

Trends and Cycles in Economic Time Series: A Bayesian Approach

Andrew C. Harvey^a, Thomas M. Trimbur^{a,b}, Herman K. van Dijk^c

^a Cambridge University, Faculty of Economics and Politics

^b U.S. Bureau of the Census, Washington DC

^c Econometric Institute, Erasmus University

July 25, 2005

Abstract

Trends and cyclical components in economic time series are modeled in a Bayesian framework. This enables prior notions about the duration of cycles to be used, while the generalized class of stochastic cycles employed allows the possibility of relatively smooth cycles being extracted. The posterior distributions of such underlying cycles can be very informative for policy makers, particularly with regard to the size and direction of the output gap and potential turning points. From the technical point of view a contribution is made in investigating the most appropriate prior distributions for the parameters in the cyclical components and in developing Markov chain Monte Carlo methods for both univariate and multivariate models. Applications to US macroeconomic series are presented.

KEYWORDS: Output gap, Kalman filter, Markov chain Monte Carlo, real-time estimation, turning points, unobserved components.

JEL classification: C11, C32, E32

1 Introduction

Decomposing time series into trends and cycles is fundamental to a good deal of macroeconomic analysis. Key features of the ‘business’ cycle, such as length and turning points, are of great interest to policy makers in industry

and government. Similarly, potential output trends and their deviations from the actual level of output yield important signals on the performance of an economy.

Trends and cycles may be modeled¹ directly as unobserved components within the framework of structural time series models (STMs); see Harvey and Jaeger (1993). The statistical treatment is based on the state space form with the components in a linear model extracted by the Kalman filter and associated smoother. For a Gaussian model, the likelihood function is obtained from the innovations produced by the Kalman filter and maximized numerically with respect to the unknown parameters. This classical procedure is implemented in the STAMP package of Koopman et al (2000). However, fitting the standard trend plus cycle plus irregular model to time series of Gross Domestic Product (GDP) often results in the irregular component disappearing with the result that the cycle is quite noisy. More generally, maximum likelihood (ML) can sometimes produce implausible parameter values, resulting perhaps in trends that are too inflexible or cycles that have too long a period. The higher order stochastic cycles introduced recently by Harvey and Trimbur (2003) tend to produce smoother extracted cycles, but problems of implausible estimates still remain. This provides one of the motivations for investigating a Bayesian approach.

A key parameter in the stochastic cycle is the period around which most of the power of the spectrum is concentrated. In building models to capture business cycles, it is not unreasonable to take on board prior notions about the period. These may be incorporated into the model in a flexible way; we do this here using a beta prior distribution. Previous work on using Bayesian methods for STMs, such as Durbin and Koopman (2002) and Koop and van Dijk (2000), has not dealt with cycles. Huerta and West (1999) study cyclical behavior indirectly using autoregressive models, but this approach does not make it easy to use prior information on periodicity.

We present a Markov chain Monte Carlo (MCMC) algorithm in order to compute posterior results on parameters, model probabilities and unobserved components. The treatment of cycles introduces a number of new issues that

¹Our concern here is with a model-based approach. Although detrending methods such as the Hodrick-Prescott filter are popular in macroeconomics, they can be misleading when used inappropriately, as argued in Harvey and Jaeger (1993) and Cogley and Nason (1995). The same is true of the more recent band-pass filter of Baxter and King (1999); see Harvey and Trimbur (2003) and Murray (2003). Furthermore, model-based approaches have the advantage that measures of uncertainty can be attached to the output.

need to be addressed. With the aid of state space modeling techniques, we set out an efficient procedure for computing the joint posterior density of parameters and components based on Gibbs sampling. In doing so we draw on earlier work on the efficient smoothing of unobserved components by, amongst others, Carter and Kohn (1994), Frühwirth-Schnatter (1994), de Jong and Shephard (1995) and Durbin and Koopman (2002).

While one of the potential advantages of a Bayesian approach is that it is able to avoid fitting implausible models, another is that it can yield more informative results. For example, many of the parameters in STMs are variances. The small sample behaviour of ML estimators of such parameters is not easy to pin down, but it is certainly the case that distributions can be very far from normality. When the true value of a variance is zero, the asymptotic theory is non-standard. One response to gauging the significance² of estimated variances is to use the bootstrap as documented in Stoffer and Wall (2004). The Bayesian approach provides another, possibly more attractive, line of attack by offering the opportunity to examine posterior distributions.

Another motive for investigating a Bayesian approach is that it allows for parameter uncertainty in the posterior distributions of components. This is important in the present context since one of our concerns is to present information on the size of the output gap as represented by the cycle. It is also straightforward to compute statistics such as the probability that the cycle, or its rate of change, is negative. Of course, a Bayesian approach also allows for parameter uncertainty in the predictive distributions of future observations.

Applications are based on quarterly US macroeconomic time series from 1947 onwards. Univariate methods are illustrated using real GDP and features of the cycle such as its duration, turning points and time-varying amplitude are analyzed. We also demonstrate on-line analysis of the size and direction of the output gap. A multivariate model is then fitted to consumption, investment and GDP, so providing a contrast with the well-known study of King, Plosser, Stock, and Watson (1991). Finally, a bivariate model of inflation and output, developed from the ideas of Kuttner (1994), is used to estimate the output gap by exploiting the Phillips curve relationship.

The paper is arranged as follows. Section 2 reviews the extension of the class of structural time series models to include higher order cyclical

²Valid tests of the null hypothesis that the variance is zero can be carried out, but this usually requires that the restricted model be estimated; see Harvey (2001).

components and discusses ways of capturing their features and assessing the size and direction of the output gap. The Bayesian treatment is developed in section 3, while section 4 applies the methods to US GDP. The multivariate applications are in section 5 and section 6 concludes. Technical details on the state space form and the Markov chain Monte Carlo algorithms for the univariate and multivariate cases are laid out in an appendix.

2 Structural time series model for trends and cycles

Define the $N \times 1$ vector of observations \mathbf{y}_t , where $\mathbf{y}_t = (y_t^1, \dots, y_t^N)'$. The class of multivariate structural time series models under consideration consists of trend, cycle and irregular components, denoted by $N \times 1$ vectors $\boldsymbol{\mu}_t$, $\boldsymbol{\psi}_{n,t}$ and $\boldsymbol{\varepsilon}_t$ respectively. Thus

$$\mathbf{y}_t = \boldsymbol{\mu}_t + \boldsymbol{\psi}_{n,t} + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim NID(\mathbf{0}, \boldsymbol{\Sigma}_\varepsilon), \quad t = 1, \dots, T, \quad (1)$$

where $NID(\mathbf{0}, \boldsymbol{\Sigma}_\varepsilon)$ denotes that the vector is serially independent and normally distributed with zero mean vector and $N \times N$ positive semi-definite covariance matrix, $\boldsymbol{\Sigma}_\varepsilon$. The stochastic trend is a multivariate integrated random walk

$$\begin{aligned} \boldsymbol{\mu}_t &= \boldsymbol{\mu}_{t-1} + \boldsymbol{\beta}_{t-1}, \\ \boldsymbol{\beta}_t &= \boldsymbol{\beta}_{t-1} + \boldsymbol{\zeta}_t, \quad \boldsymbol{\zeta}_t \sim NID(0, \boldsymbol{\Sigma}_\zeta), \end{aligned} \quad (2)$$

where $\boldsymbol{\beta}_t$ is the vector of slopes. A seasonal component can easily be added if appropriate.

The vector $\boldsymbol{\psi}_{n,t}$ is a generalization of the similar cycle model of Harvey and Koopman (1997). The aim of the generalization, originally proposed by Harvey and Trimbur (2003), is to include higher order models that tend to produce smoother extracted cycles.

2.1 Generalized stochastic cycles

An $n - th$ order univariate cycle is defined by

$$\begin{aligned} \begin{bmatrix} \psi_{1,t} \\ \psi_{1,t}^* \end{bmatrix} &= \rho \begin{bmatrix} \cos \lambda_c & \sin \lambda_c \\ -\sin \lambda_c & \cos \lambda_c \end{bmatrix} \begin{bmatrix} \psi_{1,t-1} \\ \psi_{1,t-1}^* \end{bmatrix} + \begin{bmatrix} \kappa_t \\ \kappa_t^* \end{bmatrix}, \\ \begin{bmatrix} \kappa_t \\ \kappa_t^* \end{bmatrix} &\sim NID \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_\kappa^2 & 0 \\ 0 & \sigma_\kappa^2 \end{bmatrix} \right) \end{aligned} \quad (3)$$

$$\begin{bmatrix} \psi_{i,t} \\ \psi_{i,t}^* \end{bmatrix} = \rho \begin{bmatrix} \cos \lambda_c & \sin \lambda_c \\ -\sin \lambda_c & \cos \lambda_c \end{bmatrix} \begin{bmatrix} \psi_{i,t-1} \\ \psi_{i,t-1}^* \end{bmatrix} + \begin{bmatrix} \psi_{i-1,t-1} \\ \psi_{i-1,t-1}^* \end{bmatrix}, \quad i = 2, \dots, n \quad (4)$$

The parameter λ_c denotes frequency in radians while ρ is a damping factor lying between zero and one; if it is equal to one, the cycle is nonstationary. The disturbances, κ_t and κ_t^* , are assumed to be uncorrelated with each other and with the disturbances driving the other components. Harvey and Trimbur (2003) show that as n increases the signal extraction filter for a cycle plus noise model tends towards a band pass filter as in Baxter and King (1999). General expressions for the variance, autocovariances and spectrum are given in Trimbur (2005).

In a similar cycle model ρ and λ_c are the same across all series. Therefore, the cycles have the same dynamic properties in the sense that their autocorrelation functions and spectral densities are identical. However, the cycles themselves are not, in general, identical. The similar cycle model, originally formulated for $n = 1$, may be extended to higher order cycles by defining a $2nN \times 1$ state vector

$$\boldsymbol{\psi}_t = [\psi_{n,t}^1, \dots, \psi_{n,t}^N, \psi_{n,t}^{1*}, \dots, \psi_{n,t}^{N*}, \dots, \psi_{1,t}^1, \dots, \psi_{1,t}^N, \psi_{1,t}^{1*}, \dots, \psi_{1,t}^{N*}]' \quad (5)$$

where the sub-vector $\boldsymbol{\psi}_{n,t} = [\psi_{n,t}^1, \dots, \psi_{n,t}^N]'$ appears in (1). Define the matrix

$$\mathbf{T}_n = \mathbf{I}_n \otimes \mathbf{T} + \mathbf{S}_n \otimes \mathbf{I}_2 \quad (6)$$

where

$$\mathbf{T} = \rho \begin{bmatrix} \cos \lambda_c & \sin \lambda_c \\ -\sin \lambda_c & \cos \lambda_c \end{bmatrix} \quad (7)$$

and \mathbf{S}_n is $n \times n$ with ones on the off-diagonal strip that lies adjacent to the main diagonal on the right hand side and zeros everywhere else; that is, the row i , column $i + 1$ element of \mathbf{S}_n equals 1 for $i = 1, \dots, n - 1$, and all

other elements equal 0. Define \mathbf{c}_n to be an $n \times 1$ vector with one in the last position and zeroes elsewhere. Then

$$\boldsymbol{\psi}_t = (\mathbf{T}_n \otimes \mathbf{I}_N)\boldsymbol{\psi}_{t-1} + \mathbf{c}_n \otimes \begin{bmatrix} \boldsymbol{\kappa}_t \\ \boldsymbol{\kappa}_t^* \end{bmatrix}, \quad (8)$$

where the assumptions on the $N \times 1$ vectors of Gaussian disturbance, $\boldsymbol{\kappa}_t$ and $\boldsymbol{\kappa}_t^*$, are

$$E(\boldsymbol{\kappa}_t) = \mathbf{0}, \quad E(\boldsymbol{\kappa}_t \boldsymbol{\kappa}_t') = E(\boldsymbol{\kappa}_t^* \boldsymbol{\kappa}_t^{*'}) = \boldsymbol{\Sigma}_\kappa, \quad E(\boldsymbol{\kappa}_s \boldsymbol{\kappa}_t') = \mathbf{0}, \quad \text{for } s \neq t$$

with $\boldsymbol{\Sigma}_\kappa$ an $N \times N$ covariance matrix and $E(\boldsymbol{\kappa}_s \boldsymbol{\kappa}_t^{*'}) = \mathbf{0}$ for all $s, t = 1, \dots, T$.

2.2 The changing output gap and turning points in the cycle

The trend in GDP is often regarded as a permanent component, while the cycle is transitory, but serially correlated; see, for example, Blanchard and Fischer (1989). The STM makes the concept operational in that the trend is the component that yields the long-run forecasts; see Harvey (1989, p 284-6). This avoids the ad hoc nature of a trend based on the Hodrick-Prescott or band pass filter. In what follows we will associate the trend with potential output and the cycle with the output gap.

Having fitted a model, we are interested in studying the characteristics of the cycle from the smoothed estimates. For example, although the expected value of the square of the amplitude is $E(\psi_t^2 + \psi_t^{*2}) = 2\sigma_\psi^2$, it may be of interest to plot estimates of the amplitude through time to see if it is changing in any way. Of course the full distribution of the output gap at any time may be of prime interest and we may wish to use this to produce statistics such as the probability that the economy is below potential output.

One characteristic of a cyclical series is its turning points. (For instance, the change from top to bottom in a downturn of the cycle gives a measure of the severity of the contraction, but to measure this we clearly need to know the top and bottom). Here turning points are identified from the extracted cycle rather than from the application of a criterion directly to the series itself. A fairly straightforward approach is to follow Zellner, Hong and Gulati (1990) in labeling a time point t as a peak, or downturn, if

$$\widehat{\psi}_{n,t-b}, \widehat{\psi}_{n,t-b+1}, \dots, \widehat{\psi}_{n,t-2}, \widehat{\psi}_{n,t-1} < \widehat{\psi}_{n,t} > \widehat{\psi}_{n,t+1}, \widehat{\psi}_{n,t+2}, \dots, \widehat{\psi}_{n,t+a} \quad (9)$$

where a and b are positive integers, and defining a trough, or upturn, in an analogous fashion. The smoother is the extracted cycle, $\widehat{\psi}_{n,t}$, the easier it should be to identify meaningful turning points.

Turning points may also be assessed directly from the fitted stochastic cycle as the elements in the state (5) contain information on the rate of change of the cycle. To see this it is necessary to move to a continuous time formulation. It is shown in appendix A that the discrete time analogue of the expected incremental change of a first-order cycle is given by

$$D\psi_t = (\log \rho)\psi_t + \lambda_c\psi_t^* \quad (10)$$

The second-order case is more complicated but it is argued in appendix A that for the discrete time model formulated in the previous sub-section an appropriate measure of change is

$$D\psi_{2,t} = (\log \rho)\psi_{2,t} + \lambda_c\psi_{2,t}^* + \rho^{-1}(\psi_{1,t} \cos \lambda_c - \psi_{1,t}^* \sin \lambda_c). \quad (11)$$

This formula can also be used for higher order cycles with the subscripts 2 and 1 replaced by n and $n - 1$ respectively. Plotting an estimate of $D\psi_{2,t}$ should be highly informative. So too should the series on the probability that $D\psi_{2,t}$ is negative (positive), this being the probability that the cycle is moving down (up). Turning points can be associated with points at which the estimate of $D\psi_{2,t}$ changes sign.

The NBER turning points are defined with respect to the level of the series. Harding and Pagan (2002) show that a rather simple dating rule applied to the differences in GDP reproduces the NBER peaks and troughs quite closely. A trough at time t is defined by $\{\Delta_2 y_t < 0, \Delta y_t < 0, \Delta y_{t+1} > 0, \Delta_2 y_{t+2} > 0\}$ where $\Delta_2 y_{t+2} = y_{t+2} - y_t = \Delta y_{t+2} + \Delta y_{t+1}$, and a peak similarly. Expansions and contractions are defined from these peaks and troughs and used as the basis for recognizing a cycle and measuring its characteristics. The same criterion could be applied to the extracted level and cycle components by looking at the change in estimates of $\mu_t + \psi_{n,t}$.

One might, of course, question the whole notion of the identification of changing points and the associated binary classification into expansions and recessions. The series of estimates of $\psi_{2,t}$ and $D\psi_{2,t}$ contain more information as they measure the size and direction of the output gap at all points in time. Even more information is contained in the full posterior distributions.

With a model-based approach, more information is obtained on a component at a particular point in time as more observations become available.

Thus information on the current cycle is obtained by filtering, but more accurate information is given by smoothing as new observations become available. On-line tracking of components is of considerable importance for policy makers and it is important to be aware of the uncertainty attached to any estimates; see the discussion in Orphanides and van Norden (2002).

3 Bayesian treatment

For convenience, we present the univariate case in this section and refer to the appendix for the multivariate case. The three variance parameters and two cyclical parameters are arranged in the vector $\boldsymbol{\theta} = \{\sigma_{\zeta}^2, \sigma_{\kappa}^2, \sigma_{\varepsilon}^2, \rho, \lambda_c\}$. The goal is to analyze the posterior distribution, $p(\boldsymbol{\theta}|\mathbf{y})$, where $\mathbf{y} = \{y_1, \dots, y_T\}$ denotes the observations. The joint posterior of the trend and cycle components is produced as a by-product of the Markov chain Monte Carlo routine.

3.1 Priors

The direct interpretation of the cycle parameters in the STM makes it straightforward to design priors that reflect knowledge of the business cycle. Thus for quarterly data we consider priors for frequency, λ_c , based on beta distributions with a mode at $2\pi/20$, corresponding to a period of five years. Figure 1 shows three such priors, labeled wide, intermediate and sharp. For technical details see appendix B.

The parameter ρ is linked to the order of the cycle. In the first order case ρ is the rate of decay of the cycle, but for higher orders the interpretation of ρ changes somewhat so that different values are appropriate. Although it seems, on the basis of empirical work, that ρ falls as n increases, it is difficult to be precise about the form of the relationship so we use a uniform prior on ρ over the interval $[0,1]$.

For the variance parameters independent flat priors on $[0, \infty]$ are assumed. The use of flat priors has the same advantage as inverted gammas of allowing one to sample directly, but in applications where expectations on the values of variance parameters are rather vague, a flat prior ensures that the likelihood surface is not distorted near zero³.

³The unobserved components models we consider in this paper have essentially the same fundamental structure as hierarchical models for studying group effects. As Gelman (2005) notes for the case of a basic hierarchical model, any noninformative prior on the

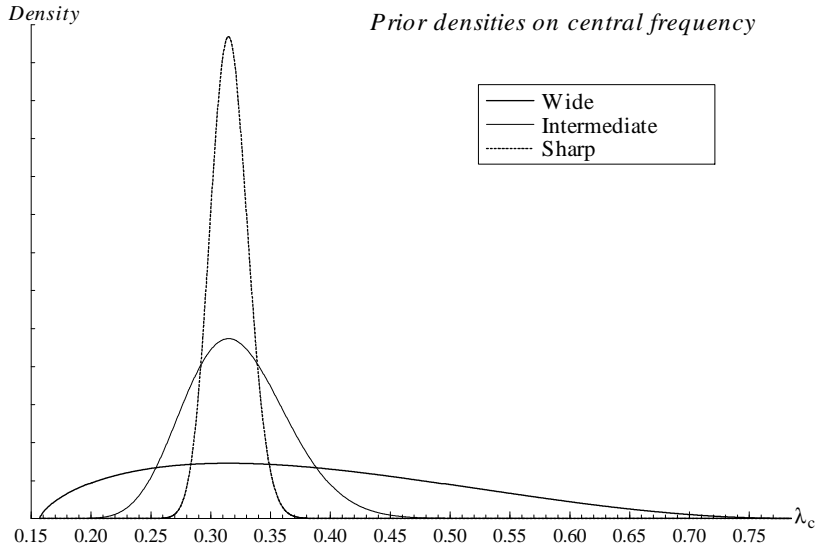


Figure 1: Beta-based priors on λ_c , with mode equal to $2\pi/20$ (five-year period for quarterly data).

3.2 Posterior

The posterior distribution is obtained as

$$p(\boldsymbol{\theta}|\mathbf{y}) = L(\boldsymbol{\theta}; \mathbf{y})p(\boldsymbol{\theta}),$$

where the likelihood function, $L(\boldsymbol{\theta}; \mathbf{y})$, is evaluated using the Kalman filter; see, for example, Harvey (1989, p. 126). The posterior is difficult to work with directly as the constant of proportionality is not available analytically. A strategy is therefore needed for analyzing its properties. In the applications that follow we deal with up to twenty-dimensional parameter vectors and MCMC methods offer an efficient way to sample (pseudo-random) parameter drawings from the posterior. This also allows us to produce drawings of regular functions of the parameters, such as periods of cycles and signal-noise ratios, and to compare finite sample results on posterior moments with ML estimates.

group level variance gives a proper posterior when sufficient data are available, as is the case here.

The parameter space is extended to include the components and associated auxiliary processes in (1), which together form the state vector, $\boldsymbol{\alpha}_t$, for the model, taken over all observation times $t = 1, \dots, T$. The state space form is discussed in appendix C. In this way, we design an MCMC routine that is able to capitalize on recent developments in state space modeling. Specifically, the simulation smoother for drawing from the conditional density of the state vector, as in de Jong and Shephard (1995), and, more recently, Durbin and Koopman (2002), may be embedded in a MCMC sampling setup.

Thus, the basic idea is to set up the algorithm to produce drawings from the expanded multivariate density $p(\boldsymbol{\alpha}, \boldsymbol{\theta} | \mathbf{y})$, where $\boldsymbol{\alpha}$ denotes the set of state vector elements over the entire sample period; the definition of $\boldsymbol{\alpha}$ is given in equations (17) and (18) in appendix C. This provides an efficient route for obtaining draws from $p(\boldsymbol{\theta} | \mathbf{y})$, and it also gives additional information that is useful for studying the trend and cycle.

3.3 Signal Extraction on Characteristics of Cyclical and Trend Components

The MCMC scheme produces draws from the joint posterior of the trend and cyclical components. The estimated component series are obtained by averaging over the J state draws, for example $\hat{\mu}_t = \sum_{j=1}^J \mu_t^{(j)} / J$, where $\mu_t^{(j)}$ denotes the j -th draw for the trend at time t . The standard deviation of the trend estimate at each time point is given by the square root of $\sum \mu_t^{2(j)} / J - \hat{\mu}_t^2$ and higher-order moments can be similarly constructed. The amplitude of the cycle at time t is estimated by

$$A_t = \frac{1}{J} \sum_{j=1}^J \sqrt{\hat{\psi}_{n,t}^{2(j)} + \hat{\psi}_{n,t}^{*2(j)}}, \quad t = 1, \dots, T \quad (12)$$

The probabilities that the cycle and its change, (11), are negative are easily estimated.

In keeping with state space terminology we will refer to posterior means computed over the whole sample as smoothed estimates while the corresponding estimates based only on current and past observations will be called filtered estimates.

3.4 Model evaluation

The marginal likelihood $m(\mathbf{y})$ is given as

$$m(\mathbf{y}) = \int L(\boldsymbol{\theta}; \mathbf{y})p(\boldsymbol{\theta})d\boldsymbol{\theta} \quad (13)$$

Bayes factors are computed as the ratio of marginal likelihoods for different model specifications $\{L(\boldsymbol{\theta}; \mathbf{y}), p(\boldsymbol{\theta})\}$. Posterior odds may be formed by multiplying Bayes factors by prior probability ratios that give the relative preference, *ex ante*, for the various likelihood-prior model structures. These form the basis for decision-making in a Bayesian framework; see Kass and Raftery (1995). The method we use for estimating marginal likelihoods is discussed in Appendix C.

4 US GDP

In this section we fit univariate models, with different orders for the cycle component, to the logarithms of quarterly real Gross Domestic Product from 1947Q1 to 2004Q4 (Source: Bureau of Economic Analysis, US Department of Commerce). The aim of the analysis is to present results on the mapping from prior to posterior of model parameters including a sensitivity analysis with respect to the choice of the prior on the frequency as shown in figure 1. The properties of extracted cycles are analysed by looking at turning points and changes in amplitude, as well as size and direction. Table 1 summarizes the results of fitting models in terms of posterior means and marginal likelihoods.

n	<i>Prior</i>	σ_ζ^2	σ_κ^2	σ_ε^2	ρ	λ_c	$2\pi/\lambda_c$	σ_ψ^2	q	$m(y)$
1	Wide	46.1	466	32	0.884	0.409	16.02	2336	0.0276	698.1
	Intermediate	28.3	539	23	0.894	0.341	18.66	2911	0.0117	698.4
	Sharp	24.5	561	21	0.898	0.320	19.69	3110	0.0092	698.4
2	Wide	17.1	363	111	0.697	0.272	24.62	4603	0.0058	704.0
	Intermediate	20.2	322	117	0.695	0.308	20.71	3854	0.0068	704.1
	Sharp	20.5	314	118	0.694	0.313	20.09	3706	0.0070	704.0
3	Wide	26.5	218	148	0.560	0.291	23.29	4097	0.0115	703.5
	Intermediate	26.9	199	150	0.562	0.311	20.54	3723	0.0112	703.7
	Sharp	26.6	196	151	0.563	0.314	20.05	3648	0.0108	703.8
4	Wide	43.0	159	157	0.461	0.310	22.13	3804	0.0287	702.3
	Intermediate	37.8	151	159	0.467	0.314	20.36	3759	0.0202	702.8
	Sharp	37.5	150	159	0.468	0.314	20.04	3728	0.0206	702.9

Table 1: Posterior means for logarithms of quarterly US real GDP from 1947:1 to 2004:4 for $n = 1$ to 4 with different priors on λ_c . $2\pi/\lambda_c$ is the period in quarters, and σ_ψ^2 is the variance of the cycle. The estimated logarithm of the marginal likelihood is denoted by $m(y)$. The signal-ratio $q = \sigma_\zeta^2/(\sigma_\psi^2 + \sigma_\varepsilon^2)$. All variance parameters are multiplied by 10^7 .

4.1 Priors and posteriors

Marginal posterior densities for $n = 1$, with the wide prior on λ_c , are shown in figures 2 and 3, along with 95% HPD (Highest Posterior Density) regions. Each estimated density function from the MCMC routine represents a standard approximation based on local Gaussian kernels, as implemented in the Ox programming language of Doornik (1999). The 95% HPD region is defined as the interval of minimum length that contains 95% of the probability mass for which the upper and lower boundary have equal density. The HPD regions resemble classical confidence intervals, but have a different interpretation in that they represent a direct probability statement about the value of an uncertain parameter or hidden component.

Figure 2 shows that the marginal posteriors of both the irregular variance, σ_ε^2 , and the slope variance, σ_ζ^2 , are skewed, but while the density of σ_ε^2 is concentrated near zero, that of σ_ζ^2 displays a clear peak away from zero, giving clear evidence for stochastic variation in the trend. As regards the cycle parameters, the marginal posterior for ρ , which is based on a uniform prior on $[0, 1]$, peaks near 0.9, while the density of the cyclical error variance σ_κ^2 appears symmetric. The prior and posterior densities for the frequency, λ_c , and period, $2\pi/\lambda_c$, appear in figure 3. The posteriors indicate a clear peak around a four- to five-year period even with a relatively noninformative prior. These results suggest that the likelihood surface in the first order model has a more or less regular shape so that it is relatively straightforward to pick out a business cycle component.

The underlying growth rate (slope of the trend) and the cycle, estimated as described in sub-section 3.3, are displayed in figures 4 and 5. The shaded regions in figure 5 denote recessions as identified by the NBER. The marginal posterior densities of the slope and cycle at 2003Q1 are shown in the lower panels of figure 3. The cycle will be analysed in the next sub-section. Figure 4 suggests the intriguing possibility of a cycle in the growth rate but this idea will not be pursued further in this paper.

For $n = 1$ a flat prior on λ_c gives nearly the same results as for the wide prior, since the likelihood surface with respect to λ_c has a clear peak. For higher order cycles, the results are more sensitive to the prior. With a flat prior the posterior of λ_c for $n = 2$ has most of its probability mass at very low values, with no clearly discernible peak; see figure 18 in appendix E. A corresponding problem arises in computing maximum likelihood estimates; Harvey and Trimbur (2003) had to resort to fixing λ_c in this case. Using an

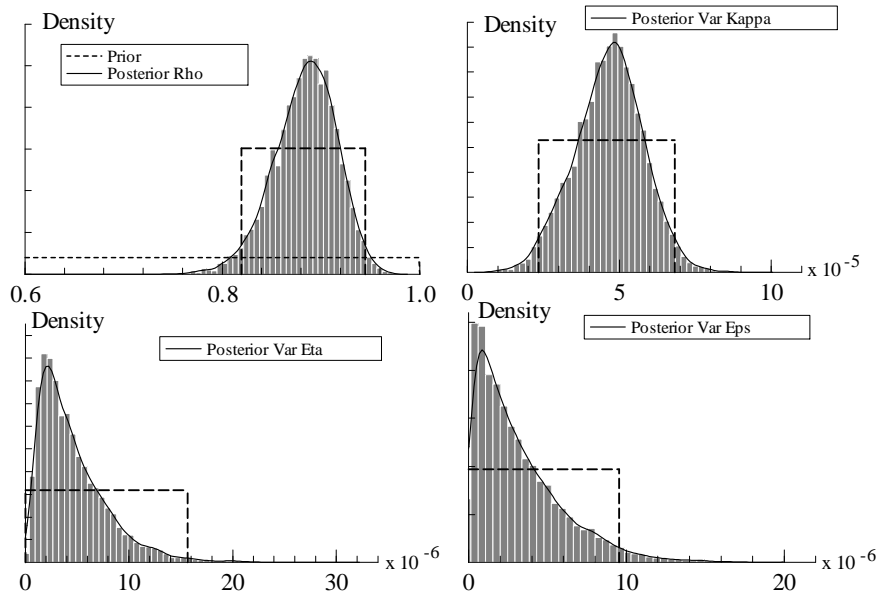


Figure 2: Marginal posterior densities of ρ , σ_κ^2 , σ_ζ^2 , and σ_ε^2 for $n = 1$, with least informative prior on λ_c , for quarterly US real GDP (logarithms) from 1947Q1 to 2004Q4. The dashed lines indicate the 95% Highest Posterior Density intervals; the rectangular regions shown on the graphs have a height that gives a total area of unity.

informative prior addresses the issue in a more flexible way. The results in table 1 indicate that the differences between the outcomes for wide, intermediate and sharp priors are not great and so it is probably safest to stick to the wide prior.

Figure 6 shows marginal posteriors of the cycle parameters for $n = 2$ with the wide prior. The posterior means are now around six years. The posteriors of σ_ε^2 and the cycle variance, σ_ψ^2 , indicate⁴ that more high frequency movement is consigned to the irregular. Figure 7 also shows the posterior distribution of the signal-noise ratio $q = \sigma_\zeta^2 / (\sigma_\psi^2 + \sigma_\varepsilon^2)$. This measures the relative variation in the nonstationary and stationary parts of the model. The density in figure 7 is skewed with a 95% HPD region that extends up to

⁴The posterior of σ_ψ^2 was obtained by computing the cycle variance for each set of draws $\{\rho^{(j)}, \lambda_c^{(j)}, \sigma_\kappa^2\}$.

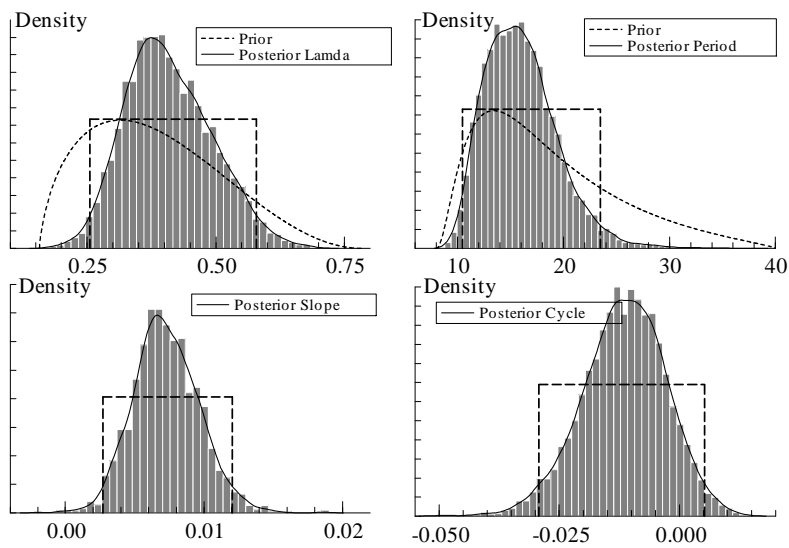


Figure 3: (Top) Marginal posterior densities of λ_c and $2\pi/\lambda_c$ for $n = 1$ with wide informative prior on λ_c . Results are based on quarterly data from 1947Q1 to 2004Q4, with sample size $T = 232$. (Bottom) Marginal posterior densities of the slope and cyclical components $\beta_{t,T}$ and $\psi_{t,T}$ at the observation time $t = 225$, which corresponds to 2003Q1.

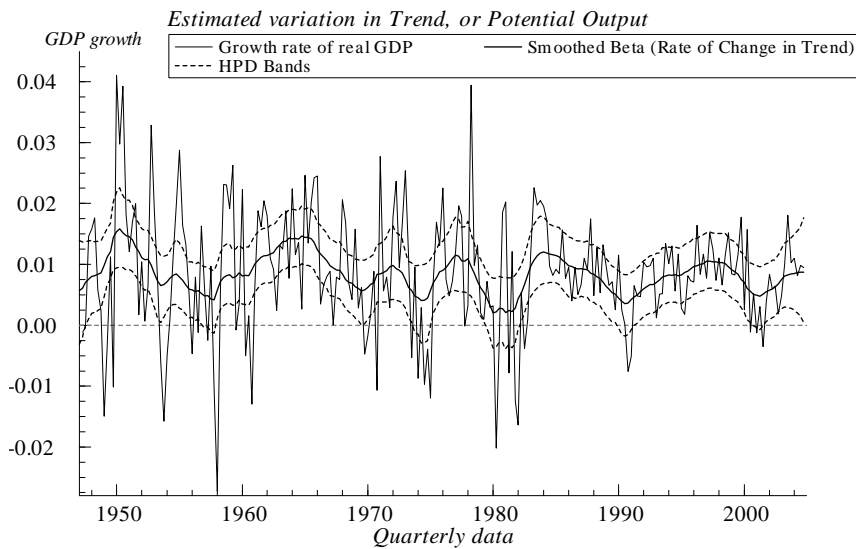


Figure 4: Smoothed slope (growth rate) in trend of US GDP, with 95% HPD bands, estimated with $n = 1$ and a wide prior on λ_c .

0.02.

Similar analysis could be carried out for $n = 3$ and 4. However, the main contrast is between $n = 2$ and the standard first-order case. Furthermore, the marginal likelihood shows the second-order cycle to be the preferred one⁵. The analysis in the next two sub-sections is based on the second-order cycle with a wide informative prior on the frequency.

4.2 Characteristics of the cycle

The relative smoothness of the second order cycle shown in figure 8 makes it easier to track movements in the business cycle and to identify peaks and troughs. We show the turning points evaluated from definition (9), with $b = 10$ and $a = 8$. Our turning points are related to many of the NBER

⁵There is tricky theoretical issue here because of the use of diffuse priors on certain parameters. However, it is generally accepted that comparisons are valid so long as the parameters in question are not restricted and occur in all models; see Gelman (2005). In the present context our preference for $n = 2$ is supported by the plausibility of posterior moments, the properties of the extracted cycles and forecasting performance.

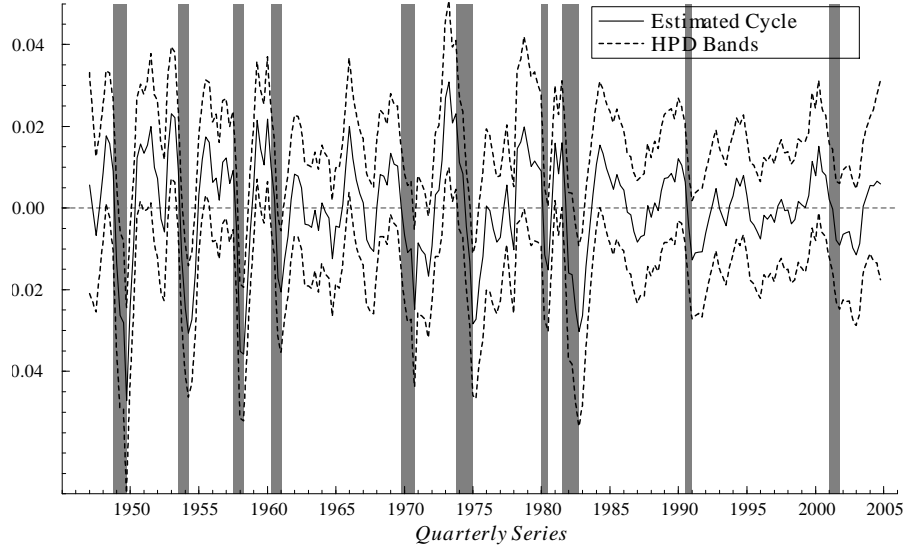


Figure 5: First-order cycle in quarterly US real GDP, with 95% HPD bands, estimated using a wide informative prior on λ_c .

expansions and recessions: the peaks occur a little before the start of an NBER recession while the troughs tend to match the end of the recessionary periods. However, as was observed in sub-section 2.2, focussing on the binary labelling of business cycle phases means that one loses sight of the finer detail offered by the extracted cycle. For example, the smoothed cycle in figure 8 shows clearly that some recessions are deeper than others. A plot of the changing amplitude of the cycle, using formula (12), allows one to see how the intensity of business cycles has changed over the last half-century. Figure 9 displays the evolving amplitude for the second order model. The graph shows that the strength of the cycle has been more moderate, as well as less volatile, since the mid 1980's. However, it also shows a slight upturn in the last few years, indicating that, contrary to the view of some economists, the business cycle is still relevant.

Figure 10, shows a plot of the posterior mean of $D\psi_{2,t}$, together with a corresponding estimate based on $(\Delta\psi_t + \Delta\psi_{t+1})/2$. This second estimate yields a series very close to $D\psi_{2,t}$, though the averaging makes it slightly less

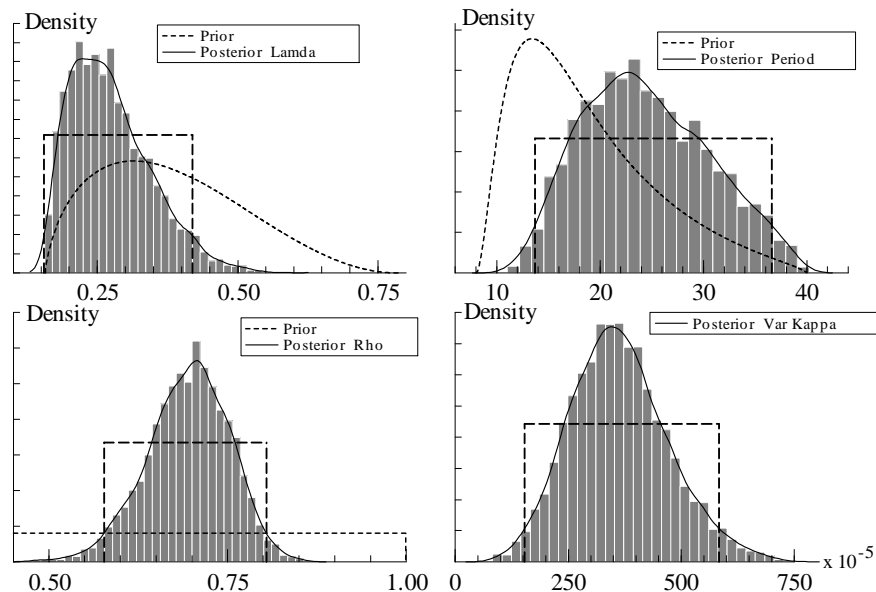


Figure 6: Marginal posterior densities of cycle parameters for $n = 2$, with wide informative prior on λ_c for quarterly US real GDP (logarithms) from 1947Q1 to 2004Q4.

volatile. The attraction of $D\psi_{2,t}$ is that it can be calculated at time t , rather than at $t + 1$, and this is important at the end of the sample.

The bottom-right graph of figure 7 shows the marginal posterior of the change in the cycle, $D\psi_{2,t}$, at observation point 1973Q4. As is clear from figure 8 this quarter marks the onset of the deep recession induced by the first oil price shock of the 1970's. Table 2 shows the probability that $D\psi_{2,t}$ is negative over the four year period surrounding the 1973-4 recession.

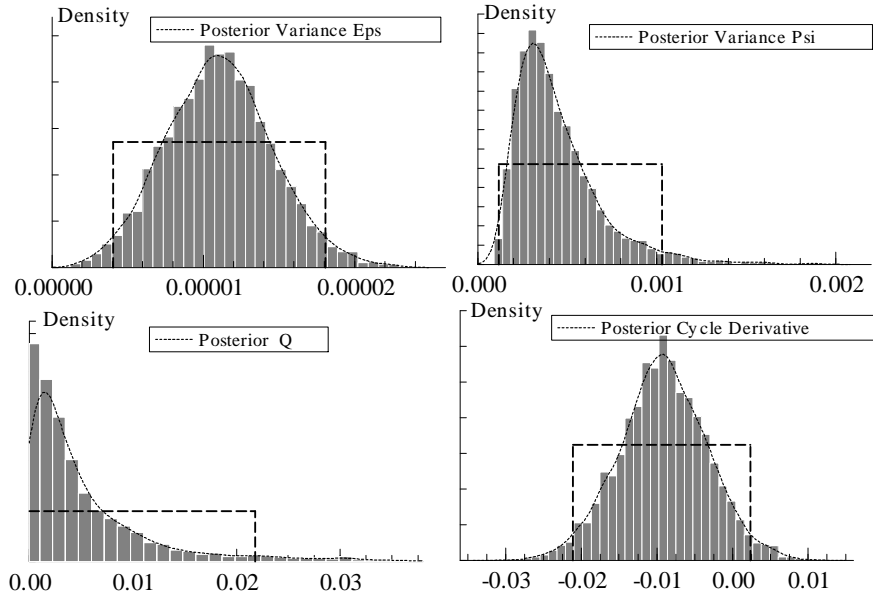


Figure 7: (Top) Marginal posterior densities of irregular and cycle variances for $n = 2$ with wide prior on λ_c . Data is quarterly US real GDP (logarithms) from 1947Q1 to 2004Q4. (Bottom, Left) Marginal posterior of signal-noise ratio $q = \sigma_\zeta^2 / (\sigma_\psi^2 + \sigma_\varepsilon^2)$. (Bottom, Right) Marginal posterior of $D\psi_{2,t}$ at observation time $t = 108$, which corresponds to 1973Q4.

Date	$Pr(D\psi_{2,t} < 0)$	Date	$Pr(D\psi_{2,t} < 0)$
1972Q1	0.0124	1974Q1	0.8536
1972Q2	0.1689	1974Q2	0.9924
1972Q3	0.0252	1974Q3	0.9954
1972Q4	0.0002	1974Q4	1.000
1973Q1	0.0438	1975Q1	0.8734
1973Q2	0.8114	1975Q2	0.1330
1973Q3	0.5200	1975Q3	0.0796
1973Q4	0.9426	1975Q4	0.0068

Table 2 : Probability that the rate of change in the cycle is negative for the period 1972Q1 to 1975Q4.

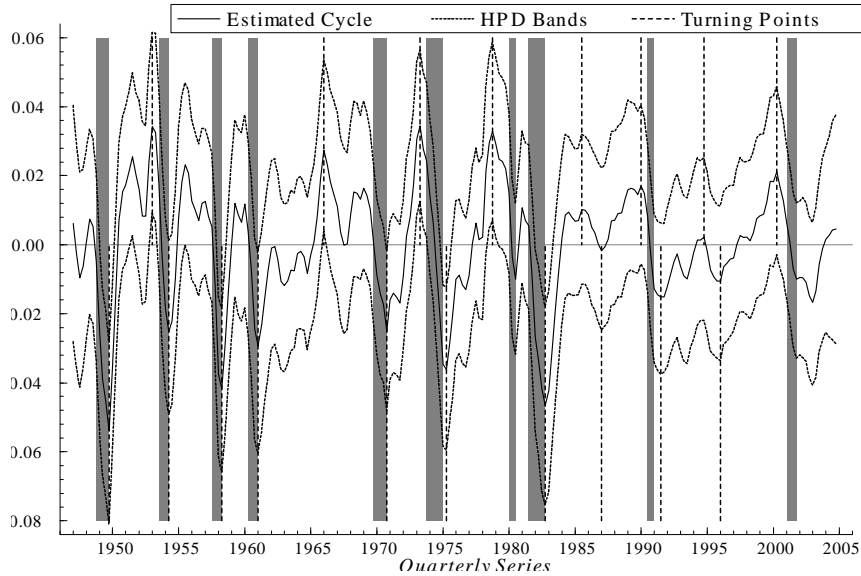


Figure 8: Second-order cycle in quarterly US real GDP, with 95% HPD bands, estimated using a wide informative prior on λ_c .

4.3 On line tracking of the output gap

While smoothed estimates for the whole series provide a historical perspective, what is most important for policy makers is real time estimates the current state of the economy. Figure 11 tracks the filtered cycle from the first quarter of 1999 to the end of 2004. An associated series of the probability of the cycle being negative could also be produced.

Figure 12 shows the estimated on-line change in the cycle based on filtered estimates of $D\psi_{2,t}$, while figure 13 shows the estimated probability that $D\psi_{2,t}$ is negative. The most interesting feature of these graphs is that they show how $D\psi_{2,t}$ anticipates changes in direction of the cycle. Thus in 2002Q1 the cycle is still moving down but $D\psi_{2,t}$ is positive: in the next period the output gap has narrowed. Similar behaviour can be seen in 2002Q3 and 2003Q2.

Figure 14 shows the filtered marginal posterior for the cycle at 2001Q4 together with some of the subsequent (smoothed) posteriors. The idea is to show how the uncertainty is reduced with the arrival of new data. The posterior is already much tighter after two periods. One year later, there is a sizeable gain in precision and by the end of 2003 the distribution is relatively

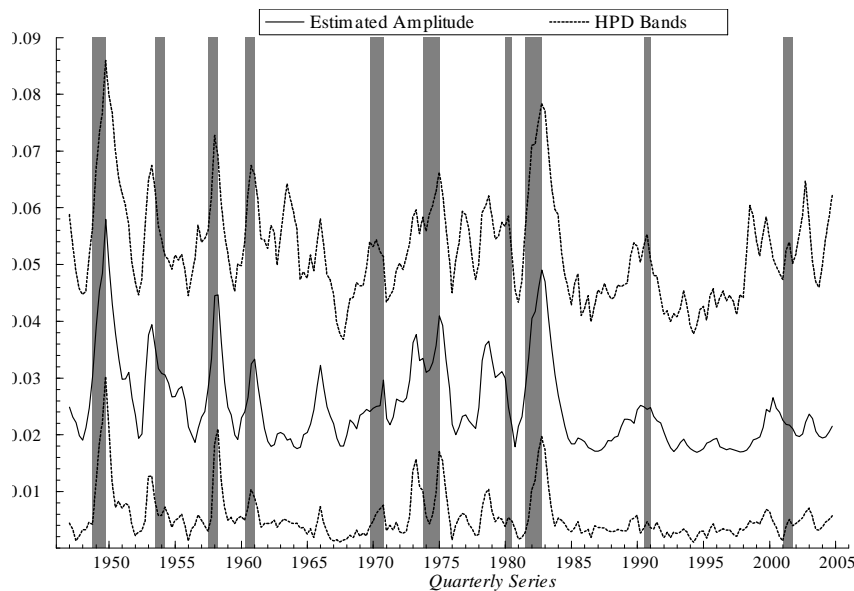


Figure 9: Estimated amplitude of the cycle in quarterly US real GDP for $n = 2$ and a wide informative prior on λ_c , shown with 95% HPD bands.

compact and symmetric with a peak close to -0.01 .

5 Multivariate models

In this section we investigate whether modeling GDP jointly with other US macroeconomic series can help to give a more accurate picture of the business cycle. The first sub-section fits a model of the form (1) to consumption, investment and GDP. Although most of the benefit to business cycle analysis is likely to come from the inclusion of investment, it is interesting to fit the model to the three series, firstly to make a comparison with the well-known study of King, Plosser, Stock and Watson (1991) and secondly to see whether there are technical difficulties in applying our method in the trivariate case. The second sub-section reports an extension of the modeling framework in which lagged cyclical components are included in an inflation equation.

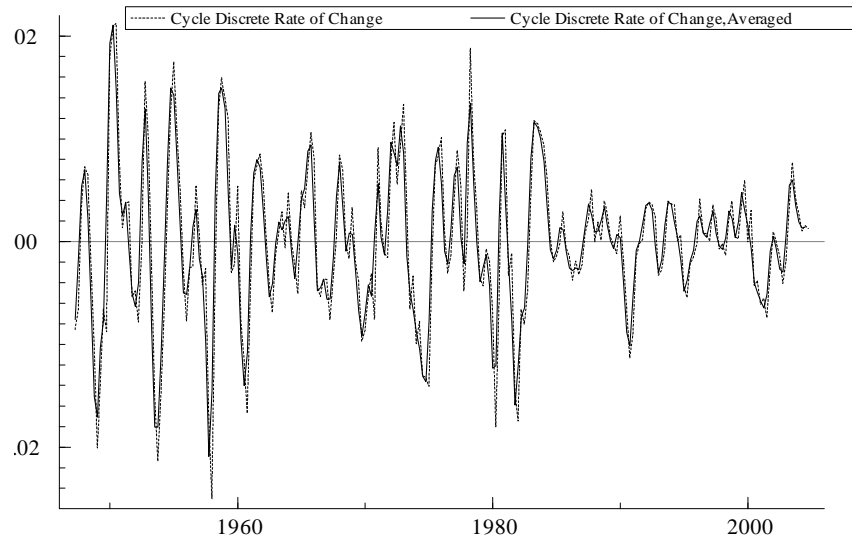


Figure 10: $D\psi_{2,t}$ and average of adjacent differences

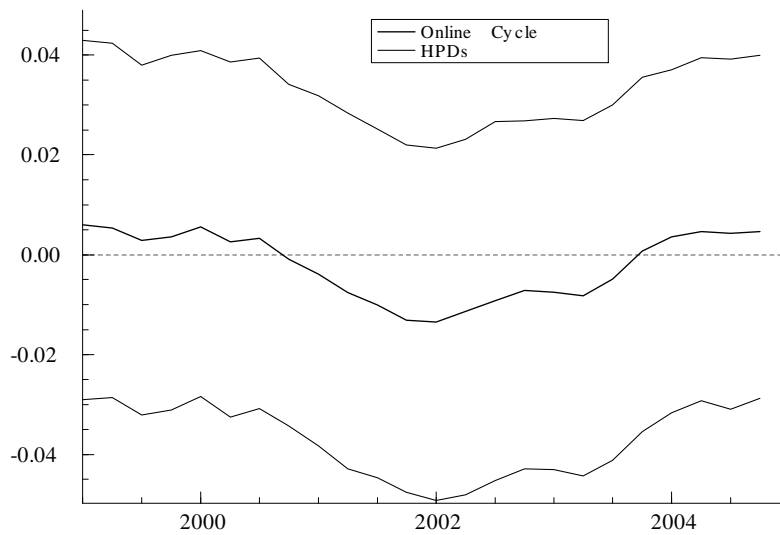


Figure 11: Filtered cycle from 1999Q1 with 95% HPD bands

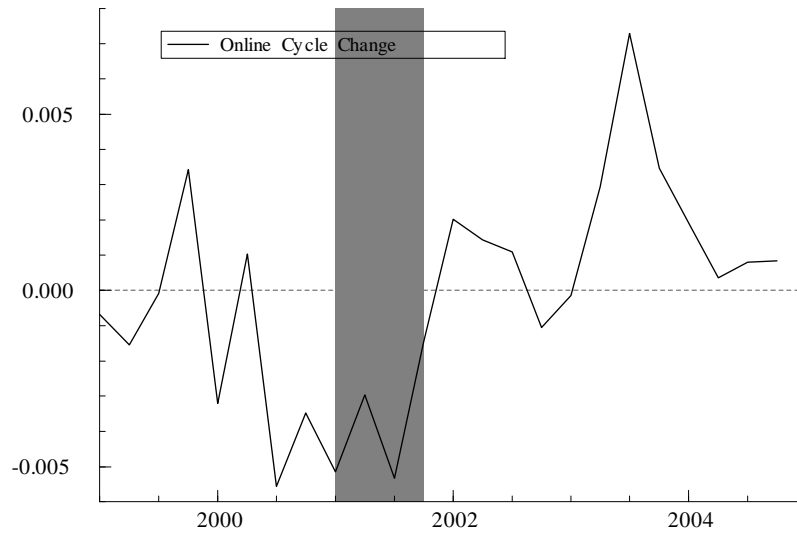


Figure 12: Estimated change in the filtered cycle

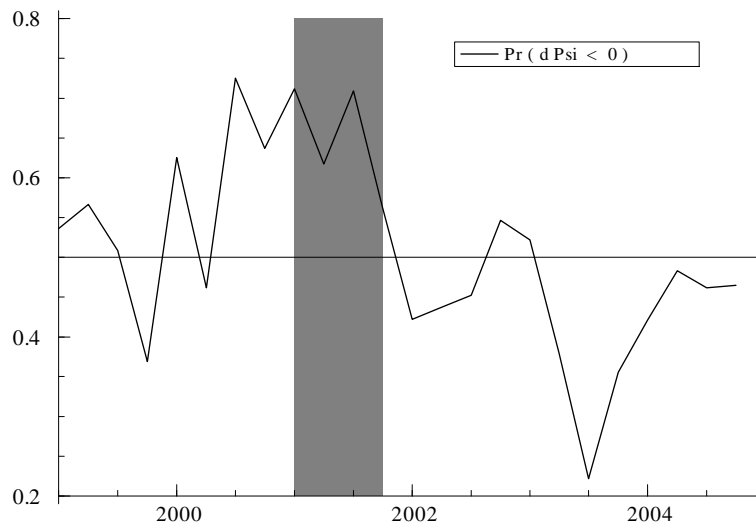


Figure 13: Probability that change in the filtered cycle, $D\psi_{2,t}$, is negative

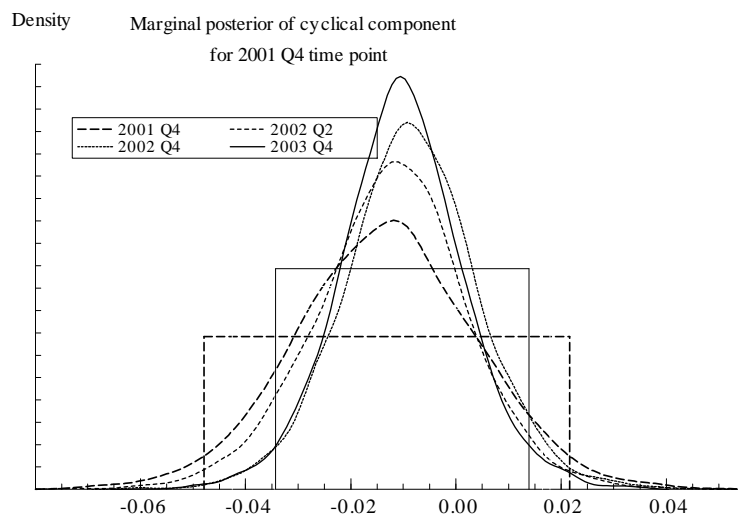


Figure 14: Marginal posterior densities of $\psi_{2,t}$ (with a wide informative prior on λ_c) for 2001Q4 using data up to 2001Q4 and then using more observations. HPD regions are shown for the filtered estimate (at 2001Q4) and the smoothed estimate at the end of the series (2003Q4).

5.1 C-I-Y for the US economy

The data consist of quarterly series, expressed as real quantities with base year 2000, on Personal Consumption Expenditures, Gross Private Domestic Investment, and GDP from 1947Q1 to 2004Q4 (Source: Bureau of Economic Analysis). The similar cycle assumption of a shared cyclical period is perfectly sensible, and it may lead to efficiency gains. The MCMC routine for the Bayesian treatment of the general model (1) is closely related to the univariate algorithm, and the details of the multivariate extension are shown in the appendix.

To demonstrate the effectiveness of pooling the data, we first assumed a flat prior on λ_c . (The uniform prior on $[0,1]$ is again used for ρ). To avoid potential distortions of the multivariate likelihood surface due to the influence of the inverted Wishart prior, we used flat priors on the variance matrices. The posterior density for λ_c with $n = 2$, shown in figure 19 in appendix E, reached a clear peak where the implied period is about twenty-one quarters.

If we use an informative prior for $n = 2$ there seems little point in going beyond the wide prior given the results of the previous paragraph - in any case the wide prior seemed to be adequate even in the univariate case. Table 3 shows posterior means and figure 20, in the appendix, displays estimated marginal posteriors for a number of variance parameters. Random drawings for the correlations, denoted v , between the components across different series are directly constructed from the draws for the variance matrices.

n	Series: Consumption			Series: Investment			Series: Output		
	σ_ζ^2	σ_κ^2	σ_ε^2	σ_ζ^2	σ_κ^2	σ_ε^2	σ_ζ^2	σ_κ^2	σ_ε^2
1	70.1	96.3	167	1640	7890	4430	106	279	82.1
2	40.5	99.6	193	338	7170	6160	51.8	260	137
3	39.4	64	208	296	4650	7221	66.6	168	169
4	40.9	33	218	300	1999	8220	86.5	75	191

Correlations and shared cyclical parameters

n	ρ	λ_c	$2\pi/\lambda_c$	v_ζ	v_ε	v_κ	$m(y)$
1	0.886	0.545	11.6	0.927	-0.679	0.654	-1842.9
2	0.705	0.334	20.3	0.732	-0.493	0.824	-1724.9
3	0.603	0.266	34.6	0.709	-0.388	0.851	-1680.0
4	0.554	0.281	31.5	0.620	-0.345	0.895	-1663.1

Table 3: Posterior means for a trivariate model for quarterly US real GDP, consumption, and investment from 1947:1 to 2004:4 using flat priors on the variance matrices and a wide prior on λ_c . All variance parameters are multiplied by 10^7 . The parameters ρ and λ_c are common to all series and v denotes correlation.

The estimated second order cycle in GDP - shown in figure 21 in appendix E - is similar to the one produced by the univariate case. However, the HPD bands are approximately 25% smaller than the univariate band during the 1990s. The filtered estimates corresponding to those reported at the end of the previous section will likewise be more accurate.

Instead of looking at recent filtered estimates, as in the previous section, we examine forecasting performance. Figure 15 shows the multi-step predictions⁶ for the cycle made at the end of 2002 for the subsequent two-year period. The forecast function indicates an upturn in the cycle and this is consistent with the smoothed estimates made at the end of 2004 and shown in figure 21 in the appendix (the corresponding univariate estimates shown in figure 8 are similar).

⁶The subsequent observations are not being used to construct a series of one-step ahead predictions.

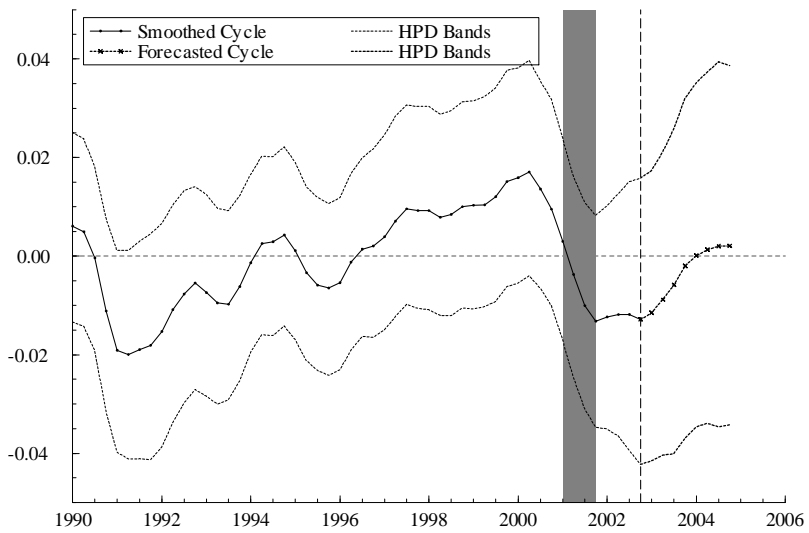


Figure 15: Forecast of cyclical component, with 95% HPD bands, of quarterly US real GDP (logarithms) for the two-year period 2003Q1 to 2004Q4, based on the trivariate C-I-Y model with $n = 2$ using a wide prior on λ_c . Also shown from 1990Q1 to 2002Q4, is the smoothed cycle.

5.2 Inflation and the output gap

Kuttner (1994) and Planas and Rossi (2004) argue that the link between inflation and the output gap, as reflected in the Phillips curve, may be exploited to produce more reliable estimates of the output gap. Thus an unobserved components model for GDP is combined with an equation in which inflation depends on lagged values of the output gap, as measured by the cycle in GDP.

The model we fit is

$$\begin{bmatrix} \pi_t \\ y_t \end{bmatrix} = \begin{bmatrix} \mu_t^\pi \\ \mu_t^y \end{bmatrix} + \begin{bmatrix} \psi_t^\pi \\ \psi_t^y \end{bmatrix} + \begin{bmatrix} p_t \\ 0 \end{bmatrix} + \begin{bmatrix} \varepsilon_t^\pi \\ \varepsilon_t^y \end{bmatrix}$$

where y_t is the logarithm of quarterly real U.S. GDP, π_t is the CPI rate of inflation and

$$p_t = c_1\psi_{t-1}^y + c_2\psi_{t-2}^y$$

denotes the price pressure due to recent output gap levels. Note that a simple transformation of the similar cycle model allows the cycle in inflation to be broken down into two independent parts, one of which depends on the GDP cycle, that is $\psi_t^\pi = c\psi_t^y + \psi_t^+$. Putting the model into state space form is straightforward; see appendix D. Posterior results for cycle orders from 1 to 4 are shown in table 4.

The above inflation equation is somewhat different from the ones in Kuttner (1994) and Planas and Rossi (2004) in that we drop the lagged growth rate of GDP and include a stochastic trend. The reason for including the stochastic trend is that it is difficult to find a stable relationship between inflation and output without it. Figure 16 shows the trend and cycles obtained by a model of the form (1) with $n = 2$. Plotting the cycles on the same graph, figure 17, provides an indication that GDP leads inflation, particularly in the 1970s.

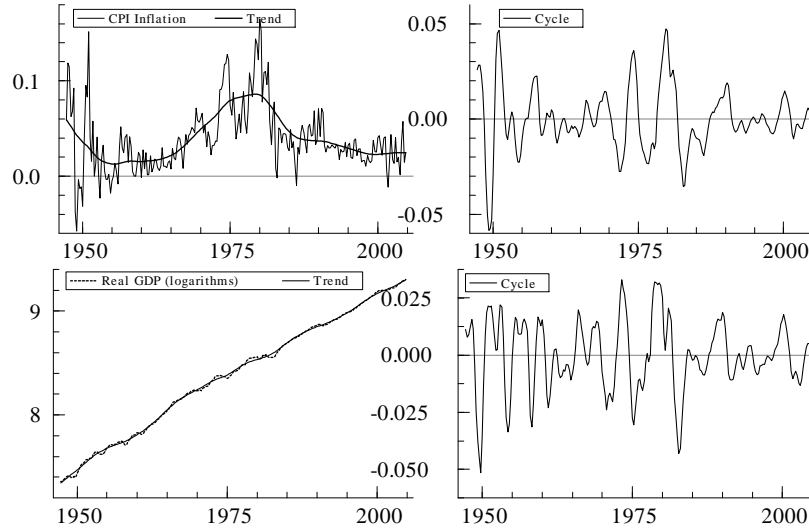


Figure 16: Estimated trends and cycles in quarterly CPI inflation and US real GDP (logarithms) for $n = 2$ with wide prior on λ_c .

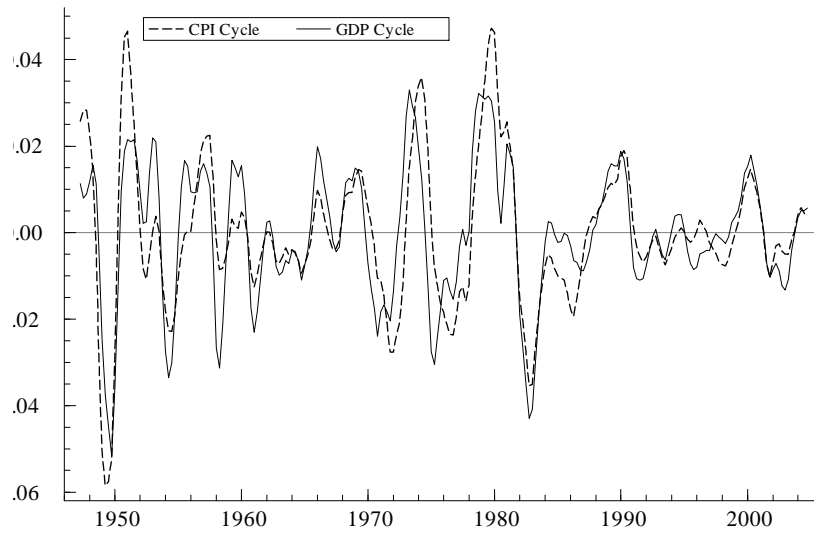


Figure 17: Estimated second-order cycles in quarterly CPI inflation and US real GDP.

n	Series: Inflation			Series: GDP		
	σ_ζ^2	σ_κ^2	σ_ε^2	σ_ζ^2	σ_κ^2	σ_ε^2
1	5.97	337	5241	44.4	497	39.6
2	3.77	112	5875	25.8	272	140
3	4.76	39.3	6315	62.6	131	173
4	4.25	19.3	6594	96.1	82	180

Correlations, shared cyclical parameters, and inflation response coefficients

n	ρ	λ_c	$2\pi/\lambda_c$	v_ζ	v_ε	v_κ	c_1	c_2	$m(y)$
1	0.902	0.349	18.25	0.045	-0.008	0.016	0.710	0.061	-1136.0
2	0.725	0.295	22.36	-0.221	-0.028	0.040	0.723	0.018	-974.1
3	0.596	0.302	22.09	-0.081	-0.029	0.168	0.876	0.098	-948.2
4	0.501	0.313	21.54	0.105	-0.023	0.243	1.01	0.155	-944.5

Table 4: Posterior means for a bivariate model of quarterly US CPI inflation and real GDP from 1947:2 to 2004:4 using flat priors on the variance matrices and a wide prior on the central frequency λ_c . All variance parameters are multiplied by 10^7 . The parameter v_ζ denotes the correlation between the slope disturbances, and ρ and λ_c are the shared cyclical parameters. The parameters c_1 and c_2 represent the response of inflation to the GDP cycle at lags of one and two quarters, respectively.

In moving from first to second order cycles, the marginal likelihood rises markedly, and the coefficients that relate the response of inflation to lagged output gap increase. The posteriors for c_1 and c_2 - shown in appendix E - appear more or less symmetric with the 95% HPD bounds being [0.144, 1.380] and [-0.486, 0.470] respectively. Furthermore the correlation between the cycle disturbances is close to zero, implying a small coefficient on the current cycle if the model is reparameterized in the way indicated earlier⁷. The conclusion is that a one percent increase in the GDP cycle is expected to foreshadow a rise of about 0.7 of a percentage point in the rate of inflation in the next quarter. Including a second lag, as is done by Planas and Rossi (2004) but not by Kuttner (1994), is probably not necessary.

The GDP cycle - shown in figure 23 in appendix E - resembles corresponding series from the univariate model, but the HPD bands are narrower. Of course, it would be even more useful if one had an equation that featured leads in the cycle rather than lags. Nevertheless given that it takes two or three periods to recognize a turning point, a one period lag is not entirely useless.⁸

6 Conclusion

This article further extends the model-based methodology for the estimation of trends and cycles in macroeconomic time series. The preferred models use the second order cycles introduced recently by Harvey and Trimbur (2003) since these tend to be smoother than first-order cycles, with relatively more noise consigned to the irregular component. We suggest various ways in which the information obtained by fitting such models can be used to describe past movements of the cycle and to focus attention on such features such as changing volatility and turning points.

There are two main attractions to a Bayesian approach. Firstly, flexible restrictions can be placed on key parameters, such as the frequency parameter in the stochastic cycle, and this avoids fitting implausible models. Secondly, parameter uncertainty is taken account of in providing information about extracted components and their forecasts. The disadvantage is that the computational requirements are heavier than for maximum likelihood.

⁷ $c = 0.040\sqrt{112/272} = 0.027$

⁸Of course the model might also be useful for forecasting inflation.

However, although Bayesian estimation typically requires several minutes⁹, as opposed to a few seconds, this hardly renders it infeasible. The Markov chain Monte Carlo routines we describe should be of value for future research, particularly for multivariate series where the computations are non-trivial.

Cycles were successfully extracted from US GDP using a univariate model. Such cycles have a simple interpretation in terms of the percentage by which they exceed or fall below the long-term level of potential output. Associated measures track the size and direction of the cycle. The posterior distributions clearly indicate the degree of uncertainty that arises from signal extraction and the fact that the parameters are unknown. Fitting multivariate models offers the possibility of reducing this uncertainty. Used on-line, measures such as the probability that the output gap is increasing or decreasing may be of considerable practical value.

ACKNOWLEDGMENTS

Earlier versions of this paper were presented at the Cambridge PhD workshop, the NAKE Conference, ESAM2004 and JRC in Italy. Trimbur wishes to thank the Cambridge Commonwealth Trust and the Richard Kahn Fund for financial support, and the Tinbergen Institute for its hospitality and financial support in autumn 2001. He also grateful for support from an ASA/NSF Research Fellowship/Grant at the U.S. Census Bureau. Harvey thanks the Economic and Social Research Council (ESRC) for support as part of a project on Dynamic Common Factor Models for Regional Time Series, grant number L138 25 1008. Van Dijk thanks the Erasmus Research Institute for Management for financial assistance. We are grateful to Bill Bell, Simon Godsill, Neil Shephard and Arnold Zellner for helpful comments.

References

- Baxter, M. and King, R. G., 1999. Measuring business cycles: approximate band-pass filters for economic time series, *Review of Economics and Statistics* 81, 575-93.
- Blanchard, O.J. and Fischer, S. 1989, *Lectures in Macroeconomics*, (Cambridge, MA: MIT Press)

⁹Univariate : 5000 iterations takes about 40 seconds for $n = 2$ and about 80 seconds for $n = 4$. Output gap / Bivariate : 5000 iterations takes 4 minutes for $n = 2$. Trivariate C-I-Y : 5000 iterations takes about 5 minutes, 45 seconds for $n = 2$. Recall that when we take 5000 draws with 10 iterations per draw and 10,000 burned there are 60,000 iterations. Microprocessor of computer is Pentium 4 (2.26 GHz).

- Carter, C. K. and Kohn R., 1994. On Gibbs sampling for state space models, *Biometrika* 81, 541-53.
- Chib, S. and Greenberg, E., 1996. Markov chain Monte Carlo simulation methods in econometrics, *Econometric Theory* 12, 409-31.
- Cogley, T. and Nason, J.M., 1995. Effects of the Hodrick-Prescott filter on trend and difference stationary time series: implications for business cycle research, *Journal of Economic Dynamics and Control* 19, 253-78.
- de Jong, P. and Shephard, N., 1995. The simulation smoother for time series models, *Biometrika* 82, 339-50.
- Doornik, J. A., 1999. *Ox: An Object-Oriented Matrix Programming Language* (London: Timberlake Consultants Ltd.).
- Durbin, J., and Koopman, S.J., 2002. A simple and efficient simulation smoother, *Biometrika* 89, 603-16.
- Frühwirth-Schnatter, S., 1994. Data augmentation and dynamic linear models, *Journal of Time Series Analysis*, 15, 183-202.
- Gelman, A., 2005. Prior distributions for variance parameters in hierarchical models, *Journal of Bayesian Analysis*, forthcoming.
- Harding, D. and Pagan, A. 2002. Dissecting the cycle: a methodological investigation. *Journal of Monetary Economics*, 49, 365-81.
- Harvey A.C., 1989. *Forecasting, structural time series models and the Kalman filter* (Cambridge: Cambridge University Press).
- Harvey A.C., 2001. Testing in unobserved components models, *Journal of Forecasting*, 20, 1-19.
- Harvey, A. C., and Jaeger, A., 1993. Detrending, stylised facts and the business cycle, *Journal of Applied Econometrics* 8, 231-47.
- Harvey, A. C., and Koopman, S.J., 1997. Multivariate structural time series models. In C. Heij et al.(eds), *System dynamics in economic and financial models* (Chichester: Wiley and Sons).

- Harvey, A. C., and Trimbur, T.M., 2003. General model-based filters for extracting trends and cycles in economic time series, *Review of Economics and Statistics*, 85, 244-55.
- Huerta, G and West, M, 1999. Priors and component structures in autoregressive time series models. *Journal of the Royal Statistical Society, Series B*, 61, 881-99.
- Hodrick, R. J. and Prescott, E.C., 1997. Postwar US business cycles: an empirical investigation, *Journal of Money, Credit and Banking* 24, 1-16.
- Kass, R. E. and Raftery, A.E., 1995. Bayes factors, *Journal of the American Statistical Association* 24, 773-95.
- King, R., Plosser, C., Stock, J., and Watson, M., 1991. Stochastic trends and economic fluctuations, *American Economic Review* 81, 819-40.
- Kitagawa, G., and Gersch, W., 1996. *Smoothness priors analysis of time series* (Berlin: Springer-Verlag).
- Kleijn R., and van Dijk, H.K., 2005. Bayes model averaging of cyclical components within economic time series, *Journal of Applied Econometrics*, forthcoming.
- Kohn, R., C. Ansley, and Wong, C.H., 1992. Nonparametric spline regression with autoregressive moving average errors, *Biometrika* 79, 335-46.
- Koop, G. and van Dijk, H.K., 2000. Testing for integration using evolving trend and seasonals models: a Bayesian approach, *Journal of Econometrics* 97, 261-91.
- Koopman, S. J., Shephard N., and Doornik, J., 1999. Statistical algorithms for models in state space using SsfPack 2.2, *Econometrics Journal* 2, 113-66.
- Koopman, S. J., Harvey A.C., Doornik J.A., and Shephard, N., 2000. STAMP 6.0 Structural Time Series Analysis Modeller and Predictor (London: Timberlake Consultants Ltd.).
- Kuttner, K., 1994. Estimating potential output as a latent variable, *Journal of Business and Economic Statistics* 12, 361-8.

- Magnus, J. and Neudecker, H., 1999. *Matrix Differential Calculus with Applications in Statistics and Econometrics* (Chichester: John Wiley & Sons).
- Murray, C. J., 2003. Cyclical properties of Baxter-King filtered time series, *Review of Economics and Statistics*, 85, 471-6.
- Planas, C. and Rossi, A., 2004. Can inflation data improve the real-time reliability of output gap estimates ? *Journal of Applied Econometrics*, 19, 121-33.
- Poirier, D., 1995. *Intermediate Statistics and Econometrics* (Cambridge, MA: MIT Press).
- Orphanides, A., and van Norden, S., 2002. The unreliability of output gap estimates in real-time, *Review of Economics and Statistics*, 84, 4, 569-83.
- Stoffer, D. and Wall, K., (2004). Resampling in State Space Models, in *State Space and Unobserved Component Models*, ed Harvey, A.C, Koopman, S.J and N. Shephard, 171-202. (Cambridge: Cambridge University Press).
- Trimbur, T. M., 2005. Properties of higher order stochastic cycles, *Journal of Time Series Analysis*, forthcoming.
- Zellner, A., 1971. *An introduction to Bayesian inference in econometrics* (New York: John Wiley & Sons).
- Zellner, A., Hong, C., and Gulati, G., 1990. Turning points in economic time series, loss structures, and Bayesian forecasting. In S. Geisser et al.(eds), *Bayesian and Likelihood Methods in Statistics and Econometrics* (North-Holland: Elsevier Science Publishers).

A Rate of change in the cycle

A first-order cycle in continuous time is written

$$d\boldsymbol{\psi}(t) = \mathbf{A}\boldsymbol{\psi}(t)dt + d\mathbf{W}_{\kappa}(t)$$

where $\boldsymbol{\psi}(t) = \{\psi(t) \ \psi^*(t)\}$, $\mathbf{W}_\kappa(t)$ is a 2×1 vector of Brownian motion and \mathbf{A} is the matrix

$$A = \begin{bmatrix} \ln \rho & \lambda_c \\ -\lambda_c & \ln \rho \end{bmatrix}$$

The expected incremental change in $\psi(t)$ is

$$d\psi(t) = (\log \rho)\psi(t)dt + \lambda_c\psi^*(t)dt$$

and the discrete time model is as in (3); see Harvey (1989, p 487). The discrete time expression corresponding to $d\psi(t)$ is therefore as in (10). There is a slight complication with flow variables in that the discrete time model has correlated measurement and transition equation noise; see Harvey (1989, p 494). However, we would still suggest using (10) as an approximation.

The second-order cycle is

$$\begin{aligned} d\boldsymbol{\psi}_2(t) &= \mathbf{A}\boldsymbol{\psi}_2(t)dt + \boldsymbol{\psi}_1(t)dt \\ d\boldsymbol{\psi}_1(t) &= \mathbf{A}\boldsymbol{\psi}_1(t)dt + d\mathbf{W}_\kappa(t) \end{aligned}$$

with expected incremental change

$$d\boldsymbol{\psi}_2(t) = (\log \rho)\boldsymbol{\psi}_2(t)dt + \lambda_c\boldsymbol{\psi}_2^*(t)dt + \boldsymbol{\psi}_1(t)dt \quad (14)$$

The model in in (4), that is

$$\boldsymbol{\psi}_{2,t} = \mathbf{T}\boldsymbol{\psi}_{2,t-1} + \boldsymbol{\psi}_{1,t-1}, \quad (15)$$

$$\boldsymbol{\psi}_{1,t} = \mathbf{T}\boldsymbol{\psi}_{1,t-1} + \boldsymbol{\kappa}_t, \quad \boldsymbol{\kappa}_t \sim NID(\mathbf{0}, \sigma_\kappa^2 \mathbf{I}) \quad (16)$$

with \mathbf{T} as in (7) and $\boldsymbol{\psi}_{1,t} = \mathbf{T}\boldsymbol{\psi}_{1,t}^\dagger$, where $\boldsymbol{\psi}_{1,t}^\dagger$ corresponds to the continuous time variable $\boldsymbol{\psi}_1(t)$, is an approximation to the discrete time formulation as there will be a disturbance, correlated with $\boldsymbol{\kappa}_t$, attached to the first equation and, for a flow, both disturbances will be correlated with the measurement equation disturbance. If the approximation to the continuous time model is accepted, the appropriate formula for change is as in (11) because of the transformation from $\boldsymbol{\psi}_{1,t}^\dagger$ to $\boldsymbol{\psi}_{1,t}$.

B Priors

The univariate case

The priors on λ_c derive from the beta family of densities and are based on expectations of period lying in the business cycle range. For details on the beta family we refer to Poirier (1995). The priors on λ_c are all constructed to have a mode of $2\pi/20$ and to lie between $\pi/20$ (10 year period) and $\pi/4$ (2 year period). In this way, the prior mass peaks around a period of five years and implausible average periods are preempted. If one were to use a constant mean of $\pi/10$ as a basis, then for large spreads, the skewness of the density means that the prior would peak at lower frequencies. By concentrating on the mode, one ensures the prior probability density is at a maximum in a region representative of the business cycle period. The three priors in figure (1) cover a range of spreads to account for different degrees of informativeness. For the widest prior the standard deviation is $\sigma_\lambda = 2\pi/50$ and the proportional spread is $\sigma_\lambda/\hat{\mu}_\lambda = 40\%$. For the intermediate prior $\sigma_\lambda = 2\pi/150$, $\sigma_\lambda/\hat{\mu}_\lambda = 13\%$, and for the sharpest prior $2\pi/400$, $\sigma_\lambda/\hat{\mu}_\lambda = 5\%$. The priors are designed by setting the beta density parameters and interval bounds to give the desired characteristics.

The flat priors on the variance parameters $\sigma_\zeta^2, \sigma_\kappa^2, \sigma_\varepsilon^2$ ensure that the posterior is unaffected by the prior shape around zero. As shown below, the variance parameters have independent inverted gamma conditional posteriors when independent flat priors are used, so direct simulation is possible, which enhances the efficiency of the algorithm. We will give the expressions for the scale and shape of each conditional posterior density in the following appendix. For details on the inverted gamma family we refer to Poirier (1995).

Another possibility would be to use inverted gamma densities. In both cases, as shown in the appendix, the conditional posteriors are inverted gamma, enabling direct simulation within the Gibbs sampler routine. The use of inverted gamma priors, which are conditionally natural conjugate for the variance parameters, allows for implementing any prior knowledge that may be available. However, difficulties may arise when using inverted gamma densities with very low shape and scale, to represent vague prior notions. If the likelihood function does not fall away rapidly enough, for low values of the variance parameters, then the posterior may be significantly influenced by the way the prior rises toward zero. Low values are not uncommon for innovation variances in components models, particularly for the trend, and in some cases for the irregular variance as well.

For these reasons, flat priors are implemented in all applications. In the univariate case, we use uniform densities on the interval $[\epsilon, a]$ where

$\epsilon > 0$ is very small and a is sufficiently large, so that the posterior domain is completely covered. In practice, with the logarithms of macroeconomic variables being modeled, choosing values such as $\epsilon = 10^{-100}$, $a = 1$ gives an interval wide enough to envelope the likelihood function to a high degree of accuracy. Specifying each prior as a bounded uniform density gives a well-defined basis for the implementation as it ensures propriety of posteriors. In this case the conditional posterior is inverted gamma with the shape and scale entirely determined by the data. The draws for $\sigma_\zeta^2, \sigma_\kappa^2, \sigma_\epsilon^2$ in the MCMC routine may then be obtained by standard sampling, as discussed in the next appendix.

The multivariate case

Similarly, in the multivariate case, we work with the generalization of the inverted gamma distribution, i.e., the Inverted Wishart (IW) distribution; for details see e.g. Zellner (1971). Using independent inverted Wishart densities or flat priors on the variance matrices $\Sigma_\kappa, \Sigma_\zeta, \Sigma_\epsilon$ produces independent conditional posteriors of the *IW* form. To reflect informative expectations, one may use independent *IW* priors for the variance matrices and choose the scale and shape parameters appropriately; the natural conjugate property holds as in the univariate case.

However, if one assumes relatively vague prior knowledge, then flat priors are advisable. Using an *IW* density with small values for the scale and shape would represent a poor general strategy since the same difficulties, described for the inverted gamma priors in the univariate case, apply to the inverted Wishart class. As the number of series being modeled increases, the dimensionality of the vector of variance and covariance parameters rises, so the pattern of distortions in the posterior results may become rather complex. We prefer noninformative priors that will reflect the shape of the multivariate likelihood surface, and therefore we use flat priors for the variance matrices $\Sigma_\kappa, \Sigma_\zeta, \Sigma_\epsilon$ defined over suitably large regions.

C MCMC routine

In this appendix an MCMC routine, known as a ‘Metropolis-Hastings within Gibbs’ sampling algorithm, is presented for obtaining posterior draws for the class of models in (1). The algorithm produces pseudo-random drawings from the combined joint density of state and parameter vectors. The state space form of (1) was set out in Trimbur (2005).

We start by deriving the density of the states conditional on the hyper-parameters, which plays a key role in the MCMC routine. This conditional density is obtained from the state space form and initial conditions. The model defined in equations (1) to (4) expresses the observed process as a sum of unobserved components. The $(2+2n) \times 1$ state vector contains the trend, cycle, and the associated processes used in their definition, that is

$$\boldsymbol{\alpha}_t = (\mu_t \quad \beta_t \quad \psi_{n,t} \quad \psi_{n,t}^* \quad \psi_{n-1,t} \quad \cdot \quad \cdot \quad \cdot \quad \psi_{1,t} \quad \psi_{1,t}^*)' \quad (17)$$

We will use the notation $\boldsymbol{\alpha}_t = (\boldsymbol{\mu}'_t, \boldsymbol{\psi}'_{n,t})$ to refer to the cycle and trend portions of the state vector. Denote by $\boldsymbol{\alpha}$ the $(m + 2n) \times T$ matrix formed by stacking the complete set of state vectors over the sample period:

$$\boldsymbol{\alpha} = [\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_T] \quad (18)$$

This will be referred to as the state matrix. The full density of the state matrix, given the parameters is

$$p(\boldsymbol{\alpha}|\boldsymbol{\theta}) = p(\boldsymbol{\alpha}_1|\boldsymbol{\theta}) \prod_{t=2}^T p(\boldsymbol{\alpha}_t|\boldsymbol{\alpha}_{t-1}, \boldsymbol{\theta}) \quad (19)$$

The nonstationary trend state vector is initialized with a diffuse prior; thus the variance matrix $var(\boldsymbol{\mu}_1)$ is equal to the limit $\sigma_*^2 I_2$ as $\sigma_*^2 \rightarrow \infty$. The initialization of the cyclical part is based on the unconditional distribution of the cyclical state vector $\boldsymbol{\psi}_{n,t}$, which is Gaussian with mean zero and known covariance matrix; see Trimbur (2005). The density of the initial state, given diffuse initial conditions for the trend part, depends only on $\boldsymbol{\psi}_1$ (which contains the starting values of the elements of the cyclical state vector) and the cyclical parameters :

$$p(\boldsymbol{\alpha}_1|\sigma_\kappa^2, \rho, \lambda_c) \propto |\boldsymbol{\Gamma}|^{-1/2} \exp\left(-\frac{1}{2}\boldsymbol{\psi}'_1 \boldsymbol{\Gamma}^{-1} \boldsymbol{\psi}_1\right) \quad (20)$$

In (20), the notation $\boldsymbol{\Gamma}$ stands for $var(\boldsymbol{\psi}_{n,t})$; this matrix is a function of the cycle order and parameters; see Trimbur (2005).

Conditional on $\boldsymbol{\alpha}_{t-1}$, some elements of $\boldsymbol{\alpha}_t$ are known so we define the reduced state vector, transition matrix, and covariance matrix by:

$$\boldsymbol{\alpha}_t^* = \begin{bmatrix} \beta_t \\ \psi_{1,t} \\ \psi_{1,t}^* \end{bmatrix}, \quad \mathbf{T}^* = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \rho \cos \lambda_c & \rho \sin \lambda_c \\ 0 & -\rho \sin \lambda_c & \rho \cos \lambda_c \end{bmatrix}, \quad \boldsymbol{\Omega}^* = \begin{bmatrix} \sigma_\zeta^2 & 0 & 0 \\ 0 & \sigma_\kappa^2 & 0 \\ 0 & 0 & \sigma_\kappa^2 \end{bmatrix}$$

Then the one-step ahead density is

$$p(\boldsymbol{\alpha}_t | \boldsymbol{\alpha}_{t-1}, \boldsymbol{\theta}) = |\boldsymbol{\Omega}^*|^{-1/2} \exp \left\{ -\frac{1}{2} (\boldsymbol{\alpha}_t^* - \mathbf{T}^* \boldsymbol{\alpha}_{t-1}^*)' \boldsymbol{\Omega}^{*-1} (\boldsymbol{\alpha}_t^* - \mathbf{T}^* \boldsymbol{\alpha}_{t-1}^*) \right\}$$

Therefore,

$$p(\boldsymbol{\alpha} | \boldsymbol{\theta}) \propto |\boldsymbol{\Gamma}|^{-1/2} |\boldsymbol{\Omega}^*|^{-(T-1)/2} \exp \left\{ -\frac{1}{2} \boldsymbol{\psi}'_1 \boldsymbol{\Gamma}^{-1} \boldsymbol{\psi}_1 - \frac{1}{2} \sum_{t=2}^T (\boldsymbol{\alpha}_t^* - \mathbf{T}^* \boldsymbol{\alpha}_{t-1}^*)' \boldsymbol{\Omega}^{*-1} (\boldsymbol{\alpha}_t^* - \mathbf{T}^* \boldsymbol{\alpha}_{t-1}^*) \right\}$$

Given the partitioned structure of \mathbf{T}^* , this can be rewritten as

$$p(\boldsymbol{\alpha} | \boldsymbol{\theta}) \propto |\boldsymbol{\Gamma}|^{-1/2} \sigma_\kappa^{-2(T-1)} \sigma_\zeta^{-(T-1)} \exp \left\{ -\frac{1}{2} \boldsymbol{\psi}'_1 \boldsymbol{\Gamma}^{-1} \boldsymbol{\psi}_1 - \frac{1}{2\sigma_\kappa^2} \sum_{t=2}^T c_t \right\} \quad (21)$$

$$\times \exp \left\{ -\frac{1}{2\sigma_\zeta^2} \sum_{t=2}^T (\beta_t - \beta_{t-1})^2 \right\}$$

where

$$c_t = (\psi_{1,t} - \rho \cos \lambda_c \psi_{1,t-1} - \rho \sin \lambda_c \psi_{1,t-1}^*)^2 + (\psi_{1,t}^* + \rho \sin \lambda_c \psi_{1,t-1} - \rho \cos \lambda_c \psi_{1,t-1}^*)^2$$

In designing an efficient MCMC routine in order to produce draws $\{\boldsymbol{\theta}^{(i)}, \boldsymbol{\alpha}^{(i)}\}$ from the joint posterior $p(\boldsymbol{\theta}, \boldsymbol{\alpha} | \mathbf{y})$, it is convenient that the conditional densities of the $\boldsymbol{\alpha}$ and the variance parameters $\{\sigma_\kappa^2, \sigma_\zeta^2, \sigma_\varepsilon^2\}$ are known up to proportionality and can be handled using Gibbs methods. We label this steps 1 and 2. The remaining parameters ρ and λ_c are generated using an M-H method in steps 3 and 4, respectively. Similar algorithms, which capitalize on the state space framework, may be found in Carter and Kohn (1994), Koop and van Dijk (2000), and Kleijn and van Dijk (2005).

The four sections of the routine correspond to a set of complete conditional densities for the joint posterior density $p(\boldsymbol{\theta}, \boldsymbol{\alpha} | \mathbf{y})$. Taken on their

own, the variates $\{\boldsymbol{\theta}^{(i)}\}$ represent draws from the marginal posterior $p(\boldsymbol{\theta}|\mathbf{y})$. Similarly, the $\{\boldsymbol{\alpha}^{(i)}\}$ serve as draws from $p(\boldsymbol{\alpha}|\mathbf{y})$.

We start with an initial value for the parameter vector, $\boldsymbol{\theta}^{(0)}$. Each iteration i involves the following sequence of draws, organised into four steps:

1. $\boldsymbol{\alpha}^{(i)}$ is drawn from $p(\boldsymbol{\alpha}|\boldsymbol{\theta}^{(i-1)}, \mathbf{y})$:

The state space form of the model was shown in the previous appendix. This enables the direct application of the general simulation smoother of Durbin and Koopman (2002).

2. The variances $\{\sigma_\kappa^2, \sigma_\zeta^2, \sigma_\varepsilon^2\}$ are sampled as a group from the joint density $p(\sigma_\kappa^2, \sigma_\zeta^2, \sigma_\varepsilon^2 | \rho, \lambda_c, \boldsymbol{\alpha}^{(i)}, \mathbf{y})$:

As the priors on all parameters are assumed independent, we start with

$$p(\sigma_\kappa^2, \sigma_\zeta^2, \sigma_\varepsilon^2 | \rho, \lambda_c, \boldsymbol{\alpha}, \mathbf{y}) \propto p(\mathbf{y}|\boldsymbol{\alpha}, \boldsymbol{\theta})p(\boldsymbol{\alpha}|\boldsymbol{\theta})p(\sigma_\kappa^2)p(\sigma_\zeta^2)p(\sigma_\varepsilon^2) \quad (22)$$

Conditional on ρ and λ_c , the reduced state transition matrix \mathbf{T}^* is fixed. The unconditional variance matrix of the cyclical part of the state can be written as $\mathbf{\Gamma} = \sigma_\kappa^2 \mathbf{\Gamma}^*$ where $\mathbf{\Gamma}^*$ is a $2n \times 2n$ matrix that depends only on ρ and λ_c for given n . The determinant of $\mathbf{\Gamma}$ is therefore proportional to σ_κ^{4n} , and it follows that

$$\begin{aligned} p(\boldsymbol{\alpha}|\boldsymbol{\theta}) &\propto \sigma_\kappa^{-(T+n-1)} \exp \left\{ -\frac{1}{2\sigma_\kappa^2} \boldsymbol{\psi}'_1 \mathbf{\Gamma}^{*-1} \boldsymbol{\psi}_1 - \frac{1}{2\sigma_\kappa^2} \sum_{t=2}^T c_t \right\} \\ &\times \sigma_\zeta^{-(T-1)/2} \exp \left\{ -\frac{1}{2\sigma_\zeta^2} \sum_{t=2}^T (\beta_t - \beta_{t-1})^2 \right\} \end{aligned}$$

Furthermore,

$$p(\mathbf{y}|\boldsymbol{\alpha}, \boldsymbol{\theta}) \propto \sigma_\varepsilon^{-T/2} \exp \left\{ -\frac{1}{2\sigma_\varepsilon^2} \sum_{t=1}^T (y_t - z_t \boldsymbol{\alpha}_t)^2 \right\} \quad (23)$$

Given independent flat priors on the variances, the joint conditional posterior factors into three independent *IG* densities. Comparing the above expressions with the standard form of the inverted gamma, we see that the scales $\{c_\kappa, c_\zeta, c_\varepsilon\}$ and shape parameters $\{S_\kappa, S_\zeta, S_\varepsilon\}$ of the posteriors are

$$c_\kappa = 2(T + n - 2), \quad c_\zeta = T - 3, \quad c_\varepsilon = T - 2$$

$$S_\kappa = \boldsymbol{\psi}_1 \boldsymbol{\Gamma}^{*-1} \boldsymbol{\psi}_1 + \sum_{t=2}^T c_t, \quad S_\zeta = \sum_{t=2}^T (\beta_t - \beta_{t-1})^2, \quad S_\varepsilon = \sum_{t=1}^T (y_t - z_t \boldsymbol{\alpha}_t)^2$$

For our application with quarterly US real GDP and investment, the flat priors we used helped avoid distortion of the posterior shape. In particular, the slope innovation variance tended to be low, corresponding to a relatively smooth trend. Our experience with simulations showed that the use of an inverted gamma prior density, with shape and scale set to low values to aim at an effectively noninformative, could distort the resulting estimates; in particular, the marginal posterior was concentrated near $\sigma_\zeta^2 = 0$. This reflected the way in which the *IG* prior rose indefinitely toward zero. The use of flat prior better reflected the variation in the trend apparent in the data, as represented in the likelihood surface. This was also the case for the irregular variance for $n = 1$. In all applications in the current paper, we used flat priors on the variance parameters to represent noninformative expectations. A Bayesian treatment of a state space class of models which relied on inverted gamma priors was introduced in Fruhwirth-Schnatter (1994). The extended class of models we investigate include the presence of cyclical components, hence it is necessary to handle ρ and λ_c in the analysis by embedding them in the Gibbs sampler.

3. $\{\rho^{(i)}\}$ is drawn from $p(\rho | \sigma_\kappa^2, \sigma_\zeta^2, \sigma_\varepsilon^2, \lambda_c, \boldsymbol{\alpha}^{(i)}, \mathbf{y})$:

The parameter ρ determines how the white noise shocks feed through to the cyclical component over time. A uniform prior is used to ensure the value of ρ lie between zero and one, but otherwise gives no information about its location within the interval. Denote the prior by $p(\rho)$. Using Bayes theorem for densities,

$$p(\rho | \sigma_\kappa^2, \sigma_\zeta^2, \sigma_\varepsilon^2, \lambda_c, \boldsymbol{\alpha}, Y) \propto p(\rho) p(Y | \boldsymbol{\alpha}, \boldsymbol{\theta}) p(\boldsymbol{\alpha} | \boldsymbol{\theta}) \propto p(\rho) p(\boldsymbol{\alpha} | \boldsymbol{\theta})$$

$$p(\rho | \sigma_\kappa^2, \sigma_\zeta^2, \sigma_\varepsilon^2, \lambda_c, \boldsymbol{\alpha}, Y) \propto p(\rho) |\boldsymbol{\Gamma}|^{-1/2} \exp \left\{ -\frac{1}{2} \boldsymbol{\psi}_1' \boldsymbol{\Gamma}^{*-1} \boldsymbol{\psi}_1 - \frac{1}{2\sigma_\kappa^2} \sum_{t=2}^T c_t \right\} \quad (24)$$

Recall from (20) and (21) that c_t and $\boldsymbol{\Gamma}$ both depend on ρ . The use of an M-H step made it possible to handle (24) in a straightforward manner. A

random walk proposal density was used. Note that if informative priors on ρ were used, for instance to capture the relationship between ρ and the cyclical order n , then the algorithm would remain the same, with the scale in the M-H step adjusted.

4. Analogously to the case for ρ , for the frequency parameter we have

$$p(\lambda_c | \sigma_\kappa^2, \sigma_\zeta^2, \sigma_\varepsilon^2, \rho, \boldsymbol{\alpha}, Y) \propto p(\lambda_c) p(Y | \boldsymbol{\alpha}, \boldsymbol{\theta}) p(\boldsymbol{\alpha} | \boldsymbol{\theta}) \propto p(\lambda_c) p(\boldsymbol{\alpha} | \boldsymbol{\theta}) \quad (25)$$

$$p(\lambda_c | \sigma_\kappa^2, \sigma_\zeta^2, \sigma_\varepsilon^2, \rho, \boldsymbol{\alpha}, Y) \propto p(\lambda_c) |\Gamma|^{-1/2} \exp \left\{ -\frac{1}{2} \boldsymbol{\psi}'_1 \boldsymbol{\Gamma}^{*-1} \boldsymbol{\psi}_1 - \frac{1}{2\sigma_\kappa^2} \sum_{t=2}^T c_t \right\} \quad (26)$$

With a beta density for $p(\lambda_c)$, the conditional posterior is nonstandard as trigonometric functions of λ_c occur within the exponent. Again, an M-H step was applied, with candidates generated by a random walk, see e.g., Chib and Greenberg (1996), Koop and van Dijk (2000), and Bos, Mahieu, and van Dijk (2000). The scales of the proposal densities for ρ and λ_c were calibrated for each model and prior based on the suggested rule of thumb noted in Chib and Greenberg (1996). That is, the variances for the random walk innovations were set to attain acceptance probabilities of 30-40%. The simulations performed well within this range, though the exact choices for the scale parameters are not crucial. By adapting the M-H properties to different priors, cyclical orders, and datasets, the efficiency of the posterior computations was enhanced. Calibration of the scales for ρ and λ_c was based on trial runs of the Gibbs sampler with a small number of iterations and typically required less than a minute.

The ultimate set of drawings $\{\boldsymbol{\theta}^{(j)}, \boldsymbol{\alpha}^{(j)}\}$, $j = 1, \dots, J$, used in the final posterior sample were a subset of the $\{\boldsymbol{\theta}^{(i)}, \boldsymbol{\alpha}^{(i)}\}$ produced by the MCMC routine. Specifically, to improve the efficiency of the analysis, correlations between successive posterior draws were reduced by running multiple iterations for each draw and by discarding, or ‘burning’, a number of initial iterations. The starting values for the parameters had a negligible impact on the final results assuming that plausible values were used. In the univariate case, we ran twenty repetitions of the Gibbs sampler to produce each draw, and we burned the first 5000 iterates as the chain moved toward convergence. The multivariate applications used five iterations per posterior draw after burning the first 2000. The final posterior sample consisted of 5,000 draws for

parameters and components of multivariate models and 10,000 draws in the univariate examples. Similarly, in the bivariate inflation-output application, after burning the first 5000 iterates, the final posterior sample consisted of 5000 draws, with ten iterations per draw. The autocorrelations for each sequence of parameter draws typically fell to near zero after just a few lags. The slope innovation variances generally had the highest degree of correlation between successive parameter draws; the correlogram for the period typically appeared similar to that of the frequency. Enhancing the informational content of the final set of posterior draws led to more efficient estimation of trend and cyclical properties.

Evaluation of marginal likelihood

Marginal likelihoods may be estimated using the posterior output. Kass and Raftery (1995) and DiCiccio, Kass, Raftery, and Wasserman (1997) discuss computational approaches based on the Laplace method. This method, which relies on a multivariate Gaussian approximating density to the posterior, is appealing due to its simplicity. For our particular MCMC study, calibration of M-H steps and the use of multiple iterations per draw helped to produce a posterior sample with favourable properties. In other situations, for example, if the chain shows a tendency to veer away for long intervals, one may use an estimator of the posterior covariance matrix that is robust to multivariate outliers.

The Laplace approximation requires the evaluation of the prior and likelihood, preferably at a point of high density. We used the posterior mean, which is simple to compute. Alternatively, the posterior mode or median could be estimated. For the basic Laplace method, the estimated posterior ordinate is then obtained directly from the sample posterior covariance matrix. This approximation, based on a Gaussian kernel, is appropriate for comparing marginal likelihoods in cases where the posterior is relatively well-behaved, for instance it is not multimodal in a high density region for any of the parameters.

State matrix and Bayesian smoother

Taken on their own, the draws $\{\boldsymbol{\alpha}^{(j)}\}$ are variates from $p(\boldsymbol{\alpha}|Y)$. The posterior density of $\boldsymbol{\alpha}$ involves a substantial amount of information as it describes the relationships among subsets of the state vector at various points in time. The associated marginal densities of the components are of primary interest for current purposes, but one could in theory examine virtually any aspect of the joint density of the cyclical estimates by averaging over different functions of draws. This permits a great deal of flexibility in analyzing the

cycle and its relative positions at different points in time. The conditional mean of the cyclical component, given the data, represents the Bayesian counterpart of the classical smoother.

Conditional posteriors in multivariate case

In extending the Bayesian analysis of model (1) to the multivariate setup where $N > 1$, the algorithm remains essentially the same as for the univariate case. The main novelty lies in the treatment of the N -dimensional variance matrices.

We now set out the exact expressions for the conditional densities, that combine to form the Gibbs sampler for the multivariate model. The density of the initial state is now

$$p(\boldsymbol{\alpha}_1 | \boldsymbol{\Sigma}_\kappa, \rho, \lambda_c) \propto |\boldsymbol{\Sigma}_\psi|^{-1/2} \exp\left(-\frac{1}{2} \boldsymbol{\psi}'_1 \boldsymbol{\Sigma}_\psi^{-1} \boldsymbol{\psi}_1\right)$$

The conditional density of the state matrix, given the parameters, is

$$\begin{aligned} p(\boldsymbol{\alpha} | \boldsymbol{\Sigma}_\zeta, \boldsymbol{\Sigma}_\kappa, \rho, \lambda_c) &\propto |\boldsymbol{\Sigma}_\psi|^{-1/2} |\mathbf{I}_2 \otimes \boldsymbol{\Sigma}_\kappa|^{-(T-1)/2} |\boldsymbol{\Sigma}_\zeta|^{-(T-1)/2} & (27) \\ &\times \exp\left\{-\frac{1}{2} \boldsymbol{\psi}'_1 \boldsymbol{\Sigma}_\psi^{-1} \boldsymbol{\psi}_1 - \frac{1}{2} \sum_{t=2}^T \mathbf{c}'_t (\mathbf{I}_2 \otimes \boldsymbol{\Sigma}_\kappa)^{-1} \mathbf{c}_t\right\} \\ &\times \exp\left\{-\frac{1}{2} \sum_{t=2}^T (\boldsymbol{\beta}_t - \boldsymbol{\beta}_{t-1})' \boldsymbol{\Sigma}_\zeta^{-1} (\boldsymbol{\beta}_t - \boldsymbol{\beta}_{t-1})\right\} \end{aligned}$$

where $\mathbf{c}_t = [\boldsymbol{\psi}'_{1,t}, \boldsymbol{\psi}^*_{1,t}]' - (\mathbf{T} \otimes \mathbf{I}_N) [\boldsymbol{\psi}'_{1,t-1}, \boldsymbol{\psi}^*_{1,t-1}]'$ with $\boldsymbol{\psi}_{1,t} = [\psi_{1,t}^1, \dots, \psi_{1,t}^N]'$, $\boldsymbol{\psi}^*_{1,t} = [\psi_{1,t}^{1*}, \dots, \psi_{1,t}^{N*}]'$.

The summation in the cyclical part may be written as

$$\begin{aligned} \sum_{t=2}^T \mathbf{c}'_t (\mathbf{I}_2 \otimes \boldsymbol{\Sigma}_\kappa)^{-1} \mathbf{c}_t &= \sum_{t=2}^T \text{tr} [(\mathbf{I}_2 \otimes \boldsymbol{\Sigma}_\kappa^{-1}) \mathbf{c}_t \mathbf{c}'_t] \\ &= \text{tr}(\boldsymbol{\Sigma}_\kappa^{-1} \mathbf{c}_{1,t} \mathbf{c}'_{1,t}) + \text{tr}(\boldsymbol{\Sigma}_\kappa^{-1} \mathbf{c}_{2,t} \mathbf{c}'_{2,t}) \end{aligned}$$

where $\mathbf{c}'_t = [\mathbf{c}'_{1,t} \quad \mathbf{c}'_{2,t}]$. For this partition of \mathbf{c}_t , the upper half $\mathbf{c}_{1,t}$ ($N \times N$) corresponds to the N first order cycles $[\psi_{1,t}^1, \dots, \psi_{1,t}^N]'$, and the lower half $\mathbf{c}_{2,t}$ to the auxiliaries, $[\psi_{1,t}^{1*}, \dots, \psi_{1,t}^{N*}]'$.

Thus the cyclical part of the conditional density of $\boldsymbol{\alpha}$ may be written as

$\mathbf{\Gamma} = \sigma_\kappa^2 \mathbf{\Gamma}^*$ where $\mathbf{\Gamma}^*$ is a $2n \times 2n$ matrix that depends only on ρ and λ_c for given n .

$$p(\boldsymbol{\alpha} | \boldsymbol{\Sigma}_\kappa, \rho, \lambda_c) \propto |\mathbf{\Gamma}^* \otimes \boldsymbol{\Sigma}_\kappa|^{-1/2} |\mathbf{I}_2 \otimes \boldsymbol{\Sigma}_\kappa|^{-(T-1)/2} \\ \times \exp \left\{ -\frac{1}{2} \boldsymbol{\psi}'_1 [V(\rho, \lambda_c, n) \otimes \boldsymbol{\Sigma}_\kappa]^{-1} \boldsymbol{\psi}_1 - \frac{1}{2} \text{tr} (\boldsymbol{\Sigma}_\kappa^{-1} \mathbf{G}) \right\}$$

where

$$\mathbf{G} = \sum_{t=2}^T (\mathbf{c}_{1,t} \mathbf{c}'_{1,t} + \mathbf{c}_{2,t} \mathbf{c}'_{2,t})$$

Since

$$|\mathbf{\Gamma}^* \otimes \boldsymbol{\Sigma}_\kappa| = |\mathbf{\Gamma}^*|^N |\boldsymbol{\Sigma}_\kappa|^{2n}$$

and $|\mathbf{I}_2 \otimes \boldsymbol{\Sigma}_\kappa| = |\boldsymbol{\Sigma}_\kappa|^2$ we have

$$|\mathbf{\Gamma}^* \otimes \boldsymbol{\Sigma}_\kappa|^{-1/2} |\mathbf{I}_2 \otimes \boldsymbol{\Sigma}_\kappa|^{-(T-1)/2} = |\mathbf{\Gamma}^*|^{-N/2} |\boldsymbol{\Sigma}_\kappa|^{-(n+T-1)}$$

Let the elements of block i, j of $\mathbf{\Gamma}^*$ be denoted by

$$\{\mathbf{\Gamma}^*\}_{i,j} = \begin{bmatrix} \mathbf{\Gamma}^*_{i,j} & \mathbf{\Gamma}^*_{i,j^*} \\ \mathbf{\Gamma}^*_{i^*,j} & \mathbf{\Gamma}^*_{i^*,j^*} \end{bmatrix}$$

The complete cyclical state vector at each t is partitioned as

$$\boldsymbol{\psi}_t = [\psi_{n,t}^1, \dots, \psi_{n,t}^N, \psi_{n,t}^{1*}, \dots, \psi_{n,t}^{N*}, \dots, \psi_{1,t}^1, \dots, \psi_{1,t}^N, \psi_{1,t}^{1*}, \dots, \psi_{1,t}^{N*}]' \\ = [\boldsymbol{\psi}'_{n,t}, \boldsymbol{\psi}^*_{n,t}, \boldsymbol{\psi}'_{n-1,t}, \dots, \boldsymbol{\psi}'_{1,t}, \boldsymbol{\psi}^*_{1,t}]'$$

The term involving the initial state vector in the exponent may be written as

$$\boldsymbol{\psi}'_1 [\mathbf{\Gamma}^* \otimes \boldsymbol{\Sigma}_\kappa]^{-1} \boldsymbol{\psi}_1 = \frac{1}{2} \text{tr} (\boldsymbol{\Sigma}_\kappa^{-1} \mathbf{H}) \quad (28)$$

where the $(N \times N)$ matrix \mathbf{H} is given by

$$\mathbf{H} = \sum_{i=1}^n \left(\sum_{j=1}^n \mathbf{\Gamma}^*_{i,j}^{-1} \boldsymbol{\psi}_{j,1} \boldsymbol{\psi}'_{i,1} + \sum_{j=1}^n \mathbf{\Gamma}^*_{i,j^*}^{-1} \boldsymbol{\psi}^*_{j,1} \boldsymbol{\psi}'_{i,1} \right) \\ + \sum_{i=1}^n \left(\sum_{j=1}^n \mathbf{\Gamma}^*_{i^*,j}^{-1} \boldsymbol{\psi}_{j,1} \boldsymbol{\psi}^*_{i,1} + \sum_{j=1}^n \mathbf{\Gamma}^*_{i^*,j^*}^{-1} \boldsymbol{\psi}^*_{j,1} \boldsymbol{\psi}^*_{i,1} \right)$$

This expression is derived by first noting that

$$[\mathbf{\Gamma}^* \otimes \mathbf{\Sigma}_\kappa]^{-1} = tr [(\mathbf{\Gamma}^{*-1} \otimes \mathbf{\Sigma}_\kappa^{-1})\boldsymbol{\psi}_1\boldsymbol{\psi}'_1]$$

Once the Kronecker product and multiplication are applied, the sum of the diagonal elements gives (28).

The summation for the trend part of the conditional density can similarly be written as

$$\sum_{t=2}^T (\boldsymbol{\beta}_t - \boldsymbol{\beta}_{t-1})' \boldsymbol{\Sigma}_\zeta^{-1} (\boldsymbol{\beta}_t - \boldsymbol{\beta}_{t-1}) = \sum_{t=2}^T tr [\boldsymbol{\Sigma}_\zeta^{-1} (\boldsymbol{\beta}_t - \boldsymbol{\beta}_{t-1}) (\boldsymbol{\beta}_t - \boldsymbol{\beta}_{t-1})']$$

In summary, the conditional density of the state matrix is

$$p(\boldsymbol{\alpha} | \mathbf{\Sigma}_\kappa, \rho, \lambda_c) \propto |\mathbf{\Gamma}^*|^{-N/2} |\mathbf{\Sigma}_\kappa|^{-(n+T-1)} \exp \left[-\frac{1}{2} tr (\boldsymbol{\Sigma}_\kappa^{-1} [\mathbf{G} + \mathbf{H}]) \right] \quad (29)$$

$$\times |\boldsymbol{\Sigma}_\zeta|^{-(T-1)/2} \exp \left\{ -\frac{1}{2} tr \left(\boldsymbol{\Sigma}_\zeta^{-1} \left[\sum_{t=2}^T (\boldsymbol{\beta}_t - \boldsymbol{\beta}_{t-1}) (\boldsymbol{\beta}_t - \boldsymbol{\beta}_{t-1})' \right] \right) \right\}$$

The conditional density of $\boldsymbol{\alpha}$ is shown in (29). The joint density of the observations given the hyperparameters and $\boldsymbol{\alpha}$ is a direct extension of the univariate expression:

$$p(Y | \boldsymbol{\alpha}, \boldsymbol{\theta}) \propto |\boldsymbol{\Sigma}_\varepsilon|^{-T/2} \exp \left\{ -\frac{1}{2} \sum_{t=1}^T tr [\boldsymbol{\Sigma}_\varepsilon^{-1} (y_t - z_t \boldsymbol{\alpha}_t) (y_t - z_t \boldsymbol{\alpha}_t)'] \right\}$$

The joint posterior of $\{\boldsymbol{\Sigma}_\zeta, \mathbf{\Sigma}_\kappa, \boldsymbol{\Sigma}_\varepsilon\}$ factors into a product of inverted Wishart densities with shape and scale matrix parameters given by

$$\boldsymbol{\Sigma}_\zeta : c_\zeta^* = c_\zeta + (T - 1), \quad \mathbf{S}_\zeta^* = \mathbf{S}_\zeta + \sum_{t=2}^T (\boldsymbol{\beta}_t - \boldsymbol{\beta}_{t-1}) (\boldsymbol{\beta}_t - \boldsymbol{\beta}_{t-1})'$$

$$\mathbf{\Sigma}_\kappa : c_\kappa^* = c_\kappa + 2(n + T - 1), \quad \mathbf{S}_\kappa^* = \mathbf{S}_\kappa + \mathbf{G} + \mathbf{H}$$

$$\boldsymbol{\Sigma}_\varepsilon : c_\varepsilon^* = c_\varepsilon + T, \quad \mathbf{S}_\varepsilon^* = \mathbf{S}_\varepsilon + \sum_{t=1}^T (y_t - z_t \boldsymbol{\alpha}_t) (y_t - z_t \boldsymbol{\alpha}_t)'$$

D Extension of state space model for bivariate estimation of output gap series

Here we show the modified state space form for the cyclical part of the inflation-output model for $n = 1$. The state vector is now

$$\alpha_t = [\psi_t^\pi \quad \psi_t^y \quad \psi_t^{\pi*} \quad \psi_t^{y*} \quad \psi_{t-1}^y \quad \psi_{t-2}^y]'$$

with transition equation

$$\begin{bmatrix} \psi_t^\pi \\ \psi_t^y \\ \psi_t^{\pi*} \\ \psi_t^{y*} \\ \psi_{t-1}^y \\ \psi_{t-2}^y \end{bmatrix} = \begin{bmatrix} \rho \cos \lambda_c & 0 & \rho \sin \lambda_c & 0 & 0 & 0 \\ 0 & \rho \cos \lambda_c & 0 & \rho \sin \lambda_c & 0 & 0 \\ -\rho \sin \lambda_c & 0 & \rho \cos \lambda_c & 0 & 0 & 0 \\ 0 & -\rho \sin \lambda_c & 0 & \rho \cos \lambda_c & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \psi_{t-1}^\pi \\ \psi_{t-1}^y \\ \psi_{t-1}^{\pi*} \\ \psi_{t-1}^{y*} \\ \psi_{t-2}^y \\ \psi_{t-3}^y \end{bmatrix} + \begin{bmatrix} \kappa_t^\pi \\ \kappa_t^y \\ \kappa_t^{\pi*} \\ \kappa_t^{y*} \\ 0 \\ 0 \end{bmatrix}$$

The corresponding part of the observation equation is

$$\begin{bmatrix} \pi_t \\ y_t \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & c_1 & c_2 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \psi_t^\pi \\ \psi_t^y \\ \psi_t^{\pi*} \\ \psi_t^{y*} \\ \psi_{t-1}^y \\ \psi_{t-2}^y \end{bmatrix} + \begin{bmatrix} \varepsilon_t^\pi \\ \varepsilon_t^y \end{bmatrix}$$

The form of the matrices for $n > 1$ is analogous.

E Additional figures

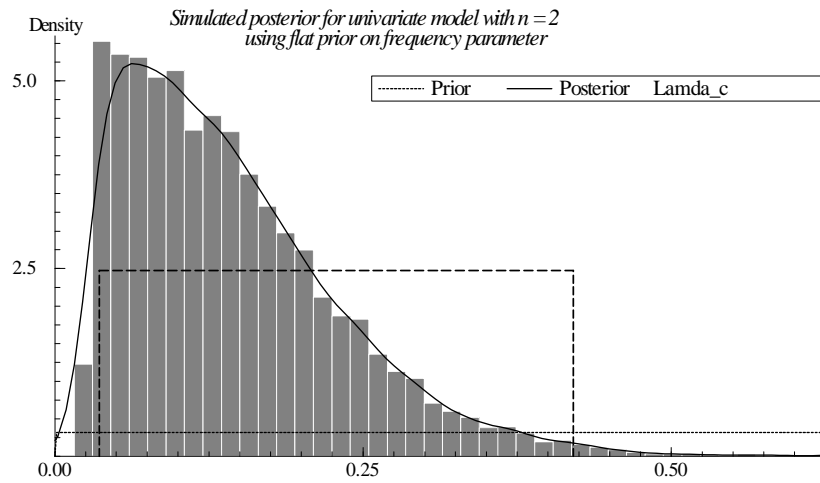


Figure 18: Marginal posterior of the cyclical frequency λ_c for $n = 2$, using a uniform prior on $[0, \pi]$. The series is quarterly US real GDP from 1947Q1 to 2004Q4. The graph shows the density function over the domain 0 to $\pi/5$; the midpoint corresponds to a five-year period.

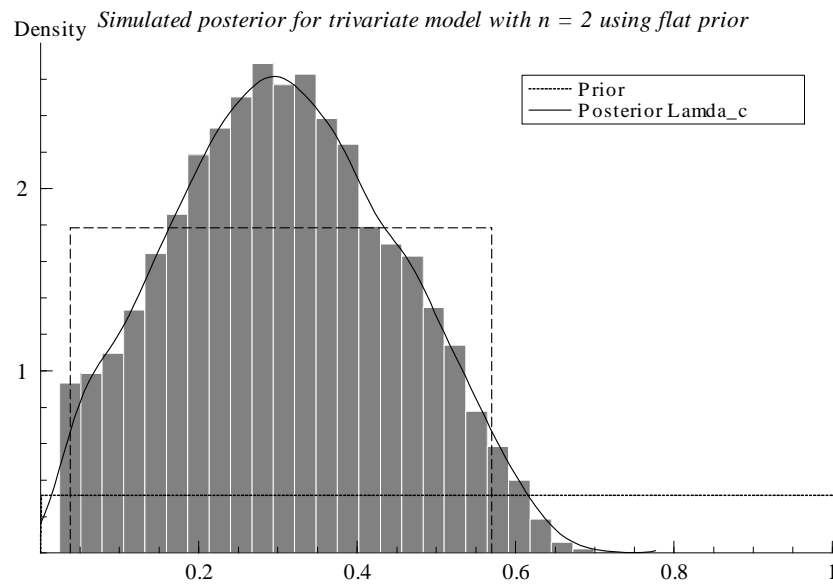


Figure 19: Estimated posterior density for the cyclical frequency parameter λ_c based on a trivariate model with quarterly series on consumption (C), investment (I), and gross domestic product (Y). The sample period is 1947Q1 to 2004Q4 and a flat prior on λ_c was used.

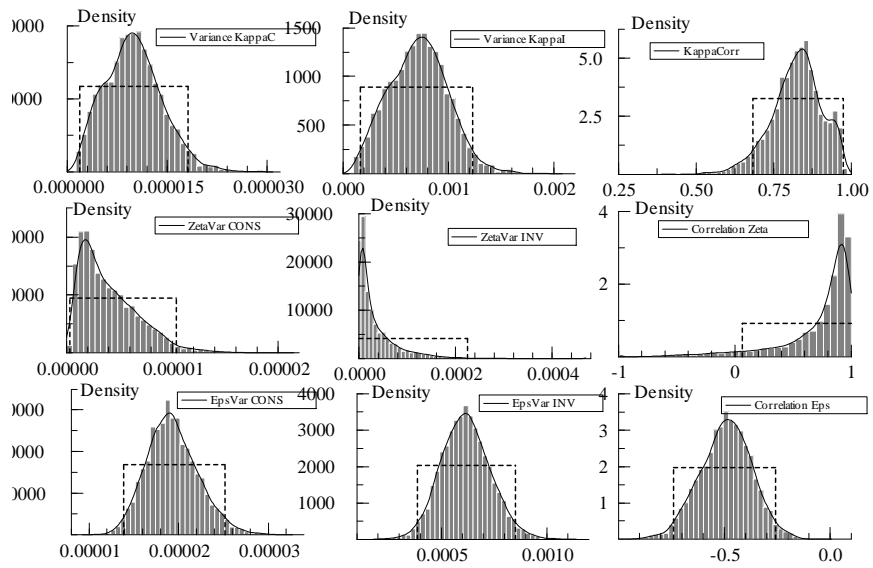


Figure 20: Marginal posteriors for variance parameters and correlations for consumption (C) and investment (I) based on the trivariate $C-I-Y$ model. The sample period is 1947Q1 to 2004Q4 and a flat prior on λ_c was used.

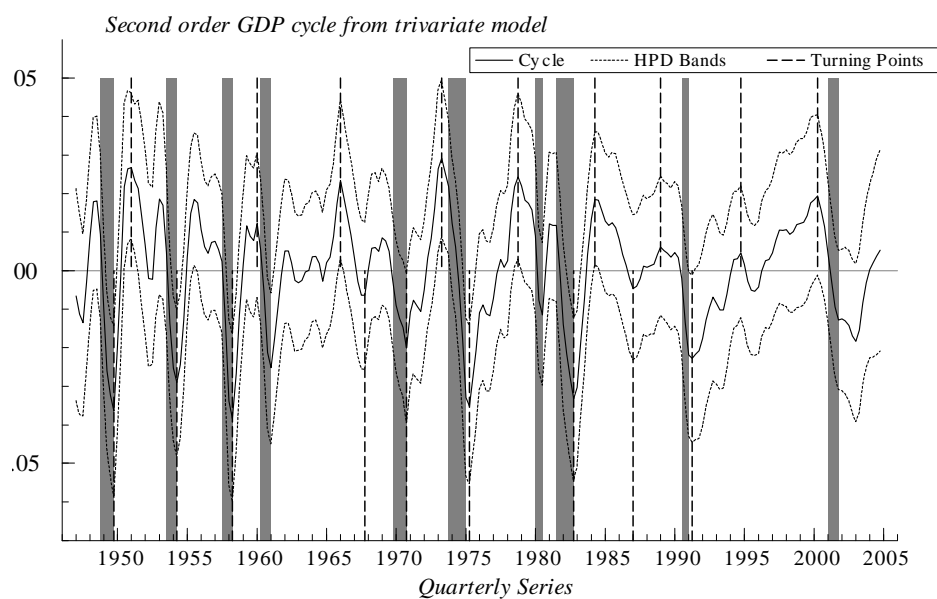


Figure 21: Estimated cycle in quarterly US real GDP (logarithms) from the trivariate C-I-Y model shown with 95% HPD bands for $n = 2$. The wide informative prior on λ_c was used. The sample period is 1947Q1 to 2004Q4.

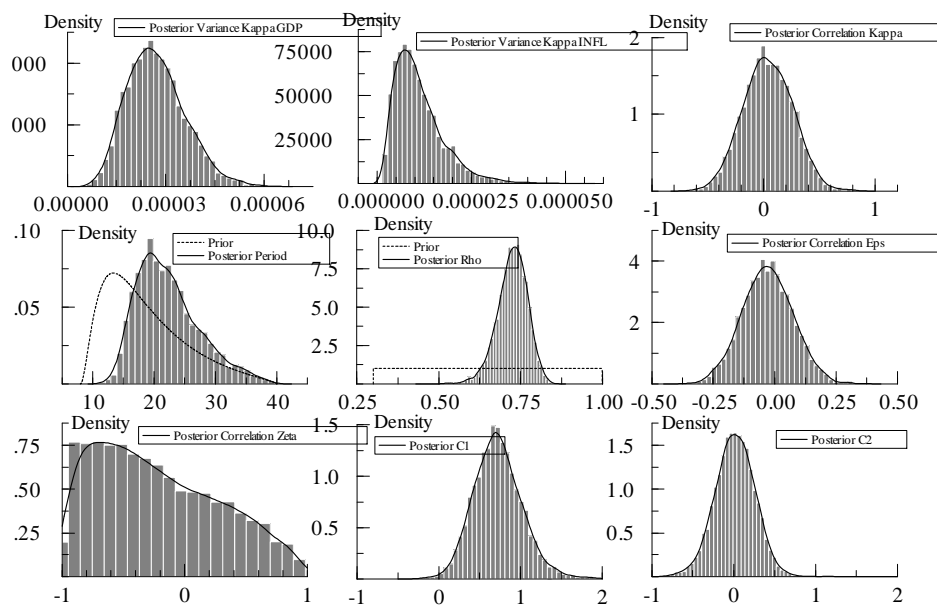


Figure 22: Posteriors for bivariate model of CPI inflation and quarterly US real GDP (logarithms) for sample period 1947Q2 to to 2004Q4, with $n = 2$ and a wide prior on λ_c .

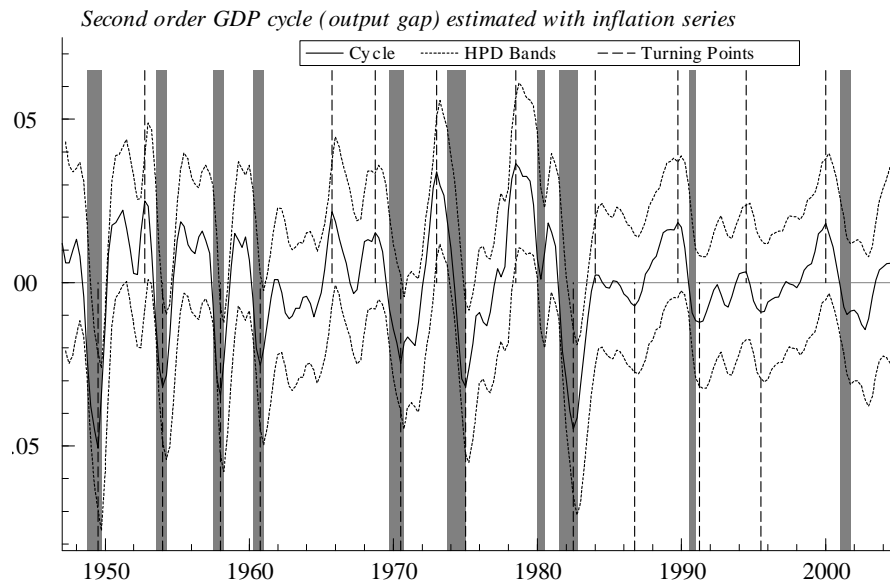


Figure 23: Estimated cyclical component (output gap) of quarterly US real GDP (logarithms) for sample period 1947Q2 to 2004Q4, based on the inflation-output bivariate model with $n = 2$ and a wide prior on λ_c . The series is shown with 95% HPD bands.