Department of Economics

A Bayesian Analysis of Unit Roots and Structural Breaks in the Level and the Error Variance of Autoregressive Models

Loukia Meligkotsidou, Elias Tzavalis and Ioannis D. Vrontos

Working Paper No. 514 July 2004 ISSN 1473-0278

Queen Mary
University of London

A Bayesian analysis of unit roots and structural breaks in the level and the error variance of autoregressive models

Loukia Meligkotsidou* Elias Tzavalis† Ioannis D. Vrontos‡

Abstract

In this paper, a Bayesian approach is suggested to compare unit root models with stationary models when both the level and the error variance are subject to structural changes (known as breaks) of an unknown date. The paper utilizes analytic and Monte Carlo integration techniques for calculating the marginal likelihood of the models under consideration, in order to compute the posterior model probabilities. The performance of the method is assessed by simulation experiments. Some empirical applications of the method are conducted with the aim to investigate if it can detect structural breaks in financial series, with changes in the error variance.

JEL Classification: C11, C22, G10

Keywords: Bayesian inference, model comparison, autoregressive models, unit roots, structural breaks

*Department of Mathematics and Statistics, Lancaster University, Lancaster LA1 4YF, UK. *Email*: l.meligotsidou@lancaster.ac.uk

†Department of Economics, Queen Mary, University of London, London E1 4NS, UK. *Email*: E.Tzavalis@qmul.ac.uk

‡Department of Statistics, Athens University of Economics and Business, Pattision 76, Athens 134 76, Greece. *Email*: vrontos@aueb.gr

1 Introduction

The unit root hypothesis has received a lot of attention in time series literature after the findings of Nelson and Plosser (1982), that shocks have permanent effects on the level of most of economic series. Although earlier evidence of units roots in the literature has been considerably challenged by the development of unit root testing procedures which allow for structural changes (known as breaks) in the mean of the series [see Perron (1989), Christiano (1992), Perron and Vogelsang (1992), Zivot and Andrews (1992), and Lumsdaine and Papell (1997), *inter alia*], there are economic series (especially, financial) which strongly favour the unit root hypothesis despite the lack of economic intuition or the occurrence of structural changes in the economy. Typical examples of such series are the nominal interest rates [see Hall, Anderson and Granger (1992), and Tzavalis and Wickens (1997)] and real exchange rates [see Papell (1997, 2002), for a survey]. This evidence may be attributed to the fact that most of the testing procedures employed to test the unit root hypothesis do not allow for changes in the error variance, which are apparent in financial series.

The main difficulty for developing testing procedures allowing for changes in the variance within the classical framework stems from the fact that, in general, it is difficult to identify some of the nuisance parameters (including the error variance with a break-point) under the null [e.g. Garcia (1998), and Kim and Nelson (1999)]. This problem becomes more severe when the break-point is treated as unknown. The above difficulty can be overcome within the Bayesian model comparison framework. Furthermore, within this framework, all the hypotheses under consideration (corresponding to either unit root or stationary models) are treated symmetrically. This can increase the power of the Bayesian procedure to detect the correct model generating the data. The main issue in the Bayesian approach is is to detect which out of a number of competing hypotheses is the most likely to be consistent with the data. This can be done by calculating the marginal likelihood of the data under each model by integrating the parameters of the model out of the posterior density, in order to compute the posterior model probabilities. Finally, another advantage of the Bayesian approach is that, in this framework, the identification of a break-point in the series is considered as a part of the model selection problem.

Despite the plethora of studies considering Bayesian methods to compare unit root models with stationary models, dating from Sims (1988)[1], only recently Marriott and Newbold (2000) have shown that these methods can be succesfully used to distinguish unit root models from stationary models with a break-point. However, this study does not consider breaks in the error variance of the models, which is the focus of our paper.[2] Ignoring a break in the error variance of the models may increase the strength of evidence for a unit root, independently on whether the models ignore breaks in the level of the series, or not. The Bayesian approach that we suggest in this paper is based on analytic and Monte Carlo integration techniques for the calculation of the marginal likelihood of the data under unit root

[1] See also Schotman and van Dijk (1991a), DeJong and Whiteman (1989,1991a,b), Sims and Uhlig (1991) and Koop (1992, 1994).

[2] Bayesian methods diagnosing the presence of breaks in the mean or error variance have been developed by Wang and Zivot (2000). But, these authors have not considered unit roots models.

and stationary models allowing for breaks. To appraise the performance of our method, we conduct a simulation study considering for values of the autoregressive parameters closed to unity. As empirical applications of our method, we investigate if unit root evidence on US nominal Treasury bill rates and some European real exchange rates against US dollar can be challenged by the presence of breaks in the mean or the error variance of the series.

The remaining of the paper is organised as follows. In Section 2, we give all the necessary notation and introduce the models of interest. In Sections 3 and 4, we discuss the Bayesian approach to inference and we show how to calculate the marginal likelihood functions of the models under consideration, respectively. In Section 5 we report the simulation results and in Section 6 we present the results of our empirical applications. Finally, Section 6 concludes the paper.

2 The autoregressive model with a structural break in parameters

Consider a non-linear autoregressive model of order one allowing for a break in the intercept, the autoregressive coefficient or the error variance at an unknown time point (referred to as break-point) T_0

$$y_t = (\gamma_1 + \delta_t)(1 - \phi_{s_t}) + \phi_{s_t} y_{t-1} + \sigma_{s_t}\epsilon_t, \quad t = 1, \ldots\ldots, T \tag{1}$$

where y_t, $t = 1, ..., T$ is a sample of T consecutive observations, $\{s_t\}$, $t = 1, \ldots..., T$, is a binary process with $s_t = 1$, if $t \leq T_0$, and $s_t = 2$, otherwise, indicating the subsample (segment) to which each observation y_t is assigned, $\delta_t = \begin{cases} 0, & \text{if} \quad t \leq T_0 \\ \delta, & \text{if} \quad t > T_0 \end{cases}$ is a dummy variable determining the change in the mean of y_t and ϵ_t are assumed to be independent and identically normally distributed. The initial observation of the series $\{y_t\}$, y_0, is assumed to be known. This is a standard assumption in unit roots or structural breaks literature, where deviations from the initial conditions of the series are of non-stationary nature. For the autoregressive parameters ϕ_{s_t}, across the two segments of the sample, we assume that $\phi_{s_t} \in \Omega \cup \{1\}$, where $\Omega = \{\phi_{s_t} \mid -1 < \ell_{s_t} \leq \phi_{s_t} < 1\}$, $s_t \in \{1, 2\}$.

Writing model (1) as

$$y_t = \begin{cases} \gamma_1 + \phi_1(y_{t-1} - \gamma_1) + \sigma_1\epsilon_t, & t \leq T_0 \\ \gamma_2 + \phi_2(y_{t-1} - \gamma_2) + \sigma_2\epsilon_t, & t > T_0, \end{cases} \tag{2}$$

where $\gamma_2 = \gamma_1 + \delta$, it can be seen that this model can nest different models which can correspond to economic hypotheses of interest for empirical work. The likelihood function for a sample of T observations, collected in the vector $\mathbf{y} = (y_1, \ldots\ldots, y_T)$, under model (2) with parameter vector $\theta = (\gamma_1, \gamma_2, \phi_1, \phi_2, \sigma_1^2, \sigma_2^2, T_0)$ is given by

$$\begin{aligned} \ell(\mathbf{y} \mid \theta) &= (2\pi)^{-T/2}\sigma_1^{-T_0}\sigma_2^{-(T-T_0)} \exp\left\{-\frac{1}{2\sigma_1^2}\sum_{t=1}^{T_0}[y_t - \gamma_1 - \phi_1(y_{t-1} - \gamma_1)]^2\right\} \\ &\times \exp\left\{-\frac{1}{2\sigma_2^2}\sum_{t=T_0+1}^{T}[y_t - \gamma_2 - \phi_2(y_{t-1} - \gamma_2)]^2\right\}. \end{aligned} \tag{3}$$

In a hypothesis testing setting, it is the unit root hypothesis, $\phi_{s_t} = 1$, that is used as the null hypothesis against the alternative of a stationary model, i.e. $\ell_1 \leq \phi_1 < 1$ and $\ell_2 \leq \phi_2 < 1$. Under the unit root hypothesis, model (2) reduces to

$$y_t = \begin{cases} y_{t-1} + \sigma_1 \epsilon_t, & t \leq T_0 \\ y_{t-1} + \sigma_2 \epsilon_t, & t > T_0. \end{cases} \tag{4}$$

This model assumes that the observations in the two segments of the sample (before and after the break-point T_0) follow random walk processes, but with different variances. We will refer to this model as the broken-variance random walk model. The change in the error variance under the null hypothesis may be attributed to an exogenous event, e.g. a monetary regime change announcement. If there is no structural break in the error variance, i.e. $\sigma_1^2 = \sigma_2^2$, then model (4) reduces to the standard (pure) random walk model

$$y_t = y_{t-1} + \sigma \epsilon_t. \tag{5}$$

In a model comparison setting, we may consider several alternative hypotheses to the unit root hypothesis which can be represented by models (4), or (5). For example, we can consider a model with different error variances, across the two segments of the sample, but the same unconditional mean ($\gamma = \gamma_1 = \gamma_2$), i.e.

$$y_t = \gamma \left(1 - \phi_{s_t}\right) + \phi_{s_t} y_{t-1} + \sigma_{s_t} \epsilon_t. \tag{6}$$

Further, we can assume that the autoregressive coefficient ϕ is also the same ($\phi = \phi_1 = \phi_2$), thus leading to the autoregressive model with different error variances

$$y_t = \gamma \left(1 - \phi\right) + \phi y_{t-1} + \sigma_{s_t} \epsilon_t. \tag{7}$$

On the other hand, we can consider the case where the variances are equal across the two segments, while the unconditional means (gammas) are different and the autoregressive coefficients are either different, i.e.

$$y_t = \left(\gamma_1 + \delta_t\right) \left(1 - \phi_{s_t}\right) + \phi_{s_t} y_{t-1} + \sigma \epsilon_t, \tag{8}$$

or equal , i.e.

$$y_t = \left(\gamma_1 + \delta_t\right) \left(1 - \phi\right) + \phi y_{t-1} + \sigma \epsilon_t. \tag{9}$$

Finally, if there is not a structural break, then, under the alternative hypothesis of stationarity, model (2) reduces to the standard autoregressive model

$$y_t = \gamma \left(1 - \phi\right) + \phi y_{t-1} + \sigma \epsilon_t. \tag{10}$$

For each of the reduced specifications of model (2), the likelihood function (3) can be modified according to the respective vector of parameters.

3 Bayesian Inference and Model Comparison

3.1 Bayesian Inference

The Bayesian approach to inference requires specifying a prior distribution for the unknown parameter vector θ. After observing data $\mathbf{y}$, our knowledge about θ is updated using information in the likelihood function $\ell(\mathbf{y} \mid \theta)$. Then, inference on the unknown parameter vector θ, given the data, is made from the joint posterior distribution of θ, which is given, up to a constant of proportionality, by

$$p(\theta \mid \mathbf{y}) \propto \ell(\mathbf{y} \mid \theta)p(\theta).$$

In complex problems, analytic calculation of the normalizing constant of $p(\theta \mid \mathbf{y})$ is not possible. Instead, computationally intensive methods, such as Monte Carlo integration, Importance sampling and Markov chain Monte Carlo (MCMC) methods, are used to simulate draws from the posterior distribution. MCMC methods (see Gilks et. al., 1996) are based on the construction of an irreducible and aperiodic Markov chain, with realizations $\theta^{(1)}, \theta^{(2)}, \ldots$ in the parameter space, which has $p(\theta \mid \mathbf{y})$ as its stationary distribution. Under mild regularity conditions, the realizations of this Markov chain converge to draws from the posterior distribution of interest [see Roberts and Smith (1994)]. In many statistical applications of the MCMC method a sample from the posterior distribution of interest is obtained via the Gibbs sampler [see, for instance, Geman and Geman (1984), and Gelfand and Smith (1990)]. The Gibbs sampler updates the components of θ one at a time by iteratively generating values from the complete set of their full conditional distributions.

3.2 Bayesian Model Comparison

The objective of the Bayesian approach in the model comparison setting is to determine how probable one model (corresponding to a hypothesis of interest) is relative to another, or various other alternative models. Consider having K competing models $m_1, \ldots, m_K$, each of which corresponds to a different hypothesis. The posterior probability of model m_k, $k = 1, \ldots, K$, (or equivalently of hypothesis H_k) is given by

$$p(m_k \mid \mathbf{y}) = \frac{p(\mathbf{y} \mid m_k)p(m_k)}{\sum_{j=1}^{K} p(\mathbf{y} \mid m_j)p(m_j)},$$

where

$$p(\mathbf{y} \mid m_k) = \int \ell(\mathbf{y} \mid \theta_k, m_k) dp_k(\theta_k) \tag{11}$$

is the marginal likelihood of the vector of observations $\mathbf{y}$ under model m_k, θ_k denotes the model specific parameter vector for model m_k, $\ell(\mathbf{y} \mid \theta_k, m_k)$ is the likelihood function given model m_k, $p_k(\theta_k) = p(\theta_k \mid m_k)$ is the prior density of θ_k under model m_k, and $p(m_k)$ is the prior probability of model m_k. It can be easily seen from (11), that the marginal likelihood under model m_k is just the likelihood function integrated over the specified prior distribution for that model, provided that the integration is feasible. Equivalently, it can be seen as the normalizing constant of the posterior distribution of θ_k, defined as the integral of the product likelihood times prior, which is known as the unnormalized posterior.

In the context of hypothesis testing, inference about the comparison of two different models (say m_{k_1}, corresponding to hypothesis H_{k_1}, and m_{k_2}, corresponding to hypothesis H_{k_2}) can be made using the Bayes Factor (BF) of model m_{k_1} against model m_{k_2} given by

$$BF = \frac{p(\mathbf{y} \mid m_{k_1})}{p(\mathbf{y} \mid m_{k_2})} = \frac{p(m_{k_1} \mid \mathbf{y})}{p(m_{k_2} \mid \mathbf{y})} \frac{p(m_{k_2})}{p(m_{k_1})}.$$

The Bayes Factor is the ratio of the posterior odds to the prior odds. For comparing more than two models, the posterior probabilities of all models can be used as a measure of how probable each model is relative to the others. We will follow the latter approach to compare the stationary and non-stationary models presented in section 2.

4 Calculation of the marginal likelihood

The calculation of the posterior probabilities requires the evaluation of the marginal densities $p(\mathbf{y} \mid m_k)$, defined in (11). Such integrals are in general difficult to calculate; Kass and Raftery (1995) provide an extensive description of available numerical strategies. However, if the prior specification is conjugate to the likelihood function, at least some of the model parameters can be integrated out of the posterior distribution analytically. Under our choice of prior, the marginal likelihood of the pure random walk and of the random walk with broken-variance model can be easily calculated analytically by integration. For the autoregressive break-point models under consideration, most (but not all) of the integrations in (11) can be calculated analytically, while Monte Carlo integration can be used for the rest.

4.1 Prior specification

As stated before, implementation of the Bayesian methodology requires a prior specification for the model parameters. This is very crucial in model comparison, as the choice of the prior can affect the marginal likelihoods of the different models considered for generating the data. As a general principle, note that flat priors tend to penalize more the models which are more complex [see Bernardo and Smith (1994), chapter 6]. In the unit root problem, choosing an appropriate prior distribution is not an easy task. Sims (1988) used a flat prior as an non-informative prior for testing the unit root hypothesis, while Phillips (1991) argued that a flat prior is actually informative and proposed an ignorance prior (or Jeffreys invariant prior). For the AR(1) model with constant term, Schotman and van Dijk (1991a,b) used a proper and weakly informative Normal prior for the constant term, which is centered around the initial observation y_0 and its variance is determined by the other parameters of the model, a uniform prior for the autoregressive coefficient, and a non-informative and improper prior for σ^2.[3] Other authors have proposed prior distributions for the autoregressive coefficient with support that includes a region beyond unity; for a comparative discussion see Bauwens, Lubrano and Richard (1999).

[3]For a detailed discussion on particular choices of prior distributions and their affection on inference and model comparison see, for example, Sims (1991), Leamer (1991), Schotman and van Dijk (1991a,b), Koop and Steel (1991), DeJong and Whiteman (1991c), Phillips (1991).

In a Bayesian analysis of autoregressive models with structural breaks, the presence of different constant terms γ_{s_t} and/or different autoregressive coefficients ϕ_{s_t} will complicate inference in the unit root problem. Dealing with the comparison of unit root models with stationary models in this complex setting, we choose a prior specification in the line of Schotman and van Dijk (1991a). We use a proper prior distribution for the parameters of each of the models under consideration, which is a choice that enables us to integrate the likelihood function with respect to the corresponding probability measure in order to obtain the marginal likelihood of the model. In general, we use quite/weakly informative priors, which are appropriate in a Bayesian model comparison setting. Except for the prior specification within each model, in order to calculate the posterior model probabilities $p(m_k \mid \mathbf{y})$ one has to assign a prior probability $p(m_k)$ to each model m_k. We choose to assign equal prior probability to each of the models under consideration, as a non-informative choice of prior.

For the autoregressive coefficient ϕ_{s_t}, we assume a uniform prior distribution over a part $(\ell_{s_t}, 1)$ of the stationary region, i.e. $p(\phi_{s_t}) = 1/(1-\ell_{s_t})$, $s_t = 1, 2$. Choosing the lower bound ℓ_{s_t} of the support of ϕ_{s_t} is a crucial step of the prior specification. If the range $(\ell_{s_t}, 1)$ is not supported by the data, i.e. ℓ_{s_t} is small enough to include a significant interval with zero density to the left of the support of the posterior, then the posterior model probabilities are biased against stationarity. With respect to the break point T_0, we use a discrete uniform distribution on the integers $\{1, \ldots\ldots, T-1\}$. For the error variances $\sigma^2_{s_t}$, we assume the conjugate inverted Gamma priors $IG(c_{s_t}, d_{s_t})$, $s_t = 1, 2$. Finally, for γ_{s_t}, we adopt Normal prior distributions $N(\mu_{\gamma_{s_t}}, \tau^2_{s_t}\sigma^2_{s_t})$, with hyperparameters $\mu_{\gamma_{s_t}}$ and $\tau^2_{s_t}$, $s_t = 1, 2$.

In the reduced specification models with common error variances $\sigma^2_{s_t} = \sigma^2$ [see equations (5), (8), (9) and (10)], we assume that the prior distribution for the different γ_{s_t} is $N(\mu_{\gamma_{s_t}}, \tau^2_{s_t}\sigma^2)$ and the prior for σ^2 is $IG(c, d)$. For the rest of the parameters of these models, we use prior specifications similar with that of the most general model (2). For the reduced models with common γ and different error variances [(6) and (7)], we assume the prior $N(\mu_\gamma, \sigma^2_\gamma)$ for γ, while for the rest of the parameters the priors are assumed to be the same with those of model (2).

4.2 Analytic marginal likelihood calculation for Random walk models

For the broken-variance random walk model (4), with parameter vector $\theta = (\sigma^2_1, \sigma^2_2, T_0)$, the unnormalized posterior distribution $\ell(\mathbf{y} \mid \theta)p(\theta)$ can be integrated analytically with respect to the nuisance parameters σ^2_1 and σ^2_2. Then, the marginal likelihood $p(\mathbf{y})$ for model (4) can be taken by the sum of the resulting expression over the discrete support of the break-point T_0, i.e.

$$\sum_{T_0=1}^{T-1} \left\{ \frac{d_1^{c_1} d_2^{c_2}}{\Gamma(c_1)\Gamma(c_2)\,(T-1)\,(2\pi)^{T/2}} \frac{\Gamma\left(c_1 + \frac{T_0}{2}\right)}{\left[d_1 + \frac{1}{2}\sum\limits_{t=1}^{T_0}(y_t - y_{t-1})^2\right]^{c_1 + \frac{T_0}{2}}} \frac{\Gamma\left(c_2 + \frac{T-T_0}{2}\right)}{\left[d_2 + \frac{1}{2}\sum\limits_{t=T_0+1}^{T}(y_t - y_{t-1})^2\right]^{c_2 + \frac{T-T_0}{2}}} \right\}.$$

Note that the expression in brackets is the unnormalized marginal posterior distribution of the break-point T_0, with the sum being its normalizing constant. Therefore, for the Random walk break-variance model, the discrete marginal posterior distribution of T_0 can be also calculated analytically.

For the pure Random walk model (5) with constant error variance, i.e. $\sigma^2 = \sigma_1^2 = \sigma_2^2$, the marginal likelihood $p(\mathbf{y})$ is given by

$$p(\mathbf{y}) = \frac{d^c}{\Gamma(c)(2\pi)^{T/2}} \frac{\Gamma\left(c + \frac{T}{2}\right)}{\left[d + \frac{1}{2}\sum\limits_{t=1}^{T}(y_t - y_{t-1})^2\right]^{c+\frac{T}{2}}}.$$

4.3 Marginal likelihood calculation for autoregressive models with a break point

For the autoregressive model (2), with $\phi_{s_t} \in \Omega$, the unnormalized posterior can be integrated analytically with respect to the model parameters γ_1, γ_2, σ_1^2 and σ_2^2. The steps of the multi-dimensional integration are presented in Appendix A. From the final step, we can obtain the unnormalized marginal posterior of (ϕ_1, ϕ_2), i.e. $\ell(\mathbf{y}|\phi_1, \phi_2)p(\phi_1, \phi_2)$, given by

$$\sum_{T_0=1}^{T-1} \left\{ \frac{d_1^{c_1} d_2^{c_2}}{\Gamma(c_1)\Gamma(c_2)(2\pi)^{T/2}(1-l_1)(1-l_2)(T-1)} \frac{\Gamma\left(c_1 + \frac{T_0}{2}\right)}{\sqrt{1 + T_0\tau_1^2(\phi_1 - 1)^2}} \frac{\Gamma\left(c_2 + \frac{T-T_0}{2}\right)}{\sqrt{1 + \tau_2^2(T-T_0)(\phi_2 - 1)^2}} \right.$$

$$\times \left[d_1 + \frac{1}{2\tau_1^2}\left(\mu_{\gamma_1}^2 + \tau_1^2 \sum_{t=1}^{T_0}(\phi_1 y_{t-1} - y_t)^2 - \frac{\left(\mu_{\gamma_1} + \tau_1^2(\phi_1 - 1)\sum_{t=1}^{T_0}(\phi_1 y_{t-1} - y_t)\right)^2}{1 + T_0\tau_1^2(\phi_1 - 1)^2} \right) \right]^{-\left(c_1 + \frac{T_0}{2}\right)}$$

$$\left. \times \left[d_2 + \frac{1}{2\tau_2^2}\left(\mu_{\gamma_2}^2 + \tau_2^2 \sum_{t=T_0+1}^{T}(\phi_2 y_{t-1} - y_t)^2 - \frac{\left(\mu_{\gamma_2} + \tau_2^2(\phi_2 - 1)\sum_{t=T_0+1}^{T}(\phi_2 y_{t-1} - y_t)\right)^2}{1 + \tau_2^2(T-T_0)(\phi_2 - 1)^2} \right) \right]^{-\left(c_2 + \frac{T-T_0}{2}\right)} \right\}.$$

The marginal likelihood for model (2) can be obtained by Monte Carlo integration of $\ell(\mathbf{y}|\phi_1, \phi_2)p(\phi_1, \phi_2)$ with respect to the autoregressive parameters ϕ_1 and ϕ_2. Following Kass and Raftery (1995), we can use Monte Carlo integration with Importance sampling to evaluate the integral $\int p(\phi_1, \phi_2)\ell(\mathbf{y} \mid \phi_1, \phi_2)d\phi_1 d\phi_2$. After specifying an importance function, denoted as $g(\phi_1, \phi_2)$, the importance weight of a random draw $(\phi_1^{(i)}, \phi_2^{(i)})$ from $g(\phi_1, \phi_2)$ can be defined as $w^{(i)} = p\left(\phi_1^{(i)}, \phi_2^{(i)}\right)/g\left(\phi_1^{(i)}, \phi_2^{(i)}\right)$. Then, an estimate of the marginal likelihood $p(\mathbf{y})$, denoted as $\hat{p}(\mathbf{y})$, can be calculated as

$$\hat{p}(\mathbf{y}) = \frac{\sum_{i=1}^{N} w^{(i)}\ell\left(\mathbf{y} \mid \phi_1^{(i)}, \phi_2^{(i)}\right)}{\sum_{i=1}^{N} w^{(i)}},$$

with all normalizing constants included, where $(\phi_1^{(i)}, \phi_2^{(i)})$, $i = 1, \ldots, N$, denote a sample of N draws from the importance function $g(\phi_1, \phi_2)$. As importance function, we will adopt the product of two normal densities $N(\mu_{\phi_i}, \sigma_{\phi_i}^2)$ truncated at the support of ϕ_1 and ϕ_2, respectively. A natural choice for μ_{ϕ_i} can be $\mu_{\phi_i} = (\ell_i + 1)/2$, that is the center of the support $(\ell_i, 1)$ of ϕ_i. $\sigma_{\phi_i}^2$ should be chosen to be large enough, so that the tails of the importance function to be sufficiently heavy.

The above procedure of calculating the marginal likelihood for the stationary autoregressive model (2) can be applied to the other specifications of the break-point models nested in the general model (2).

For model (6), which assumes a common mean, i.e. $\gamma = \gamma_1 = \gamma_2$, we can integrate the unnormalized posterior with respect to the parameters σ_1^2, σ_2^2, ϕ_1 and ϕ_2. This yields the following formula for the unnormalized marginal posterior $\ell(\mathbf{y}|\gamma)p(\gamma)$

$$\sum_{T_0=1}^{T-1} \left\{ \frac{d_1^{c_1} d_2^{c_2} \pi}{\Gamma(c_1)\Gamma(c_2)(2\pi)^{\frac{T+1}{2}} \sigma_\gamma \left(1-\ell_1\right)\left(1-\ell_2\right)\left(T-1\right)} \left(\frac{1}{2}\right)^{-\left(c_1+c_2+\frac{T}{2}\right)} \right.$$

$$\times \Gamma\left[\frac{2c_1+T_0-1}{2}\right] \Gamma\left[\frac{2c_2+T-T_0-1}{2}\right] \left(F_{t_1}\left(1\right) - F_{t_1}\left(\ell_1\right)\right)\left(F_{t_2}\left(1\right) - F_{t_2}\left(\ell_2\right)\right)$$

$$\times \exp\left[-\frac{1}{2\sigma_\gamma^2}\left(\gamma-\mu_\gamma\right)^2\right] \left[\sum_{t=1}^{T_0}\left(y_{t-1}-\gamma\right)^2\right]^{-1/2} \left[\sum_{t=T_0+1}^{T}\left(y_{t-1}-\gamma\right)^2\right]^{-1/2}$$

$$\times \left[\sum_{t=1}^{T_0}\left(y_t-\gamma\right)^2 + 2d_1 - \frac{\left[\sum_{t=1}^{T_0}\left(y_{t-1}-\gamma\right)\left(y_t-\gamma\right)\right]^2}{\sum_{t=1}^{T_0}\left(y_{t-1}-\gamma\right)^2}\right]^{-\left(c_1+\frac{T_0-1}{2}\right)}$$

$$\left. \times \left[\sum_{t=T_0+1}^{T}\left(y_t-\gamma\right)^2 + 2d_2 - \frac{\left[\sum_{t=T_0+1}^{T}\left(y_{t-1}-\gamma\right)\left(y_t-\gamma\right)\right]^2}{\sum_{t=T_0+1}^{T}\left(y_{t-1}-\gamma\right)^2}\right]^{-\left(c_2+\frac{T-T_0-1}{2}\right)} \right\}.$$

In the last relationship, $F_{t_1}(\cdot)$ stands for the non-central student-t distribution function with $2c_1+T_0-1$ degrees of freedom, non-centrality parameter given by

$$\frac{\sum_{t=1}^{T_0}\left(y_{t-1}-\gamma\right)\left(y_t-\gamma\right)}{\sum_{t=1}^{T_0}\left(y_{t-1}-\gamma\right)^2}$$

and scale parameter given by

$$\frac{1}{\left(2c_1+T_0-1\right)\sum_{t=1}^{T_0}\left(y_{t-1}-\gamma\right)^2} \left[\sum_{t=1}^{T_0}\left(y_t-\gamma\right)^2 + 2d_1 - \frac{\left[\sum_{t=1}^{T_0}\left(y_{t-1}-\gamma\right)\left(y_t-\gamma\right)\right]^2}{\sum_{t=1}^{T_0}\left(y_{t-1}-\gamma\right)^2}\right].$$

$F_{t_2}(\cdot)$ stands for the non-central student-t distribution function with $2c_2+T-T_0-1$ degrees of freedom, non-centrality parameter given by

$$\frac{\sum_{t=T_0+1}^{T}\left(y_{t-1}-\gamma\right)\left(y_t-\gamma\right)}{\sum_{t=T_0+1}^{T}\left(y_{t-1}-\gamma\right)^2}$$

and scale parameter given by

$$\frac{1}{\left(2c_2+T-T_0-1\right)\sum_{t=T_0+1}^{T}\left(y_{t-1}-\gamma\right)^2} \left[\sum_{t=T_0+1}^{T}\left(y_t-\gamma\right)^2 + 2d_2 - \frac{\left[\sum_{t=T_0+1}^{T}\left(y_{t-1}-\gamma\right)\left(y_t-\gamma\right)\right]^2}{\sum_{t=T_0+1}^{T}\left(y_{t-1}-\gamma\right)^2}\right].$$

As for model (2), the marginal likelihood for model (6) can be obtained by Monte Carlo integration of $\ell(\mathbf{y}|\gamma)p(\gamma)$ with respect to γ using the Importance function $N(\mu_\gamma, \sigma_\gamma^2)$, where μ_γ is the sample mean of y_t, denoted as $\bar{y}$, and σ_γ^2 takes a sufficiently large value.

Since the posterior distribution for the general model (2), or its reduced specifications, can not be integrated over some subset of the model parameters, the marginal posterior distribution of the break-point T_0 cannot be calculated analytically for those models. Instead, a sample from the posterior distribution of T_0 has to be obtained within an MCMC sampling scheme. For each specification, a Gibbs sampler can be constructed to generate a sample from the conditional posterior distributions of the model parameters including T_0. Details of the sampler for model (2) can be found in Appendix B.

5 Simulation Study

In this section, we conduct a simulation study with the aim to assess the performance of the Bayesian approach suggested in the previous sections to detect the correct data generating process and to identify a break-point in the data. This exercise will enable us to see whether the Bayesian approach can distinguish the unit root models from stationary autoregressive models allowing for a break in the mean or the error variance of the models. To this end, we estimate the posterior model probabilities of the following models, presented in Section 2,

$$
\begin{aligned}
&m_1 : y_t = (\gamma_1 + \delta_t)(1 - \phi_{s_t}) + \phi_{s_t} y_{t-1} + \sigma_{s_t}\epsilon_t \quad \text{(general model (1))}\\
&m_2 : y_t = \gamma(1 - \phi) + \phi y_{t-1} + \sigma_{s_t}\epsilon_t \quad \text{(common } \gamma \text{ -common } \phi)\\
&m_3 : y_t = (\gamma_1 + \delta_t)(1 - \phi) + \phi y_{t-1} + \sigma\epsilon_t \quad \text{(common } \sigma^2\text{-common } \phi)\\
&m_4 : y_t = \gamma(1 - \phi_{s_t}) + \phi_{s_t} y_{t-1} + \sigma_{s_t}\epsilon_t \quad \text{(common } \gamma)\\
&m_5 : y_t = (\gamma_1 + \delta_t)(1 - \phi_{s_t}) + \phi_{s_t} y_{t-1} + \sigma\epsilon_t \quad \text{(common } \sigma^2)\\
&m_6 : y_t = y_{t-1} + \sigma_{s_t}\epsilon_t \quad \text{(random walk with broken variance)}\\
&m_7 : y_t = y_{t-1} + \sigma\epsilon_t \quad \text{(pure random walk)}\\
&m_8 : y_t = \gamma(1 - \phi) + \phi y_{t-1} + \sigma\epsilon_t \ \text{(standard AR(1) model)},
\end{aligned}
$$

when data are generated from the models of interest $m_1 - m_7$. In our experiments, we consider representative time series samples of size $T = \{200, 300, 500\}$, with break-points at $T_0 = \{T/4, T/2, 3T/4\}$. For the autoregressive parameters, we consider changes near to the unit root, i.e. $\phi_1 = 0.90$ and $\phi_2 = 0.95$, while for the nuisance parameters we assume that $\gamma_1 = 0.01$ and $\gamma_2 = 0.02$, or $\gamma_2 = 0.03$ (see Tables 1-7), and $\sigma_1^2 = 0.009$ and $\sigma_2^2 = 0.03$, across the two segments of the sample. For the initial observation y_0, we assume that $y_0 = 0$. The above values of the parameters are chosen so that to reflect bigger changes in the variance than the mean of the series, which are consistent with many financial economic series.

The prior specifications that we consider in our simulation study are of the general form described in subsection 4.1. For γ_{s_t}, $s_t = 1, 2$, we adopt Normal prior distributions $N\left(\mu_{\gamma_{s_t}}, \tau_{s_t}^2 \sigma_{s_t}^2\right)$. The value of $\mu_{\gamma_{s_t}}$ determines the center of the prior distribution of γ_{s_t}, while $\tau_{s_t}^2$ determines the magnitude of the prior variance of γ_{s_t}. We consider values of $\tau_{s_t}^2$ in the interval $\left(\frac{1}{1-\ell_{s_t}^2}, +\infty\right)$, where ℓ_{s_t} is the lower bound of the support of ϕ_{s_t}. As an initial choice, we take $\mu_{\gamma_1} = y_0$, $\mu_{\gamma_2} = \bar{y}$ and $\tau_1^2 = \tau_2^2 = 4$. This prior is quite informative with respect to the centers $\mu_{\gamma_{s_t}}$. In particular, the data dependent choices $\mu_{\gamma_1} = y_0$ and

$\mu_{\gamma_2} = \bar{y}$ can help to identify a change in the mean of the series between the two sample segments. To see if the Bayesian approach remains robust to alternative prior specifications for γ_{s_t}, we also consider a prior for γ_{s_t} with a weakly informative mean $\mu_{\gamma_{s_t}} = y_0$ and $\tau_{s_t}^2 = \frac{1}{1-\ell_{s_t}^{*2}}$, with $\ell_{s_t}^* = \frac{1+\ell_{s_t}}{2}$, where ℓ_{s_t} is the prior and, hence, posterior, lower bound of the support of ϕ_{s_t}. This prior specification is in the line of Schotman and van Dijk's (1991a,b) prior. It implies a non data dependent prior which has the nice feature that $\tau_{s_t}^2$ is determined by $\ell_{s_t}^*$, that is by the center of the support of ϕ_{s_t}.[4]

For the remaining parameters of the models $m1$-$m8$, we consider the following priors. For the error variance $\sigma_{s_t}^2$, we adopt $IG(0.01, 0.01)$ priors. These distributions are rather defuse, non-informative priors. For the autoregressive coefficients ϕ_{s_t}, we assume uniform prior distributions over the interval $(\ell_{s_t}, 1)$, with $\ell_{s_t} = 0.8$ when the data are simulated from the true value $\phi_{s_t} = 0.9$, and with $\ell_{s_t} = 0.9$ when the data are simulated from the true value $\phi_{s_t} = 0.95$. Note, at this point, that the choice of ℓ_{s_t} is crucial because it can affect the posterior model probabilities. ℓ_{s_t} should not be far away from unity so that to include a significant interval with zero density to the left of the support of the posterior. But, it should not be however closed enough to unity so that to leave an interval with non-zero density out of the support of the posterior. To see if the posterior probabilities remain robust to different values of ℓ_{s_t} than the above, a sensitivity analysis is conducted.

Tables 1-7 present the simulation results on the posterior model probabilities. Panel A presents the results using the quite informative $N(\mu_{\gamma_{s_t}}, \tau_{s_t}^2 \sigma_{s_t}^2)$ prior for $\mu_{\gamma_{s_t}}$, while Panel B presents the results using the weakly informative $N\left(y_0, \tau_{s_t}^2 \sigma_{s_t}^2\right)$ prior, with $\tau_{s_t}^2 = \frac{1}{1-\ell_{s_t}^{*2}}$. The results of the tables lead to the following main conclusions:

First, the Bayesian approach can adequately distinguish the correct models generating the data, especially the unit root models from the stationary ones. Note that the performance of the method significantly increases with the sample size, T.

Second, the alternative prior specifications considered for γ_{s_t} do not seem to significantly affect the performance of the Bayesian method to select the correct model of the data. Both the quite and weakly informative priors for γ_{s_t} seem to select the correct model of the data with almost the same posterior probabilities, with the weakly informative priors to perform slightly better in detecting the unit root models.

In Tables 8 and 9 we present the results of the sensitivity analysis, for different values of ℓ_{s_t}, $s_t = 1, 2$. For reasons of space, we report results of the posterior probabilities only for the general model m_1 against the random walk model with broken variance m_6 when the data are generated from model m_1 (see Table 8) and, conversely, for model m_6 against model m_1 when the data are generated from model m_6 (see Table 9). For both cases, we used the weakly informative prior for γ_{s_t}. The results of the tables indicate that the posterior probabilities remain robust to the alternative values of ℓ_{s_t}, even for values of ℓ_{s_t} far

[4]We also considered a $N\left(y_0, \frac{\sigma_{s_t}^2}{1-\phi_{s_t}^2}\right)$ prior for γ_{s_t}, which corresponds to Schotman and van Dijk's prior, allowing for a break point. But, we found that the posterior model probabilities are not different from those found for the $N\left(y_0, \tau_{s_t}^2 \sigma_{s_t}^2\right)$, with $\tau_{s_t}^2 = \frac{1}{1-\ell_{s_t}^{*2}}$. These results are available upon request.

below than unity.[5] As expected, when the range $(\ell_{s_t}, 1)$ is large and it is not supported by the data the posterior odds of the two models is biased against stationarity and this is reflected on the posterior model probabilities. But as ℓ_{st} gets close to unity the posterior probabilities of the random walk model decrease.

Next, we turn into answering the question whether the Bayesian approach can efficiently identify the break-point of the series. To this end, in Figure 1 we present the histograms of the posterior distributions of the break-point T_0 for simulated data of size $T = \{200, 300, 500\}$ from the general model, m_1, with break-points at the fractions $\{T/4, T/2, 3T/4\}$ of the sample. In Figure 2 we present the exact marginal posterior distributions of the break-point T_0 for series of size $T = \{200, 300, 500\}$ from the random walk model with broken variance, m_6, with break-points at the fractions $\{T/4, T/2, 3T/4\}$ of the sample. Inspection of the histograms indicate that the Bayesian methodology can efficiently identify the correct break-point of the data for almost all the fractions of the sample. They show that the posterior mass functions give very narrow ranges, which cover the true break-points. This becomes more apparent, as the sample size increases. However, for smaller samples the performance of the Bayesian approach is satisfactory when the posterior of T_0 is analytically evaluated. This can be seen by comparing the posterior distributions of $T_0 = T/4$, for the two models.

6 Applications to financial data

In this section, we give some empirical applications of our Bayesian method using financial economic series. In particular, we investigate whether evidence of unit roots in nominal US interest rates and CPI-based real bilateral exchange rates against the US dollar can be attributed to structural breaks. In implementing our method, we use prior distributions of the type described in section 4.1. For the break-time T_0, we use a discrete uniform distribution, while for the error variances $\sigma^2_{s_t}$, $s_t = 1, 2$ we adopt $IG(0.01, 0.01)$ priors. For the uniform priors of the autoregressive coefficients ϕ_{s_t}, we choose hyperparameters values $\ell_{s_t} = 0.8$, $s_t = 1, 2$, to reflect our prior ignorance about ϕ_{s_t}, as well as the fact that the estimates of the autoregressive coefficients are expected to be near unity. Finally, for the constant terms we adopt $N(\mu_{\gamma_{s_t}}, \tau^2_{s_t}\sigma^2_{s_t})$ priors, with $\mu_{\gamma_{s_t}} = y_0$ and $\tau_{s_t} = 2.3$ for the nominal interest rates and $\tau_{s_t} = 10$ for the real exchange rates. These priors are quite informative with respect to the centers $\mu_{\gamma_{s_t}}$, while the tuning hyperparameters τ_{s_t} are chosen so that the prior for the error variances to be consistent with the magnitude of the data.

6.1 Structural changes in US interest rates

Our empirical application on the US nominal interest rates uses monthly data from 1970:01 to 1991:01 for the one-month and the one-year to maturity interest rates. These data were obtained from McCulloch and Kown's (1993) term structure data base and cover the period from 1979:09 to 1982:09, during which

[5] Analogous results can be obtained using the informative prior specification for $\mu_{\gamma_{s_t}}$.

interest rates fluctuated freely due to changes in the monetary operating procedures by the Fed. These procedures were abolished in 1982:09, when a pegged system of interest rates was adopted.

In Table 10a we present the posterior probabilities for all models $m1$-$m8$, while in Figure 3 we present the graphs of the series together with the histogram of the break-point T_0 of the most probable model (here $m2$). To see if classical unit root sequential procedures can capture a potential break in the data of unknown date, we also present plots of the Zivot-Andrews' (1992) test statistic -henceforth ZA. This statistic allows for a break only in the levels of the series. In contrast to the results of the ZA test statistic, the results of our Bayesian method reveal that the unit root hypothesis can not be supported by the data and that there is a break in the data. For all interest rates, the posterior probabilities for the random walk models $m6$ or $m7$ are very small, compared with the stationary models, while the standard autoregressive model m_8 has zero posterior probability.

In Table 10b, we present the mean and standard deviations (in parentheses) of the posterior distributions of the parameters' estimates of model m_2, which is found to be the most probable model of the data. For T_0, we also report the mode of the posterior distribution. This model allows for a break only in the error variance. The results of the table and Figure 3 clearly indicate that the most likely break-point date is very close (or coincides, for the one-month) to the date 1982:09 of the monetary regime change. After this date, the variance of interest rates significantly reduced due to the changes in the monetary operating procedures.

6.2 Structural changes in real exchange rates

Our empirical analysis for real exchange rates is based on the following countries' currencies against the US dollar: UK, Denmark, France and Netherlands, and it covers the period from 1974:07 to 1999:05. Some of these series were analyzed by Schotman and van Dijk (1991a), who applied a Bayesian approach to compare the pure random walk model with the standard stationary model for real exchange rates.

In Table 11a, we present the posterior probabilities of all models $m1$-$m8$, while in Figure 4 we plot the real exchange series together with the histograms of the break-point T_0 which is found to be the most likely stationary model (here $m5$). As in Figure 3, we also plot the estimates of the sequential ZA test statistic. The results of the table indicate that model $m5$, with a break in the mean and the autoregressive coefficients, can better explain the real exchange rates series examined compared with the random walk model. The only exception is the real exchange rate for the UK currency for which the random walk model with broken variance seems to be the most probable model.Compared with Schotman and van Dijk's (1991a), our results indicate that it is more likely to reject the unit root hypothesis for real exchange rates in a Bayesian framework which allows for structural breaks.

In Table 11b, we report the mean and standard deviations of the posterior probabilities of model's $m5$ parameters, together with the mode of T_0. The results of this table and Figure 4 indicate that there is a substantial shift in the autoregression coefficient after 1985:01, implying higher speed of convergence to the equilibrium real exchange rates after that date and a substantial appreciation of the US dollar, compared with the European currencies. This may be associated to Volcker's tight monetary, stabilization programme in early eighties which abruptly reduced US inflation.

6.3 Conclusions

In this paper, we have suggested a Bayesian approach to compare the random walk model with, or without, structural changes in the error variance with the first-order stationary autoregressive model allowing for a potential structural change (break) of an unknown date in any of its parameters. The allowance for a structural change in the error variance, alongside the other parameters of the model, is consistent with evidence on many financial series.

The Bayesian approach that we have followed in the paper employs analytic and Monte Carlo integration techniques for calculating the marginal likelihoods of the models, which are necessary for calculating the posterior model probabilities. This is done under a choice of prior which is quite flexible to accommodate informative or less informative priors, often used in the unit root Bayesian literature. Simulation results have shown that our method can adequately distinguish stationarity from unit roots, for moderately large sample sizes. Empirical applications of our method to US nominal interest rates and real bilateral exchange rates of European currencies against the US dollar have shown that, in contrast to classical testing procedures, the Bayesian approach favors stationary models allowing for structural breaks, compared with unit root models.

Appendix A

To compute the marginal likelihood (11) for the single change-point autoregressive model (2) we have to integrate the unnormalized posterior with respect to the model parameters. We assume a Discrete Uniform prior distribution for the time break T_0, a Uniform $U(\ell_{s_t}, 1)$ prior for the autoregressive coefficient ϕ_{s_t}, Inverted Gamma $IG(c_{s_t}, d_{s_t})$ and Normal $N(\mu_{\gamma_{s_t}}, \tau^2_{s_t}\sigma^2_{s_t})$ priors for the variances of the error processes $\sigma^2_{s_t}$, and the constant terms γ_{s_t}, respectively. We perform analytic integration with respect to the parameters γ_1, γ_2, σ_1^2, and σ_2^2, then we take the sum over the discrete support of T_0, and finally we use Monte Carlo integration for ϕ_1 and ϕ_2 to obtain the marginal likelihood. The calculation of the integrals with respect to γ_1, γ_2, σ_1^2, and σ_2^2 are presented below.

The integral $\int \ell(\mathbf{y}|\theta)p(\theta)d\gamma_1 d\gamma_2$ of the unnormalized posterior with respect to the parameters γ_1 and γ_2 results in

$$\frac{1}{(2\pi)^{\frac{T}{2}}}\frac{d_1^{c_1}}{\Gamma(c_1)}\frac{d_2^{c_2}}{\Gamma(c_2)}\frac{1}{T-1}\frac{1}{1-\ell_1}\frac{1}{1-\ell_2}\frac{1}{\sqrt{1+T_0\tau_1^2(\phi_1-1)^2}}\frac{1}{\sqrt{1+(T-T_0)\tau_2^2(\phi_2-1)^2}}\left(\sigma_1^{-2}\right)^{\frac{T_0}{2}+c_1-1}\left(\sigma_2^{-2}\right)^{\frac{T-T_0}{2}+c_2-1}$$

$$\times\exp\left\{-\left(\sigma_1^{-2}\right)\left[d_1+\frac{1}{2\tau_1^2}\left(\mu_{\gamma_1}^2+\tau_1^2\sum_{t=1}^{T_0}(\phi_1 y_{t-1}-y_t)^2-\frac{\left(\mu_{\gamma_1}+\tau_1^2(\phi_1-1)\sum_{t=1}^{T_0}(\phi_1 y_{t-1}-y_t)\right)^2}{1+T_0\tau_1^2(\phi_1-1)^2}\right)\right]\right\}$$

$$\times\exp\left\{-\left(\sigma_2^{-2}\right)\left[d_2+\frac{1}{2\tau_2^2}\left(\mu_{\gamma_2}^2+\tau_2^2\sum_{t=T_0+1}^{T}(\phi_2 y_{t-1}-y_t)^2-\frac{\left(\mu_{\gamma_2}+\tau_2^2(\phi_2-1)\sum_{t=T_0+1}^{T}(\phi_2 y_{t-1}-y_t)\right)^2}{1+(T-T_0)\tau_2^2(\phi_2-1)^2}\right)\right]\right\}.$$

After integrating the above expression with respect to σ_1^2, and σ_2^2 we obtain the integral $\int \ell(\mathbf{y}|\theta)p(\theta)d\gamma_1 d\gamma_2 d\sigma_1^{-2} d\sigma_2^{-2}$ which is given by

$$\frac{1}{(2\pi)^{\frac{T}{2}}}\frac{d_1^{c_1}}{\Gamma(c_1)}\frac{d_2^{c_2}}{\Gamma(c_2)}\frac{1}{T-1}\frac{1}{1-\ell_1}\frac{1}{1-\ell_2}\frac{1}{\sqrt{1+T_0\tau_1^2(\phi_1-1)^2}}\frac{1}{\sqrt{1+\tau_2^2(T-T_0)(\phi_2-1)^2}}$$

$$\times\frac{\Gamma\left[c_1+\frac{T_0}{2}\right]}{\left[d_1+\frac{1}{2\tau_1^2}\left(\mu_{\gamma_1}^2+\tau_1^2\sum_{t=1}^{T_0}(\phi_1 y_{t-1}-y_t)^2-\frac{\left(\mu_{\gamma_1}+\tau_1^2(\phi_1-1)\sum_{t=1}^{T_0}(\phi_1 y_{t-1}-y_t)\right)^2}{1+T_0\tau_1^2(\phi_1-1)^2}\right)\right]^{c_1+\frac{T_0}{2}}}$$

$$\times\frac{\Gamma\left[c_2+\frac{T-T_0}{2}\right]}{\left[d_2+\frac{1}{2\tau_2^2}\left(\mu_{\gamma_2}^2+\tau_2^2\sum_{t=T_0+1}^{T}(\phi_2 y_{t-1}-y_t)^2-\frac{\left(\mu_{\gamma_2}+\tau_2^2(\phi_2-1)\sum_{t=T_0+1}^{T}(\phi_2 y_{t-1}-y_t)\right)^2}{1+\tau_2^2(T-T_0)(\phi_2-1)^2}\right)\right]^{c_2+\frac{T-T_0}{2}}}.$$

Finally, we obtain $\ell(\mathbf{y}|\phi_1,\phi_2)p(\phi_1,\phi_2)$ by taking the sum of the above expression over the discrete support of T_0.

Appendix B

The MCMC sampling scheme for obtaining draws from the posterior distribution of the parameters of the break-point autoregressive model (2) is presented here. All the conditional distributions used are of known functional form and, therefore, the parameters can be updated by using the Gibbs sampler. Note that at steps 1-5 certain parameters have been integrated out of the conditional posterior distribution of the parameter that is updated at that step. These marginalizations help to reduce the correlations among the components of the produced Markov chain.

- *Step 1:* Update the time break T_0, from its conditional distribution $p\left(T_0 \mid \mathbf{y}, \phi_1, \phi_2\right)$, which is a mass function proportional to

$$\frac{\Gamma\left[c_1+\frac{T_0}{2}\right]\Gamma\left[c_2+\frac{T-T_0}{2}\right]}{\sqrt{1+T_0\tau_1^2(\phi_1-1)^2}\sqrt{1+\tau_2^2(T-T_0)(\phi_2-1)^2}}\times$$

$$\left[d_1+\frac{1}{2\tau_1^2}\left(\mu_{\gamma_1}^2+\tau_1^2\sum_{t=1}^{T_0}\left(\phi_1 y_{t-1}-y_t\right)^2-\frac{\left(\mu_{\gamma_1}+\tau_1^2\left(\phi_1-1\right)\sum_{t=1}^{T_0}\left(\phi_1 y_{t-1}-y_t\right)\right)^2}{1+T_0\tau_1^2(\phi_1-1)^2}\right)\right]^{-\left(c_1+\frac{T_0}{2}\right)}\times$$

$$\left[d_2+\frac{1}{2\tau_2^2}\left(\mu_{\gamma_2}^2+\tau_2^2\sum_{t=T_0+1}^{T}\left(\phi_2 y_{t-1}-y_t\right)^2-\frac{\left(\mu_{\gamma_2}+\tau_2^2\left(\phi_2-1\right)\sum_{t=T_0+1}^{T}\left(\phi_2 y_{t-1}-y_t\right)\right)^2}{1+\left(T-T_0\right)\tau_2^2(\phi_2-1)^2}\right)\right]^{-\left(c_2+\frac{T-T_0}{2}\right)}.$$

- *Step 2:* Update parameter γ_1 from its conditional posterior distribution $p\left(\gamma_1 \mid \mathbf{y}, T_0, \phi_1\right)$, which is a non-central Student-t distribution with $2c_1+T_0$ degrees of freedom, non-centrality parameter

$$\frac{\mu_{\gamma_1}+\tau_1^2(\phi_1-1)\sum_{t=1}^{T_0}(\phi_1 y_{t-1}-y_t)}{1+\tau_1^2(\phi_1-1)^2 T_0}$$

and scale parameter

$$\frac{1}{(2c_1+T_0)(1+\tau_1^2(\phi_1-1)^2 T_0)}\times$$
$$\left[2\tau_1^2 d_1+\mu_{\gamma_1}^2+\tau_1^2\sum_{t=1}^{T_0}\left(\phi_1 y_{t-1}-y_t\right)^2-\frac{\left(\mu_{\gamma_1}+\tau_1^2\left(\phi_1-1\right)\sum_{t=1}^{T_0}\left(\phi_1 y_{t-1}-y_t\right)\right)^2}{1+T_0\tau_1^2(\phi_1-1)^2}\right].$$

- *Step 3:* Update parameter γ_2 from its conditional posterior distribution $p\left(\gamma_2 \mid \mathbf{y}, T_0, \phi_2\right)$, which is a non-central Student-t distribution with $2c_2+T-T_0$ degrees of freedom, non-centrality parameter

$$\frac{\mu_{\gamma_2}+\tau_2^2(\phi_2-1)\sum_{t=T_0+1}^{T}(\phi_2 y_{t-1}-y_t)}{1+\tau_2^2(\phi_2-1)^2(T-T_0)}$$

and scale parameter

$$\frac{1}{(2c_2+T-T_0)(1+\tau_2^2(\phi_2-1)^2(T-T_0))}\times$$
$$\left[2\tau_2^2 d_2+\mu_{\gamma_2}^2+\tau_2^2\sum_{t=T_0+1}^{T}\left(\phi_2 y_{t-1}-y_t\right)^2-\frac{\left(\mu_{\gamma_2}+\tau_2^2\left(\phi_2-1\right)\sum_{t=T_0+1}^{T}\left(\phi_2 y_{t-1}-y_t\right)\right)^2}{1+\left(T-T_0\right)\tau_2^2(\phi_2-1)^2}\right].$$

- *Step 4:* Update parameter σ_1^2 from its conditional posterior distribution $p\left(\sigma_1^2 \mid \mathbf{y}, T_0, \phi_1\right)$ which is an Inverted Gamma distribution with shape parameter

$$c_1+\frac{T_0}{2}$$

and rate parameter

$$d_1 + \frac{1}{2\tau_1^2}\left(\mu_{\gamma_1}^2 + \tau_1^2 \sum_{t=1}^{T_0} (\phi_1 y_{t-1} - y_t)^2 - \frac{\left(\mu_{\gamma_1} + \tau_1^2 (\phi_1 - 1) \sum_{t=1}^{T_0} (\phi_1 y_{t-1} - y_t)\right)^2}{1 + T_0 \tau_1^2 (\phi_1 - 1)^2}\right).$$

- *Step 5:* Update parameter σ_2^2 from its conditional posterior distributions $p\left(\sigma_2^2 \mid \mathbf{y}, T_0, \phi_2\right)$ which is an Inverted Gamma distribution with shape parameter

$$c_2 + \frac{T - T_0}{2}$$

and rate parameter

$$d_2 + \frac{1}{2\tau_2^2}\left(\mu_{\gamma_2}^2 + \tau_2^2 \sum_{t=T_0+1}^{T} (\phi_2 y_{t-1} - y_t)^2 - \frac{\left(\mu_{\gamma_2} + \tau_2^2 (\phi_2 - 1) \sum_{t=T_0+1}^{T} (\phi_2 y_{t-1} - y_t)\right)^2}{1 + (T - T_0)\, \tau_2^2 (\phi_2 - 1)^2}\right).$$

- *Step 6:* Update ϕ_1 from its full conditional distribution, $p\left(\phi_1 \mid \mathbf{y}, T_0, \sigma_1^2, \gamma_1\right)$ which is a truncated Normal distribution on the interval $(\ell_1, 1)$ with mean

$$\frac{\sum_{t=1}^{T_0}(y_t - \gamma_1)(y_{t-1} - \gamma_1)}{\sum_{t=1}^{T_0}(y_{t-1} - \gamma_1)^2}$$

and variance

$$\frac{\sigma_1^2}{\sum_{t=1}^{T_0}(y_{t-1} - \gamma_1)^2}.$$

- *Step 7:* Update ϕ_2 from its full conditional distribution, $p\left(\phi_2 \mid \mathbf{y}, T_0, \sigma_2^2, \gamma_2\right)$ which is a truncated Normal distribution on the interval $(\ell_2, 1)$ with mean

$$\frac{\sum_{t=T_0+1}^{T}(y_t - \gamma_2)(y_{t-1} - \gamma_2)}{\sum_{t=T_0+1}^{T}(y_{t-1} - \gamma_2)^2}$$

and variance

$$\frac{\sigma_2^2}{\sum_{t=T_0+1}^{T}(y_{t-1} - \gamma_2)^2}.$$

References

Bauwens, L., Lubrano, M., and J. Richard (1999). *Bayesian Inference in Dynamic Econometric Models*, Oxford University press.

Bernardo, J.M., and A.F.M. Smith (1994). *Bayesian Theory*, Wiley Series in Applied Probability and Statistics.

Christiano, L.J. (1992). Searching for a Break in GNP, *Journal of Business and Economic Statistics*, **10**, 237-250.

DeJong, D.N., and C.H. Whiteman (1989). Trends and cycles as unobserved components in US real GNP: A Bayesian perspective, *Proceedings of the American Statistical Association, Business and Economics Section*, pp.63-70.

DeJong, D.N., and C.H. Whiteman (1991a). Reconsidering 'Trends and random walks in macroeconomic time series', *Journal of Monetary Economics*, **28**, 221-254.

DeJong, D.N., and C.H. Whiteman (1991b). The temporal stability of dividends and stock prices: evidence from the likelihood function, *American Economic Review*, **81**, 600-617.

DeJong, D.N., and C.H. Whiteman (1991c).The case for trend-stationarity is stronger than we thought. *Journal of Applied Econometrics*, **6**, 413-421.

Garcia, R. (1998). Asymptotic null distribution of the likelihood ratio test in Markov-switching models, *International Economic Review*, **39**, 763-788.

Gelfand, A., and A.F.M. Smith (1990). Sampling-based approaches to calculating marginal densities, *Journal of the American Statistical Association*, **85**, 398-409.

Geman, S., and D. Geman (1984). Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images, *IEEE Transactions in Pattern Analysis and Machine Intelligence*, **6**, 721-741.

Gilks, W.R., S. Richardson, and D.J. Spiegelhalter, eds. (1996). *Markov Chain Monte Carlo in Practice*, Chapman and Hall.

Hall, A.D., H.M. Anderson, and C.W.J. Granger (1992). Treasury bill yield curves and cointegration, *Review of Economics and Statistics*, **74**, 116-126.

Kass, R.E., and A.E. Raftery (1995). Bayes Factors, *Journal of the American Statistical Association*, **90**, 773-795.

Kim, C.J., amd C.R. Nelson (1999). Has the US economy become more stable? A Bayesian approach based on a Markov-switching model of the business cycle, *The review of Economics and Statistics*, **81**, 608-616.

Koop, G. (1992). ‘Objective’ Bayesian unit root tests, *Journal of Applied Econometrics*, **7**, 65-82.

Koop, G. (1994). Recent progress in applied Bayesian econometrics, *Journal of Economic Survey*, **8**, 1-34.

Koop, G., and M.F.J. Steel (1991). A comment on: ’To criticize the critics: An objective Bayesian analysis of stochastic trends’, by Peter C.B. Phillips, *Journal of Applied Econometrics*, **6**, 365-370.

Leamer, E.E. (1991). Comment on ’To criticize the critics’, *Journal of Applied Econometrics*, **6**, 371-373.

Lumsdaine, R.L., and D.H. Papell (1997). Multiple trend breaks and the unit-root hypothesis, *The Review of Economics and Statistics*, **79**, 212-218.

Marriott, J., and P. Newbold (2000). The strength of evidence for unit autoregressive roots and structural breaks: A Bayesian perspective, *Journal of Econometrics*, **98**, 1-25.

McCulloch, H.J. and Heon-Chol Kown (1993). US term structure data, Working paper 93-6, Ohio State University.

Nelson, C.R., and C.I. Plosser (1982). Trends and Random walks in Macroeconomic Time Series: some evidence and implications, *Journal of Monetary Economics*, **10**, 139-162.

Papell, D.H. (1997). Searching for stationarity: Purchasing power parity under the current float, *Journal of International Economics*, **43**, 313-332.

Papell, D.H. (2002). The great appreciation, the great depreciation, and the purchasing power hypothesis, *Journal of International Economics*, 51-82.

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.

Phillips, P.C.B. (1991). To criticize the critics: An objective Bayesian analysis of stochastic trends, *Journal of Applied Econometrics*, **6**, 333-364.

Roberts, G.O., and A.F.M. Smith (1994). Simple conditions for the convergence of the Gibbs sampler and Metropolis-Hastings algorithms, *Stochastic Processes and their applications*, **49**, 207-216.

Schotman, P.C., and H. van Dijk (1991a). A Bayesian analysis of the unit root in real exchange rates, *Journal of Econometrics*, **49**, 195-238.

Schotman, P.C., and H. van Dijk (1991b). On Bayesian routes to unit roots, *Journal of Econometrics*, **6**, 387-401.

Sims, C.A. (1988). Bayesian scepticism on unit root econometrics, *Journal of Economic dynamics and Control*, **12**, 463-474.

Sims, C.A. (1991). Comment by Christopher A. Sims on 'To criticize the critics', by Peter C.B. Phillips, *Journal of Applied Econometrics*, **6**, 423-434.

Sims, C.A., and H. Uhlig (1991). Understanding unit rooters: a helicopter tour, *Econometrica*, **59**, 1591-1599.

Tzavalis, E. and M.R. Wickens (1997). Explaining the failures of the term spread models of the rational expectations hypothesis of the term structure, *Journal of Money, Credit, and banking*, **29**, 364-380.

Wang, J. and E, Zivot (2000). A Bayesian time series model of multiple structural changes in level, trend and variance, *Journal of Business & Economic Statistics*, **18**, 374-386.

Zivot, E., and D.W.K. Andrews (1992). Further evidence of the great crash, the oil price shock, and the unit root hypothesis, *Journal of Business and Economic Statistics*, **10**, 251-270.

Table 1: Posterior model probabilities. Simulated data from model (2) $[m_1]$ using $\gamma_1 = 0.01$, $\gamma_2 = 0.02$, $\phi_1 = 0.90$, $\phi_2 = 0.95$, $\sigma_1^2 = 0.009$, and $\sigma_2^2 = 0.03$.

T	T_0	m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8
200	50	0.3317	0.0381	0.1350	0.1141	0.2293	0.0183	0.0382	0.0953
200	100	0.7532	0.0403	0.0000	0.1915	0.0000	0.0151	0.0000	0.0000
200	150	0.4043	0.2899	0.0000	0.2814	0.0000	0.0243	0.0000	0.0000
300	75	0.7793	0.0522	0.0001	0.1601	0.0002	0.0081	0.0000	0.0001
300	150	0.4573	0.2205	0.0000	0.3124	0.0000	0.0098	0.0000	0.0000
300	225	0.5512	0.0962	0.0000	0.3489	0.0000	0.0037	0.0000	0.0000
500	125	0.4669	0.2135	0.0000	0.3106	0.0000	0.0089	0.0000	0.0000
500	250	0.6377	0.0584	0.0000	0.3009	0.0000	0.0029	0.0000	0.0000
500	375	0.6306	0.0135	0.0000	0.3556	0.0000	0.0003	0.0000	0.0000
200	50	0.2750	0.0466	0.1409	0.1557	0.2023	0.0236	0.0490	0.1070
200	100	0.6537	0.0491	0.0000	0.2761	0.0000	0.0211	0.0000	0.0000
200	150	0.3696	0.2939	0.0000	0.3067	0.0000	0.0298	0.0000	0.0000
300	75	0.6653	0.0741	0.0001	0.2481	0.0002	0.0121	0.0000	0.0001
300	150	0.3745	0.2397	0.0000	0.3732	0.0000	0.0126	0.0000	0.0000
300	225	0.5047	0.1001	0.0000	0.3904	0.0000	0.0049	0.0000	0.0000
500	125	0.3748	0.2524	0.0000	0.3607	0.0000	0.0121	0.0000	0.0000
500	250	0.5759	0.0709	0.0000	0.3492	0.0000	0.0040	0.0000	0.0000
500	375	0.6149	0.0146	0.0000	0.3701	0.0000	0.0004	0.0000	0.0000

Table 2: Posterior model probabilities. Simulated data from model (7) $[m_2]$ using $\gamma_1 = 0.01$, $\gamma_2 = 0.01$, $\phi_1 = \phi_2 = 0.95$, $\sigma_1^2 = 0.009$, and $\sigma_2^2 = 0.03$.

T	T_0	m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8
200	50	0.1144	0.1710	0.1809	0.1441	0.1682	0.0074	0.0101	0.2039
200	100	0.2876	0.3685	0.0029	0.3163	0.0028	0.0184	0.0002	0.0033
200	150	0.0309	0.0334	0.2950	0.0277	0.2481	0.0017	0.0101	0.3532
300	75	0.1952	0.2551	0.1046	0.2139	0.1007	0.0034	0.0016	0.1255
300	150	0.2791	0.4105	0.0019	0.2990	0.0018	0.0054	0.0000	0.0023
300	225	0.3631	0.3294	0.0000	0.3028	0.0000	0.0047	0.0000	0.0000
500	125	0.2782	0.2853	0.0000	0.4337	0.0000	0.0028	0.0000	0.0000
500	250	0.4813	0.1670	0.0000	0.3496	0.0000	0.0022	0.0000	0.0000
500	375	0.2867	0.3917	0.0000	0.3178	0.0000	0.0039	0.0000	0.0000
200	50	0.1116	0.1738	0.1715	0.1426	0.1641	0.0115	0.0157	0.2093
200	100	0.2729	0.3761	0.0028	0.3127	0.0028	0.0290	0.0003	0.0035
200	150	0.0302	0.0353	0.2756	0.0289	0.2424	0.0027	0.0163	0.3686
300	75	0.1919	0.2610	0.0948	0.2192	0.0941	0.0054	0.0026	0.1311
300	150	0.2417	0.4302	0.0019	0.3132	0.0018	0.0088	0.0000	0.0024
300	225	0.3367	0.3466	0.0000	0.3090	0.0000	0.0076	0.0000	0.0000
500	125	0.2704	0.2911	0.0000	0.4338	0.0000	0.0046	0.0000	0.0000
500	250	0.4289	0.1795	0.0000	0.3879	0.0000	0.0037	0.0000	0.0000
500	375	0.2740	0.3975	0.0000	0.3223	0.0000	0.0062	0.0000	0.0000

Table 3: Posterior model probabilities. Simulated data from model (9) [m_3] using $\gamma_1 = 0.01$, $\gamma_2 = 0.03$, $\phi_1 = \phi_2 = 0.95$, $\sigma_1^2 = 0.009$, and $\sigma_2^2 = 0.009$.

T	T_0	m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8
200	50	0.0013	0.0015	0.3319	0.0014	0.2897	0.0000	0.0111	0.3631
200	100	0.0013	0.0015	0.3352	0.0014	0.2934	0.0000	0.0113	0.3559
200	150	0.0013	0.0015	0.3454	0.0014	0.2936	0.0001	0.0120	0.3448
300	75	0.0011	0.0014	0.3400	0.0011	0.2824	0.0000	0.0023	0.3716
300	150	0.0011	0.0014	0.3493	0.0012	0.2884	0.0000	0.0025	0.3561
300	225	0.0012	0.0013	0.3456	0.0012	0.2921	0.0000	0.0027	0.3560
500	125	0.0243	0.0314	0.2843	0.0269	0.2164	0.0000	0.0001	0.4166
500	250	0.0236	0.0317	0.2879	0.0269	0.2220	0.0000	0.0001	0.4079
500	375	0.0251	0.0320	0.2918	0.0264	0.2218	0.0000	0.0001	0.4028
200	50	0.0013	0.0016	0.3242	0.0014	0.2919	0.0001	0.0169	0.3627
200	100	0.0013	0.0015	0.3246	0.0014	0.2942	0.0001	0.0173	0.3595
200	150	0.0013	0.0016	0.3342	0.0014	0.2935	0.0001	0.0182	0.3498
300	75	0.0011	0.0015	0.3208	0.0013	0.2774	0.0000	0.0038	0.3941
300	150	0.0011	0.0015	0.3327	0.0012	0.2799	0.0000	0.0041	0.3794
300	225	0.0012	0.0014	0.3254	0.0012	0.2870	0.0000	0.0044	0.3794
500	125	0.0241	0.0354	0.2520	0.0302	0.2009	0.0000	0.0002	0.4572
500	250	0.0240	0.0345	0.2523	0.0291	0.2024	0.0000	0.0002	0.4574
500	375	0.0246	0.0370	0.2596	0.0290	0.2045	0.0000	0.0002	0.4450

Table 4: Posterior model probabilities. Simulated data from model (6) [m_4] using $\gamma_1 = \gamma_2 = 0.01$, $\phi_1 = 0.90$, $\phi_2 = 0.95$, $\sigma_1^2 = 0.009$, and $\sigma_2^2 = 0.03$.

T	T_0	m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8
200	50	0.3393	0.0370	0.1335	0.1121	0.2285	0.0183	0.0381	0.0930
200	100	0.7593	0.0388	0.0000	0.1867	0.0000	0.0151	0.0000	0.0000
200	150	0.4046	0.2891	0.0000	0.2818	0.0000	0.0245	0.0000	0.0000
300	75	0.7894	0.0498	0.0001	0.1524	0.0002	0.0081	0.0000	0.0001
300	150	0.4637	0.2170	0.0000	0.3093	0.0000	0.0100	0.0000	0.0000
300	225	0.5528	0.0949	0.0000	0.3486	0.0000	0.0037	0.0000	0.0000
500	125	0.4760	0.2105	0.0000	0.3043	0.0000	0.0092	0.0000	0.0000
500	250	0.6458	0.0563	0.0000	0.2949	0.0000	0.0030	0.0000	0.0000
500	375	0.6328	0.0130	0.0000	0.3538	0.0000	0.0003	0.0000	0.0000
200	50	0.2872	0.0437	0.1350	0.1530	0.2037	0.0238	0.0493	0.1044
200	100	0.6523	0.0503	0.0000	0.2766	0.0000	0.0209	0.0000	0.0000
200	150	0.3711	0.3001	0.0000	0.2991	0.0000	0.0297	0.0000	0.0000
300	75	0.6775	0.0683	0.0001	0.2417	0.0002	0.0121	0.0000	0.0001
300	150	0.3877	0.2449	0.0000	0.3544	0.0000	0.0130	0.0000	0.0000
300	225	0.5141	0.1042	0.0000	0.3771	0.0000	0.0047	0.0000	0.0000
500	125	0.3865	0.2364	0.0000	0.3646	0.0000	0.0125	0.0000	0.0000
500	250	0.5702	0.0673	0.0000	0.3584	0.0000	0.0040	0.0000	0.0000
500	375	0.6045	0.0140	0.0000	0.3811	0.0000	0.0004	0.0000	0.0000

Table 5: Posterior model probabilities. Simulated data from model (8) [m_5] using $\gamma_1 = 0.01$, $\gamma_2 = 0.03$, $\phi_1 = 0.90$, $\phi_2 = 0.95$, $\sigma_1^2 = \sigma_2^2 = 0.009$.

T	T_0	m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8
200	50	0.0014	0.0007	0.2616	0.0011	0.5038	0.0002	0.0434	0.1878
200	100	0.0013	0.0007	0.2575	0.0010	0.5157	0.0002	0.0407	0.1830
200	150	0.0012	0.0010	0.3379	0.0011	0.3829	0.0001	0.0136	0.2622
300	75	0.0011	0.0006	0.2542	0.0008	0.4976	0.0000	0.0149	0.2307
300	150	0.0011	0.0006	0.2645	0.0009	0.4706	0.0000	0.0062	0.2562
300	225	0.0011	0.0007	0.2507	0.0011	0.4719	0.0000	0.0050	0.2696
500	125	0.0010	0.0004	0.2587	0.0009	0.5358	0.0000	0.0049	0.1983
500	250	0.0010	0.0003	0.2455	0.0009	0.5904	0.0000	0.0029	0.1590
500	375	0.0012	0.0002	0.2229	0.0010	0.6703	0.0000	0.0012	0.1032
200	50	0.0013	0.0008	0.2730	0.0013	0.4534	0.0002	0.0564	0.2136
200	100	0.0013	0.0008	0.2694	0.0012	0.4642	0.0002	0.0538	0.2091
200	150	0.0012	0.0011	0.3499	0.0013	0.3393	0.0001	0.0171	0.2900
300	75	0.0010	0.0007	0.2713	0.0010	0.4371	0.0001	0.0208	0.2680
300	150	0.0010	0.0007	0.2780	0.0011	0.4151	0.0000	0.0085	0.2956
300	225	0.0010	0.0008	0.2655	0.0013	0.4160	0.0000	0.0065	0.3089
500	125	0.0010	0.0005	0.2489	0.0010	0.5227	0.0000	0.0065	0.2195
500	250	0.0010	0.0004	0.2374	0.0011	0.5828	0.0000	0.0038	0.1736
500	375	0.0012	0.0002	0.2150	0.0011	0.6745	0.0000	0.0014	0.1066

Table 6: Posterior model probabilities. Simulated data from model (4) [m_6] using $\gamma_1 = \gamma_2 = 0$, $\phi_1 = \phi_2 = 1$, $\sigma_1^2 = 0.009$, and $\sigma_2^2 = 0.03$.

T	T_0	m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8
200	50	0.0765	0.0514	0.1134	0.0264	0.0929	0.2286	0.3254	0.0853
200	100	0.3495	0.0734	0.0000	0.0459	0.0000	0.5312	0.0000	0.0000
200	150	0.1347	0.1765	0.0000	0.1099	0.0000	0.5790	0.0000	0.0000
300	75	0.2207	0.0527	0.0001	0.0156	0.0000	0.7104	0.0004	0.0001
300	150	0.0768	0.1262	0.0000	0.0417	0.0000	0.7553	0.0000	0.0000
300	225	0.0330	0.1570	0.0000	0.0497	0.0000	0.7603	0.0000	0.0000
500	125	0.0413	0.0535	0.0000	0.0068	0.0000	0.8983	0.0000	0.0000
500	250	0.0125	0.0602	0.0000	0.0088	0.0000	0.9185	0.0000	0.0000
500	375	0.0031	0.0751	0.0000	0.0116	0.0000	0.9102	0.0000	0.0000
200	50	0.0819	0.0432	0.0968	0.0373	0.0860	0.2405	0.3424	0.0719
200	100	0.3104	0.0776	0.0000	0.0966	0.0000	0.5154	0.0000	0.0000
200	150	0.1263	0.1584	0.0000	0.0967	0.0000	0.6186	0.0000	0.0000
300	75	0.0676	0.0502	0.0000	0.0426	0.0000	0.8391	0.0005	0.0000
300	150	0.0331	0.0967	0.0000	0.0277	0.0000	0.8426	0.0000	0.0000
300	225	0.0239	0.1100	0.0000	0.0306	0.0000	0.8354	0.0000	0.0000
500	125	0.0136	0.0177	0.0000	0.0078	0.0000	0.9608	0.0000	0.0000
500	250	0.0054	0.0599	0.0000	0.0057	0.0000	0.9290	0.0000	0.0000
500	375	0.0027	0.0361	0.0000	0.0040	0.0000	0.9573	0.0000	0.0000

Table 7: Posterior model probabilities. Simulated data from model (5) [m_7] using $\gamma_1 = \gamma_2 = 0$, $\phi_1 = \phi_2 = 1$, $\sigma_1^2 = \sigma_2^2 = 0.03$.

T	T_0	m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8
200	50	0.0011	0.0011	0.1943	0.0009	0.1634	0.0038	0.4867	0.1486
200	100	0.0010	0.0011	0.1966	0.0009	0.1622	0.0038	0.4831	0.1514
200	150	0.0011	0.0011	0.1855	0.0009	0.1655	0.0038	0.4874	0.1547
300	75	0.0004	0.0005	0.1356	0.0003	0.0863	0.0037	0.6684	0.1047
300	150	0.0005	0.0005	0.1318	0.0002	0.0895	0.0037	0.6681	0.1056
300	225	0.0005	0.0006	0.1345	0.0002	0.0888	0.0037	0.6687	0.1030
500	125	0.0001	0.0003	0.0476	0.0001	0.0235	0.0043	0.8684	0.0558
500	250	0.0001	0.0003	0.0531	0.0001	0.0252	0.0043	0.8647	0.0522
500	375	0.0001	0.0003	0.0510	0.0001	0.0242	0.0043	0.8684	0.0516
200	50	0.0011	0.0010	0.2097	0.0008	0.1973	0.0036	0.4644	0.1221
200	100	0.0011	0.0009	0.2097	0.0008	0.1948	0.0036	0.4631	0.1260
200	150	0.0012	0.0009	0.2018	0.0008	0.1988	0.0036	0.4651	0.1278
300	75	0.0003	0.0004	0.0721	0.0002	0.0441	0.0045	0.8111	0.0674
300	150	0.0002	0.0003	0.0691	0.0002	0.0442	0.0045	0.8132	0.0682
300	225	0.0003	0.0004	0.0718	0.0002	0.0450	0.0045	0.8123	0.0655
500	125	0.0000	0.0000	0.0186	0.0000	0.0071	0.0047	0.9507	0.0188
500	250	0.0000	0.0001	0.0218	0.0000	0.0081	0.0047	0.9463	0.0190
500	375	0.0000	0.0001	0.0204	0.0000	0.0076	0.0047	0.9490	0.0181

Table 8: Posterior probabilities of the AR change-point model (2) for different values of ℓ_1 and ℓ_2, compared to the broken-variance Random walk model (4). Simulated data from model (2) using $\gamma_1 = 0.01$, $\gamma_2 = 0.02$, $\phi_1 = 0.90$, $\phi_2 = 0.95$, $\sigma_1^2 = 0.009$, and $\sigma_2^2 = 0.03$.

		$T = 200$			$T = 300$			$T = 500$		
ℓ_1	ℓ_2	$T_0 = 50$	$T_0 = 100$	$T_0 = 150$	$T_0 = 75$	$T_0 = 150$	$T_0 = 225$	$T_0 = 125$	$T_0 = 250$	$T_0 = 375$
0.5	0.6	0.7968	0.9137	0.7593	0.9340	0.8621	0.9493	0.8547	0.9568	0.9956
0.5	0.7	0.8291	0.9343	0.8131	0.9474	0.8845	0.9629	0.8808	0.9703	0.9969
0.5	0.8	0.8830	0.9563	0.8500	0.9640	0.9183	0.9703	0.9166	0.9772	0.9979
0.5	0.9	0.9190	0.9685	0.8674	0.9775	0.9450	0.9829	0.9464	0.9882	0.9989
0.6	0.6	0.8096	0.9252	0.7955	0.9420	0.8794	0.9548	0.8748	0.9694	0.9965
0.6	0.7	0.8574	0.9430	0.8403	0.9529	0.8971	0.9651	0.8973	0.9735	0.9973
0.6	0.8	0.8986	0.9615	0.8685	0.9699	0.9270	0.9758	0.9249	0.9826	0.9982
0.6	0.9	0.9306	0.9734	0.8893	0.9814	0.9516	0.9861	0.9537	0.9893	0.9990
0.7	0.6	0.8404	0.9330	0.8346	0.9521	0.8962	0.9621	0.8963	0.9726	0.9969
0.7	0.7	0.8771	0.9485	0.8617	0.9636	0.9153	0.9737	0.9142	0.9793	0.9978
0.7	0.8	0.9093	0.9667	0.8919	0.9738	0.9372	0.9797	0.9334	0.9838	0.9984
0.7	0.9	0.9389	0.9762	0.9066	0.9843	0.9600	0.9885	0.9618	0.9917	0.9992
0.8	0.6	0.8025	0.9169	0.8609	0.9457	0.9215	0.9729	0.9225	0.9770	0.9976
0.8	0.7	0.8434	0.9356	0.8885	0.9586	0.9351	0.9764	0.9332	0.9824	0.9982
0.8	0.8	0.8874	0.9558	0.9137	0.9706	0.9512	0.9842	0.9497	0.9877	0.9987
0.8	0.9	0.9224	0.9689	0.9259	0.9826	0.9675	0.9908	0.9686	0.9931	0.9993

Table 9: Posterior probabilities of the broken-variance Random walk model (4) for different values of ℓ_1 and ℓ_2, compared to the AR change-point model (2). Simulated data from model (4) using $\gamma_1 = \gamma_2 = 0$, $\phi_1 = \phi_2 = 1$, $\sigma_1^2 = 0.009$, and $\sigma_2^2 = 0.03$.

		$T = 200$			$T = 300$			$T = 500$		
ℓ_1	ℓ_2	$T_0 = 50$	$T_0 = 100$	$T_0 = 150$	$T_0 = 75$	$T_0 = 150$	$T_0 = 225$	$T_0 = 125$	$T_0 = 250$	$T_0 = 375$
0.6	0.6	0.9878	0.9915	0.9887	0.9969	0.9976	0.9986	0.9995	0.9997	0.9996
0.6	0.7	0.9817	0.9870	0.9836	0.9960	0.9964	0.9982	0.9988	0.9998	0.9999
0.6	0.8	0.9720	0.9783	0.9739	0.9938	0.9955	0.9971	0.9981	0.9997	0.9998
0.6	0.9	0.9324	0.9216	0.9453	0.9880	0.9899	0.9934	0.9965	0.9991	0.9995
0.7	0.6	0.9809	0.9854	0.9858	0.9968	0.9975	0.9981	0.9981	0.9998	0.9999
0.7	0.7	0.9729	0.9784	0.9770	0.9933	0.9955	0.9968	0.9984	0.9993	0.9998
0.7	0.8	0.9576	0.9628	0.9640	0.9911	0.9938	0.9950	0.9966	0.9990	0.9997
0.7	0.9	0.9029	0.8777	0.9257	0.9771	0.9847	0.9922	0.9933	0.9983	0.9992
0.8	0.6	0.9717	0.9738	0.9817	0.9896	0.9952	0.9972	0.9976	0.9993	0.9995
0.8	0.7	0.9582	0.9596	0.9697	0.9881	0.9936	0.9962	0.9976	0.9991	0.9996
0.8	0.8	0.9350	0.9323	0.9464	0.9816	0.9918	0.9934	0.9976	0.9986	0.9994
0.8	0.9	0.8533	0.7879	0.9076	0.9589	0.9781	0.9867	0.9913	0.9980	0.9991
0.9	0.6	0.9441	0.9339	0.9609	0.9812	0.9894	0.9923	0.9966	0.9979	0.9993
0.9	0.7	0.9258	0.9132	0.9426	0.9795	0.9902	0.9920	0.9961	0.9986	0.9993
0.9	0.8	0.8885	0.8609	0.9137	0.9672	0.9843	0.9882	0.9941	0.9976	0.9987
0.9	0.9	0.7491	0.6203	0.8339	0.9257	0.9581	0.9715	0.9869	0.9954	0.9980

Table 10a: Interest Rates. Posterior model probabilities.

	m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8
One-month	0.1722	0.4606	0.0000	0.2211	0.0000	0.1461	0.0000	0.0000
One-year	0.2626	0.3086	0.0000	0.2132	0.0000	0.2156	0.0000	0.0000

Table 10b: Interest Rates. Posterior means and standard errors (in parenthesis) of the parameters of the most probable model $m_2 : y_t = \gamma(1-\phi) + \phi y_{t-1} + \sigma_{s_t}\epsilon_t$ (common γ -common ϕ)

	T_0 (mode)	T_0 (mean)	γ	σ_1^2	σ_2^2	ϕ
One-month	1982:09	1982:11	7.209	0.953	0.271	0.953
		(4.3)	(0.881)	(0.121)	(0.061)	(0.023)
One-year	T_0 (mode)	T_0 (mean)	γ	σ_1^2	σ_2^2	ϕ
	1982:11	1983:09	7.669	0.681	0.149	0.961
		(11.5)	(0.915)	(0.082)	(0.029)	(0.024)

Table 11a: Real Exchange Rates. Posterior model probabilities.

	m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8
United Kingdom	0.0000	0.0000	0.0754	0.0000	0.3860	0.0000	0.4308	0.1078
Denmark	0.0244	0.0609	0.0151	0.0260	0.4272	0.3321	0.0990	0.0153
France	0.0090	0.0180	0.0356	0.0084	0.5640	0.0996	0.2265	0.0389
Netherlands	0.0020	0.0030	0.0337	0.0019	0.7033	0.0161	0.2030	0.0369

Table 11b: Real Exchange Rates. Posterior means and standard errors (in parenthesis) of the parameters of the most probable model $m_5 : y_t = (\gamma_1 + \delta_t)(1-\phi_{s_t}) + \phi_{s_t} y_{t-1} + \sigma\epsilon_t$ (common σ^2)

	T_0 (mode)	T_0 (mean)	γ_1	γ_2	σ^2	ϕ_1	ϕ_2
United Kingdom	1985:01	1985:03	0.459	0.357	0.0002	0.984	0.879
		(28.9)	(0.029)	(0.006)	(0.00002)	(0.018)	(0.033)
Denmark	1985:01	1985:02	7.384	5.535	0.035	0.991	0.912
		(8.3)	(0.548)	(0.073)	(0.003)	(0.007)	(0.018)
France	1985:01	1985:01	5.718	4.585	0.021	0.991	0.907
		(16.8)	(0.458)	(0.096)	(0.002)	(0.014)	(0.021)
Netherlands	1985:01	1985:02	3.239	2.603	0.008	0.992	0.905
		(11.9)	(0.256)	(0.036)	(0.0006)	(0.0097)	(0.021)

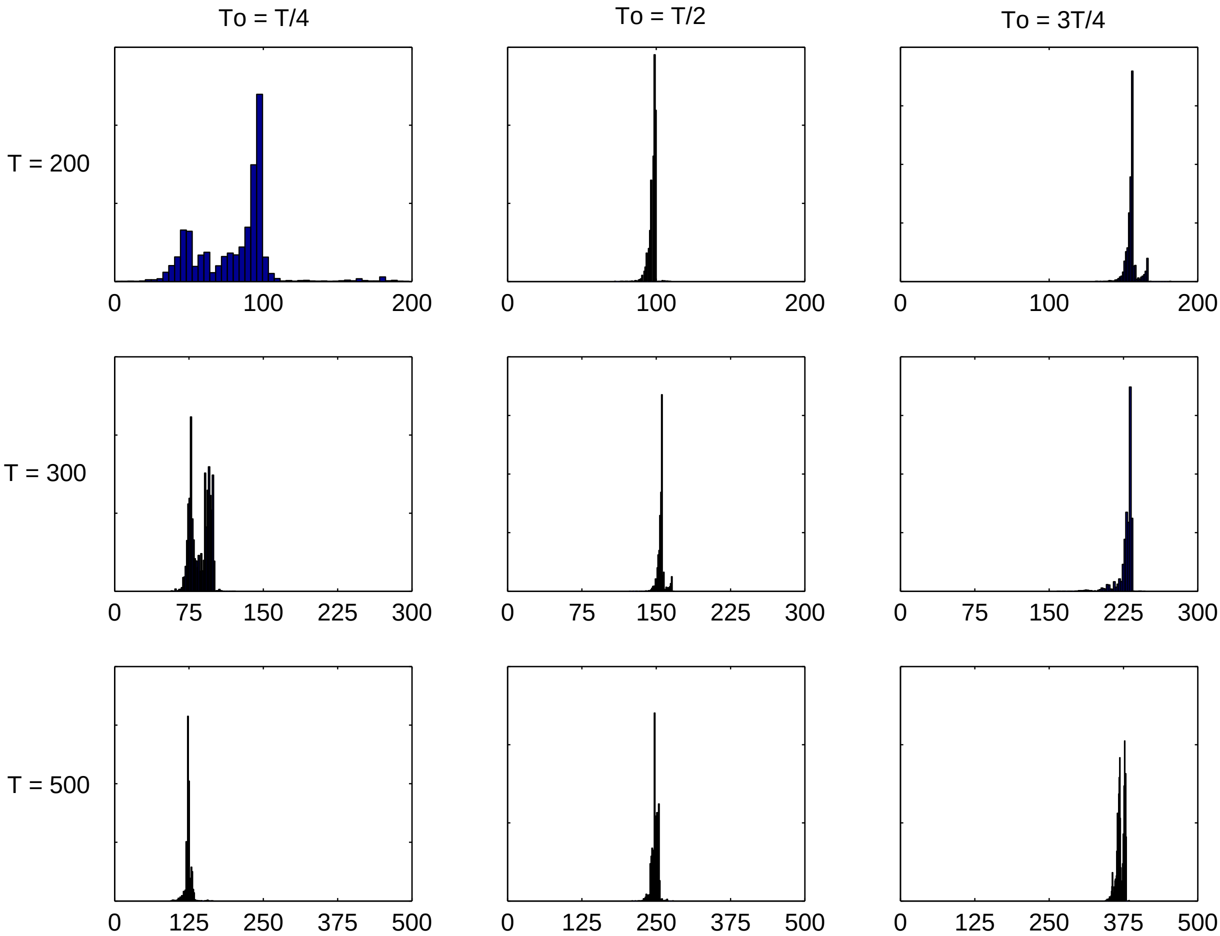

Figure 1: Histograms of the posterior distributions of the break-point T_0 for simulated data of size $T = \{200, 300, 500\}$ from the general model m_1 with true values $T_0 = \{T/4, T/2, 3T/4\}$.

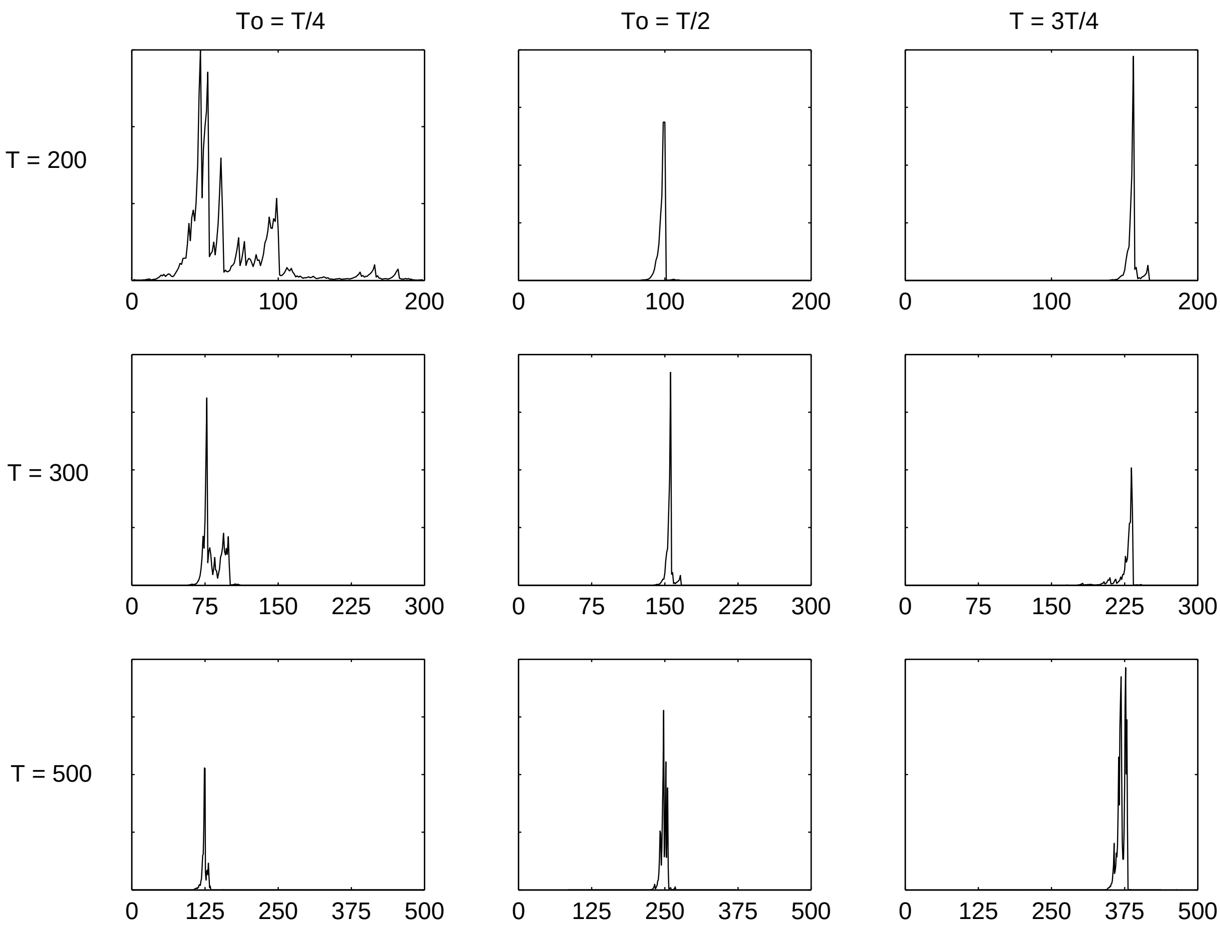

Figure 2: Exact marginal posterior distributions of the break-point T_0 for simulated data of size $T = \{200, 300, 500\}$ from the random walk model with broken variance, m_6, with true value $T_0 = \{T/4, T/2, 3T/4\}$.

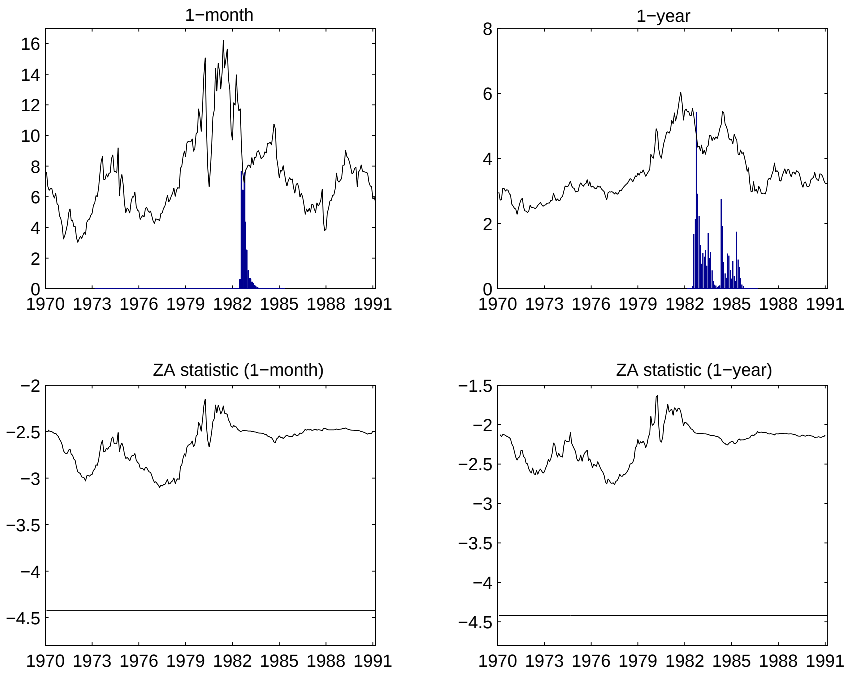

Figure 3: Real interest rates series with posterior distribution of the break date and ZA test statistic.

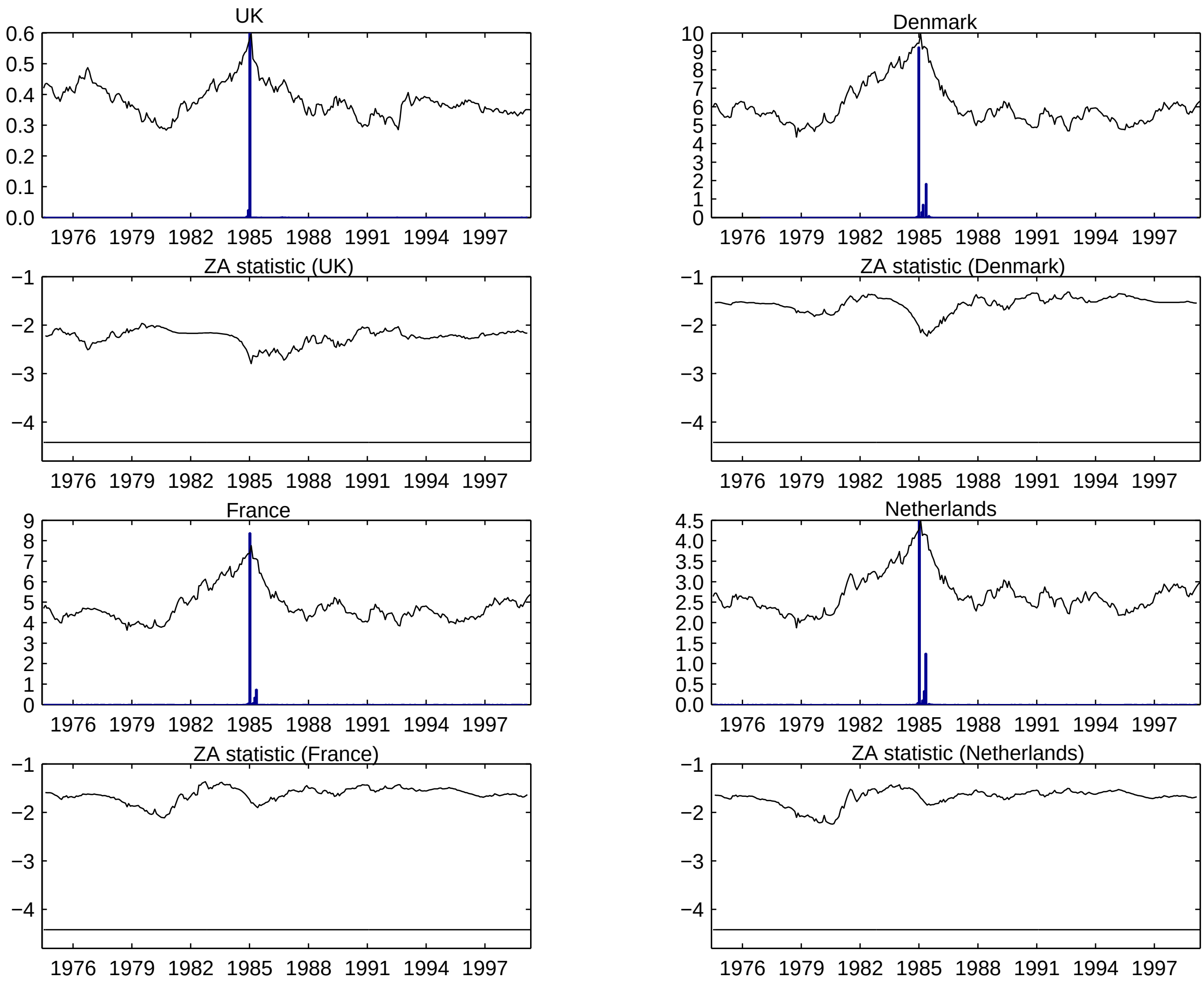

Figure 4: Real exchange rates series with posterior distribution of the break date and ZA test statistic.

Queen Mary
University of London

This working paper has been produced by
the Department of Economics at
Queen Mary, University of London

Department of Economics
Queen Mary, University of London
Mile End Road
London E1 4NS
Tel: +44 (0)20 7882 5096
Fax: +44 (0)20 8983 3580
Website: www.econ.qmul.ac.uk/papers/wp.htm