Bayes Estimates of Markov Trends in Possibly Cointegrated Series: An Application to US Consumption and Income

Richard Paap*
Rotterdam Institute for Business Economic Studies
Erasmus University Rotterdam

and

Herman K. van Dijk
Econometric Institute
Erasmus University Rotterdam

Abstract

Stylized facts show that the average growth rates of US per capita consumption and income differ in recession and expansion periods. Since a linear combination of such series does not have to be a constant mean process, standard cointegration analysis between the variables, to examine the permanent income hypothesis, may not be valid. To model the changing growth rates in both series, we introduce a multivariate Markov trend model, which allows for different growth rates in consumption and income during expansions and recessions. The deviations from the multivariate Markov trend are modelled by a vector autoregressive model. Bayes estimates of this model are obtained using Markov chain Monte Carlo methods. The empirical results suggest that there exist a cointegration relation between US per capita disposable income and consumption, after correction for a multivariate Markov trend.

Key words: multivariate Markov trend, cointegration, MCMC, permanent income hypothesis

*Part of this research was sponsored by the Netherlands Foundation for Scientific Research (N.W.O.). We thank Frank Kleibergen, Peter Boswijk and Philip Hans Franses for helpful comments on an earlier version of this paper. Correspondence to: Richard Paap, RIBES, Erasmus University Rotterdam, P.O. Box 1738, NL-3000 DR Rotterdam, The Netherlands, e-mail: paap@few.eur.nl

1 Introduction

The permanent income hypothesis implies that there exists a long run relation between consumption and disposable income, see *e.g.* Hall (1978). If we translate this theoretical result to time series properties, it implies that a linear combination of consumption and disposable income series has to be a stationary process. Since most studies on the univariate properties of consumption and income series suggest that they are integrated processes, see for instance Dickey and Fuller (1979), both series have to be cointegrated for the permanent income hypothesis to hold. As a result, recent empirical research on the permanent income hypothesis focuses on cointegration analysis between consumption and income, see Campbell (1987) and Jin (1995) among others.

In these studies it is usually assumed that the logarithm of real income is a linear process. However, Goodwin (1993), Potter (1995) and Peel and Speight (1998) among others argue that the log of many real income series contain a nonlinear cycle. This nonlinear cycle is often interpreted as the business cycle in real income. A popular model to describe the business cycle in time series is the Markov switching model of Hamilton (1989). This model allows for different average growth rates in income during expansion and recession periods, where the transitions between expansions and recessions and *vice versa* are modeled by an unobserved first-order Markov process. We will refer to the trend that models this specific behavior as a Markov trend. Hall *et al.* (1997) consider the permanent income hypothesis, while they assume that real income contains a Markov trend. They show that in that case the difference between log consumption and log real income is affected by changes in the mean, caused by the changes in the growth rate of the real income series. The difference between the log consumption and income series is not a constant mean process and standard cointegration analysis in linear vector autoregressive models may indicate incorrectly the absence of cointegration.

In this paper, we analyze the long run relationship between quarterly seasonally adjusted aggregate consumption and disposable income for the United States, where we allow for the possibility of a Markov trend in the income series. Our paper differs from previous studies in several ways. We consider a full system cointegration analysis. Cointegration is tested in a vector autoregression, which models the deviation of log per capita consumption and income from a multivariate Markov trend. This differs from Hall *et al.* (1997), who consider a single equation analysis and use an *ad hoc* procedure for cointegration analysis. Our model is a multivariate generalization of Hamilton's (1989) model and nests the theoretical results in Hall *et al.* (1997). Furthermore, the model allows that the growth rate of consumption may be different than the growth rate in income in each stage of the business cycle as suggested by a simple stylized facts analysis. Hence, the analysis for the presence of a cointegration relation between the consumption and income series is done, while we allow for different growth rates in expansions and recession periods via the multivariate Markov trend. Finally, we use a Bayesian point of view to analyze the presence of a stable long run relation between per capita consumption and income. We use Markov chain Monte Carlo methods to evaluate posterior distributions and construct Bayes factors to determine the cointegration rank. This Bayesian cointegration analysis is based on Kleibergen and Paap (1998).

The outline of this paper is as follows. In Section 2 we give a short review of the permanent income hypothesis in case income contains a Markov trend. In Section 3 we discuss some stylized facts of US per capita income and consumption series. In Section 4 we propose the multivariate Markov trend model and discuss its interpretation. Section 5 deals with prior specification. To obtain posterior results, we propose in Section 6 a Markov chain Monte Carlo algorithm to sample from the posterior distribution. Section 7 deals with Bayes factors to determine the cointegration rank. In Section 8 we apply our multivariate Markov model on the US series and relate the posterior results to suggestions made by economic theory and the stylized facts analysis. We conclude in Section 9.

2 Permanent Income Hypothesis and a Markov Trend

The permanent income hypothesis states that current aggregate consumption c_t can be written as

$$c_t = \frac{r}{1+r} \sum_{j=0}^{\infty} \frac{1}{(1+r)^j} \mathrm{E}[y_{t+j} | \Omega_t], \tag{1}$$

where y_t is real disposable income, r is the interest rate and Ω_t denotes the information set that is available to economic agents at time t. Straightforward algebra shows that (1) is the forward solution of the following expectational difference equation

$$c_t = \frac{r}{1+r}\mathrm{E}[y_t|\Omega_t] + \frac{1}{1+r}\mathrm{E}[c_{t+1}|\Omega_t]. \tag{2}$$

In most cases one assumes that the logarithm of real income is a random walk process. This assumption and (2) imply the existence of a stationary relation between current log consumption and income, see for instance Campbell (1987). We will however proceed in a similar way as Hall *et al.* (1997), who assume that the log of real income contains a Markov trend as suggested by Hamilton (1989). The Markov trend is a stochastic segmented trend with two slopes, which model the different growth rates in the stages of the business cycle. The direction of the slope in every period depends on the value of an unobserved first-order two-state Markov process. In other words, the logarithm of real income is written as

$$\ln(y_t) = n_t + z_t, \tag{3}$$

where n_t is a so-called univariate Markov trend and z_t models the deviation from the trend. The Markov trend n_t is defined as

$$n_t = n_{t-1} + \gamma_0 + \gamma_1 s_t, \tag{4}$$

where γ_0 and γ_1 are parameters and s_t is a binary random variable, which follows an unobserved first-order Markov process with transition probabilities

$$\begin{array}{ll} \Pr[s_t = 0|s_{t-1} = 0] = p, & \Pr[s_t = 1|s_{t-1} = 0] = 1 - p, \\ \Pr[s_t = 1|s_{t-1} = 1] = q, & \Pr[s_t = 0|s_{t-1} = 1] = 1 - q. \end{array} \tag{5}$$

The deviations from the trend are usually assumed to be an integrated autoregressive [AR] model, see Hamilton (1989). Here we assume for simplicity that z_t is a random walk process

$$z_t = z_{t-1} + \epsilon_t, \tag{6}$$

where $\epsilon_t \sim \mathrm{NID}(0, \sigma^2)$ such that the growth rate in real income at time t equals $\gamma_0 + \gamma_1 s_t + \epsilon_t$.

As Hall *et al.* (1997) show equations (2) and (3) with (4) imply that for $s_t = 0$, $c_t = e^{\kappa_0} y_t$ and that for $s_t = 1$ we have the relation $c_t = e^{\kappa_0+\kappa_1} y_t$ with

$$\begin{aligned} \kappa_0 &= \ln\left(\frac{r + qe^{\kappa_0}E_0 + (1-q)e^{\kappa_0+\kappa_1}E_1}{1+r}\right) \\ \kappa_0 + \kappa_1 &= \ln\left(\frac{r + (1-p)e^{\kappa_0}E_0 + pe^{\kappa_0+\kappa_1}E_1}{1+r}\right), \end{aligned} \tag{7}$$

and where $E_0 = e^{\gamma_0+\frac{1}{2}\sigma}$ and $E_1 = e^{\gamma_0+\gamma_1+\frac{1}{2}\sigma}$. As s_t is an unobserved process, we have the following relation between log consumption and log income

$$\ln(c_t) = \kappa_0 + \kappa_1 s_t + \ln(y_t), \tag{8}$$

where κ_0 and κ_1 follow from the solution of (7)

$$\begin{aligned} \kappa_0 &= \ln\left(\frac{r(1 + (1-p-q)(1+r)^{-1}E_1)}{(1 + r - qE_0 - pE_1) - (1+r)^{-1}(1-p-q)E_0E_1}\right) \\ \kappa_1 &= \ln\left(\frac{(1+r) + (1-p-q)E_0}{(1+r) + (1-p-q)E_1}\right). \end{aligned} \tag{9}$$

The consumption-income relation (8) implies that the log of consumption can be written as

$$\ln(c_t) = n_t + \kappa_0 + \kappa_1 s_t + z_t, \tag{10}$$

where n_t and z_t are defined in (4) and (6), respectively. Equation (10) shows that log consumption has the same Markov trend as log income and hence the growth rates of consumption and income are the same during expansions and recessions. However, the difference between log consumption and income, given by $\kappa_0 + \kappa_1 s_t$, is different during expansions and recessions, see also (8).

In Section 4 we propose a multivariate Markov trend model, which allows us to test the validity of the permanent income hypothesis, when the log of real income contains a Markov trend. Since the economic theory in this section may be too simplistic in describing reality, we allow for a more flexible structure than the theory suggests. This flexible structure will be based on a simple stylized facts analysis of the US per capita income and consumption series in the next section.

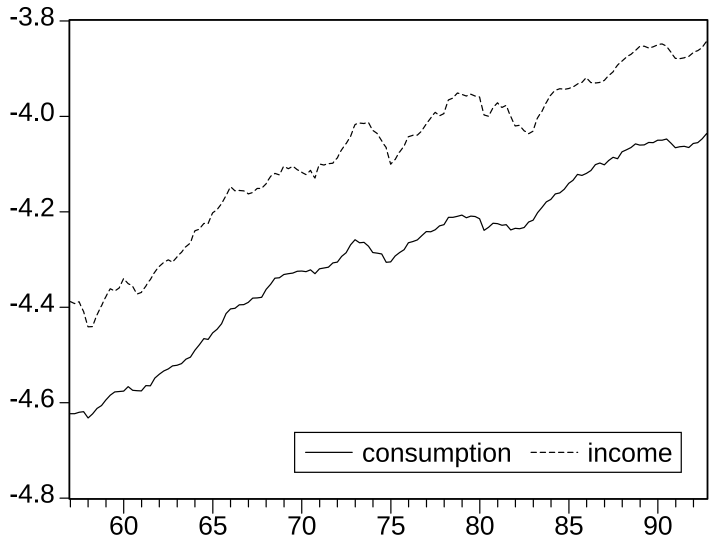

Figure 1: The logarithm of US per capita consumption and income, 1957.1–1992.4.

3 Stylized Facts

Figure 1 shows a plot of the logarithm of quarterly observed seasonally adjusted per capita real disposable income and private consumption of the United States, 1957.1–1992.4. The series are obtained from Citibase. Both series are increasing over the sample period with short periods of decline, for instance in the middle and the end of the 1970s. These periods of decline are more pronounced in the income series than in the consumption series but seem to occur roughly simultaneously. The average quarterly growth rate of the income series is 0.38% per quarter. For the consumption series the average quarterly growth rates equals 0.41%. A naive likelihood ratio [LR] test statistic for equal average growth rates is not significant at the 5% level[1], which indicates that the growth rates in both series are roughly the same.

To analyze the effect of the business cycle on real per capita income and consumption we split the sample in two parts. The first part corresponds to quarters which are labelled as a recession according to the NBER peaks and troughs, see the final two columns of

[1]The LR test for equal growth rates is based on the assumption that the growth rates of per capita income and consumption are bivariate normally distributed with non-zero mean and a general covariance matrix. The test is naive since we do not correct for serial correlation in the growth rates.

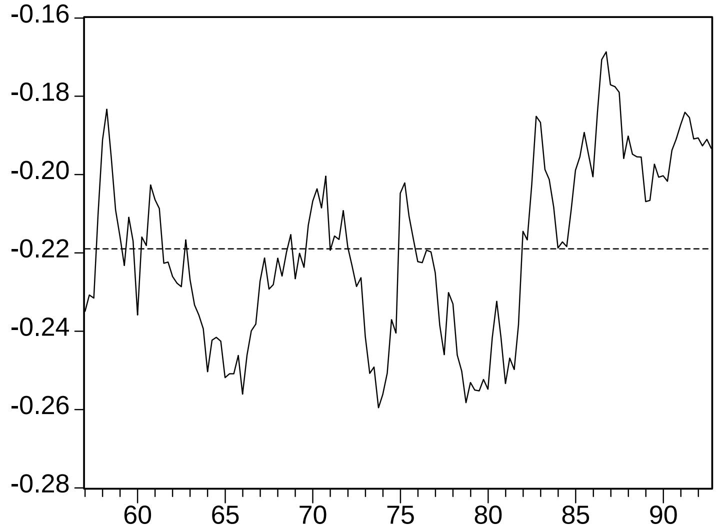

Figure 2: Difference between log US per capita income and consumption, 1957.I–1992.IV.

Table 1. The average quarterly growth rates of per capita income during recessions equals -1.25%, while for consumption the average growth rate equals -0.34%. The naive classical LR test statistic for equal quarterly growth rates in per capita income during recessions is significant at the 5% level and hence it suggest different growth rates in per capita consumption and income during recessions. The second part contains quarters, which corresponds to expansions. During expansions, the average quarterly growth rate in per capita income is 0.71%, while the average quarterly growth rate in per capita consumption is 0.56%. Again, the naive LR classical test statistic for equal quarterly growth rates is significant at the 5% level and hence the growth rates during expansion periods seem to be different.

The differences in the average growth rates in the consumption and income series in recessions and expansions may have consequences for analyzing the permanent income hypothesis. A simple cointegration analysis in a linear (vector) autoregressive model as for instance in Jin (1995) may lead to the wrong conclusion. If the growth rates in both series are different in both stages of the business cycle it is unlikely that a linear combination of the two series has a constant mean. To make this more clear we consider in Figure 2 the difference between the logarithm of per capita consumption and income. The graph shows that the mean of this difference is not constant over time but displays a more or less changing regime pattern. This switching patterns seems to coincide with the business cycle. The mean of the difference between per capita consumption and income during expansions equals -0.22. During recessions it is 0.006 larger. A naive classical F-test statistic for equal means during recessions and expansions equals turns out to be 1.60, which is not significant at the 5% level[2].

This result suggests that a switch in the constant of the consumption-income relation (8) may not be relevant. Furthermore, relation (8) also implies that the growth rate in per capita consumption and income during recessions and expansions have to be the same, which contradicts our earlier test results. A consumption-income relation, which allows for different growth rates in consumption and income during expansions and recessions is given by

$$\ln(c_t) = \kappa_0 + \kappa_1 s_t + \beta_2 \ln(y_t). \tag{11}$$

This implies that the trend in consumption equals $\beta_2 n_t$, where n_t is the trend in log income defined in (4). If $\beta_2 < 0$ the growth rate in consumption during expansions is smaller than in income, while during recessions it is larger, which corresponds to our earlier findings. Note that (11) corresponds to a nonlinear relation between consumption and income $c_t = e^{\kappa_0 + \kappa_1 s_t} y_t^{\beta_2}$.

To analyze the permanent income hypothesis for the US consumption and income series, we propose in the next section a multivariate Markov trend model. This multivariate model is an extension of Hamilton's univariate model, see also Paap (1997, Chapters 5 and 7). The models contains a multivariate Markov trend, which allows for different growth rates in the consumption and income series during recessions and expansions. The deviations from the Markov trend are modelled by a vector autoregressive model. To analyze the presence of a consumption-income relation, we perform a cointegration analysis on these deviations from the multivariate Markov trend. Additionally, we check whether the mean of the possible cointegration relation is affected by changes in the business cycle as suggested by the economic theory in Section 2.

[2]Note that this test is only valid if the difference between the logarithm of per capita consumption and income is stationary.

4 The Multivariate Markov Trend Model

In this section we propose the multivariate Markov trend model. In Section 4.1 we discuss representation, while in Section 4.2 we deal with model interpretation. In Section 4.3 we derive the likelihood function of the model.

4.1 Representation

Let $\{Y_t\}_{t=1}^{T}$ denote a 2-dimensional time series containing the log of the per capita consumption and income series. Assume that $Y_t = (\ln(c_t) \;\; \ln(y_t))'$ can be decomposed as

$$Y_t = N_t + R_t + Z_t, \tag{12}$$

where N_t represents a trend component, R_t allows for possible level shifts and Z_t represents the deviations from the trend component N_t, and R_t. The 2-dimensional trend component N_t is a multivariate generalization of the univariate Markov trend (4)

$$N_t = N_{t-1} + ,_0 + ,_1 s_t, \tag{13}$$

where $,_0$ and $,_1$ are (2×1) parameter vectors, s_t is an unobserved first-order Markov process with transition probabilities given in (5). The value of the unobserved state variable s_t determines the stage of the business cycle. If $s_t = 0$ the slope of the Markov trend is $,_0$, while for $s_t = 1$ the slope equals $,_0 + ,_1$, see also Hamilton (1989). In this paper we choose $s_t = 1$ to correspond to a recession and $s_t = 0$ to an expansion. The values of the slopes of the trends in the individual series in Y_t do not have to be the same although the changes in the value of the slope occur at the same time. Note that different slope values for each series do not have to lead to divergent Markov trends in the long run. The expected slope value of the Markov trend equals $,_0 + ,_1(1-p)/(2-p-q)$, see *e.g.* Hamilton (1989). Hence, two series can have different slopes values in each regimes but with the same expected slope value. The backward solution of (13) equals

$$N_t = ,_0 (t-1) + ,_1 \sum_{i=2}^{t} s_i + N_1, \tag{14}$$

where N_1 denotes the initial value of the Markov trend, which is independent of t. Hence, the Markov trend consists of a deterministic trend with slope $,_0$ and a stochastic trend $\sum_{i=2}^{t} s_i$ with impact vector $,_1$.

The component R_t models possible level shifts in the first element of Y_t during recessions

$$R_t = \begin{pmatrix} \delta_1 \\ 0 \end{pmatrix} s_t = \delta s_t, \tag{15}$$

such that $\delta = (\delta_1\ 0)'$. This term takes care of level shifts in the consumptions series during recessions as suggested by the theory in Section 2, see also Krolzig (1997, Chapter 13) for a similar discussion about the role of this term. At the end of this section we show how δ_1 is related to the κ_1 parameter in (8).

The deviations from the Markov trend and R_t, *i.e.* Z_t are assumed to be an vector autoregressive process of order k [VAR(k)]

$$Z_t = \sum_{i=1}^{k} \Phi_i Z_{t-i} + \varepsilon_t, \tag{16}$$

or using the lag polynomial $\Phi(L) = (\mathbf{I} - \Phi_1 L - \cdots - \Phi_k L^k)$

$$(\mathbf{I} - \Phi_1 L - \cdots - \Phi_k L^k) Z_t = \varepsilon_t, \tag{17}$$

where ε_t is a 2-dimensional vector normally distributed process with zero mean and (2×2) positive definite symmetric covariance matrix Σ, and Φ_i, $i = 1, \ldots, k$, are (2×2) parameter matrices.

4.2 Model Interpretation

For our analysis it is convenient to write (16) in error correction form

$$\Delta Z_t = \Pi Z_{t-1} + \sum_{j=1}^{k-1} \bar{\Phi}_j \Delta Z_{t-j} + \varepsilon_t, \tag{18}$$

where $\Pi = \sum_{j=1}^{k} \Phi_j - \mathbf{I}_2$ and $\bar{\Phi}_i = -\sum_{j=i+1}^{k} \Phi_j$, $i = 1, \ldots, k-1$. The characteristic equation of the Z_t process is given by

$$|\mathbf{I} - \Phi_1 z - \cdots - \Phi_k z^k| = 0. \tag{19}$$

If the roots of (19) are outside the unit circle the process Z_t is stationary and hence

Y_t is a stationary process around a multivariate Markov trend. We can write

$$(\Delta Y_t - ,_0 - ,_1 s_t - \delta \Delta s_t) = \Pi(Y_{t-1} - ,_0(t-2) - ,_1 \sum_{i=2}^{t-1} s_i - N_1 - \delta s_{t-1}) +$$
$$\sum_{i=1}^{k-1} \bar{\Phi}_i(\Delta Y_{t-i} - ,_0 - ,_1 s_{t-i} - \delta \Delta s_{t-i}) + \varepsilon_t, \quad (20)$$

with Π a full rank matrix. The vectors $,_0$ and $,_0 + ,_1$ contain the slopes of the trend in Y_t during expansions and recessions, respectively. The initial value of the Markov trend N_1 is unknown and plays the role of an intercept parameter vector. The δ_1 parameter models a level shift in the intercept of the Markov trend during recessions for the log consumption series. If $s_t = 0$ the initial value of the Markov trend equals N_1, while for $s_t = 1$ this value equals $N_1 + \delta s_t$.

Stochastic trends enter Z_t and therefore Y_t if at least one of the roots of (19) is on the unit circle. Since $\Phi(1) = -\Pi$ by definition, unit roots and therefore the presence of stochastic trends imply rank reduction in Π, see Johansen (1995) for an introduction into cointegration.

We first consider the case of two unit roots in (19) with the remaining roots outside the unit circle. In that case $\Pi = \mathbf{0}$ and (20) becomes

$$(\Delta Y_t - ,_0 - ,_1 s_t - \delta \Delta s_t) = \sum_{i=1}^{k-1} \bar{\Phi}_i(\Delta Y_{t-i} - ,_0 - ,_1 s_{t-i} - \delta \Delta s_{t-i}) + \varepsilon_t. \quad (21)$$

The first difference of Y_t is a stationary VAR process with a stochastically changing mean $(= ,_0 + ,_1 s_t)$. Note that the initial value of the Markov trend N_1 drops out of the model. If $s_t = s_{t-1}$, ΔY_t is not affected by R_t. However if $s_t \neq s_{t-1}$ the growth rate in consumption is δ_1 larger or smaller than the growth rate in income, so a change in the stage of the business cycle leads to a one time extra adjustment in the growth rate of per capita consumption. Of course this adjustment is absent if $\delta_1 = 0$.

If the series in Z_t are cointegrated, only one of the roots equals unity. The series in Z_t contain a common stochastic trend. Since in that case the rank of Π is one, we can write Π as $\alpha\beta'$, where α and β are (2×1) vectors. The vector β represents the cointegration relation between the elements of Z_t and hence $\beta' Z_t$ is a stationary process. The α vector contains the adjustment parameters. Since the number of free parameters in α and β is larger than in Π under rank reduction, the parameters in α or β have to be restricted to become estimable. We choose here for the following restriction: $\beta = (1 \ -\beta_2)'$. Under the cointegration specification the model becomes

$$(\Delta Y_t - {,}\ _0 - {,}\ _1 s_t - \delta \Delta s_t) = \alpha\beta'(Y_{t-1} - {,}\ _0(t-2) - {,}\ _1 \sum_{i=2}^{t-1} s_i - N_1 - \delta s_t) + \\ \sum_{i=1}^{k-1} \bar{\Phi}_i(\Delta Y_{t-i} - {,}\ _0 - {,}\ _1 s_{t-i} - \delta \Delta s_{t-i}) + \varepsilon_t. \quad (22)$$

Under certain parameter restrictions the cointegration relation $\beta' Y_t = \beta'(N_t + R_t + Z_t)$ corresponds to the consumption-income relations given in (8) and (11). For $\beta'{,}\ _0 = \beta'{,}\ _1 = \mathbf{0}$, $\kappa_0 = \beta' N_1$ and $\kappa_1 = \beta'\delta$ we obtain the relation (11). The extra condition $\beta_2 = 1$ leads to relation (8). Finally note that the restriction $\beta'{,}\ _1 = 0$ removes the Markov trend from the cointegration relation. Dwyer and Potter (1996) refer to this phenomenon as reduced rank Markov trend cointegration. Note that in their model $\delta_1 = 0$.

4.3 The Likelihood Function

To analyze the multivariate Markov trend model we derive the likelihood function. First, we consider the likelihood function least restricted Markov trend stationary model (20) conditional on the states s_t. The conditional density of Y_t for this model given the past and current states $s^t = \{s_1, \ldots, s_t\}$ and given the past observations $Y^{t-1} = \{Y_1, \ldots, Y_{t-1}\}$ is given by

$$f(Y_t|Y^{t-1}, s^t, {,}\ _0, {,}\ _1, N_1, \delta_1, \Sigma, \Pi, \bar{\Phi}) = \frac{1}{(\sqrt{2\pi})^2} |\Sigma|^{-\frac{1}{2}} \exp(-\frac{1}{2}\varepsilon_t' \Sigma^{-1} \varepsilon_t), \quad (23)$$

where ε_t is given in (20) and $\bar{\Phi} = \{\bar{\Phi}_1, \ldots, \bar{\Phi}_{k-1}\}$. Hence the likelihood function for model (20) conditional on the states s^T and the first k initial observations Y^k equals

$$\mathcal{L}_2(Y^T|Y^k, s^T, \Theta_2) = p^{\mathcal{N}_{0,0}} (1-p)^{\mathcal{N}_{0,1}} q^{\mathcal{N}_{1,1}} (1-q)^{\mathcal{N}_{1,0}} \\ \prod_{t=k+1}^{T} f(Y_t|Y^{t-1}, s^t, {,}\ _0, {,}\ _1, N_1, \Sigma, \Pi, \bar{\Phi}), \quad (24)$$

where $\Theta_2 = \{{,}\ _0, {,}\ _1, N_1, \delta_1, \Sigma, \Pi, \bar{\Phi}, p, q\}$ and where $\mathcal{N}_{i,j}$ denotes the number of transitions from state i to state j. The unconditional likelihood function $\mathcal{L}_2(Y^T|Y^k, \Theta_2)$ can be obtained by summing over all possible realizations of s^T

$$\mathcal{L}_2(Y^T|Y^k, \Theta_2) = \sum_{s_1}\sum_{s_2}\cdots\sum_{s_T}\mathcal{L}_2(Y^T|Y^k, s^T, \Theta_2). \tag{25}$$

The unconditional likelihood function for the Markov trend model with one cointegration relation (22) denoted by $\mathcal{L}_1$ and without cointegration (21) denoted by $\mathcal{L}_2$ follow directly from (25)

$$\begin{aligned} \mathcal{L}_1(Y^T|Y^k, \Theta_1) &= \mathcal{L}_2(Y^T|Y^k, \Theta_2)|_{\Pi=\alpha\beta'} \\ \mathcal{L}_0(Y^T|Y^k, \Theta_0) &= \mathcal{L}_2(Y^T|Y^k, \Theta_2)|_{\Pi=\mathbf{0}} \end{aligned} \tag{26}$$

with $\Theta_1 = \{, {}_0, , {}_1, N_1, \delta_1, \Sigma, \alpha, \beta_2, \bar{\Phi}, p, q\}$ and $\Theta_0 = \{, {}_0, , {}_1, N_1, \delta_1, \Sigma, \bar{\Phi}, p, q\}$. Note that the subscript r on Θ_r and $\mathcal{L}_r$ refer to the number of cointegration relations in Z_t.

In the next section we discuss the prior distributions for the model parameters of the multivariate Markov trend model presented in this section.

5 Prior Specification

The multivariate Markov trend is non-linear in certain parameters. This phenomenon often leads to local non-identification for certain parameters in the model. For instance, under rank reduction in Π, the parameter N_1 is not fully identified. Specifying a diffuse prior on N_1 implies that the conditional posterior of N_1 given Π is constant and non-zero in the point of rank reduction. The integral over this conditional posterior in the point of rank reduction is therefore infinity, favoring rank reduction. To circumvent this identification problem we follow the prior specification of Zivot (1994), see also Hoek (1997). The prior distribution for N_1 conditional on the first observation Y_1 and Σ is normal with mean Y_1 and covariance Σ

$$N_1|Y_1, \Sigma \sim \mathrm{N}(Y_1, \Sigma). \tag{27}$$

For Σ we take a standard diffuse prior

$$p(\Sigma) \propto |\Sigma|^{-\frac{1}{2}}. \tag{28}$$

The prior distributions for the transition probabilities p and q are independent and uniform on the unit interval $(0, 1)$

$$\begin{aligned} p(p) &= \mathbb{I}_{(0,1)} \\ p(q) &= \mathbb{I}_{(0,1)}, \end{aligned} \tag{29}$$

where $\mathbb{I}_{(0,1)}$ represents an indicator function which is one on the interval (0,1) and zero elsewhere. Under flat priors for p and q special attention must be payed to the priors for $,_0$ and $,_1$. It is easy to show that under $\Pi = \mathbf{0}$ the likelihood has the same value if we switch the role of the states and change the values of $,_0$, $,_1$, δ, p and q into $,_0 + ,_1$, $-,_1$, $-\delta$, q and p respectively. This complicates proper posterior analysis if we specify uninformative priors on $,_0$ and $,_1$. To circumvent this problem we have to define priors for $,_0$ and $,_1$ on subspaces $\mathcal{G}_0$ and $\mathcal{G}_1$ which uniquely identify the regimes for all specifications of the model,

$$\begin{aligned} p(,_0) &\propto \begin{cases} 1 & \text{if } ,_0 \in \mathcal{G}_0 \\ 0 & \text{elsewhere,} \end{cases} \\ p(,_1|,_0) &\propto \begin{cases} 1 & \text{if } ,_1 \in \mathcal{G}_1 \\ 0 & \text{elsewhere,} \end{cases} \end{aligned} \tag{30}$$

where $\mathcal{G}_0 = \{,_0 \in \mathbb{R}^2 |,_0 > \mathbf{0}\}$ and $\mathcal{G}_1 = \{,_1 \in \mathbb{R}^2 |,_0 + ,_1 \leq \mathbf{0}\}$. Another option to circumvent the identification problem is to specify appropriate matrix normal prior distributions for $,_0$ and $,_1$. Since we identify the two regimes by a prior on $,_0$ and $,_1$ we may use an improper prior for δ_1

$$p(\delta_1) \propto 1. \tag{31}$$

For the autoregressive parameters of the model we also use flat priors

$$\begin{aligned} p(\bar{\Phi}_i) &\propto 1, \qquad i = 1, \ldots, k-1, \\ p(\Pi) &\propto 1. \end{aligned} \tag{32}$$

The priors for α and β_2 parameters follow directly from the prior for Π and the following decomposition

$$\Pi = \alpha\beta' + \alpha_\perp \lambda \beta_\perp', \tag{33}$$

where $\beta = (1 \;\; -\beta_2)'$ and $\alpha_\perp$ and $\beta_\perp$ are specified such that $\alpha_\perp'\alpha = 0$ with $\alpha_\perp'\alpha_\perp = 1$ and $\beta_\perp'\beta = 0$ with $\beta_\perp'\beta_\perp = 1$, see Kleibergen and Paap (1998) for details. The scalar λ

models the deviation from the cointegration specification and hence it can be used to test for cointegration. If $\lambda = 0$, cointegration occurs. Note that the row- and columnspace of the matrix $(\alpha_\perp \lambda \beta'_\perp)$, which models the deviation from the cointegration specification $\alpha\beta'$, are spanned by the orthogonal complements of the vector of adjustment parameters α and the cointegrating vector β, respectively.

Decomposition (33) is equal to a singular value decomposition on Π,

$$\Pi = USV' \tag{34}$$

where U and V are (2×2) orthonormal matrices, S is an (2×2) diagonal matrix containing the positive singular values of Π (in decreasing order), see *e.g.* Magnus and Neudecker (1988). If we write

$$U = \begin{pmatrix} u_{11} & v_{12} \\ u_{21} & u_{22} \end{pmatrix}, \qquad S = \begin{pmatrix} s_{11} & 0 \\ 0 & s_{22} \end{pmatrix} \text{ and } V = \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix} \tag{35}$$

with u_{ij}, s_{ij}, v_{ij}, $i = 1, 2$, $j = 1, 2$ scalars, we obtain the following expressions for α, λ and β_2

$$\begin{aligned} \alpha &= u_{11}\, s_{11}\, (v_{11}\ v_{21})' \\ \lambda &= \text{sign}(u_{22}v_{22})s_{22} \\ \beta_2 &= -u_{21}\, u_{11}^{-1}, \end{aligned} \tag{36}$$

where sign$(\cdot)$ denotes the sign of the argument. The number of non-zero eigenvalues of a matrix determines the rank of a matrix. The singular value decomposition (35) shows that λ is identified through the smallest singular value of Π, which ends up in s_{22}. The scalar λ can be positive and negative in contrast to the singular value s_{22} which is always positive.

Using the decomposition (33), the prior for Π in (32) implies a joint prior for α, λ and β_2 in the following way

$$p(\alpha, \lambda, \beta_2) \propto p(\Pi)|_{\Pi=\alpha\beta'+\alpha_\perp\lambda\beta'_\perp} |J(\alpha, \lambda, \beta_2)|, \tag{37}$$

where $|J(\alpha, \lambda, \beta_2)|$ is the Jacobian of the transformation from Π to $(\alpha, \lambda, \beta_2)$. The derivation and expression of this Jacobian is given in Appendix A. Now, the joint prior for α and β_2 in case of cointegration ($\lambda = 0$) is simply the prior (37) evaluated in $\lambda = 0$ or

$$\begin{aligned} p(\alpha, \beta_2) &\propto p(\alpha, \lambda, \beta_2)|_{\lambda=0} \\ &\propto p(\Pi)|_{\Pi=\alpha\beta'} |J(\alpha, \lambda, \beta_2)|_{\lambda=0}. \end{aligned} \tag{38}$$

The joint priors for the Markov trend models with different number of unit roots follows from the marginal priors in this section. The joint prior for the Markov trend stationary model (20) $p_2(\Theta_2)$ is given by the product of (27)–(32). The prior for the Markov trend model with one cointegration relation (22) $p_1(\Theta_1)$ is the product of (27)–(31) and (38), while the prior for the model without cointegration (21) $p_0(\Theta_0)$ is simply the product of (27)–(31).

6 Posterior Distributions

The posterior distributions for the model parameters of the multivariate Markov trend models is proportional to the product of the priors $p_r(\Theta_r)$ and the unconditional likelihood functions $\mathcal{L}_r(Y^T|Y^k, \Theta_r)$, $r = 0, 1, 2$. These posterior distributions are too complicated to derive analytical posterior results. As Albert and Chib (1993), McCulloch and Tsay (1994) and Chib (1996) demonstrate, the Gibbs sampling algorithm of Geman and Geman (1984) is very useful tool for the computation of posterior results for models with unobserved states. The state variables $\{s_t\}_{t=1}^T$ can be treated as unknown parameters and simulated alongside the model parameters. This technique is known as data augmentation, see Tanner and Wong (1987).

The Gibbs sampler is an iterative algorithm, where one consecutively samples from the full conditional posterior distribution of the model parameters. This produces a Markov chain, which converges under mild conditions. The in this way obtained draws can be seen as a sample from the posterior distribution. For an introduction and details about the Gibbs sampling algorithm we refer to Smith and Roberts (1993) and Tierney (1994). In Appendix B we derive the full conditional posterior distributions which are necessary in the Gibbs sampler. We focus on the most general Markov trend stationary model (20). The full conditional posterior distributions of the other models can be derived in a similar way. Unfortunately, the full conditional distributions of the α and the β_2 parameters are not of a known type. To sample these parameters we need to build in a Metropolis-Hasting step in the Gibbs sampler, see also Chib and Greenberg (1995) for a discussion about introducing a Metropolis-Hasting step in a Gibbs sampler.

7 Determining the Cointegration Rank

The analysis of the determination of the cointegration rank starts with assigning prior probabilities to every possible rank of Π

$$\Pr[\text{rank} = r], \qquad r = 0, 1, 2, \tag{39}$$

which defines prior probabilities to the number of cointegration relations r. These prior probabilities imply the following prior odds ratios [PROR]

$$\text{PROR}(r|2) = \frac{\Pr[\text{rank} = r]}{\Pr[\text{rank} = 2]}, \qquad r = 0, 1, 2. \tag{40}$$

The Bayes factor to compare rank r with rank 2 equals

$$\text{BF}(r|2) = \frac{\int \mathcal{L}_r(Y^T|Y^k, \Theta_r)\, p_r(\Theta_r)\, d\Theta_r}{\int \mathcal{L}_2(Y^T|Y^k, \Theta_2)\, p_2(\Theta_2)\, d\Theta_2}, \qquad r = 0, 1 \tag{41}$$

where $\mathcal{L}_r(Y^T|Y^k, \Theta_r)$ and $p_r(\Theta_r)$ denote the unconditional likelihood function and the joint prior of the model with rank r. The posterior odds ratios to compare rank r with rank 2 equals prior odds ratio times the Bayes factor, $\text{POR}(r|2) = \text{PROR}(r|2) \times \text{BF}(r|2)$, and the posterior probabilities for every rank are simply

$$\Pr[\text{rank} = r|Y^T] = \frac{\text{POR}(r|n)}{\sum_{i=0}^{2} \text{POR}(i|2)}, \qquad r = 0, 1, 2. \tag{42}$$

The Bayes factors (41) are in fact Bayes factors for $\lambda = 0$ and $\Pi = \mathbf{0}$. They can be computed using the Savage-Dickey density ratio of Dickey (1971), which states that the Bayes factor for $\lambda = 0$ (or $\Pi = \mathbf{0}$) equals the ratio of the marginal posterior density and the marginal prior density of λ (Π), both evaluated in $\lambda = 0$ ($\Pi = \mathbf{0}$)

$$\begin{aligned} \text{BF}(1|2) &= \frac{p(\lambda|Y^T)|_{\lambda=0}}{p(\lambda)|_{\lambda=0}} \\ \text{BF}(0|2) &= \frac{p(\Pi|Y^T)|_{\Pi=\mathbf{0}}}{p(\Pi)|_{\Pi=\mathbf{0}}}. \end{aligned} \tag{43}$$

This means that we need the marginal posterior densities of λ and Π to compute this Savage-Dickey density ratio. The marginal posterior density of Π can be computed directly from the Gibbs output by averaging the full conditional posterior distribution of Π in the point $\mathbf{0}$ over the sampled model parameters, see Gelfand and Smith (1990). This trick cannot be used for λ, since the full conditional distribution of λ is of an unknown type. To compute the height of the marginal posterior of λ we may use a kernel estimator on simulated λ values, see *e.g.* Silverman (1986). Another possibility is to use an approximation of the full conditional posterior of λ in combination with importance weights, see Chen (1994). Kleibergen and Paap (1998) argue that the density function $g(\lambda|\Theta_2\backslash\{\lambda\}, Y^T)$ defined in (64) is a good approximation. This results in the following expression to compute the marginal posterior height in $\lambda = 0$

$$p(\lambda|Y^T)|_{\lambda=0} \approx \frac{1}{N}\sum_{i=1}^{N} \frac{|J(\alpha^i, \lambda, \beta_2^i)|_{\lambda=\mathbf{0}}|}{|J(\alpha^i, \lambda^i, \beta_2^i)|} g(\lambda|\Theta_2^i\backslash\{\lambda\}, Y^T)|_{\lambda=0} \tag{44}$$

where N denotes the number of simulations and Π in Θ_2 is written as $\alpha\beta' + \alpha_\perp\lambda\beta_\perp'$.

As we have specified diffuse priors for λ and Π the height of the marginal prior in $\lambda = 0$ and $\Pi = \mathbf{0}$ is not defined. The Bayes factor is therefore not defined in case of diffuse priors. The experiments in Kleibergen and Paap (1998) however show, that a Bayesian cointegration analysis with a diffuse prior specification on Π works fine if one replaces the marginal prior height by a penalty function. They suggest to replace the prior height by the factor $(2\pi)^{-\frac{1}{2}(2-r)^2}$, which leads to a Bayes factor that corresponds to the posterior information criterion [PIC] of Phillips and Ploberger (1994). We will opt for the same solution in this paper.

8 US Consumption and Income

In this section we analyze the presence of a long run relation between the US per capita consumption and income series considered in Section 3. We first start in Section 8.1 with a simple analysis of cointegration between the two series in a vector autoregression with a linear deterministic trend to illustrate the effects of neglecting a Markov trend in the series. In Section 8.2 we analyze the presence of a long run relation between consumption and income using the multivariate Markov trend model specification of Section 4.

8.1 A VAR model without Markov Trend

If we restrict , $_1$ and δ_1 in the Markov trend model (20) to zero, we end up with a vector autoregression for Y_t with only a linear deterministic trend. In this subsection we analyze the presence of a cointegration relation between US per capita consumption and income in this vector autoregression for $Y_t = 100 \times (\ln(c_t), \ln(y_t))'$. The priors for the model parameters of the linear VAR are given by (27), (28) and (32). For $, {}_0$ we take a flat prior: $p(, {}_0) \propto 1$. Under this prior specification the posterior means of the model parameters with no cointegration imposed are given by

$$
\begin{aligned}
Y_t &= N_t + Z_t, \\
N_t &= -\begin{pmatrix} \underset{(0.7)}{462.3} \\ \underset{(1.1)}{438.7} \end{pmatrix} + \begin{pmatrix} \underset{(0.07)}{0.43} \\ \underset{(0.09)}{0.40} \end{pmatrix} (t-1) \\
\Delta Z_t &= \begin{pmatrix} \underset{(0.04)}{0.07} & \underset{(0.03)}{-0.06} \\ \underset{(0.07)}{0.20} & \underset{(0.05)}{-0.16} \end{pmatrix} Z_{t-1} + \varepsilon_t, \text{ with } \Sigma = \begin{pmatrix} \underset{(0.06)}{0.50} & \underset{(0.08)}{0.53} \\ \underset{(0.08)}{0.53} & \underset{(0.16)}{1.34} \end{pmatrix}.
\end{aligned}
\tag{45}
$$

where posterior standard deviations appear in parentheses. The posterior means of the slopes of the deterministic trends in the consumption and income series are 0.43% and 0.40% respectively. They differ only 0.02% from the average quarterly growth rates reported in Section 3, which are well within the two posterior standard deviation regions around the posterior mean.

For the analysis of the presence of a cointegration relation between the two series, we assign equal probabilities to the possible cointegration ranks, *i.e.* $\Pr[\text{rank} = r] = \frac{1}{3}$ for $r = 0, 1, 2$. The prior for α and β_2 for the cointegration specification (rank=1) is given by (38). Since we have specified a diffuse prior for Π, the Bayes factors for rank reduction are not defined. Therefore, we consider PIC based Bayes factors for $\Pi = \mathbf{0}$ and $\lambda = 0$ in the decomposition (33), where we replace the prior height by the penalty function $(2\pi)^{-\frac{1}{2}(2-r)^2}$ as suggested by Kleibergen and Paap (1998). This leads to the following Bayes factors and posterior probabilities:

rank=r	$\ln(\text{BF}(r\|2))$	$\Pr[\text{rank} = r\|Y^T]$
0	10.10	0.99
1	4.90	0.01
2	0.00	0.00

The results show that a model with rank 0 or rank 1 is preferred to a model with full rank. The Bayes factor to compare the model with rank 0 versus the model with rank

1 equals $\ln(\text{BF}(0|1)) = 10.10 - 4.90 = 5.20$ such that the model without cointegration is preferred. The Bayes factors lead to assigning 99% posterior probability to the model with no cointegration relation. Hence, there is no evidence for a long run equilibrium between US per capita consumption and income in a VAR model with only a deterministic trend[3].

8.2 A Multivariate Markov Trend Model

The VAR model with a deterministic trend assumes that the quarterly growth rates of consumption and income are constant over time. However, the stylized facts suggest that the long run average quarterly growth rates are roughly the same, but that there may be different growth rates in both series during expansions and recessions. To correct for possible different growth rates in consumption and income during recessions and expansions, we consider the Markov trend model (20). The prior for the model parameters is given by (27)–(32). To identify the regimes we impose the restriction $,_0 > 0$ and $,_0 + ,_1 < 0$ as suggested in Section 5. The posterior means of the model parameters of the multivariate Markov trend model with no cointegration imposed are given by

$$
\begin{aligned}
Y_t &= N_t + R_t + Z_t, \\
N_t &= -\begin{pmatrix} \underset{(0.6)}{462.0} \\ \underset{(0.8)}{439.1} \end{pmatrix} + \begin{pmatrix} \underset{(0.14)}{0.73} \\ \underset{(0.15)}{0.98} \end{pmatrix} (t-1) - \begin{pmatrix} \underset{(0.16)}{0.82} \\ \underset{(0.23)}{1.61} \end{pmatrix} \sum_{i=2}^{t} s_i, \\
R_t &= \begin{pmatrix} \underset{(0.25)}{0.16} \\ 0 \end{pmatrix} s_t, \\
\Delta Z_t &= \begin{pmatrix} \underset{(0.12)}{0.26} & \underset{(0.08)}{-0.23} \\ \underset{(0.26)}{0.63} & \underset{(0.17)}{-0.51} \end{pmatrix} Z_{t-1} + \varepsilon_t, \text{ with } \Sigma = \begin{pmatrix} \underset{(0.06)}{0.40} & \underset{(0.09)}{0.27} \\ \underset{(0.09)}{0.27} & \underset{(0.16)}{0.67} \end{pmatrix},
\end{aligned} \tag{46}
$$

where posterior standard deviations are in parentheses. The posterior means of the transition probabilities equal

$$p = 0.86\ (0.07) \text{ and } q = 0.76\ (0.09).$$

The posterior mean of the per capita growth rates of disposable income are 0.98% during expansions and -0.63% $(0.98 - 1.61)$ during recessions. For the consumption

[3] Also the standard Johansen trace tests for rank reduction do not indicate the presence of a cointegration relation between the two series.

series the growth rates are 0.73% and 0.09% (0.73 − 0.82), respectively. The expected slope of the Markov trend is given by $,_0 + ,_1(1-p)/(2-p-q)$. The posterior means of this expected slope equals 0.42% for the consumption series and 0.39% for the income series. Note that these slope values are almost the same as the posterior means of the slope parameters of the deterministic trend in (45). Finally, note that the posterior mean of the δ_1 parameter, which equals 0.16, lies within one posterior standard deviations from zero.

Again, we perform a cointegration analysis but now we analyze the presence of a cointegration relation in the deviations from a Markov trend instead of a deterministic trend. We assign equal probabilities to the possible cointegration ranks, *i.e.* $\Pr[\text{rank} = r] = \frac{1}{3}$ for $r = 0, 1, 2$. The prior for α and β_2 for the cointegration specification (rank=1) is given by (38). Using decomposition (33) we compute Bayes factors for $\Pi = \mathbf{0}$ and $\lambda = 0$, as in (43), where we replace the prior height by the penalty function $(2\pi)^{-\frac{1}{2}(2-r)^2}$ since we are dealing with an uninformative prior for Π. This leads to the following Bayes factors and posterior probabilities:

rank=r	$\ln(\text{BF}(r\|2))$	$\Pr[\text{rank} = r\|Y^T]$
0	2.72	0.13
1	4.62	0.86
2	0.00	0.01

Again, rank reduction is chosen above a full rank model. However, the Bayes factor to compare a multivariate Markov trend model with rank 1 versus rank 0 equals $\ln(\text{BF}(0|1)) = 2.72 - 4.62 = -1.90$ and hence the model with one cointegration relation is preferred. The posterior probabilities assign 86% probability to the model with one cointegration relation.

Cointegration Specification

The Bayes factors suggest that the multivariate Markov trend model with one cointegration relation (22) is suitable to model the logarithm of US per capita consumption and income. The prior for the model parameters is given by (27)–(31) and (38). The posterior means of the model parameters based on this prior are given by

$$
\begin{aligned}
Y_t &= N_t + Z_t, \\
N_t &= -\begin{pmatrix} \underset{(0.6)}{462.1} \\ \underset{(0.8)}{439.1} \end{pmatrix} + \begin{pmatrix} \underset{(0.18)}{0.68} \\ \underset{(0.21)}{0.97} \end{pmatrix} (t-1) - \begin{pmatrix} \underset{(0.18)}{0.78} \\ \underset{(0.21)}{1.65} \end{pmatrix} \sum_{i=2}^{t} s_i, \\
R_t &= \begin{pmatrix} \underset{(0.20)}{0.11} \\ 0 \end{pmatrix} s_t, \\
\Delta Z_t &= \begin{pmatrix} \underset{(0.10)}{0.31} \\ \underset{(0.21)}{0.68} \end{pmatrix} \begin{pmatrix} 1 & \underset{(0.19)}{-0.84} \end{pmatrix} Z_{t-1} + \varepsilon_t, \text{ with } \Sigma = \begin{pmatrix} \underset{(0.06)}{0.41} & \underset{(0.07)}{0.26} \\ \underset{(0.07)}{0.26} & \underset{(0.12)}{0.64} \end{pmatrix},
\end{aligned} \tag{47}
$$

where again posterior standard deviations appear in parentheses. The posterior means of the transition probabilities equal

$$p = 0.86 \ (0.05) \text{ and } q = 0.76 \ (0.09)$$

The posterior results are obtained by including a Metropolis-Hasting step in the Gibbs sampler to sample α and β_2, see Appendix B. The candidate draw for α and β_2 was accepted in 80% of the iterations.

The posterior mean of cointegration relation parameter $\beta_2 = -0.84$ does not differ more than two posterior standard deviations from -1. In fact, a PIC based Bayes factor for $\beta_2 = -1$ equals 1.69 and hence the consumption-income relation (8) may be valid. The Bayes factor is computed using the Savage-Dickey density ratio, where we use the $1/\sqrt{2\pi}$ as a penalty function, since we have used an uninformative prior for β_2, see also Section 7 for a similar approach. The adjustment parameters α are both positive, which indicates that there is no adjustment towards the equilibrium for the consumption equation. This phenomenon is not due to the non-linear Markov trend in the model, since unreported results show that this also arises in a simple VAR model with and even without a deterministic trend instead of a Markov trend, see also (45). Note that this does not imply that the series move away from the equilibrium, since the adjustment of income towards the equilibrium is larger than the non-adjustment in consumption, see also Johansen (1995, p. 39–42).

The posterior mean of the δ_1 parameter equals 0.11 with a posterior standard deviation of 0.21. A PIC based Bayes factor for $\delta_1 = 0$ equals 3.93 and hence it is very likely that δ_1

equals 0. The posterior means of the quarterly growth rates of the income series are 0.97% during an expansion regime and -0.68% $(= 0.97 - 1.65)$ during a contraction regime. For the consumption series we get 0.68% and -0.10% $(= 0.68 - 0.78)$, respectively. Hence, during recessions the negative growth rate in consumption is smaller than the negative growth rate in income. To correct for this difference in the growth rates, the positive growth rate in income has to be larger than the positive growth rate in consumption during expansions. The expected slope of the Markov trend is given by $,_0 + ,_1(1-p)/(2-p-q)$. The posterior means of this expected slope equals 0.40% for the consumption series and 0.37% for the income series and hence the expected long run slopes in both series are roughly the same. Reduced rank Markov trend cointegration is not likely since based on the posterior mean of $\beta' ,_1$ equals 0.59 with a relatively small posterior standard deviation. Hence, the existence of a consumption-income (8) which requires that both $\beta' ,_1$ and $\beta' ,_0$ equal 0 is not likely. On the other hand the results suggest that during recession periods there is less decline in consumption than in income, which is compensated in the expansion periods where income grows faster than consumption.

Finally, we analyze how the estimated Markov trend relates to the NBER business cycle. The posterior mean of the probability of staying in the expansion regime is 0.86, which is larger than the posterior mean of the probability of staying in a recession 0.76. The posterior probability that p is larger than q is 0.92, which indicates the existence of an asymmetric cycle. The posterior expectations of the states variables $\mathrm{E}[s_t|Y^T]$ are shown in Figure 3. Values of this expectation which are close to 1, correspond to recessionary periods. Figure 4 shows the difference of the logarithm of US income and consumption. The shaded areas correspond to the periods where the growth rate in consumption is larger than the growth rate in income. In these periods the posterior mean of the growth rate for the difference equals $-0.78 - 1.65 = 0.87\%$. The negative slope of 0.29% $(= 0.69 - 0.97)$ in the other periods results from the larger positive slope of the Markov trend for the income series than for the consumption series. As we already have discussed the posterior mean of the expected slopes of the Markov trend is 0.37% for the income series and 0.40% for the consumption series. Hence, the posterior mean of the unconditional expectation of the slope of the Markov trend in the $(1, -1)$-cointegration relation is about -0.03%, which is almost zero.

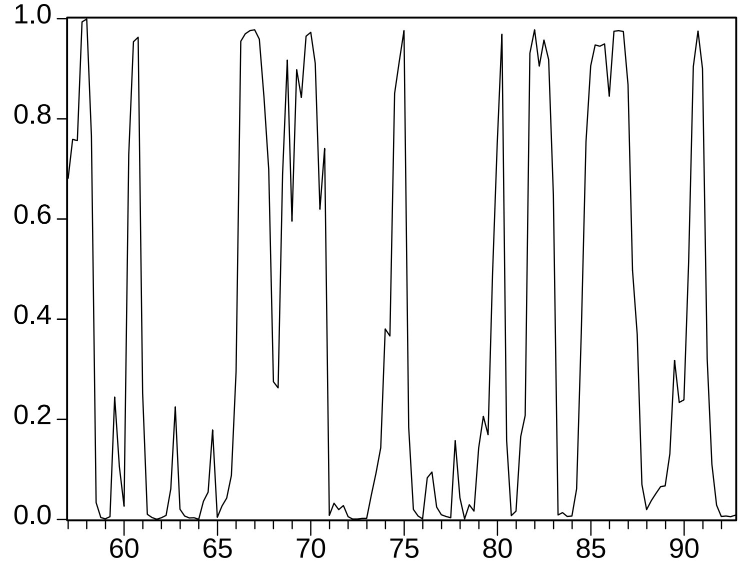

Figure 3: Posterior expectations of the state variables $E[s_t|Y^T]$.

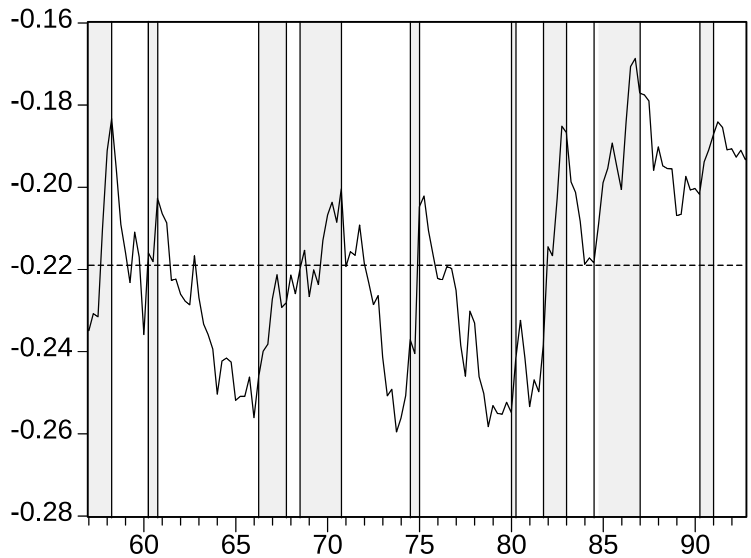

Figure 4: Difference between log US per capita consumption and income. The shaded areas correspond to periods where $E[s_t|Y^T] > 0.5$.

Table 1: Peaks and troughs based on the posterior expectations of the states[1].

US		NBER	
peak	trough	peak	trough
–	1958.2	1957.3	1958.2
1960.1	1960.4	1960.2	1961.1
1966.1	1967.4		
1968.2	1970.4	1969.4	1970.4
1974.2	1975.1	1973.4	1975.1
1979.4	1980.2	1980.1	1980.3
1981.3	1983.1	1981.3	1982.4
1984.3	1987.1		
1990.1	1991.1	1990.3	1991.1

[1] A recession is defined by 2 consecutive quarters for which $\mathrm{E}[s_t|Y^T] > 0.5$. A peak corresponds with the last expansion observation before a recession and a trough with the last observation in a recession.

Table 1 shows the estimated peaks and troughs based on the posterior expectation of the states variables together with the official NBER peaks and troughs. We define a recession by 2 consecutive quarters for which $\mathrm{E}[s_t|Y^T] > 0.5$. A peak is defined by the last expansion observation before a recession. A trough is defined by the last observation in a recession. We see that the estimated turning points correspond very well with the official NBER peaks and troughs. However, we detect two extra recessionary periods, which do not correspond to official reported recessions. Note that the consumption income analysis in this paper is based on per capita disposable income. If we look at the government purchases on goods and services, which are used to create the disposable income series, we see that government expenses increase during recessions resulting in an extra decrease in disposable income. However, there was also a large increase in government expenses during the two periods which are incorrectly reported as recession. This resulted in a small decline or a smaller growth in disposable income during these two periods, which explains the detection of the two extra recessions in our data.

In summary, the multivariate Markov trend model provides a good description for the US per capita income and consumption series. The multivariate Markov trend captures the different growth rates in both series during recession and expansion periods. After detrending with the Markov trend we detect a stationary linear combination between per capita income and consumption. This cointegration relation is not found if we use a regular deterministic trend instead of a Markov trend for detrending.

9 Conclusion

In this paper we have proposed a multivariate Markov trend model to analyze the possible existence of a long run relation between per capita consumption and income of the United States. The model specification has been based on suggestions by simple economic theory and a simple stylized facts analysis on both series. The model contains a multivariate Markov trend specification, which allows for different growth rates during recessions and expansions. The deviations from the multivariate Markov trend are modelled by a vector autoregressive model. To analyze the US series with the multivariate Markov trend model, we have chosen for a Bayesian approach. Bayes factors are proposed to analyze the presence of a cointegration relation in the deviations of the series from the multivariate Markov trend.

The posterior results suggest that there exist a stationary linear relation between log per capita consumption and income after correcting for a Markov trend. The Markov trend models the different growth rates in both series during recessions and expansions. The negative growth rate in consumption is smaller in absolute value than the negative growth rate in income during recessions. To compensate this difference the growth rate in income is larger than the growth rate in consumption during expansion periods. If we replace the Markov trend by a deterministic linear trend posterior results do not indicate the presence of a stationary linear relation between both series.

We end this conclusion with some suggestion for further research. The multivariate Markov trend model we proposed in this paper is linear in deviation from the Markov trend. Possible cointegrating vectors and adjustment parameters are not affected by regime changes. We may however also allow that the adjustment parameters or the cointegrating vector have different values over the business cycle. This implies a nonlinear error correction mechanism in consumption and income, see also Peel (1992). It is then even possible that the series are only cointegrated in expansions and not in recessions. Testing for the presence of cointegration in the different regimes may however be difficult since the number of observations for recessionary periods is usually very small. Furthermore, the dynamic properties of such models are not easy to derive, see Holst *et al.* (1994) and Warne (1996). Finally, we may also consider alternative multivariate nonlinear models to analyze the consumption and income series, like threshold models, see *e.g.* Granger and Teräsvirta (1993).

A Jacobian Transformation

In this appendix we derive the Jacobian of the transformation from Π to $(\alpha, \lambda, \beta_2)$ for a 2-dimensional vector autoregressive model. For larger dimensions, see Kleibergen and Paap (1998). Define $\alpha = (\alpha_1, \alpha_2)$, where α_1 and α_2 are scalars and $\theta_2 = -\alpha_2/\alpha_1$ such that $\alpha = \alpha_1\theta$ with $\theta = (1 \quad -\theta_2)'$. The derivation of the Jacobian of the complete transformation from Π to $(\alpha_1, \alpha_2, \lambda, \beta_2)$ is for notional convenience split up in the Jacobian of the transformation of Π to $(\alpha_1, \theta_2, \lambda, \beta_2)$ and then the transformation of θ_2 to α_2. Since $\theta_\perp \in \alpha_\perp$ we can write

$$\begin{aligned}
\Pi &= \alpha\beta' + \alpha_\perp\lambda\beta'_\perp \\
&= (\alpha\ \alpha_\perp)\begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix}\begin{pmatrix} \beta' \\ \beta'_\perp \end{pmatrix} \\
&= \begin{pmatrix} 1 & \theta_2/\sqrt{1+\theta_2^2} \\ -\theta_2 & 1/\sqrt{1+\theta_2^2} \end{pmatrix}\begin{pmatrix} \alpha_1 & 0 \\ 0 & \lambda \end{pmatrix}\begin{pmatrix} 1 & -\beta_2 \\ -\beta_2/\sqrt{1+\beta_2^2} & 1/\sqrt{1+\beta_2^2} \end{pmatrix} \\
&= \alpha_1\begin{pmatrix} 1 & -\beta_2 \\ -\theta_2 & \theta_2\beta_2 \end{pmatrix} + \frac{\lambda}{\sqrt{(1+\theta_2^2)(1+\beta_2^2)}}\begin{pmatrix} -\theta_2\beta_2 & \theta_2 \\ -\beta_2 & 1 \end{pmatrix}
\end{aligned} \tag{48}$$

The derivatives of Π with respect to α_1, θ_2, λ and β_2 read

$$\begin{aligned}
J_1 &= \frac{\partial\,\mathrm{vec}(\Pi)}{\partial\alpha_1} = \begin{pmatrix} 1 \\ -\theta_2 \\ -\beta_2 \\ \theta_2\beta_2 \end{pmatrix} \\
J_2 &= \frac{\partial\,\mathrm{vec}(\Pi)}{\partial\theta_2} = \begin{pmatrix} 0 \\ -\alpha_1 \\ 0 \\ \alpha_1\beta_2 \end{pmatrix} + \frac{\lambda}{\sqrt{(1+\theta_2^2)(1+\beta_2^2)}}\begin{pmatrix} -\beta_2 + \theta_2^2\beta_2/(1+\theta_2^2) \\ \theta_2\beta_2/(1+\theta_2^2) \\ 1-\theta_2^2/(1+\theta_2^2) \\ -\theta_2/(1+\theta_2^2) \end{pmatrix} \\
J_3 &= \frac{\partial\,\mathrm{vec}(\Pi)}{\partial\lambda} = \frac{1}{\sqrt{(1+\theta_2^2)(1+\beta_2^2)}}\begin{pmatrix} -\theta_2\beta_2 \\ -\beta_2 \\ \theta_2 \\ 1 \end{pmatrix} \\
J_4 &= \frac{\partial\,\mathrm{vec}(\Pi)}{\partial\beta_2} = \begin{pmatrix} 0 \\ 0 \\ -\alpha_1 \\ \alpha_1\theta_2 \end{pmatrix} + \frac{\lambda}{\sqrt{(1+\theta_2^2)(1+\beta_2^2)}}\begin{pmatrix} -\theta_2 + \theta_2\beta_2^2/(1+\beta_2^2)] \\ 1-\beta_2^2/(1+\beta_2^2) \\ -\theta_2\beta_2/(1+\beta_2^2) \\ -\beta_2/(1+\beta_2^2) \end{pmatrix}.
\end{aligned} \tag{49}$$

The Jacobian from θ_2 to α_2 is simply

$$G = \left|\frac{\partial\theta_2}{\partial\alpha_2}\right| = -\frac{1}{\alpha_1} \tag{50}$$

Hence, the Jacobian for the total transformation equals

$$J(\alpha, \lambda, \beta_2) = |J_1 \; J_2 \; J_3 \; J_4| \; |G|. \tag{51}$$

B Full Conditional Posterior Distributions

Full Conditional Posterior of the States

To sample the states, we need the full conditional posterior density of s_t, denoted by $p(s_t|s^{-t}, \Theta_2, Y^T)$, $t = 1, \ldots, T$, where $s^{-t} = s^T \backslash \{s_t\}$. Since s_t follows a first-order Markov process, it is easily seen that

$$p(s_t|s^{-t}) \propto p(s_t|s_{t-1}) \, p(s_{t+1}|s_t), \tag{52}$$

due to the Markov property. Following Albert and Chib (1993), we can write

$$\begin{aligned} p(s_t|s^{-t}, \Theta_2, Y^T) &= \frac{p(s_t|s^{-t}, \Theta_2, Y^t) \, f(Y_{t+1}, \ldots, Y_T|Y^t, s^{-t}, s_t, \Theta_2)}{f(Y_{t+1}, \ldots, Y_T|Y^t, s^{-t}, \Theta_2)} \\ &\propto p(s_t|s^{-t}, \Theta_2, Y^t) \, f(Y_{t+1}, \ldots, Y_T|Y^t, s^{-t}, s_t, \Theta_2). \end{aligned} \tag{53}$$

Using the rules of conditional probability, the first term of (53) can be simplified as

$$\begin{aligned} p(s_t|s^{-t}, \Theta_2, Y^t) &\propto p(s_t|s^{-t}, \Theta_2, Y^{t-1}) \, f(Y_t, s_{t+1}, \ldots, s_T|Y^{t-1}, s^t, \Theta_2) \\ &\propto p(s_t|s_{t-1}, \Theta_2) \, f(Y_t|Y^{t-1}, s^t, \Theta_2) \\ &\qquad p(s_{t+1}|s^t, \Theta_2, Y^t) \, p(s_{t+2}, \ldots, s_T|s^{t+1}, \Theta_2, Y^t) \\ &\propto p(s_t|s_{t-1}, \Theta_2) \, f(Y_t|Y^{t-1}, s^t, \Theta_2) \, p(s_{t+1}|s_t, \Theta_2), \end{aligned} \tag{54}$$

where we use the fact that $\{s_{t+2}, \ldots, s_T\}$ is independent of s_t given s_{t+1}. The second term of (53) is proportional to

$$f(Y_{t+1}, \ldots, Y_T|Y^t, s^t, \Theta_2) \propto \prod_{i=t+1}^{T} f(Y_i|Y^{i-1}, s^i, \Theta_2). \tag{55}$$

Next, using (54) and (55) the full conditional distribution of s_t for $t = k+1, \ldots, T$ is given by

$$p(s_t|s^{-t}, \Theta_2, Y^T) \propto p(s_t|s_{t-1}, \Theta_2) \, p(s_{t+1}|s_t, \Theta_2) \prod_{i=t}^{T} f(Y_i|Y^{i-1}, s^i, \Theta_2), \tag{56}$$

where $f(Y_t|Y^{t-1}, s^t, \Theta_2)$ is defined in (23) and the constant of proportionality can be obtained by summing over the two possible values of s_t. At time $t = T$ the term $p(s_{T+1}|s_T, \Theta_2)$ drops out. The first k states can be sampled from the full conditional distribution

$$p(s_t|s^{-t}, \Theta_2, Y^T) \propto p(s_t|s_{t-1}, \Theta_2)\, p(s_{t+1}|s_t, \Theta_2) \prod_{i=k+1}^{T} f(Y_i|Y^{i-1}, s^i, \Theta_2), \tag{57}$$

for $t = 1, \ldots, k$, where at time $t = 1$ the term $p(s_t|s_{t-1}, \Theta_2)$ is replaced by the unconditional density $p(s_1|\Theta_2)$, which is a binomial density with probability $(1 - p)/(2 - p - q)$.

As Albert and Chib (1993) show, sampling of the state variables is easier if $\Pi = \mathbf{0}$. Under this restriction only the first $(k - 1)$ future conditional densities of Y_t depend on s_t instead of all future conditional densities. However, sampling is possible in the same way: take the most recent value of s^T and sample the states backward in time, one after another, starting with s_T. After each step, the t-th element of s^T is replaced by its most recent draw.

Full Conditional Posterior of p and q

From the conditional likelihood function (24) if follows that the full conditional posterior densities of the transition parameters are given by

$$\begin{aligned} p(p|s^T, \Theta_2\backslash\{p\}, Y^T) &\propto p^{\mathcal{N}_{0,0}}(1 - p)^{\mathcal{N}_{0,1}} \\ p(q|s^T, \Theta_2\backslash\{q\}, Y^T) &\propto q^{\mathcal{N}_{1,1}}(1 - q)^{\mathcal{N}_{1,0}}, \end{aligned} \tag{58}$$

where $\mathcal{N}_{i,j}$ again denotes the number of transitions from state i to state j. This implies that the transition probabilities can be sampled from beta distributions.

Full Conditional Posterior of Σ

It is easy to see from the conditional likelihood (24) that the full conditional posterior of Σ is proportional to

$$p(\Sigma|s^T, \Theta_2\backslash\Sigma, Y^T) \propto |\Sigma|^{-\frac{1}{2}(T-k+2)} \exp(-\frac{1}{2}\mathrm{tr}(\Sigma^{-1}((Y_1 - N_1)(Y_1 - N_1)' + \sum_{t=k+1}^{T} \epsilon_t\epsilon_t')) \tag{59}$$

and hence the covariance matrix Σ can be sampled from an inverted Wishart distribution, see Zellner (1971, p. 395).

Full Conditional Posterior of N_1, $,_0$ and $,_1$

To derive the full conditional posterior distribution of N_1, $,_0$ and $,_1$ we write (20) as

$$\begin{aligned}
\Sigma^{-\frac{1}{2}}\Phi(L)Y_t &= \Sigma^{-\frac{1}{2}}\Phi(L)(,_0(t-1) + ,_1 \sum_{i=2}^{t} s_i + N_1) + \Sigma^{-\frac{1}{2}}\varepsilon_t \\
&= -\Sigma^{-\frac{1}{2}} \sum_{j=1}^{k} \Phi_j(,_0 \quad ,_1 \quad N_1) \begin{pmatrix} L^j(t-1) \\ L^j \sum_{i=2}^{t} s_i \\ 1 \end{pmatrix} + \Sigma^{-\frac{1}{2}}\varepsilon_t,
\end{aligned} \tag{60}$$

where $\Phi_0 = -\mathbf{I}$. Without the Φ_j matrices, we have a multivariate regression model in the parameters N_1, $,_0$ and $,_1$ and the full conditional distribution would be matrix normal. To reverse the order of $\Phi(L)$ and the parameters $(,_0 \quad ,_0 \quad N_1)$, we apply the vec operator to both sides of (60). Using the vec notation and the fact that $\text{vec}(ABC) = (C' \otimes A)\text{vec}(B)$, we can write (60) as a linear regression model and hence the full conditional distributions of $\text{vec}(N_1)$, $\text{vec}(,_0)$ and $\text{vec}(,_1)$ are normal.

Full Conditional Posterior of $\boldsymbol{\delta_1}$

We write (20) as

$$\Sigma^{-\frac{1}{2}}\Phi(L)(Y_t - N_t) = \Sigma^{-\frac{1}{2}}\phi(L)\delta R_t + \Sigma^{-\frac{1}{2}}\varepsilon_t \tag{61}$$

with $\Phi_0 = -\mathbf{I}$. Applying the vec operator to both sides leads to a standard regression model with regression parameter δ_1. The full conditional posterior of δ_1 is therefore normal.

Full Conditional Posterior of Π and $\bar{\Phi}$

To sample from the full conditional posterior of the autoregressive parameters we use that conditional on $,_0$, $,_1$, N_1 and the states $\{s_t\}_{t=1}^{T}$, equation (20) can be seen as a multivariate regression model in the parameters Π and $\bar{\Phi}$. From Zellner (1971, chapter VIII) if follows that the full conditional posterior distribution of the parameter matrices are matrix normal. A draw from the full conditional distribution of λ can be obtained by performing a singular value decomposition on the sampled Π and solving for λ using (36).

Sampling of α and β_2

To derive the full conditional posterior distributions for α and β_2 we rewrite (22) such that conditional on $\bar{\Phi}$, N_1, $,_0$, $,_1$ and the states $\{s_T\}_{t=1}^T$ it resembles a simple VAR(1) model. Using $Z_t = Y_t - N_t - R_{t-1}$ we can write

$$\begin{aligned}\Delta Z_t - \sum_{i=1}^{k-1} \bar{\Phi}_i \Delta Z_{t-i} &= \alpha\beta' Z_{t-1}, + \varepsilon_t \\ \Delta Z_t^* &= \alpha\beta' Z_{t-1}^* + \varepsilon_t,\end{aligned} \tag{62}$$

where $\Delta Z_t^* = \Delta Z_t - \sum_{i=1}^{k-1} \bar{\Phi}_i \Delta Z_{t-i}$ and $Z_{t-1}^* = Y_{t-1} - N_{t-1} - R_{t-1}$. Kleibergen and Paap (1998) propose a Metropolis-Hasting algorithm to sample α and β_2 in this simple VAR model. Chib and Greenberg (1994, 1995) show that it is possible to build such a Metropolis-Hasting algorithm into the Gibbs sampling procedure. The Metropolis-Hasting algorithm step works as follows. First, draw in iteration i of the Gibbs sampler Π^i from its full conditional posterior distribution, see above. Perform a singular value decomposition on Π and solve for α^i, λ^i and β_2^i. Now accept this draw of α^i and β_2^i with probability $\min\left(\frac{w(\alpha^i,\lambda^i,\beta_2^i)}{w(\alpha^{i-1},\lambda^{i-1},\beta_2^{i-1})}, 1\right)$, where i denotes the current draw, $i-1$ the previous draw and

$$w(\alpha, \lambda, \beta_2) = \frac{|J(\alpha,\lambda,\beta_2)|_{\lambda=0}|}{|J(\alpha,\lambda,\beta_2)|} g(\lambda|\Theta_2\backslash\{\lambda\}, Y^T)|_{\lambda=0} \tag{63}$$

where

$$\begin{aligned}g(\lambda|\Theta_2\backslash\{\lambda\}, Y^T) = (2\pi)^{-\frac{1}{2}} |\alpha_\perp' \Sigma^{-1} \alpha_\perp|^{\frac{1}{2}} |\beta_\perp' (Z_{-1}^{*\prime} Z_{-1}^*) \beta_\perp|^{\frac{1}{2}} \\ \exp(-\frac{1}{2} \text{tr}((\beta_\perp' (Z_{-1}^{*\prime} Z_{-1}^*) \beta_\perp)(\lambda - \tilde{\lambda})(\alpha_\perp' \Sigma^{-1} \alpha_\perp)(\lambda - \tilde{\lambda})))\end{aligned} \tag{64}$$

with

$$\tilde{\lambda} = (\beta_\perp' (Z_{-1}^{*\prime} Z_{-1}^*) \beta_\perp)^{-1} \beta_\perp' Z_{-1}^{*\prime} (\Delta Z^* - Z_{-1}^* \beta\alpha') \Sigma^{-1} \alpha_\perp (\alpha_\perp' \Sigma^{-1} \alpha_\perp)^{-1} \tag{65}$$

and $Z_{-1}^* = (Z_k^* \dots Z_{T-1}^*)'$, $\Delta Z^* = (\Delta Z_{k+1}^* \dots \Delta Z_T^*)'$. If the draw of α^i and β_2^i is rejected, one has to take the previous draw *i.e.* $\alpha^i = \alpha^{i-1}$ and $\beta_2^i = \beta_2^{i-1}$, see Kleibergen and Paap (1998) for more details on this sampling algorithm.

References

Albert, J.H. and S. Chib, 1993, Bayes Inference via Gibbs Sampling of Autoregressive Time Series subject to Markov Mean and Variance Shifts, *Journal of Business & Economic Statistics* **11**, 1–15.

Campbell, J.Y., 1987, Does Saving Anticipate Declining Labor Income? An Alternative Test of the Permanent Income Hypothesis, *Econometrica* **55**, 1249–1273.

Chen, M-H., 1994, Importance-Weighted Marginal Bayesian Posterior Density Estimation, *Journal of the American Statistical Association* **89**, 818–824.

Chib, S., 1996, Calculating Posterior Distributions and Model Estimates in Markov Mixture Models, *Journal of Econometrics* **75**, 79–97.

Chib, S. and E. Greenberg, 1994, Bayes Inference in Regression Models with ARMA(p,q) Errors, *Journal of Econometrics* **64**, 183–206.

Chib, S. and E. Greenberg, 1995, Understanding the Metropolis-Hastings Algorithm, *American Statistician* **49**, 327–335.

Dickey, D.A. and W.A. Fuller, 1979, Distribution of Estimators for Autoregressive Time Series with a Unit Root, *Journal of the American Statistical Association* **74**, 427–431.

Dickey, J., 1971, The Weighted Likelihood Ratio, Linear Hypothesis on Normal Location Parameters, *The Annals of Mathematical Statistics* **42**, 240–223.

Dwyer, M.J. and S.M. Potter, 1996, Markov Trend Cointegration, unpublished working paper, University of California.

Gelfand, A.E. 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 Relaxations, Gibbs Distributions, and the Bayesian Restoration of Images, *IEEE Transaction on Pattern Analysis and Machine Intelligence* **6**, 721–741.

Goodwin, T.H., 1993, Business-Cycle Analysis with a Markov-Switching Model, *Journal of Business & Economic Statistics* **11**, 331–339.

Granger, C.W.J. and T. Teräsvirta, 1993, *Modelling Nonlinear Economic Relations*, Oxford University Press, Oxford.

Hall, R.E., 1978, Stochastic Implications of the Life-Cycle Permanent Income Hypothesis: Theory and Evidence, *Journal of Political Economy* **86**, 971–987.

Hall, S.G., Z. Psaradakis and M. Sola, 1997, Cointegration and Changes in Regime: The Japanese Consumption Function, *Journal of Applied Econometrics* **12**, 151–168.

Hamilton, J.D., 1989, A New Approach to the Econometric Analysis of Nonstationary Time Series and Business Cycles, *Econometrica* **57**, 357–384.

Hoek, H., 1997, *Variable Trends: A Bayesian Perspective*, PhD thesis, Tinbergen Institute, Thesis Publishers, Amsterdam.

Holst, U., G. Lindgren, J. Holst and M. Thuvesholmen, 1994, Recursive Estimation in Switching Autoregressions with a Markov Regime, *Journal of Time Series Analysis* **15**, 489–506.

Jin, F., 1995, Cointegration of Consumption and Disposable Income: Evidence from Twelve OECD Countries, *Southern Economic Journal* **62**, 77–88.

Johansen, S., 1995, *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press, Oxford.

Kleibergen, F. and R. Paap, 1998, Priors, Posterior and Bayes Factors for a Bayesian Analysis of Cointegration, *Econometric Institute Report 9821/A*, Erasmus University Rotterdam.

Krolzig, H-M., 1997, *Markov Switching Vector Autoregressions: Modelling, Statistical Inference, and an Application to Business Cycle Inference*, Springer, Berlin.

Magnus, J.R. and H. Neudecker, 1988, *Matrix Differential Calculus with Applications in Statistics and Econometrics*, Wiley & Sons, Chichester.

McCulloch, R.E. and R.S. Tsay, 1994, Statistical Analysis of Economic Time Series via Markov Switching Models, *Journal of Time Series Analysis* **15**, 523–540.

Paap, R., 1997, *Markov Trends in Macroeconomic Time Series*, PhD thesis, Tinbergen Institute, Thesis Publishers, Amsterdam.

Peel, D.A., 1992, Some Analysis of the Long-run Time Series Properties of Consumption and Income in the U.K., *Economics Letters* **39**, 173–178.

Peel, D.A. and A.E.H. Speight, 1998, Threshold Nonlinearities in Output: Some International Evidence, *Applied Economics* **30**, 323–334.

Phillips, P.C.B. and W. Ploberger, 1994, Posterior Odds Testing for a Unit Root with Data-based Model Selection, *Econometric Theory* **10**, 774–808.

Potter, S.M., 1995, A Nonlinear Approach to US GNP, *Journal of Applied Econometrics* **10**, 109–125.

Silverman, B.W., 1986, *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, London.

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 B* **55**, 3–23.

Tanner, M.A. and W.H. Wong, 1987, The Calculation of Posterior Distributions by Data Augmentation, *Journal of the American Statistical Association* **82**, 528–550.

Tierney, L., 1994, Markov Chains for Exploring Posterior Distributions, *The Annals of Statistics* **22**, 1701–1762.

Warne, A., 1996, Autocovariance Functions and Maximum Likelihood in a VAR Model under Markov Switching, unpublished working paper, Stockholm University.

Zellner, A., 1971, *An Introduction to Bayesian Inference in Econometrics*, Wiley, New York.

Zivot, E., 1994, A Bayesian Analysis of the Unit Root Hypothesis within an Unobserved Components Model, *Econometric Theory* **10**, 552–578.