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

Richard PAAP and Herman K. VAN DIJK
Econometric Institute, Erasmus University, Rotterdam

Stylized facts show that average growth rates of U.S. per capita consumption and income differ in recession and expansion periods. Because 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 that accounts for different growth rates in consumption and income during expansions and recessions and across variables within both regimes. The deviations from the multivariate Markov trend are modeled by a vector autoregression (VAR) model. Bayes estimates of this model are obtained using Markov chain Monte Carlo methods. The empirical results suggest the existence of a cointegration relation between U.S. per capita disposable income and consumption, after correction for a multivariate Markov trend. This result is also obtained when per capita investment is added to the VAR.

KEY WORDS: Cointegration; Markov chain Monte Carlo; Multivariate Markov trend; Permanent income hypothesis.

1. INTRODUCTION

The *permanent income hypothesis* implies that there exists a long-run relation between consumption and disposable income (see, e.g., Flavin 1981). This theoretical result may be translated to time series properties. Most studies on the univariate properties of consumption and income series suggest that they are integrated processes (see the applications following Dickey and Fuller 1979). Hence both series must 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, e.g., Campbell 1987; Jin 1995).

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, have argued that the logarithm of many real income series contains a nonlinear cycle. This nonlinear cycle is often interpreted as the business cycle in real income. A popular model used to describe the business cycle in time series is the Markov switching model of Hamilton (1989), which allows for different average growth rates in income during expansion and recession periods, where the transitions between expansions and recessions are modeled by an unobserved first-order Markov process. We refer to the trend that models this specific behavior as a *Markov trend*. Hall, Psaradakis, and Sola (1997) considered the permanent income hypothesis under the assumption that real income contains a Markov trend. They showed that in this case, the difference between log consumption and log income is affected by changes in the mean, caused by changes in the growth rate of the real income series. The difference between log consumption and income series is not a constant mean process such that standard cointegration analysis in linear vector autoregressive (VAR) models may wrongly indicate the absence of cointegration (see Nelson, Piger, and Zivot 2001; Psaradakis 2001, 2002 for some results in univariate time series).

Several studies have considered the effects of deterministic shifts on cointegration relations (see, e.g., Gregory and Hansen 1996; Hansen and Johansen 1999; Martin 2000). In this article 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 both the income and consumption series. Our work differs from previous studies in several ways. We consider a full system cointegration analysis in a nonlinear model. We test cointegration in a VAR, which models the deviation of log per capita consumption and income from a multivariate Markov trend. This differs from the approach of Hall et al. (1997), who considered a single equation analysis and used an ad hoc procedure for cointegration analysis. Our model is a multivariate generalization of Hamilton's (1989) model and nests the theoretical results of Hall et al. (1997). Furthermore, the model allows the growth rate of consumption to be different from the growth rate in income at each stage of the business cycle, as suggested by a simple stylized facts analysis. Hence we analyze the presence of a cointegration relation between consumption and income series while allowing for different growth rates in expansion and recession periods via the multivariate Markov trend. We investigate the robustness of our results by including an investment variable in the model.

To perform econometric inference on the presence of a stable long-run relation between per capita consumption and income, we follow a Bayesian approach. We apply Markov chain Monte Carlo (MCMC) methods to evaluate posterior distributions and construct Bayes factors (BFs) to determine the cointegration rank. Our Bayesian cointegration analysis is an extension of the techniques of Kleibergen and Paap (2002) and Kleibergen and van Dijk (1998) to the case of a nonlinear VAR model containing a Markov trend.

The outline of the article is as follows. In Section 2 we give a short review on the current state of the literature on the permanent income hypothesis in cases where income is assumed

Journal of Business & Economic Statistics
October 2003, Vol. 21, No. 4
DOI 10.1198/073500103288619296

to contain a Markov trend. In Section 3 we discuss some stylized facts of U.S. per capita income and consumption series. In Section 4 we propose the multivariate Markov trend model and discuss its interpretation. Given the main application of this article, we limit the discussion to a bivariate model, but it can be easily extended to more dimensions, as shown in Section 9. In Section 5 we discuss prior specification, and in Section 6 we propose a MCMC algorithm to sample from the posterior distribution. We deal with BFs used to determine the cointegration rank in Section 7. In Section 8 we apply our multivariate Markov model to the U.S. series and relate the posterior results to suggestions made by economic theory and the stylized facts analysis. To analyze the robustness of our results, in Section 9 we consider a Markov trend model for U.S. per capita income, consumption, and investment series. We conclude the article in Section 10.

2. THE PERMANENT INCOME HYPOTHESIS AND A MARKOV TREND

The permanent income hypothesis states that current aggregate consumption is equal to a weighted average of expected future real disposable incomes. More precisely, 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 at time t, r is the interest rate, and Ω_t denotes the information set that is available to economic agents at time t. In Flavin's (1981) formulation, y_t denotes labor income solely, in which case one must add real wealth to (1). We follow Sargent's (1978) assumption that the annuity value of future capital income is equal to the value of real financial wealth (see Flavin 1981 for a critique of this assumption). Straightforward algebra shows that (1) is the forward solution of the expectational difference equation

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

Subtracting y_t from both sides of (1), we obtain

$$c_t - y_t = \frac{r}{1+r}\sum_{j=0}^{\infty}\frac{1}{(1+r)^j}\mathrm{E}[y_{t+j} - y_t|\Omega_t], \tag{3}$$

which shows that there exists a stationary relation between current consumption and income if the first difference of y_t is stationary (see, e.g., Campbell 1987). In many studies it is therefore assumed that real income follows a random-walk process (see, e.g., Jin 1995). Several other studies, however, suggest that the log income series contains a nonlinear cycle that corresponds to the business cycle (see, e.g., Goodwin 1993; Potter 1995; Peel and Speight 1998). To capture this business cycle, one often assumes that log real income is the sum of a random-walk process and a Markov trend, as suggested by Hamilton (1989). To explain the role of the Markov trend on the permanent income hypothesis, we now briefly discuss the approach of Hall et al. (1997).

The logarithm of real income is written as

$$\ln y_t = n_t + z_t, \tag{4}$$

where z_t is a standard random-walk process

$$z_t = z_{t-1} + \epsilon_t, \tag{5}$$

with $\epsilon_t \sim \mathrm{NID}(0, \sigma^2)$ and where n_t is a so-called *univariate Markov trend*. This Markov trend is defined as

$$n_t = n_{t-1} + \gamma_0 + \gamma_1 s_t, \tag{6}$$

where γ_0 and γ_1 are parameters and s_t is an unobserved binary random variable that models the business cycle. In the remainder of this article, we assume that $s_t = 0$ corresponds to an expansion observation and $s_t = 1$ corresponds to a recession. This implies that during an expansion, the slope of the Markov trend equals γ_0, whereas during a recession, the slope is given by $\gamma_0 + \gamma_1$. The random variable s_t is assumed to follow a first-order, two-state Markov process with transition probabilities

$$\begin{aligned} &\Pr[s_t = 0|s_{t-1} = 0] = p, \qquad \Pr[s_t = 1|s_{t-1} = 0] = 1 - p,\\ &\Pr[s_t = 1|s_{t-1} = 1] = q, \qquad \Pr[s_t = 0|s_{t-1} = 1] = 1 - q \end{aligned} \tag{7}$$

(see Hamilton 1989 for details).

As Hall et al. (1997) showed, (2) and (4)–(6) with $\Omega_t = \{y_t, y_{t-1}, \ldots, s_t, s_{t-1}, \ldots\}$ imply that $c_t = e^{\kappa_0} y_t$ for $s_t = 0$ and $c_t = e^{\kappa_0+\kappa_1} y_t$ for $s_t = 1$, with

$$\begin{aligned} \kappa_0 &= \ln\left(\frac{r + pe^{\kappa_0}E_0 + (1-p)e^{\kappa_0+\kappa_1}E_1}{1+r}\right),\\ \kappa_0 + \kappa_1 &= \ln\left(\frac{r + (1-q)e^{\kappa_0}E_0 + qe^{\kappa_0+\kappa_1}E_1}{1+r}\right), \end{aligned} \tag{8}$$

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

$$\ln c_t = \kappa_0 + \kappa_1 s_t + \ln y_t, \tag{9}$$

where κ_0 and κ_1 result from (8); that is,

$$\begin{aligned} \kappa_0 &= \ln\left(\frac{r(1 + (1-p-q)(1+r)^{-1}E_1)}{(1 + r - pE_0 - qE_1) - (1+r)^{-1}(1-p-q)E_0E_1}\right),\\ \kappa_1 &= \ln\left(\frac{(1+r) + (1-p-q)E_0}{(1+r) + (1-p-q)E_1}\right). \end{aligned} \tag{10}$$

Substituting (4) in the consumption–income relation (9), we note that the log of consumption can be written as

$$\ln c_t = n_t + \kappa_0 + \kappa_1 s_t + z_t, \tag{11}$$

where z_t and n_t are defined in (5) and (6). It follows that log consumption is built up of the same Markov trend as log income, and hence it corresponds to the idea that growth rates of consumption and income are the same during expansions and recessions. Note that (4) and (11) with (5) and (6) imply a stochastically singular system for y_t and c_t. To describe consumption and income series with this model, one must add extra noise to (11). Equation (9) implies that the difference between log consumption and log income is different across the phases of the business cycle and is described by the process $w_t = \kappa_0 + \kappa_1 s_t$. This process can be written as $w_t = \mu + \rho w_{t-1} + \kappa_1 v_t$, where $\mu = (1-\rho)\kappa_0 + \kappa_1(1-p)$, $\rho = (-1 + p + q)$, and v_t is a martingale difference sequence (see Hamilton 1989). This implies that (9) can be seen as a cointegration relation between log consumption and income with non-Gaussian innovations. If the transition probabilities p and q are near 1 (i.e., if both regimes are persistent), then it may be difficult to distinguish the process w_t from a random-walk process (see also Nelson et al. 2001; Psaradakis 2001, 2002). In turn, this may complicate detection of a stationary relation between log consumption and income using a standard cointegration analysis approach.

To test the presence of a stationary relation between U.S. log consumption and income, when the log of real income contains a Markov trend, we propose in Section 4 a multivariate Markov trend model. Because 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 is based on a simple stylized facts analysis of the U.S. per capita income and consumption series given in the next section.

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, 1959.1–1999.4. The series were obtained from the Federal Reserve Bank of St. Louis. Both series are increasing over the sample period with short periods of decline, for example, 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 .67% per quarter. For the consumption series, the average quarterly growth rates equals .62%. The growth rates in both series seem roughly the same.

To analyze the effect of the business cycle on real per capita income and consumption, we split the sample in two subsamples. The first subsample corresponds to quarters labeled a recession according to the National Bureau Economic of Research (NBER) peaks and troughs (see *http://www.nber.org/cycles.html*). During recessions, the average quarterly growth rate of per capita income is -1.03% and it is $-.24\%$ for consumption. The second subsample contains quarters corresponding to expansion observations. During expansions, the average quarterly growth rate in per capita income is .93%, and the average quarterly growth rate in per capita consumption is .75%. Although the average quarterly growth rates based on the whole sample are roughly the same across the two series, the average growth rates seem different in both subsamples.

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 (e.g., Jin 1995) may lead to the wrong conclusion. If the growth rates in both series are different in both stages of the business cycle, then it is unlikely that a linear combination of the two series has a constant mean. To make this more clear, Figure 2 depicts the difference between log per capita consumption and log per capita income. The graph shows that the mean of this possible cointegration relation is not constant over time, but rather displays a more-or-less changing regime pattern. This switching pattern seems to coincide with the NBER-defined business cycle.

Relating the stylized facts to the simple model in Section 2, we note that the possible changes in the mean of the difference between log consumption and income are captured by the switching constant $\kappa_0 + \kappa_1 s_t$ in (9). But the differences in growth rates of both series in each stage of the business cycle are not captured by the model, because relation (11) implies that the growth rates in both series during recessions and expansions must be the same. A consumption–income relation that allows for the former behavior is given by

$$\ln c_t = \kappa_0 + \kappa_1 s_t + \beta_2 \ln y_t. \qquad (12)$$

The trend in consumption now equals $\beta_2 n_t$, where n_t is the Markov trend in log income defined in (6). If $\beta_2 < 1$, $\kappa_0 > 0$, and $\kappa_0 + \kappa_1 < 0$, then the growth rate is smaller in consumption than in income during expansions and larger during recessions, which corresponds to our earlier findings. We note that relation (12) corresponds to a nonlinear relation between consumption and income, that is, $c_t = e^{\kappa_0 + \kappa_1 s_t} y_t^{\beta_2}$.

To analyze the permanent income hypothesis for the U.S. consumption and income series, we propose a multivariate Markov trend model in the next section. This multivariate model is an extension of Hamilton's univariate model. The

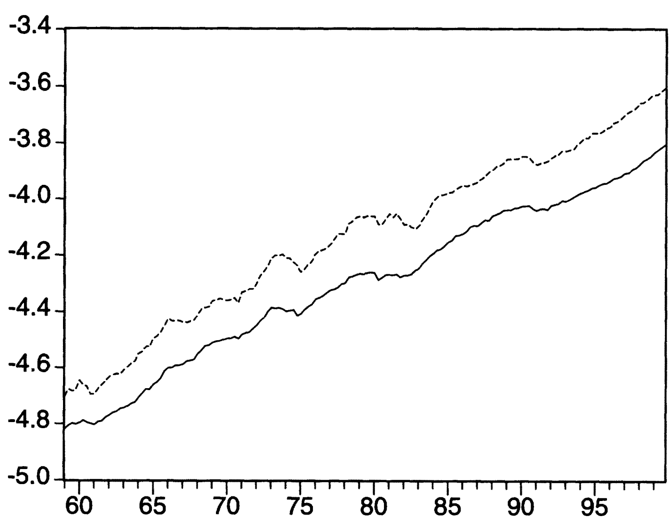

Figure 1. Logarithm of U.S. Per Capita Consumption (——) and Income (- - - -), 1959.1–1999.4.

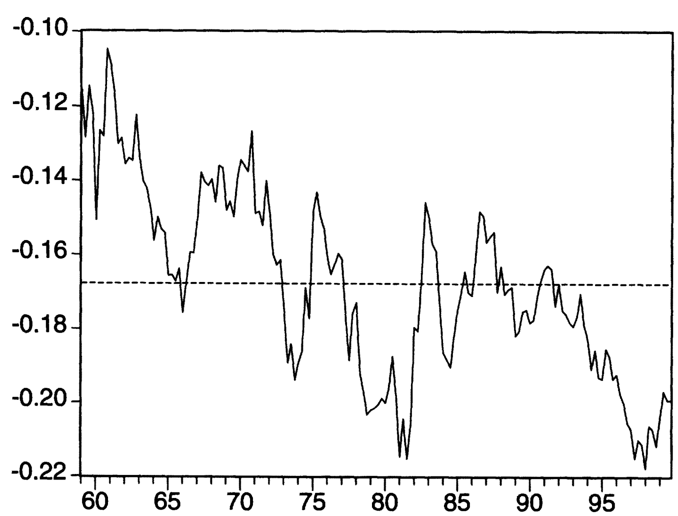

Figure 2. Difference Between Log Per Capita Consumption and Log Per Capita Income, 1959.1–1999.4.

model contains a multivariate Markov trend that allows for different growth rates in the consumption and income series during recessions and expansions. The deviations from the Markov trend are modeled by a VAR model. To analyze the presence of a consumption–income relation, we perform a cointegration analysis on the deviations from the multivariate Markov trend. In addition, we investigate 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.

4. THE MULTIVARIATE MARKOV TREND MODEL

In this section we propose the multivariate Markov trend model on which we base our analysis of the consumption–income relation. This model is a multivariate generalization of the model proposed by Hamilton (1989), where the slope of the multivariate Markov trend is different across series and across the regimes. The regime changes occur simultaneously in all series. The deviations from the Markov trend are modeled by a VAR model, which may contain unit roots. A similar representation was suggested by Dwyer and Potter (1996).

In Section 4.1 we discuss representation, and in Section 4.2 we deal with model interpretation. In Section 4.3 we derive the likelihood function of the model. Although we explain the model for bivariate time series, the discussion can be easily extended to more than two time series, as shown in Section 9.

4.1 Representation

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

$$Y_t = N_t + R_t + Z_t, \tag{13}$$

where N_t represents a trend component, R_t allows for possible level shifts, and Z_t represents the deviations from N_t and R_t. The two-dimensional trend component N_t is a multivariate generalization of the univariate Markov trend (6), that is,

$$N_t = N_{t-1} + \Gamma_0 + \Gamma_1 s_t, \tag{14}$$

where Γ_0 and Γ_1 are (2×1) parameter vectors and s_t is an unobserved first-order Markov process with transition probabilities given in (7). Kim and Yoo (1995) added an extra normally distributed error term to (14), but we do not pursue this here, because it a priori imposes a unit root in the series Y_t (see also Luginbuhl and de Vos 1999). We allow unit roots to enter Y_t only through Z_t; see also Section 4.2. The value of the unobserved state variable s_t models the stages of the business cycle. If $s_t = 0$ (expansion), then the slope of the Markov trend is Γ_0, whereas for $s_t = 1$ (recession), the slope equals $\Gamma_0 + \Gamma_1$ (see also Hamilton 1989). The values of the slopes of the trends in the individual series in Y_t do not have to be the same even though the changes in the value of the slope occur simultaneously. The latter assumption can be relaxed (see, e.g., Phillips 1991), but this extension is not necessary for the application in this article. The expected slope value of the Markov trend equals $\Gamma_0 + \Gamma_1(1-p)/(2-p-q)$ (see Hamilton 1989). Hence one may have different slopes values in each regime but the same expected slope. The backward solution of (14) equals

$$N_t = \Gamma_0(t-1) + \Gamma_1 \sum_{i=2}^{t} s_i + N_1, \tag{15}$$

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 Γ_0 and a stochastic trend $\sum_{i=2}^t s_i$ with impact vector Γ_1.

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

$$R_t = \begin{pmatrix} \delta_1 \\ 0 \end{pmatrix} s_t = \delta s_t, \tag{16}$$

such that $\delta = (\delta_1 \ 0)'$. This term takes care of level shifts in the consumption series during recessions, as suggested by the theory in Section 2. (See Krolzig 1997, chap. 13, for a similar discussion about the role of this term.) The parameter δ_1 turns out to be related to the κ_1 parameter in (9), as discussed at the end of Section 4.2.

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

$$Z_t = \sum_{i=1}^{k} \Phi_i Z_{t-i} + \varepsilon_t \tag{17}$$

or, using the lag polynomial notation,

$$\Phi(L) Z_t = (I - \Phi_1 L - \cdots - \Phi_k L^k) Z_t = \varepsilon_t, \tag{18}$$

where ε_t is a two-dimensional vector normally distributed process with mean 0 and a (2×2) positive-definite symmetric covariance matrix Σ, and where Φ_i, $i = 1, \ldots, k$, are (2×2) parameter matrices.

4.2 Model Interpretation

For our analysis of a potentially stationary relation between log consumption and income, it is convenient to write (17) in error-correction form,

$$\Delta Z_t = \Pi Z_{t-1} + \sum_{j=1}^{k-1} \bar{\Phi}_j \Delta Z_{t-j} + \varepsilon_t, \tag{19}$$

where $\Pi = \sum_{j=1}^k \Phi_j - I$ 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

$$\left| I - \Phi_1 z - \cdots - \Phi_k z^k \right| = 0. \tag{20}$$

We can now distinguish three cases depending on the number of unit root solutions of the characteristic equation (20). The first case corresponds to the situation where the solutions of (20) are outside the unit circle. The process Z_t is stationary, and hence Y_t is a stationary process around a multivariate Markov trend. This is in fact the multivariate extension of the model proposed by Lam (1990). We can write

$$\begin{aligned}
&(\Delta Y_t - \Gamma_0 - \Gamma_1 s_t - \delta \Delta s_t) \\
&\quad = \Pi \left(Y_{t-1} - \Gamma_0(t-2) - \Gamma_1 \sum_{i=2}^{t-1} s_i - N_1 - \delta s_{t-1} \right) \\
&\qquad + \sum_{i=1}^{k-1} \bar{\Phi}_i (\Delta Y_{t-i} - \Gamma_0 - \Gamma_1 s_{t-i} - \delta \Delta s_{t-i}) + \varepsilon_t,
\end{aligned} \tag{21}$$

where Π a full-rank matrix. The vectors Γ_0 and $\Gamma_0 + \Gamma_1$ contain the slopes of the trend in Y_t during expansions and recessions. 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 $s_t = 0$, then the initial value of the Markov trend equals N_1, whereas for $s_t = 1$, this value equals $N_1 + \delta s_t$.

The second case concerns the situation of two unit root solutions of (20) with the remaining roots outside the unit circle. In that case, $\Pi = 0$, and (21) becomes

$$(\Delta Y_t - \Gamma_0 - \Gamma_1 s_t - \delta \Delta s_t) = \sum_{i=1}^{k-1} \bar{\Phi}_i (\Delta Y_{t-i} - \Gamma_0 - \Gamma_1 s_{t-i} - \delta \Delta s_{t-i}) + \varepsilon_t. \quad (22)$$

The first difference of Y_t is a stationary VAR process with a stochastically changing mean ($= \Gamma_0 + \Gamma_1 s_t$). Note that the initial value of the Markov trend N_1 drops out of the model. If $s_t = s_{t-1}$, then ΔY_t is not affected by R_t. If, however, $s_t \neq s_{t-1}$, then the growth rate in consumption is δ_1 larger or smaller than the growth rate in income. A change in the stage of the business cycle leads to a one-time extra adjustment in the growth rate of per capita consumption. This adjustment is absent if $\delta_1 = 0$, in which case the model simplifies to the one considered by Kim and Nelson (1999a), who a priori imposed that $\Pi = 0$ (see also Hamilton and Perez-Quiros 1996).

The third case corresponds to the situation where only one of the roots equals unity and the other roots are outside the unit circle. The series in Z_t are said to be cointegrated (see Johansen 1995 for a discussion on cointegration). Under cointegration, the rank of Π equals 1, and we can write Π as $\alpha\beta'$, where α and β are (2×1) vectors. The β vector describes 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. Because the number of free parameters in α and β is larger than in Π under rank reduction, the parameters in α or β must be restricted to become estimable. We choose to impose the restriction $\beta = (1 \ \ -\beta_2)'$. Under cointegration, model (21) becomes

$$\begin{aligned}(\Delta Y_t - \Gamma_0 - \Gamma_1 s_t - \delta \Delta s_t) &= \alpha\beta' \left(Y_{t-1} - \Gamma_0 (t-2) - \Gamma_1 \sum_{i=2}^{t-1} s_i - N_1 - \delta s_t \right) \\ &\quad + \sum_{i=1}^{k-1} \bar{\Phi}_i (\Delta Y_{t-i} - \Gamma_0 - \Gamma_1 s_{t-i} - \delta \Delta s_{t-i}) + \varepsilon_t. \quad (23)\end{aligned}$$

The cointegration relation is given by $\beta' Y_t = \beta'(N_t + R_t + Z_t)$. For $\beta'\Gamma_0 = \beta'\Gamma_1 = 0$, $\kappa_0 = \beta' N_1$, and $\kappa_1 = \beta'\delta$, we obtain the consumption–income relation (12). The extra condition $\beta_2 = 1$ leads to relation (9). Finally, note that the restriction $\beta'\Gamma_1 = 0$ removes the Markov trend from the cointegration relation. Dwyer and Potter (1996) refered 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 of the least-restricted Markov trend stationary model [(21)] 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

$$f(Y_t | Y^{t-1}, s^t, \Gamma_0, \Gamma_1, N_1, \delta_1, \Sigma, \Pi, \bar{\Phi}) = \frac{1}{(\sqrt{2\pi})^2} |\Sigma|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \varepsilon_t' \Sigma^{-1} \varepsilon_t \right), \quad (24)$$

where ε_t is given in (21) and $\bar{\Phi} = \{\bar{\Phi}_1, \ldots, \bar{\Phi}_{k-1}\}$. Hence the likelihood function for model (21), conditional on the states s^T and the first k initial observations Y^k, is

$$\mathcal{L}_2(Y^T | Y^k, s^T, \Theta_2) = p^{\mathcal{N}_{0,0}} (1-p)^{\mathcal{N}_{0,1}} q^{\mathcal{N}_{1,1}} (1-q)^{\mathcal{N}_{1,0}} \times \prod_{t=k+1}^{T} f(Y_t | Y^{t-1}, s^t, \Gamma_0, \Gamma_1, N_1, \Sigma, \Pi, \bar{\Phi}), \quad (25)$$

where $\Theta_2 = \{\Gamma_0, \Gamma_1, N_1, \delta_1, \Sigma, \Pi, \bar{\Phi}, p, q\}$ 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,

$$\mathcal{L}_2(Y^T | Y^k, \Theta_2) = \sum_{s_1} \sum_{s_2} \cdots \sum_{s_T} \mathcal{L}_2(Y^T | Y^k, s^T, \Theta_2). \quad (26)$$

The unconditional likelihood function for the Markov trend model with one cointegration relation [(23)] follows directly from (26),

$$\mathcal{L}_1(Y^T | Y^k, \Theta_1) = \mathcal{L}_2(Y^T | Y^k, \Theta_2)|_{\Pi = \alpha\beta'}, \quad (27)$$

with $\Theta_1 = \{\Gamma_0, \Gamma_1, N_1, \delta_1, \Sigma, \alpha, \beta_2, \bar{\Phi}, p, q\}$. In case of no cointegration [(22)], the unconditional likelihood function is given by

$$\mathcal{L}_0(Y^T | Y^k, \Theta_0) = \mathcal{L}_2(Y^T | Y^k, \Theta_2)|_{\Pi = 0}, \quad (28)$$

with $\Theta_0 = \{\Gamma_0, \Gamma_1, N_1, \delta_1, \Sigma, \bar{\Phi}, p, q\}$. Note that the subscript "r" of Θ_r and $\mathcal{L}_r$ refers 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

To perform inference on the parameters of the multivariate Markov trend model and on the presence of a stationary relation between consumption and income, we opted for a Bayesian approach. We chose to impose prior information, which is relatively uninformative compared with the information in the likelihood. The Markov trend model is nonlinear in certain parameters, which leads to local nonidentification for certain parameters in the model. In sum, we must deal with three types of identification issues: the initial value identification (N_1), the regime identification (Γ_0 and Γ_1) and the identification of β_2 in the reduced-rank model (23). To tackle these identification problems, we proceed as follows.

It follows from (21) that the parameter N_1 drops out of the model in case $\Pi = 0$. Even in the case of rank reduction in Π it follows from (23) that we can identify only βN_1. Specifying a diffuse prior on N_1 implies that the conditional posterior of N_1 given Π is constant and nonzero at the point of rank reduction. The integral over this conditional posterior at the point of rank reduction is, therefore, infinity, favoring rank reduction (see Schotman and van Dijk 1991a,b for a related discussion of identification problems associated with the intercept term in univariate autoregressions). To circumvent this identification problem, we follow the prior specification of Zivot (1994) (see also Hoek 1997, chap. 2). The prior distribution for N_1 conditional on both Σ and the first observation Y_1 is normal with mean Y_1 and covariance Σ,

$$N_1|Y_1,\Sigma \sim \mathrm{N}(Y_1,\Sigma). \tag{29}$$

For Σ, we take a standard inverted Wishart prior with scale parameter S and degrees of freedom ν,

$$p(\Sigma) \propto |S|^{\frac{1}{2}\nu}|\Sigma|^{-\frac{1}{2}(\nu+3)}\exp\left(-\frac{1}{2}\Sigma^{-\frac{1}{2}}S\right). \tag{30}$$

If we do not want to impose an informative prior for Σ, then we opt for $p(\Sigma)\propto|\Sigma|^{-1}$, which results from (30) by letting the degrees of freedom approach 0 (see Geisser 1965).

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

$$\begin{aligned} p(p) &= \mathbb{I}_{(0,1)}, \\ p(q) &= \mathbb{I}_{(0,1)}, \end{aligned} \tag{31}$$

where $\mathbb{I}_{(0,1)}$ represents an indicator function that is 1 on the interval (0,1) and 0 elsewhere. Under flat priors for p and q, special attention must be given to the priors for Γ_0 and Γ_1. It is easy to show that under $\Pi=0$, the likelihood has the same value if we switch the role of the states and change the values of Γ_0, Γ_1, δ, p, and q into $\Gamma_0+\Gamma_1$, $-\Gamma_1$, $-\delta$, q, and p. This complicates proper posterior analysis if we specify uninformative priors on Γ_0 and Γ_1. There are several ways of identifying the parameters. One could, for example, specify appropriate matrix normal prior distributions for Γ_0 and Γ_1. But we define priors for Γ_0 and Γ_1 on subspaces that identify the regimes for all specifications of the model. Several specifications for these subspaces are possible. With our present application in mind, we restrict the growth rates in the income series to be positive during expansions and negative in recessions. This results in the prior specification

$$\begin{aligned} p(\Gamma_0) &\propto \begin{cases} 1 & \text{if } \Gamma_0\in\{\Gamma_0\in\mathbb{R}^2|\Gamma_{0,2}>0\} \\ 0 & \text{elsewhere,} \end{cases} \\ p(\Gamma_1|\Gamma_0) &\propto \begin{cases} 1 & \text{if } \Gamma_1\in\{\Gamma_1\in\mathbb{R}^2|\Gamma_{0,2}+\Gamma_{1,2}\le 0\} \\ 0 & \text{elsewhere.} \end{cases} \end{aligned} \tag{32}$$

Note that because we have identified the two regimes by the prior on Γ_0 and Γ_1, we may use an improper prior for δ_1,

$$p(\delta_1)\propto 1. \tag{33}$$

For the autoregressive parameters apart from Π, we also use flat priors,

$$p(\bar{\Phi}_i)\propto 1, \qquad i=1,\dots,k-1. \tag{34}$$

The three model specifications are different with respect to the rank of Π. Under cointegration, the rank of Π equals 1, and we can write $\Pi=\alpha\beta'$. It is easy to see that if $\alpha=0$, then β_2 is not identified (see Kleibergen and van Dijk 1994 for a general discussion). To solve this identification problem, we follow the approach of Kleibergen and Paap (2002) (see also Kleibergen and van Dijk 1998 for a similar approach in simultaneous equation models). A convenient byproduct of this approach is a Bayesian posterior odds analysis for the rank of Π; see also Section 7. The analysis is based on the following decomposition of the matrix Π:

$$\Pi=\alpha\beta'+\alpha_\perp\lambda\beta_\perp', \tag{35}$$

where $\alpha_\perp$ and $\beta_\perp$ are specified such that $\alpha_\perp'\alpha=0$ with $\alpha_\perp'\alpha_\perp=1$ and $\beta_\perp'\beta=0$ with $\beta_\perp'\beta_\perp=1$. It is easy to see that cointegration (i.e., rank reduction in Π) occurs if $\lambda=0$, and hence the parameter λ can be used to test for cointegration. The matrix $(\alpha_\perp\lambda\beta_\perp')$ models the deviation from cointegration. The row and column spaces of this matrix are spanned by the orthogonal complements of the vector of adjustment parameters α and the cointegrating vector β. The decomposition in (35) is not unique, however. To identify α and β, we impose the condition that $\beta=(1\ \ -\beta_2)'$, as is often done in cointegration analysis. To identify λ, $\alpha_\perp$, and $\beta_\perp$ in $\alpha_\perp\lambda\beta_\perp'$, we relate λ to the smallest singular value of Π. Note that singular values determine the rank of Π in an unambiguous way.

The singular value decomposition of Π is given by

$$\Pi=USV', \tag{36}$$

where U and V are (2×2) orthonormal matrices and S is an (2×2) diagonal matrix containing the positive singular values of Π (in decreasing order) (see Golub and van Loan 1989, p. 70). If we write

$$U=\begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22}\end{pmatrix}, \qquad S=\begin{pmatrix} s_{11} & 0 \\ 0 & s_{22}\end{pmatrix}, \qquad \text{and} \qquad V=\begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22}\end{pmatrix}, \tag{37}$$

with u_{ij}, s_{ij}, v_{ij}, $i=1,2$, $j=1,2$, scalars and use that

$$(\alpha\ \alpha_\perp)\begin{pmatrix} 1 & 0 \\ 0 & \lambda\end{pmatrix}(\beta'\ \beta_\perp')=USV', \tag{38}$$

then we obtain the following expressions for α and β_2:

$$\begin{aligned} \alpha &= \begin{pmatrix} u_{11}s_{11}v_{11} \\ u_{21}s_{11}v_{11}\end{pmatrix} \\ \beta_2 &= -v_{21}/v_{11}. \end{aligned} \tag{39}$$

Identification of λ follows from the fact that we must express $\alpha_\perp$ and $\beta_\perp$ in terms of u_{11}, u_{21}, s_{11}, v_{11} and v_{21} to obtain a one-to-one relation with the singular value decomposition. Kleibergen and Paap (2002) showed that if we take

$$\alpha_\perp=\sqrt{u_{22}^2}\begin{pmatrix} u_{12}u_{22}^{-1} \\ 1\end{pmatrix} \qquad \text{and} \qquad \beta_\perp=\sqrt{v_{22}^2}\begin{pmatrix} v_{22}^{-1}v_{12} \\ 1\end{pmatrix}, \tag{40}$$

then λ is identified by

$$\lambda=\frac{u_{22}s_{22}v_{22}}{\sqrt{u_{22}^2}\sqrt{v_{22}^2}}=\mathrm{sign}(u_{22}v_{22})s_{22}, \tag{41}$$

where "sign(·)" denotes the sign of the argument. Hence the absolute value of λ is equal to the smallest singular value of Π that corresponds to s_{22}. Note that λ can be positive and negative, in contrast with the singular value s_{22}, which is always positive. Golub and van Loan (1989) showed that the number of nonzero eigenvalues of a matrix completely determine the rank of a matrix. Restricting the scalar λ to equal 0 is, therefore, an unambiguous way of restricting the rank of Π and imposing cointegration.

To construct priors for the α and β_2 parameters that take into account the identification problem, we take as starting point the prior for Π given Σ, denoted by $p(\Pi|\Sigma)$. Because the matrix Π can be decomposed using (35), $p(\Pi|\Sigma)$ implies the following joint prior for α, λ, and β_2 given Σ:

$$p(\alpha, \lambda, \beta_2|\Sigma) \propto p(\Pi|\Sigma)|_{\Pi=\alpha\beta'+\alpha_\perp\lambda\beta'_\perp} |J(\alpha, \lambda, \beta_2)|, \quad (42)$$

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 are given in Appendix A. Because restricting λ to equal 0 is an unambiguous way of restricting the rank of Π and imposing cointegration, we construct the joint prior for α and β_2 by restricting (42) in $\lambda = 0$,

$$\begin{aligned} p(\alpha, \beta_2|\Sigma) &\propto p(\alpha, \lambda, \beta_2|\Sigma)|_{\lambda=0} \\ &\propto p(\Pi|\Sigma)|_{\Pi=\alpha\beta'} |J(\alpha, \lambda, \beta_2)|_{\lambda=0}. \end{aligned} \quad (43)$$

The posterior resulting from this prior leads to proper posterior distributions for α and β_2 (see Kleibergen and Paap 2002). The posteriors are also unique in the sense that they do not depend on the ordering of the variables in the system and the normalization to identify α and β [in our case, $\beta = (1 \;\; -\beta_2)'$]. The proposed strategy for prior construction for α and β_2 can be carried out for a proper or an improper prior specification on Π given Σ. In this article we opt for a normal prior on Π given Σ with mean P and covariance matrix $(\Sigma \otimes A^{-1})$,

$$p(\Pi|\Sigma) \propto |\Sigma|^{-1}|A| \exp\left(-\frac{1}{2}\,\mathrm{tr}\big(\Sigma^{-1}(\Pi - P)'A(\Pi - P)\big)\right). \quad (44)$$

Hence the prior for α and β_2 is given by

$$\begin{aligned} &p(\alpha, \beta_2|\Sigma) \\ &\quad \propto |\Sigma|^{-1}|A| \exp\left(-\frac{1}{2}\,\mathrm{tr}\left(\Sigma^{-1}(\alpha\beta' - P)'A(\alpha\beta' - P)\right)\right) \\ &\quad \times |J(\alpha, \lambda, \beta_2)|_{\lambda=0}. \end{aligned} \quad (45)$$

If one prefers a noninformative prior, then one may consider $p(\Pi|\Sigma) \propto 1$ in combination with $p(\Sigma) \propto |\Sigma|^{-1}$. In that case, the resulting prior for α and β_2, given Σ, is $p(\alpha, \beta_2|\Sigma) \propto |J(\alpha, \lambda, \beta_2)|_{\lambda=0}$.

The joint priors for the Markov trend models with different numbers of unit roots follow from the marginal priors in this section. The joint prior for the Markov trend stationary model [(21)], $p_2(\Theta_2)$, is given by the product of (29)–(34) and (44). The prior for the Markov trend model with one cointegration relation [(23)], $p_1(\Theta_1)$, is the product of (29)–(34) and (45), whereas the prior for the model without cointegration [(22)], $p_0(\Theta_0)$, is simply the product of (29)–(34).

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 enable the analytical derivation of posterior results. As Albert and Chib (1993), McCulloch and Tsay (1994), Chib (1996), and Kim and Nelson (1999b) have demonstrated, the Gibbs sampling algorithm of Geman and Geman (1984) is a 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 in which one consecutively samples from the full conditional posterior distributions of the model parameters. This produces a Markov chain that converges under mild conditions. The resulting draws can be considered as a sample from the posterior distribution. (For details on the Gibbs sampling algorithm, see Smith and Roberts 1993; Tierney 1994.) In Appendix B we derive the full conditional posterior distributions associated with the most general Markov trend stationary model [(21)]. The full conditional posterior distributions associated with 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 a Metropolis–Hastings step into the Gibbs sampler (see Chib and Greenberg 1995 for a discussion).

7. DETERMINING THE COINTEGRATION RANK

To determine the cointegration rank, we begin by assigning prior probabilities to every possible rank of Π

$$\Pr[\mathrm{rank} = r], \qquad r = 0, 1, 2. \quad (46)$$

This is equivalent to assigning prior probabilities to the different possible number of cointegration relations, r. The prior probabilities imply the following prior odds ratios (PRORs):

$$\mathrm{PROR}(r|2) = \frac{\Pr[\mathrm{rank} = r]}{\Pr[\mathrm{rank} = 2]}, \qquad r = 0, 1, 2. \quad (47)$$

The Bayes factor (BF) for comparing rank r with rank 2 equals

$$\mathrm{BF}(r|2) = \frac{\int \mathcal{L}_r(Y^T|Y^k, \Theta_r)p_r(\Theta_r)d\Theta_r}{\int \mathcal{L}_2(Y^T|Y^k, \Theta_2)p_2(\Theta_2)d\Theta_2}, \qquad r = 0, 1, \quad (48)$$

where $\mathcal{L}_r(Y^T|Y^k, \Theta_r)$ denotes the unconditional likelihood function and $p_r(\Theta_r)$ denotes the joint prior of the model with rank r. The posterior odds ratio (POR) for comparing rank r with rank 2 equals the PROR times the BF, $\mathrm{POR}(r|2) = \mathrm{PROR}(r|2) \times \mathrm{BF}(r|2)$, and the posterior probabilities for each rank are simply

$$\Pr[\mathrm{rank} = r|Y^T] = \frac{\mathrm{POR}(r|n)}{\sum_{i=0}^2 \mathrm{POR}(i|2)}, \qquad r = 0, 1, 2. \quad (49)$$

The BFs in (48) are in fact BFs for $\Pi = 0$ and $\lambda = 0$. They can be computed using the Savage–Dickey density ratio of Dickey (1971), which states that the BF for $\Pi = 0$ (or $\lambda = 0$)

equals the ratio of the marginal posterior density and the marginal prior density of Π (λ), both evaluated in $\Pi = 0$ ($\lambda = 0$),

$$\mathrm{BF}(0|2) = \frac{p(\Pi|Y^T)|_{\Pi=0}}{p(\Pi)|_{\Pi=0}},$$
$$\mathrm{BF}(1|2) = \frac{p(\lambda|Y^T)|_{\lambda=0}}{p(\lambda)|_{\lambda=0}}. \tag{50}$$

This means that we need the marginal posterior densities of Π and λ to compute these Savage–Dickey density ratios. 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 0 over the sampled model parameters (see Gelfand and Smith 1990). This approach cannot be used for λ, because 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 (2002) argued that the density function $g(\lambda|\Theta_1, Y^T)$ defined in (B.13) is a good approximation. This results in the following expression to compute the marginal posterior height at $\lambda = 0$:

$$p(\lambda|Y^T)|_{\lambda=0} \approx \frac{1}{N}\sum_{i=1}^{N} \frac{|J(\alpha^i, \lambda, \beta_2^i)|_{\lambda=0}|}{|J(\alpha^i, \lambda^i, \beta_2^i)|} g(\lambda|\Theta_1^i, Y^T)|_{\lambda=0}, \tag{51}$$

where N denotes the number of simulations. Note that we can avoid the importance weights by using numerical integration to determine the integrating constant of the full posterior conditional distribution of λ in every Gibbs step.

Because we have a closed form for the prior density of Π, we can compute the prior height of Π at $\Pi = 0$ directly. To compute the prior height of λ, we follow a similar procedure as for the posterior height. First, we sample from the prior of Σ and Π, given Σ. Next, we perform a singular value decomposition on the sampled Π^i (37), resulting in λ^i α^i and β_2^i. To compute the marginal prior height of λ at $\lambda = 0$, we may use a kernel estimator on the sampled λ^i. Again, it is possible to use an approximation of the full conditional prior of λ in combination with importance weights. The prior height can be computed as

$$p(\lambda)|_{\lambda=0} \approx \frac{1}{N}\sum_{i=1}^{N} \frac{|J(\alpha^i, \lambda, \beta_2^i)|_{\lambda=0}|}{|J(\alpha^i, \lambda^i, \beta_2^i)|} h(\lambda|\Theta_1^i)|_{\lambda=0}, \tag{52}$$

where $h(\lambda|\Theta_1)$ is an approximation of the full conditional prior distribution of λ. An appropriate candidate for h turns out to be

$$h(\lambda|\Theta_1) = (2\pi)^{-\frac{1}{2}} |\alpha_\perp \Sigma^{-1} \alpha_\perp'|^{\frac{1}{2}} |\beta_\perp' A \beta_\perp|^{\frac{1}{2}} \times \exp\left(-\frac{1}{2}\operatorname{tr}\left(\beta_\perp' A \beta_\perp (\lambda - l)\alpha_\perp \Sigma^{-1} \alpha_\perp' (\lambda - l)'\right)\right), \tag{53}$$

with $l = (\beta_\perp' A \beta_\perp)^{-1} \beta_\perp' A (P - \beta\alpha)\Sigma^{-1}\alpha_\perp' (\alpha_\perp \Sigma^{-1} \alpha_\perp')^{-1}$.

Finally, if we specifies an improper prior for Π and λ, then the height of the marginal prior at $\Pi = 0$ and $\lambda = 0$ is not defined. Therefore, the BFs in (50) are not properly defined in cases of diffuse priors. Kleibergen and Paap (2002) argued that a Bayesian cointegration analysis under a diffuse prior specification on Π is possible if one replaces the prior height by the factor $(2\pi)^{-\frac{1}{2}(2-r)^2}$. This leads to a BF that corresponds to the posterior information criterion (PIC) of Phillips and Ploberger (1994). We opt for the same solution in this article.

8. U.S. CONSUMPTION AND INCOME

In this section we analyze the presence of a long-run relation between the U.S. 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 VAR with a linear deterministic trend to illustrate the effects of neglecting the presence of a possible 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 proposed in Section 4.

8.1 A Vector Autoregressive Model Without Markov Trend

If we restrict Γ_1 and δ_1 in the Markov trend model (21) to 0, then we end up with a VAR for Y_t with only a linear deterministic trend. In this section we analyze the presence of a cointegration relation between U.S. per capita consumption and income in this VAR for $Y_t = 100 \times (\ln c_t, \ln y_t)'$. The priors for N_1 and Σ are given by (29) and (30), with $S = I$ and $\nu = 3$. For Π, given Σ, we opt for a g-type prior (see Zellner 1986). This prior is given in (44) with $P = 0$ and $A = \tau/T\sum_{t=1}^{T} \bar{Y}_t'\bar{Y}_t$ for different values of τ, where $\bar{Y}_t$ denotes the demeaned and detrended value of Y_t. Because we are dealing with nonstationary time series, we divide by the number of observations, T (see Kleibergen and Paap 2002 for a similar approach). A smaller value of τ implies less precision in the prior information on $\Pi|\Sigma$. For Γ_0 and $\bar{\Phi}_i$ we take flat priors $p(\Gamma_0) \propto 1$ and $p(\bar{\Phi}_i) \propto 1$.

Before beginning our analysis, we must choose the lag order k of the VAR model. To determine the lag order, we sequentially test for the significance of an extra lag using PIC-based Bayes factors starting with $k = 1$. Given this strategy we find that $k = 2$. We note that the same lag order is found using the Bayes information criterion (BIC) of Schwarz (1978) to determine k. For the cointegration analysis, we assign equal prior probabilities to the possible cointegration ranks (46), that is, $\Pr[\text{rank} = r] = \frac{1}{3}$ for $r = 0, 1, 2$. The prior for α and β_2 for the cointegration specification (rank=1) is given by (45).

Columns 2–7 in the first panel of Table 1 shows log BFs and posterior probabilities for the cointegration rank r for different values of τ. The results show that a model with a rank of Π of 0 or 1 is preferred over a model with full rank for Π. The log BFs computed for the model with rank 0 versus the model with rank 1 are 4.20 (6.69–2.49), 7.64, and 11.09 for τ equal to 1, .1, and .01. Hence the model with no cointegration relation is preferred over the model with 1 cointegration relation. The BFs lead to the assignment of 98% posterior probability to the model with no cointegration relation if $\tau = 1$ and 100% for the other values of τ. In sum, there is no evidence for a long-run equilibrium between U.S. per capita consumption and income in a VAR model with only a linear deterministic trend. [The standard trace tests for rank reduction (Johansen 1995) also do

Table 1. Log BFs and Posterior Probabilities for the Cointegration Rank in a Linear VAR Model (k = 2) and the Multivariate Markov Trend Model (k = 1)

	$\tau = 1$		$\tau = .1$		$\tau = .01$		PIC	
r	$\ln \mathrm{BF}(r\|2)$	$\Pr[r\|Y^T]$	$\ln \mathrm{BF}(r\|2)$	$\Pr[r\|Y^T]$	$\ln \mathrm{BF}(r\|2)$	$\Pr[r\|Y^T]$	$\ln \mathrm{BF}(r\|2)$	$\Pr[r\|Y^T]$
Linear VAR model								
0	6.69	.98	11.28	1.00	15.88	1.00	11.66	1.00
1	2.49	.02	3.64	0	4.79	0	4.55	0
2	0	0	0	0	0	0	0	0
Multivariate Markov trend model								
0	< -5	0	< -5	0	< -5	1.00	< -5	0
1	1.82	.86	2.98	.95	4.27	.99	3.96	.98
2	0	.14	0	.05	0	.01	0	.02

NOTE: A log BF, $\ln \mathrm{BF}(r|2) > 0$, denotes that a cointegration model with r cointegration relations is more likely than a model with two cointegration relations. The posterior probability of the cointegration rank, $\Pr[r|Y^T]$, is defined in (49) and based on equal prior probabilities (46) for every rank r. Posterior results are based on 400,000 iterations with the Gibbs sampler, neglecting the first 100,000 draws.

not indicate the presence of a cointegration relation between the two series.] Unreported results show that this finding is robust with respect to the chosen lag order. The log BFs for models with order $2 < k \leq 5$ are very similar to those reported in Table 1.

The results in Table 1 show that if we increase the prior variance of Π by decreasing τ, then the evidence for rank reduction, and hence the presence of unit roots, increases. This is due to the fact that our prior is centered at $\Pi = 0$. When we increase the prior variance, the prior height at $\Pi = 0$ decreases. The posterior height at $\Pi = 0$ remains almost the same, because the value of τ is so small that the prior has only a minimal effect on the posterior. From Section 7, we have seen that the BF for $\Pi = 0$ equals the ratio of the posterior and prior heights at $\Pi = 0$, and hence that too small a value of τ leads to rank reduction being favored, no matter what the nature of the sample evidence. This phenomenon is known as the *Lindley paradox* (see Zellner 1971). In the second-to-last column of the first panel of Table 1, we report the log BFs for improper priors on Π and Σ, that is, $p(\Pi, \Sigma) \propto |\Sigma|^{-1}$. Under this prior specification, BFs are not defined properly. Instead, we report a PIC-based BF, where we replace the prior heights in (50) by the penalty function $(2\pi)^{-\frac{1}{2}(2-r)^2}$. These BFs again indicate that rank reduction is preferred to the full-rank case and lead to the assignment of 100% posterior probability to the model with no cointegration relation.

With no cointegration imposed, the estimated VAR model is

$$\hat{Y}_t = \hat{N}_t + \hat{Z}_t,$$

$$\hat{N}_t = -\begin{pmatrix} 481.70 \\ (.69) \\ 469.61 \\ (1.03) \end{pmatrix} + \begin{pmatrix} .63 \\ (.06) \\ .68 \\ (.07) \end{pmatrix} (t-1),$$

$$\hat{Z}_t = \begin{pmatrix} 1.11 & .08 \\ (.10) & (.06) \\ .59 & .94 \\ (.16) & (.10) \end{pmatrix} Z_{t-1} + \begin{pmatrix} -.09 & -.11 \\ (.10) & (.06) \\ -.47 & -.07 \\ (.16) & (.10) \end{pmatrix} Z_{t-2} \qquad (54)$$

$$+ \hat{\varepsilon}_t, \quad \text{with}$$

$$\hat{\Sigma} = \begin{pmatrix} .48 & .45 \\ (.06) & (.07) \\ .45 & 1.14 \\ (.07) & (.13) \end{pmatrix},$$

where the point estimates are posterior means based on the improper prior specification discussed earlier and posterior standard deviations are given in parentheses. Note that this model is equal to (13)–(17) with $\Gamma_1 = 0$, $\delta_1 = 0$, and $k = 1$. The posterior means of the slopes of the deterministic trends in the consumption and income series are .63% and .68%. They differ by about .01% from the average quarterly growth rates reported in Section 3. Note that this difference is small compared with the posterior standard deviations of the slopes.

8.2 A Bivariate 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 there may be different growth rates in both series during expansions and recessions. To allow for the possibility of different growth rates in consumption and income during recessions and expansions, we consider the Markov trend model (21). The prior for the model parameters is given by (29)–(34) with $S = I$ and $\nu = 3$. For Π given Σ, we again use the same g-type prior as for the non-Markov model. The prior is given in (44) with $P = 0$ and $A = \tau/T \sum_{t=1}^{T} \tilde{Y}_t' \tilde{Y}_t$, where $\tilde{Y}_t$ denotes the demeaned and detrended value of Y_t.

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. To determine the lag order of the VAR part of the model, we use the same strategy as for the non-Markov model. It turns out that one lag is sufficient, and hence we impose $k = 1$. We assign equal probabilities to the possible cointegration ranks, that is, $\Pr[\text{rank} = r] = \frac{1}{3}$ for $r = 0, 1, 2$. The prior for α and β_2 for the cointegration specification ($r = 1$) is given by (45). Columns 2–7 of the second panel of Table 1 report the log BFs and posterior probabilities for the rank of Π for different values of τ. Comparing the corresponding results in the first panel, where we show the results for the model without Markov trend, we see that all log BFs are smaller. Not surprisingly, there is more posterior evidence for rank reduction if we allow for a Markov trend instead of a deterministic trend. For all values of τ, the model with two unit roots ($r = 0$) is clearly rejected against both the cointegration ($r = 1$) and the Markov trend stationary ($r = 2$) specifications. The posterior probabilities assign more than 86% posterior probability to the cointegration specification. For $\tau = .01$, we find the least evidence for cointegration, although the evidence is certainly not weak. As discussed earlier, under this prior specification we a priori favor the presence of two unit roots and no cointegration, because the prior height at $\Pi = 0$ in the second BF in (50) is relatively small. The final two columns of the second panel of Table 1 refer to the case where we impose an improper prior on Π and Σ. We report again a PIC-based log BF, where replace the prior heights in (50) by the penalty function $(2\pi)^{-\frac{1}{2}(2-r)^2}$. The log BFs imply an assignment of 98% posterior probability to the cointegration specification.

Overall, the BF analysis suggests that the multivariate Markov trend model with one cointegration relation [(23)] is suitable for modeling the logarithm of U.S. per capita consumption and income. The estimated model is given by

$$\hat{Y}_t = \hat{N}_t + \hat{R}_t + \hat{Z}_t,$$
$$\hat{N}_t = -\begin{pmatrix} 481.89 \\ (.63) \\ 469.85 \\ (.81) \end{pmatrix} + \begin{pmatrix} .83 \\ (.13) \\ 1.16 \\ (.17) \end{pmatrix}(t-1) - \begin{pmatrix} .60 \\ (.21) \\ 1.33 \\ (.20) \end{pmatrix}\sum_{i=2}^{t} s_i,$$
$$\hat{R}_t = \begin{pmatrix} .15 \\ (.22) \\ 0 \end{pmatrix} s_t, \tag{55}$$
$$\Delta\hat{Z}_t = \begin{pmatrix} .24 \\ (.08) \\ .55 \\ (.19) \end{pmatrix}(1 \quad -.81)Z_{t-1} + \hat{\varepsilon}_t, \quad \text{with}$$
$$\hat{\Sigma} = \begin{pmatrix} .40 & .26 \\ (.06) & (.07) \\ .26 & .66 \\ (.07) & (.11) \end{pmatrix},$$

where the point estimates are posterior means and posterior standard deviations appear in parentheses. Because the posterior distribution of β_2 may have Cauchy-type tails, we report the posterior mode. (This is also done for other posterior quantities involving β_2.) The posterior means of the transition probabilities equal

$$\hat{p} = .86\ (.05) \qquad \text{and} \qquad \hat{q} = .76\ (.10).$$

The posterior results are based on the prior specification (29), (31)–(34), $p(\Sigma) \propto |\Sigma|^{-1}$, and $p(\alpha, \beta_2|\Sigma) \propto |J(\alpha, \lambda, \beta_2)|_{\lambda=0}$ and are obtained by including a Metropolis–Hastings 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 about 70% of the iterations. Note that a noninformative prior does not lead to problems if one just wants to estimate the model parameters without testing the rank.

Figure 3 shows the posterior density of β_2. The posterior mode of the cointegration relation parameter is $-.81$. The 95% highest posterior density (HPD) region for β_2 is $(-1.05, -.65)$, and hence -1 is included just in this region. There is only weak evidence for the consumption–income relation (9). The adjustment parameters .24 and .55 are both positive, which indicates that there is no adjustment toward equilibrium for the consumption equation. Note that this does not imply that the series move away from the equilibrium, because the adjustment of income toward equilibrium is larger than the nonadjustment in consumption (see also Johansen 1995, pp. 39–42).

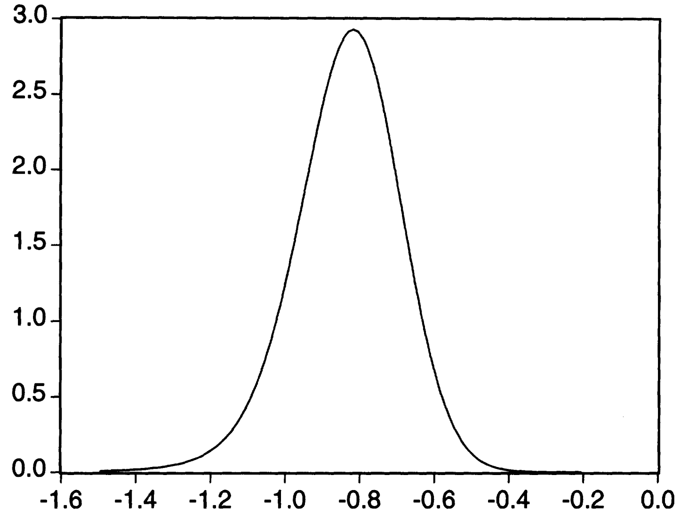

Figure 3. Posterior Density of β_2.

The posterior mean of the δ_1 parameter equals .15. The 95% HPD region for this parameter is $(-.31, .59)$, and hence it is very likely that δ_1 equals 0. The posterior means of the quarterly growth rates of the income series are 1.16% during an expansion regime and $-.17\%$ (1.16–1.33) during a contraction regime. For the consumption series, we get .83% and .23% (.83–.60). Hence during recessions, the growth rate in consumption is larger than the negative growth rate in income. To correct for this difference in the growth rates, the growth rate in income has to be larger than the growth rate in consumption during expansions.

Reduced-rank Markov trend cointegration ($\beta'\Gamma_1 = 0$) is not likely, because the posterior mode of $\beta'\Gamma_1$ equals .44 and its 95% HPD region is (.20, .89). The 95% HPD region of $\beta'\Gamma_0$ is $(-.39, .37)$ with a posterior mode of $-.08$. Hence the existence of a consumption–income relation (9) requiring that both $\beta'\Gamma_1$ and $\beta'\Gamma_0$ equal 0 is not likely. On the other hand, the results suggest that during recession periods, the growth rate in consumption is larger than in income, which is compensated for in the expansion periods, where income grows faster than consumption.

The expected slope of the Markov trend equals $\Gamma_0 + \Gamma_1(1-p)/(2-p-q)$ (see Hamilton 1989). The posterior mean of the expected slope of the Markov trend is .65% for the income series and .60% for the consumption series. These values differ by only .02 from the average quarterly growth rates reported in Section 3. The 95% HPD region of the expected slope of the Markov trend in the cointegration relation is $(-.10, .46)$, and the posterior mode equals .08. During recessions the posterior mode of the growth of the cointegration relation $\beta'(\Gamma_0 + \Gamma_1)$ is .34 (.21, .62), whereas during expansions it equals $-.08$ $(-.39, .37)$ as reported before.

Finally, we analyze how the estimated Markov trend relates to the NBER turning points. The posterior mean of the probability of staying in the expansion regime is .86, which is larger than the posterior mean of the probability of staying in a recession, .76. The posterior probability that p is larger than q

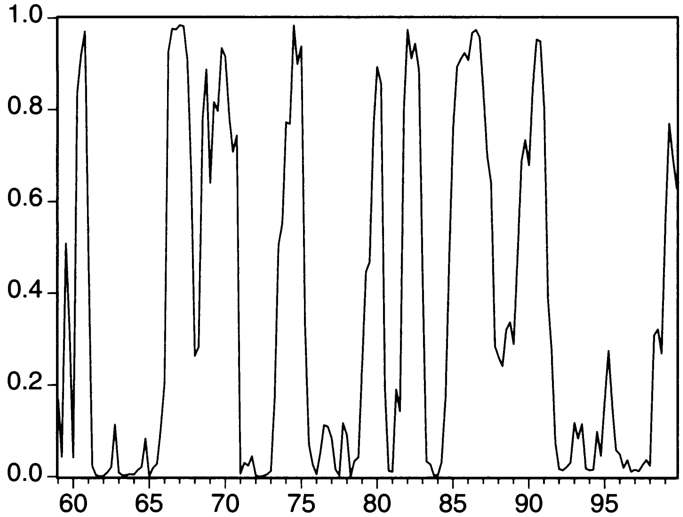

Figure 4. Posterior Expectations of the State Variables $E[s_t | Y^T]$.

is .88, which indicates the existence of an asymmetric cycle. Figure 4 shows the posterior expectations of the states variables, $E[s_t|Y^T]$. Values of these expectation close to 1 correspond to recessionary periods. Figure 5 shows the difference between the logarithm of U.S. income and consumption. The shaded areas correspond to the recessionary periods, where the growth rate in consumption is larger than the growth rate in income.

Table 2 indicates 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 two consecutive quarters for which $E[s_t|Y^T] > .5$. A peak is defined by the last expansion observation before a recession; a trough, 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 that do not correspond with official reported recessions. Remember that the consumption income analysis in this article is based on per capita disposable income. Looking at the government purchases on goods and services used to create the disposable income series, we see that government

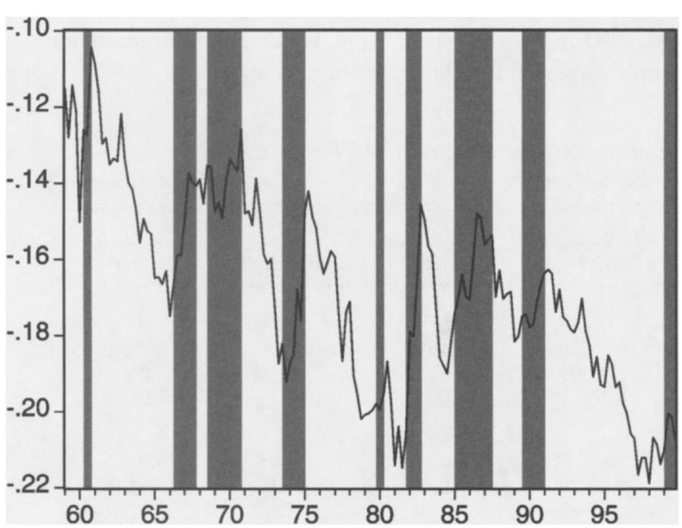

Figure 5. Difference Between Log U.S. Per Capita Consumption and Income. The shaded areas correspond to recessionary periods.

Table 2. Peaks and Troughs Based on the Posterior Expectations of the Unobserved State Variables

U.S.		NBER	
Peak	*Trough*	*Peak*	*Trough*
1960.1	1960.4	1960.2	1961.1
1966.1	1967.4		
1968.2	1970.4	1969.4	1970.4
1973.2	1975.1	1973.4	1975.1
1979.3	1980.2	1980.1	1980.3
1981.3	1982.4	1981.3	1982.4
1984.4	1987.3		
1989.2	1991.1	1990.3	1991.1

NOTE: A recession is defined by two consecutive quarters for which $E[s_t|Y^T] > .5$. A peak corresponds with the last expansion observation before a recession, a trough, with the last observation in a recession.

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 incorrectly reported as recessions. 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 U.S. 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 log 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. U.S. CONSUMPTION, INCOME, AND INVESTMENT

In the previous section we showed the existence of a cointegration relation between log per capita income and consumption only if a Markov trend is allowed for. We may investigate whether the inclusion of a third variable with a more pronounced cyclical pattern can help improve the model. Therefore, in this section we consider a multivariate Markov trend model for per capita real disposable income, private consumption, and private investment in the U.S., 1959.1–1999.4. The consumption and income series are the same as in the previous section. The investment series was also obtained from the Federal Reserve Bank of St. Louis. Figure 6 shows a plot of the log of the three series. The investment series clearly demonstrates a more pronounced cyclical pattern than the other two series.

To describe the three series, in Section 9.1 we consider a VAR model without a Markov trend. In Section 9.2 we introduce the multivariate Markov trend in the model, where we allow the growth rates in the three series to be different across the series and across the stages of the business cycle.

9.1 A Vector Autoregressive Model Without Markov Trend

In this section we analyze the presence of a cointegration relations in a VAR model without Markov trend [$\Gamma_1 = 0$ and

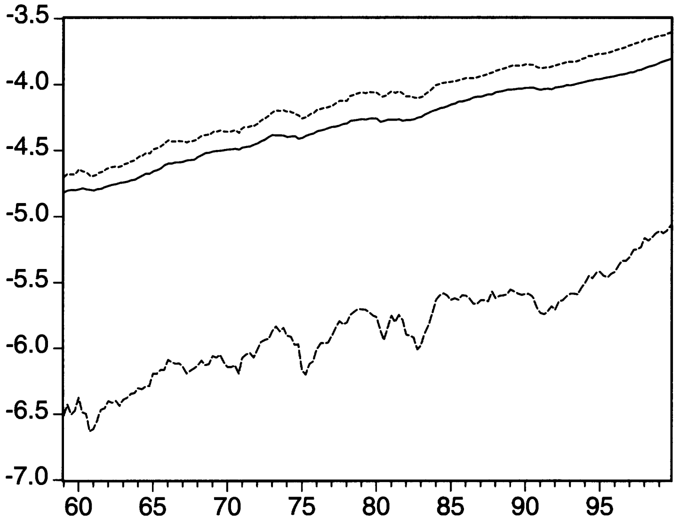

Figure 6. Logarithm of U.S. Per Capita Consumption (——), Income (· · · ·), and Investment (- - - -), 1959.1–1999.4.

$\delta_1 = 0$ in (21)] for $Y_t = 100 \times (\ln c_t, \ln y_t, \ln i_t)'$, where i_t denotes per capita investment series. The priors for model parameters are the same as in the previous example. The priors for N_1 and Σ are given by (29) and (30), with $S = I$ and $\nu = 4$. The g-type prior for Π given Σ is given in (44) with $P = 0$ and $A = \tau/T \sum_{t=1}^{T} \bar{Y}_t' \bar{Y}_t$ for different values of τ, where $\bar{Y}_t$ denotes the demeaned and detrended value of Y_t. For Γ_0 and $\bar{\Phi}_i$ we take flat priors $p(\Gamma_0) \propto 1$ and $p(\bar{\Phi}_i) \propto 1$.

The lag order determination is done in the same way as in Section 8. The resulting order is 2, which is also obtained if the BIC is used to determine k. The Bayesian cointegration analysis is a multivariate extension of the analysis in Section 7. The prior for α and β_2 for the cointegration specifications (ranks 1 and 2) are similar to (45). We assign equal prior probabilities to the possible cointegration ranks (46), that is, $\Pr[\text{rank} = r] = \frac{1}{4}$ for $r = 0, 1, 2, 3$.

The first panel of Table 3 displays the log BFs together with the posterior probabilities for different values of τ. For all values of τ, the BFs lead to 100% probability of a VAR model with 1 cointegration relation. [The standard Johansen (1995) trace tests do not indicate the presence of a cointegration relation between the three series if the deterministic trend is restricted within the cointegration space.] This is also true if one chooses to consider the PIC-based BFs.

The posterior results suggests that we must consider a VAR(2) model with 1 cointegration relation. The estimated model is given by

$$\hat{Y}_t = \hat{N}_t + \hat{Z}_t,$$

$$\hat{N}_t = -\begin{pmatrix} 481.78 \\ (.69) \\ 470.22 \\ (1.08) \\ 651.48 \\ (4.28) \end{pmatrix} + \begin{pmatrix} .70 \\ (.13) \\ .74 \\ (.17) \\ .84 \\ (.52) \end{pmatrix} (t-1),$$

$$\Delta\hat{Z}_t = \begin{pmatrix} .03 \\ (.04) \\ .01 \\ (.07) \\ -.22 \\ (.34) \end{pmatrix} (1 \quad -.95 \quad .24) Z_{t-1} \tag{56}$$

$$+ \begin{pmatrix} .20 & -.07 & .04 \\ (.14) & (.16) & (.03) \\ .68 & -.17 & .04 \\ (.21) & (.25) & (.05) \\ 3.80 & -.98 & .18 \\ (.86) & (.98) & (.20) \end{pmatrix} \Delta Z_{t-1} + \hat{\varepsilon}_t, \quad \text{with}$$

$$\hat{\Sigma} = \begin{pmatrix} .49 & .46 & .61 \\ (.06) & (.07) & (.25) \\ .46 & 1.19 & 3.91 \\ (.07) & (.14) & (.50) \\ .61 & 3.91 & 18.37 \\ (.25) & (.50) & (2.15) \end{pmatrix},$$

where again the point estimates are posterior means (except for the cointegration relation parameters) and posterior standard deviations are given in parentheses. The posterior results are based on a diffuse prior specification. The posterior modes of the cointegration relation parameters are $-.95$ for the consumption series and .24 for the investment series. The corresponding 95% HPD regions are $(-.47, 1.07)$ and $(-2.19, .04)$. Note that the HPD regions are quite large, which is due to the relatively small values of the adjustment parameters.

Table 3. Log BFs and Posterior Probabilities for the Cointegration Rank in a Linear VAR Model (k = 2) and the Multivariate Markov Trend Model (k = 1)

	$\tau = 1$		$\tau = .1$		$\tau = .01$		PIC	
r	$\ln \text{BF}(r\|3)$	$\Pr[r\|Y^T]$	$\ln \text{BF}(r\|3)$	$\Pr[r\|Y^T]$	$\ln \text{BF}(r\|3)$	$\Pr[r\|Y^T]$	$\ln \text{BF}(r\|3)$	$\Pr[r\|Y^T]$
Linear VAR model								
0	< -5	0	< -5	0	< -5	0	< -5	0
1	14.75	1.00	10.15	1.00	5.58	1.00	15.23	1.00
2	2.30	0	1.15	0	.01	0	3.06	0
3	0	0	0	0	0	0	0	0
Multivariate Markov trend model								
0	< -5	0	< -5	0	< -5	1.00	< -5	0
1	15.54	1.00	9.58	1.00	5.28	0.99	14.93	1.00
2	2.47	0	1.39	0	.27	.01	3.33	0
3	0	0	0	0	0	0	0	0

NOTE: A log BF, $\ln \text{BF}(r|3) > 0$, denotes that a cointegration model with r cointegration relations is more likely than a model with three cointegration relations. Posterior results are based on 400,000 iterations with the Gibbs sampler, neglecting the first 100,000 draws.

The posterior means of the slope parameters of the consumption and income series are somewhat larger than for the bivariate model discussed in Section 8.1. The posterior mean of the slope parameter of the investment series corresponds reasonably well with the average quarterly growth rate of the series equal to .89%.

9.2 A Multivariate Markov Trend Model

To allow for the possibility of different growth rates across the series and across the stages of the business cycle, we consider the Markov trend model (21). We take similar prior distributions as for the bivariate model in Section 8.2. Hence the prior distributions for the model parameters are given by (29)–(44) with $S = I$, $\nu = 4$, $P = 0$, and $A = \tau/T \sum_{t=1}^{T} \bar{Y}_t'\bar{Y}_t$, where $\bar{Y}_t$ denotes the demeaned and detrended value of Y_t.

The lag order selection procedure for the VAR part of the model results in $k = 1$. The priors for α and β_2 for the cointegration specifications are similar to (45). We again assign equal probabilities to the possible cointegration ranks, that is, $\Pr[\text{rank} = r] = \frac{1}{4}$ for $r = 0, 1, 2, 3$. The second panel of Table 3 reports the log BFs and posterior probabilities for the rank of Π for different values of τ. The values of the log BFs are similar to the values in the first panel of the table. Hence adding a Markov trend to the model does not change the posterior probabilities concerning the number of cointegration relations.

The selected model by the BF is a VAR(1) model with one cointegration relation. The estimated model is given by

$$\hat{Y}_t = \hat{N}_t + \hat{R}_t + \hat{Z}_t,$$

$$\hat{N}_t = -\begin{pmatrix} 481.95 \\ (.61) \\ 469.78 \\ (.79) \\ 648.54 \\ (3.03) \end{pmatrix} + \begin{pmatrix} 0.86 \\ (.11) \\ 1.15 \\ (.15) \\ 2.80 \\ (.49) \end{pmatrix}(t-1) - \begin{pmatrix} .66 \\ (.16) \\ 1.32 \\ (.16) \\ 5.12 \\ (.62) \end{pmatrix} \sum_{i=2}^{t} s_i,$$

$$\hat{R}_t = \begin{pmatrix} .34 \\ (.14) \\ 0 \\ 0 \end{pmatrix} s_t, \tag{57}$$

$$\Delta\hat{Z}_t = \begin{pmatrix} .22 \\ (.07) \\ .50 \\ (.14) \\ 2.02 \\ (.64) \end{pmatrix} (1 \quad -.71 \quad -.06) Z_{t-1} + \hat{\varepsilon}_t, \quad \text{with}$$

$$\hat{\Sigma} = \begin{pmatrix} .40 & .27 & -.15 \\ (.05) & (.06) & (.22) \\ .27 & .72 & 2.21 \\ (.06) & (.11) & (.41) \\ -.15 & 2.21 & 12.72 \\ (.11) & (.41) & (1.93) \end{pmatrix},$$

where again the point estimates are posterior means and posterior standard deviations are given in parentheses. The posterior results are based on a diffuse prior specification. The posterior means of the transition probabilities equal

$$\hat{p} = .86\ (.05) \qquad \text{and} \qquad \hat{q} = .76\ (.10),$$

which are equal to those of the bivariate Markov trend model in Section 8.2.

The posterior modes of the cointegration relation parameters are $-.71$ for the consumption series and $-.06$ for the investment series. The corresponding HPD regions are $(-1.00, -.35)$ and $(-.18, .06)$, which are clearly smaller than those for the linear VAR specification. The adjustment parameters are more than two posterior standard deviations away from 0, and hence the cointegration relation seems more relevant than in the model without the Markov trend. The HPD region of the cointegration relation parameter for investment contains 0, suggesting that any contribution of investment to the cointegration relation is of minor importance.

The posterior means of the Markov trend parameters of the consumption and income series are almost the same as those for the bivariate model in Section 8.1. For the investment series, the posterior mean of the quarterly growth rate is 2.80% during expansions and -2.32% during recessions (2.80–5.12). Reduced-rank Markov trend cointegration ($\beta'\Gamma_1 = 0$) is again not very likely, because the posterior mode of $\beta'\Gamma_1$ equals .45, and its 95% HPD region is (.19, .99).

In sum, we have seen that BFs suggest one out of three possible cointegration relations in a VAR model with deterministic trend for per capita consumption, income, and investment. This implies that there are still two unit roots remaining in the system, as was also the case in our bivariate specification in Section 8. Although BFs suggest the presence of one cointegration relation, the relevance of the error correction term is small. If we turn to a multivariate Markov trend model, then the error correction term becomes more relevant, and the contribution of investment to the cointegration relation is negligible. The inclusion of a Markov trend now does not lead to a decrease in the number of unit roots in the system as in the bivariate case. Although investment seems to partly replace the role of the Markov trend in the linear VAR, the posterior results of the Markov trend model role suggest that the Markov trend remains important.

10. CONCLUSION

In this article we have proposed using a multivariate Markov trend model to analyze the possible existence of a long-run relation between U.S. per capita consumption and income. The model specification was based on suggestions by simple economic theory and a simple stylized facts analysis on both series. The model contains a multivariate Markov trend specification that allows for different growth rates in the series and different growth rates during recessions and expansions. The deviations from the multivariate Markov trend are modeled by a VAR model. We have chosen a Bayesian approach to analyze U.S. series with the multivariate Markov trend model. BFs 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 a stationary linear relation exists 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 growth rate in consumption is larger than the negative growth rate in income during recessions. To compensate for this difference, the growth rate in income is larger than the growth rate in consumption during expansion periods. Replacing the Markov trend with a deterministic linear trend, the posterior results do not indicate the presence of a stationary linear relation between both series.

To analyze the robustness of our approach, we included per capita investment in the model, because this series has a more pronounced cyclical pattern. Hence we considered a multivariate Markov trend model for log per capita consumption, income, and investment series. The posterior results suggest the presence of only one cointegration relation between the three series. This result is found for both the Markov trend and the linear deterministic trend specification. Hence, adding a possible nonstationary variable to the Markov trend model does not increase the number of cointegration relations in the system. Although BFs suggest cointegration in the linear VAR model with deterministic trend, the posterior standard deviations of the adjustment parameters show that the cointegration relation is of minor importance. In the multivariate Markov trend model, the error correction term is more relevant, and investment does not make a significant contribution to the cointegration relation.

We end the article with some suggestion for further research. The multivariate Markov trend model that we have proposed 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 the adjustment parameters or the cointegrating vector to 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 cointegrated only in expansions and not in recessions. Testing for the presence of cointegration in the different regimes may be difficult, however, because 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, Lindgren, Holst, and Thuvesholmen 1994; Warne 1996). Finally, we may also consider alternative multivariate nonlinear models, like threshold models, to analyze the consumption and income series (see, e.g., Granger and Teräsvirta 1993; Balke and Fomby 1997).

APPENDIX A: JACOBIAN TRANSFORMATION

Here we derive the Jacobian of the transformation from Π to $(\alpha, \lambda, \beta_2)$ for a two-dimensional VAR model. (For larger dimensions, see Kleibergen and Paap 2002.) 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 $\theta = (1\ \ -\theta_2)'$. For notional convenience, the derivation of the Jacobian of the complete transformation from Π to $(\alpha_1, \alpha_2, \lambda, \beta_2)$ is 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. Because $\theta_\perp \in \alpha_\perp$, we can write

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

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

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

The Jacobian from θ_2 to α_2 is simply

$$G = \left|\frac{\partial\theta_2}{\partial\alpha_2}\right| = -\frac{1}{\alpha_1}. \qquad \text{(A.3)}$$

Hence the Jacobian for the total transformation is

$$J(\alpha, \lambda, \beta_2) = |J_1\ J_2\ J_3\ J_4|\ |G|. \qquad \text{(A.4)}$$

APPENDIX B: FULL CONDITIONAL POSTERIOR DISTRIBUTIONS

B.1 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 $s^{-t} = s^T\backslash\{s_t\}$. Because s_t follows a first-order Markov process, it is easily seen that

$$p(s_t|s^{-t}) \propto p(s_t|s_{t-1})p(s_{t+1}|s_t), \qquad \text{(B.1)}$$

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

$$\begin{aligned}
&p(s_t|s^{-t}, \Theta_2, Y^T) \\
&\quad = \frac{p(s_t|s^{-t}, \Theta_2, Y^t)f(Y_{t+1}, \ldots, Y_T|Y^t, s^{-t}, s_t, \Theta_2)}{f(Y_{t+1}, \ldots, Y_T|Y^t, s^{-t}, \Theta_2)} \\
&\quad \propto p(s_t|s^{-t}, \Theta_2, Y^t)f(Y_{t+1}, \ldots, Y_T|Y^t, s^{-t}, s_t, \Theta_2). \qquad \text{(B.2)}
\end{aligned}$$

Using the rules of conditional probability, the first term of (B.2) can be simplified as

$$\begin{aligned} & p(s_t|s^{-t}, \Theta_2, Y^t) \\ & \quad \propto p(s_t|s^{-t}, \Theta_2, Y^{t-1}) f(Y_t, s_{t+1}, \ldots, s_T|Y^{t-1}, s^t, \Theta_2) \\ & \quad \propto p(s_t|s_{t-1}, \Theta_2) f(Y_t|Y^{t-1}, s^t, \Theta_2) \\ & \qquad \times p(s_{t+1}|s^t, \Theta_2, Y^t) p(s_{t+2}, \ldots, s_T|s^{t+1}, \Theta_2, Y^t) \\ & \quad \propto p(s_t|s_{t-1}, \Theta_2) f(Y_t|Y^{t-1}, s^t, \Theta_2) p(s_{t+1}|s_t, \Theta_2), \end{aligned} \quad \text{(B.3)}$$

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 (B.2) is proportional to

$$f(Y_{t+1}, \ldots, Y_T|Y^t, s^t, \Theta_2) \propto \prod_{i=t+1}^{T} f(Y_i|Y^{i-1}, s^i, \Theta_2). \quad \text{(B.4)}$$

Next, using (B.3) and (B.4), the full conditional distribution of s_t for $t = k+1, \ldots, T$ is given by

$$\begin{aligned} & p(s_t|s^{-t}, \Theta_2, Y^T) \\ & \quad \propto p(s_t|s_{t-1}, \Theta_2) p(s_{t+1}|s_t, \Theta_2) \prod_{i=t}^{T} f(Y_i|Y^{i-1}, s^i, \Theta_2), \end{aligned} \quad \text{(B.5)}$$

where $f(Y_t|Y^{t-1}, s^t, \Theta_2)$ is defined in (24) 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,

$$\begin{aligned} p(s_t|s^{-t}, \Theta_2, Y^T) & \propto p(s_t|s_{t-1}, \Theta_2) p(s_{t+1}|s_t, \Theta_2) \\ & \quad \times \prod_{i=k+1}^{T} f(Y_i|Y^{i-1}, s^i, \Theta_2), \end{aligned} \quad \text{(B.6)}$$

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) showed, sampling of the state variables is easier if $\Pi = 0$. 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 tth element of s^T is replaced by its most recent draw.

B.2 Full Conditional Posterior of p and q

From the conditional likelihood function (25), it follows that the full conditional posterior densities of the transition parameters are given by

$$\begin{aligned} p(p|s^T, \Theta_2\backslash\{p\}, Y^T) & \propto p^{\mathcal{N}_{0,0}}(1-p)^{\mathcal{N}_{0,1}}, \\ p(q|s^T, \Theta_2\backslash\{q\}, Y^T) & \propto q^{\mathcal{N}_{1,1}}(1-q)^{\mathcal{N}_{1,0}}, \end{aligned} \quad \text{(B.7)}$$

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.

B.3 Full Conditional Posterior of Σ

It is easy to see from the conditional likelihood (25) that the full conditional posterior of Σ is proportional to

$$\begin{aligned} & p(\Sigma|s^T, \Theta_2\backslash\Sigma, Y^T) \propto |\Sigma|^{-\frac{1}{2}(T-k+2+\lambda)} \\ & \quad \times \exp\left(-\frac{1}{2}\text{tr}\left(\Sigma^{-1}(S + (Y_1 - N_1)(Y_1 - N_1)' + \sum_{t=k+1}^{T} \epsilon_t\epsilon_t'\right)\right), \end{aligned} \quad \text{(B.8)}$$

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

B.4 Full Conditional Posterior of N_1, Γ_0, and Γ_1

To derive the full conditional posterior distribution of N_1, Γ_0, and Γ_1, we write (21) as

$$\begin{aligned} \Sigma^{-\frac{1}{2}}\Phi(L)Y_t & = \Sigma^{-\frac{1}{2}}\Phi(L)\left(\Gamma_0(t-1) + \Gamma_1\sum_{i=2}^{t} s_i + N_1\right) + \Sigma^{-\frac{1}{2}}\varepsilon_t \\ & = -\Sigma^{-\frac{1}{2}}\sum_{j=1}^{k}\Phi_j(\Gamma_0\ \Gamma_1\ N_1)\begin{pmatrix} L^j(t-1) \\ L^j\sum_{i=2}^{t} s_i \\ 1 \end{pmatrix} \\ & \quad + \Sigma^{-\frac{1}{2}}\varepsilon_t, \end{aligned} \quad \text{(B.9)}$$

where $\Phi_0 = -I$. Without the Φ_j matrices, we have a multivariate regression model in the parameters N_1, Γ_0, and Γ_1, and the full conditional distribution would be matrix normal. To reverse the order of $\Phi(L)$ and the parameters $(\Gamma_0\ \Gamma_0\ N_1)$, we apply the vec operator to both sides of (B.9). Using the vec notation and the fact that $\text{vec}(ABC) = (C' \otimes A)\,\text{vec}(B)$, we can write (B.9) as a linear regression model, and hence the full conditional distributions of $\text{vec}(N_1)$, $\text{vec}(\Gamma_0)$, and $\text{vec}(\Gamma_1)$ are normal.

B.5 Full Conditional Posterior of δ_1

We write (21) as

$$\Sigma^{-\frac{1}{2}}\Phi(L)(Y_t - N_t) = \Sigma^{-\frac{1}{2}}\phi(L)\delta R_t + \Sigma^{-\frac{1}{2}}\varepsilon_t, \quad \text{(B.10)}$$

with $\Phi_0 = -I$. 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.

B.6 Full Conditional Posterior of Π and $\bar{\Phi}$

To sample from the full conditional posterior of the autoregressive parameters, we use the fact that conditional on Γ_0, Γ_1, N_1, and the states $\{s_t\}_{t=1}^{T}$, (21) can be seen as a multivariate regression model in the parameters Π and $\bar{\Phi}$. From Zellner (1971, chap. 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 (39).

B.7 Sampling of α and β_2

To derive the full conditional posterior distributions for α and β_2 we rewrite (23) such that conditional on $\bar{\Phi}$, N_1, Γ_0, Γ_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

$$\Delta Z_t - \sum_{i=1}^{k-1} \bar{\Phi}_i \Delta Z_{t-i} = \alpha\beta' Z_{t-1} + \varepsilon_t,$$
$$\Delta Z_t^* = \alpha\beta' Z_{t-1}^* + \varepsilon_t, \qquad \text{(B.11)}$$

where $\Delta Z_t^* = \Delta Z_t - \sum_{i=1}^{k-1} \bar{\Phi}_i \Delta Z_{t-i}$ and $Z_{t-1}^* = Y_{t-1} - N_{t-1} - R_{t-1}$. It is easy to see that the full conditional posterior distributions of α and β_2 are nonstandard. Therefore, Kleibergen and Paap (2002) proposed a Metropolis–Hastings algorithm to sample α and β_2 in this simple VAR model. Chib and Greenberg (1994, 1995) showed that it is possible to build such a Metropolis–Hastings algorithm into the Gibbs sampling procedure. The Metropolis–Hastings algorithm step works as follows. First, draw in iteration i of the Gibbs sampler Π^i from its full conditional posterior distribution. Perform a singular value decomposition on Π and solve for α^i, λ^i and β_2^i using (39). Now accept this draw of α^i and β_2^i with probability $\min(\frac{w(\alpha^i,\lambda^i,\beta_2^i)}{w(\alpha^{i-1},\lambda^{i-1},\beta_2^{i-1})}, 1)$, where i denotes the current draw, $i-1$ denotes the previous draw, and

$$w(\alpha, \lambda, \beta_2) = \frac{|J(\alpha, \lambda, \beta_2)|_{\lambda=0}|}{|J(\alpha, \lambda, \beta_2)|} g(\lambda|\Theta_1, Y^T)|_{\lambda=0}, \qquad \text{(B.12)}$$

where

$$g(\lambda|\Theta_1, Y^T) = (2\pi)^{-\frac{1}{2}} |\alpha'_\perp \Sigma^{-1} \alpha_\perp|^{\frac{1}{2}} |\beta'_\perp (A + Z_{-1}^{*\prime} Z_{-1}^*)\beta_\perp|^{\frac{1}{2}} \times \exp\left(-\frac{1}{2}\operatorname{tr}\left((\beta'_\perp (A + Z_{-1}^{*\prime} Z_{-1}^*)\beta_\perp)(\lambda - \tilde{\lambda}) \times (\alpha'_\perp \Sigma^{-1}\alpha_\perp)(\lambda - \tilde{\lambda})\right)\right), \qquad \text{(B.13)}$$

with

$$\tilde{\lambda} = \left(\beta'_\perp (A + Z_{-1}^{*\prime} Z_{-1}^*)\beta_\perp\right)^{-1} \beta'_\perp \left(A(P - \beta\alpha') + Z_{-1}^{*\prime}(\Delta Z^* - Z_{-1}^*\beta\alpha')\right)\Sigma^{-1}\alpha_\perp(\alpha'_\perp\Sigma^{-1}\alpha_\perp)^{-1}, \qquad \text{(B.14)}$$

$Z_{-1}^* = (Z_k^* \ \dots \ Z_{T-1}^*)'$, and $\Delta Z^* = (\Delta Z_{k+1}^* \ \dots \ \Delta Z_T^*)'$. If the draw of α^i and β_2^i is rejected, then one must take the previous draw, that is, $\alpha^i = \alpha^{i-1}$ and $\beta_2^i = \beta_2^{i-1}$.

ACKNOWLEDGMENTS

The first author thanks the Netherlands Organization for Scientific Research for its financial support. The authors thank three anonymous referees for very helpful comments that substantially improved the article. They also thank Peter Boswijk, Philip Hans Franses, Niels Haldrup, Frank Kleibergen, Katsuhiro Sugita, and Peter Summers.

[Received December 2000. Revised February 2003.]

REFERENCES

Albert, J. H., and Chib, S. (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.

Balke, N. S., and Fomby, T. B. (1997), "Threshold Cointegration," *International Economic Review*, 38, 627–646.

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

——— (1995), "Understanding the Metropolis–Hastings Algorithm," *The American Statistician*, 49, 327–335.

Dickey, D. A., and Fuller, W. A. (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 Potter, S. M. (1996), "Markov Trend Cointegration," unpublished working paper, University of California.

Flavin, M. A. (1981), "The Adjustment of Consumption to Changing Expectations About Future Income," *Journal of Political Economy*, 89, 974–1009.

Geisser, S. (1965), "A Bayes Approach for Combining Correlated Estimates," *Journal of the American Statistical Association*, 60, 602–607.

Gelfand, A. E., and Smith, A. F. M. (1990), "Sampling-Based Approaches to Calculating Marginal Densities," *Journal of the American Statistical Association*, 85, 398–409.

Geman, S., and Geman, D. (1984), "Stochastic Relaxations, Gibbs Distributions, and the Bayesian Restoration of Images," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 6, 721–741.

Golub, G. H., and van Loan, C. F. (1989), *Matrix Computations*, Baltimore, MD: John Hopkins University Press.

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 Teräsvirta, T. (1993), *Modelling Nonlinear Economic Relations*, Oxford, U.K.: Oxford University Press.

Gregory, A. G., and Hansen, B. E. (1996), "Residual-Based Tests for Cointegration in Models With Regime Shifts'," *Journal of Econometrics*, 70, 99–126.

Hall, S. G., Psaradakis, Z., and Sola, M. (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.

Hamilton, J. D., and Perez-Quiros, G. (1996), "What Do the Leading Indicators Lead," *Journal of Business*, 69, 27–49.

Hansen, H., and Johansen, S. (1999), "Some Tests for Parameter Constancy in Cointegrated VAR Models," *Econometrics Journal*, 2, 306–333.

Hoek, H. (1997), "Variable Trends: A Bayesian Perspective," unpublished doctoral thesis, Tinbergen Institute, Amsterdam.

Holst, U., Lindgren, G., Holst, J., and Thuvesholmen, M. (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, U.K.: Oxford University Press.

Kim, C.-J., and Nelson, C. R. (1999a), "Has the U.S. Economy Become More Stable?: A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle," *Review of Economics and Statistics*, 81, 608–616.

——— (1999b), *State-Space Models With Regime Switching: Classical and Gibbs-Sampling Approaches With Applications*, Cambridge, MA: MIT Press.

Kim, M.-J., and Yoo, J.-S. (1995), "New Index of Coincident Indicators: A Multivariate Markov-Switching Factor Model Approach," *Journal of Monetary Economics*, 36, 607–630.

Kleibergen, F., and Paap, R. (2002), "Priors, Posterior and Bayes Factors for a Bayesian Analysis of Cointegration," *Journal of Econometrics*, 111, 223–249.

Kleibergen, F., and van Dijk, H. K. (1994), "On the Shape of the Likelihood/Posterior in Cointegration Models," *Econometric Theory*, 10, 514–551.

——— (1998), "Bayesian Simultaneous Equation Analysis Using Reduced-Rank Regression," *Econometric Theory*, 14, 701–743.

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

Lam, P.-S. (1990), "The Hamilton Model With a General Autoregressive Component," *Journal of Monetary Economics*, 26, 409–432.

Luginbuhl, R., and de Vos, A. (1999), "Bayesian Analysis of an Unobserved-Component Time Series Model of GDP With Markov-Switching and Time-Varying Growths," *Journal of Business & Economic Statistics*, 17, 456–465.

Martin, G. M. (2000), "U.S. Deficit Sustainability: A New Approach Based on Multiple Endogeneous Breaks," *Journal of Applied Econometrics*, 15, 83–105.

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

Nelson, C. R., Piger, J., and Zivot, E. (2001), "Markov Regime Switching and Unit-Root Tests," *Journal of Business & Economic Statistics*, 19, 404–415.

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 Speight, A. E. H. (1998), "Threshold Nonlinearities in Output: Some International Evidence," *Applied Economics*, 30, 323–334.

Phillips, K. L. (1991), "A Two-Country Model for Stochastic Output With Changes in Regime," *Journal of International Economics*, 31, 121–142.

Phillips, P. C. B., and Ploberger, W. (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 U.S. GNP," *Journal of Applied Econometrics*, 10, 109–125.

Psaradakis, Z. (2001), "Markov-Level Shifts and the Unit-Root Hypothesis," *Econometrics Journal*, 4, 225–241.

——— (2002), "On the Asymptotic Behaviour of Unit-Root Tests in the Presence of a Markov Trend," *Statistics & Probability Letters*, 57, 101–109.

Sargent, T. J. (1978), "Rational Expectations, Econometric Exogeneity, and Consumption," *Journal of Political Economy*, 86, 673–700.

Schotman, P. C., and van Dijk, H. K. (1991a), "A Bayesian Analysis of the Unit Root in Real Exchange Rates," *Journal of Econometrics*, 49, 195–238.

——— (1991b), "On Bayesian Routes to Unit Roots," *Journal of Applied Econometrics*, 6, 387–401.

Schwarz, G. (1978), "Estimating the Dimension of a Model," *The Annals of Statistics*, 6, 461–464.

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

Smith, A. F. M., and Roberts, G. O. (1993), "Bayesian Computation via the Gibbs Sampler and Related Markov Chain Monte Carlo Methods," *Journal of the Royal Statistical Society*, Ser. B, 55, 3–23.

Tanner, M. A., and Wong, W. H. (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*, New York: Wiley.

——— (1986), "On Assesing Prior Distributions and Bayesian Regression Analysis With g-Prior Distributions," in *Bayesian Inference and Decesion Techniques: Essays in Honor of Bruno de Finetti*, eds. P. Goel and A. Zellner, Amsterdam: North-Holland, pp. 233–243.

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