THE UNIVERSITY OF MELBOURNE

FACULTY OF BUSINESS & ECONOMICS

Melbourne Institute Working Paper Series

Working Paper No. 27/14

A Bayesian Approach to Modelling Bivariate Time-Varying Cointegration and Cointegrating Rank

Chew Lian Chua and Sarantis Tsiaplias

MELBOURNE INSTITUTE®
of Applied Economic and Social Research

A Bayesian Approach to Modelling Bivariate Time-Varying Cointegration and Cointegrating Rank*

Chew Lian Chua and Sarantis Tsiaplias
Melbourne Institute of Applied Economic and Social Research
The University of Melbourne

Melbourne Institute Working Paper No. 27/14

ISSN 1328-4991 (Print)

ISSN 1447-5863 (Online)

ISBN 978-0-7340-4367-2

December 2014

* For correspondence, contact Dr Tsiaplias at <sarantis.tsiaplias@unimelb.edu.au>.

Melbourne Institute of Applied Economic and Social Research
The University of Melbourne
Victoria 3010 Australia
***Telephone* (03) 8344 2100**
***Fax* (03) 8344 2111**
***Email* melb-inst@unimelb.edu.au**
***WWW Address* http://www.melbourneinstitute.com**

Abstract

A bivariate model that allows for both a time-varying cointegrating matrix and time-varying cointegrating rank is presented. The model addresses the issue that, in real data, the validity of a constant cointegrating relationship may be questionable. The model nests the sub-models implied by alternative cointegrating matrix ranks and allows for transitions between stationarity and non-stationarity, and cointegrating and non-cointegrating relationships in accordance with the observed behaviour of the data. A Bayesian test of cointegration is also developed. The model is used to assess the validity of the Fisher effect and is also applied to equity market data.

JEL classification: C11, C32, C51, C52

Keywords: Error correction models, singular value decomposition, cointegration tests

1 Introduction

Standard models of bivariate cointegration allow for the possibility that a constant linear transformation can be applied to the pair of non-stationary variables to produce a stationary variable (Engle and Granger, 1987). In accordance with such models, numerous tests have been developed to determine the validity of a constant cointegrating relationship between the two variables (Johansen, 1991). In practice, however, structural change or regime shifts are frequently present in economic or financial data thereby rendering inappropriate the assumption of a constant cointegrating relationship.

To address this issue, cointegration tests in the presence of structural breaks and regime shifts have been proposed (Gregory and Hansen, 1996a,b; Johansen et al, 2000; Andrade et al, 2005). A drawback with such tests is that their applicability is limited to a small number of breaks or shifts. To avoid this limitation, researchers have developed approaches aimed at identifying cointegration in the presence of interlocutors such as multiple structural breaks or regimes; in lieu of the breaks or shifts in the data, the variables of interest share a common stochastic trend. This research includes allowing for multiple breaks (Martin, 2000; Hansen, 2003), regime-switching behaviour (Paap and van Dijk, 2003), or time-varying cointegrating vectors (Saikonnen and Choi, 2004; Bierens and Martins, 2010; Koop et al, 2011).

A common objective underlying the above-mentioned research is identifying a cointegrating relationship in a noisy, structurally dynamic environment. A generalisation of this approach involves the possibility of alternating cointegrated and non-cointegrated states. This generalisation has two noteable benefits. First, the presence of such alternating behaviour allows for a range of cointegrating scenarios. Two variables that are typically cointegrated may behave in a manner inconsistent with cointegration for considerable periods of time. Alternatively, changes in preferences may produce cyclical behaviour such that variables share a common stochastic trend in some periods and not in others.

Second, the econometrician is always unaware of the true data generating process. Even where variables are cointegrated, the econometrician never observes the cointegrating equation. In practice, cointegration tests are applied that often provide conflicting results. Relatively minor amendments to the postulated form of the cointegrating relationship can yield significantly different conclusions regarding cointegration. In many cases, the addition of new data or the choice of the starting period for a dataset impact significantly on the assessment of cointegration. By allowing for stationarity, non-stationarity, and cointegration, the a priori imposition of a cointegrating rank is avoided. Instead, if the econometrician has evidence regarding any cointegrating relationship, this is embedded into the model through the choice of an appropriate prior.

We develop a bivariate model that can alternate between states of cointegration and non-cointegration. This is achieved by relaxing the restriction that the rank of the cointegrating matrix in an error correction model (ECM) is time-invariant. Since the rank of the cointegrating matrix is allowed to vary over time, the model allows for all three possible states identified by the rank-condition in the cointegrating matrix: two variables may be I(1) with a single cointegrating relationship, I(1) with no cointegrating relationship, or I(0). This generalisation is important since, even where two variables are typically cointegrated, there may be periods where the variables do not appear to share a common stochastic trend or where the variables behave as stationary processes. Indeed, this type of switching behaviour is observed in the two applications undertaken in this paper covering the relationship between US interest and inflation rates, and the relationship between the FTSE 100 and S&P 100 stockmarket indices.

Sugita (2006) and Jochmann and Koop (2011) also allow for time-varying changes to the cointegrating space. Sugita (2006) focuses on the identification of structural breaks in the vector autoregressive-ECM (VECM), but allows for the cointegrating rank to change with each structural break. The cointegrating rank is determined using an approach proposed by Strachan (2003) that relies on the singular value decomposition (SVD) of the cointegrating matrix (Kleibergen and van Dijk, 1998). Jochmann and Koop (2011) highlight the computational difficulties associated with the estimation of the Sugita (2006) model. With anything more than a relatively small number of structural breaks, model estimation becomes intractable; as a result, only a small number of rank changes may occur in the data.[1]

Jochmann and Koop (2011) overcome this limitation by allowing for a Markovian structure in the latent (discrete) regime structure. By switching between regimes, their approach allows for cyclical changes in cointegrating rank, in addition to permanent break-dependent changes in rank. The modelling approach adopted in our paper, however, differs significantly from that adopted in Jochmann and Koop (2011). By treating the individual elements of the SVD of Π as dynamic latent processes, we are able observe the time-dependent path of the elements of the decomposition, in addition to being able to specify the functional form of the latent processes. This provides significant flexibility in modelling the time path of Π. For example, stickiness in the rate of adjustment parameters or in the cointegrating relationship is straightforward to accommodate by specifying AR forms for the appropriate latent processes.

In turn, and in contrast to Jochmann and Koop (2011), we adopt a methodology that activates or deactivates the column space of the matrices in the SVD of Π in accordance with the prevailing Markovian regime. Pursuant to this methodology, the latent processes that are used to construct Π continue to be estimated even where the cointegrating rank is zero. As such, the latent processes shadow the cointegrating structure when the cointegrating rank is zero, and are re-admitted into the cointegrating structure when a non-zero rank is chosen. There are marked benefits to this approach, including allowing the econometrician to observe the strength of the time-varying signals in the elements of Π.

[1] Sugita (2006) considers up to 4 breaks in US term structure data.

The approach also allows us to parsimoniously nest the time-varying parameter (TVP) VAR for I(0) data, the VAR for I(1) data and the TVP cointegrated VAR for I(1) data into our model. This property is used to propose a straightforward test to estimate whether the data are I(0), I(1) with no cointegrating relationship, or I(1) with a single cointegrating relationship. The test provides estimates of the probabilities associated with each of the competing relationships that can be obtained directly from the model parameters and does not require the derivation of marginal likelihoods.

This paper is structured as follows. Section 2 presents the standard bivariate ECM and formulates extensions in the form of a time-varying SVD of the cointegrating matrix and a time-varying cointegrating rank. Section 3 presents the Metropolis-in-Gibbs sampler used to estimate the model. Section 4 evaluates the model using simulated data, while Section 5 applies the model to interest rate and inflation data, and to equity market indices. Section 6 concludes the paper.

2 The bivariate ECM

The standard bivariate ECM is represented as follows[2]

$$
\begin{aligned}
\triangle y_t &= c^* + t\eta^* + y_{t-1}\Pi + \triangle y_{t-1}B + \varepsilon_t \qquad t = 1, ..., T \\
& \varepsilon_t \overset{iid}{\sim} N(0, \Sigma)
\end{aligned}
\tag{1}
$$

where $y_t = [\ y_{1t} \quad y_{2t}\ ]$ is a vector of I(0) or I(1) variables, c^*, η^* are (1×2) vectors, Π is a (2×2) cointegrating matrix, B is a (2×2) coefficient matrix for $\triangle y_{t-1}$ and ε_t is a (1×2) independent bivariate Gaussian process with positive definite covariance matrix Σ.

Kleibergen and van Dijk (1998) and Kleibergen and Paap (2002) have shown that the

[2] For ease of exposition, we work with first-order ECMs in this paper. The extension to higher-order ECMs is straightforward.

local non-identification issue in Π, and the variant problem in the cointegrating vector β (pursuant to the standard decomposition $\Pi = \beta\alpha$) due to the ordering of equations, can be resolved by the singular value decomposition of Π into two orthonormal matrices and a diagonal matrix of singular values

$$\Pi = U\Lambda V' \tag{2}$$

where U and V are (2×2) orthonormal matrices such that $U'U = V'V = I_2$, and Λ is a (2×2) diagonal matrix that contains (in descending order) non-negative singular values.

Since the number of non-zero singular values reflects the rank of Π, the existence of a cointegrating relationship between y_{1t} and y_{2t} is consistent with the lower diagonal element of Λ being constrained to zero. In addition, the smallest singular value provides information about the degree of rank deficiency in Π.

We relate α and β to the SVD in (2) as follows:

1. $\alpha = \Lambda V'$ and $\beta = U$ for rank$(\Pi) = 2$, and
2. $\alpha = w_1 V_1'$ and $\beta = U_1$ for rank$(\Pi) = 1$. w_1 is the first diagonal element of Λ, V_1 is the first column of V and U_1 is the first column of U.

In the case of rank$(\Pi) = 1$, α can be interpreted as a speed of adjustment vector and β as the cointegrating vector. In this case we set $c^* = c - \eta_0\alpha$ and $\eta^* = -\eta_1\alpha$ such that the cointegrating vector $y_{t-1}\beta$ is associated with the attractor $(\eta_0 + t\eta_1)$ (see, further, Franses, 2001).

2.1 Time-variation in the cointegrating matrix

In line with (2), the SVD applied to Π will obtain the orthnormal matrices U, V and the diagonal matrix Λ (see Strachan, 2003; Strachan and Inder, 2004; Villani, 2006). This paper does not, however, adopt an explicit SVD of Π to obtain $U = \{u_{ij}\}$, $V = \{v_{ij}\}$, and $\Lambda = \{w_j\}$. Instead, we treat $S \subset \{U, V, \Lambda\}$ as a set of estimable latent processes. Time-variation in Π can, therefore, be implemented by assuming a dynamic form for S.

Equation (2) therefore becomes

$$\Pi_t = U_t \Lambda_t V_t' \tag{3}$$

where $\Lambda_t = \begin{bmatrix} w_{1t} & 0 \\ 0 & w_{2t} \end{bmatrix}$, with the ordered pair (w_{1t}, w_{2t}) following $w_{1t} \geq w_{2t} \geq 0$. In the special case of rank$(\Pi) = 1$, (3) implies that the speed of adjustment and the cointegrating vector are time-varying pursuant to $\alpha_t = w_{1t} V_{1t}'$ and $\beta_t = U_{1t}$.

For any (2×2) orthonormal matrix, the elements within the matrix are interdependent functions and may follow two possible trigonometric representations (rotation or reflection). In terms of estimation, the representation adopted is arbitrary and, in this paper, U_t and V_t are treated as rotation matrices

$$U_t = \begin{bmatrix} \cos\theta_{ut} & \sin\theta_{ut} \\ -\sin\theta_{ut} & \cos\theta_{ut} \end{bmatrix}, V_t = \begin{bmatrix} \cos\theta_{vt} & \sin\theta_{vt} \\ -\sin\theta_{vt} & \cos\theta_{vt} \end{bmatrix}. \tag{4}$$

The elements in U_t and V_t are, respectively, functions of real-valued frequency parameters θ_u and θ_{vt}, and are expressed in radians. Rather than estimating θ_{ut}, θ_{vt}, w_{1t} and w_{2t}, however, it is computationally convenient to estimate the product of V_t and Λ_t

$$\begin{aligned} D_t &= V_t \Lambda_t = \begin{bmatrix} d_{1t} & d_{3t} \\ -d_{2t} & d_{4t} \end{bmatrix} \\ &= \begin{bmatrix} w_{1t}\cos\theta_{vt} & w_{2t}\sin\theta_{vt} \\ -w_{1t}\sin\theta_{vt} & w_{2t}\cos\theta_{vt} \end{bmatrix} \end{aligned} \tag{5}$$

where D_t follows

$$D_t'D_t = \begin{bmatrix} w_{1t}^2 & 0 \\ 0 & w_{2t}^2 \end{bmatrix}. \tag{6}$$

Without loss of generality, we assume that $d_{3t} = \kappa_t d_{2t}$ and $d_{4t} = \kappa_t d_{1t}$. To enable identification of $\kappa_t \forall t$, we impose $w_{1t} \geq \varepsilon^*$ where ε^* is an arbitrarily small positive number. Using (5) it can be shown that $\kappa_t = w_{2t}/w_{1t}$ such that κ_t necessarily lies in the unit space. Estimating d_{1t}, d_{2t} and κ_t is therefore sufficient to recover the parameters in D_t as

$$\begin{aligned} w_{1t} &= \sqrt{d_{1t}^2 + d_{2t}^2} && (7) \\ w_{2t} &= \sqrt{d_{3t}^2 + d_{4t}^2} && (8) \\ d_{3t} &= \kappa_t d_{2t} && (9) \\ d_{4t} &= \kappa_t d_{1t} && (10) \\ \theta_{vt} &= \cos^{-1}\left(\frac{d_{1t}}{w_{1t}}\right) && (11) \end{aligned}$$

A range of functional forms may be adopted for d_{1t}, d_{2t}, θ_{ut} and κ_t. We adopt the following broad specifications which, with appropriate restrictions, accommodate a wide range of processes including time-invariant parameters, the random walk, and the stationary autoregressive model.

$$d_{1t} = \mu_1 + \rho_1 d_{1t-1} + v_{1t}, \qquad v_{1t} \sim N(0, \sigma_1) \tag{12}$$

$$d_{2t} = \mu_2 + \rho_2 d_{2t-1} + v_{2t}, \qquad v_{2t} \sim N(0, \sigma_2) \tag{13}$$

$$\theta_{ut} = \mu_3 + \rho_3 \theta_{u,t-1} + v_{3t}, \qquad v_{3t} \sim N(0, \sigma_3)^3 \tag{14}$$

[3]Rather than following the standard approach of normalising $\sqrt{\sigma_3}$ to 1, we normalise $\sqrt{\sigma_3}$ to 1 degree ($\frac{\pi}{180}$ radians) in line with θ_{ut}'s status as a frequency measure.

$$\kappa_t = \kappa_{t-1} + v_{4t}, \qquad v_{4t} \sim N(0, \sigma_4). \tag{15}$$

We find it convenient to set $\mu_1 = \mu_2 = 0$ and $\rho_1 = \rho_2 = 1$ such that d_{1t}, d_{2t} follow random walks, thereby incorporating a relatively agnostic path dependency in the parameters belonging to Π_t. This appears to be a reasonable attribute given its implication that the relationship between Δy_t and y_{t-1} will typically not change from period to period in a manner unrelated to its path.

Our general specification provides flexibility in setting a range of priors pertaining to the rate of adjustment and cointegration parameters. Adopting tight zero-mean normal priors for ρ_1, ρ_2 and inverse gamma priors for σ_1, σ_2 with location close to 0, for example, imposes the prior assumption that the rate of adjustment is largely constant. Alternatively, specifying an AR(2) specification for d_{1t}, d_{2t} with non-zero priors for the AR parameters, provides a simple means for imposing a prior assumption of (some) cyclicality in the rate of adjustment.

Typically, however, economic theory introduces restrictions on the cointegrating relationship. Restrictions of this nature are straightforward to accommodate. Setting a zero-mean normal prior on μ_3, for example, asserts a prior belief in a cointegrating vector located around $\begin{bmatrix} 1 & 0 \end{bmatrix}$. In the analysis of the Fisher effect undertaken in Section 5, a normal prior for μ_3 located at $\frac{\pi(1-\rho_3)}{4}$ is consistent with the prior location $\begin{bmatrix} 1 & -1 \end{bmatrix}$ for the cointegrating relationship as implied by theory. In general, adopting the prior $\mu_3 \sim N\left(\frac{\pi(1-\rho_3)}{4}, \sigma^2_{\mu_3}\right)$ incorporates the belief that the simple difference $y_{1t} - y_{2t}$ is I(0).

2.2 Time-variation in the cointegrating rank

In this sub-section, we augment the model with a latent discrete Markov-switching component to enable time-variation in the rank of Π_t. This is achieved by incorporating a (2×2) idempotent matrix $I(S_t)$ into (3). The idempotent matrix $I(S_t)$ is used to engage or disengage the columns in the orthogonal matrices (U_t, V_t) and the elements in the singular value matrix Λ_t. Consequently, any rank reduction in Π_t is made feasible by appropriately controlling the admission of the elements in (U_t, V_t, Λ_t) that constitute the matrix Π_t

Equation (3) is re-formulated as

$$\begin{aligned} \Pi_t &= U_t \Lambda_t V_t' \\ &= U_t I(S_t) I(S_t) \Lambda_t V_t' \end{aligned} \tag{16}$$

where

$$I(S_t) = \begin{bmatrix} (1-s_{1t})(s_{2t}+s_{3t}) & 0 \\ 0 & (1-s_{1t})s_{3t} \end{bmatrix}. \tag{17}$$

Note that s_{jt}, $j = 1,2,3$, is an indicator variable such that $s_{jt} = 1$ if $S_t = j$ and $s_{jt} = 0$ if $S_t \neq j$. When s_{jt} is equal to 1, (16) implies that the rank of Π_t is $(j-1)$.

The discrete state S_t is assumed to follow a Markovian transition probability for a move from state i to state j

$$p_{ij} = \Pr(S_t = j | S_{t-1} = i), \qquad \sum_{j=1}^{3} p_{ij} = 1. \tag{18}$$

The transition (18) can be represented in matrix form as

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix}. \tag{19}$$

The choice of a Markovian transition probability allows for persistence in the rank condition. The combination of a Markovian transition matrix (19) and the path dependent specifications for $\theta_{ut}, d_{1t}, d_{2t}$ and κ_t implies that both the cointegrating rank and the elements of Π_t (even where the cointegrating rank is constant) depend on their historical values.

Pursuant to the association between S_t and the rank of Π_t, the steady state probabilities associated with P (being the solution to $x = P'x$) represent the posterior probabilities associated with each rank. These probabilities are readily obtained as the normalised eigenvector associated with the largest eigenvalue (having the value unity) of the matrix P'. The following limiting conditions hold:

1. $p_{11} = p_{21} = p_{31} = 1$ eliminates any cointegration between the I(1) vector y_t and the resulting model is a VAR in $\triangle y_t$;

2. $p_{12} = p_{22} = p_{32} = 1$ implies that the rank of Π_t $\forall t$ must be equal to $r = 1$ and y_t is cointegrated by reference to the cointegrating matrix $\beta_t = U_{1t}$. The resulting model is a VECM;

3. $p_{13} = p_{23} = p_{33} = 1$ implies that there is no reduced rank for Π_t $\forall t$ and $y_t \sim I(0)$. The resulting model is a VAR for stationary data.

Accordingly, the model can be viewed as a meta-structure which nests a VAR, a TVP-VECM, and a TVP model for stationary data (TVP-S). Denote the VAR, TVP-VECM and TVP-S models as models M_1, M_2, and M_3 respectively. The steady state probabilities associated with P provide the posterior probabilities associated with each model, $\widehat{P}(M_i)$ for $i = 1, 2$, and 3.

It is clear that prior beliefs about the existence of cointegration in y_t can be incorporated into the model as priors on the parameters p_{ij}. Define the 3×3 transition matrix P_j containing a vector of ones in the jth column and zeros everywhere else. The standard approach of ex ante choosing among the models defined by the different rank conditions is straightforward to obtain by imposing a degenerative prior such that $P = P_j$, thereby imposing rank$(\Pi) = j - 1$. In the converse situation, a diffuse prior on P will result in probabilities for the possible ranks of Π that depend only on the observed data. Alternatively, a relatively informative Dirichlet prior on P can be used to accommodate the econometrician's prior belief without forcing a choice among competing models.

3 Bayesian model estimation

Estimation of the model is undertaken in a Bayesian context using a Metropolis-in-Gibbs sampler that relies on the Kalman filter/smoother and the auxilliary particle filter. The sampler involves 12 steps and is summarised below.

The following notation applies: $\widetilde{d}_T = \left(\widetilde{d}_{1T}, \widetilde{d}_{2T}\right)$; $\widetilde{d}_{iT} = (d_{i1}, d_{i2}, ..., \ d_{iT})$ for $i = 1, 2$; $\widetilde{\theta}_{uT} = (\theta_{u1}, \theta_{u2}, ..., \theta_{uT})$; $\widetilde{\kappa}_T = (\kappa_1, \kappa_2, ..., \kappa_T)$; $\widetilde{S}_T = (S_1, S_2, ..., S_T)$; $p = (p_{11}, p_{12}, ..., p_{32}, p_{33})$; $\mu = (\mu_1, \mu_2, \mu_3, \mu_4)$; $\rho = (\rho_1, \rho_2, \rho_3, \rho_4)$; $\sigma = (\sigma_1, \sigma_2, \sigma_3, \sigma_4)$; $\eta = \begin{bmatrix} \eta_0 & \eta_1 \end{bmatrix}'$; $C = \begin{bmatrix} c_1 & c_2 & c_3 \end{bmatrix}'$ and $g = vec\left(\begin{bmatrix} B_1' & B_2' & B_3' \end{bmatrix}'\right)$.

Step 1: Draw the Markovian regimes $\widetilde{S}_T$ given $\widetilde{y}_T, \widetilde{d}_T, \widetilde{\theta}_{uT}, \widetilde{\kappa}_T, g, \Sigma, \eta, C$ and σ.

Step 2: Draw the transition probabilites p given $\widetilde{S}_T$.

Step 3: Draw the latent variable $\widetilde{d}_{1T}$ given $\widetilde{y}_T, \widetilde{S}_T, \widetilde{d}_{2T}, \widetilde{\theta}_{uT}, \widetilde{\kappa}_T, g, \Sigma, \eta, C, \mu, \rho$ and σ.

Step 4: Draw the latent variable $\widetilde{d}_{2T}$ given $\widetilde{y}_T, \widetilde{S}_T, \widetilde{d}_{1T}, \widetilde{\theta}_{uT}, \widetilde{\kappa}_T, g, \Sigma, \eta, C, \mu, \rho$ and σ.

Step 5: Draw the latent variable $\widetilde{\theta}_{uT}$ given $\widetilde{y}_T, \widetilde{S}_T, \widetilde{d}_T, \widetilde{\kappa}_T, g, \Sigma, \eta, C, \mu, \rho$ and σ.

Step 6: Draw the latent variable $\widetilde{\kappa}_T$ subject to $\kappa_t \in [0,1] \forall t$ and given $\widetilde{y}_T, \widetilde{S}_T, \widetilde{d}_T, g, \Sigma, \eta, C, \mu, \rho$ and σ.

Step 7: Draw the latent variable variances σ given $\widetilde{\kappa}_T$, $\widetilde{d}_T, \mu$ and ρ.[4]

Step 8: Draw the the intercept and autoregressive parameters μ, ρ for each of the $\widetilde{d}_{1T}, \widetilde{d}_{2T}$ and $\widetilde{\theta}_{uT}$ given σ.

[4] Note that σ_3 is normalised to $\left(\frac{\pi}{180}\right)^2$.

Step 9: Draw the autoregressive parameters (for $\triangle y_{t-1}$) g given $\widetilde{y}_T, \widetilde{S}_T, \widetilde{d}_T, \widetilde{\theta}_{uT}, \widetilde{\kappa}_T, \eta, C$ and Σ.

Step 10: Draw the intercept C given $\widetilde{y}_T, \widetilde{S}_T, \widetilde{d}_T, \widetilde{\theta}_{uT}, \widetilde{\kappa}_T, g, \eta$ and Σ.

Step 11: Draw the intercept and trend η pertaining to the cointegrating relation given $\widetilde{y}_T, \widetilde{S}_T, \widetilde{d}_T, \widetilde{\theta}_{uT}, \widetilde{\kappa}_T, g, C$ and Σ.

Step 12: Draw the variance Σ given $\widetilde{y}_T, \widetilde{S}_T, \widetilde{d}_T, \widetilde{\theta}_{uT}, \widetilde{\kappa}_T, \eta, C$ and g.

We assume the following proper and independent prior densities for the model parameters

$$g \sim N(\underline{g}, \Sigma_g)$$

$$C \sim N(\underline{c}, \Sigma_c) \quad i = 1, 2, 3$$

$$\eta \sim N(\underline{\eta}, \Sigma_\eta)$$

$$\Sigma_i \sim IW(\underline{v}_i, \underline{D}_i), \quad i = 1, 2, 3$$

$$\rho_i \sim N(\underline{\rho}, \sigma_\rho), \qquad i = 1, 2, 3.4$$

$$\mu_i \sim N(\underline{\mu}, \sigma_\mu), \qquad i = 1, 2, 3, 4$$

$$\sigma_i \sim IG(\frac{\vartheta_i}{2}, \frac{\vartheta_i f_i}{2}), \qquad i = 1, 2 \text{ and } 4$$

$$p_{ii} \sim beta(\tau_{ii}, \overline{\tau}_{ii}), \qquad i = 1, 2, 3$$

$$\overline{p}_{ij} \sim beta(\tau_{ij}, \tau_{ik}), \qquad i, j, k = 1, 2, 3 \text{ and } i \neq j \neq k.$$

$$d_{10}, d_{20} \sim N\left(\underline{d}, \sigma_d\right), \ \theta_{u0} \sim U\left(0, \pi\right), \ \kappa_0 \sim beta(\underline{\kappa}_a, \underline{\kappa}_b).$$

The latent variables $\widetilde{d}_{1T}$, $\widetilde{d}_{2T}$ are simulated using the Kalman filter and smoother (Carter and Kohn, 1994), while draws of the truncated latent variables $\widetilde{\theta}_{uT}$, $\widetilde{\kappa}_T$ are obtained using the auxiliary particle filter (Pitt and Shephard, 1999).[5] We follow Carter and Kohn (1994) and Kim, Nelson and Startz (1998) in drawing the Markovian regimes $\widetilde{S}_T$ and the the associated transition parameters p. Conditional on the unobserved variables and the adoption of a multivariate normal prior, the regression coefficients g, C and η are multivariate normal (see, for example, Chib and Greenberg, 1996). Conditional on $\widetilde{\theta}_{uT}$ and the chosen prior, the intercept and AR(1) coefficients μ_3, ρ_3 are also normally distributed and straightforward to obtain.[6] Finally, the variance parameters σ, Σ are conditionally distributed as inverse gamma and inverse Wishart respectively. Details regarding each step of the sampler are provided in the Appendix.

4 Simulated evidence

We specify the following DGP to assess the sampler constructed in Section 3. Observations of $y_t = \begin{bmatrix} y_{1t} & y_{2t} \end{bmatrix}$, $t = 1, 2, ..., 1000$, are obtained by generating 1500 observations from (20) - (25) and discarding the initial 500 observations.

$$\triangle y_t = c_{S_t} - \eta_0 \alpha_t - t\eta_1 \alpha_t + y_{t-1}\beta_t \alpha_t + \triangle y_{t-1} B_{S_t} + \varepsilon_t \tag{20}$$

$$\varepsilon_t \overset{iid}{\sim} MVN(0, \Sigma_{S_t}) \tag{21}$$

$$d_{1t} = d_{1t-1} + v_{1t}, \qquad v_{1t} \overset{iid}{\sim} N\left(0, 0.001^2\right) \tag{22}$$

$$d_{2t} = d_{2t-1} + v_{2t}, \qquad v_{2t} \overset{iid}{\sim} N\left(0, 0.001^2\right) \tag{23}$$

$$\theta_{ut} = \theta_{u,t-1} + v_{3t}, \qquad v_{3t} \overset{iid}{\sim} N\left(0, 0.001^2\right) \tag{24}$$

$$\kappa_t = \kappa_{t-1} + v_{4t}, \qquad v_{4t} \overset{iid}{\sim} N\left(0, 0.01^2\right) \tag{25}$$

[5] We have also obtained $\widetilde{\kappa}_T$ using the method proposed in Chan, Koop and Potter (2013).

[6] In our applications we set $(\mu_1 = 0, \rho_1 = 1)$, $(\mu_2 = 0, \rho_2 = 1)$. However, the aforementioned parameters are obtained in the same manner as μ_3, ρ_3.

where $\Sigma_{S_t} = s_{1t} \begin{bmatrix} 0.15 & -0.1 \\ -0.1 & 0.15 \end{bmatrix} + s_{2t} \begin{bmatrix} 0.55 & 0.1 \\ 0.1 & 0.45 \end{bmatrix} + s_{3t} \begin{bmatrix} 0.2 & 0.01 \\ 0.01 & 0.3 \end{bmatrix}$, $B_{S_t} = s_{1t} \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} + s_{2t} \begin{bmatrix} 0.1 & 0.2 \\ 0.1 & 0.15 \end{bmatrix} + s_{3t} \begin{bmatrix} 0.5 & 0.2 \\ 0.2 & 0.2 \end{bmatrix}$, $c_i = \begin{bmatrix} 0.0 & 0.0 \end{bmatrix}$ for $i \in \{1, 2, 3\}$, $\eta_0 = \begin{bmatrix} 0.01 & 0.01 \end{bmatrix}$, $\eta_1 = \begin{bmatrix} 0.03 & 0.03 \end{bmatrix}$ and $I(S_t)$ is generated using the transition parameters $p_{11} = 0.98$, $p_{22} = p_{33} = 0.99$, $p_{12} = p_{13} = 0.01$ and $p_{21} = p_{23} = p_{31} = p_{32} = 0.005$.

The DGP assumes that the cointegrating rank is sticky. If the rank of $\Pi_t = U_t I(S_t) I(S_t) \Lambda_t V_t' = \beta_t \alpha_t$ is j, then Π_{t+k} is also likely to be of rank j except where k is large. Any transition is, however, allowed and the system is not restricted to step-wise rank transitions (e.g. Π_t can move from rank $j = 0$ to rank $j = 1$ or $j = 2$ at time $t + 1$). Accordingly, y_t switches between cointegrated and non-cointegrated states, and between stationarity and non-stationarity, although these switches are not common. The first column of U_t can be interpreted as the shadow cointegrating vector, whose parameters are functions of θ_{ut} which follows a slow- changing random walk. The cointegrating relationship holds when $S_t = 2$ and is associated with the attractor $(\eta_0 + t\eta_1)$ which, for simplicity, is not assumed to be regime dependent. The matrix $D_t = V_t \Lambda_t$ is determined by d_{1t}, d_{2t} and κ_t which are modelled as slow-changing random walk processes. In the event of rank$(\Pi_t) = 1$, the first column of D_t can be interpreted as a speed of adjustment vector. In this respect, the slow-changing d_{1t},d_{2t} variables imply that the speed of adjustment is unlikely to change dramatically from t to $t + 1$.

To estimate the model parameters, we produce 50,000 draws and apply a burn-in of the first 10,000 draws. We undertake this process for two distinct sets of priors: the first is fairly diffuse and the second is more informative. In each case we obtain similar results which suggests that the estimates are insensitive to the choice of prior. Consequently, we restrict our attention to estimates obtained using the first set of priors which is fairly diffuse:

$$\underline{g} = \begin{bmatrix} 0 & 0 & ... & 0 & 0 \end{bmatrix}'$$

$$\Sigma_g = diag\left(\begin{bmatrix} 100 & 100 & ... & 100 & 100 \end{bmatrix}\right)$$

$$\underline{\eta} = \begin{bmatrix} 0 & ... & 0 \end{bmatrix}'$$

$$\Sigma_\eta = diag\left(\begin{bmatrix} 100 & ... & 100 \end{bmatrix}\right)$$

$$\underline{D}_i = 0.036 \begin{bmatrix} var(y_{1t}) & 0 \\ 0 & var(y_{2t}) \end{bmatrix}$$

and $\underline{v}_i = 2$, $\underline{\rho} = 0.5$, $\sigma_\rho = 100$, $\underline{\mu} = 0$, $\sigma_\mu = 100$ $\vartheta_i = 0.0196$, $f_i = 1$, $\tau_{ii} = \overline{\tau}_{ii} = \tau_{ij} = \tau_{ik} = 2$ for $i, j = 1, 2, 3$ and $i \neq j \neq k$. We also adopt the following diffuse parameters for the initial states $\underline{d} = 0$, $\sigma_d = 100$, $\underline{\kappa}_a = \underline{\kappa}_b = 1$.

Figure 1 shows the actual and expected values for the regimes s_{1t}, s_{2t} and s_{3t}. Pursuant to the figure, the DGP switches considerably between periods of cointegration ($s_{2t} = 1$), non-stationarity without cointegration ($s_{1t} = 1$), and stationarity ($s_{3t} = 1$). The posterior regime probabilities are estimated as the sample mean of s_{jt}, $j \in \{1, 2, 3\}$, across the draws obtained using the sampler. It is readily evident that the sampler accurately captures both the persistence of each regime and the transitions between the regimes.[7] Deviation between the actual and estimated regime is limited to a period between $t = 110$ and $t = 180$, where fairly similar probabilities for regimes 2 and 3 are observed in the first half of the period, with regime 3 declining in favour of regime 2 as $t \to 180$. During this period, the posterior probabilities correctly result in a sharp drop associated with the first regime. However, a corresponding rise in the probability of regime 2 is observed rather than a rise in both regimes 2 and 3. Given the strong

[7] The model was also estimated subject to time-invariant B and Σ and was able to accurately capture rank transitions, with the model typically producing the correct rank within 5 periods.

tracking of actual regimes, the estimates of the transition parameters are close to their true values, with p_{11}, p_{22} and p_{33} all being close to unity.[8]

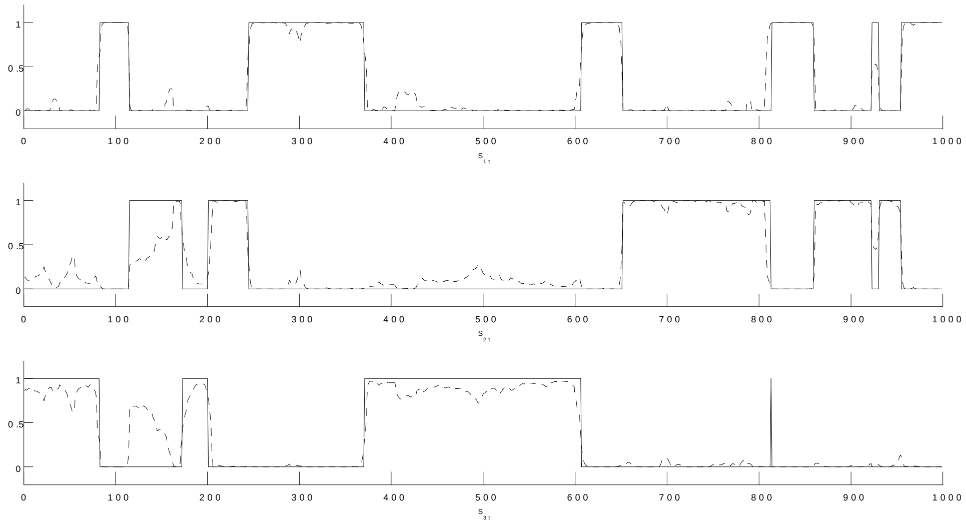


Figure 1. Actual and expected values for regimes s_{1t}, s_{2t} and s_{3t}. The dotted lines are posterior expected probabilities for the regimes s_{1t}, s_{2t} and s_{3t}, while the solid lines are their actual values.

Figure 2 shows the actual values of d_{1t}, d_{2t}, θ_{ut} and κ_t, together with 95 percent credible intervals obtained using draws from the sampler. In each case the 95% credible intervals encapsulate the actual values of the time-varying parameters as expected. The results suggest that the sampler readily discerns the location of the latent parameters. Testing of the model with alternative DGP structures suggests that uncertainty regarding the latent parameters increases with the rapidity of regime shifts. This is unsurprising since numerous, successive regime shifts imply significant uncertainty about the true

[8] The posterior means of p_{11}, p_{22} and p_{33} are 0.9717, 0.9698 and 0.9719 respectively. To preserve space, the estimates and associated statistics for the set of DGP parameters are not presented. The estimates are available on request from the authors.

model at any given time period.

We also obtain the steady state probabilities of the transition matrix to compare the prevalence of M_1, M_2 and M_3. The results of the test show that $\widehat{P}(M_1) = 0.34$, $\widehat{P}(M_2) = 0.35$, and $\widehat{P}(M_3) = 0.30$. These probabilities correctly indicate that no single model is valid for the entire time period, with each regime prevailing for a similar period of time.

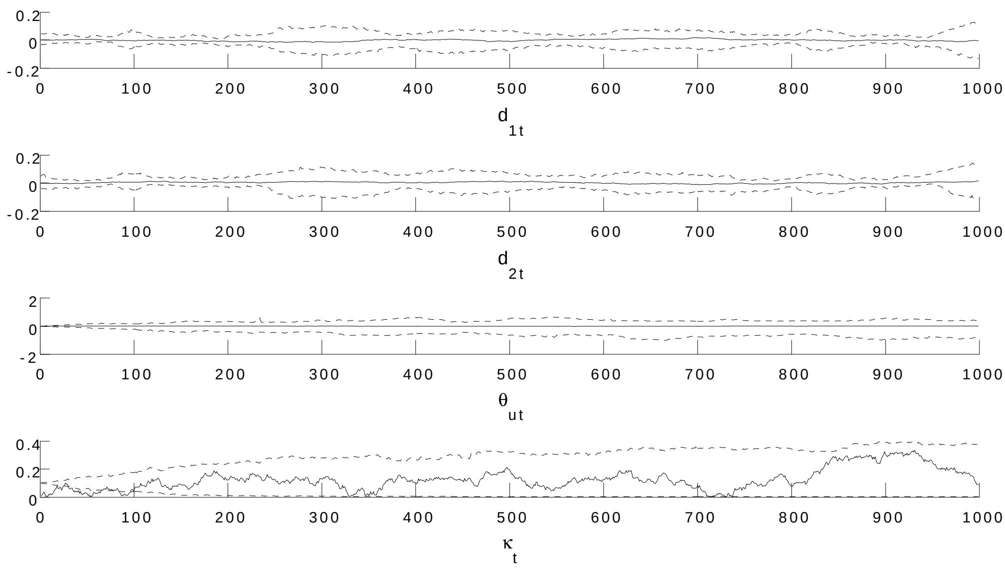


Figure 2. Actual values of d_{1t}, d_{2t}, θ_{ut} and κ_t and their 95 percent credible intervals. The solid lines are actual values and the dotted lines are the credible intervals.

5 Applications

5.1 Short term interest and inflation rates: The Fisher effect

The Fisher equation asserts that the real interest rate is determined by the difference between the nominal interest rate and the expected inflation rate given information at time t

$$r_t = i_t - \pi_t^e \tag{26}$$

$$= i_t - \pi_t + e_t \tag{27}$$

where e_t is a stationary zero-mean innovation term, and (27) follows from (26) subject to rational expectations. Assuming that the real interest rate is stationary, if i_t and π_t are I(1) variables, (27) implies that the nominal interest rate and the inflation rate are cointegrated with cointegrating vector $\begin{bmatrix} 1 & \beta = -1 \end{bmatrix}$. This relationship is also known as the Fisher effect.

The validity of the Fisher effect has been studied extensively, with conflicting results. Koustas and Serletis (1999) assess the relationship between interest and inflation rates for a number of countries and find that the data generally reject the Fisher effect. Koop et al. (2011), on the other hand, find evidence of a single cointegrating relationship between UK interest and inflation rates. However, they find that the cointegrating vector changes over time, with β typically less than -1. Neely and Rapach (2008) undertake a comprehensive review of the relevant literature.

To evaluate the Fisher effect, we apply our model to US short term interest rates (y_{1t}) and annualised inflation rates (y_{2t}).[9] Monthly data are used spanning the period January 1948 to June 2012 ($n = 774$). We produce 50,000 draws from an adjusted version of the sampler in Section 4, discarding the first 5000 draws. Since the data are not consistent

[9] We use the 3-Month Treasury Bill (secondary market rate). Data are from the Federal Reserve Bank of St Louis' FRED database (http://research.stlouisfed.org/fred2/).

with the presence of a trend in the cointegrating equation, we follow Jochmann and Koop (2011) in setting $\eta_1 = \begin{bmatrix} 0 & 0 \end{bmatrix}$. We also set $C = 0$ such that the intercept is given by $\eta_0 \alpha_t$ rather than $c_{S_t} - \eta_0 \alpha_t$. We find that the adoption of an unrestricted C produces erratic estimates of the normalised cointegrating vector and therefore limit our discussion to the restricted specification, which appears to better represent the properties of the data. The diffuse priors adopted in the simulated example are also largely applied here. We make a slight adjustment, however, to (14) and estimate $\theta_{ut} = \mu_3^* (1 - \rho_3) + \rho_3 \theta_{u,t-1} + v_{3t}$ such that μ_3^* is the expected value of θ_{ut}. Two priors are adopted for the intitial state θ_{u0} being $\theta_{u0} \sim U(0, \pi)$ and $\theta_{u0} \sim N\left(\frac{\pi}{4}, 0.09\right)$ with the latter prior imposing the belief that the initial state of θ_{ut} is consistent with the cointegrating relationship implied by the Fisher effect. The results exhibit little sensitivity to the choice of prior for θ_{u0}, therefore only the results pertaining to the informative initial prior are presented.[10]

Table 1 provides the posterior estimates of the model parameters. The model probabilities, pursuant to the steady state values of the transition matrix P, are provided in Table 2. Figures 3 to 5 present, respectively, the time-varying regimes S_t, the parameters d_{1t}, d_{2t}, θ_{ut} and κ_t used to construct Π_t, and the β_t parameter stemming from the (normalised) cointegrating vector. We obtain an estimate of β from the normalised cointegrating vector by settting $\beta = -\sin\theta_{ut} / \cos\theta_{ut}$.

The regime parameters indicate a high level of stickiness in the rank condition, especially when $r = 1$ or $r = 2$. When $r = 0$ there is a relatively greater likelihood of moving towards a cointegrating relationship with $p_{12} = 0.1122$. The first and second regimes prevail for all t with the exception of the beginning and towards the end of the sample, where $r = 2$ prevails. Clearly, the model indicates that US interest and inflation rates are cointegrated for the majority of the period under evaluation, with some periods where the two rates behave as non-cointegrated I(1) processes or I(0) processes. In the early 70s and early to mid 80s, and in 2008, the model suggests a rejection of the Fisher

[10] Several runs of the sampler were undertaken using significantly different initial conditions with little change to the posterior estimates.

effect's implication that interest and inflation rates share a cointegrating relationship. These results are similar to those observed by Neely and Rapach (2003), who find evidence of instability in the real interest rate in the 1970s and early 1980s (see, also, Bai and Perron, 2003).

	Mean	Std Dev		Mean	Std Dev		Mean	Std Dev
p_{11}	0.8609	0.0643	σ_3	$(\frac{\pi}{180})^2$	0	$B_{1,11}$	0.3905	0.1245
p_{12}	0.1122	0.0621	σ_4	0.0024	0.0010	$B_{1,12}$	0.0784	0.2045
p_{13}	0.0269	0.0182	η_{01}	0.0070	0.1721	$B_{1,21}$	0.0666	0.0574
p_{21}	0.0239	0.0151	η_{02}	-0.2990	2.3622	$B_{1,22}$	0.2984	0.1071
p_{22}	0.9628	0.0163	$\Sigma_{1,11}$	0.2770	0.0616	$B_{2,11}$	0.1931	0.0531
p_{23}	0.0132	0.0073	$\Sigma_{1,12}$	0.0285	0.0589	$B_{2,12}$	0.0688	0.0380
p_{31}	0.0141	0.0101	$\Sigma_{1,22}$	1.0112	0.2141	$B_{2,21}$	0.1032	0.0626
p_{32}	0.0353	0.0192	$\Sigma_{2,11}$	0.0648	0.0059	$B_{2,22}$	0.3453	0.0589
p_{33}	0.9506	0.0219	$\Sigma_{2,12}$	0.0015	0.0029	$B_{3,11}$	0.4430	0.0696
μ_3^*	0.0092	0.0373	$\Sigma_{2,22}$	0.0435	0.0087	$B_{3,12}$	0.0040	0.0091
ρ_3	0.9982	0.0014	$\Sigma_{3,11}$	0.4249	0.0658	$B_{3,21}$	0.2015	0.6820
σ_1	0.0101	0.0030	$\Sigma_{3,12}$	0.0088	0.0047	$B_{3,22}$	-0.0302	0.1314
σ_2	0.0102	0.0028	$\Sigma_{3,22}$	0.0047	0.011			

Table 1. Posterior parameter estimates: short term interest and inflation rates. Note that $\Sigma_{i,jk}$ refers to the j,kth element of the regime dependent variance Σ_i $(i = 1, 2, 3)$. Conversely, $B_{i,jk}$ refers to the j,kth element of B_i.

	$\widehat{P}(M_1)$	$\widehat{P}(M_2)$	$\widehat{P}(M_3)$
Probability	0.1322	0.6277	0.2401

Table 2. Model probabilities: short term interest and inflation rates

The third regime indicates stationarity and is dominant in the post WW2 period and during the GFC crisis. Both these periods are associated with distinctly flat US short-term interest rates. US short-term interest rates varied little from 1948 to 1950, whereas from 2009 to the end of the sample US short-term rates have been close to zero. In terms of the latter, the US Federal Reserve has also cultivated expectations that short-term rates will be close to zero for a potentially lengthy period of time (Bernanke, 2009).

One of the difficulties encountered in assessing the Fisher effect is the presence of unknown breaks and shifts in the data that pose problems in both detecting cointegration and estimating the value of β. These issues are largely obviated by our model given its allowance for transitions in the rank of Π_t that can depend on shifts in the data through the regime-dependent terms. Although there are clear regime shifts in the data, our approach provides support for the hypothesis of a cointegrating relationship between US short-term interest rates and inflation, with a steady state probability for M_2 of 62.8 per cent. We note that the probability in favour of a single cointegrating relationship does not imply that the Fisher effect holds.

The Fisher effect also asserts that the (normalised) cointegrating vector takes on the values $\begin{bmatrix} 1 & \beta = -1 \end{bmatrix}$. An oft-omitted implication of this assertion is that the the cointegrating vector is assumed to be time-invariant. The property of time-invariance does not appear to hold with θ_{ut}, pursuant to which the cointegrating vector $\begin{bmatrix} \cos\theta_{ut} & -\sin\theta_{ut} \end{bmatrix}'$ is estimated, typically declining over time. Overall, θ_{ut} has a persistence close to unity ($\rho_3 = 0.9982$) and negligible drift (given $\mu_3^* = 0.0092$) indicating that it essentially behaves as a Martingale process.

Our results suggest that β typically does not equal -1. The posterior mean of β varies from approximately -1 at the beginning of the sample to around -3 towards the end of the sample. The 90 percent credible interval for β encompasses -1 approximately 30.1 percent of the time providing only limited support for the cointegrating relationship implied by the Fisher effect. Interestingly, Koop et al. (2011) obtain a fairly similar result for the UK, with the 90 per cent credible interval for β including -1 around 34 percent of the time. Our results, coupled with those of Koop et al (2011), suggest that in recent years the value of β is clearly below -1.[11]

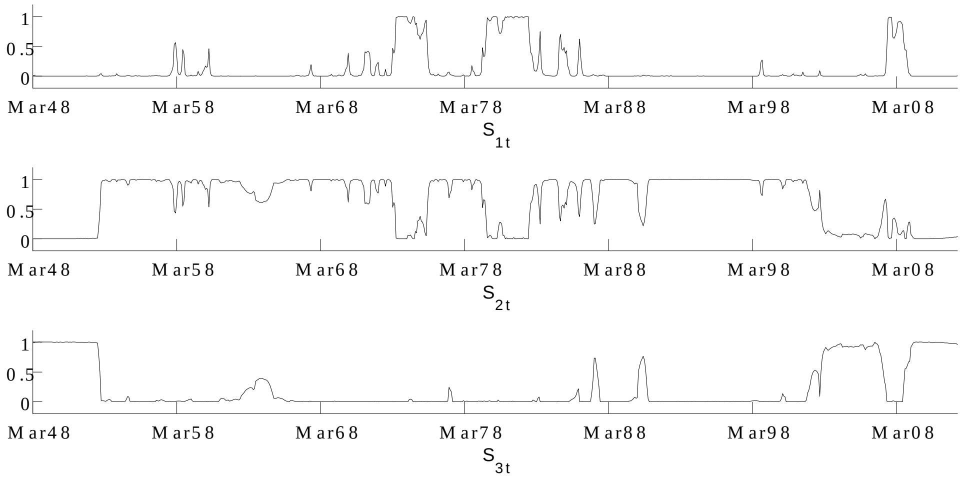
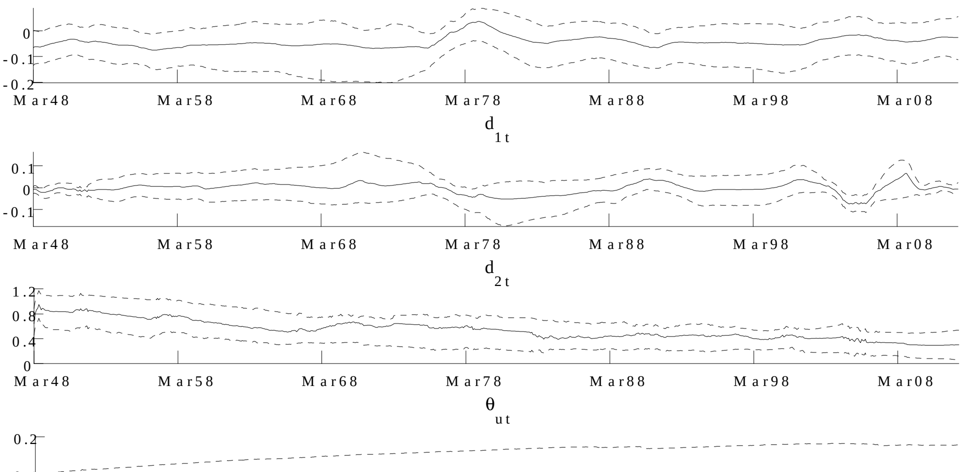


Figure 3. Posterior means of s_{1t}, s_{2t}, and s_{3t}.



[11] Our 90 per cent credible interval is entirely below -1 during the period 1980 - 2012. Similarly, Koop et al's 90 per cent credible interval for UK data is entirely below -1 for almost the entire period 1980 - 2009.

Figure 4. Estimated path of d_{1t}, d_{2t}, θ_{ut} and κ_t. The dotted and solid lines are respectively the 95 percent credible intervals and posterior means of d_{1t}, d_{2t}, θ_{ut} and κ_t.

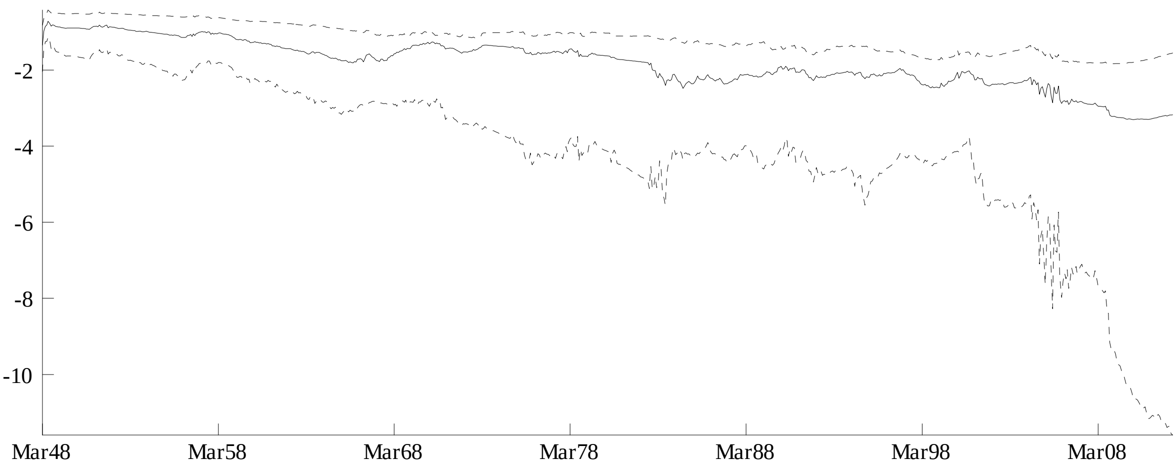


Figure 5. Estimated path of β_t. β_t is obtained from the normalised cointegrating vector, $\beta_t = -\sin\theta_{ut}/\cos\theta_{ut}$. The dotted and solid lines are, respectively, the 90 percent credible intervals and posterior mean of β_t.

5.2 Equity market data

We obtain FTSE 100 and S&P 100 index values at the weekly frequency over the period 6 October 1997 to 13 August 2012 ($n = 775$). Johansen's test of cointegration on the index values suggest that there is no cointegration, although the results are not conclusive. The trace statistic implies the presence of a single cointegrating relationship at the .10 level. The test, however, rejects $r = 1$ at the .05 level. In contrast, the eigen statistic for $r = 0$ is not rejected at the .10 level implying that the indices are I(1) but not cointegrated. Overall, there is some uncertainty as to whether a cointegrating relationship is present.

We produce 20,000 draws from the sampler in Section 4, discarding the first 4000 draws.[12] Since the data provide little evidence regarding an intercept or time trend in the cointegrating equation, we set $\eta_0 = \eta_1 = \begin{bmatrix} 0 & 0 \end{bmatrix}$. The model is estimated on the index data divided by 100. Table 3 provides the posterior estimates of the parameters, with model probabilities in Table 4. Figure 6 presents the time-varying regimes S_t, and Figure 7 presents the parameters d_{1t}, d_{2t}, θ_{ut}, κ_t used to construct Π_t.

The first regime S_{1t}, consistent with $r = 0$, typically prevails from 2001 to the end of the sample. The second regime prevails consistently in the first three years suggesting that the data exhibit a shift from a cointegrating relationship towards non-cointegrated I(1) processes at about 2001. During the prevalence of S_{2t}, the behaviour of θ_{ut} is relatively tightly bound suggesting that the cointegrating relationship is fairly stable albeit exhibiting very little persistence ($\rho_3 = 0.0327$ and is not significantly different to zero). Although 2001 coincides with the demise of the dot-com bubble, the timing of the structural shift (being the week ending September 24, 2001) indicates that the shift from a cointegrated I(1) relationship to a non-cointegrated I(1) relationship is associated with the financial market's reaction to the September 11 terrorist attack.

Following the attack, the behaviour of the indices is consistent with $r = 0$ save for two periods in support of $r = 1$. The first period occurs about June and July 2002 and is associated with the stockmarket crash of 2002. The second period occurs in 2008 and reflects a general reaction to the beginning of the GFC and the collapse or bailout of numerous financial institutions in the US. There is almost no evidence in favour of the prevalence of the third regime S_{3t}, which is consistent with $r = 2$, suggesting that the hypothesis of stationarity is never supported. This is not particularly surprising and is consistent with the notion of efficient markets producing individual weekly asset prices that always behave as martingale processes.

[12] The diffuse priors adopted in the simulated example are also applied here. Two runs of the sampler were undertaken using significantly different initial conditions with little change to the posterior estimates.

	Mean	Std		Mean	Std		Mean	Std
p_{11}	0.9866	0.0065	$c_{1,1}$	0.0086	0.0053	$\Sigma_{3,22}$	697340	661360
p_{12}	0.0106	0.0059	$c_{1,2}$	-0.0010	0.0039	$B_{1,11}$	-0.0896	0.0775
p_{13}	0.0028	0.0020	$c_{2,1}$	0.7187	0.0832	$B_{1,12}$	-0.0047	0.0077
p_{21}	0.0301	0.0156	$c_{2,2}$	0.7324	0.1597	$B_{1,21}$	0.3354	0.7994
p_{22}	0.9626	0.0170	$c_{3,1}$	-0.0220	9.9651	$B_{1,22}$	-0.1451	0.0772
p_{23}	0.0073	0.0053	$c_{3,2}$	0.0034	10.0120	$B_{2,11}$	-0.3267	0.0500
p_{31}	0.2502	0.1664	$\Sigma_{1,11}$	0.0143	0.0009	$B_{2,12}$	0.0677	0.0062
p_{32}	0.2493	0.1658	$\Sigma_{1,12}$	0.1177	0.0086	$B_{2,21}$	-0.1144	0.5960
p_{33}	0.5005	0.2239	$\Sigma_{1,22}$	1.3937	0.0979	$B_{2,22}$	-0.1916	0.0855
ρ	0.0327	0.0321	$\Sigma_{2,11}$	0.0005	0.0003	$B_{3,11}$	-0.0421	9.9667
σ_1	0.0159	0.0017	$\Sigma_{2,12}$	-0.0171	0.0131	$B_{3,12}$	0.1126	9.9987
σ_2	0.0553	0.0151	$\Sigma_{2,22}$	2.0929	1.7203	$B_{3,21}$	0.0209	10.0104
σ_3	$(\frac{\pi}{180})^2$	0	$\Sigma_{3,11}$	2648.9	2876.1	$B_{3,22}$	0.0824	10.0021
σ_4	0.0152	0.0034	$\Sigma_{3,12}$	6647.8	15797			

Table 3. Posterior parameter estimates: equity market data. Note that $\Sigma_{i,jk}$ refers to the j,kth element of Σ_i $(i = 1, 2, 3)$. $B_{i,jk}$ refers to the j,kth element of B_i. $c_{0,ij}$ refers to the intercept for variable j $(j = 1, 2)$ in regime i $(i = 1, 2, 3)$.

	$\widehat{P}(M_1)$	$\widehat{P}(M_2)$	$\widehat{P}(M_3)$
Probability	0.7320	0.2601	0.0079

Table 4. Model probabilities: equity market data

As is evident from Figure 6, the regimes S_{1t}, S_{2t} are highly persistent and this is reflected in the transition estimates $p_{11} = 0.9866$ and $p_{22} = 0.9626$. These estimates suggest that the cointegrating rank does not change dramatically over short periods. Consequently, the model produces infrequent changes in regime for the two indices. Overall, our test of the support for M_1, M_2 and M_3 indicates a greater support for M_1, with $\widehat{P}(M_1) = 0.7320$, relative to M_2 ($\widehat{P}(M_2) = 0.2601$), implying that the two indices are typically not cointegrated. Although there is a non-negligible weight associated with M_2, it is clear that the strict adoption of a model with a permanent cointegrating equation for the two indices is not supported by the data. In this respect, given the persistence of S_{1t}, the imposition of a permanent cointegrating relationship is likely to produce erroneous estimates about the behaviour of the data for significant periods of time.

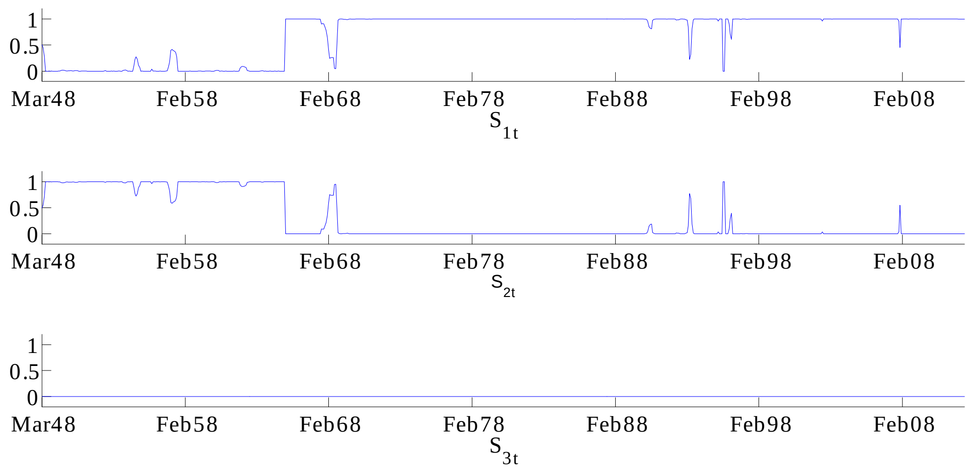


Figure 6. Posterior means of s_{1t}, s_{2t}, and s_{3t}.

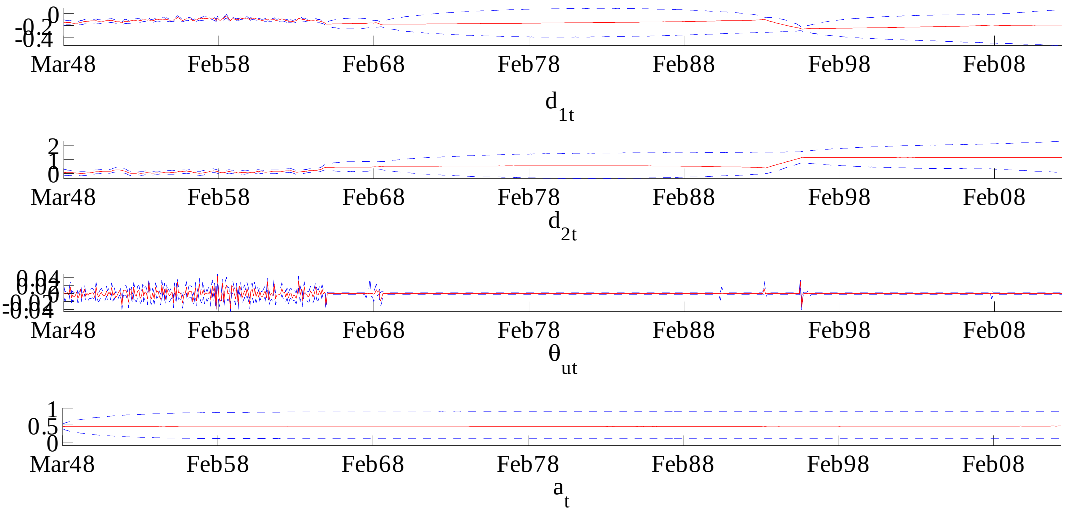


Figure 7. Estimated path of d_{1t}, d_{2t}, θ_{ut} and κ_t. The dotted and solid lines are, respectively, the 95 percent credible intervals and posterior means of d_{1t}, d_{2t}, θ_{ut} and κ_t.

6 Conclusion

We develop a bivariate model that can alternate between states of cointegration and non-cointegration, and between stationarity and non-stationarity. The model allows for all three competing relationships identified by the rank of Π; two variables may be I(1) with a single cointegrating relationship, I(1) with no cointegrating relationship, or I(0). In addition, since the rank of Π is allowed to vary over time, the model allows for data that behave in a cointegrated manner in some periods and not in others. It also allows for data that behave as stationary in some periods and non-stationary in others. As such, the econometrician can avoid ex ante judgments regarding the presence of cointegration or the timing of shifts in the rank of Π, both of which are often unclear in practice.

We utilise the properties of our model to develop a straightforward test among the

the three competing relationships. The test identifies the probabilities associated with each relationship without requiring the estimation of marginal likelihoods.

The model is applied to assess the Fisher effect. The results indicate that the rank of Π changes over time for US interest rates and inflation rates. For the majority of time periods, interest and inflation rates are cointegrated. There are periods, however, when the two rates behave as non-cointegrated I(1) variables. Interestingly, during the GFC crisis, our model suggests that interest rates and inflation rates are best modelled as I(0) variables. Over the entire course of the sample period, our model suggests that the Fisher effect's implication of a cointegrating vector $\begin{bmatrix} 1 & -1 \end{bmatrix}$ between interest and inflation rates is generally not supported.

The model is also applied to FTSE 100 and S&P 100 index values and indicates a shift from a long-term cointegrating relationship between the two indices prior to 2001, to a non-cointegrating relationship post-2001.

Appendix: Sampling methodology

The Metropolis-in-Gibbs sampler used in the paper consists of the 12 steps detailed in this Appendix. The following notation holds for each step.

$$g = vec\left(\begin{bmatrix} B'_{11} & B'_{12} & B'_{13} \end{bmatrix}'\right)$$

$$C = \begin{bmatrix} c_1 & c_2 & c_3 \end{bmatrix}'$$

$$\eta = \begin{bmatrix} \eta_0 & \eta_1 \end{bmatrix}'$$

$$\eta_0 = \begin{bmatrix} \eta_{01} & \eta_{02} \end{bmatrix}$$

$$\eta_1 = \begin{bmatrix} \eta_{11} & \eta_{12} \end{bmatrix}$$

$$\iota_t = I_2 \otimes \begin{bmatrix} s_{1t} & s_{2t} & s_{3t} \end{bmatrix}$$

$$U_t = \begin{bmatrix} \cos\theta_{ut} & \sin\theta_{ut} \\ -\sin\theta_{ut} & \cos\theta_{ut} \end{bmatrix}$$

$$x_t = I_2 \otimes \begin{bmatrix} s_{1t}\triangle y_{t-1} & s_{2t}\triangle y_{t-1} & s_{3t}\triangle y_{t-1} \end{bmatrix}$$

$$\begin{bmatrix} y^*_{1,t-1} & y^*_{2,t-1} \end{bmatrix} = y_{t-1}U_t$$

6.1 Step 1: Draw $\widetilde{S}_T$ given $\widetilde{y}_T, \widetilde{d}_T, \widetilde{\theta}_{uT}, \widetilde{\kappa}_T, g, \Sigma, \eta$,and C

Draws of $\widetilde{S}_T$ are based on the multi-move Gibbs Sampling algorithm of Carter and Kohn (1994) (see, also: Chib, 1996; Kim and Nelson, 1998). The conditional posterior probability for $S_t = j$, $j \in \{1, 2, 3\}$, is

$$\Pr(S_t = j|\widetilde{y}_t, d_t, \theta_{ut}, \kappa_t, g, \eta, C, \Sigma, S_{t+1}) = \frac{p(S_t = j|\widetilde{y}_T, d_t, \theta_{ut}, \kappa_t, g, \eta, C, \Sigma, S_{t+1})}{\sum_{i=1}^{3} p(S_t = i|\widetilde{y}_T, d_t, \theta_{ut}, \kappa_t, g, \eta, C, \Sigma, S_{t+1})}$$

$$p(S_t|\widetilde{y}_T, d_t, \theta_{ut}, \kappa_t, g, \eta, C, \Sigma, S_{t+1}) \propto p(S_t|\widetilde{y}_t, d_t, \theta_{ut}, \kappa_t, g, \eta, C, \Sigma)p(S_{t+1}|S_t)$$

where $p(S_{t+1}|S_t)$ is the transition probability and $p(S_t|\widetilde{y}_t, d_t, \theta_{ut}, \kappa_t, g, \eta, C, \Sigma)$ is obtained using Hamilton's (1989) basic filter. s_{jt} is set to 0 or 1 according to:

- $s_{1t} = 1$ if a draw u from the $U(0, 1)$ distribution is less than or equal to $\Pr(S_t = 1|\widetilde{y}_T, d_t, \theta_{ut}, g, \eta, C, \Sigma, \widetilde{S}_{\neq t})$.

- $s_{2t} = 1$ if u is between $\Pr(S_t = 1|\widetilde{y}_T, d_t, \theta_{ut}, g, \eta, C, \Sigma, \widetilde{S}_{\neq t})$ and $\Pr(S_t = 1|\widetilde{y}_T, d_t, \theta_{ut}, g, \eta, C, \Sigma, \widetilde{S}_{\neq t})$ $+ \Pr(S_t = 2|\widetilde{y}_T, d_t, \theta_{ut}, g, \eta, C, \Sigma, \widetilde{S}_{\neq t})$.

- $s_{3t} = 1$ if u is between $\Pr(S_t = 1|\widetilde{y}_T, d_t, \theta_{ut}, g, \eta, C, \Sigma, \widetilde{S}_{\neq t})$+$\Pr(S_t = 2|\widetilde{y}_T, d_t, \theta_{ut}, g, \eta, C, \Sigma, \widetilde{S}_{\neq t})$ and 1.

Step 2: Draw p given $\widetilde{S}_T$

Draws of p are based on the approach proposed in Kim, Nelson and Startz (1998). Conditional on $\widetilde{S}_T$, the transition parameters are independent of the remaining parameters. Draws from the transition probabilities proceed as follows

- Draw p_{ii} from

$$p_{ii}|\widetilde{S}_T \sim beta(\tau_{ii} + \vartheta_{ii}, \overline{\tau}_{ii} + \overline{\vartheta}_{ii}) \qquad i = 1, 2, 3.$$

- Draw $\overline{p}_{ij}$ from

$$\overline{p}_{ij}|\widetilde{S}_T \sim beta(\tau_{ij} + \vartheta_{ij}, \tau_{ik} + \vartheta_{ik}) \quad i \neq j \neq k \text{ and } i, j, k = 1, 2, 3.$$

- Compute p_{ij} and p_{ik}, $i \neq j \neq k$, as

$$p_{ij} = \overline{p}_{ij}(1 - p_{ii}),$$

$$p_{ik} = 1 - p_{ii} - p_{ij}.$$

Note that $\overline{p}_{ij} = \Pr(S_t = j|S_{t-1} = i, S_t \neq i)$, ϑ_{ij} is the total number of transitions from $S_{t-1} = i$ to $S_t = j$, $(i, j = 1, 2, 3$ and $t = 2, 3, ..., T)$. $\overline{\vartheta}_{ii}$, on the other hand, is the total number of transitions from $S_{t-1} = i$ to $S_t \neq i$. $\tau_{ii}, \overline{\tau}_{ii}, \tau_{ik}, \tau_{ij}$ are prior hyperparameters.

Step 3: Draw $\widetilde{d}_{1T}$ given $\widetilde{y}_T, \widetilde{S}_T, \widetilde{d}_{2T}, \widetilde{\theta}_{uT}, \widetilde{\kappa}_T, g, \Sigma, \eta, C, \mu, \rho$ and σ

Given $\widetilde{d}_{2T}$, $\widetilde{\kappa}_T$ and the rest of the parameters, $\widetilde{d}_{1T}$ can be estimated pursuant to the following linear, Gaussian state space model

$$z_t = \begin{bmatrix} \left(-\eta_{01} - t\eta_{11} + y^*_{1,t-1}\right)(1-s_{1t})(s_{2t}+s_{3t}) \\ \kappa_t\left(-\eta_{02} - t\eta_{02} + y^*_{2,t-1}\right)(1-s_{1t})s_{3t} \end{bmatrix} d_{1t} + \varepsilon'_t$$

$$d_{1t} = d_{1t-1} + v_{1t}$$

where $z_t = \triangle y'_t - \iota_t C - x_t g - \begin{bmatrix} \kappa_t\left(-\eta_{02} - t\eta_{02} + y^*_{2,t-1}\right)(1-s_{1t})s_{3t} \\ \left(\eta_{01} + t\eta_{01} - y^*_{1,t-1}\right)(1-s_{1t})(s_{2t}+s_{3t}) \end{bmatrix} d_{2t}$. Accordingly, the Kalman filter and smoother can be used to obtain the conditional density of $\widetilde{d}_{1T}$. Draws are obtained pursuant to Carter and Kohn's (1994) multi-move Gibbs sampler.

Step 4: Draw $\widetilde{d}_{2T}$ given $\widetilde{y}_T, \widetilde{S}_T, \widetilde{d}_{1T}, \widetilde{\theta}_{uT}, \widetilde{\kappa}_T, g, \Sigma, \eta, C, \mu, \rho$ and σ

The draws for $\widetilde{d}_{2T}$ are obtained using the same approach as for $\widetilde{d}_{1T}$. Given $\widetilde{d}_{1T}$, $\widetilde{\kappa}_T$ and the rest of the parameters, the following state space model can be used to estimate $\widetilde{d}_{2T}$

$$z^*_t = \begin{bmatrix} \kappa_t\left(-\eta_{02} - t\eta_{12} + y^*_{2,t-1}\right)(1-s_{1t})s_{3t} \\ \left(\eta_{01} + t\eta_{11} - y^*_{1,t-1}\right)(1-s_{1t})(s_{2t}+s_{3t}) \end{bmatrix} d_{2t} + \varepsilon'_t$$

$$d_{2t} = d_{2t-1} + v_{2t}$$

where $z^*_t = \triangle y'_t - \iota_t C - x_t g - \begin{bmatrix} \left(-\eta_{01} - t\eta_{11} + y^*_{1,t-1}\right)(1-s_{1t})(s_{2t}+s_{3t}) \\ \kappa_t\left(-\eta_{02} - t\eta_{12} + y^*_{2,t-1}\right)(1-s_{1t})s_{3t} \end{bmatrix} d_{1t}$.

Step 5: Draw $\widetilde{\theta}_{uT}$ given $\widetilde{y}_T, \widetilde{S}_T, \widetilde{d}_T, \widetilde{\kappa}_T, g, \Sigma, \eta, C, \mu, \rho$ and σ

Conditional on the latent variables, regimes and other parameters, θ_{ut} can be represented using the following non-linear, Gaussian state space form

$$\triangle r_t = \left(\left(y_{t-1}U_tI(S_t)\right)I(S_t)\otimes I_2\right)\begin{bmatrix} d_{1t} \\ -d_{2t} \\ \kappa_t d_{2t} \\ \kappa_t d_{1t} \end{bmatrix} + \varepsilon_t'$$

$$\theta_{ut} = \mu_3 + \rho_3\theta_{ut-1} + v_{3t}$$

where $\triangle r_t = \triangle y_t' - \iota_t C + (\eta_0 I(S_t)\otimes I_2 + t\eta_1 I(S_t)\otimes I_2)\begin{bmatrix} d_{1t} \\ -d_{2t} \\ \kappa_t d_{2t} \\ \kappa_t d_{1t} \end{bmatrix} - x_t g$. Pitt and Shephard's (1999) auxiliary particle filter is used to produce draws of $\widetilde{\theta}_{uT}$.

6.2 Step 6: Draw $\widetilde{\kappa}_T$ given $\widetilde{S}_T, \widetilde{y}_T, \widetilde{\theta}_{uT}, \widetilde{d}_T, g, \eta, C, \Sigma, , \mu, \rho$ and σ

Conditional on the other parameters, $\widetilde{\kappa}_T$ can be represented using the following state space form

$$\widehat{z}_t = \begin{bmatrix} \left(-\eta_{02} - t\eta_{12} + y^*_{2,t-1}\right)(1-s_{1t})s_{3t}d_{2t} \\ \left(-\eta_{02} - t\eta_{12} + y^*_{2,t-1}\right)(1-s_{1t})s_{3t}d_{1t} \end{bmatrix}\kappa_t + \varepsilon_t'$$

$$\kappa_t = \kappa_{t-1} + v_{4t}$$

where $\widehat{z}_t = \triangle y_t' - \iota_t C - x_t g - \begin{bmatrix} \left(-\eta_{01} - t\eta_{11} + y^*_{1,t-1}\right)(1-s_{1t})(s_{2t}+s_{3t})d_{1t} \\ -\left(-\eta_{01} - t\eta_{11} + y^*_{1,t-1}\right)(1-s_{1t})(s_{2t}+s_{3t})d_{2t} \end{bmatrix}$. Given the restriction $\kappa_t \in [0,1]$, the error term v_{4t} must be truncated ensure κ_t always lies in the restricted space. The resulting model can be regarded as a nonlinear state space model. As such, we use the auxiliary particle filter (Pitt and Shephard, 1999) to draw $\widetilde{\kappa}_T$. The method proposed in Chan, Koop and Potter (2013) may also be used.

Step 7: Draw σ given $\widetilde{d}_T, \widetilde{\kappa}_T, \mu$ and ρ

Conditional on $\widetilde{d}_T$, it can be shown that the conditional posterior densities of σ_1 and σ_2 are inverse gamma

$$\sigma_i|\widetilde{d}_{iT} \sim IG\left(0.5\left(\vartheta_i + T\right),\ 0.5\left(\vartheta_i f_i + \sum_2^T (d_{it} - d_{it-1})^2\right)\right), \quad i = 1, 2.$$

Due to the restriction on κ_t's error term v_{4t}, the conditional posterior density of σ_4 does not follow a standard distribution. Accordingly, the independent Metropolis-Hastings algorithm is used to draw σ_4. To implement the algorithm, we adopt the following inverse gamma proposal density

$$\sigma_4 \sim IG\left(v_4, k\widehat{\sigma}_4\right)$$

where $k = 1.5$, $v_4 = 2$ and $\widehat{\sigma}_4$ is a variance estimate obtained by OLS regression of κ_t on its lag κ_{t-1}.

Step 8: Draw μ, ρ given $\widetilde{d}_{1T}, \widetilde{d}_{2T}, \widetilde{\theta}_{uT}$ and σ

Since $cov\left(v_{it}, v_{jt}\right) = 0$ for $i \neq j$, the draws μ, ρ for each of the $\widetilde{d}_{1T}, \widetilde{d}_{2T}$, and $\widetilde{\theta}_{uT}$ can be obtained separately. Draws of μ_3, ρ_3 are obtained as follows, with draws for (μ_1, ρ_1) and (μ_2, ρ_2) being obtained in an analogous manner.

Given $\widetilde{\theta}_{uT}$, set $X^* = \begin{bmatrix} 1_{T-1} & \begin{bmatrix} \theta_{u1} & ... & \theta_{uT-1} \end{bmatrix}' \end{bmatrix}$ and $y^* = \begin{bmatrix} \theta_{u2} & ... & \theta_{uT} \end{bmatrix}'$. The prior for $(\mu_3, \rho_3)'$ is normal with mean $\underline{\mu} = \left[\underline{\mu_3}, \underline{\rho_3}\right]'$ and variance $\underline{H} = diag\left(\sigma_{\mu_3}, \sigma_{\rho_3}\right)$. The conditional posterior density of $(\mu_3, \rho_3)'$ is given by

$$(\mu_3, \rho_3)' \,|\widetilde{\theta}_{uT}, \sigma_3 \sim MVN\left(\overline{\mu}, \overline{H}\right) I(\rho_3)$$

where $\overline{H} = \left(\sigma_3^{-1} X^{*\prime} X^* + \underline{H}^{-1}\right)^{-1}$ and $\overline{\mu} = \overline{H}\left(\underline{H}^{-1}\underline{\mu} + \sigma_3^{-1} X^{*\prime} y^*\right)$. $I(\rho_3)$ is an indicator variable taking on the value unity *iff* $|\rho_3| < 1$.

Step 9: Draw g given $\widetilde{y}_T, \widetilde{S}_T, \widetilde{d}_T, \widetilde{\theta}_{uT}, \widetilde{\kappa}_T, \eta, C$ and Σ

Conditional on the remaining parameters, the mean and variance of g can be obtained from the following linear, Gaussian SUR equation

$$\triangle y_t^* = x_t g + \varepsilon_t \tag{28}$$

where $\triangle y_t^* = \triangle y_t' - \iota_t C - \alpha_t' \beta' y_{t-1}' + \alpha_t' \eta_0' + \alpha_t' \eta_1' t$. Given the adoption of a multivariate normal prior for g, the conditional posterior density of g is also multivariate normal

$$g|\widetilde{\theta}_{uT}, \widetilde{y}_T, \widetilde{d}_T, \widetilde{S}_T, \eta, C, \Sigma \sim MVN(\overline{g}, \overline{\Sigma}_g)$$

with $\overline{\Sigma}_g = \left(X_g'(\Sigma^{-1} \otimes I_T)X_g + \Sigma_g\right)^{-1}$, $\overline{g} = \overline{\Sigma}_g \left(\Sigma_g^{-1}\underline{g} + X_g'(\Sigma^{-1} \otimes I_T)\triangle Y^*\right)$, $X_g = I_2 \otimes \begin{bmatrix} \triangle Y_{-1} & ... & \triangle Y_{-l} \end{bmatrix}$, $\triangle Y^* = \begin{bmatrix} \triangle y_1^* \\ \triangle y_2^* \end{bmatrix}$, $\triangle y_i^* = \begin{bmatrix} \triangle y_{i,1}^* & ... & \triangle y_{i,T}^* \end{bmatrix}'$ and $\triangle Y_{-j} = \begin{bmatrix} \triangle y_{1,-j} & ... & \triangle y_{1,T-j} \\ \triangle y_{2,-j} & ... & \triangle y_{2,T-j} \end{bmatrix}'$.

Step 10: Draw C given $\widetilde{y}_T, \widetilde{S}_T, \widetilde{d}_T, \widetilde{\theta}_{uT}, \widetilde{\kappa}_T, \eta, g$ and Σ

The conditional mean and variance of C can be obtained from

$$\triangle \widehat{y}_t^* = \iota_t C + \varepsilon_t' \tag{29}$$

where $\triangle \widehat{y}_t^* = \triangle y_t' - \alpha_t' \beta_t' y_{t-1}' - x_t g + \alpha_t' \eta_0' + \alpha_t' \eta_1' t$. As per (28), (29) takes the linear, Gaussian SUR form. Given a multivariate normal prior, the conditional posterior density of C is also multivariate normal

$$C|\widetilde{\theta}_{uT}, \widetilde{y}_T, \widetilde{d}_T, \widetilde{S}_T, \eta, g, \Sigma \sim MVN(\overline{c}', \overline{\Sigma}_c)$$

with $\overline{\Sigma}_c = \left(X_c'(\Sigma^{-1} \otimes I_T)X_c + \Sigma_c\right)^{-1}$, $\overline{c} = \overline{\Sigma}_c \left(\Sigma_c^{-1}\underline{c} + X_c'(\Sigma^{-1} \otimes I_T)\triangle\widehat{Y}^*\right)$, $X_c = \begin{bmatrix} X_{1c} \\ X_{2c} \end{bmatrix}$,
$\triangle\widehat{Y}^* = \begin{bmatrix} \triangle\widehat{y}_1^* \\ \triangle\widehat{y}_2^* \end{bmatrix}$, $\triangle\widehat{y}_i^* = \begin{bmatrix} \triangle\widehat{y}_{i,1}^* & ... & \triangle\widehat{y}_{i,T}^* \end{bmatrix}'$, $X_{ic} = \begin{bmatrix} \iota_{i,t}' & ... & \iota_{i,T}' \end{bmatrix}'$ for $i \in \{1, 2\}$, and $\iota_{i,t}$ is the i-th row of ι_t.

Step 11: Draw η given $\widetilde{y}_T, \widetilde{S}_T, \widetilde{d}_T, \widetilde{\theta}_{uT}, \widetilde{\kappa}_T, g, C$ and Σ

The conditional mean and variance of η can be obtained from

$$\triangle\widehat{y}_t^* = x_t^*\eta + \varepsilon_t' \tag{30}$$

where $\triangle\widehat{y}_t^* = \triangle y_t' - \iota_t C - \alpha_t'\beta_t'y_{t-1}' - x_t g$ and $x_t^* = \begin{bmatrix} -\alpha_t' & -t\alpha_t' \end{bmatrix}$. (30) takes the linear, Gaussian SUR form. Given that the prior for η is multivariate normal, its conditional posterior is also multivariate normal

$$\eta|\widetilde{\theta}_{uT}, \widetilde{y}_T, \widetilde{d}_T, \widetilde{S}_T, g, \Sigma \sim MVN(\overline{\eta}', \overline{\Sigma}_\eta)$$

with $\overline{\Sigma}_\eta = \left(X_\eta'(\Sigma^{-1} \otimes I_T)X_\eta + \Sigma_\eta\right)^{-1}$, $\overline{\eta} = \overline{\Sigma}_\eta \left(\Sigma_\eta^{-1}\underline{\eta} + X_\eta'(\Sigma^{-1} \otimes I_T)\triangle\widehat{Y}^*\right)$, $X_\eta = \begin{bmatrix} X_{1\eta} \\ X_{2\eta} \end{bmatrix}$, $\triangle\widehat{Y}^* = \begin{bmatrix} \triangle\widehat{y}_1^* \\ \triangle\widehat{y}_2^* \end{bmatrix}$, $\triangle\widehat{y}_i^* = \begin{bmatrix} \triangle\widehat{y}_{i,1}^* & ... & \triangle\widehat{y}_{i,T}^* \end{bmatrix}'$, $X_{i\eta} = \begin{bmatrix} x_{i,t}^{*\prime} & ... & x_{i,T}^{*\prime} \end{bmatrix}'$ for $i \in \{1, 2\}$, and $x_{i,t}^*$ is the i-th row of x_t^*.

Step 12: Draw Σ given $\widetilde{y}_T, \widetilde{S}_T, \widetilde{d}_T, \widetilde{\theta}_{uT}, \widetilde{\kappa}_T, g, \eta$ and C

Given the adoption of an inverse Wishart prior for Σ_i, $i = 1, 2, 3$ such that $\Sigma_{S_t} = \Sigma_1 s_{1t} + \Sigma_2 s_{2t} + \Sigma_3 s_{3t}$, the conditional posterior of Σ_i is also inverse Wishart with

$$\Sigma_i | \theta_u, \widetilde{y}_T, \widetilde{d}_T, \widetilde{S}_T, g, \eta, C \sim IW(\underline{v}_i + T_i, \underline{D}_i + \overline{A}_i)$$

where $\overline{A}_i = \sum_{\forall t} s_{it} (\varepsilon_t' \varepsilon_t)$, $\varepsilon_t = \triangle y_t' - \iota_t C + \alpha_t' \eta_0' + \alpha_t' \eta_1' t - \alpha_t' \beta_t' y_{t-1}' - x_t g$ and T_i is the number of times regime i prevails over $t = 1$ to T.

References

[1] Andrade, P., Bruneau, C. and Gregoir, S. (2005), "Testing for the cointegration rank when some cointegrating directions are changing", Journal of Econometrics, 124, 269–310.

[2] Bai, J. and Perron, P. (2005), "Computation and Analysis of Multiple Structural Change Models", Journal of Applied Econometrics, 18(1), 1-22.

[3] Bernanke, B.S., 2009. The Crisis and the Policy Response. Speech given on January 13 at the Stamp Lecture, London School of Economics, London, England.

[4] Bierens, H. J. and Martins, L. F. (2010), "Time-varying cointegration," Econometric Theory, 26, 1453–1490.

[5] Carter, C.K. and Kohn, R. (1994), "On Gibbs sampling for state space models", Biometrika, 81, 541–553.

[6] Chan, J.G., Koop, G. and Potter, S. (2013), "A New Model of Trend Inflation", Journal of Business and Economic Statistics, 31(1), 94-106.

[7] Chib, S. (1996), "Calculating posterior distributions and modal estimates in Markov mixture models", Journal of Econometrics, 75, 79–97.

[8] Chib, S and Greenberg, E. (1996), "Markov chain Monte Carlo simulation methods in econometrics", Econometric Theory, 12, 409–431.

[9] Engle, R. F. and Granger, C. W. J. (1987), "Co-integration and error correction: Representation, estimation and testing", Econometrica, 55, 251–276.

[10] Franses, P.H. (2001). "How to deal with intercept and trend in practical cointegration analysis?," Applied Economics, 33, 577–579.

[11] Gregory, A.W. and Hansen, B.E. (1996a), "Residual-based tests for cointegration in models with regime shifts", Journal of Econometrics, 70, 99–126.

[12] Gregory, A.W. and Hansen, B.E. (1996b), "Tests for cointegration in models with regime and trend shifts", Oxford Bulletin of Economics and Statistics, 58, 555–560.

[13] Hansen, P.R. (2003), "Structural changes in the cointegrated vector autoregressive model", Journal of Econometrics 114, 261–295.

[14] Hamilton, J (1989), "A new approach to the economic analysis of nonstationary time series and the business cycle", Econometrica, 57, 357–384.

[15] Hoff, P.D. (2012), "Simulation of the matrix Bingham-von-Mises-Fisher distribution, with applications to multivariate and relational data", Journal of Computational and Graphical Statistics, 18, 438-456.

[16] Johansen, S. (1991), "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models", Econometrica, 59, 1551–1580.

[17] Johansen, S., Mosconi, R. and Nielsen (2000), "Cointegration analysis in the presence of structural breaks in the deterministic trend", The Econometrics Journal, 3, 216–249.

[18] Jochmann, M. and Koop, G. (2011), "Regime-Switching Cointegration", Working Papers 1125, University of Strathclyde Business School, Department of Economics.

[19] Kim, C.J. and Nelson, C.R. (1998), "Business cycle turning points, a new coincident index, and tests of duration dependence based on a dynamic factor model with regime switching", Review of Economics and Statistics, 80, 188–201.

[20] Kim, C.J., Nelson, C.R. and Startz, R. (1998), "Testing for mean reversion in heteroskedastic data based on Gibbs-sampling-augmented randomization," Journal of Empirical Finance, 5, 131–154.

[21] Kleibergen, F. and Paap, R. (2002), "Priors, posteriors and bayes factors for a Bayesian analysis of cointegration", Journal of Econometrics, 111, 223–249.

[22] Kleibergen, F. and van Dijk, H.K. (1998), "Bayesian simultaneous equation analysis using reduced rank structures", Econometric Theory, 14, 701–743.

[23] Koop, G. Leon-Gonzalez, R., and Strachan, R. (2011), "Bayesian inference in a time varying cointegration model", Journal of Econometrics, 165, 210–220.

[24] Koustas, Z., and Serletis, A. (1999), "On the Fisher Effect", Journal of Monetary Economics, 44(1), 105–130.

[25] Lutkepohl, H. (1996), Handbook of Matrices, Wiley.

[26] Neely, C.J., and Rapach, D.E. (2008), "Real Interest Rate Persistence: Evidence and Implications", Federal Reserve Bank of St Louis Review, 90(6), 609–641.

[27] Martin, G. M. (2000), "US deficit sustainability: a new approach based on multiple endogenous breaks", Journal of Applied Econometrics,15, 83–105.

[28] Paap, R. and van Dijk, H.K. (2003), "Bayes estimates of Markov trends in possibly cointegrated series: an application to US consumption and income", Journal of Business Economics and Statistics, 21, 547–563.

[29] Pitt, M. and Shephard, N. (1999), "Filtering via simulation: auxiliary particle filter", Journal of the American Statisitical Association, 94, 590–559.

[30] Saikkonen P., and Choi, I. (2004), "Cointegrating smooth transition regressions", Econometric Theory, 20, 301–340.

[31] Strachan, R. and Inder, B.A. (2004) “Bayesian analysis of the error correction model”, Journal of Econometrics 123, 307–325.

[32] Strachan, R. (2003), "Valid Bayesian estimation of the cointegrating error correction model", Journal of Business and Economic Statistics, 21(1), 185–195.

[33] Sugita, K. (2006), "Bayesian analysis of dynamic multivariate models with multiple structural breaks," Discussion paper No 2006-14, Graduate School of Economics, Hitotsubashi University.

[34] Villani, M. (2006), “Bayesian point estimation of the cointegration space”, Journal of Econometrics, 134, 645–664.