p = D_1 y - D_1 c)$. Since $f(c|y) = f(c_1|D_1 c, y) \times f(D_1 c|y)$, it can be seen that given the increments $D_1 c$, the data do not bring further information about the starting point c_1, i.e. $f(c_1|D_1 c, y) = f(c_1|D_1 c)$. Given the t marginal prior for c, the term $f(c_1|D_1 c)$ can be obtained as a Student density. Multiplying it by the distribution in Corollary 1 yields the kernel of the posterior distribution of the unobserved components.

This makes possible the use of Richard and Tompa's (1980) results to draw posterior samples of the unobservables marginally to the variance parameters in two steps: first the increments, for instance following Appendix B.4.6 in Bauwens et al. (1999, p.321), and then the starting point given the increments. Program simplicity would be the main appeal: neither diffuse Kalman Filter initialization (see deJong, 1991) nor smoothing algorithm is needed. The cost however would be a substantial computing time delay due to the resorting to matrix computations. If the sampling of the state variable is inserted into a MCMC scheme, for instance when Corollary 1 holds conditionally on any other random quantity, such a delay can become prohibitive. For these cases, a recursive scheme such as the Carter and Kohn (1994) state-sampler remains preferable. At least so long a procedure for factorizing poly-t 2-0 densities is not available.

2.3 Second-order random walk trends

For some macroeconomic variables like unemployment, the hypothesis of constant growth is unrealistic. Letting the slope be an integrated process gives more flexibility. The trend equation becomes:

$$\begin{aligned} \Delta p_t &= \mu_{t-1} + a_{pt} \qquad & a_{pt}|V_p \sim N(0, V_p) \\ \Delta \mu_t &= a_{\mu t} \qquad & a_{\mu t}|V_\mu \sim N(0, V_\mu) \end{aligned} \tag{2.8}$$

Model (2.8) is known as the local linear trend (see Harvey, 2006). If $V_p = 0$, it reduces to the I(2) plus noise process that is implicitly considered in Hodrick-Prescott (HP) filtering (see Hodrick and Prescott, 1997, and Harvey and Jaeger, 1993). Equation (2.8) introduces one more latent variable, μ_t with associated variance parameter V_μ for which we assume the IG-prior distribution:

$$V_\mu \sim IG(s_{\mu 0}, \nu_{\mu 0}) \tag{2.9}$$

Like previously, the IG-prior hypothesis implies that marginally to the variance parameters, the slope's shocks a_μ re-scaled as $\tau_\mu = a_\mu (s_{\mu 0}/\nu_{\mu 0})^{-1/2}$ follow a standard t-density. Let D_2 denote the $T-2 \times T$ second-difference matrix. In terms of the τ-variables, the measurement equation can be written as:

$$D_2 y = (s_{\mu 0}/\nu_{\mu 0})^{1/2} \tau_\mu + (s_{p0}/\nu_{p0})^{1/2} D_1 \tau_p + (s_{c0}/\nu_{c0})^{1/2} D_2 \tau_c$$

Clearly, for such I(2) models the D_1-matrix dimension is lessened to $T-2 \times T-1$. We have now:

Lemma 3 The marginal likelihood of the local linear trend (2.1)-(2.8) with IG-priors (2.2) and (2.9) verifies:

$$\begin{aligned} f_D(y) &= \pi^{-\frac{T-2}{2}} \Gamma(\frac{\nu}{2}) \Gamma(\frac{\nu_{\mu 0}}{2})^{-1} \Gamma(\frac{\nu_{p0}}{2})^{-1} \Gamma(\frac{\nu_{c0}}{2})^{-1} \\ &\times \int_0^1 \int_0^1 u_1^{\nu_{\mu 0}/2-1} u_2^{\nu_{p0}/2-1} (1-u_1-u_2)^{\nu_{c0}/2-1} \\ &\times \mid \frac{s_{\mu 0}}{u_1} M_\mu + \frac{s_{p0}}{u_2} M_p + \frac{s_{c0}}{1-u_1-u_2} M_c \mid^{-1/2} \\ &\times [1 + D_2 y' (\frac{s_{\mu 0}}{u_1} M_\mu + \frac{s_{p0}}{u_2} M_p + \frac{s_{c0}}{1-u_1-u_2} M_c)^{-1} D_2 y]^{-\frac{\nu}{2}} du_1 du_2 \end{aligned} \tag{2.10}$$

with $\nu = \nu_{\mu 0} + \nu_{p0} + \nu_{c0} + T - 2$, $M_\mu = I_{T-2}$, $M_p = D_1 D_1'$, and $M_c = D_2 D_2'$.∎

Because simultaneous diagonalisations do not extend to the three-matrix case, the double-integration over $(0,1)$ is computationally more demanding than for first-order random walk models. The marginal likelihood for the I(2) plus noise model is obtained by imposing $M_p = 0_{T-2}$, the integration reducing to one dimension. Lemma 3 can be generalized to the m-th order trend plus noise models discussed by Harvey and Trimbur (2003), but we do not develop this point here as these models have been proposed as tools for signal extraction rather than for fitting data.

In the second-order random walk plus noise model, the posterior distribution of the unobservable components marginally to the variance parameters remains a poly-t 2-0 density only when conditioning on one unobservable. Their joint posterior distribution is of unknown form: Corollary 1 does not extend straightforwardly. It is one reason why we prefer to use Dickey's results instead of Richard and Tompa (1980)'s work about poly-t densities. We now turn to STS models with autoregressive dynamics.

3 Dickey-Laplace marginal likelihood for STS models with autoregressive dynamics and IG-variance parameters

When the structural processes contain some autoregressive dynamics, Lemmas 1-3 give the marginal likelihood only conditionally on some parameters. For instance, the damped trend model assumes a stationary 0-mean autoregressive slope such as (see Harvey, 1989, p.46):

$$\mu_t = \phi_\mu \mu_{t-1} + a_{\mu t}$$

Let Σ_μ denote the variance-covariance matrix of the slope up to V_μ, i.e. $\Sigma_\mu = V(\mu_1^{T-1})/V_\mu$. The distribution of μ given ϕ_μ marginally to V_μ is a Student density with precision matrix Σ_μ^{-1}. We have in this case:

Lemma 4 For the damped trend plus noise model with IG-variance priors (2.2) and (2.9), the marginal likelihood of y conditional on ϕ_μ is like in Lemma 3, equation (2.10), with $M_p = I_{T-1}$, $M_c = D_1 D_1'$, $M_\mu = \Sigma_\mu$, and using $D_1 y$ instead of $D_2 y$.■

More often however the dynamic is introduced in the short-term component. Indeed many macroeconomic series display recurrent short-term movements, usually in relationship with the business cycle, and for such series the STS model must complement the long-term trend with a cyclical component. The regularity of the cyclical fluctuations can be reproduced using an AR(2) process with complex roots parameterized in terms of amplitude A and periodicity Per as in:

$$(1 - 2A\cos\frac{2\pi}{Per}L + A^2L^2)c_t = a_{ct} \qquad a_{ct}|V_c \sim N(0, V_c) \tag{3.1}$$

This specification is closely related to the stochastic cycle discussed in Harvey (1989, p.46). The amplitude-periodicity parameterization is appealing as it suits well the prior information available about the business cycle (see Planas, Rossi and Fiorentini, 2007). Here we can let undefined the prior distribution of the polar coordinates, only the distribution of V_c needs to be specified; we keep the assumption (2.2). Let Σ_c denote the variance-covariance matrix of the cycle up to V_c, i.e. $\Sigma_c = V(c)/V_c$. Its inverse gives the precision matrix of the t-marginal distribution of c given A and Per. The following lemma extends the previous results to models with short-term dynamics.

Lemma 5 For STS models such as first and second-order random walk trends plus AR(2)-cycle with IG-variance priors, the marginal likelihood of y given the polar coordinates A and Per is like in Lemmas 1-3 with either $M_c = D_1 \Sigma_c D_1'$ or $M_c = D_2 \Sigma_c D_2'$ according to the model integration order. For the damped trend plus AR(2)-cycle model, the marginal likelihood of y given A, Per and ϕ_μ is like in Lemma 4 with $M_c = D_1 \Sigma_c D_1'$.∎

We thus obtain the marginal likelihood for trend plus cycle decompositions by simple integration over $(0, 1)$, conditionally on the autoregressive parameters, say $\Lambda = (A, Per)$ or (A, Per, ϕ_μ). For integrating Λ out, we suggest to adapt the Laplace method (see Tierney and Kadane, 1986). This method has been used in the STS context for instance by Harvey et al. (2007). It solves the marginal likelihood integral in the neighborhood of the posterior mode using a normal estimates of the posterior density: the more precise the normal approximation around the mode, the better the marginal likelihood evaluation. The strategy we put forward here aims at improving the normal approximation by integrating out the variance parameters using Lemmas 1-5. An improvement is expected because variance parameters are typically the main responsible for the posteriors' departure from normality (see for instance Figure 2 in Harvey et al., 2007). Of course, given the asymptotics at work, the smaller the sample size the larger should be the gain in accuracy compared to standard Laplace applications.

A requirement is that posterior samples of model parameters are available. They can be obtained following the MCMC schemes proposed for instance in Harvey et al. (2007) or in Planas et al. (2007). Let $\tilde{\Lambda}$ denote the posterior mode of the parameters Λ and let $\Sigma(\tilde{\Lambda})$ represent minus the inverted Hessian matrix of the logarithm of the non-normalized marginal posterior $f_D(y|\Lambda)f(\Lambda)$ evaluated at $\tilde{\Lambda}$:

$$\Sigma(\tilde{\Lambda}) = -[\frac{\partial^2 \ log\{f_D(y|\Lambda)f(\Lambda)\}}{\partial\Lambda'\partial\Lambda}\ |_{\theta=\tilde{\Lambda}}]^{-1} \qquad (3.2)$$

For normal posteriors, (3.2) coincides with the parameters variance-covariance matrix. The second-order expansion of the non-normalized log-posterior around its mode is such that:

$$log\ \{f_D(y|\Lambda)f(\Lambda)\} \simeq log\ \{f_D(y|\tilde{\Lambda})f(\tilde{\Lambda})\} - \frac{1}{2}(\Lambda - \tilde{\Lambda})\Sigma(\tilde{\Lambda})^{-1}(\Lambda - \tilde{\Lambda})'$$

The last term above takes the form of the kernel of a normal distribution with mean $\tilde{\Lambda}$ and variance-covariance matrix $\Sigma(\tilde{\Lambda})$. Exponentiating and integrating out Λ yields:

$$f_{DL}(y) = (2\pi)^{d/2}|\Sigma(\tilde{\Lambda})|^{1/2} f_D(y|\tilde{\Lambda})f(\tilde{\Lambda}) \qquad (3.3)$$

where d is the dimension of Λ. We shall refer to equation (3.3) as the Dickey-Laplace marginal likelihood estimates. In (3.3), the term $f(\tilde{\Lambda})$ assigns a prior weight to the posterior mode while $f_D(y|\tilde{\Lambda})$ is the model likelihood marginally to the IG-variance parameters as given by Lemmas 1-5.

4 Comparison with bridge sampling

We evaluate the Dickey (D), Laplace (LP), and Dickey-Laplace (D-LP) marginal likelihoods against the Meng-Wong (MW, 1996) estimates in a simulation exercise. During these last two decades, econometricians have often resorted to importance sampling for computing marginal likelihoods (see Kloek and Van Dijck, 1978; Geweke, 1989). MW's technique is an extension that re-weights both the importance function and the posterior density through a bridge function. Given that a consensus seems to be emerging about the potential superiority of MW's technique over the other estimators available, we adopt here this method as benchmark (see Meng and Schilling, 1996; diCiccio et al., 1997; and Fruhwirth-Schnatter, 2004).

Let S and S_q denote the support of the parameter posterior distribution and of an importance function, say $q(\theta)$. Let also $h(\theta)$ represent a function defined over $S \bigcap S_q$. The MW marginal likelihood estimate is obtained from (see also Gelman and Wong, 1998):

$$f(y) = \frac{\int_{S_q} \frac{h(\theta)}{q(\theta)} dq(\theta)}{\int_S \frac{h(\theta)}{f(y|\theta)f(\theta)} df(\theta|y)}$$

Equivalent formulations are sometimes given in terms of a function $\gamma(\theta)$ such that $\gamma(\theta)q(\theta)f(\theta|y) = h(\theta)$ (see Fruhwirth-Schnatter, 2004). The bridge function $h(\theta)$ reduces the estimation error when located at an intermediate position between the importance function and the parameter posterior distribution. MW propose as optimal choice a recursive procedure based on:

$$h(\theta) \propto \frac{q(\theta)f(\theta|y)}{n_q q(\theta) + n_y f(\theta|y)}$$

where the constants n_q and n_y refer to the number of draws from the importance function and from the posterior density, respectively. The recursions are introduced through the term $f(\theta|y)$ that involves a preliminary marginal likelihood estimate. We initialize the algorithm using the Laplace approximation and then iterate for ten rounds; no further sampling is needed for iterating. The MW estimator can also be built around likelihood functions marginal to the variance coefficients, i.e. using $f_D(y|\Lambda)$ in place of $f(y|\theta)$. Because Dickey's integral would need to be evaluated for every sample out of the importance function, we discard this possibility for its computational cost.

We simulate three series of respective length $T = 50, 100, 250$ from a random walk with drift plus AR(2) cycle like in (2.6)-(3.1). The coefficients are set to $\mu_p = .1$, $A = .8$, $Per = 10$, $V_p = .01$ and $V_c = .05$. The marginal likelihood of the simulated series is estimated using the LP, D-LP and MW methods for eight models obtained as combinations of four trend models, i.e. I(2), integrated random walk (irw), random walk plus drift (rw), and damped trend (dt), with two models for the cycle, i.e. the white noise (wn) and the autoregressive model (ar2) in (3.1). The prior distributions are omitted for the sake of space. For each model, we record two thousand samples from the parameters posterior distribution out of two hundred thousands simulations using the Gibbs sampling scheme detailed in Planas et al. (2007), after a burn-in of ten thousand iterations. The sparse recording serves at lowering correlations. This MCMC output is then used to compute the marginal likelihoods, and the whole computations are repeated twenty times in order to get numerical averages and standard deviations. Lemmas 1-5 integrals are calculated over grids of one thousand points in dimension one, and over squares of four hundred points in each side in dimension two; their standard deviation is neglected as of irrelevant size. Notice that when the STS model includes some autoregressive dynamics, the normalizing constants of the full conditional distributions are not entirely known, so Chib (1995)'s marginal likelihood estimator does not apply.

Table 1 displays the results. The average marginal likelihoods are reported with a minus sign and the numerical standard deviations lie between brackets. The models are ranked in increasing number of parameters, from the I(2) plus noise model with 2 parameters to the damped trend plus AR(2) cycle with 6 parameters. For models with only variance coefficients, Dickey and MW estimates are in close agreement: the differences are of the three-digit order, whatever the sample size. When dynamic parameters are introduced, the deviations get to the one-digit order. The error in the LP estimates can instead reach a unit, especially in short sample. As can be seen, marginalizing out the variance parameters always improves the approximation. For the models and sample sizes considered, the improvement is such as to make the D-LP estimate almost as accurate as the MW one. This result is interesting because no further sampling from an importance function is needed with the D-LP approach.

All marginal likelihood estimates point to the random walk with drift plus AR(2) cycle as the most adequate model. Mispecifying the short-term dynamics implies quite a large drop in the marginal likelihood. It could be argued that the Laplace estimator remains useful for model discrimination in spite of the approximation errors, but such a conclusion depends on the discrepancies between the alternatives considered. We shall see in the next Section that discriminating between models with comparable properties can become difficult with this estimator. Moreover, when the model misses some important pattern such as the short-term dynamic, the mispecification can yield posterior distributions with bi-modal characteristics. In such cases the LP marginal likelihood is unreliable. For this experiment we contained this problem by carefully tuning the prior distributions and, in a few cases, by trimming the output.

Table 1 Minus average log marginal likelihood

Trend	Cycle	n = 25					
		MW		LP		D-LP	
i2	wn	23.772	[.015]	23.143	[.092]	23.765	[—]
irw	wn	23.200	[.022]	20.925	[.101]	23.202	[—]
rw	wn	21.378	[.021]	19.510	[.169]	21.378	[—]
dt	wn	21.368	[.023]	19.484	[.123]	21.504	[.088]
i2	ar2	11.012	[.025]	9.693	[.106]	10.286	[.084]
irw	ar2	11.230	[.027]	9.113	[.138]	10.532	[.084]
rw	ar2	9.075	[.026]	7.406	[.179]	8.383	[.084]
dt	ar2	9.700	[.028]	7.915	[.115]	9.170	[.090]
		n = 100					
		MW		LP		D-LP	
i2	wn	77.607	[.011]	77.512	[.079]	77.606	[—]
irw	wn	68.518	[.014]	68.009	[.298]	68.518	[—]
rw	wn	64.946	[.011]	64.840	[.058]	64.946	[—]
dt	wn	66.744	[.080]	64.191	[.311]	66.930	[.140]
i2	ar2	21.893	[.016]	21.428	[.116]	21.707	[.080]
irw	ar2	21.702	[.021]	20.872	[.090]	21.551	[.090]
rw	ar2	15.627	[.021]	14.928	[.148]	15.492	[.076]
dt	ar2	24.612	[.021]	23.270	[.176]	24.230	[.132]
		n = 250					
		MW		LP		D-LP	
i2	wn	141.749	[.009]	141.763	[.082]	141.746	[—]
irw	wn	124.670	[.013]	124.444	[.077]	124.697	[—]
rw	wn	114.732	[.011]	114.751	[.109]	114.730	[—]
dt	wn	114.727	[.018]	114.348	[.119]	114.717	[.034]
i2	ar2	54.653	[.015]	54.345	[.131]	54.468	[.080]
irw	ar2	54.817	[.018]	54.192	[.158]	54.720	[.125]
rw	ar2	37.354	[.016]	36.860	[.100]	37.251	[.084]
dt	ar2	57.289	[.019]	56.439	[.174]	57.065	[.091]

Notes: MW Meng-Wong, LP Laplace, D-LP Dickey-Laplace. Models: rw random walk, irw integrated rw, i2 I(2), dt damped trend, wn white noise. Standard deviations between brackets.

5 Application to the euro area and US NAIRUs

We apply this methodology to the analysis of the NAIRU in the euro area and in the US. The NAIRU is of particular interest because it is related to the imperfect equilibrium of the labor market. The European Commission uses it for evaluating the potential growth of Member States and for the cyclical adjustment of their budget balances, in application of the Stability and Growth Pact (see Denis, Grenouilleau, Mc Morrow and Roeger, 2006). Central Banks also scrutinize the NAIRU but for assessing the inflation pressures, following Phillips curve theory (see for instance Stiglitz, 1997, and Ball and Mankiw, 2004).

Characterizing the NAIRU is however difficult, mainly because of its unobserved and changing nature. Staiger, Stock and Watson (1997) underlined the lack of precision of estimates obtained with standard specifications. Also, although its changing nature is now well-accepted, not much is known about its actual variability. For instance, the widely-used HP filter requires a prior hypothesis about the signal to noise ratio, but this hypothesis is rarely confronted to the data. Here we take advantage of the Bayesian framework to address the following questions: which STS models best describe the euro area and the US NAIRUs? How smooth are they? And how precise can be their univariate STS estimates?

The euro area unemployment series has been collected from AMECO, the national accounts database of the EC Directorate General Economic and Financial Affairs, available at *europa.eu.int/comm/economy_finance* following the link Indicators. The US data have been downloaded from the Bureau of Labor Statistics web-site *www.bls.gov*. Both series are annual averages over 1960-2007, the last figure being preliminary. Following standard practice in the NAIRU literature, we describe these two series as made up of a cycle plus a trend. The cyclical dynamics are represented with an AR(2) process parameterized as in (3.1), with amplitude and periodicity parameters assumed to be Beta-distributed. The prior distribution of the former is tuned so as to yield an average amplitude of 0.8 for the euro-area and of 0.7 for the US, in agreement with empirical business cycle studies (see for instance Kuttner, 1994, Gerlach and Smets, 1999). The standard deviations are set to one-tenth of the mean so as to not impose too much precision. Namely, we use a $Beta(19.2, 4.8)$ for the euro-area cycle amplitude and a $Beta(29.3, 12.6)$ for the US one. The periodicity parameter is also assumed to be Beta-distributed, with support translated to $[2, 48]$ given the sample length. Still according to business cycle studies, we tune the periodicity prior distribution so as to get cycles of mean length 9 years for the euro-area and 8 years for the US, with standard deviations of 2 and 1.5 time periods, respectively. This is obtained with the distributions $(Per-2)/(48-2) \sim Beta(10.2, 57)$ for the euro-area and $Beta(13.8, 91.9)$ for the US. Finally, the IG-distribution for the short-term shocks variance has been set so as to add a mean deviation of 0.5 for euro and 0.7 for US, with the distributions $IG(1.9, 9.6)$ and $IG(2.6, 7.1)$ respectively.

For the NAIRU we consider the four specifications discussed in Section 2, namely the driftless random walk, the I(2), the integrated random walk and the damped trend. Three different prior distributions are used for the variance parameters V_p and V_μ: IG(.08,6), IG(.28,6) and IG(.80,6). These three priors imply increasing means at 0.02, 0.07, and 0.20; we shall refer to them as low (L), medium (M) and high (H). As can be seen in Figure 1, they cover quite a wide range of patterns. For its empirical relevance we also let the damped trend model have no level shocks, i.e. $V_p = 0$. Finally, the prior for the slope autoregressive parameter ϕ_μ has been set to the Normal distribution $N(.85, 1/30)I_{(0,1)}$ truncated to the stationary positive region. Altogether, the combination of the four trend specifications with the different variance priors yields twenty-seven models for the NAIRU.

Table 2 reports the posterior probabilities of each model marginally to the model parameters, i.e. $p(M_i|Y)$, $i = 1, ...27$. These posterior probabilities have been computed using the Dickey-Laplace approximation to the marginal likelihood discussed in Section 3. Table 2 displays the models in decreasing order of relevance. For both euro area and US series, the first five models receive a total posterior weight greater than 50%. Of these best fitting models, all but one are integrated of order 1: the data strongly support the I(1) hypothesis. With all I(2) models ranked last, the evidence is particularly striking for the US. This result can be related to the failure of I(2) models to produce reasonable long-term forecasts of unemployment rate series. The data also express an overwhelming preference for the damped trend model, i.e. the model that accounts for a time-varying slope with moderate persistence. Finally, the euro area NAIRU seems to have received larger shocks on its slope than on its level, perhaps explaining why for euro area some I(2) models receive a relevant posterior weight. On the contrary, the shocks on US structural unemployment seem to have hit mostly its level, an observation that would plead against the use of the HP filter for inferring about the US NAIRU.

Table 2 Posterior model probabilities

	euro-area			US		
Rank	Trend	V_p-V_μ	$p(M_i \mid Y)$	Trend	V_p-V_μ	$p(M_i \mid Y)$
1	**dt**	**0-M**	.126	**rw**	**H-0**	.141
2	**dt**	**0-L**	.114	**dt**	**H-L**	.113
3	**dt**	**L-L**	.108	**dt**	**M-L**	.098
4	**dt**	**L-M**	.102	**dt**	**L-L**	.085
5	**i2**	**0-L**	.089	**dt**	**0-L**	.081
6	**irw**	**L-L**	.081	**rw**	**M-0**	.065
7	**dt**	**M-L**	.058	**dt**	**0-M**	.064
8	**dt**	**M-M**	.052	**dt**	**M-M**	.062
9	**irw**	**M-L**	.047	**dt**	**H-M**	.061
10	**dt**	**0-H**	.043	**dt**	**L-M**	.055
11	**i2**	**0-M**	.034	**dt**	**L-H**	.027
12	**irw**	**L-M**	.033	**dt**	**M-H**	.024
13	**dt**	**L-H**	.029	**dt**	**0-H**	.024
14	**irw**	**M-M**	.017	**rw**	**L-0**	.019
15	**dt**	**M-H**	.014	**dt**	**H-H**	.019
16	**dt**	**H-L**	.014	**irw**	**H-L**	.012
17	**irw**	**H-L**	.010	**irw**	**M-L**	.011
18	**dt**	**H-M**	.010	**irw**	**L-L**	.011
19	**i2**	**0-H**	.006	**i2**	**0-L**	.010
20	**irw**	**L-H**	.005	**irw**	**M-M**	.004
21	**irw**	**H-M**	.003	**irw**	**H-M**	.004
22	**dt**	**H-H**	.002	**irw**	**L-M**	.003
23	**irw**	**M-H**	.002	**i2**	**0-M**	.003
24	**irw**	**H-H**	.000	**irw**	**M-H**	.001
25	**rw**	**H-0**	.000	**irw**	**L-H**	.001
26	**rw**	**M-0**	.000	**irw**	**H-H**	.001
27	**rw**	**L-0**	.000	**i2**	**0-H**	.001

Notes: Priors for variance parameters: for $\ell = p, \mu$, L $\Leftrightarrow V_\ell \sim IG(.08, 6)$; M $\Leftrightarrow V_\ell \sim IG(.28, 6)$; H $\Leftrightarrow V_\ell \sim IG(.80, 6)$.

Given Section 3 simulations, the robustness of the model classification to the marginal likelihood estimator is worth verifying. For the first five models, Table 3 reports the posterior weights and rankings obtained with the MW and LP estimators. The D-LP results are also displayed for easing comparison. As can be seen, D-LP and MW are in close agreement: the differences between the posterior weights are less than 10% of the estimates and the classification is changed only in one occasion, two successive models being permuted. We can thus be confident about Table 2 results. The LP outcome is instead quite different: the posterior weights show variations that can reach 100% of the estimates and the ranking is upset. Hence for these models we consider the LP marginal likelihood as unreliable. Probably because the series sample size is not large enough to make the posterior distribution of the variance parameters approximately Normal.

Table 3

euro-area

Trend	V_p-V_μ	D-LP		MW		LP	
		Rank	$P(M_i\|Y)$	Rank	$P(M_i\|Y)$	Rank	$P(M_i\|Y)$
dt	**0-M**	1	.126	1	.132	5	.088
dt	**0-L**	2	.114	2	.107	3	.112
dt	**L-L**	3	.108	4	.097	1	.167
dt	**L-M**	4	.102	3	.105	2	.121
i2	**0-L**	5	.089	5	.088	6	.063

US

Trend	V_p-V_μ	D-LP		MW		LP	
		Rank	$P(M_i\|Y)$	Rank	$P(M_i\|Y)$	Rank	$P(M_i\|Y)$
rw	**H-0**	1	.141	1	.138	6	.062
dt	**H-L**	2	.113	2	.106	3	.110
dt	**M-L**	3	.098	3	.101	4	.115
dt	**L-L**	4	.085	4	.088	1	.200
dt	**0-L**	5	.081	5	.076	7	.061

The best fitting model is, for euro area, the damped trend with no level shocks and for US, the random walk without drift. This last has been frequently used in empirical studies of the US NAIRU, for instance by Staiger, Stock and Watson (1997) and by Gordon (1998). Figure 2 shows the corresponding estimates. As can be seen, the euro area NAIRU is continuously decreasing since the mid-1990's peak. It is tempting to see here the effect of new regulations for increasing the flexibility of euro area labor markets, following the European Employment Strategy (1997) within the Lisbon agenda. The US NAIRU seems to be almost constant in the last ten years, after fifteen years of steady decrease between 1982 and 1997. As expected given the respective labor markets flexibility, it embodies more short-term dynamics than the euro-area one.

In order to analyze the respective smoothness, we compare the ratios between the variance of the cycle and the variance of the trend second difference, i.e. $V(c_t)/V(\Delta^2 p_t)$. This quantity is a slight generalization of the inverse signal to noise ratio typically considered in HP filtering. More elaborate measures of smoothness have been proposed in the literature (see for instance Froeba and Koyak, 1994), but the acquaintance of economists with the HP filter gives such variance ratios the advantage of immediacy. Figure 3 shows the posterior distribution of the generalized inverse signal to noise ratio; the continuous line refers to the best model and the dashed one is obtained marginally to the model choice. As can be seen, the US NAIRU participates more to the unemployment fluctuations than the euro area one, and this evidence is strong enough to hold marginally to the model choice: the variance ratio mode is about 2 for the US against 8.0 for the euro area with the best model, 5.0 after model marginalizing. These are the ratios we would recommend should the HP-filter be used for detrending euro area and US annual unemployment; they are not too far from the values advised by Ravn and Uhlig (2002). Notice that the posterior distribution of the variance ratio is quite diffuse for the euro area, perhaps reflecting a substantial time-varying behavior.

Finally, Figure 4 shows the posterior distribution of the 2007 NAIRUs for euro area and the US. Again, the continuous line refers to the best model and the dashed one corresponds to the model average. The current NAIRU is measured about 7.5 for euro area and about 5-5.1 for the US. This result is obtained with both the best model and marginally to the model specification. A 95% confidence band around the modes covers about 2.6 points, with the interval (6.2,8.9) for euro area and (3.7,6.3) for the US. This is comparable with the uncertainty that Staiger et al. (1997) reported for the US NAIRU in 1991's first quarter using also inflation data. There is a close matching between the posterior distribution obtained with the best model and the one obtained marginally to the model specification, mainly because the models that receive the highest posterior weights yield similar NAIRU estimates. Hence, so long a reasonable model is used like for instance the first five in Table 2, researchers should not worry too much about model uncertainty.

6 Conclusion

We obtain simple expressions for evaluating the marginal likelihood of STS models by taking benefit of the model properties and of the results in Dickey (1968). For trend plus noise models, they only involve an integration over the (0,1)-support. For trend plus cycle models, we show that coupling this approach with the Laplace method yields a substantial gain in accuracy with respect to traditional Laplace marginal likelihood estimator. Overall the precision is comparable to that of the MW estimator, without requiring any importance sampling.

We apply this discrimination tool to the analysis of the euro area and US NAIRU. As best model, we found a damped trend for the euro area and the driftless random walk for the US; these would be our recommendation to practitioners. The NAIRU smoothness seems in broad agreement with the inverse signal to noise ratio suggested by Ravhn and Uhlig (2002) for HP-detrending annual data. Model uncertainty does not seem to add much variation to the NAIRU estimates, at least so long a reasonable model is used. We could see that conducting this analysis with the traditional version of the Laplace marginal likelihood gives misleading results, perhaps because of the limited sample size. The methodology we propose can be extended to STS models including a third unobserved variable such as the irregular component, and also to bivariate system such as the Kuttner (1994) Phillips-curve augmented model for output gap.

References

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

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

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

CHIB S. (1995), "Marginal likelihood from the Gibbs output", *Journal of the American Statistical Association*, 90, 432, 1313-1321.

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

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

COGLEY T. AND SARGENT T.J (2005), "Drifts and volatilities: monetary policies and outcomes in the post war US", *Review of Economic Dynamics*, 8, 262-302.

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

DENIS C., GRENOUILLEAU D., MC MORROW K. AND ROEGER W. (2006), "Calculating potential growth rates and output gaps - A revised production function approach", European Economy, Economic Papers nr. 246, European Commission.

DICICCIO T.J, KASS R.E., RAFTERY A., AND WASSERMAN L. (1997), "Computing Bayes factors by combining simulation and asymptotic approximations", *Journal of the American Statistical Association*, 91, 435, 1254-1267.

DICKEY J. (1968), "Three multidimensional-integral identities with Bayesian applications", *Annals of Mathematical Statistics*, 39, 5, 1615-1637.

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

FROEBA L. AND KOYAK R. (1994), "Measuring and comparing smoothness in time series: the production smoothing hypothesis", *Journal of Econometrics*, 64, 1-2, 97-122.

FRUHWIRTH-SCHNATTER S. (1994), "Data augmentation and dynamic linear models", *Journal of Time Series Analysis*, 15, 2, 183-202.

FRUHWIRTH-SCHNATTER S. (2004), "Estimating marginal likelihoods for mixture and

Markov switching models using bridge sampling techniques", *Econometrics Journal*, 7, 143-167.

GELMAN A. AND WONG X.L. (1998), "Simulating normalizing constants: from importance sampling to bridge sampling to path sampling", *Statistical Science*, 13, 2, 163-185.

GEWEKE J. (1989), "Bayesian inference in econometric models using Monte-Carlo integration", *Econometrica*, 57, 1317-1339.

HALL R.E. AND MISHKIN F.S. (1982), "The sensitivity of consumption to transitory income: estimates from panel data on households", *Econometrica*, 50, 2, 461-481.

HANSEN G. (1997), "Technical progress and aggregate fluctuations", *Journal of Economic Dynamics and Control*, 21, 6, 1005-1023.

HARVEY A.C. (1989), *Forecasting Structural Time Series Models and the Kalman Filter*, Cambridge University Press: Cambridge.

HARVEY, A.C. (2001), "Testing in unobserved components models", *Journal of Forecasting*, 20, 1-19.

HARVEY, A.C. (2006), "Forecasting with unobserved components time series", Chapter 7 of *Handbook of Economic Forecasting*, ed. by G.Elliot, Granger C.W. and A. Timmermann, Elsevier.

HARVEY, A.C. AND JAEGER, A. (1993), "Detrending, stylised Facts and the business cycles", *Journal of Applied Econometrics*, 8, 231-247.

HARVEY, A.C. AND TRIMBUR T.M. (2003), "General model-based filters for extracting cycles and trends in economic time series", *Review of Economics and Statistics*, 85, 2, 244-255.

HARVEY A.C., TRIMBUR T.M. AND VAN DIJK H.K. (2007), "Trends and cycles in economic time series: a Bayesian approach", *Journal of Econometrics*, 140, 2, 618-649.

HODRICK R.J. AND E.C. PRESCOTT (1997), "Post-war U.S. business cycle: an empirical investigation", *Journal of Money, Credit and Banking*, 29, 1, 1-16.

KLOEK T. AND VAN DIJCK H.K (1978), "Bayesian estimates of equation system parameters: an application of integration by Monte Carlo", *Econometrica*, 46, 1-19.

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

MAGNUS J.R. AND NEUDECKER H. (1988), *Matrix Differential Calculus with Applications in Statistics and Econometrics*. New York: Wiley.

Meng X.L. and Schilling S. (1996), "Fitting full-information item factor models and an empirical investigation of bridge sampling", *Journal of the American Statistical Association*, 92, 439, 903-915.

Meng X.L. and Wong W.H. (1996), "Simulating ratios of normalizing constants via a simple identity: a theoretical exploration", *Statistica Sinica*, 6, 831-860.

Pagan A. (1975), "A note on the extraction of components from time series", *Econometrica*, 43, 1, 163-183.

Planas C., Roeger W. and Rossi A. (2007), "How much has labour taxation contributed to European structural unemployment?", *Journal of Economic Dynamics and Control*, 31, 1359-1375.

Planas C., Rossi A. and Fiorentini G. (2007), "Bayesian analysis of the output gap", *Journal of Business & Economic Statistics*, forthcoming.

Richard J.F. and Tompa H. (1980), "On the evaluation of poly-t density functions", *Journal of Econometrics*, 12, 335-351.

Staiger D., Stock J.H. and Watson M.W. (1997), "How precise are estimates of the natural rate of unemployment?", in *Reducing Inflation: Motivation and Strategy*, by C.D. Romer and D.H. Romer, eds. Chicago: University of Chicago Press, 195-246.

Stiglitz, J. (1997), "Reflections on the natural rate hypothesis", *Journal of Economic Perspectives*, 11, 1, 3-10.

Stock J.H. and Watson M.W. (1988), "Variables trends in economic time series", *Journal of Economic Perspectives*, 2, 147-174.

Stock J.H. and Watson M.W. (2007), "Why has US inflation become harder to forecast?", *Journal of Money, Credit and Banking*, 39, 1, 3-33.

Tierney L. and Kadane J.B. (1986), "Accurate approximations for posterior moments and marginal densities", *Journal of the American Statistical Association*, 81, 393, 82-86.

Figure 1
Prior distributions for variance parameters V_p and V_μ

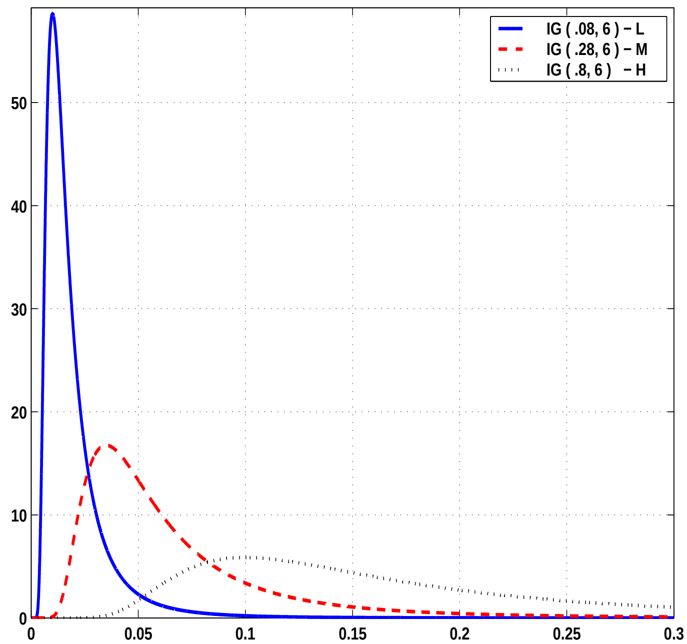


Figure 2
Unemployment, NAIRU (——) and 95% confidence bands (– –)

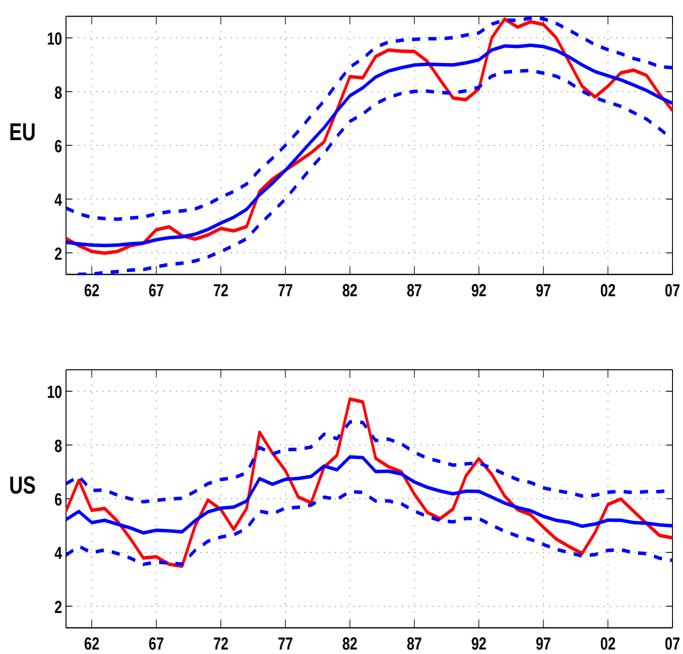
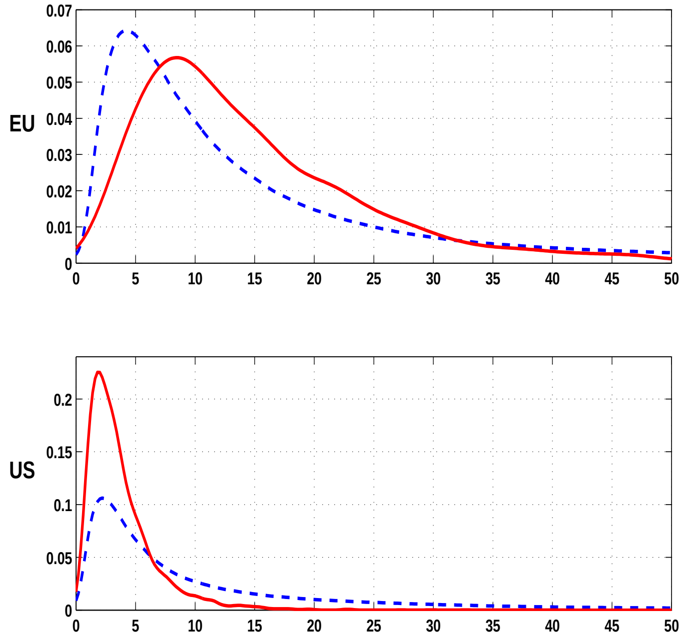


Figure 3

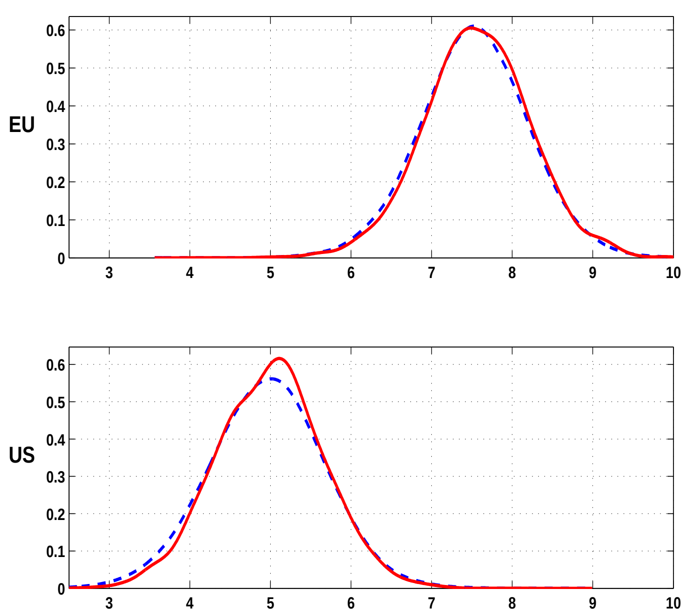
Variance ratio $V(c)/V(\Delta^2 p)$ posterior distribution



—— Best model – – Model average

Figure 4

Posterior distribution of 2007 NAIRU



—— Best model – – Model average