

# **STRATHCLYDE**

**DISCUSSION PAPERS IN ECONOMICS**



**REGIME-SWITCHING COINTEGRATION\***

**BY**

**MARKUS JOCHMANN AND GARY KOOP**

**No. 11-25**

**DEPARTMENT OF ECONOMICS  
UNIVERSITY OF STRATHCLYDE  
GLASGOW**

# Regime-Switching Cointegration\*

Markus Jochmann  
Newcastle University  
Rimini Centre for Economic Analysis

Gary Koop  
University of Strathclyde  
Rimini Centre for Economic Analysis

May 2011

## Abstract

We develop methods for Bayesian inference in vector error correction models which are subject to a variety of switches in regime (e.g. Markov switches in regime or structural breaks). An important aspect of our approach is that we allow both the cointegrating vectors and the number of cointegrating relationships to change when the regime changes. We show how Bayesian model averaging or model selection methods can be used to deal with the high-dimensional model space that results. Our methods are used in an empirical study of the Fisher effect.

Keywords: Bayesian, Markov switching, structural breaks, cointegration, model averaging

JEL codes: C11, C32, C52

---

\*All authors are Fellows at the Rimini Centre for Economic Analysis. We acknowledge financial support from the Leverhulme Trust under Grant F/00 273/J.

# 1 Introduction

Two of the most important challenges of modern empirical macroeconomics involve the wish to incorporate restrictions suggested by economic theory and the empirical need to allow for parameter change in multivariate time series models. With regard to the former, cointegration has played an important role as economic theory often suggests particular cointegrating relationships which the researcher may wish to impose or test for. As one example, consider the UK macroeconomic model of Garratt, Lee, Pesaran and Shin (2003). This uses the purchasing power parity relationship, an interest rate parity condition, a neoclassical growth model, the Fisher hypothesis and a theory of portfolio balance to build a macroeconometric model involving five cointegrating relationships. With regard to the latter, papers such as Ang and Bekaert (2002) and Stock and Watson (1996) document widespread evidence of parameter change in many macroeconomic time series. In the field of cointegration, there are a large number of theoretical and empirical papers that model breaks or other forms of nonlinearity in cointegrating relationships, present empirical results relating to cointegration work using subsamples of the data or attribute failures of cointegration tests to parameter change (see, among many others, Michael, Nobay and Peel, 1997, Quintos, 1997, Park and Hahn, 1999, Lettau and Ludvigson, 2004, Saikkonen and Choi, 2004, Andrade, Bruneau and Gregoir, 2005, Beyer, Haug and Dewald, 2009 and Bierens and Martins, 2010).

All this work provides evidence of widespread empirical and theoretical interest in cointegration models with changing cointegrating spaces. However, with few exceptions (e.g. Martin, 2000, Paap and van Dijk, 2003 and Sugita, 2006 and Koop, León-González and Strachan, 2008) this work is non-Bayesian. One purpose of the present paper is to provide a set of Bayesian tools for working with Vector Error Correction Models (VECMs) in the presence of changes in regime. The previous Bayesian work with time-varying cointegration typically assumes cointegrating rank is constant across regimes (e.g. Koop, León-González and Strachan, 2008a) or works with much simpler model spaces than the one considered here (e.g. Martin, 2000, Paap and van Dijk, 2003 and Sugita, 2006).

A second purpose of this paper is to address the issues that arises with cointegration models due to the fact that the model space can be large. The researcher will typically consider models with different cointegrating ranks, different restrictions imposed on the cointegrating relationships, different lag lengths, different treatment of deterministic terms, etc. In previous work (see Koop, Potter and Strachan, 2008 and Jochmann, Koop, León-González and Strachan, 2011), we have developed Bayesian methods to navigate through such high-dimensional model spaces in the constant coefficient VECM. In this paper, we work with models with regime change and, in these, the dimension of the model space is greatly increased. For instance, we may wish to allow the cointegrating rank to differ across regimes or a restriction implied by economic theory to hold at some time periods but not others (e.g. we might have

purchasing power parity holding in the 1970s but not the 1980s). Furthermore, in practice it is typically unclear what determines changes in regime. Of models that allow for regime change, structural break models assume breaks occur at specific points of time and regimes do not recur. Markov switching models allow for regimes to recur (i.e. the model switches between expansionary and recessionary dynamics). It is empirically-sensible to work with a model space that allows for a range of such possibilities. Thus, a final contribution of this paper lies in the fact we offer a richer treatment of regime change, allowing for both structural break and Markov switching behavior. We show how, regardless of whether the researcher wishes to do Bayesian model averaging (BMA) or select a single model, the Bayesian approach is an attractive one in model spaces of this dimension.

Our methods are applied in an empirical exercise investigating the Fisher effect.

## 2 VECMs with Regime Switching

### 2.1 A General Framework

An unrestricted VECM for an  $n$ -dimensional vector  $y_t$  can be written as:

$$\Delta y_t = \alpha' \beta' y_{t-1} + \sum_{j=1}^{p-1} \gamma_j \Delta y_{t-j} + \varepsilon_t, \quad (1)$$

where  $\alpha$  is a full rank  $r \times n$  matrix,  $\beta$  is a full rank  $n \times r$  matrix,  $\gamma_j$  is  $n \times n$  and  $\varepsilon_t \sim N(0, \Sigma)$ .  $r$  and  $p$  are the number of cointegrating relationships and lag length, respectively. For notational simplicity, we have not included deterministic terms in (1). See, e.g., Johansen (1995, Section 5.7) or Franses (2001) for a discussion of deterministic terms in VECMs.

A wide range of regime switching VECMs can be obtained by adding  $s_t$  subscripts to the parameters in (1), leading to:

$$\Delta y_t = \alpha'_{s_t} \beta'_{s_t} y_{t-1} + \sum_{j=1}^{p-1} \gamma_{j,s_t} \Delta y_{t-j} + \varepsilon_t, \quad (2)$$

with  $\varepsilon_t \sim N(0, \Sigma_{s_t})$  and  $s_t \in \{1, \dots, M\}$  indicating which of  $M$  regimes applies at time  $t$ . Importantly, we assume  $\alpha_{s_t}$  is  $r_{s_t} \times n$  and  $\beta_{s_t}$  are  $n \times r_{s_t}$  so that the cointegrating rank can change when the regime changes.

Examples of models that can be put in this framework include Markov switching, other regime switching models such as endogenous threshold models, structural break models and time-varying parameter models. In the Bayesian multivariate time series literature, the emphasis has been on extensions to Vector Autoregressive (VAR) models. Prominent examples include the Markov switching VAR of Sims and Zha (2006)

and the time varying parameter (TVP) VARs of Cogley and Sargent (2005) and Primiceri (2005). However, VARs are parameter-rich models and VARs with regime change are even more parameter rich. This has led to approaches which attempt to mitigate over-parametrization worries by using shrinkage priors and impose restrictions. Cointegration provides a good source of potential restrictions (often motivated by economic theory) which can help achieve parsimony.

## 2.2 Modeling the Regime Switching Process

Many specifications for  $S^T = (s_1, \dots, s_T)'$  are possible. For instance, Koop, León-González and Strachan (2008a) set  $s_t = t$  and  $M = T$ , resulting in a TVP-VECM. However, their model assumes a common cointegrating rank at all points in time, an assumption we wish to relax in the present paper. Sugita (2006) assumes a Uniform prior over break dates which involves the assumption that  $s_t$  is sequentially increasing (i.e.  $s_t = s_{t-1} + 1$  if a break occurs at time  $t$ ). Such an approach can be computationally daunting in the case of multiple breaks. That is, with one break in a sample of size  $T$  there are on the order of  $T$  possible break dates, but with  $M - 1$  breaks this increases to the order of  $T^M$  which can lead to a serious computational burden if additional structure is not placed on  $S^T$ .

Many approaches in the literature can be interpreted as placing a particular structure on  $S^T$  using hierarchical priors. In this paper, we consider one class of hierarchical priors using Markov specifications for  $S^T$ . These are empirically popular in many contexts and convenient and computationally efficient MCMC algorithms exist (e.g., Chib, 1996). A standard Markov switching specification of the sort used, e.g., in Sims and Zha (2006) has:

$$\Pr(s_t = j | s_{t-1} = i) = \xi_{ij}, \quad i, j = 1, \dots, M, \quad (3)$$

where  $\xi_{ij}$  is the probability of switching from regime  $i$  to regime  $j$ . In the Markov switching model no restrictions (other than the ones implied by probabilities summing to one) are placed on the  $\xi_{ij}$ .

Chib (1998) notes that a Markov switching model can be turned into a structural break model by placing restrictions on the  $\xi_{ij}$ . In particular, he sets  $\xi_{ij} = 0$  for all  $i$  and  $j$  except for the following:

$$\begin{aligned} \Pr(s_t = i | s_{t-1} = i) &= \xi_{ii}, \quad i = 1, \dots, M - 1, \\ \Pr(s_t = i + 1 | s_{t-1} = i) &= 1 - \xi_{ii}, \quad i = 1, \dots, M - 1, \\ \Pr(s_t = M | s_{t-1} = M) &= 1. \end{aligned} \quad (4)$$

It can be seen that this leads to a model with  $M - 1$  structural breaks. That is, if regime  $i$  holds at time  $t - 1$ , then at time  $t$  the process can either remain in regime  $i$  (with probability  $\xi_{ii}$ ) or a break occurs and the process moves to regime  $i + 1$  (with

probability  $1 - \xi_{ii}$ ). The process moves through regimes sequentially (i.e. it cannot jump from regime  $i$  to regime  $i + 2$ ). Once a break occurs, the process cannot revert to an old regime (i.e. it cannot jump from regime  $i$  to regime  $i - 1$ ).

By modelling  $S^T$  in terms of a Markov process we obtain a computationally feasible model (using the algorithm of Chib, 1996) and can allow for regime switching behavior of various sorts. We can have a conventional Markov switching formulation where VECM coefficients vary over the business cycle (or in some other manner) or a structural break model where coefficients change at particular points in time. These are the two specifications for the break process considered in this paper. However, any specification for the  $\xi_{ij}$  can be used with the methods outlined in this paper and only trivial alterations would be required to accommodate other specifications for  $S^T$ .

## 2.3 Model Space

The previous material outlines a general modeling framework for regime-switching VECMs. The resulting model space can be large since we allow for both  $\beta_{s_t}$  and  $r_{s_t}$  to differ across regimes. Furthermore, we may wish to consider models which impose restrictions on  $\beta_{s_t}$ . For instance, in our empirical work, we consider versions of the model which impose the restriction  $\beta_{s_t} = (1, -1)'$  which is the value implied by Fisher's hypothesis. The cointegration rank  $r_{s_t}$  can be either be 0 or 1 and we consider lag lengths  $p = 1, 2, 3$ . In the case of structural breaks we analyze models with 2 and 3 regimes which already gives us 81 models. This does not even include modeling choices such as the treatment of deterministic terms which will increase the model space even more. With model spaces of this size, sequential hypothesis testing procedures can be risky. BMA (which averages over all models with weights proportional to posterior model probabilities<sup>1</sup>) or model selection (which chooses the single model with the highest posterior model probability) are attractive alternatives. But this suggests the need for efficient posterior simulation and marginal likelihood calculation methods and it is to these we now turn.

## 3 Bayesian Inference in Regime Switching VECMs

The Appendix contains complete details on priors, posterior simulation and marginal likelihood calculation. Here we provide a summary of the main ideas involved in each.

---

<sup>1</sup>In the case where all models, a priori, are given equal weight, posterior model probabilities are proportional to marginal likelihoods.

### 3.1 Prior Distributions

We let the vector  $\theta$  collect all model parameters. It contains the VECM parameters  $\{\alpha_i\}$ ,  $\{\beta_i\}$ ,  $\{\Gamma_i\}$  and  $\{\Sigma_i\}$ ,  $i = 1, \dots, M$  and the switching probabilities  $\{\xi_{ij}\}$ ,  $i, j = 1, \dots, M$ . For the latter it is common (e.g. Chib, 1998) to use Beta priors and we follow this practice. Our priors for the VECM parameters are the same as those used in previous work and are in all cases proper (thus, allowing for valid calculation of marginal likelihoods). We assume the priors in different regimes are independent of one another. The reader interested in a detailed motivation is referred to the previous literature (see, e.g., Strachan, 2003, Strachan and Inder, 2004, Koop, Potter and Strachan, 2008 and Koop, León-González and Strachan, 2010) with precise formulae being given in the Appendix. Briefly, for  $\{\alpha_i\}$  and  $\{\Gamma_i\}$  Normal shrinkage priors are used with similar properties to Minnesota priors. They reduce worries associated with over-fitting. For  $\{\Sigma_i\}$  inverted Wishart priors are used. Typically, in a cointegration analysis it is the priors for  $\{\beta_i\}$  which are most important. The basic idea of this prior is that, given the lack of identification of the VECM due to the product structure of the terms  $\{\alpha'_i\beta'_i\}$ , it is only the space spanned by the cointegrating vectors which is identified. In our empirical work, in regimes where cointegration is present, we use two different priors. In situations in which we do not want to restrict the cointegration space in a regime we use a Uniform prior over the cointegration space. In contrast, if we want the Fisher effect to hold in a regime we assume an informative prior centered over the space implied by the Fisher effect. It can be shown that these are both proper priors and, thus, valid marginal likelihoods can be obtained.

### 3.2 Posterior Simulation and Marginal Likelihood Calculation

Efficient posterior simulation in the VECM with the aforementioned prior can be implemented using the algorithm developed in Koop, León-González and Strachan (2010).  $S^T$  divides the sample into regimes. Thus, conditional on  $S^T$ , we can use this algorithm to draw the VECM coefficients in each regime. Conditional on posterior draws of the VECM coefficients, the algorithm of Chib (1996) (restricted as in Chib, 1998, for the structural break case), can be used to draw  $S^T$ .

Marginal likelihood calculation can be difficult in multivariate state space models such as the VECM. This has led to the use of approximations (e.g. the Laplace approximation of Strachan and Inder, 2004 or the information criteria of Koop, Potter and Strachan, 2008), methods based on the Savage-Dickey density ratio (e.g. Koop, León-González and Strachan, 2008b), methods which do not explicitly calculate the marginal likelihood in each model (e.g. the stochastic search variable selection approach of Jochmann, Koop, León-González and Strachan, 2011) or alternatives to the marginal likelihood such as the predictive likelihood (e.g. Geweke, 1996). Given a desire to directly use marginal likelihoods and avoid approximations, in this paper

we use a bridge sampler to calculate the marginal likelihood. See Gelman and Meng (1998) for a general treatment of bridge sampling and Frühwirth-Schnatter (2004) for bridge sampling in Markov switching models. Frühwirth-Schnatter (2004) compares bridge sampling with other methods and finds the former to be much more reliable and efficient.

Complete details of prior, posterior computation and bridge sampling are provided in the Appendix.

## 4 Application: The Fisher Effect

The Fisher effect is the name given to the theory which implies that a permanent change in inflation will, in the long run, cause an equal change in the nominal interest rate. Or, equivalently, monetary shocks will have no effect on the real interest rate in the long run. This can be taken to imply a cointegrating relationship between inflation,  $\pi_t$ , and the interest rate,  $i_t$ , with cointegrating vector  $(1, -1)'$ . This relationship has been investigated in numerous papers for numerous countries and is often found not to hold. Beyer, Haug and Dewalt (2009) offer a discussion of this literature and investigate whether structural breaks exist in the cointegrating relationship in a cross-country study.

In our empirical work, we look at the case of France. For this country, Beyer, Haug and Dewalt (2009) analyze quarterly data from 1970:Q1 to 2004:Q3 and find evidence of unit roots in  $\pi_t$  and  $i_t$  using classical unit root tests. However, both the Johansen trace and eigenvalue tests for cointegration indicate cointegration is not present and, thus, the Fisher effect appears not to hold. They next do a classical test where the null hypothesis is that cointegration is present, but with a structural break at an unknown point in time. This test does not reject the null hypothesis and finds a break in 1981:Q4. However, when Johansen tests are done using sub-samples (before and after 1981:Q4), the trace test finds cointegration in the second sub-sample but not in the first, whereas the eigenvalue test finds cointegration in both sub-samples. We take this as an interesting case where the evidence of previous work suggests there is a great deal of model uncertainty, both about the presence of cointegration and about the break process.

Our data on French CPI inflation (quarterly inflation at an annualized rate) and the 3 month treasury bill rate runs from 1970Q1-2010Q2 and is shown in Figure 1.



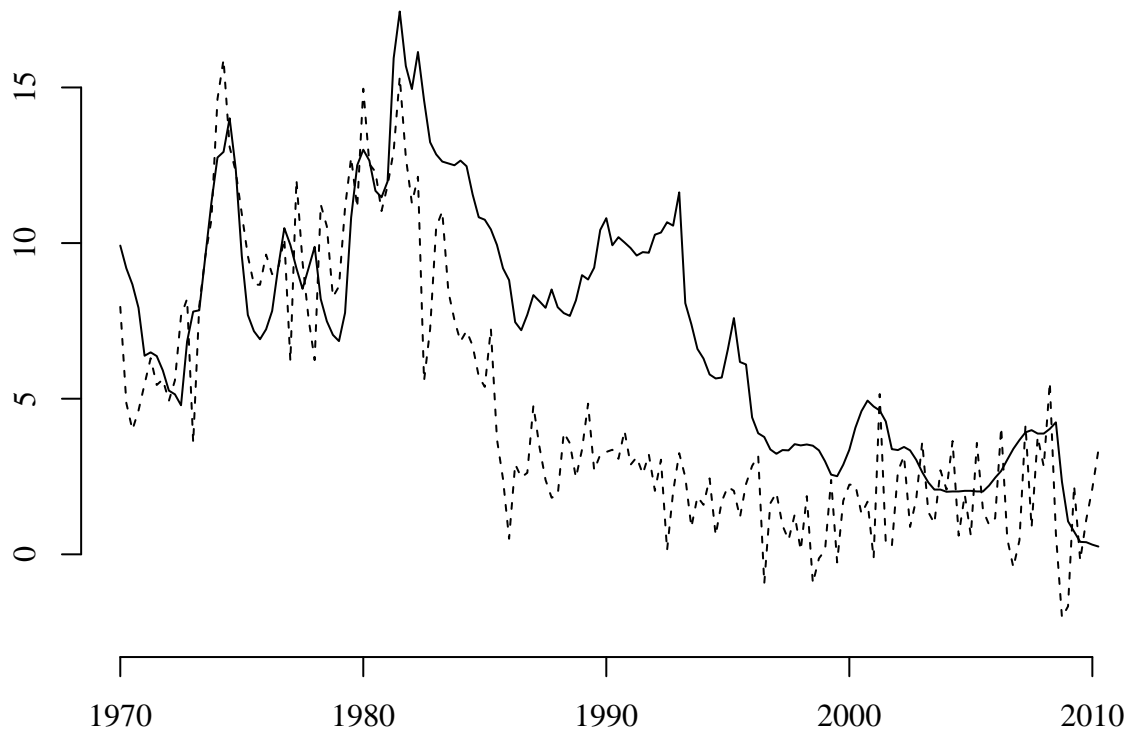


Figure 1:  
CPI inflation (dashed line) and three-month interest rate (solid line)

Concerning the inclusion of deterministic trends, we only put a constant in the cointegration part of the model since neither the inflation series nor the interest rate series display a trending pattern (see Franses, 2001, for justification of that choice). In each regime the cointegration relationship between the two variables follows one of the following three cases: i) they are not cointegrated (we denote this case by  $b=0$ ), ii) the cointegration rank is one but the cointegration space is not constrained ( $b=1$ ), or iii) the cointegration rank is one and the cointegration space is restricted to cases that are implied by Fisher's hypothesis ( $b=1F$ ).<sup>2</sup> For the lag length we consider the cases  $p = 1, 2, 3$ .

Our empirical results strongly favor Markov switching VECMs over structural break or constant coefficients VECMs. In fact, in a BMA exercise Markov switching models would receive virtually all of the weight. For the Markov switching case, there is never any evidence for more than 2 regimes. Accordingly, our empirical results focus on the Markov switching models with  $M=2$ . However, to illustrate the

---

<sup>2</sup>This case is obtained by using a prior centered tightly over the restriction (see the discussion of the prior distributions in the Appendix for more details and Jochmann, Koop, León-González and Strachan, 2011, for justification of this approach).

properties of our approach, we also present results for the models with structural breaks (even though there is little support for these models). For these, we do find evidence for three regimes and, accordingly, present results for structural break models with  $M = 2$  and 3.

For brevity's sake, we do not present any results for constant coefficient VECMs ( $M=1$ ). For such VECMs with the various combinations of  $b$  and  $p$ , we never find a log marginal likelihood higher than -533.9 which, compared with the results presented below, means constant coefficient models receive negligible support.

## Markov Switching Models

First, we look at results for the Markov switching case with two regimes ( $M=2$ ). We impose an identification restriction which specifies that the variance of the interest rate equation in the first regime is bigger than the variance in the second regime.<sup>3</sup> Table 1 gives logarithms of marginal likelihoods for models with different cointegration relationships in the two regimes and different lag lengths. The model with the highest marginal likelihood has a lag length of two and specifies that both regimes are cointegrated but the Fisher effect restriction only holds in the first regime.

Cointegration		Lag length		
$b_1$	$b_2$	$p = 1$	$p = 2$	$p = 3$
0	0	-536.4	-515.6	-515.4
0	1	-525.4	-516.1	-517.0
0	1F	-531.5	-517.1	-517.3
1	0	-531.7	-515.6	-516.7
1	1	-521.5	-515.4	-519.0
1	1F	-527.5	-517.3	-519.4
1F	0	-530.3	-514.0	-515.5
1F	1	-520.0	-513.7	-517.8
1F	1F	-525.9	-515.6	-518.3

Table 1:  
Logarithms of marginal likelihoods for the Markov switching case

For this “best model” Figure 2 plots the posterior probability that the regime where the Fisher effect holds occurs. It can be seen that this probability is very high during the 1970s and in the beginning of the 1980s. After that the probability is very low for much of the time. If this were the full story, then we would expect a structural break model to work well, with a break occurring around 1983. The

<sup>3</sup>Imposing the identification restriction means that we do not have to worry about the label-switching problem. We checked the restriction's appropriateness by examining draws from the unconstrained posterior obtained with the permutation sampler of Frühwirth-Schnatter (2001).

timing of the break is similar to that reported in Beyer, Haug and Dewalt (2009). However, there are three, relatively short, time periods where the Fisher effect seems to hold again (in the mid 1990s and at the end of the sample). This kind of behavior is more consistent with a Markov switching process than a structural break model and this is why Markov switching models perform so well in our analysis.

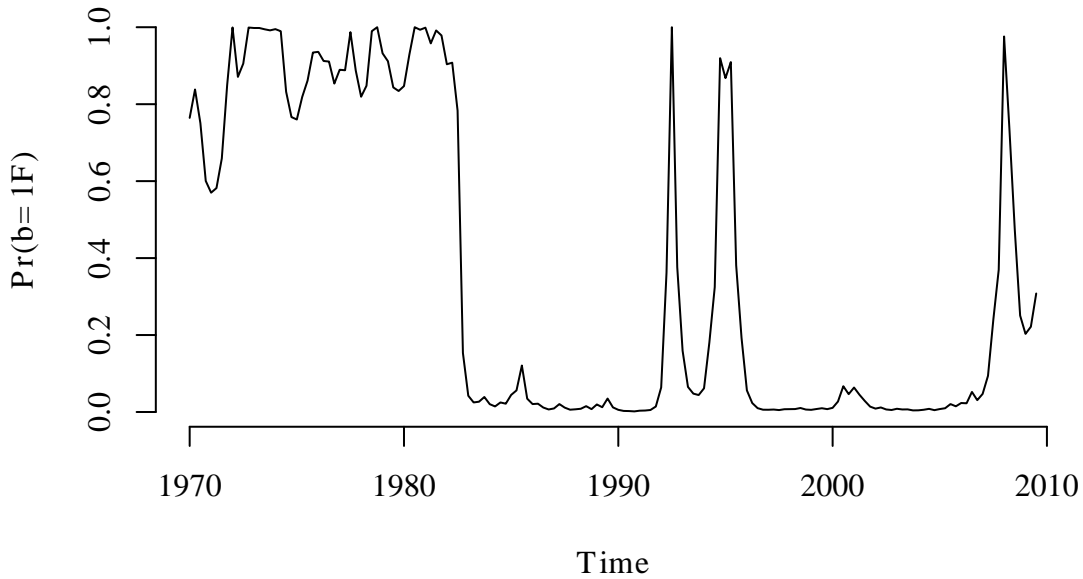


Figure 2:  
Posterior probability that the Fisher effect holds in the “best model”

The same conclusion can be drawn from Figure 3. Here, the cointegration space is normalized to be a vector  $(\tilde{\beta}_1, 1, \tilde{\beta}_3)$ . In this normalization,  $\tilde{\beta}_1$  is the normalized intercept and the Fisher hypothesis tells us that  $\tilde{\beta}_3$  should be  $-1$ . The posterior median of  $\tilde{\beta}_3$  and its 16% and 84% posterior quantiles are drawn. As expected, the posterior median of  $\tilde{\beta}_3$  is close to  $-1$  at the same times that Figure 2 says there is a high probability that the Fisher effect holds.

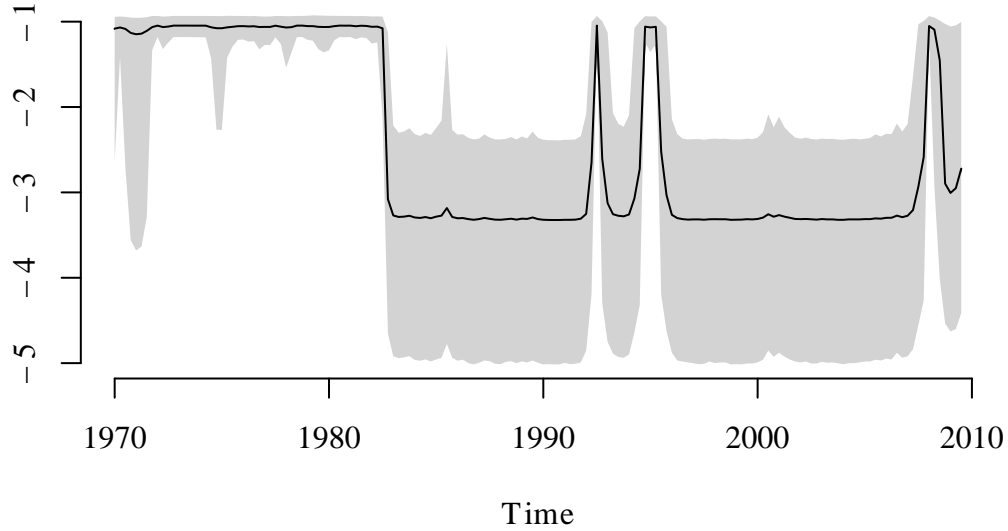


Figure 3: Posterior median and 16% and 84% posterior quantiles of  $\tilde{\beta}_3$  in the “best model”

So far, we have presented results for the single model with highest marginal likelihood. However, there are many other models whose marginal likelihoods are only slightly smaller than that of the “best model”. For example, the close second best model is the one where the first regime is cointegrated and the Fisher effect holds but with no cointegration in the second regime. Faced with such model uncertainty, the researcher may wish to do BMA. Figure 4 gives the results of a BMA exercise. It plots the posterior probabilities of the three cointegration cases at each point in time, averaged across all the models in Table 1. The story told by Figure 4 is similar to that in Figures 2 and 3. Up until 1983, in the mid 1990s and at the end of the sample the Fisher effect is supported. But elsewhere it is not. Furthermore, in the periods where the Fisher effect is not supported, there is great uncertainty over whether cointegration occurs or not.

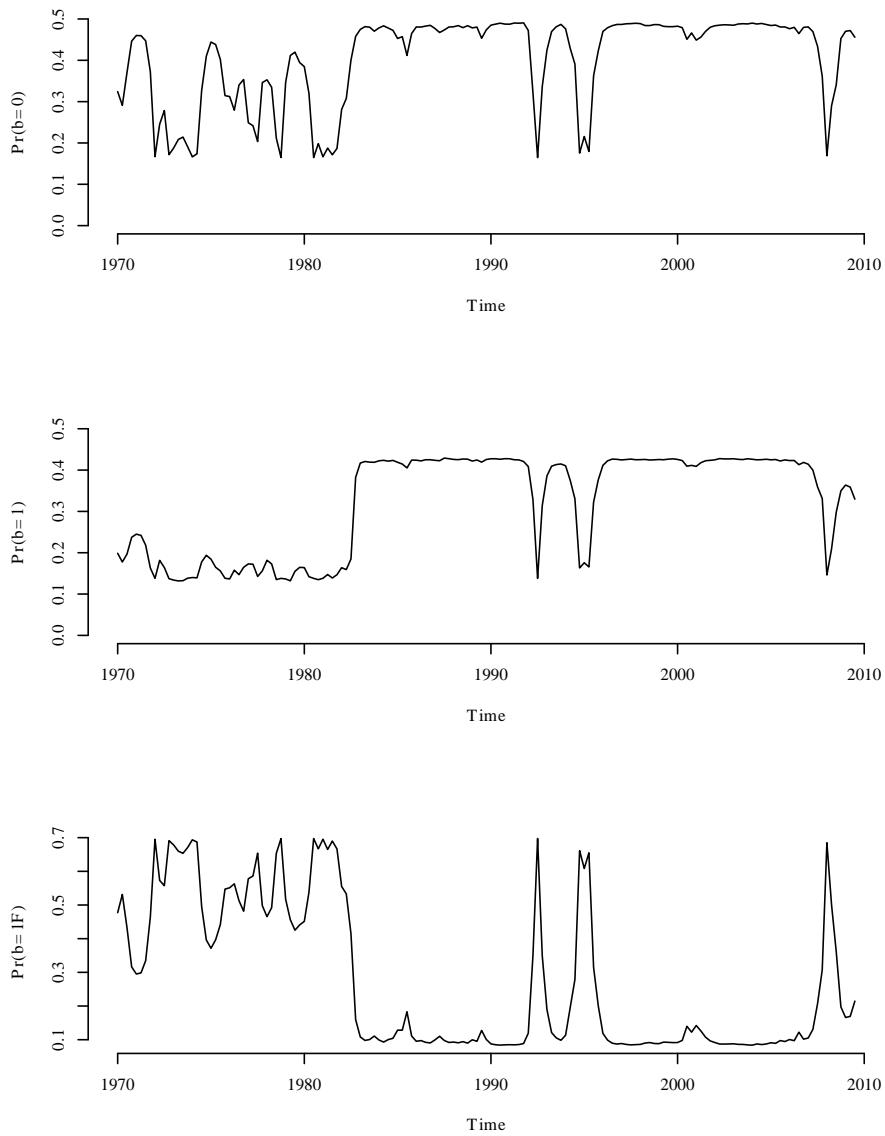


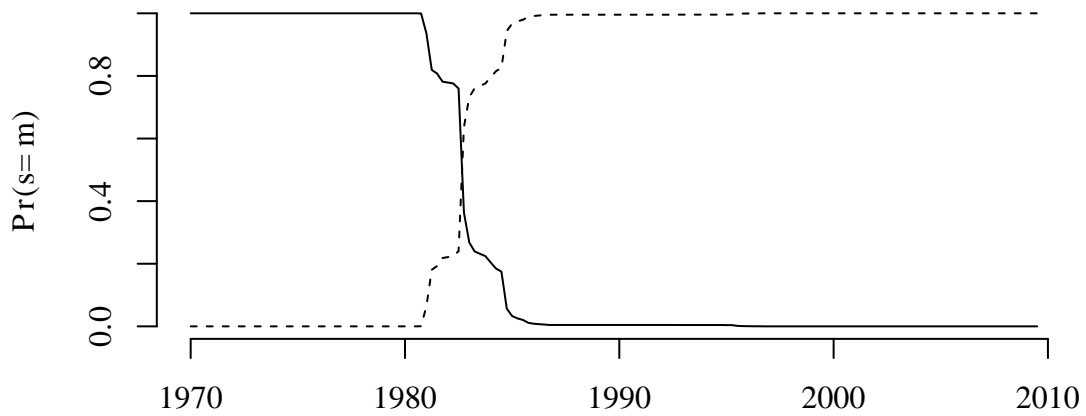
Figure 4: Posterior probabilities of the three cointegration cases averaged over all models

### Structural Break Models

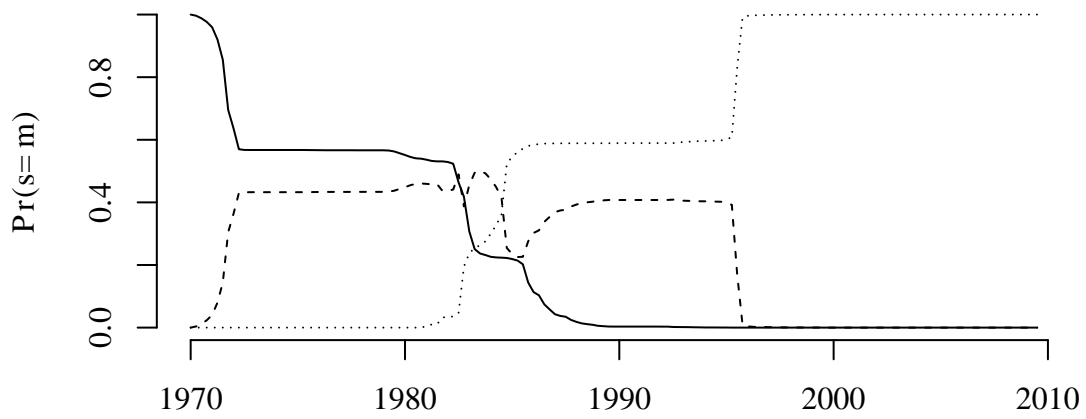
Now we discuss results for the structural break case with two and three regimes ( $M = 2, 3$ ). Table 2 gives the logarithms of marginal likelihoods for the different models. As discussed previously, these are much lower than for the Markov switching models and we include these structural break models for illustrative purposes only.

The “best model” with two regimes has a lag length of two. Both regimes are cointegrated with the Fisher effect holding in the first regime. The “best model” with three regimes also has a lag length of two. Here, the first regime is cointegrated and the Fisher effect holds, the second regime is not cointegrated and the third regime is cointegrated again but the Fisher effect does not hold. Note that conventional models of cointegration with structural breaks could not handle such a case where the cointegrating space switches between cointegrating ranks as well as switching between restricted and unrestricted cointegrating spaces. This illustration shows that such cases are empirically relevant, highlighting the importance of a modelling approach which allows for such possibilities.

Figure 5 plots the posterior probabilities for each regime to occur for these two “best models”. It can be seen that the structural breaks models are trying (poorly) to approximate the Markov switching properties of Figure 2.



a) 2 regimes, solid line:  $\Pr(s=1)$ , dashed line:  $\Pr(s=2)$



b) 3 regimes, solid line:  $\Pr(s=1)$ , dashed line:  $\Pr(s=2)$ , dotted line:  $\Pr(s=3)$

Figure 5: Posterior probabilities of regimes

		Lag length			
		$p = 1$	$p = 2$	$p = 3$	
$b_1$	$b_2$				
0	0	-546.1	-531.2	-531.1	
0	1	-530.0	-526.5	-527.7	
0	1F	-536.5	-530.2	-530.8	
1	0	-544.2	-532.3	-532.8	
1	1	-528.3	-526.9	-529.3	
1	1F	-534.5	-530.6	-532.4	
1F	0	-541.9	-530.7	-531.5	
1F	1	-526.5	-524.9	-527.9	
1F	1F	-532.4	-528.3	-530.9	
$b_1$	$b_2$	$b_3$			
0	0	0	-544.5	-528.1	-527.1
0	0	1	-530.2	-525.7	-526.1
0	0	1F	-536.7	-527.8	-527.6
0	1	0	-532.5	-528.4	-528.2
0	1	1	-528.3	-526.2	-527.5
0	1	1F	-532.9	-528.2	-528.7
0	1F	0	-538.9	-529.4	-528.5
0	1F	1	-526.6	-527.0	-527.7
0	1F	1F	-536.9	-528.8	-529.2
1	0	0	-542.5	-529.1	-529.0
1	0	1	-528.0	-526.7	-528.1
1	0	1F	-534.6	-528.8	-529.6
1	1	0	-530.8	-529.0	-530.2
1	1	1	-528.8	-527.1	-530.0
1	1	1F	-531.6	-529.0	-531.1
1	1F	0	-536.9	-530.1	-530.5
1	1F	1	-528.8	-527.8	-530.1
1	1F	1F	-532.8	-530.2	-531.5
1F	0	0	-540.4	-527.6	-528.1
1F	0	1	-525.9	-525.2	-527.2
1F	0	1F	-532.5	-527.5	-528.8
1F	1	0	-529.0	-527.3	-529.2
1F	1	1	-527.1	-525.4	-529.0
1F	1	1F	-531.0	-527.3	-530.3
1F	1F	0	-534.9	-528.5	-529.7
1F	1F	1	-526.0	-526.4	-529.1
1F	1F	1F	-531.9	-528.5	-530.9

Table 2:  
Logarithms of marginal likelihoods for the structural break case



## 5 Conclusions

This paper sets out a framework for Bayesian cointegration analysis which allows for regime-switching. We allow for both cointegrating rank and the exact cointegrating space to change when the regime changes. We consider two processes for regime change, leading to structural break and Markov switching VECMs. BMA or model selection using marginal likelihoods can be used to deal with the problems caused by the high-dimensional model space. We develop methods for Bayesian inference are developed and bridge sampling methods are used to calculate the marginal likelihoods. An empirical application involving the Fisher effect shows the usefulness of our approach.

## References

- 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.
- Ang, A. and Bekaert, G. (2002), "Regime Switches in Interest Rates," *Journal of Business and Economic Statistics*, 20, 163-182.
- Beyer, A., Haug, A. and Dewald, W. (2009), "Structural Breaks, Cointegration and the Fisher Effect," ECB working paper 1013.
- Bierens, H.J. and Martins, L.F. (2010), "Time Varying Cointegration," *Econometric Theory*, 26, 1453-1490.
- Chib, S. (1996), "Calculating Posterior Distributions and Modal Estimates in Markov Mixture Models," *Journal of Econometrics*, 75, 79-97.
- Chib, S. (1998), "Estimation and Comparison of Multiple Change-point Models," *Journal of Econometrics*, 86, 221-241.
- Cogley, T. and Sargent, T. (2005), "Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S," *Review of Economic Dynamics*, 8, 262-302.
- Franses, P.H. (2001), "How to Deal with Intercept and Trend in Practical Cointegration Analysis," *Applied Economics*, 33, 577-579.
- Frühwirth-Schnatter, S. (2001), "MCMC Estimation of Classical and Dynamic Switching and Mixture Models," *Journal of the American Statistical Association*, 96, 194-209.
- Frühwirth-Schnatter, S. (2004), "Estimating Marginal Likelihoods for Mixture and Markov Switching Models Using Bridge Sampling Techniques," *Econometrics Journal*, 7, 143-167.
- Garratt, A., Lee, K., Pesaran, M.H. and Shin, Y. (2003), "A Long-run Structural Macroeconometric Model of the UK Economy," *Economic Journal*, 113, 412-455.
- Gelman, A. and Meng, X. (1998), "Simulating Normalizing Constants: From Importance Sampling to Bridge Sampling to Path Sampling," *Statistical Science*, 13, 163-185.
- Geweke, J. (1996), "Bayesian Reduced Rank Regression in Econometrics," *Journal of Econometrics*, 75, 121-146.
- Jochmann, M., Koop, G., León-González, R. and Strachan, R. (2011), "Stochastic Search Variable Selection in Vector Error Correction Models with an Application to a Model of the UK Macroeconomy," forthcoming in *Journal of Applied Econometrics*.
- Johansen, S. (1995), *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford: Oxford University Press.
- Koop, G., León-González, R. and Strachan, R. (2008a), "Bayesian Inference in the Time Varying Cointegration Model," Rimini Centre for Economic Analysis Working Paper 23-08.
- Koop G., León-González, R. and Strachan R. (2008b), "Bayesian Inference in a Cointegrating Panel Data Model," *Advances in Econometrics*, 23, 433-469.

Koop, G., León-González, R. and Strachan, R. (2010), "Efficient Posterior Simulation for Cointegrated Models with Priors on the Cointegration Space," *Econometric Reviews*, 29, 224-242.

Koop, G., Potter, S. and Strachan, R. (2008), "Re-examining the Consumption-Wealth Relationship: The Role of Model Uncertainty," *Journal of Money, Credit and Banking*, 40, 341-367.

Lettau, M. and Ludvigson, S. (2004), "Understanding Trend and Cycle in Asset Values: Reevaluating the Wealth Effect on Consumption," *American Economic Review*, 94, 276-299.

Martin, G. (2000), "US Deficit Sustainability: A New Approach Based on Multiple Endogenous Breaks," *Journal of Applied Econometrics*, 15, 83-105.

Michael, P., Nobay, A. and Peel, D. (1997), "Transactions Costs and Nonlinear Adjustment in Real Exchange Rates: An Empirical Investigation," *Journal of Political Economy*, 105, 862-879.

Paap, R., and van Dijk, H. (2003), "Bayes Estimates of Markov Trends in Possibly Cointegrated Series: An Application to U.S. Consumption and Income," *Journal of Business Economics and Statistics*, 21, 547-563.

Park, J., and Hahn, H. (1999), "Cointegrating Regressions with Time Varying Coefficients," *Econometric Theory*, 15, 664-703.

Primiceri, G. (2005), "Time Varying Structural Vector Autoregressions and Monetary Policy," *Review of Economic Studies*, 72, 821-852.

Quintos, C.E. (1997), "Stability Tests in Error Correction Models," *Journal of Econometrics*, 82, 289-315.

Saikkonen, P. and Choi, I. (2004), "Cointegrating Smooth Transition Regressions," *Econometric Theory*, 20, 301-340.

Sims, C. and Zha, T. (2006), "Were There Regime Switches in Macroeconomic Policy?" *American Economic Review*, 96, 54-81.

Stock, J. and Watson, M. (1996), "Evidence on Structural Instability in Macroeconomic Time Series Relations," *Journal of Business and Economic Statistics*, 14, 11-30.

Strachan, R. (2003), "Valid Bayesian Estimation of the Cointegrating Error Correction Model," *Journal of Business and Economic Statistics*, 21, 185-195.

Strachan, R. and Inder, B. (2004), "Bayesian Analysis of the Error Correction Model," *Journal of Econometrics*, 123, 307-325.

Sugita, K. (2006), "Bayesian Analysis of Dynamic Multivariate Models with Multiple Structural Breaks," Discussion paper No 2006-14, Graduate School of Economics, Hitotsubashi University.

## Appendix

This appendix describes how the model can be written in matrix notation and introduces a reparametrization that we use. Furthermore, the prior distributions and the algorithm for posterior simulation are discussed. Finally, we show how marginal likelihoods are computed with the bridge sampler.

### Model in Matrix Form

The model in each regime can be written in the following way:

$$Y_i = X_i\beta_i\alpha_i + W_i\Gamma_i + E_i, \quad i = 1, \dots, M, \quad (\text{A1})$$

where  $Y_i$ ,  $X_i$  and  $W_i$  collect the observations belonging to regime  $i$ .  $Y_i$  is  $T_i \times n$  with the rows given by  $\Delta y'_{\tilde{t}_i}$ ,  $X_i$  is  $T_i \times n$  with the rows given by  $y'_{\tilde{t}_i-1}$  and  $W_i$  is  $T_i \times [n(p-1)]$  with the rows given by  $(\Delta y'_{\tilde{t}_i-1}, \dots, \Delta y'_{\tilde{t}_i-p+1})$ , where  $\tilde{t}_i$  denotes the  $t^{\text{th}}$  observation in regime  $i$  and  $T_i$  gives the number of observations in regime  $i$ .  $E_i$  is  $T_i \times n$  with  $\text{vec}(E_i) \sim \text{N}(0, \Sigma \otimes I)$ .  $\alpha_i$  is  $r \times n$ ,  $\beta_i$  is  $n \times r$  and  $\Gamma_i$  is  $[n(p-1)] \times n$ .

Following Koop, León-González and Strachan (2010) we next introduce non-identified  $r \times r$  symmetric positive definite matrices  $D_i$  and define  $\alpha_i^* = D_i^{-1}\alpha_i$  and  $\beta_i^* = \beta_i D_i$  where  $\alpha_i^*$  is  $r \times n$  and  $\beta_i^*$  is  $n \times r$ . Since  $\alpha_i$  and  $\beta_i$  always occur in product form and  $\beta_i\alpha_i = \beta_i D_i D_i^{-1}\alpha_i = \beta_i^*\alpha_i^*$  this does not affect the model which now can be written as:

$$Y_i = X_i\beta_i^*\alpha_i^* + W_i\Gamma_i + E_i, \quad i = 1, \dots, M. \quad (\text{A2})$$

### Prior Distributions

1.  $\{\alpha_i^*\}$ : In regimes with no cointegration the  $\{\alpha_i^*\}$  are set to zero. For cointegrated regimes, we assume the following shrinkage prior:

$$a_i^* \equiv \text{vec}(\alpha_i^*) \sim \text{N}\left(0, \underline{\eta}_\alpha^{-1} I_{rn}\right), \quad i = 1, \dots, M. \quad (\text{A3})$$

In our application we set  $\underline{\eta}_\alpha = 10$ .

2.  $\{\beta_i^*\}$ : In regimes with no cointegration the  $\{\beta_i^*\}$  are set to zero. Our prior in the case of cointegration but with the Fisher effect not imposed is:

$$b_i^* \equiv \text{vec}(\beta_i^*) \sim \text{N}\left(0, \underline{\tau}^{\frac{1}{rn}} I_{rn}\right), \quad i = 1, \dots, M. \quad (\text{A4})$$

The prior we use for restricting the cointegration space in accordance with the Fisher effect is:

$$b_i^* \equiv \text{vec}(\beta_i^*) \sim \text{N}\left[0, I_r \otimes (HH' + \underline{\tau} H_\perp H'_\perp)\right], \quad i = 1, \dots, M. \quad (\text{A5})$$

The matrix  $H$  that imposes the Fisher effect is given by  $H = \tilde{H}(\tilde{H}'\tilde{H})^{-\frac{1}{2}}$  with  $\tilde{H} = (1, -1)'$ . The transformation makes  $H$  orthogonal. The dispersion of the prior is controlled by the scalar  $\underline{\tau} \in [0, 1]$  with  $\underline{\tau} = 0$  dogmatically imposing the restrictions expressed by  $H$  on the cointegrating space and  $\underline{\tau} = 1$  being noninformative. As discussed in Jochmann, Koop, León-González and Strachan (2011), in some cases there are some problems with setting  $\underline{\tau} = 0$  and accordingly we recommend setting  $\underline{\tau}$  to a small value in order to “almost impose” a restriction. In our application we choose  $\underline{\tau} = 0.05$ . Note that  $\underline{\tau}$  appears in (A4) so that both priors have the same scale.

3.  $\{\Gamma_i\}$ : We assume the following shrinkage prior for  $\{\Gamma_i\}$ :

$$c_i^* \equiv \text{vec}(\Gamma_i) \sim N\left(0, \underline{\eta}_\Gamma^{-1} I_{n^2(p-1)}\right), \quad i = 1, \dots, M. \quad (\text{A6})$$

For our application we choose  $\underline{\eta}_\Gamma = 10$ .

4.  $\{\Sigma_i\}$ : We use an inverted Wishart prior:

$$\Sigma_i \sim \text{InvWishart}(\underline{\nu}, \underline{S}), \quad i = 1, \dots, M. \quad (\text{A7})$$

In the application we set  $\underline{\nu} = 13$  and  $\underline{S} = 10I_2$ . It follows that  $E(\Sigma_i) = I_2$ .

5.  $\{\xi_{ij}\}$ : In the case of structural breaks we use the following Beta prior distributions:

$$\xi_{ii} \sim \text{Beta}(\underline{a}, \underline{b}), \quad i = 1, \dots, M - 1. \quad (\text{A8})$$

In our application we set  $\underline{a} = 10$  and  $\underline{b} = 0.1$ . For the Markov switching models we assume

$$\xi_i \sim \text{Dirichlet}(\underline{c}_{i1}, \dots, \underline{c}_{iM}), \quad i = 1, \dots, M. \quad (\text{A9})$$

In our application we choose  $\underline{c}_{ij} = 10$  if  $i = j$  and  $\underline{c}_{ij} = 1$  otherwise.

## Posterior Simulation

We sample from the posterior distribution with a Gibbs sampler. Given initial conditions, the data, and in each block the other parameters, the algorithm comprises the following steps:

1. *Structural break model:*  
Draw  $S^T$  using Chib’s (1998) algorithm.
- Markov switching model:*  
Draw  $S^T$  using Chib’s (1996) algorithm.

2. *Structural break model:*

Draw  $\xi_i$  from Beta  $[\underline{a} + N_{ii}(S^T), \underline{b} + 1]$  for  $i = 1, \dots, M - 1$ , where  $N_{ii}(S^T)$  is the number of one-step transitions from state  $i$  to state  $i$  in the sequence  $S^T$ .

*Markov switching model:*

Draw  $\xi_i$  from Dirichlet  $[\underline{c}_{i1} + N_{i1}(S^T), \dots, \underline{c}_{iM} + N_{iM}(S^T)]$  for  $i = 1, \dots, M$ , where  $N_{ij}(S^T)$  is the number of one-step transitions from state  $i$  to state  $j$  in the sequence  $S^T$ .

3. Draw  $a_i^*$  for  $i = 1, \dots, M$ : If regime  $i$  is not cointegrated, set  $a_i^*$  equal to zero. Otherwise, draw  $a_i^*$  from  $N(\bar{a}_i, \bar{A}_i)$  with

$$\bar{A}_i = \left[ (\Sigma_i^{-1} \otimes \beta_i^{*'} X_i' X_i \beta_i^*) + \underline{\eta}_\alpha I_{rn} \right]^{-1} \quad (\text{A10})$$

and

$$\bar{a}_i = \bar{A}_i (\Sigma_i^{-1} \otimes \beta_i^{*'} X_i') \text{vec}(Y_i - W_i \Gamma_i). \quad (\text{A11})$$

4. Draw  $b_i^*$  for  $i = 1, \dots, M$ : If regime  $i$  is not cointegrated, set  $b_i^*$  equal to zero. Otherwise, draw  $b_i^*$  from  $N(\bar{b}_i, \bar{B}_i)$  with

$$\bar{B}_i = \left[ (\alpha_i^* \Sigma_i^{-1} \alpha_i^{*'}) \otimes (X_i' X_i) + P^{-1} \right]^{-1} \quad (\text{A12})$$

and

$$\bar{b}_i = \bar{B}_i (\alpha_i^* \Sigma_i^{-1} \otimes X_i') \text{vec}(Y_i - W_i \Gamma_i). \quad (\text{A13})$$

In the case of no restriction on the cointegration space  $P$  is defined as:

$$P = \underline{\tau} \frac{1}{rn} I_{rn}, \quad (\text{A14})$$

with the Fisher effect imposed it is defined as:

$$P = I_r \otimes (HH' + \underline{\tau} H_\perp H_\perp'). \quad (\text{A15})$$

5. Draw  $c_i^*$  for  $i = 1, \dots, M$  from  $N(\bar{c}_i, \bar{C}_i)$  with

$$\bar{C}_i = \left[ (\Sigma_i^{-1} \otimes W_i' W_i) + \eta_\Gamma I_{n^2(p-1)} \right]^{-1} \quad (\text{A16})$$

and

$$\bar{c}_i = \bar{C}_i (\Sigma_i^{-1} \otimes W_i') \text{vec}(Y_i - X_i \beta_i^* \alpha_i^*). \quad (\text{A17})$$

6. Draw  $\Sigma_i$  for  $i = 1, \dots, M$  from  $\text{InvWishart}(\bar{\nu}_i, \bar{S}_i)$  with

$$\bar{\nu}_i = \underline{\nu} + T_i \quad (\text{A18})$$

and

$$\bar{S}_i = \underline{S} + (Y_i - X_i \beta_i^* \alpha_i^* - W_i \Gamma_i)' (Y_i - X_i \beta_i^* \alpha_i^* - W_i \Gamma_i). \quad (\text{A19})$$

## Marginal Likelihood Calculation using the Bridge Sampler

Given  $B$  MCMC sampler draws  $\{\theta_K^{(b)}\}$ ,  $b = 1, \dots, B$ , from the posterior distribution  $p(\theta_K|y, \mathcal{M}_K)$ , the bridge sampler consists of the following steps:

1. *Simulation.* Construct the unsupervised importance density  $q(\theta_K)$  discussed in Frühwirth-Schnatter (2004, Section 3.4). Draw  $L$  draws  $\{\tilde{\theta}_K^{(l)}\}$ ,  $l = 1, \dots, L$ , from this importance density.
2. *Evaluation.* Calculate both the non-normalized posterior  $p^*(\theta_K|y, \mathcal{M}_K)$  and the importance density  $q(\theta_K)$  at the draws both from the posterior and the importance density.
3. *Iteration.* Get a starting value for the estimate of the marginal likelihood  $\hat{p}_0$  (for example the importance sampling estimator). Run the following recursion until convergence has been achieved:

$$\hat{p}_t = \frac{\frac{1}{L} \sum_{l=1}^L \frac{p^*(\tilde{\theta}_K^{(l)}|y, \mathcal{M}_K)}{Lq(\tilde{\theta}_K^{(l)}) + Bp^*(\tilde{\theta}_K^{(l)}|y, \mathcal{M}_K) / \hat{p}_{t-1}}}{\frac{1}{B} \sum_{b=1}^B \frac{q(\theta_K^{(b)})}{Lq(\theta_K^{(b)}) + Bp^*(\theta_K^{(b)}|y, \mathcal{M}_K) / \hat{p}_{t-1}}}. \quad (\text{A20})$$