ELSEVIER Journal of Econometrics 78 (1997) 359–380

JOURNAL OF Econometrics

Bayesian analysis of seasonal unit roots and seasonal mean shifts

Philip Hans Franses[a], Henk Hoek[b,*], Richard Paap[b]

[a] Econometric Institute, Erasmus University Rotterdam, The Netherlands
[b] Tinbergen Institute, Erasmus University Rotterdam, NL-3000 DR Rotterdam, The Netherlands

Received July 1995; final version received July 1996

Abstract

In this paper we propose a Bayesian analysis of seasonal unit roots in quarterly observed time series. Seasonal unit root processes are useful to describe economic series with changing seasonal fluctuations. A natural alternative model for similar purposes contains deterministic seasonal mean shifts instead of seasonal stochastic trends. This leads to analysing seasonal unit roots in the presence of mean shifts using Bayesian techniques. Our method is illustrated using several simulated and empirical data.

Key words: Unit roots; Bayesian analysis; Seasonality; Structural breaks
JEL classification: C11, C12, C22.

1. Introduction

An empirical regularity of many quarterly observed macroeconomic time series is that the seasonal fluctuations do not seem constant over time. A class of models that is useful to describe such series is an autoregressive (AR) model with one or more so-called seasonal unit roots, see Hylleberg et al. (1990) (HEGY). In HEGY a test procedure for seasonal and nonseasonal unit roots in quarterly data is proposed. Seasonal unit roots correspond to the presence of stochastic trends at the seasonal frequencies. Since the usual purpose of univariate time series analysis is to obtain an indication of how to construct multivariate models like,

* Corresponding author.

The first author thanks the Royal Netherlands Academy of Arts and Sciences for its financial support. The third author thanks the Economic Research Foundation, which is part of the Netherlands Organization for Scientific Research (N.W.O.). Comments from two anonymous referees, Søren Johansen and participants of the 7th World Congress of the Econometric Society in Tokyo are gratefully acknowledged.

PII S0304-4076(97)00018-8

e.g., the seasonal cointegration model in Engle et al. (1993), it is important to have an adequate impression of the number of seasonal unit roots in individual series. Practical experience with statistical tests for seasonal unit roots reveals that seasonal unit roots are detected in many macroeconomic time series, see, e.g., Hylleberg et al. (1993).

A particular alternative model, for which tests for seasonal unit roots can be expected to have a low rejection frequency, is the AR model with one or more deterministic shifts in the seasonal means, see, e.g., Ghysels (1994). This conjecture is based on the results for nonseasonal time series discussed in, e.g., Perron (1989), where tests for zero frequency unit roots break down in the presence of shifts in mean or trend. The seasonal mean shift model may be useful in case statistical agencies start measuring economic quantities differently at some point in time or when two or more sources of data are combined into one single time series. The latter can occur for such variables as GNP and Employment. From an economic point of view, the seasonal mean shift model may reflect that economic agents change their behaviour instantaneously and permanently because of perceived exogenous shocks. Such shocks can be generated, for example, by changes in policies because of a government change, by a decision (halfway the observed sample) to execute tax changes in a certain season only, or (in case of nondurable consumption) by the fact that holiday periods sometimes change abruptly over time such that there appears a strong tendency to have holidays twice a year. From a statistical point of view, it is important to have reasonably precise knowledge of the properties of univariate time series, since such knowledge usually forms the basis of methods for constructing multivariate models. Finally, from a forecasting point of view, the seasonal unit root model and the seasonal mean shift model can result in widely different point-forecasts, and foremost, forecast intervals; see Paap et al. (1997). In this paper we analyse univariate quarterly time series processes for seasonal unit roots in the presence of seasonal mean shifts. We assume no a priori knowledge of the timing of such shifts, and no a priori knowledge of the presence of seasonal unit roots. We choose to use Bayesian techniques for this analysis.

The outline of our paper is as follows. In Section 2 we discuss the seasonal unit roots model and the model with seasonal mean shifts. In Section 3 we consider Bayesian analysis of the joint model, i.e., the model that nests both the seasonal unit roots representation and the seasonal mean shifts model. In this section we also elaborate on how our method differs from that advocated by Koop and Pitarakis (1992). A focal point in this section is the appropriate representation of the nesting model. Furthermore, we consider the specification of the prior distributions of the various parameters. In Section 4 we discuss some computational issues regarding the application of Bayesian techniques. In Section 5 we apply our approach to some simulated time series and three quarterly observed macroeconomic time series, i.e., total consumption in Sweden and in the United Kingdom and nondurable consumption in the United States. It appears that the evidence for seasonal unit roots becomes less pronounced for the Swedish and US series if we allow for seasonal mean shifts. Finally, in Section 6 we conclude this paper with some remarks and suggestions for further research.

2. Seasonal unit roots and mean shifts

A typical differencing filter that is applied to a quarterly observed time series, $y_t, t = 1, \ldots, T$, is the seasonal differencing filter $\Delta_4 = (1 - B^4)$, where B is the backward shift operator defined by $B^m y_t = y_{t-m}$ and where $\Delta_m \equiv (1 - B^m)$, $m = 1, 2, \ldots$. Since the polynomial $(1 - B^4)$ can be decomposed as

$$\begin{aligned}(1 - B^4) &= (1 - B)(1 + B)(1 - iB)(1 + iB)\\ &= (1 - B)(1 + B)(1 + B^2)\\ &= (1 - B)(1 + B + B^2 + B^3),\end{aligned} \tag{1}$$

it is clear that a time series which needs fourth differences to obtain stationarity has four roots on the unit circle. Such a series is said to be seasonally integrated. The nonseasonal root at the zero frequency (1) corresponds to a nonseasonal stochastic trend. The seasonal unit root at the frequency $\frac{1}{2}$ (-1) corresponds to two cycles per year and the seasonal unit roots at the frequencies $\frac{1}{4}$ and $\frac{3}{4}$ (i and $-i$) correspond to one cycle per year. Notice from (1) that the $(1 - B^4)$ filter can be decomposed in a part with a nonseasonal unit root and a part with three seasonal unit roots.

A procedure to test for nonseasonal and seasonal unit roots in a quarterly time series is developed by Hylleberg et al. (1990). This procedure is based on the auxiliary regression model

$$\begin{aligned}\Delta_4 y_t = {}& \sum_{s=1}^{4} d_s D_{st} + cT_t + \pi_1 y_{1,t-1} + \pi_2 y_{2,t-1} + \pi_3 y_{3,t-2} + \pi_4 y_{3,t-1}\\ &+ \sum_{i=1}^{k} \phi_i \Delta_4 y_{t-i} + \varepsilon_t,\end{aligned} \tag{2}$$

where D_{st} represent the usual seasonal dummies, where ε_t is assumed to be a standard white noise process, where T_t is a deterministic trend term ($T_t = 0, 1, 2, \ldots$) and where

$$\begin{aligned}y_{1,t} &= (1 + B + B^2 + B^3)y_t = y_t + y_{t-1} + y_{t-2} + y_{t-3},\\ y_{2,t} &= (-1 + B - B^2 + B^3)y_t = -y_t + y_{t-1} - y_{t-2} + y_{t-3},\\ y_{3,t} &= (-1 + B^2)y_t = -y_t + y_{t-2}.\end{aligned} \tag{3}$$

In practice, the value of k in (2) is unknown and has to be determined. The parameters d_s, $s = 1, \ldots, 4$, c, $\pi_j, j = 1, \ldots, 4$ and $\phi_i, i = 1, \ldots, k$, can be estimated using ordinary least squares (OLS). For unit root testing, the π_j parameters are the most relevant. In fact, if $\pi_1 = 0$ the series contains a unit root at the zero frequency. A unit root at the frequency $\frac{1}{2}$ (-1) corresponds to $\pi_2 = 0$. If $\pi_3 = 0$ and $\pi_4 = 0$ the series contains the roots i and $-i$. After applying OLS to (2) t- and F-tests are performed to check for the significance of the π_j parameters. Critical values of t-tests for the significance of the π_j's and an F-test for the significance of π_3 and π_4 are tabulated in Hylleberg et al. (1990). The asymptotic distributions of the various tests are discussed in Hylleberg et al. (1990) and Engle et al. (1993).

2.1. Seasonal mean shifts

The auxiliary test regression (2) and the estimated t- and F-values for the π_j parameters can be used to investigate the presence of seasonal and nonseasonal stochastic trends in y_t. An implication of a seasonal stochastic trend is that the seasonal fluctuations in y_t can change over time. There may however be alternative models for y_t that are useful in practice to describe changing seasonal fluctuations. For nonseasonal time series, it is shown in, e.g., Perron (1989) and Perron and Vogelsang (1992) that a mean shift biases unit root statistics towards nonrejection. Similarly, changing seasonal patterns can be generated by a model like

$$\Delta_1 y_t = \sum_{s=1}^{4} d_s D_{st} + \sum_{s=1}^{4} d_s^* D_{st}[I_{t \geqslant \tau}] + \varepsilon_t, \tag{4}$$

where $[I]$ is an indicator function and where seasonal mean shifts (from d_s to $d_s + d_s^*$) occur at time τ. In fact, given the results in Perron (1989), one may expect that the seasonal unit root statistics based on (2) to be biased towards nonrejection when (4) is the data-generating process (DGP). In (unreported) simulation experiments (see also Paap et al., 1997) we indeed find that neglecting seasonal mean shifts yields evidence of seasonal unit roots. In particular, when (4) is the DGP and (2) is estimated (using classical methods), we tend to find much evidence in favour of the seasonal unit root -1 (the bi-annual frequency), and a relatively smaller increase of evidence for the seasonal unit roots $\pm i$. Below, in Section 5, we report similar findings based on Bayesian techniques.

The purpose of the present paper is now to analyse (2) in the presence of seasonal mean shifts as in (4) using Bayesian techniques. We wish to confine our Bayesian analysis to the relevance of the $\pi_1, \ldots, \pi_4$ parameters in (2). This implies that we condition on a priori knowledge of the value of k, the number of lagged $\Delta_4 y_t$ variables in (2). Furthermore, we adopt a nesting framework in the sense that we somehow incorporate model (4) within (2), and then we focus on $\pi_1, \ldots, \pi_4$. In subsequent work we aim to analyse (2) when k is also a parameter of the model using a Bayesian method and to compare (2) and (4) in a nonnested framework using Bayes factors. At present, we consider these two extensions beyond the scope of this paper.

3. Bayesian analysis

For our Bayesian analysis we consider the series y_t in deviation from seasonal intercepts and the deterministic trend, i.e., we define

$$\tilde{y}_t = y_t - \delta_s D_{st} - \gamma T_t. \tag{5}$$

An advantage of this parameterization over that in (2) is that the parameters δ_s and γ in (5) have a more natural interpretation than the parameters d_s and c in (2). In fact, the parameters d_s and c in (2) have different interpretations under different unit root hypotheses. This point can be illustrated using the following two parameterizations of the AR(1) model with unknown mean μ:

$$y_t - \mu = \rho(y_{t-1} - \mu) + \varepsilon_t \tag{6}$$

and

$$y_t = \alpha + \rho y_{t-1} + \varepsilon_t. \tag{7}$$

If $|\rho| < 1$, μ has the interpretation of the mean of the process y_t, while in terms of Eq. (7), the mean is given by $\alpha/(1-\rho)$. If $\rho = 1$ however, α can be interpreted as a drift term, while μ is no longer identified. This reflects that a random walk process has no unconditional mean.

A similar identification problem arises in the analysis of the HEGY model

$$\Delta_4 \tilde{y}_t = \pi_1 \tilde{y}_{1,t-1} + \pi_2 \tilde{y}_{2,t-1} + \pi_3 \tilde{y}_{3,t-2} + \pi_4 \tilde{y}_{3,t-1} + \sum_{i=1}^{k} \phi_i \Delta_4 \tilde{y}_{t-i} + \varepsilon_t, \tag{8}$$

where $\tilde{y}_t$ is given in Eq. (5) and $\tilde{y}_{1,t}, \tilde{y}_{2,t}, \tilde{y}_{3,t}$ are defined by (3). To make the identification problem explicit, consider the following one-to-one transformation of the seasonal mean parameters δ_s into the parameters β_s, $s = 1,2,3,4$:

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \end{pmatrix}, \tag{9}$$

or $\beta = L\delta$, where $\beta = (\beta_1, \beta_2, \beta_3, \beta_4)'$, $\delta = (\delta_1, \delta_2, \delta_3, \delta_4)'$ and L a (4×4) transformation matrix. Using this transformation[1] and defining for notational

[1] Note that γ is always identified and has, independent of the hypothesis under consideration, the interpretation of a growth rate: $\mathrm{E}(\Delta_4 y_t) = 4\gamma$.

convenience, $\bar{y}_t = y_t - \gamma T_t$, we can write (8) as

$$\begin{aligned}\Delta_4 \bar{y}_t = {} & \pi_1(\bar{y}_{1,t-1} - \beta_1) + \pi_2(\bar{y}_{2,t-1} - \beta_2(-1)^t) \\ & + \pi_3(\bar{y}_{3,t-2} - \beta_3\kappa_{t-1} - \beta_4\kappa_{t-2}) + \pi_4(\bar{y}_{3,t-1} - \beta_3\kappa_t - \beta_4\kappa_{t-1}) \\ & + \sum_{i=1}^{k} \phi_i \Delta_4 \bar{y}_{t-i} + \varepsilon_t,\end{aligned} \tag{10}$$

where $\kappa_t = \frac{1}{2}(\mathrm{i}^t + (-\mathrm{i})^t)$ and where we assume that $t = 1$ corresponds to a first quarter observation. It is easy to see that in (10) the parameter β_1 is not identified if $\pi_1 = 0$. Analogously, if $\pi_2 = 0$ the parameter β_2 is not identified. Roots at the frequencies $\frac{1}{4}$ and $\frac{3}{4}$ ($\pi_3 = \pi_4 = 0$) imply that β_3 and β_4 are not identified.

Under the assumption $\varepsilon_t \sim \mathrm{N}(0, \sigma^2)$, the likelihood function for model (10) conditional on the initial observations $Y_0 = \{y_{-3-k}, y_{-2-k}, \ldots, y_0\}$, is given by

$$\mathcal{L}(Y_T \mid Y_0; \theta) = \prod_{t=1}^{T} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\varepsilon_t^2}{2\sigma^2}\right), \tag{11}$$

where $Y_T = \{y_1, y_2, \ldots, y_T\}$, $\theta = (\pi, \beta, \gamma, \sigma, \phi)'$, $\pi = (\pi_1, \pi_2, \pi_3, \pi_4)'$ and $\phi = (\phi_1, \ldots, \phi_k)'$.

Similar to the HEGY test procedure, we compare the following hypotheses for model (8) with k assumed known a priori:

$$\begin{aligned} &\mathrm{H} &&: \pi \in \Omega, \\ &\mathrm{H}_1 &&: \pi_1 = 0,\ \{\pi_2, \pi_3, \pi_4\} \in \Omega_1, \\ &\mathrm{H}_2 &&: \pi_2 = 0,\ \{\pi_1, \pi_3, \pi_4\} \in \Omega_2, \\ &\mathrm{H}_{34} &&: \pi_3 = \pi_4 = 0,\ \{\pi_1, \pi_2\} \in \Omega_{34}. \end{aligned} \tag{12}$$

Under the first hypothesis H the π parameters are restricted to the region Ω, where all the roots of (10) are outside the unit circle, i.e., $\Omega = \{\pi \mid$ all roots are outside the unit circle$\}$.[2] The H_1 hypothesis denotes the presence of the unit root at the zero frequency, which corresponds to the restriction $\pi_1 = 0$. The remaining π parameters are restricted to the region Ω_1, which ensures that the remaining roots in (10) are outside the unit circle. The presence of the root -1 is contained in hypothesis H_2. The corresponding parameter restriction is $\pi_2 = 0$ and $\{\pi_1, \pi_3, \pi_4\} \in \Omega_2$, which restrict the π parameters such that the remaining roots in (10) are outside the unit circle. The restriction $\pi_3 = \pi_4 = 0$ in the final hypothesis H_{34} is used to test for the unit roots at the frequencies $\frac{1}{4}$ and $\frac{3}{4}$. Again the Ω_{34} region ensures that the remaining roots in the model are outside the unit circle: $\Omega_{34} = \{(\pi_1, \pi_2)' \mid \pi_1 < 0, \pi_2 < 0, \pi_1 + \pi_2 > -2\}$. The subscripts for H and Ω correspond to the null restrictions on the π parameters. The same type of notation will be used to specify the prior densities under the different hypotheses.

[2] It is implicitly assumed that all the roots of $\phi(L) = 1 - \phi_1 L - \cdots - \phi_k L^k$ are outside the unit circle.

It should be mentioned that our analysis can naturally be extended by combining, e.g., H_2 and H_{34} into a single hypothesis.

3.1. Prior specification under hypothesis H

The prior specifications under the different hypotheses are based on recent developments in Bayesian analysis of stationary AR(MA) processes, which consider the exact likelihood instead of the conditional likelihood function (see, e.g., Chib and Greenberg, 1994). We assume that the initial observations are generated according to the unconditional distribution of the process. In our framework, the assumption of stationarity and of $\phi_i = 0$ for $i = 1, ..., k$, where the latter assumption[3] is made for the moment mainly for analytical convenience, results in the following model for the initial observations:

$$\begin{pmatrix} y_{-3} - (-3)\gamma \\ y_{-2} - (-2)\gamma \\ y_{-1} - (-1)\gamma \\ y_0 - (0)\gamma \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \end{pmatrix} + \begin{pmatrix} u_{-3} \\ u_{-2} \\ u_{-1} \\ u_0 \end{pmatrix} \tag{13}$$

or in matrix notation,

$$\bar{Y}_0 = I\delta + \boldsymbol{u}, \tag{14}$$

where $\boldsymbol{u} \sim \mathrm{N}(\boldsymbol{0}, V)$, with V the unconditional covariance matrix of the stationary AR(4) process in $\bar{y}_t$ described by (10) with $\phi_i = 0, i = 1, \ldots, k$. Using the (4×4) transformation matrix L in Eq. (9), the model in (14) in terms of β is

$$\bar{Y}_0 = L^{-1}\beta + \boldsymbol{u}. \tag{15}$$

Following Schotman and van Dijk (1993), this model for the initial observations can also be interpreted as a prior density for β, i.e.,

$$p(\beta|\pi, \gamma, \sigma, Y_0) = \mathrm{N}(L\bar{Y}_0, LVL'), \tag{16}$$

where Y_0 are the initial observations. Note that the covariance matrix V is a function of the π parameters. If one or more of the π parameters approaches zero, the corresponding elements of V diverge to infinity. See, for example, the AR(1) case in (6), where V would be $\sigma^2/(1 - \rho^2)$. This reflects the fact that the β parameters are unidentified when some π parameters are zero, see (10).

For the other parameters of the model we choose the following priors:

$$\begin{aligned} p(\gamma, \sigma) &\propto \sigma^{-1} \\ p(\pi) &\propto 1 \quad \text{for } \pi \in \Omega, \end{aligned} \tag{17}$$

[3] This assumption is similar to that in Schotman and van Dijk (1993) for analysing a unit root at the zero frequency. We assume that information in the data a posteriori excludes the possibility that nonzero values for these parameters introduce another unit root.

where Ω represents the parameter space for π in which all roots are outside the unit circle. In case $\phi \neq \mathbf{0}$ we also include a flat prior on these parameters

$$p(\phi) \propto 1. \tag{18}$$

The total prior for the parameters under hypothesis H, $p(\theta) = p(\pi, \beta, \gamma, \sigma, \phi)$ is proportional to the product of (16)–(18).

3.2. Prior specification under hypothesis H_{34}

To illustrate the prior specification under one of the unit root hypotheses, we consider the hypothesis H_{34} in which $\pi_3 = \pi_4 = 0$, i.e., the seasonal unit roots i and $-i$ are present. If $\pi_3 = \pi_4 = 0$ and k is assumed to be 0, the HEGY model (10) becomes

$$\Delta_4 \bar{y}_t = \pi_1(\bar{y}_{1,t-1} - \beta_1) + \pi_2(\bar{y}_{2,t-1} - \beta_2(-1)^t) + \varepsilon_t, \tag{19}$$

which corresponds to a stationary AR(2) model for $\bar{y}_{3,t}$. Let V_{34} denote the unconditional covariance matrix of this AR(2) model, which is a function of π_1 and π_2 only. The prior for β_1, β_2 (note that β_3 and β_4 are not present in this model) can now be written in the form:[4]

$$p_{34}(\beta_1, \beta_2 \mid \pi_1, \pi_2, \gamma, \sigma, Y_0) = \mathrm{N}\left(\begin{pmatrix} \bar{y}_{1,0} \\ -\bar{y}_{2,0} \end{pmatrix}, L_{34} V_{34} L_{34}'\right), \tag{20}$$

where L_{34} is the (2×2) left upper corner submatrix of the (4×4) transformation matrix L. Priors for the remaining parameters are given by

$$\begin{aligned} p_{34}(\phi, \gamma, \sigma) &\propto \sigma^{-1} \\ p_{34}(\pi_1, \pi_2) &\propto \quad 1 \text{ for } \{\pi_1, \pi_2\} \in \Omega_{34}, \end{aligned} \tag{21}$$

where $\Omega_{34} = \{(\pi_1, \pi_2)' \mid \pi_1 < 0, \pi_2 < 0, \pi_1 + \pi_2 > -2\}$, which corresponds to the parameter space where the two remaining roots are outside the unit circle. The total prior for the parameters under the hypothesis H_{34}, $p_{34}(\theta_{34}) = p_{34}(\pi_1, \pi_2, \beta_1, \beta_2, \gamma, \sigma, \phi)$ is proportional to the product of (20), (21) and (18) in case $k \neq 0$.

3.3. Prior specification under hypotheses H_1 and H_2

The prior specifications for the two remaining hypotheses are similar to the prior specification under the H_{34} hypothesis. Under the restriction $\pi_1 = 0$ ($\pi_2 = 0$) the HEGY model (10) becomes an AR(3) for $\bar{y}_{1,t}$ ($\bar{y}_{2,t}$) in which β_1 (β_2) drops out. The unconditional covariance matrix of this AR(3) model is used to specify the prior on the remaining β parameters like in (20). The prior densities for ϕ,

[4] The subscript corresponds to the restriction $\pi_3 = \pi_4 = 0$.

γ and σ are the same as in (21) and for the remaining π parameters flat on the stationary region Ω_1 (Ω_2).

To compare the different hypotheses in (12), we compute posterior odds. Assuming that the hypotheses under consideration are, a priori, equally likely, the posterior odds ratio equals the Bayes factor (see, e.g., Zellner, 1971). The Bayes factor to compare hypothesis H_{34} with H is given by

$$K_{34} = \frac{\Pr[H_{34} \mid Y_T, Y_0]}{\Pr[H \mid Y_T, Y_0]} = \frac{\int p_{34}(\theta_{34})\mathcal{L}(Y_T \mid Y_0; \theta_{34})\,d\theta_{34}}{\int p(\theta)\mathcal{L}(Y_T \mid Y_0; \theta)\,d\theta}, \tag{22}$$

where the subscipt on K refers to the H_{34} hypothesis. Similar Bayes factors can be defined for testing the presence of the other unit roots, i.e., for the hypotheses H_1 and H_2. We note that the posterior density may have negligible probability mass on some regions of the parameter space. Since the numerator and the denominator of the Bayes factor can be interpreted as average heights, this can lead to misleading conclusions. To avoid this problem, we a posteriori restrict the parameter regions Ω to the 99% highest posterior density (HPD) region (see also Schotman and van Dijk, 1993). Computation of Bayes factors has received considerable attention in the recent literature. In our analysis we apply the approach of Chib (1995), where the output of the Gibbs sampler (see Section 4 below) is used to compute the marginal likelihoods needed in (22).

3.4. Seasonal mean shifts

A structural break in the seasonal pattern can be incorporated by replacing δ_s in (5) by $\delta_s + \delta_s^*[I_{t\geqslant\tau}]$. Since we assume that the breakpoint τ is unknown, we treat it as an extra parameter with the following noninformative prior density:

$$p(\tau) = \frac{1}{T-4}, \quad \tau = 1, \ldots, T-4. \tag{23}$$

Concerning the mean shift parameters δ_s^*, we note that, unlike the seasonal mean parameters themselves, their interpretation is independent of the number of unit roots. Therefore, the following noninformative prior is used:

$$p(\delta_1^*, \delta_2^*, \delta_3^*, \delta_4^*) \propto 1. \tag{24}$$

Extending the earlier analysis by including these breakpoint parameters is straightforward.

It should be mentioned that our modelling of seasonal mean shifts implies a sudden adjustment in the level of the series. The transformed parameters β_s, see (9), change only stepwise to their new values. Hence, in the terminology of Perron (1989), we only consider the 'additive outlier' framework.

3.5. Remarks

An alternative Bayesian approach to investigate seasonal unit roots with the possibility of a structural break can be found in Koop and Pitarakis (1992). Their analysis differs from ours in several respects. First, they consider the linear parameterization of the HEGY model given in Eq. (2) combined with a flat prior. Hence, they follow the Bayesian unit root analysis of, e.g., DeJong and Whiteman (1991). Furthermore, they consider an informative prior suggested by Zellner and Siow (1980). The main advantage of this approach is that it is easily implemented: posterior odds and HPD regions can be obtained analytically. However, as we noted earlier, the intercept and trend parameters in model (2) are difficult to interpret since their interpretation changes with the hypothesis under consideration. This also applies to the interpretation of the seasonal mean shift parameters. Second, Koop and Pitarakis (1992) do not explicitly add the breakpoint τ as an extra parameter. Instead, posterior odds are computed to test for no structural change ($H_0 : \delta_i^* = 0,\ i = 1,2,3,4$) for all possible values of τ. They recommend pretesting for structural change before applying the HEGY test procedure. Instead of pretesting, we propose to compute Bayes factors to test for (seasonal) roots both under the hypothesis of structural change and no structural change. The third and final difference is that in our framework we obtain the posterior density of the breakpoint parameter τ. This may enable an interpretation of the timing of the shifts.

4. Sampling of the posterior distribution

Recently, Markov chain Monte Carlo (MCMC) sampling techniques have proved to be a useful tool to analyse posterior distributions. Two MCMC sampling algorithms will be used to evaluate the posterior distributions which result from combining the prior and likelihood function described in the previous section: the Gibbs sampling and the Metropolis–Hastings algorithm. In the appendix we give a short description of both techniques.

Following Chib and Greenberg (1994), it is straightforward to show that the full conditional distributions of our parameters of interest β, γ, ϕ and δ^* are normal and that σ has an inverted gamma-2 distribution conditional on the other parameters. The full conditional density of the π parameters, under the assumption of stationarity, is of the form

$$P(\pi|\tau, \theta\backslash\{\pi\}) \propto \Psi(\pi) \exp\left(-\tfrac{1}{2}(\pi - \hat{\pi})' \Sigma_{\hat{\pi}}^{-1}(\pi - \hat{\pi})\right) [I_\Omega], \tag{25}$$

where $\Psi(\pi)$ is the prior defined on β in (20) and $[I_\Omega]$ is the indicator function defined on the stationary region Ω. Like Chib and Greenberg (1994), we recognize that the second part of (25) corresponds to the kernel of a (truncated) normal distribution. The mean and variance of this normal distribution are given by $\hat{\pi}$

and $\Sigma_{\hat{\pi}}$, which correspond to the OLS estimate of π in (10) and the corresponding estimated covariance matrix, respectively.[5] Let $\tilde{\pi}$ denote a draw from this distribution at iteration i of the Gibbs sampler. Next, we apply the Metropolis–Hastings step with acceptance probability $\min\left(\Psi(\tilde{\pi})/\Psi(\pi^{i-1}), 1\right)$, with π^{i-1} denoting the draw from the previous iteration, to give π^i.

To draw from the conditional distribution of the breakpoint parameter τ, we note that this parameter only takes discrete values. Therefore, the distribution function of τ given the parameter vector θ can easily be computed. Using this distribution function a draw of the breakpoint parameter can be obtained using an inversion method. The conditional distribution of δ^* is normal.

5. Applications

To demonstrate the Bayesian analysis of seasonal unit roots and seasonal mean shifts, we consider simulated and empirical quarterly time series.

5.1. Simulated series

We start with four simulated data generating processes:

$$\begin{aligned}
&\text{DGP I} &&: \; y_t = \sum_{s=1}^{4} \alpha_s D_{st} + \varepsilon_t,\\
&\text{DGP II} &&: \; \Delta_4 y_t = \mu + \varepsilon_t,\\
&\text{DGP III} &&: \; \Delta_1 y_t = \sum_{s=1}^{4} \psi_s D_{st} + \varepsilon_t,\\
&\text{DGP IV} &&: \; \Delta_1 y_t = \sum_{s=1}^{4} \psi_s D_{st} + \psi_s^* D_{st}[I_{t\geqslant\tau}] + \varepsilon_t,
\end{aligned} \tag{26}$$

where $\varepsilon_t \sim \mathrm{IN}(0, 0.5)$, $t = 1, 2, \ldots, 164$, $\alpha = (1, 2, 3, 4)'$, $\mu = 1$, $\psi = (1, -2, -1, 3)'$ and $\psi^* = (-4, 5, -3, 2)'$. Notice that we do not compare DGPs I–IV relative to each other, but merely we investigate the effect of (not) allowing for seasonal mean shifts when analysing seasonal unit roots. The first 44 observations are used as starting values and deleted from the sample, resulting in samples of 120 observations. The breakpoint τ is fixed at observation 40 of the 120 observations. The first DGP contains no unit roots, the second DGP contains one nonseasonal and three seasonal unit roots. The last two DGPs only contain a unit root at the zero frequency. The series are analysed using the Bayesian approach suggested in the previous sections. Posterior outcomes are based on 21,000 iterations of the MCMC sampler, discarding the first 1000 drawings as burn-in period. As starting values we take the OLS estimates of the parameters. The first two DGPs are analysed without allowing for the possibility of a structural break in the seasonal

[5] Note that, conditional on the other parameters, Eq. (10) is linear in π.

Table 1
Posterior means and standard errors between parentheses for the π parameters, all roots in stationary region, and posterior odds outcomes for the simulated series

	DGP I	DGP II	DGP III		DGP IV	
			No break	Break	No break	Break
lags k	0	0	0	0	0	0
π_1	−0.179	−0.048	−0.008	−0.008	−0.010	−0.008
	(0.049)	(0.031)	(0.006)	(0.006)	(0.007)	(0.006)
π_2	−0.201	−0.023	−0.650	−0.638	−0.042	−0.642
	(0.055)	(0.017)	(0.093)	(0.108)	(0.026)	(0.099)
π_3	−0.441	−0.031	−0.349	−0.341	−0.429	−0.357
	(0.073)	(0.020)	(0.075)	(0.079)	(0.078)	(0.078)
π_4	−0.020	0.014	−0.452	−0.468	−0.382	−0.458
	(0.069)	(0.025)	(0.076)	(0.086)	(0.079)	(0.078)
			Posterior odds ratios			
K_1	0.01	1.42	2.07	2.17	2.52	2.00
K_2	0.00	2.19	0.00	0.00	1.17	0.00
K_{34}	0.00	3.78	0.00	0.00	0.00	0.00

Lags denote the number k of lagged $\Delta_4 y_t$ included in the model. Results are based on 20,000 iterations of the MCMC sampler. K denotes the posterior odds ratio, where the subscripts correspond to the π parameters (for example, K_{34} represents the odds ratio for testing H_{34}: $\pi_3 = \pi_4 = 0$ against H: all roots in the stationary region). An odds ratio exceeding one implies that the null hypothesis is a posteriori more likely than the alternative hypothesis.

pattern. The final two DGPs are analysed with and without the possibility of a structural break. Note that DGP III does not have structural mean shifts. Table 1 shows the posterior means and standard errors of the π parameters together with the posterior odds ratio for the unit root tests. It should be mentioned here that because our DGPs in (26) do not include additional lags, we expect to find no evidence for seasonal unit roots for DGPs I and III, and hence the relevant odds ratios will be about zero.

For the first DGP we see that the means of the marginal posterior of π_1, π_2 and π_3 differ more than two standard errors from zero. However, since these posterior densities are truncated, care must be exercised in interpreting the standard errors. In Fig. 1 we depict the marginal posterior densities of the π parameters. The modes of the distributions are far away from zero, except for π_4. The latter exception corresponds to the results for the π_4 parameters discussed in Hylleberg et al. (1990). The posterior odds for the joint hypothesis $\pi_3 = \pi_4 = 0$ is clearly smaller than one, see Table 1. The same is true for the hypothesis of the presence of the roots 1 and −1.

The marginal posterior densities of the π parameters for the second DGP are shown in Fig. 2. The modes of the marginal densities are near zero. The posterior

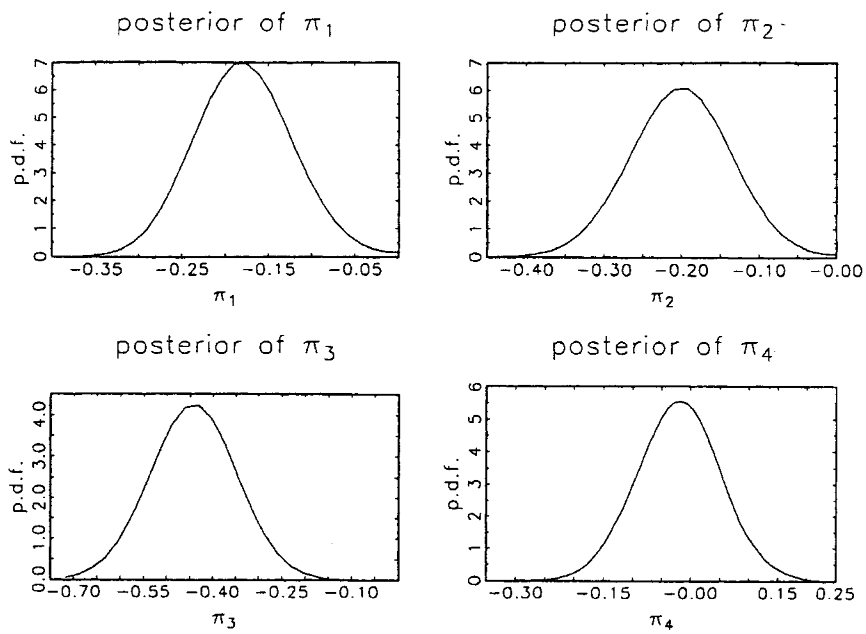

Fig. 1. Marginal posterior of the π's for DGP I, all roots in stationary region.

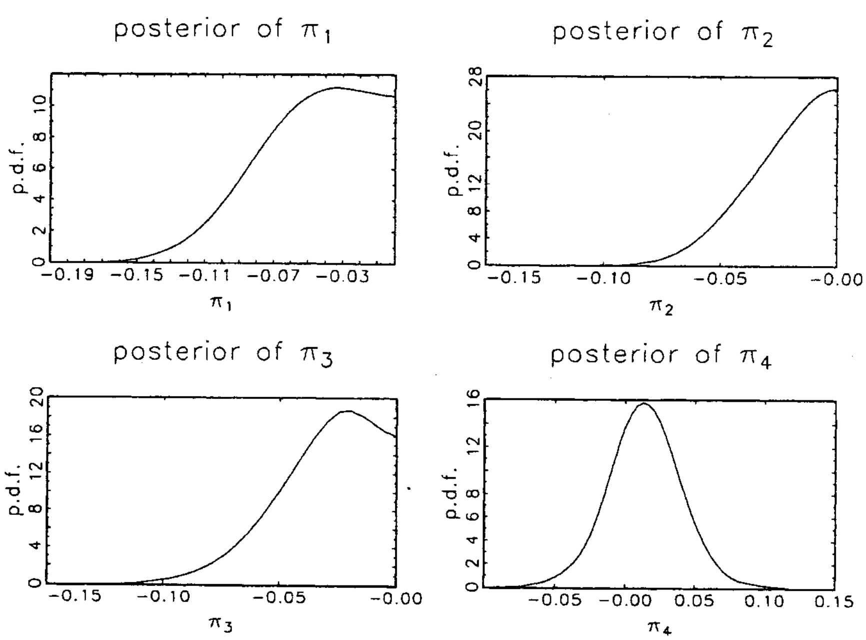

Fig. 2. Marginal posterior of the π's for DGP II, all roots in stationary region.

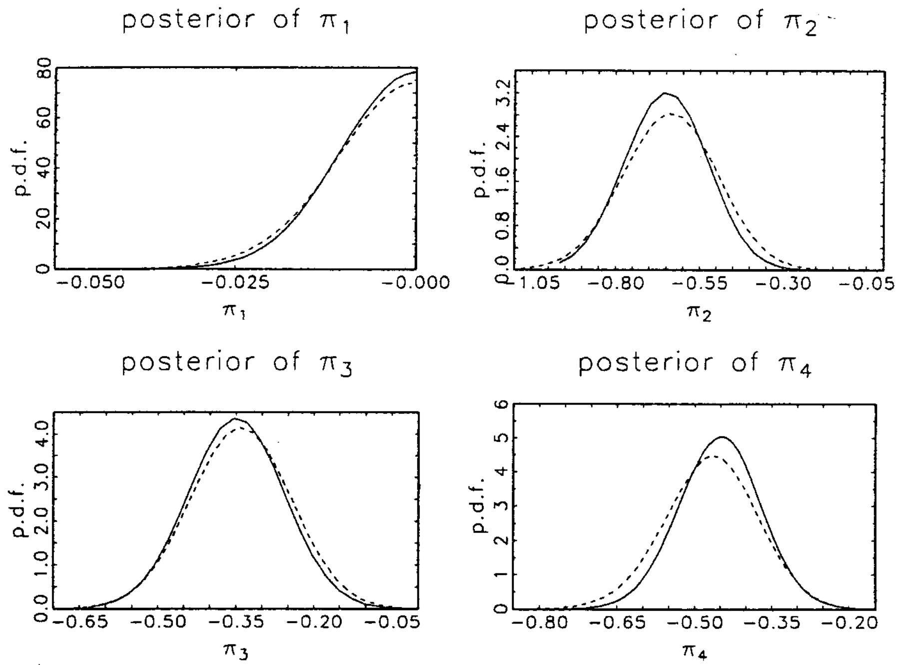

Fig. 3. Marginal posterior of the π's for DGP III, all roots in stationary region: (–): without mean shifts; (- -): with mean shifts.

odds ratio in Table 1 correctly indicates the presence of the nonseasonal and the three seasonal unit roots.

For DGP III only the posterior odds ratio for the presence of the nonseasonal unit root exceeds one. The modes of the marginal posterior densities of the π parameters are far away from zero except for π_1, see Fig. 3. As Fig. 3 also shows, the posterior results remain virtually the same when we allow for seasonal mean shifts, see also column 5 of Table 1. The marginal posterior of the parameter τ is roughly uniform on its domain with two peaks near the borders. Therefore, we can conclude that allowing for possible seasonal mean shifts, when no such shift is present, does not seem to influence the conclusions about the presence of unit roots.

The outcomes for DGP IV in column 6 of Table 1 show that the structural mean shifts can alter the conclusions about the presence of unit roots. In case we do not include the possibility of a structural mean shift, the posterior odds ratio for the hypothesis $\pi_2 = 0$ is larger than one, see Table 1. The inclusion of the structural mean shifts results in a shift of the marginal posterior density of π_2 to the left, and this leads to a posterior odds strongly smaller than one, see also Fig. 4 and the last column of Table 1. Changes in the posteriors for the other π

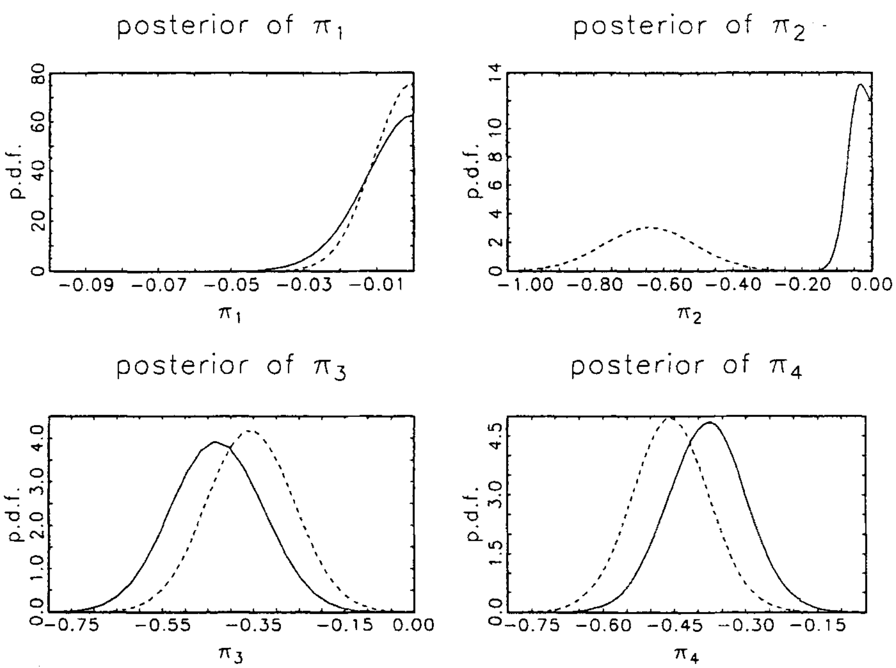

Fig. 4. Marginal posterior of the π's for DGP IV, all roots in stationary region: (–): without mean shifts; (- -): with mean shifts.

parameter are relatively small if the seasonal mean shifts are included. This result for only one experiment seems to correspond with our (unreported) findings in simulations of the HEGY tests using classical methods. The marginal posterior of the parameter τ has 57% of its probability mass at observation 39 and 43% at observation 40. This apparent precision is not surprising since we have imposed a substantial structural mean shift in the DGP. It should be mentioned here that these simulations only serve illustrative purposes. When we would reduce the size of the mean shifts for DGP IV, we would find less evidence in favour of H_2. On the other hand, when we would enlarge DGP IV with lags of $\Delta_1 y_t$, we would find larger values of K_2 and K_{34}.

In summary, the posterior outcomes of the four simulated DGPs seem to indicate the practical usefulness of our Bayesian approach. However, since only a few simulated data sets have been considered, we again stress that no general conclusion can be drawn about the performance of the approach.

5.2. Three consumption series

We now apply our Bayesian analysis of seasonal unit roots with and without structural mean shifts to three quarterly observed consumption series. These series

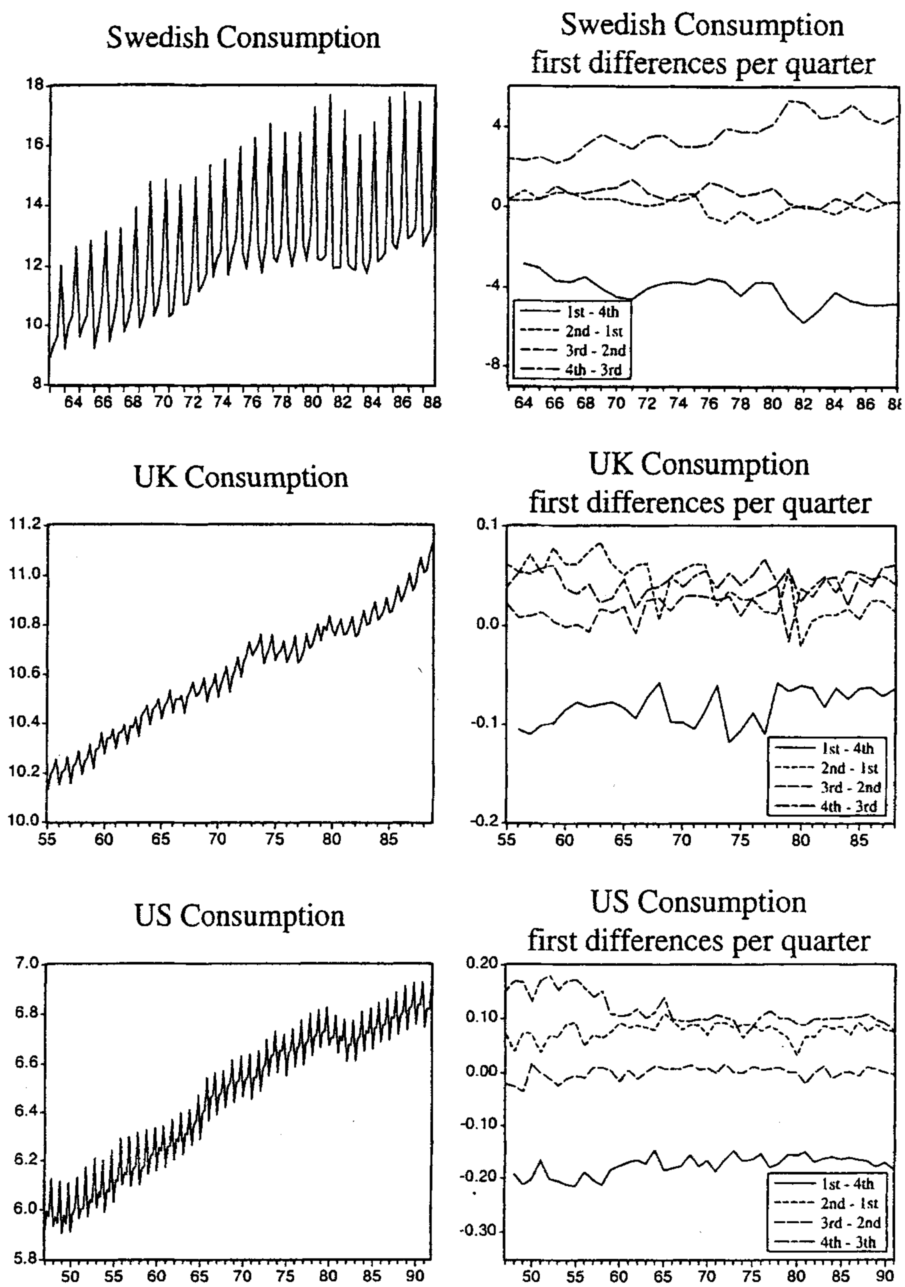

Fig. 5. Plots of total consumption of Sweden and the UK and nondurable consumption of the US.

Table 2
Posterior means and standard errors between parentheses for the π parameters, all roots in stationary region, and posterior odds outcomes for total consumption of Sweden and the UK and nondurable consumption of the US

	Total cons. Sweden		Total cons. UK		Nondur. cons. US	
	No break	Break	No break	Break	No break	Break
Lags k	8	8	8	8	8	8
π_1	−0.024	−0.027	−0.015	−0.017	−0.003	−0.003
	(0.018)	(0.021)	(0.011)	(0.013)	(0.003)	(0.003)
π_2	−0.057	−0.313	−0.069	−0.137	−0.078	−0.281
	(0.039)	(0.146)	(0.050)	(0.099)	(0.046)	(0.112)
π_3	−0.081	−0.139	−0.167	−0.229	−0.119	−0.416
	(0.046)	(0.080)	(0.082)	(0.128)	(0.055)	(0.138)
π_4	−0.069	−0.146	−0.113	−0.265	−0.050	−0.387
	(0.056)	(0.087)	(0.090)	(0.139)	(0.056)	(0.139)
			Posterior odds ratios			
K_1	2.02	2.14	1.94	2.66	2.54	2.78
K_2	1.82	0.30	2.01	2.29	1.21	0.29
K_{34}	1.95	0.02	0.39	1.43	0.21	0.00

Lags denote the number k of lagged $\Delta_4 y_t$ included in the model. Results are based on 20,000 iterations of the MCMC sampler. K denotes the posterior odds ratio, where the subscripts correspond to the π parameters (for example, K_{34} represents the odds ratio for testing H_{34}: $\pi_3 = \pi_4 = 0$ against H: all roots in the stationary region). An odds ratio exceeding one implies that the null hypothesis is a posteriori more likely than the alternative hypothesis.

are the log of real total consumption of Sweden, 1963.1–1988.4, the log of real total consumption of the UK, 1955.1–1988.4 and the log of real nondurable consumption in the US, 1947.1–1991.4.

Fig. 5 shows plots of the data. The graphs in the first column of Fig. 5 show the series. A plot of the first differences of the series split up in a series for each quarter is given in the second column. We observe from the last column of Fig. 5 that for Swedish consumption there seems to be a structural break in the beginning of 1980s, while for the US series we notice a seasonal mean shift in the first differences in the end of 1950s. The UK series, however, does not display visually obvious mean shifts.

Table 2 shows the posterior results of our Bayesian seasonal unit root analysis in the presence of structural seasonal mean shifts. A lag order of 8 for the fourth differences has been chosen for all three series. [6] In case of no break we conclude

[6] The lag order may affect conclusions concerning the unit root parameters. However, incorporating a variable lag order in our framework is outside the scope of the present paper. To get some indication of the robustness of our results with respect to the lag order, we also analysed the consumption series using $k = 4$. This only marginally affected the results and we delete details to save space.

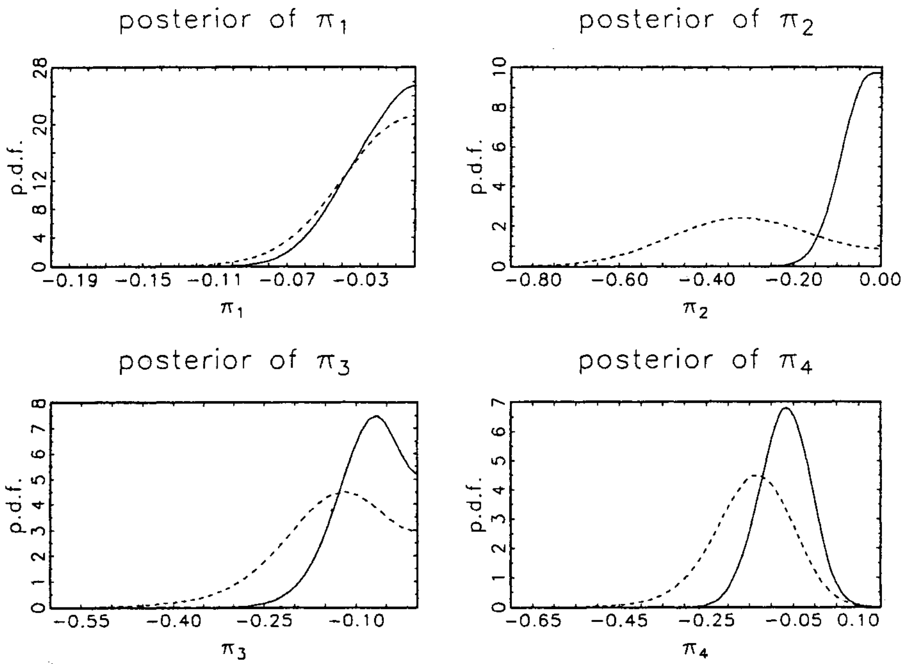

Fig. 6. Marginal posterior for total consumption of Sweden, all roots in stationary region: (–): without mean shifts; (- -): with mean shifts.

that total consumption of Sweden contains the nonseasonal and three seasonal unit roots. If we however include the possibility of structural seasonal mean shifts we observe that the marginal posterior densities of π_2, π_3 and π_4 shift to the left, see Fig. 6. The posterior odds ratio K_2 drops from 1.82 to 0.30 which results in favouring the absence of the unit root at the frequency $\frac{1}{2}$. Also, the posterior odds ratio K_{34} drops from 1.95 to 0.02, which provides further evidence that seasonal mean shifts can explain the nonrejection of seasonal unit roots when no break is included in the model. The mode of the marginal posterior of τ is in 1979.3, where we find more than 70% of the probability mass.

The posterior outcomes of total consumption of the UK indicate that the series contains the roots, 1 and -1 if we do not allow for a seasonal mean shift. Allowing a possible structural break in the series alters these results; see Table 2. In this case, support is also found for the presence of the complex roots i and $-$i. Figs. 7 show the marginal posterior distributions of the π parameters. The marginal posterior of the τ parameter displays two low peaks, the first in 1967.1, where we find 12% of the probability mass and the second in 1978.2 with 49.6% of the probability mass.

The posterior odds ratios K_1 and K_2 for the US nondurable consumption series exceed one, indicating the presence of one nonseasonal and one seasonal unit root.

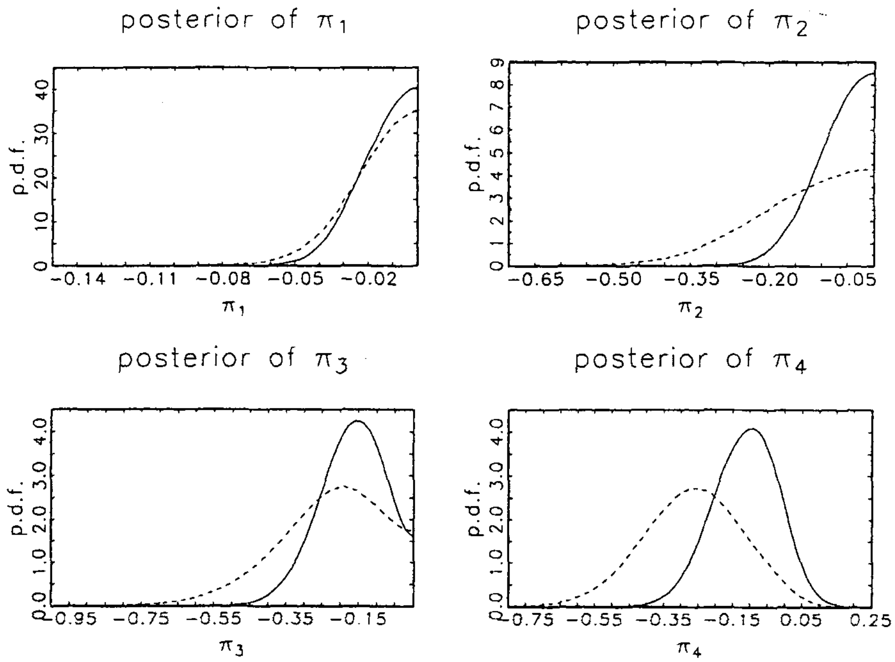

Fig. 7. Marginal posterior for total consumption of the UK, all roots in stationary region: (—): without mean shifts; (- -): with mean shifts.

From Fig. 8 it can be seen that when we allow for a possible mean shift, the posterior distributions of π_2, π_3 and π_4 move away from zero. Indeed, from the last column of Table 2 we observe that the posterior odds ratio for the presence of the root -1 drops below one. The marginal posterior density of the break parameter τ displays a clear peak around 1957.3 and 1957.4.

In summary, allowing for possible seasonal mean shifts can alter the empirical evidence in favour of seasonal unit roots. For the consumption series of Sweden and the US, the Bayesian analysis yields a rather precise estimate of the timing of seasonal mean shifts. When we allow for these shifts, the evidence for seasonal unit roots disappears. For the UK consumption series, there is no clear evidence for a deterministic seasonal mean shift. However, the inclusion of such a shift results in more evidence for seasonal unit roots.

6. Conclusions

In this paper we have presented a Bayesian approach to test for seasonal unit roots. Our analysis is based on a reparameterization of the model of Hylleberg et al. (1990), which yields parameters that appear to have a natural interpretation.

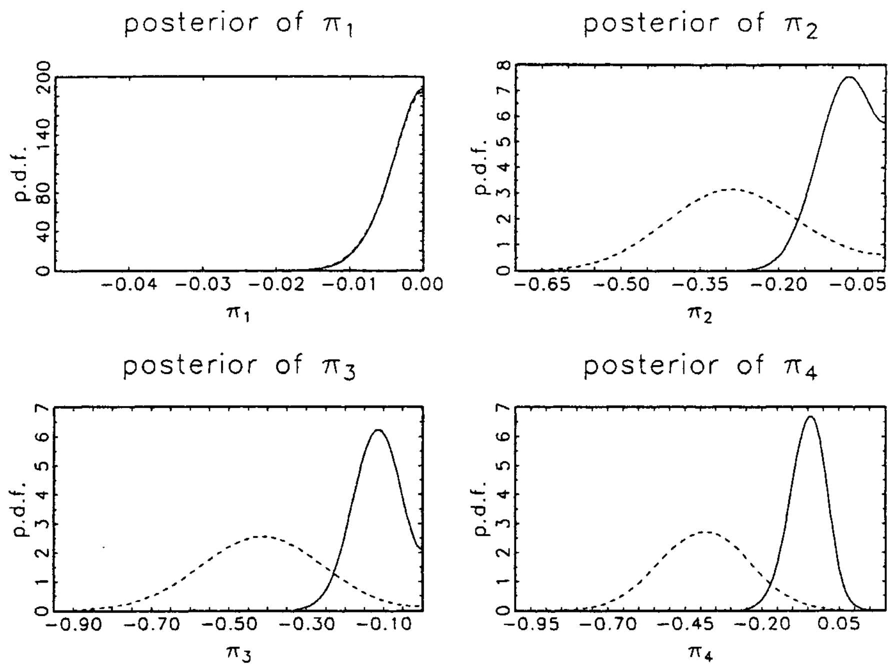

Fig. 8. Marginal posterior for nondurable consumption of the US, all roots in stationary region: (—): without mean shifts; (- -): with mean shifts.

Further, we extend the HEGY analysis by allowing for possible deterministic seasonal mean shifts.

Simulation exercises demonstrate the usefulness of our approach. In particular, it is shown that neglecting seasonal mean shifts may incorrectly suggest the presence of seasonal unit roots, and that the inclusion of an unknown breakpoint yields more appropriate results. Application of our method to a model that does not incorporate seasonal mean shifts for three consumption series results in a nonseasonal and one or more seasonal unit roots in the series. However, when we allow for seasonal mean shifts, we can 'reject' the hypothesis of the seasonal unit roots for two of the three series. Apparently, these seasonal mean shifts adequately explain the changing seasonal fluctuations in these two series.

Our focus has been to demonstrate that ignoring seasonal mean shifts may incorrectly yield evidence in favour of seasonal unit roots. Practical issues, like simultaneously determining the lag order of the process, have received little attention. However, choice of the lag order may also affect unit root inference. Therefore, this issue should be put on the research agenda.

Our analysis can also easily be extended to investigate nonseasonal and seasonal unit roots in biannual or monthly data. Furthermore, extensions to seasonally varying variances and/or t-distributed errors is straightforward. Finally, the analysis of multiple seasonal mean shifts can be based on similar methods as described in this paper.

Appendix. Markov chain Monte Carlo

In this appendix we give a short description of the two sampling techniques we use in our empirical analysis, i.e., the Gibbs sampling and the Metropolis – Hastings sampling technique.

To describe the Gibbs sampler, let x be a random vector which can be divided in d blocks $(x_1,...,x_j,...,x_d)$, Also, let $f(x_j|x_{-j})$ denote the distribution of x_j conditional on the other random variables $x_{-j} = x\backslash x_j$. The sampling method can be described as follows:

Step 1: Specify starting values $x^0 = (x_1^0,...,x_d^0)$ and set $i = 0$.
Step 2: Simulate

$$\begin{aligned}
&x_1^{i+1} \text{ from } f(x_1 \mid x_2^i,\ x_3^i,...,x_d^i),\\
&x_2^{i+1} \text{ from } f(x_2 \mid x_1^{i+1},\ x_3^i,...,x_d^i),\\
&x_3^{i+1} \text{ from } f(x_3 \mid x_1^{i+1},\ x_2^{i+1},x_4^i,...,x_d^i),\\
&\qquad\vdots\\
&x_d^{i+1} \text{ from } f(x_d \mid x_1^{i+1},\ x_2^{i+1},...,x_{d-1}^{i+1}),
\end{aligned}$$

Step 3: set $i = i + 1$, and go to step 2.

This iterative scheme generates a Markov chain, which converges under mild conditions, see, e.g., Smith and Roberts (1993) and Tierney (1994). After the chain has converged, say at H iterations, the simulated values $\{x^i, i \geqslant H\}$ can be used as a sample from the joint distribution $f(x)$ in order to compute posterior densities and expectations.

The second algorithm was introduced by Metropolis et al. (1953) and has been adapted for statistical problems by Hastings (1970). Let $f(x)$ be the target density and let $g(x,y)$ be a transition probability function. The algorithm works as follows:

Step 1: Specify starting values $x^0 = (x_1^0,...,x_d^0)$ and set $i = 0$.
Step 2: Simulate y from $g(x^i, y)$
Step 3: Define

$$\alpha(x^i, y) = \begin{cases} \min\left\{\dfrac{f(y)g(y,x^i)}{f(x^i)g(x^i,y)}, 1\right\}, & f(x^i)g(x^i,y) > 0,\\ 1, & f(x^i)g(x^i,y) = 0. \end{cases}$$

y is accepted with probability $\alpha(x^i, y)$ and $x^{i+1} = y$ and rejected with probability $1 - \alpha(x^i, y)$ and $x^{i+1} = x^i$.

Step 4: Set $i = i + 1$, and go to step 2.

Different choices for the transition probability function result in different specific forms of the algorithm. For example, if $g(x^i, y) = g(y, x^i)$ the acceptance probability simplifies to $\alpha(x^i, y) = \min\{f(y)/f(x^i), 1\}$. This describes the original Metropolis algorithm. If $g(x^i, y) = g(y)$, we get $\alpha(x^i, y) = \min\{w(y)/w(x^i), 1\}$, where $w(x^i)$ is defined by $w(x^i) := f(x^i)/g(x^i)$ (which can be interpreted as importance weights).

References

Chib, S., 1995, Marginal likelihood from the Gibbs output, Journal of the American Statistical Association 90, 1313–1321.

Chib, S. and E. Greenberg, 1994, Bayes inference in regression models with ARMA(p,q) errors, Journal of Econometrics 64, 183–206.

DeJong, D.N. and C.H. Whiteman, 1991, Reconsidering trends and random walks in macroeconomic time series, Journal of Monetary Economics 28, 221–254.

Engle, R.F., C.W.J. Granger, S. Hylleberg, and H.S. Lee, 1993, Seasonal cointegration: the Japanese consumption function. Journal of Econometrics 55, 275–298.

Ghysels, E., 1994, On the economics and econometrics of seasonality, in: C.A. Sims, ed., Advances in econometrics, sixth world conference of the econometric society (Cambridge University Press, Cambridge), 257–316.

Hastings, W.K., 1970, Monte Carlo sampling methods using Markov chain and their applications, Biometrika 57, 97–109.

Hylleberg, S., R.F. Engle, C.W.J. Granger, and B. Yoo, 1990, Seasonal integration and cointegration. Journal of Econometrics 44, 215–238.

Hylleberg, S., C. Jørgensen and N.K. Sørensen, 1993, Seasonality in macroeconomic time series, Empirical Economics 18, 321–335.

Koop, G. and J-Y. Pitarakis, 1992, A Bayesian analysis of seasonality in economic time series, Working paper (Department of Economics, University of Toronto).

Metropolis, N., A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller, 1953, Equations of state calculations by fast computing machines, Journal of Chemical Physics 21, 1087–1092.

Paap, R., P.H. Franses and H. Hoek, 1997, Mean shifts, unit roots and forecasting seasonal time series, Internation Journal of Forecasting (forthcoming).

Perron, P., 1989, The great crash, the oil price shock, and the unit root hypothesis, Econometrica 57, 1361–1401.

Perron, P. and T.J. Vogelsang, 1992, Nonstationarity and level shifts with an application to purchasing power parity, Journal of Business and Economic Statistics 10, 301–320.

Schotman, P.C. and H.K. van Dijk, 1993, Posterior analysis of possibly integrated time series with an application to real GNP, in: P. Caines et al., eds., New directions in time series analysis, part II (Springer, Berlin) 341–361.

Smith, A.F.M. and G.O. Roberts, 1993, Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods, Journal of the Royal Statistical Society Series B 55, 3–23.

Tierney, L., 1994, Markov chains for exploring posterior distributions, Annals of Statistics 22, 1701–1762.

Zellner, A., 1971, An introduction to Bayesian inference in econometrics, (Wiley, New York).

Zellner, A. and A. Siow, 1980, Postèrior odds for selected regression hypotheses, In: J. Bernardo et al., eds., Bayesian statistics, Proceedings of the First International Meeting (Valencia, Spain) May 1979, 585–603.