Ca' Foscari
University
of Venice

**Department
of Economics**

Working Paper

**Roberto Casarin, Marco
Tronzano, and Domenico
Sartore**

**Bayesian Markov Switching
Stochastic Correlation
Models**

ISSN: 1827-3580
No. 11/WP/2013

Bayesian Markov Switching Stochastic Correlation Models

Roberto Casarin

University Ca' Foscari of Venice

Marco Tronzano

University of Genova

Domenico Sartore

University Ca' Foscari of Venice

2 June 2013

Abstract

This paper builds on Asai and McAleer (2009) and develops a new multivariate Dynamic Conditional Correlation (DCC) model where the parameters of the correlation dynamics and those of the log-volatility process are driven by two latent Markov chains. We outline a suitable Bayesian inference procedure, based on sequential MCMC estimation algorithms, and discuss some preliminary results on simulated data. We then apply the model to three major cross rates against the US Dollar (Euro, Yen, Pound), using high-frequency data since the beginning of the European Monetary Union. Estimated volatility paths reveal significant increases since mid-2007, documenting the destabilizing effects of the US sub-prime crisis and of the European sovereign debt crisis. Moreover, we find strong evidence supporting the existence of a time-varying correlation structure. Correlation paths display frequent shifts along the whole sample, both in low and in high volatility phases, pointing out the existence of contagion effects closely in line with the mechanisms outlined in the recent contagion literature (Forbes and Rigobon (2002) and Corsetti et al. (2005)).

Keywords

Stochastic Correlation; Multivariate Stochastic Volatility; Markov-switching; Bayesian Inference; Monte Carlo Markov Chain.

JEL Codes

C1, C11, C15, C32, F31, G15.

Address for correspondence:

Roberto Casarin

Department of Economics
Ca' Foscari University of Venice
Cannaregio 873, Fondamenta S. Giobbe
30121 Venezia - Italy
Phone: (+39) 041 2349149
Fax: (+39) 041 2349210
e-mail: r.casarin@unive.it

This Working Paper is published under the auspices of the Department of Economics of the Ca' Foscari University of Venice. Opinions expressed herein are those of the authors and not those of the Department. The Working Paper series is designed to divulge preliminary or incomplete work, circulated to favour discussion and comments. Citation of this paper should consider its provisional character.

Bayesian Markov Switching Stochastic Correlation Models

Roberto Casarin[†] Domenico Sartore[†] Marco Tronzano[‡]

[†]University Ca' Foscari of Venice

[‡]University of Genoa

June 3, 2013

Abstract

This paper builds on Asai and McAleer (2009) and develops a new multivariate Dynamic Conditional Correlation (DCC) model where the parameters of the correlation dynamics and those of the log-volatility process are driven by two latent Markov chains. We outline a suitable Bayesian inference procedure, based on MCMC estimation algorithms, and show the effectiveness of the procedure on simulated data. We then apply the model to three major cross rates against the US Dollar (Euro, Yen, Pound), using high-frequency data since the beginning of the European Monetary Union. Estimated volatility paths reveal significant increases since mid-2007, documenting the destabilizing effects of the US sub-prime crisis and of the European sovereign debt crisis. Moreover, we find strong evidence supporting the existence of a time-varying correlation structure. Correlation paths display frequent shifts along the whole sample, both in low and in high volatility phases, pointing out the existence of contagion effects closely in line with the mechanisms outlined in the recent contagion literature (see, e.g. Forbes and Rigobon (2002) and Corsetti et al. (2005)).

KEYWORDS: Stochastic Correlation; Multivariate Stochastic Volatility; Markov-switching; Bayesian Inference; Monte Carlo Markov Chain.

AMS: 62F15; 62M10; 62P20; 91B84.

JEL : C11; C15; C32; F31; G15.

1 Introduction

Explaining and forecasting volatility co-movements in financial assets is of paramount importance in various aspects of financial management including asset pricing, optimal dynamic portfolio allocation strategies, and the computation of value-at-risk. In a more analytical perspective, moreover, accounting for the time-varying nature of

[¶]Address: Departments of Economics, University Ca' Foscari of Venice, San Giobbe 873/b, 30121 Venice, Italy. E-mail: r.casarin@unive.it

the variances and covariances of financial time series yields useful insights into some relevant issues such as contagion and volatility spillovers.

Forbes and Rigobon (2002) define contagion as a significant increase in cross-market correlation during a period of financial turmoil, whereas any continued high degree of market correlation is defined as interdependence (see also Corsetti et al. (2011) for a review of stylized facts on contagion). Drawing on this seminal paper, a large body of applied literature has investigated the existence of contagion effects on financial markets, implementing several types of multivariate dynamic conditional correlation models relying either on GARCH or on Stochastic Volatility (SV) specifications.

Explaining and forecasting the co-movements of volatility in financial assets is of paramount importance in financial management. These practices include asset pricing, optimal dynamic portfolio allocation strategies, and the computation and forecasting of value-at-risk. Moreover, accounting for the time-varying nature of the variances and covariances of financial time series may significantly improve knowledge of financial markets, thereby yielding useful insights into some relevant issues such as contagion and volatility spillovers across markets.

Multivariate GARCH models, pioneered by Kraft and Engle (1982), Bollerslev et al. (1988) and Engle and Kroner (1995) have typically raised serious technical difficulties related to the "curse of dimensionality" and the positive definiteness of the covariance matrix. The benchmark reference in this literature is represented by Bollerslev (1990) Constant Conditional Correlation (CCC) model, while Engle (2002) introduces a basic extension of this framework preserving the same decomposition of the covariance matrix, but assuming a GARCH-type structure for the Dynamic Conditional Correlations (DCC). Bauwens et al. (2006) provide a survey of some recent advances in the multivariate GARCH literature. Many refinements of the DCC model have recently been proposed. Among others, Cappiello et al. (2006), introduce an additional term in the DCC equation to account for asymmetric effects. Billio and Caporin (2009), use a BEKK structure for the conditional correlations. Interesting extensions of the DCC model makes use of a Markov-switching (MS) process to capture sudden changes in some parameter of the model. Pelletier (2006) extends the CCC model by assuming MS dynamics for the correlation matrix, and Billio and Caporin (2005) modify the DCC model by introducing an MS process in the unconditional correlation matrix and for the DCC parameters. Galeano and Ausín (2010) propose a DCC model with Gaussian mixture distribution for the standardized innovations. Finally Creal et al. (2011) introduce a new class of DCC models with a correlations dynamics based on generalized autoregressive score.

SV models offer increased flexibility over GARCH-type specifications since they assume separate innovation processes for the conditional mean and the conditional variance of the observables. Note that the use of variance-specific disturbances comes to the use of latent variables and this call for the use of suitable inference procedures such as simulation based Bayesian inference. Seminal contributions in this area date back to Taylor (1986) and Taylor (1994). Jacquier et al. (1994) propose a Bayesian inference procedure for SV models. So et al. (1998) extend the basic Bayesian

SV model, assuming that the log-volatility has a MS autoregressive dynamics, while So et al. (2002) assume a threshold autoregressive process for the log-volatility. Jacquier et al. (2004) propose a univariate SV model allowing for fat-tailed innovations in the measurement equation and for a leverage effect through the correlation between the mean and variance innovations.

In order to capture dependencies and spillover effects between the volatility of different variables, the modelling results obtained in the univariate case have been successfully extended to a multivariate setting in Harvey et al. (1994). Compared to the multivariate GARCH literature the Multivariate Stochastic Volatility (MSV) literature is much more limited mainly due to the difficulties in parameter and latent volatility matrix estimation issues. Among the existing innovative specifications see e.g. Aguilar and West (2000) and Chib et al. (2006), which consider factor MSV models for high-dimensional inference problems and Chan et al. (2005) which introduce MSV with leverage. Liesenfeld and Richard (2003) discuss efficient inference methods for MSV models. See also Asai et al. (2006) for a comprehensive survey of MSV models and Yu and Meyer (2006) for a review of general MSV models including Granger causality in volatility, time-varying correlations and heavy-tailed error distribution. The ability of these models to provide accurate density forecast is explored in Clark (2011) and Clark and Ravazzolo (2012).

All MSV models discussed above assume Constant Correlations (CC) among the variables, an assumption which appears at odds with the empirical evidence, at least for financial time series (see, e.g., Engle (2002)). Recent contributions in this field have focused on new classes of models accounting for a time-varying and stochastic correlation structure. In the literature there are two approaches to introduce *Dynamic Correlation* (DC) MSV models. The former assumes that the covariance matrix is a function of a Wishart process, which allows both the variances and the implicit correlation structure to evolve stochastically over time. Gouriéroux et al. (2004) introduce Wishart autoregressive processes for the covariance matrix, while Philipov and Glickman (2006) and Asai and McAleer (2009) assume that the inverse covariance matrix follows a Wishart distribution conditionally on the past information. In the latter approach the vector of volatilities and the correlation matrix have its own specific dynamics, thereby allowing for a more flexible representation of the dependence between series. This kind of model has been studied in Asai and McAleer (2005). Yu and Meyer (2006) introduce a bivariate SV model with stochastic correlation based on the Fisher transform (see also Amisano and Casarin (2007) for a discussion of the alternative ways of modelling stochastic correlation in a bivariate context). Asai and McAleer (2009) recently proposed new DC-MSV models and provided a comparison in terms of the number of parameters with existing MSV models.

The issue of discontinuity in dynamic correlation models would seem to remain a crucial issue in the econometric literature. As regards the threshold models, So and Choi (2008) extended the threshold SV framework of So et al. (2002) to the multivariate SV context, but did not model directly the asymmetric dynamics in the conditional correlation structure. To the best of our knowledge the issue of extending the Markov-switching approach of So et al. (1998) to stochastic correlation models has

not yet been fully explored. The aim of this paper is to fill this gap and to include the discontinuous dynamics, such as the Markov-switching dynamics, into DC-MSV models. In this paper we focus on the models proposed by Asai and McAleer (2009), where correlation and volatility have a independent dynamics. See Casarin and Sartore (2007) for a MS extension of the Philipov and Glickman (2006) and Gouriéroux et al. (2004) models.

The present paper develops new multivariate SV models aiming to fill this gap in the literature, namely to include a discontinuous dynamics, driven by a Markov-switching process, in the correlation structure. After outlining a suitable Bayesian inference procedure and discussing some results on simulated data, we apply this model to investigate the existence of financial contagion between three major exchange rates against the US Dollar since the beginning of the European Monetary Union. The main contribution to the existing literature is twofold. First, from the econometric viewpoint, we allow the parameters of the correlation dynamics to depend on a latent Markov chain, thus extending the MS correlation model of Pelletier (2006) where correlations are constant within a regime. Moreover, since our framework accounts for shifts in the correlation dynamics, we improve on the continuous autoregressive correlation models discussed in Asai and McAleer (2009). Another contribution of this paper is to propose an inference procedure, based Markov-chain Monte Carlo, for the joint estimation of the mean parameters and the variance and correlation parameters and latent processes. See Casarin and Sartore (2007) for alternative inference procedure for MS Wishart correlation processes based on sequential Monte Carlo algorithms. A latter original feature of this research is related to the applied literature on contagion on foreign exchange markets. Most research exploring contagion on foreign exchange markets focuses on Asian currencies (Dungey et al. (2004), Horen et al. (2006), Tai (2007)). Existing studies on major cross rates against the US Dollar, on the other hand, use simple DCC models and do not account for potential contagion effects on currency markets arising from the recent US financial crisis (2007-2008) and the European sovereign debt crisis (2009-2011) (e.g., see Pérez-Rodríguez (2006)). In this perspective, the present paper contributes to the literature focusing on some key exchange rates against the US Dollar and highlighting significant changes in their volatility and correlation structures related to the latest financial turmoils.

The structure of the paper is as follows. Section 2 introduces basic Markov-switching constant correlation models. Section 3 discusses different Markov-switching stochastic correlation models with a special attention to a flexible specification of the correlation dynamics and a parsimonious parameterization of the correlation and volatility processes. Section 4 details a Bayesian approach to inference based on an efficient Markov-chain Monte Carlo algorithm. Section 5 presents some simulation results. Section 6 provides a contagion analysis for three exchange rate series against the US dollar along the recent floating period.

2 Constant Correlation Markov-switching Models

In the following, boldface quantities and capital letters denote vectors and matrices, respectively, $\text{vech}(A)$ the operator which stacks the lower part of a matrix A into a vector, $\text{vecd}(A)$ the operator which creates a vector from the main diagonal elements of a matrix A , and $\text{diag}(\mathbf{x})$ the operator which defines a diagonal matrix, with the elements of the vector $\mathbf{x}$ on the main diagonal.

We present some natural extensions to the multivariate context of the univariate Markov-switching SV model given in So et al. (1998). It should be noted that the proposed models can be regarded as the MS counterpart of the MSV model in Harvey et al. (1994) which, in turn, has the CC-MSV given in Asai and McAleer (2009) as a special case.

2.1 Basic Markov-switching Constant Correlation Models

Let $\mathbf{y}_t = (y_{1t}, \dots, y_{mt})' \in \mathbb{R}^m$ be a vector-valued time series, representing the log-differences in the spot exchange rates, $\mathbf{h}_t = (h_{1t}, \dots, h_{mt})' \in \mathbb{R}^m$ the log-volatility process, $\Sigma_t \in \mathbb{R}^m \times \mathbb{R}^m$ the time-varying covariance matrix, and $s_{1,t} \in \{0, \dots, K_1 - 1\} \subset \mathbb{N}$ a time-homogeneous first-order Markov chain with K_1 states. In the applications, we will set $K_1 = 2$ in order $s_{1,t}$ to capture the change in the model parameters due to the financial crisis from the second half of 2007.

The basic MS stochastic volatility model (MS-CC-MSV) is given as

$$\mathbf{y}_t = \mathbf{a}_{s_{1,t}} + A_{s_{1,t}} \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathcal{N}_m(\mathbf{0}, \Sigma_t) \quad (1)$$

$$\mathbf{h}_t = \mathbf{b}_{s_{1,t}} + B_{s_{1,t}} \mathbf{h}_{t-1} + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim \mathcal{N}_m(\mathbf{0}, \Sigma_\eta) \quad (2)$$

with $\boldsymbol{\varepsilon}_t \perp \boldsymbol{\eta}_s \forall s, t$, and $\mathcal{N}_m(\boldsymbol{\mu}, \Sigma)$ representing the m -variate normal distribution, with mean $\boldsymbol{\mu}$ and covariance matrix Σ . The model can be further extended to account for the leverage effect, but this will be not considered here.

A suitable specification of the dynamics of the time-varying parameters will simplify the inference procedure easier. We follow So et al. (2002) and So and Choi (2008), and consider the following reparameterization:

$$\mathbf{a}_{s_{1,t}} = \mathbf{a}_{00} + \mathbf{a}_{01} s_{1,t}, \quad A_{s_{1,t}} = A_{10} + A_{11} s_{1,t} \quad (3)$$

$$\mathbf{b}_{s_{1,t}} = \mathbf{b}_{00} + \mathbf{b}_{01} s_{1,t}, \quad B_{s_{1,t}} = B_{10} + B_{11} s_{1,t} \quad (4)$$

where $\mathbf{a}_{00}$, $\mathbf{a}_{01}$, A_{10} , A_{11} , $\mathbf{b}_{00}$, $\mathbf{b}_{01}$, B_{10} and B_{11} are parameters to be estimated.

The probability law governing $s_{1,t}$ is defined by

$$s_{1,t} \sim \mathbb{P}(s_{1,t} = j | s_{1,t-1} = i) = p_{1,ij} \quad (5)$$

with $p_{1,ij}$ the element on row i and column j of the transition matrix, denoted by P_1 . Note that the CC-MSV model in Asai and McAleer (2009) corresponds to the case $\Sigma_\eta = \text{diag}\{\sigma_{1\eta}^2, \dots, \sigma_{m\eta}^2\}$, $B_{s_{1,t}} = \text{diag}\{b_{1s_{1,t}}, \dots, b_{ms_{1,t}}\}$ and $K_1 = 1$.

As regards the conditional covariance matrix Σ_t , we use the Bollerslev (1990) decomposition into conditional standard deviations and constant correlations:

$$\Sigma_t = \Lambda_t \Omega \Lambda_t', \quad (6)$$

with $\Lambda_t = \text{diag}\{\exp\{h_{1t}/2\}, \dots, \exp\{h_{kt}/2\}\}$, a diagonal matrix with the volatilities on the main diagonal and Ω the correlation matrix, defined as

$$\Omega = \begin{pmatrix} 1 & \rho_{12} & \dots & \rho_{1m} \\ \rho_{21} & 1 & \dots & \rho_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{m1} & \rho_{m2} & \dots & 1 \end{pmatrix}.$$

Note that in this model there are $(2m^2 + 3m)K_1 + K_1(K_1 - 1)$ parameters to estimate the dynamics of the observable variable, volatility and switching regimes, and $(m^2 - m)/2$ parameters in the constant correlation matrix.

2.2 Heavy-tailed model

Time series models for financial variables should account for some important departures from normality, such as excess of kurtosis and asymmetry. In order to model the excess of kurtosis, we assume that the innovations for observable variables are in the family of the Student-t distribution, as suggested for univariate ARCH models (see Bollerslev (1987)) and for MSV models (see Harvey et al. (1994)). Moreover, we choose the skewed Student-t distribution which accounts for different degrees of kurtosis for the left and right tails of the distribution (see Aas and Haff (2006) for a discussion of the alternative set of distributions for modelling skewed and heavy-tailed data).

In the literature, there exist many definitions of multivariate skewed Student-t distributions. One of the first is due to Azzalini and Dalla Valle (1996). Their results have been then generalised by Branco and Dey (2001), which introduce the class of skewed elliptical distributions. Another generalisation of the class of elliptical skewed is due to Sahu et al. (2003). In this paper we adopt a definition of multivariate skewness based on the linear transformation of univariate skewed Student-t distribution. This construction of a multidimensional distribution has been first proposed in Bauwens and Laurent (2002). In particular, we apply the constructive method due to Ferreira and Steel (2007) which has many advantages: the simulation from the distribution is simple, the existence of the moments is guaranteed by the existence of the moments of the underlying univariate distributions, and the resulting multivariate distribution allows for different magnitudes and directions of kurtosis and skewness.

As an alternative to the Gaussian MSV model, one may consider the heavy-tailed MSV model (HTMSV), which results from replacing equation (1) by the following:

$$\mathbf{y}_t = \mathbf{a}_{s_{1,t}} + A_{s_{1,t}} \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathcal{SkT}_m(\mathbf{0}, \Sigma_t, \boldsymbol{\xi}, \gamma) \quad (7)$$

with $\mathcal{SkT}_m(\mathbf{0}, \Sigma_t, \boldsymbol{\xi}, \boldsymbol{\gamma})$, the m -variate skewed Student-t distribution with null location parameter, scale matrix Σ_t , degrees of freedom parameter $\boldsymbol{\xi} = (\xi_1, \dots, \xi_m)' \in (2, +\infty)^m$, and skewness parameter $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_m)' \in \mathbb{R}_+^m$. Its density function is

$$f(\mathbf{y}; \boldsymbol{\xi}, \boldsymbol{\gamma}) = \prod_{j=1}^m g(y_j; \xi_j, \gamma_j), \quad (8)$$

where

$$g(y; \xi, \gamma) = \frac{2}{\gamma + \frac{1}{\gamma}} \left\{ \varphi_\xi \left(\frac{y}{\gamma} \right) \mathbb{I}_{[0, +\infty)}(y) + \varphi_\xi(y\gamma) \mathbb{I}_{(-\infty, 0)}(y) \right\}, \quad (9)$$

$$\varphi_\xi(x) = \frac{\Gamma((\xi + 1)/2)}{\Gamma(\xi/2)(\pi(\xi - 2))^{1/2}} \left(1 + \frac{x^2}{(\xi - 2)} \right)^{-(\xi+1)/2} \mathbb{I}_{(-\infty, +\infty)}(x). \quad (10)$$

If the skewness parameter γ is unity, then we retrieve the original symmetric density.

In the HTMSV model, there are $(2m^2 + 3m)K_1 + 2m + K_1(K_1 - 1)$ parameters for the dynamics of the observable variable, volatility and switching regimes, and $(m^2 - m)/2$ parameters in the constant correlation matrix.

3 Markov Switching Dynamic Correlation Models

We extend the basic MS-CC-MSV models given in the previous section by considering time-varying correlations. In order to capture the stochastic dependence structure in the data, we follow Asai and McAleer (2009) and consider the Bollerslev (1990) decomposition of the covariance matrix:

$$\Sigma_t = \Lambda_t \Omega_t \Lambda_t' \quad (11)$$

and a time-varying correlation matrix

$$\Omega_t = \tilde{Q}_t^{-1} Q_t \tilde{Q}_t^{-1} \quad (12)$$

with $\tilde{Q}_t = (\text{diag}\{\text{vecd } Q_t\})^{1/2}$.

In the following, we allow the parameters of the correlation dynamics to depend on a latent Markov chain. This approach can be considered as an extension of the Markov-switching correlation model in Pelletier (2006), where the correlations are constant within a regime. Note, moreover, that the proposed models account for shifts in the correlation dynamics, thereby extending the continuous autoregressive correlation models of Asai and McAleer (2009).

3.1 Homogenous Transition MS Models

The first MS dynamic correlation model, which will be referred as MS-DC-MSV1, represents an extension of the first DC-MSV given in Asai and McAleer (2009), as we include Markov-switching dynamics in both the volatility and the correlations.

In particular, we assume equations (1)-(5) of the MS-CC-MSV model, and a square symmetric positive matrix, Q_t , with the following dynamics:

$$Q_t = (1 - \vartheta)\bar{Q}_t + \vartheta Q_{t-1} + W_t, \quad W_t \sim \mathcal{W}_m(\nu, \Xi) \quad (13)$$

with $|\vartheta| < 1$ and $\mathcal{W}_m(\nu, \Xi)$, a m -dimensional Wishart distribution with degrees of freedom $\nu \in \mathbb{R}^+$ and scale parameter Ξ , which is a m -dimensional positive definite matrix. We assume that the correlation structure between the elements of $\mathbf{y}_t$ admits different equilibrium regimes. Let the matrix $\bar{Q}_t$ follow a Markov-switching process:

$$\bar{Q}_t = \sum_{k=0}^{K_2-1} \bar{D}_k \mathbb{I}_{\{k\}}(s_{2,t}) \quad (14)$$

with $\mathbb{I}_E(x)$ the indicator function, which takes value 1 if $x \in E$ and 0 otherwise, and $\bar{D}_k$, $k \in \{0, 1, \dots, K_2 - 1\}$, a sequence of positive definite matrices. The process $s_{2,t} \in \{0, 1, \dots, K_2 - 1\}$, is a Markov chain with transition probability:

$$s_{2,t} \sim \mathbb{P}(s_{2,t} = j | s_{2,t-1} = i) = p_{2,ij} \quad (15)$$

which is the (i, j) -th element of the transition matrix P_2 .

In this model, the number of parameters related to the correlation dynamics is $(m^2 + m)/2 + K_2(m^2 + m)/2 + (K_2^2 - K_2) + 2$. We follow Pelletier (2006) and consider a restricted version of the model. A parsimonious specification of the Markov-switching process can be obtained by assuming a linear combination of a state of zero correlations (that is, the m -dimensional identity matrix, I_m) and that of possibly high correlations (that is, the positive definite matrix, $\bar{D}$):

$$\bar{Q}_t = \lambda_{s_{2,t}} \bar{D} + (1 - \lambda_{s_{2,t}}) I_m.$$

In order for the matrix, $\bar{Q}_t$ to be positive definite, we require $\lambda_j \in [0, 1]$, $j = 0, \dots, K_2 - 1$. Moreover, for identifiability purposes, we consider the constraints: $\lambda_0 = 1$ and $\lambda_0 > \lambda_1 > \dots > \lambda_{K_2-1}$ (see Pelletier (2006)). The restricted model has $m^2 + m + K_2^2 + 1$ parameters.

It should be noted that in the literature on dynamic correlations, the innovations of the log-volatility process are usually independent of those of the correlation. In line with this approach, we will assume that the MS process, $s_{2,t}$, driving the parameters of the correlation is independent of $s_{1,t}$. Note however, that our model can be further extended to allow for interactions between the two chains (see Mosconi and Seri (2006) for a discussion of dependent binary processes). In the next section, we will assume that the dependence between the two chains are generated by a set of common factors which drive the transition probabilities. In the empirical application, we will consider a set of weakly exogenous variables, which could explain the joint movements of the log-volatilities and the correlations.

In a general perspective, it is possible to define the dynamics of a nonlinear autoregressive process by appropriately specifying its conditional transition

distribution. In order to obtain a positive matrix-valued process, a Wishart distribution can be used with parameters depending on the past values of the process (see Philipov and Glickman (2006) for an application of this approach to the covariance process). Asai and McAleer (2009) apply a similar approach in order to specify the dynamics of Q_t^{-1} . In the present paper, we extend the latter contribution and propose a MS counterpart of the second DC-MSV of Asai and McAleer (2009).

In our model, called MS-DC-MSV2, we assume

$$Q_t^{-1} \sim \mathcal{W}_m(\nu, S_{t-1}) \quad (16)$$

with

$$S_{t-1} = \frac{1}{\nu} Q_{t-1}^{-d/2} \bar{Q}_t Q_{t-1}^{-d/2} \quad (17)$$

$$\bar{Q}_t = \lambda_{s_{2,t}} \bar{D} + (1 - \lambda_{s_{2,t}}) I_m \quad (18)$$

where $\bar{D}$ is a positive definite symmetric matrix and d is a scalar parameter. In this model, the number of parameter related to the correlation dynamics is $(m^2 + m)/2 + K_2^2 + 1$.

In the modelling strategy followed so far, we consider distinct dynamics of volatility and correlation and assume that the above processes do not share any common latent factor. Although this approach is highly flexible, it has some limitations as some exogenous events affecting the volatility (for example the effects of a financial crisis) may also have some influence on the dynamics of the correlations among different markets or assets.

In order to capture common structural breaks in the volatility and correlation processes, we define a more general stochastic correlation model (MS-DC-MSV3). We assume

$$Q_t^{-1} \sim \mathcal{W}_m(\nu, S_{t-1}) \quad (19)$$

as in model MS-DC-MSV2, and we impose the following interaction between the latent process, Q_t^{-1} , the correlation-specific MS process $s_{2,t}$ and the MS process, $s_{1,t}$:

$$S_{t-1} = \frac{1}{\nu} Q_{t-1}^{-d/2} \bar{Q}_t Q_{t-1}^{-d/2} \quad (20)$$

$$\bar{Q}_t = [\lambda_{s_{2,t}} \bar{D}_{s_{1,t}} + (1 - \lambda_{s_{2,t}}) I_m] \quad (21)$$

where

$$\bar{D}_{s_{1,t}} = \sum_{k=0}^{K_1-1} \mathbb{I}_{\{k\}}(s_{1,t}) \bar{D}_k \quad (22)$$

where $\bar{D}_k$, $k \in \{1, \dots, K_1\}$, is a sequence of positive definite matrices, which represent the long-term dependence structures in the different regimes of $s_{1,t}$. In order for the matrix, S_{t-1} to be positive definite, we require $\lambda_j \in [0, 1]$, $j = 0, \dots, K_2 - 1$. Moreover, for identifiability purposes, we consider the constraints: $\lambda_0 = 1$ and $\lambda_0 > \lambda_1 > \dots > \lambda_{K_2-1}$. This parsimonious specification of the scale matrix S_{t-1} is much in the spirit

of Pelletier (2006). However, our specification significantly extends Pelletier (2006) allowing for changes in the sign of the correlations depending on the condition of the market. In fact, our specification combines a state of zero correlations (the identity matrix, I_m) with a mixture of two market states of non-zero correlations given by the matrices $\bar{D}_0$ and $\bar{D}_1$. Since the matrix $\bar{D}_{s_{1,t}}$ is indexed by the same chain, $s_{1,t}$, which drives the log-volatility dynamics, these two market states can be interpreted as distress and normal market conditions. See also Section 5 for a simulation-based analysis of the correlation regimes in our MS-DC-MSV3 model. The idea of nesting two switching processes is similar to the double smooth transition mechanism outlined in Silvennoinen and Teräsvirta (2008). Finally, note that model MS-DC-MSV3 has model MS-DC-MSV2 as a special case when $\bar{D}_0 = \bar{D}_1 = \dots = \bar{D}_{K_1-1}$ and involves an increase in the number of parameters to $(m^2 + m)/2K_1 + K_2^2 + 1$.

3.2 Non-Homogenous Transition MS Models

In the empirical literature, there are a few papers examining the effects of some exogenous financial and macroeconomic variables on the dynamics of the correlations between the exchange rates. Among others, Chiang et al. (2007) investigate the relationship between the stock returns of various markets and the effects of some exogenous components (specifically, the phases of the Asian financial crisis and the sovereign credit-rating changes) on the dynamic correlation. They follow a two-step procedure, first they compute the correlations for a DCC-GARCH model and then run a regression of the estimated correlations on the exogenous variables. The drawback of their paper is twofold: first, a two-step procedure is not efficient, and second the regression model in the second step does not ensure that the fitted correlation coefficients are in the $(-1, +1)$ interval. It is possible to extend their analysis to our Bayesian DC-MSV framework by assuming that a set of macroeconomic and financial variables affects the transition probability between alternative correlation regimes. In this regard the proposed approach also extends the Markov-switching correlation framework outlined in Pelletier (2006), which considers a time homogeneous Markov chain but does not address the influence of macro variables on the correlation dynamics. Moreover, this model improves the analysis of Calvet et al. (2006), who propose a Markov-switching model with time-varying conditional correlation to examine the relationship between exchange rate volatility and macroeconomic variables, but does not consider the effects on the correlation dynamics.

In line with the above discussion, we propose a non-homogeneous Markov-switching process and allow the time-varying transition probabilities, $\mathbb{P}(s_{2,t} = j | s_{2,t-1} = i) = p_{2,ij,t}$, to depend on a set of exogenous variables, $\mathbf{x}_t = (1, x_{1t}, \dots, x_{Kt})' \in \mathbb{R}^{K+1}$. We specify the following relation between $p_{2,ij,t}$ and the exogenous variables:

$$p_{2,ij,t} = \{1 + \exp\{-\beta'_{ij} \mathbf{x}_t\}\}^{-1} \quad (23)$$

with $\beta_{ij} \in \mathbb{R}^k$ a vector of parameters to be estimated. We will denote the resulting models as NHMS-DC-MSV.

4 Bayesian Inference

As model MS-DC-MSV2 is a special case of model MS-DC-MSV3, we will consider here a Bayesian inference procedure to estimate the unknown parameters and the latent variables of MS-DC-MSV3. Moreover, without loss of generality, we restrict the discussion of the Bayesian analysis to the case $K_1 = K_2 = 2$.

Define $\mathbf{y} = (\mathbf{y}'_1, \dots, \mathbf{y}'_T)'$, $\mathbf{h} = (\mathbf{h}'_0, \dots, \mathbf{h}'_T)'$, $\mathbf{s}_k = (s'_{k,0}, \dots, s'_{k,T})'$, $k = 1, 2$, and $\mathbf{q} = (\text{vech}(Q_0)', \dots, \text{vech}(Q_T)')'$. The completed likelihood function associated with the MS-DC-MSV3 model for a sample of dimension T is:

$$\begin{aligned} \mathcal{L}(\mathbf{y}, \mathbf{h}, \mathbf{q}, \mathbf{s}_1, \mathbf{s}_2 | \boldsymbol{\theta}) &= \prod_{t=1}^T \left(f(\mathbf{y}_t | \mathbf{y}_{t-1}, \mathbf{h}_t, Q_t, s_{1,t}, \boldsymbol{\theta}) f(\mathbf{h}_t | \mathbf{h}_{t-1}, s_{1,t}, \boldsymbol{\theta}) \cdot \right. \\ &\quad \cdot f(Q_t | Q_{t-1}, s_{2,t}, \boldsymbol{\theta}) f(s_{1,t} | s_{1,t-1}, \boldsymbol{\theta}) f(s_{2,t} | s_{2,t-1}, \boldsymbol{\theta}) \left. \right) f(\mathbf{y}_0 | \mathbf{h}_0, Q_0, s_{1,0}, \boldsymbol{\theta}) \cdot \\ &\quad \cdot f(\mathbf{h}_0 | s_{2,0}, \boldsymbol{\theta}) f(s_{1,0} | \boldsymbol{\theta}) f(s_{2,0} | \boldsymbol{\theta}), \end{aligned} \quad (24)$$

with

$$\begin{aligned} f(\mathbf{y}_t | \mathbf{y}_{t-1}, \mathbf{h}_t, Q_t, s_{1,t}, \boldsymbol{\theta}) &= \frac{1}{(2\pi)^{m/2} |\Sigma_t|^{1/2}} \exp \left\{ -\frac{1}{2} \boldsymbol{\varepsilon}'_t \Sigma_t^{-1} \boldsymbol{\varepsilon}_t \right\} \\ f(\mathbf{h}_t | \mathbf{h}_{t-1}, s_{1,t}, \boldsymbol{\theta}) &= \frac{1}{(2\pi)^{m/2} |\Sigma_\eta|^{1/2}} \exp \left\{ -\frac{1}{2} \boldsymbol{\eta}'_t \Sigma_\eta^{-1} \boldsymbol{\eta}_t \right\} \\ f(Q_t | Q_{t-1}, s_{1,t}, s_{2,t}, \boldsymbol{\theta}) &= 2^{-\frac{m\nu}{2}} \Gamma_m(\nu/2)^{-1} |S_{t-1}|^{-\frac{\nu}{2}} \exp \left\{ -\text{tr} \left(\frac{1}{2} S_{t-1}^{-1} Q_t^{-1} \right) \right\} |Q_t^{-1}|^{\frac{\nu-m-1}{2}} \\ f(s_{k,t} | s_{k,t-1}, \boldsymbol{\theta}) &= \left(p_{k,00}^{1-s_{k,t}} (1 - p_{k,00})^{s_{k,t}} \right)^{1-s_{k,t-1}} \left(p_{k,01}^{s_{k,t}} (1 - p_{k,01})^{1-s_{k,t}} \right)^{s_{k,t-1}}, \end{aligned}$$

and $k = 1, 2$, where $\Gamma_m(\nu/2) = \pi^{m(m-1)/4} \prod_{i=1}^m \Gamma(\frac{1}{2}(\nu - i + 1))$ is the m -variate gamma function, $\boldsymbol{\varepsilon}_t$, $\boldsymbol{\eta}_t$ and S_{t-1} are defined in equations (1), (2) and (21)-(22), respectively, and $\boldsymbol{\theta} = (\mathbf{a}'_{00}, \mathbf{a}'_{01}, \text{vec}(A_{10})', \text{vec}(A_{11})', \mathbf{b}'_{00}, \mathbf{b}'_{01}, \text{vec}(B_{10})', \text{vec}(B_{11})', \text{vech}(\Sigma_\eta)', \nu, d, \lambda_1, \text{vech}(\bar{D}_0), \text{vech}(\bar{D}_1), p_{1,11}, p_{1,22}, p_{2,11}, p_{2,22})'$ is the parameter vector. The densities of $\mathbf{y}_0$, $\mathbf{h}_0$ and $\mathbf{s}_{k,0}$ correspond to the conditional stationary distributions associated with the model. Due to the difficulty in finding the stationary distribution of the Wishart process (see Philipov and Glickman (2006)), we assume that Q_0 is known and equal to I_m , as suggested in Asai and McAleer (2009).

4.1 Prior Distributions

Consider the following partition of the parameter vector, $\boldsymbol{\theta} = (\boldsymbol{\theta}'_1, \boldsymbol{\theta}'_2, \boldsymbol{\theta}'_3, \boldsymbol{\theta}'_4)'$. $\boldsymbol{\theta}_1 = \text{vec}(\Psi)$ is the collection of parameters of the observable equation, with $\Psi = (\psi_1, \dots, \psi_m)$ a $((2m+2) \times m)$ matrix, which has in the columns the sequence of $(2m+2)$ -dimensional vectors, $\psi_j = (a_{00,j}, a_{01,j}, (A_{10,j1}, \dots, A_{10,jm}), (A_{11,j1}, \dots, A_{11,jm}))'$, $j = 1, \dots, m$. $\boldsymbol{\theta}_2 = (\boldsymbol{\phi}', \text{vech}(\Sigma_\eta)')'$ contains the parameters of the log-volatility equation, with $\boldsymbol{\phi} = \text{vec}(\Phi)$, where $\Phi = (\phi_1, \dots, \phi_m)$ is a $((2m+2) \times m)$ matrix

which has in the columns the sequence of $(2m + 2)$ -dimensional vectors $\phi_j = (b_{00,j}, b_{01,j}, (B_{10,j1}, \dots, B_{10,jm}), (B_{11,j1}, \dots, B_{11,jm}))'$, $j = 1, \dots, m$. Finally, $\boldsymbol{\theta}_3 = (\nu, d, \lambda_1, \text{vech}(\bar{D}_0), \text{vech}(\bar{D}_1))'$ and $\boldsymbol{\theta}_4 = (p_{1,00}, p_{1,11}, p_{2,00}, p_{2,11})'$ are the parameter vectors of the stochastic correlation and Markov-switching processes, respectively.

We assume the following conjugate prior for $\boldsymbol{\theta}_1$:

$$\boldsymbol{\theta}_1 \sim \mathcal{N}_{m(2m+2)}(\underline{\boldsymbol{\mu}}_1, \underline{\boldsymbol{\Upsilon}}_1^{-1})$$

with density $f(\boldsymbol{\theta}_1)$ and a fully conjugate prior for $\boldsymbol{\theta}_2$:

$$\phi | \Sigma_\eta \sim \mathcal{N}_{m(2m+2)}(\underline{\boldsymbol{\mu}}_2, \Sigma_\eta \otimes \underline{\boldsymbol{\Upsilon}}_2^{-1}), \quad \Sigma_\eta^{-1} \sim \mathcal{W}_m(\underline{\boldsymbol{\mu}}_3, \underline{\boldsymbol{\Upsilon}}_3)$$

with the associated density denoted by $f(\boldsymbol{\theta}_2)$.

Consider now the elements of $\boldsymbol{\theta}_3$. In order to ensure the stationarity condition, $|d| < 1$, and the positive definiteness of $\bar{Q}_{s2,t}$ we assume the following uniform priors for d and λ_1 :

$$d \sim \mathcal{U}_{(-1,1)}, \quad \lambda_1 \sim \mathcal{U}_{(0,1)}$$

where $\mathcal{U}_{(a,b)}$ is the uniform distribution on the interval (a, b) .

The m -dimensional Wishart distribution is defined for $\nu \geq m$, so we assume a translated gamma prior for ν with parameters $\underline{\nu}$ and $\underline{\mu}_4$:

$$\nu \sim \frac{\underline{\mu}_4^{\underline{\nu}}}{\Gamma(\underline{\nu})} (\nu - m)^{\underline{\nu}-1} \exp \left\{ -\underline{\mu}_4 (\nu - m) \right\} \mathbb{I}_{(m, +\infty)}(\nu).$$

For the precision matrices $\bar{D}_i$, $i = 0, 1$, of the Wishart process, improper priors such as the objective priors (see Robert (2001), Ch. 3), cannot be easily used. In the context of Markov-switching models, improper priors may actually yield improper posterior distributions. This may happen with a positive probability when data provide no information about the parameters of one of the regimes of the dynamic model (e.g., see Billio et al. (2012)). Thus, in this work we assume a proper fairly informative prior. Moreover, as the precision matrices must be positive definite, it is natural to assume the following inverse Wishart prior:

$$\bar{D}_i^{-1} \sim \mathcal{W}_m(\underline{\boldsymbol{\mu}}_{5+i}, \underline{\boldsymbol{\Upsilon}}_{5+i}).$$

For the elements of $\boldsymbol{\theta}_4$, we assume independent uniform priors:

$$p_{k,ii} \sim \mathcal{U}_{(0,1)}$$

for $i = 0, 1$ and $k = 1, 2$.

4.2 Markov-Chain Monte Carlo

We follow a data augmentation framework (see Tanner and Wong (1987)) and apply MCMC in order to simulate from the joint posterior distribution of the parameters $\boldsymbol{\theta}$

and latent variables $\mathbf{z} = (\mathbf{h}', \mathbf{q}', \mathbf{s}_1', \mathbf{s}_2')'$. More specifically, we consider a Gibbs sampling algorithm. Some components of the Gibbs sampler can be simulated exactly, while others will be simulated by a Metropolis-Hastings step. The resulting MCMC is a hybrid Gibbs sampler (see Chib and Greenberg (1995) and Tanner (1993)).

The iteration j , $j = 1, \dots, J$, of the Gibbs sampler includes two steps. First, we simulate the parameter vector, $\boldsymbol{\theta}^{(j)}$, from its full conditional distribution, given the value of the latent variables $\mathbf{z}^{(j-1)}$ generated at the previous step, and then simulate $\mathbf{z}^{(j)}$ from its full conditional distribution, given the updated parameter value, $\boldsymbol{\theta}^{(j)}$.

In order to simulate from the full conditional of the parameter vector, we consider the partition $\boldsymbol{\theta} = (\boldsymbol{\theta}_1', \boldsymbol{\theta}_2', \boldsymbol{\theta}_3', \boldsymbol{\theta}_4')'$, with the blocks of parameters defined in the previous section, and simulate from the full conditional distribution of $\boldsymbol{\theta}_i$ given the vector of remaining parameters denoted $\boldsymbol{\theta}_{-i}$, $i = 1, \dots, 4$, that is,

$$\boldsymbol{\theta}_1^{(j)} \sim f(\boldsymbol{\theta}_1 | \boldsymbol{\theta}_2^{(j-1)}, \boldsymbol{\theta}_3^{(j-1)}, \boldsymbol{\theta}_4^{(j-1)}, \mathbf{y}, \mathbf{z}^{(j-1)}) \quad (25)$$

$$\boldsymbol{\phi}^{(j)} \sim f(\boldsymbol{\phi} | \Sigma_\eta^{(j-1)}, \boldsymbol{\theta}_1^{(j)}, \boldsymbol{\theta}_3^{(j-1)}, \boldsymbol{\theta}_4^{(j-1)}, \mathbf{y}, \mathbf{z}^{(j-1)}) \quad (26)$$

$$\Sigma_\eta^{(j)} \sim f(\Sigma_\eta | \boldsymbol{\phi}^{(j)}, \boldsymbol{\theta}_1^{(j)}, \boldsymbol{\theta}_3^{(j-1)}, \boldsymbol{\theta}_4^{(j-1)}, \mathbf{y}, \mathbf{z}^{(j-1)}) \quad (27)$$

$$\boldsymbol{\theta}_3^{(j)} \sim f(\boldsymbol{\theta}_3 | \boldsymbol{\theta}_1^{(j)}, \boldsymbol{\theta}_2^{(j)}, \boldsymbol{\theta}_4^{(j-1)}, \mathbf{y}, \mathbf{z}^{(j-1)}) \quad (28)$$

$$\boldsymbol{\theta}_4^{(j)} \sim f(\boldsymbol{\theta}_4 | \boldsymbol{\theta}_1^{(j)}, \boldsymbol{\theta}_2^{(j)}, \boldsymbol{\theta}_3^{(j)}, \mathbf{y}, \mathbf{z}^{(j-1)}). \quad (29)$$

In the second step, we simulate the latent variables, $\mathbf{z}^{(j)}$, conditionally on the updated parameters, as follows:

$$\mathbf{h}^{(j)} \sim f(\mathbf{h} | \boldsymbol{\theta}^{(j)}, \mathbf{y}, \mathbf{q}^{(j-1)}, \mathbf{s}_1^{(j-1)}, \mathbf{s}_2^{(j-1)}) \quad (30)$$

$$\mathbf{q}^{(j)} \sim f(\mathbf{q} | \boldsymbol{\theta}^{(j)}, \mathbf{y}, \mathbf{h}^{(j)}, \mathbf{s}_1^{(j-1)}, \mathbf{s}_2^{(j-1)}) \quad (31)$$

$$\mathbf{s}_1^{(j)} \sim f(\mathbf{s}_1 | \boldsymbol{\theta}^{(j)}, \mathbf{y}, \mathbf{h}^{(j)}, \mathbf{q}^{(j)}, \mathbf{s}_2^{(j-1)}) \quad (32)$$

$$\mathbf{s}_2^{(j)} \sim f(\mathbf{s}_2 | \boldsymbol{\theta}^{(j)}, \mathbf{y}, \mathbf{h}^{(j)}, \mathbf{q}^{(j)}, \mathbf{s}_1^{(j)}). \quad (33)$$

We now present the full conditional distributions of the Gibbs sampler and discuss the associated simulation methods. Define the $(2m + 2)$ -dimensional vectors, $\mathbf{x}_t = (1, s_{1,t}, \mathbf{y}_{t-1}', s_{1,t} \mathbf{y}_{t-1}')'$ and $\mathbf{z}_t = (1, s_{1,t}, \mathbf{h}_{t-1}', s_{1,t} \mathbf{h}_{t-1}')'$, $(T \times (2m + 2))$ matrix, $Z = (\mathbf{z}_0, \dots, \mathbf{z}_{T-1})'$, and $T \times m$ matrices, $E = (\boldsymbol{\eta}_1, \dots, \boldsymbol{\eta}_T)'$, $Y = (\mathbf{y}_1, \dots, \mathbf{y}_T)'$, and $H = (\mathbf{h}_1, \dots, \mathbf{h}_T)'$. Define $\mathbf{e} = \text{vec}(E)$, $\mathbf{y} = \text{vec}(Y)$, and $\mathbf{h} = \text{vec}(H)$, then the measurement equation can be written as

$$\begin{aligned} \tilde{\mathbf{y}}_t &= (\mathbf{a}_{00}, \mathbf{a}_{01}, A_{10}, A_{11})(1, s_{1,t}, \mathbf{y}_{t-1}', s_{1,t} \mathbf{y}_{t-1}')' + \boldsymbol{\varepsilon}_t \\ &= (I_m \otimes \mathbf{x}_t') \text{vec}(\boldsymbol{\psi}) + \boldsymbol{\varepsilon}_t \end{aligned} \quad (34)$$

Given the priors defined in equation (4.1), the full conditional density of $\boldsymbol{\theta}_1$ is

$$\begin{aligned}
f(\boldsymbol{\theta}_1 | \boldsymbol{\theta}_{-1}, \mathbf{y}, \mathbf{h}, \mathbf{q}, \mathbf{s}_1, \mathbf{s}_2) &\propto \\
&\propto \exp \left\{ -\frac{1}{2} \sum_{t=1}^T \text{tr} \left[-2\mathbf{y}'_t \Sigma_t^{-1} (I_m \otimes \mathbf{x}'_t) \boldsymbol{\theta}_1 + \boldsymbol{\theta}'_1 (I_m \otimes \mathbf{x}_t) \Sigma_t^{-1} (I_m \otimes \mathbf{x}'_t) \boldsymbol{\theta}_1 \right] \right\} f(\boldsymbol{\theta}_1) \\
&\propto \exp \left\{ -\frac{1}{2} \text{tr} \left[-2 \left(\sum_{t=1}^T \mathbf{y}'_t (\Sigma_t^{-1} (I_m \otimes \mathbf{x}'_t) + \underline{\boldsymbol{\mu}}'_1 \underline{\Upsilon}_1) \right) \boldsymbol{\theta}_1 \right] \right\} \\
&\exp \left\{ -\frac{1}{2} \text{tr} \left[\boldsymbol{\theta}'_1 \left(\sum_{t=1}^T (I_m \otimes \mathbf{x}_t) \Sigma_t^{-1} (I_m \otimes \mathbf{x}'_t) + \underline{\Upsilon}_1 \right) \boldsymbol{\theta}_1 \right] \right\}.
\end{aligned}$$

Thus, the full conditional distribution is

$$(\boldsymbol{\theta}_1 | \boldsymbol{\theta}_{-1}, \mathbf{y}, \mathbf{h}, \mathbf{q}, \mathbf{s}_1, \mathbf{s}_2) \sim \mathcal{N}_{m(2m+2)}(\bar{\boldsymbol{\mu}}_1, \bar{\Upsilon}_1^{-1}) \quad (35)$$

with $\bar{\boldsymbol{\mu}}_1 = \bar{\Upsilon}_1^{-1}(\underline{\Upsilon}_1 \underline{\boldsymbol{\mu}}_1 + \sum_{t=1}^T (I_m \otimes \mathbf{x}_t) \Sigma_t^{-1} \mathbf{y}_t)$, $\bar{\Upsilon}_1 = \underline{\Upsilon}_1 + \sum_{t=1}^T (\Sigma_t^{-1} \otimes \mathbf{x}_t \mathbf{x}'_t)$. For the case a higher number regimes is used for the chain s_{1t} and the matrices of autoregressive coefficients are regime dependent this sampling procedure may be computationally inefficient as it involves the inversion of high dimensional variance covariance matrices in the posterior distribution. A more efficient sampling strategy can be used conditioning the dynamic regression on the latent variables. See for example Krolzig (1997) and Frühwirth-Schnatter (2006), ch. 11-13.

Using the notation defined above, the transition equation can be written as follows

$$\mathbf{h} = (I_m \otimes Z)\boldsymbol{\phi} + \mathbf{e}, \quad \text{with } \mathbf{e} \sim \mathcal{N}_{Tm}(\mathbf{0}, \Sigma_\eta \otimes I_T) \quad (36)$$

which is a SUR regression model in vectorized form. Given the above defined priors, and using the property $\text{tr}(ABC) = \text{vec}(A')'(I_m \otimes B)\text{vec}(C)$, with A , B and C matrices of dimensions $(m \times T)$, $(T \times T)$ and $(T \times m)$, respectively, the full conditional density of $\boldsymbol{\theta}_2$ can be written as

$$\begin{aligned}
f(\boldsymbol{\theta}_2 | \boldsymbol{\theta}_{-2}, \mathbf{y}, \mathbf{h}, \mathbf{q}, \mathbf{s}_1, \mathbf{s}_2) &\propto \\
&\propto \frac{1}{(2\pi)^{mT/2} |\Sigma_\eta|^{T/2}} \exp \left\{ -\frac{1}{2} \text{tr} \left[(\boldsymbol{\phi} - \hat{\boldsymbol{\phi}})' (\Sigma_\eta^{-1} \otimes Z'Z) (\boldsymbol{\phi} - \hat{\boldsymbol{\phi}}) \right] \right\} \\
&\exp \left\{ -\frac{1}{2} \text{tr} \left[-\mathbf{h}' (\Sigma_\eta^{-1} \otimes (Z'Z)^{-1} Z') \mathbf{h} + \mathbf{h}' (\Sigma_\eta^{-1} \otimes I_T) \mathbf{h} \right] \right\} f(\boldsymbol{\theta}_2) \\
&\propto \frac{1}{(2\pi)^{mT/2}} \exp \left\{ -\frac{1}{2} \text{tr} \left[(\boldsymbol{\phi} - \hat{\boldsymbol{\phi}})' (\Sigma_\eta^{-1} \otimes Z'Z) (\boldsymbol{\phi} - \hat{\boldsymbol{\phi}}) \right] \right\} \\
&\frac{1}{|\Sigma_\eta|^{T/2}} \exp \left\{ -\frac{1}{2} \text{tr} \left[\Sigma_\eta^{-1} (H - Z\hat{\Phi})' (H - Z\hat{\Phi}) \right] \right\} f(\boldsymbol{\theta}_2) \quad (37)
\end{aligned}$$

with $\hat{\Phi} = (Z'Z)^{-1} Z'H$ and $\hat{\boldsymbol{\phi}} = \text{vec}(\hat{\Phi})' = (I_m \otimes (Z'Z)^{-1} Z')$. Due to the conjugate priors assumption, the full conditional posterior density of $\boldsymbol{\theta}_2$ can be calculated exactly

as the product of the following Gaussian and Wishart densities:

$$(\phi | \Sigma_\eta, \boldsymbol{\theta}_{-2}, \mathbf{y}, \mathbf{h}, \mathbf{q}, \mathbf{s}_1, \mathbf{s}_2) \sim \mathcal{N}_{m(2m+2)}(\bar{\boldsymbol{\mu}}_2, \Sigma_\eta \otimes \bar{\Upsilon}_2^{-1}) \quad (38)$$

$$(\Sigma_\eta^{-1} | \boldsymbol{\theta}_{-2}, \mathbf{y}, \mathbf{h}, \mathbf{q}, \mathbf{s}_1, \mathbf{s}_2) \sim \mathcal{W}_m(\bar{\boldsymbol{\mu}}_3, \bar{\Upsilon}_3) \quad (39)$$

with $\bar{\boldsymbol{\mu}}_2 = (I_m \otimes \bar{\Upsilon}_2^{-1} \underline{\Upsilon}_2) \boldsymbol{\mu}_2 + (I_m \otimes \bar{\Upsilon}_2^{-1} Z' Z) \hat{\boldsymbol{\phi}}$, $\bar{\Upsilon}_2 = (\underline{\Upsilon}_2 + Z' Z)$, $\bar{\Upsilon}_3 = (\underline{\Upsilon}_3^{-1} + (H - Z \hat{\boldsymbol{\Phi}})'(H - Z \hat{\boldsymbol{\Phi}}))^{-1}$, and $\bar{\boldsymbol{\mu}}_3 = \underline{\boldsymbol{\mu}}_3 + T - (2m + 2)$.

We consider now the full conditional of the elements of $\boldsymbol{\theta}_3$. The full conditional of ν is

$$\begin{aligned} & f(\nu | d, \lambda_1, \bar{D}_0, \bar{D}_1, \boldsymbol{\theta}_{-3}, \mathbf{y}, \mathbf{h}, \mathbf{q}, \mathbf{s}_1, \mathbf{s}_2) \propto \\ & \propto \prod_{t=1}^T \Gamma_m\left(\frac{\nu}{2}\right)^{-1} 2^{-\frac{m\nu}{2}} |S_{t-1}|^{-\frac{\nu}{2}} \exp\left\{-\frac{1}{2} \text{tr}(S_{t-1}^{-1} Q_t^{-1})\right\} |Q_t^{-1}|^{\frac{\nu-m-1}{2}} f(\boldsymbol{\theta}_3) \\ & \propto \exp\left\{\ln(\nu - m)^{\underline{\nu}-1} + \frac{\nu T m}{2} (\ln(\nu) - \ln(2)) - T \ln \Gamma_m(\nu/2) - \frac{\nu d}{2} \sum_{t=1}^T \ln |Q_{t-1}^{-1}| \right. \\ & \quad \left. + \frac{\nu}{2} \sum_{t=1}^T \ln |Q_t^{-1}| - \frac{1}{2} \text{tr}\left(\sum_{t=1}^T S_{t-1}^{-1} Q_t^{-1}\right) - \frac{\nu}{2} \left(\sum_{t=1}^T \ln(|\bar{Q}_{s_{2,t}}|) + 2\underline{\mu}_4\right)\right\} \mathbb{I}_{(m, +\infty)}(\nu) \\ & \propto \exp\left\{\ln(\nu - m)^{\underline{\nu}-1} + \frac{\nu T m}{2} \ln(\nu) - T \ln \Gamma_m(\nu/2) - \frac{\nu}{2} U_T\right\} \mathbb{I}_{(m, +\infty)}(\nu) \end{aligned} \quad (40)$$

where

$$U_T = \sum_{t=1}^T \left[d \ln |Q_{t-1}^{-1}| - \ln |Q_t^{-1}| + \text{tr}\left(Q_{t-1}^{-\frac{d}{2}} \bar{Q}_{s_{2,t}}^{-1} Q_{t-1}^{-\frac{d}{2}} Q_t^{-1}\right) + \ln(|\bar{Q}_{s_{2,t}}|) \right] + m T \ln(2) + 2\underline{\mu}_4.$$

The degrees of freedom parameter, ν , is one of the most difficult to estimate in our model. Philipov and Glickman (2006) employ a grid sampling procedure, while Asai and McAleer (2009) use an Adaptive Rejection Metropolis-Hastings (M.-H.) algorithm. We follow here an alternative route that is less time consuming and relies on the adaptation of the proposal distribution in the M.-H. step. We apply the M.-H. algorithm given in Lenk and DeSarbo (2000) and successfully used for latent variable time series models in Billio and Casarin (2011). In their algorithm, the proposal distribution uses the local information of the log-posterior, $g(\nu)$, by considering its second-order Taylor expansion centred around the mode, $\tilde{\nu}$:

$$g(\nu) \approx g(\tilde{\nu}) + g^{(1)}(\tilde{\nu})(\nu - \tilde{\nu}) + \frac{1}{2} g^{(2)}(\tilde{\nu})(\nu - \tilde{\nu})^2$$

where the first and second derivatives of g are

$$\begin{aligned} g^{(1)}(\nu) &= \frac{\underline{\nu} - 1}{\nu - m} + \frac{T m}{2} \ln(\nu) + \frac{T m}{2} - \frac{T}{2} \sum_{j=1}^m \Psi^{(0)}\left(\frac{1}{2}(\nu - j + 1)\right) \\ g^{(2)}(\nu) &= -\frac{\underline{\nu} - 1}{(\nu - m)^2} + \frac{T m}{2\nu} - \frac{T}{4} \sum_{j=1}^m \Psi^{(1)}\left(\frac{1}{2}(\nu - j + 1)\right) \end{aligned}$$

respectively, and $\Psi^{(0)}(x)$ and $\Psi^{(1)}(x)$ are the digamma and trigamma functions, respectively. At the mode, $\tilde{\nu}$, of the full conditional, $g^{(1)}(\tilde{\nu}) = 0$. However, we do not know the mode, so we evaluate $\tilde{\nu}$ by a Newton-Rapson step. Suppose at the iteration j of the algorithm we have an estimate $\tilde{\nu}^{(j-1)}$ of the mode, then it is updated as follows:

$$\tilde{\nu}^{(j)} = \tilde{\nu}^{(j-1)} + V^{(j-1)} g^{(1)}(\tilde{\nu}^{(j-1)})$$

where $V^{(j-1)} = - (g^{(2)}(\tilde{\nu}^{(j-1)}))^{-1}$. On the j -th iteration, the proposed M.-H. algorithm generates a candidate, $\nu^{(*)}$, from a truncated normal distribution with mean $\tilde{\nu}^{(j)}$, variance $V^{(j-1)}$ and domain $(m, +\infty)$. Then the candidate is accepted with log-probability:

$$\min \left\{ 0, g(\nu^{(*)}) - g(\nu^{(j-1)}) - \frac{1}{2}(\nu^{(j-1)} - \tilde{\nu}^{(j)})^2 \left(V^{(j-1)} \right)^{-1} + \frac{1}{2}(\nu^{(*)} - \tilde{\nu}^{(j)})^2 \left(V^{(j-1)} \right)^{-1} \right\}.$$

After an initial transitory period, the sequences $\tilde{\nu}^{(j)}$ and $V^{(j)}$ stabilize and do not need to be further updated. Thus, the computational time for the M.-H. step decreases.

The full conditional of d is

$$\begin{aligned} f(d|\nu, \lambda_1, \bar{D}_0, \bar{D}_1, \boldsymbol{\theta}_{-3}, \mathbf{y}, \mathbf{h}, \mathbf{q}, \mathbf{s}_1, \mathbf{s}_2) &\propto \\ &\propto \prod_{t=1}^T |S_{t-1}|^{-\frac{\nu}{2}} \exp \left\{ -\frac{1}{2} \text{tr}(S_{t-1}^{-1} Q_t^{-1}) \right\} \mathbb{I}_{(-1,1)}(d) \\ &\propto \exp \left\{ -d \left(\frac{\nu}{2} \sum_{t=1}^T \ln |Q_{t-1}^{-1}| \right) - \frac{1}{2} \text{tr} \left(\sum_{t=1}^T Q_{t-1}^{\frac{d}{2}} \bar{Q}_t^{-1} Q_{t-1}^{\frac{d}{2}} Q_t^{-1} \right) \right\} \mathbb{I}_{(-1,1)}(d). \end{aligned} \quad (41)$$

We apply a M.-H. step in order to simulate d from the posterior. We build the proposal in a similar manner to the parameter ν .

The full conditional distribution of λ_1 can be written as

$$\begin{aligned} f(\lambda_1|\nu, d, \bar{D}_0, \bar{D}_1, \boldsymbol{\theta}_{-3}, \mathbf{y}, \mathbf{h}, \mathbf{q}, \mathbf{s}_1, \mathbf{s}_2) &\propto \exp \left\{ -\frac{\nu}{2} \sum_{t=1}^T \ln(|\bar{D}_{s_{1,t}} \lambda_1 + I_m(1 - \lambda_1)|) \mathbb{I}_{\{1\}}(s_{2,t}) \right. \\ &\quad \left. - \frac{1}{2} \text{tr} \left(\sum_{t=1}^T S_{t-1}^{-1} Q_t^{-1} \mathbb{I}_{\{1\}}(s_{2,t}) \right) \right\} \mathbb{I}_{(0,1)}(\lambda_1). \end{aligned} \quad (42)$$

We use an Metropolis-Hastings algorithm with a transformed beta-random walk as proposal: $d^* = (1 - 2\omega)$ with $\omega \sim \mathcal{Be}(\alpha(d^{(j-1)}), \beta(d^{(j-1)}))$.

The full conditional distributions of the long-run components $\bar{D}_i$, $i = 0, 1$ of the

correlation matrix are

$$\begin{aligned}
f(\bar{D}_i^{-1}|\nu, d, \lambda_1, \boldsymbol{\theta}_{-3}, \mathbf{y}, \mathbf{h}, \mathbf{q}, \mathbf{s}_1, \mathbf{s}_2) &\propto |\bar{D}_i \lambda_1 + I_m(1 - \lambda_1)|^{-\frac{\nu}{2}n_{1i}} |\bar{D}_i|^{-\frac{\mu_5 + \nu n_{0i} - m - 1}{2}} \\
&\exp \left\{ -\frac{1}{2} \text{tr} \left[\sum_{t=1}^T S_{t-1}^{-1} Q_t^{-1} + \Upsilon_5^{-1} \bar{D}_i^{-1} \right] \right\} \\
&\propto |\bar{D}_i|^{-\frac{\mu_5 + \nu n_{0i} - m - 1}{2}} \exp \left\{ -\frac{1}{2} \text{tr} \left[\left(\nu \sum_{t=1}^T Q_{t-1}^{d/2} Q_t^{-1} Q_{t-1}^{d/2} \mathbb{I}_{\{0\}}(s_{2,t}) \mathbb{I}_{\{i\}}(s_{1,t}) + \Upsilon_5^{-1} \right) \bar{D}_i^{-1} \right] \right\} \\
&|\bar{D}_i \lambda_1 + I_m(1 - \lambda_1)|^{-\frac{\nu}{2}n_{1i}} \exp \left\{ -\frac{1}{2} \text{tr} \left[\sum_{t=1}^T S_{t-1}^{-1} Q_t^{-1} \mathbb{I}_{\{1\}}(s_{2,t}) \mathbb{I}_{\{i\}}(s_{1,t}) \right] \right\} \\
&\propto |\bar{D}_i|^{-\frac{\bar{\mu}_5 - m - 1}{2}} \exp \left\{ -\frac{1}{2} \text{tr} [\bar{\Upsilon}_5^{-1} \bar{D}_i^{-1}] \right\} g(\bar{D}_i^{-1}), \quad i = 0, 1,
\end{aligned}$$

with $\bar{\mu}_{5+i} = \mu_{5+i} + \nu n_{0i}$, $\bar{\Upsilon}_{5+i}^{-1} = \Upsilon_{5+i}^{-1} + \nu \sum_{t=1}^T Q_{t-1}^{d/2} Q_t^{-1} Q_{t-1}^{d/2} \mathbb{I}_{\{0\}}(s_{2,t}) \mathbb{I}_{\{i\}}(s_{1,t})$, $n_{0i} = \sum_{t=1}^T \mathbb{I}_{\{0\}}(s_{2,t}) \mathbb{I}_{\{i\}}(s_{1,t})$, and $n_{1i} = \sum_{t=1}^T \mathbb{I}_{\{1\}}(s_{2,t}) \mathbb{I}_{\{i\}}(s_{1,t})$. Note that the posterior density of $\bar{D}_i^{-1}$ is proportional to the Wishart density with parameters $\bar{\mu}_{5+i}$ and $\bar{\Upsilon}_{5+i}^{-1}$, with proportionality factor, $g(\bar{D}_i^{-1})$, which depends on $\bar{D}_i^{-1}$. Although it is not possible to simulate exactly from the posterior, we can employ a M.-H. step with proposal

$$\bar{D}_i^{-1(*)} \sim \mathcal{W}_m(\bar{\mu}_{5+i}, \bar{\Upsilon}_{5+i}). \quad (43)$$

At the j -th iteration of the M.-H. chain, given the previous value of the chain $\bar{D}^{-1(j-1)}$, we accept the candidate, $\bar{D}_i^{-1(j)} = \bar{D}_i^{-1(*)}$, with probability

$$\rho(\bar{D}_i^{-1(*)}, \bar{D}_i^{-1(j-1)}) = \min \left\{ 1, \frac{g(\bar{D}_i^{-1(*)})}{g(\bar{D}_i^{-1(j-1)})} \right\}. \quad (44)$$

For expository purposes, in the following we will denote by

$$g(s_{k,0}) = p_{k,00}^{s_{k,0}} p_{k,10}^{1-s_{k,0}} / (p_{k,00} + p_{k,10}) \quad (45)$$

the invariant density of $s_{k,t}$ associated with $p_{k,00}$ and $p_{k,10}$, $k = 1, 2$.

The full conditional distributions of the transition probabilities of the two Markov-switching processes, $s_{k,t}$, $k = 1, 2$, are independent. Thus,

$$\begin{aligned}
f(p_{k,00}|\boldsymbol{\theta}_{-4}, \mathbf{y}, \mathbf{h}, \mathbf{q}, \mathbf{s}_1, \mathbf{s}_2) &\propto g(s_{k,0}) \prod_{t=1}^T f(s_{k,t}|s_{k,t-1}, \boldsymbol{\theta}) \\
&\propto g(s_{k,0}) p_{k,00}^{n_{k,00}} (1 - p_{k,00})^{n_{k,01}}
\end{aligned} \quad (46)$$

and

$$f(p_{k,11}|\boldsymbol{\theta}_{-4}, \mathbf{y}, \mathbf{h}, \mathbf{q}, \mathbf{s}_1, \mathbf{s}_2) \propto g(s_{k,0}) p_{k,01}^{n_{k,10}} (1 - p_{k,01})^{n_{k,11}} \quad (47)$$

with

$$\begin{aligned} n_{k,00} &= \sum_{t=1}^T \mathbb{I}_{\{0\}}(s_{k,t}) \mathbb{I}_{\{0\}}(s_{k,t-1}), & n_{k,01} &= \sum_{t=1}^T \mathbb{I}_{\{1\}}(s_{k,t}) \mathbb{I}_{\{0\}}(s_{k,t-1}) \\ n_{k,10} &= \sum_{t=1}^T \mathbb{I}_{\{1\}}(s_{k,t-1}) \mathbb{I}_{\{0\}}(s_{k,t}), & n_{k,11} &= \sum_{t=1}^T \mathbb{I}_{\{1\}}(s_{k,t-1}) \mathbb{I}_{\{1\}}(s_{k,t}). \end{aligned}$$

The full conditionals of $p_{k,00}$ and $p_{k,11}$ are proportional to the beta distributions, $\mathcal{B}e(n_{k,00}+1, n_{k,01}+1)$ and $\mathcal{B}e(n_{k,10}+1, n_{k,11}+1)$, respectively with the proportionality factor $g(s_{k,0})$ which depends on $p_{k,00}$, $p_{k,01}$, $p_{k,11}$ and $p_{k,10}$. In line with approach used in the previous step of the Gibbs sampler, we apply a M.-H. sampler with a beta proposal distribution and an acceptance probability involving the proportionality factor $g(s_{k,0})$.

For the data augmentation step, we consider the following single-move Gibbs sampler. Due to the Markov property of the process for $\{\mathbf{h}_t\}_t$, the full conditional distribution of $\mathbf{h}_t$, $t = 1, \dots, T-1$ is

$$\begin{aligned} f(\mathbf{h}_t | \boldsymbol{\theta}, \mathbf{y}, \mathbf{h}_{t+1}, \mathbf{h}_{t-1}, \mathbf{q}, \mathbf{s}_1, \mathbf{s}_2) &\propto \\ &\propto |\Sigma_t|^{-1/2} \exp \left\{ -\frac{1}{2} \text{tr} \left[\boldsymbol{\varepsilon}'_t \Sigma_t^{-1} \boldsymbol{\varepsilon}_t + \mathbf{h}'_t (\Sigma_\eta^{-1} + B'_{1s_1,t+1} \Sigma_\eta^{-1} B_{1s_1,t+1}) \mathbf{h}_t \right. \right. \\ &\quad \left. \left. - 2\mathbf{h}'_t \left(\Sigma_\eta^{-1} (\mathbf{b}_{0s_1,t} + B_{1s_1,t} \mathbf{h}_{t-1}) + B'_{1s_1,t+1} \Sigma_\eta^{-1} (\mathbf{h}_{t+1} - \mathbf{b}_{0s_1,t+1}) \right) \right] \right\} \\ &\propto \exp \left\{ -\frac{1}{2} \text{tr} [(\mathbf{h}_t - \boldsymbol{\mu}_{ht})' \Upsilon_{ht}^{-1} (\mathbf{h}_t - \boldsymbol{\mu}_{ht})] \right\} g(\mathbf{h}_t) \end{aligned} \quad (48)$$

with $\mathbf{1} = (1, \dots, 1)'$, $\boldsymbol{\mu}_{ht} = \Upsilon_{ht} (\Sigma_\eta^{-1} (\mathbf{b}_{0s_1,t} + B_{1s_1,t} \mathbf{h}_{t-1}) + B'_{1s_1,t+1} \Sigma_\eta^{-1} (\mathbf{h}_{t+1} - \mathbf{b}_{0s_1,t+1}) - \frac{1}{2} \mathbf{1})$, and $\Upsilon_{ht} = (\Sigma_\eta^{-1} + B'_{1s_1,t+1} \Sigma_\eta^{-1} B_{1s_1,t+1})^{-1}$. Note that we cannot simulate exactly from this posterior. Nevertheless, the full conditional is proportional to a Gaussian distribution with proportionality factor $g(\mathbf{h}_t) = \exp \left\{ -\frac{1}{2} \text{tr} [\boldsymbol{\varepsilon}'_t \Sigma_t^{-1} \boldsymbol{\varepsilon}_t] \right\}$ depending on $\mathbf{h}_t$. This calls for the use of a M.-H. chain with proposal

$$\mathbf{h}_t^{(*)} \sim \mathcal{N}_m(\boldsymbol{\mu}_{ht}, \Upsilon_{ht}). \quad (49)$$

At the j -th iteration of the chain, given $\mathbf{h}_t^{(j-1)}$, the acceptance probability is

$$\rho(\mathbf{h}_t^{(*)}, \mathbf{h}_t^{(j-1)}) = \min \left\{ 1, \frac{g(\mathbf{h}_t^{(*)})}{g(\mathbf{h}_t^{(j-1)})} \right\}. \quad (50)$$

The full conditional of $\mathbf{h}_T$ is

$$f(\mathbf{h}_T | \boldsymbol{\theta}, \mathbf{y}, \mathbf{h}_{T-1}, \mathbf{q}, \mathbf{s}_1, \mathbf{s}_2) \propto \exp \left\{ -\frac{1}{2} \text{tr} [(\mathbf{h}_T - \boldsymbol{\mu}_{hT})' \Upsilon_{hT}^{-1} (\mathbf{h}_T - \boldsymbol{\mu}_{hT})] \right\} g(\mathbf{h}_T). \quad (51)$$

We apply a M.-H. step with the Gaussian proposal, $\mathcal{N}_m(\boldsymbol{\mu}_{hT}, \Upsilon_{hT})$, where $\boldsymbol{\mu}_{hT} = \mathbf{b}_{0s_1,T} + B_{1s_1,T} \mathbf{h}_{T-1} - \Sigma_\eta \frac{1}{2} \mathbf{1}$ and $\Upsilon_{hT} = \Sigma_\eta$.

The full conditional of Q_t , $t = 1, \dots, T-1$, is

$$\begin{aligned}
& f(Q_t^{-1} | \boldsymbol{\theta}, \mathbf{y}, \mathbf{h}, Q_{t+1}, Q_{t-1}, \mathbf{s}_1, \mathbf{s}_2) \propto \\
& \propto \exp \left\{ -\frac{1}{2} \text{tr}(S_{t-1}^{-1} Q_t^{-1}) \right\} |Q_t^{-1}|^{\frac{\nu-m-1}{2}} |\Sigma_t|^{-1/2} \exp \left\{ -\frac{1}{2} \text{tr}(\boldsymbol{\varepsilon}_t' \Sigma_t^{-1} \boldsymbol{\varepsilon}_t) \right\} \\
& |Q_{t+1}^{-1}|^{\frac{\nu-m-1}{2}} |S_t|^{-\frac{\nu}{2}} \exp \left\{ -\frac{1}{2} \text{tr}(S_t^{-1} Q_{t+1}^{-1}) \right\} \\
& \propto \exp \left\{ -\frac{1}{2} \text{tr} \left[(S_{t-1}^{-1} + \tilde{Q}_t \Lambda_t^{-1} \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' \Lambda_t^{-1} \tilde{Q}_t) Q_t^{-1} \right] \right\} |Q_t^{-1}|^{\frac{\nu-m-1}{2}} \\
& |Q_t^{-1}|^{-\frac{d\nu}{2}} |\Lambda_t \tilde{Q}_t^{-1} Q_t \tilde{Q}_t^{-1} \Lambda_t|^{-1/2} \exp \left\{ -\frac{1}{2} \text{tr}(S_t^{-1} Q_{t+1}^{-1}) \right\} \\
& \propto |Q_t^{-1}|^{\frac{\nu+1-m-1}{2}} \exp \left\{ -\frac{1}{2} \text{tr} [S_{t-1}^{-1} Q_t^{-1}] \right\} g(Q_t^{-1}). \tag{52}
\end{aligned}$$

The density is proportional to the Wishart distribution $\mathcal{W}_m(\nu+1, S_{t-1})$, with a proportionality factor

$$g(Q_t^{-1}) = |Q_t^{-1}|^{-\frac{\nu d}{2}} |\tilde{Q}_t| \exp \left\{ -\frac{1}{2} \text{tr}(S_t^{-1} Q_{t+1}^{-1} + (\tilde{Q}_t \Lambda_t^{-1} \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' \Lambda_t^{-1} \tilde{Q}_t) Q_t^{-1}) \right\}$$

which depends on Q_t and cannot be neglected. We apply a M.-H. algorithm to sample from the full conditional. Although a possible choice for the proposal is $\mathcal{W}_m(\nu+1, S_{t-1})$, we use a more efficient proposal density, which accounts for both a proxy $\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t'$ of the covariance matrix at time t , and a proxy of the normalization factor $\tilde{Q}$ for the main diagonal of the Q_t . More specifically we generate a candidate $Q_t^{-1*} \sim \mathcal{W}_m(\mu_{Q_t}, \Upsilon_{Q_t})$, with $\mu_{Q_t} = \nu+1$ and $\Upsilon_{Q_t}^{-1} = S_{t-1}^{-1} + \tilde{Q} \Lambda_t^{-1} \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' \Lambda_t^{-1} \tilde{Q}$. At the j -th iteration of the M.-H. chain, given the previous value $Q_t^{-1(j-1)}$, the candidate $Q_t^{-1(*)}$ would be accepted, that is $Q_t^{-1(j)} = Q_t^{-1(*)}$, with probability

$$\rho(Q_t^{-1(*)}, Q_t^{-1(j-1)}) = \min \left\{ 1, \frac{g(Q_t^{-1(*)})}{g(Q_t^{-1(j-1)})} \frac{\exp \left\{ -0.5 \text{tr}(\Lambda_t^{-1} \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' \Lambda_t^{-1} \tilde{Q} Q_t^{-1(j-1)} \tilde{Q}) \right\}}{\exp \left\{ -0.5 \text{tr}(\Lambda_t^{-1} \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' \Lambda_t^{-1} \tilde{Q} Q_t^{-1(*)} \tilde{Q}) \right\}} \right\}. \tag{53}$$

In our experiments we set $\tilde{Q} = (\tilde{Q}_{t-1} + \tilde{Q}_{t+1})/2$.

For Q_T^{-1} , we use M.-H. with target density:

$$\begin{aligned}
& f(Q_T^{-1} | \boldsymbol{\theta}, \mathbf{y}, \mathbf{h}, Q_T, \mathbf{s}_1, \mathbf{s}_2) \propto \\
& \propto |Q_T^{-1}|^{\frac{\nu+1-m-1}{2}} \exp \left\{ -\frac{1}{2} \text{tr} \left[(S_{T-1}^{-1} + \tilde{Q}_T \Lambda_T^{-1} \boldsymbol{\varepsilon}_T \boldsymbol{\varepsilon}_T' \Lambda_T^{-1} \tilde{Q}_T) Q_T^{-1} \right] \right\} |\tilde{Q}_T| \tag{54}
\end{aligned}$$

and proposal density defined by: $Q_T^{-1, (*)} \sim \mathcal{W}_m(\mu_{Q_T}, \Upsilon_{Q_T})$, with $\mu_{Q_T} = \nu+1$ and $\Upsilon_{Q_T}^{-1} = S_{T-1}^{-1}$.

Then the full conditional of $s_{1,t}$ can be written as

$$\begin{aligned}
& f(s_{1,t} | \boldsymbol{\theta}, \mathbf{y}, \mathbf{h}, \mathbf{q}, s_{1,t-1}, s_{1,t+1}, \mathbf{s}_2) \propto f(\mathbf{y}_t | \mathbf{y}_{t-1}, \mathbf{h}_t, Q_t, s_{1,t}, \boldsymbol{\theta}) f(\mathbf{h}_t | \mathbf{h}_{t-1}, s_{1,t}, \boldsymbol{\theta}) \\
& f(Q_t | Q_{t-1}, s_{1,t}, s_{2,t}, \boldsymbol{\theta}) f(s_{1,t} | s_{1,t-1}, \boldsymbol{\theta}) f(s_{1,t+1} | s_{1,t}, \boldsymbol{\theta}).
\end{aligned}$$

We use a global Metropolis-Hastings step (see Billio et al. (1999) and Billio and Casarin (2011)) with proposal density $f(s_{1,t}|s_{1,t-1}, \boldsymbol{\theta})$, $t = 1, \dots, T$.

The full conditional density of $s_{2,t}$ can be written as

$$f(s_{2,t}|\boldsymbol{\theta}, \mathbf{y}, \mathbf{h}, \mathbf{q}, \mathbf{s}_1, s_{2,t-1}, s_{2,t+1}) \propto f(Q_t|Q_{t-1}, s_{1,t}, s_{2,t}, \boldsymbol{\theta}) f(s_{2,t}|s_{2,t-1}, \boldsymbol{\theta}) f(s_{2,t+1}|s_{2,t}, \boldsymbol{\theta}). \quad (55)$$

For the initial states, $s_{k,0}$, $k = 1, 2$, we use a Metropolis-Hastings step with proposal given by the invariant distribution g associated with $p_{k,00}$ and $p_{k,11}$:

$$s_{k,0} \sim g(s_{k,0}) \quad (56)$$

with $k = 1, 2$.

5 Simulation Results

5.1 Data generation

We generate 3,000 samples from our MS-DC-MSV3 with the following parameter setting. For the log-volatility process we set

$$\mathbf{b}_{00} = -1.1\boldsymbol{\iota}, \mathbf{b}_{01} = 1.09\boldsymbol{\iota}, \mathbf{B}_{10} = 0.1I_3, \mathbf{B}_{11} = 0.8I_3, \Sigma_\eta = \text{diag}((0.04, 0.05, 0.06)')$$

where $\boldsymbol{\iota} = (1, 1, 1)'$. For the correlation process we set

$$\bar{D}_0 = \begin{pmatrix} 1.01 & -0.11 & -0.15 \\ -0.11 & 1.15 & -0.11 \\ -0.15 & -0.11 & 1.10 \end{pmatrix}, \bar{D}_1 = \begin{pmatrix} 1.01 & 0.17 & -0.02 \\ 0.17 & 1.15 & 0.03 \\ -0.02 & 0.03 & 1.02 \end{pmatrix}$$

$\lambda_0 = 1$, $\lambda_1 = 0.01$, $\nu = 30$ and $d = 0.8$. The two Markov chains, s_{1t} and s_{2t} have transition probabilities

$$P_1 = \begin{pmatrix} 0.91 & 0.09 \\ 0.05 & 0.95 \end{pmatrix} \text{ and } P_2 = \begin{pmatrix} 0.98 & 0.02 \\ 0.02 & 0.98 \end{pmatrix}$$

respectively. A graphical inspection (see Figure 1) of the trajectories of the three observable variables (left column, gray line) reveals a marked feature of volatility clustering. This is an effect of the regime-switching process, $s_{1,t}$ (left column, stepwise line) on the volatility process (right column, gray line). The square of the observables (right column, black line) represents a tool to detect graphically the presence of different volatility regimes. Figure 2 shows a trajectory of the stochastic correlation process (gray lines). The trajectories exhibit shifts driven by the correlation-specific Markov chain $s_{2,t}$ and by the common log-volatility and correlation Markov chain s_{1t} . The stepwise lines show the four-regime process, s_t , that results from the following composition of the two Markov-chains: $s_t = s_{1,t} + 2s_{2,t}$. The process s_t takes values in $0, 1, 2, 3$. The constraints on λ_0 and $\mathbf{b}_{01}$ (see Section 4) lead one to identify the state

$s_t = 0$ as the correlation regime associated to a low volatility state (i.e., $s_{1,t} = 0$) and $s_t = 1$ as the one associated to the high volatility state (i.e., $s_{1,t} = 1$). The correlation components associated to low- and high-volatility regimes are $\bar{D}_0$ ($\bar{D}_1$) respectively. If $s_{2,t} = 1$ then s_t is equal to two (three) in the low (high) volatility regime and the parameter $\lambda_1 \in [0, 1]$ shrinks toward zero the correlations $\bar{D}_0$ and $\bar{D}_1$. This allows for various degrees of dependence between returns in both regimes. More specifically, the lower is the value of λ_1 , the higher is the shrinkage effect. Notwithstanding the identification constraints on the parameters, our MS-DC-MSV3 model is able to capture both negative and positive correlation signs (see gray lines in Figure 2), and is thus suitable for modelling different market conditions such as distress and normal correlation with various degrees of intensity.

The recursive estimation of the empirical correlation (black line) is able to detect correlation changes but reveal some weaknesses. For example, as one can see from the graphs, there is a delay in estimating the breaks in the correlation process. Thus, although this measure can serve as preliminary analysis, a suitable inference procedure is needed for testing the presence of significant correlation changes. Therefore we apply the MCMC algorithm presented in the previous section for approximated Bayesian inference on the posterior distribution of both the parameters and the hidden states of the MS-DC-MSV3 model.

5.2 MCMC estimation

We consider the following hyperparameters setting $\boldsymbol{\theta}_1 \sim \mathcal{N}_{24}(\mathbf{0}, 10I_{24})$, $\boldsymbol{\phi}|\Sigma_\eta \sim \mathcal{N}_{24}(\mathbf{0}, \Sigma_\eta \otimes 10I_8)$ and $\Sigma_\eta \sim \mathcal{W}_3(10, 4I_3)$ for the parameters of the observations and log-volatility process. As regard to the stochastic correlation parameters we set $d \sim \mathcal{U}_{(-1,1)}$, $\lambda_1 \sim \mathcal{U}_{(0,1)}$ and a shifted gamma prior for ν , defined on the $(3, +\infty)$ interval, with parameters $\underline{\mu}_4 = 1$ and $\underline{\nu} = 10$. Moreover, we assume $\bar{D}_i^{-1} \sim \mathcal{W}_3(10, 0.1I_3)$. For the elements of the transition matrices of $s_{1,t}$ and $s_{2,t}$ we assumed independent standard uniform prior distributions.

We run the MCMC chain for 10,000 iterations and drop the initial burn-in sample of 5,000 iterations. Figures from 3 to 4 show the raw output of the MCMC chain (solid lines) for the parameters $\boldsymbol{\theta}_i$, $i = 1, \dots, 4$, and the true value of the parameters (dotted lines). From all figures one can notice that the mixing of the MCMC chain for the parameters is good and that the true value of the parameters are recovered by the posterior means. Figure 5 shows the posterior mean and the 95% credibility intervals for the log-volatilities and correlations of the model. Figure 6 shows the estimates of the Markov-switching processes and the allocation map over the MCMC iterations. The true value of the correlation and volatility processes belongs to the 95% credibility interval, denoting the effectiveness of the proposed inference procedure. A comparison between the true and the estimated value of the Markov-switching processes, defined as the maximum a-posterior-probability, leads to the same conclusion.

The estimates of the parameters, i.e. the posterior means, are given in Tables from 1 to 4. In each table we present the standard deviation (SE), the numerical standard error (NSE), the 95% credibility interval (CI) and the p-value (CD) of the

Figure 1: Simulated samples from a MS-DC-MSV3 model. Left: 3,000 samples for the observable variables (black lines, left axis). Right: square of the observable variables (black lines, left axis). In all plots, the Markov-switching process, s_{1t} , (stepwise, right axis).

Figure 2: Simulated samples from a MS-DC-MSV3 model. In each plot the stochastic correlation process (gray line, left axis), and the sequential estimation of the empirical correlation (black line, left axis). In all plots, the Markov-switching process $s_t = s_{1,t} + 2s_{2,t}$ (stepwise line, right axis) is generated by the composition of $s_{1,t}$ with $s_{2,t}$.

Figure 3: Raw output of the Gibbs sampler for θ_1 and θ_2 .

Figure 4: Raw output of the Gibbs sampler for θ_3 and θ_4 .

Figure 5: Raw output of the Gibbs sampler for the latent processes Q_t (panel (a)) and $\mathbf{h}_t$ (panel (b)). Red lines: true value of the process. Black lines: estimated latent process. Gray area: 95% high probability density region.

Figure 6: Raw output (allocation variable heatmap and estimated latent process) of the Gibbs sampler for the latent switching processes $s_{1,t}$ and $s_{2,t}$.

convergence diagnostic statistics proposed by Geweke (1992). All these quantities have been computed on the output of 10,000 iterations of the MCMC chain.

In order to compute the convergence diagnostic statistics we split the MCMC sample of dimension n into two non-overlapping sub-samples $x^{(1)}, \dots, x^{(n_A)}$ and $x^{(n-n_B+1)}, \dots, x^{(n)}$ and consider the estimators

$$\hat{\mu}_A = \frac{1}{n_A} \sum_{t=1}^{n_A} x^{(t)} \quad (57)$$

$$\hat{\mu}_B = \frac{1}{n_A} \sum_{t=n-n_B+1}^n x^{(t)} \quad (58)$$

of the mean and the estimator

$$\frac{\hat{\sigma}_j^2}{n_j} = \hat{\gamma}(0) + \frac{2n_j}{n_j - 1} \sum_{k=1}^{h_j} K(k/h_j) \hat{\gamma}(k), \quad \text{with } j = A, B \quad (59)$$

$$\hat{\gamma}(k) = \frac{1}{n_j} \sum_{l=k+1}^{n_j} (x^{(l)} - \hat{\mu}_j)(x^{(l-k)} - \hat{\mu}_j), \quad \text{with } j = A, B \quad (60)$$

of the asymptotic variance. We applied here a non-parametric estimator with $K(x)$ a Parzen kernel (a Gaussian kernel could be alternatively be used) and $h_A = 400$ and $h_B = 400$ the bandwidths. Using these quantities we evaluate the statistics

$$CD_n = \frac{\sqrt{n}(\hat{\mu}_A - \hat{\mu}_B)}{\sqrt{\frac{\hat{\sigma}_A^2}{n_A} + \frac{\hat{\sigma}_B^2}{n_B}}} \quad (61)$$

with $n_A = \tau_A n$, $n_B = \tau_B n$ and $\tau_A + \tau_B < 1$. The statistics CD_n converges in distribution to a standard normal, under the null hypothesis that the MCMC chain has converged. For the choice τ_A and τ_B we follow Geweke (1992) and set $\tau_A = 0.1$ and $\tau_B = 0.5$.

The NSE statistics is defined as

$$NSE_n = \sqrt{\frac{1}{n} \hat{S}_n(0)} \xrightarrow{n \rightarrow +\infty} 0 \quad (62)$$

where s_n^2 and $\hat{S}_n(0)$ are the empirical variance of the MCMC samples and the empirical spectral density respectively.

In all tables the p-values of the CD statistics are greater than 0.05, and hence we can conclude that the MCMC chain has converged along the different directions of the parameter space. The parameter estimates are close to the true values and the 95% credibility intervals contain the true values.

θ_1	True	Mean	SE	NSE	95% CI	CD
$a_{00,1}$	0.3	0.258	0.030	0.025	[0.201,0.317]	0.473
$a_{00,2}$	0.3	0.337	0.029	0.028	[0.279,0.393]	0.468
$a_{01,3}$	0.3	0.272	0.028	0.029	[0.219,0.327]	0.486
$a_{01,1}$	-0.6	-0.589	0.017	0.016	[-0.621,-0.556]	0.494
$a_{01,2}$	-0.6	-0.642	0.015	0.014	[-0.672,-0.612]	0.486
$a_{01,3}$	-0.6	-0.583	0.016	0.014	[-0.612,-0.553]	0.494
$A_{10,11}$	0.6	0.577	0.036	0.034	[0.508,0.650]	0.496
$A_{10,12}$	0.2	0.205	0.035	0.035	[0.136,0.272]	0.487
$A_{10,13}$	0.1	0.132	0.012	0.013	[0.108,0.155]	0.473
$A_{10,21}$	0.2	0.145	0.036	0.036	[0.075,0.214]	0.498
$A_{10,22}$	0.6	0.619	0.035	0.034	[0.552,0.688]	0.472
$A_{10,23}$	0.1	0.118	0.011	0.012	[0.097,0.140]	0.483
$A_{10,31}$	0.1	0.077	0.035	0.029	[0.008,0.145]	0.496
$A_{10,32}$	0.2	0.203	0.034	0.036	[0.138,0.268]	0.494
$A_{10,33}$	0.6	0.623	0.010	0.010	[0.605,0.642]	0.498
$A_{11,11}$	-0.3	-0.285	0.019	0.016	[-0.323,-0.249]	0.493
$A_{11,12}$	-0.1	-0.108	0.016	0.017	[-0.141,-0.076]	0.482
$A_{11,13}$	0	-0.028	0.014	0.014	[-0.055,-0.002]	0.479
$A_{11,21}$	-0.1	-0.046	0.014	0.014	[-0.073,-0.019]	0.486
$A_{11,22}$	-0.3	-0.342	0.014	0.013	[-0.369,-0.316]	0.489
$A_{11,23}$	0	-0.003	0.011	0.010	[-0.025,0.017]	0.491
$A_{11,31}$	0	0.022	0.017	0.017	[-0.012,0.056]	0.499
$A_{11,32}$	-0.1	-0.109	0.015	0.015	[-0.139,-0.079]	0.488
$A_{11,33}$	-0.3	-0.329	0.013	0.012	[-0.354,-0.304]	0.491

Table 1: Estimates of the parameters θ_1 of the MS-DC-MSV3 model, on a simulated dataset. We draw 10,000 samples from the posterior distribution of the parameters using MCMC, discard 5,000 draws and compute the posterior mean, the standard error (SE), the numerical standard error (NSE), the 95% credibility interval (CI) and p-value of the convergence diagnostic statistics (CD).

θ_2	True	Mean	SE	NSE	95% CI	CD
$b_{00,1}$	-1.1	-1.085	0.047	0.040	[-1.172,-0.999]	0.498
$b_{00,2}$	-1.1	-1.085	0.050	0.052	[-1.180,-0.994]	0.475
$b_{01,3}$	-1.1	-1.058	0.056	0.045	[-1.163,-0.954]	0.485
$b_{01,1}$	1.09	1.075	0.018	0.008	[1.059,1.091]	0.485
$b_{01,2}$	1.09	1.074	0.018	0.008	[1.057,1.091]	0.478
$b_{01,3}$	1.09	1.038	0.018	0.009	[1.019,1.057]	0.480
$B_{10,11}$	0.1	0.103	0.041	0.039	[0.026,0.180]	0.468
$B_{10,12}$	0	-0.012	0.035	0.043	[-0.079,0.058]	0.500
$B_{10,13}$	0	0.025	0.008	0.009	[0.009,0.040]	0.484
$B_{10,21}$	0	-0.010	0.043	0.040	[-0.095,0.074]	0.478
$B_{10,22}$	0.1	0.140	0.040	0.034	[0.066,0.215]	0.495
$B_{10,23}$	0	-0.013	0.009	0.009	[-0.031,0.005]	0.482
$B_{10,31}$	0	-0.033	0.047	0.043	[-0.126,0.060]	0.486
$B_{10,32}$	0	0.065	0.043	0.038	[-0.019,0.147]	0.483
$B_{10,33}$	0.1	0.108	0.016	0.010	[0.088,0.128]	0.487
$B_{11,11}$	0.8	0.804	0.016	0.009	[0.785,0.823]	0.475
$B_{11,12}$	0	0.018	0.008	0.007	[0.002,0.034]	0.495
$B_{11,13}$	0	-0.020	0.006	0.007	[-0.032,-0.009]	0.498
$B_{11,21}$	0	0.008	0.011	0.010	[-0.013,0.029]	0.474
$B_{11,22}$	0.8	0.767	0.015	0.009	[0.750,0.784]	0.488
$B_{11,23}$	0	0.011	0.007	0.007	[-0.002,0.023]	0.477
$B_{11,31}$	0	0.045	0.012	0.011	[0.022,0.069]	0.479
$B_{11,32}$	0	-0.078	0.010	0.010	[-0.097,-0.059]	0.483
$B_{11,33}$	0.8	0.773	0.014	0.007	[0.759,0.788]	0.481
$\Sigma_{\eta,11}$	0.04	0.043	0.001	0.001	[0.041,0.046]	0.491
$\Sigma_{\eta,12}$	0	-0.001	0.001	0.001	[-0.003,0.000]	0.493
$\Sigma_{\eta,13}$	0	0.002	0.001	0.001	[0.000,0.004]	0.487
$\Sigma_{\eta,22}$	0.05	0.051	0.001	0.001	[0.049,0.054]	0.499
$\Sigma_{\eta,23}$	0	0.000	0.001	0.001	[-0.002,0.002]	0.491
$\Sigma_{\eta,33}$	0.06	0.065	0.002	0.002	[0.062,0.068]	0.496

Table 2: Estimates of the parameters θ_2 of the MS-DC-MSV3 model, on a simulated dataset. We draw 10,000 samples from the posterior distribution of the parameters using MCMC, discard 5,000 draws and compute the posterior mean, the standard error (SE), the numerical standard error (NSE), the 95% credibility interval (CI) and p-value of the convergence diagnostic statistics (CD).

θ_3	True	Mean	SE	NSE	95% CI	CD
ν	30	30.303	4.869	4.218	[21.000,40.000]	0.472
d	0.8	0.803	0.019	0.025	[0.768,0.841]	0.499
λ_1	0.01	0.215	0.014	0.014	[0.189,0.244]	0.447
$\bar{D}_{0,11}$	1.01	1.010	0.013	0.014	[0.985,1.035]	0.493
$\bar{D}_{0,12}$	-0.11	-0.116	0.009	0.008	[-0.134,-0.099]	0.495
$\bar{D}_{0,13}$	-0.15	-0.153	0.013	0.013	[-0.178,-0.127]	0.493
$\bar{D}_{0,22}$	1.15	1.150	0.014	0.016	[1.122,1.179]	0.495
$\bar{D}_{0,23}$	-0.11	-0.098	0.010	0.012	[-0.118,-0.077]	0.491
$\bar{D}_{0,33}$	1.1	1.114	0.013	0.011	[1.089,1.139]	0.499
$\bar{D}_{1,11}$	1.01	1.021	0.010	0.010	[1.002,1.040]	0.471
$\bar{D}_{1,12}$	0.17	0.167	0.013	0.017	[0.139,0.192]	0.490
$\bar{D}_{1,13}$	-0.02	-0.021	0.006	0.007	[-0.033,-0.009]	0.476
$\bar{D}_{1,22}$	1.15	1.134	0.011	0.010	[1.114,1.156]	0.492
$\bar{D}_{1,23}$	0.03	0.035	0.006	0.005	[0.024,0.046]	0.498
$\bar{D}_{1,33}$	1.02	1.009	0.009	0.009	[0.991,1.027]	0.478

Table 3: Estimates of the parameters θ_3 of the MS-DC-MSV3 model, on a simulated dataset. We draw 10,000 samples from the posterior distribution of the parameters using MCMC, discard 5,000 draws and compute the posterior mean, the standard error (SE), the numerical standard error (NSE), the 95% credibility interval (CI) and p-value of the convergence diagnostic statistics (CD).

θ_4	True	Mean	SE	NSE	95% CI	CD
$p_{1,11}$	0.91	0.902	0.010	0.009	[0.881,0.921]	0.487
$p_{1,22}$	0.95	0.959	0.004	0.004	[0.950,0.967]	0.491
$p_{2,11}$	0.98	0.984	0.003	0.003	[0.978,0.989]	0.495
$p_{2,22}$	0.98	0.977	0.004	0.004	[0.968,0.985]	0.482

Table 4: Estimates of the parameters θ_4 of the MS-DC-MSV3 model, on a simulated dataset. We draw 10,000 sample from the posterior distribution of the parameters using MCMC, discard 5,000 draws and compute the posterior mean, the standard error (SE), the numerical standard error (NSE), the 95% credibility interval (CI) and p-value of the convergence diagnostic statistics (CD).

6 An Application to Exchange Rates

6.1 Data Description

We consider daily closing values for three exchange rates against the US\$, namely Euro, Yen and Pound (see Figure 7), that are well studied in the literature (e.g., see Calvet et al. (2006) and Pelletier (2006)). Among the currencies considered in the present paper, the Yen displays the highest speculative features given its major involvement in carry trade activities¹. Bank for International Settlements (BIS) statistics tracking currency flows in the international banking system confirm the leading role of the Yen as a "funding" currency in the last decade, with total Yen-denominated claims oscillating around one trillion US Dollars (Galati et al. (2007), Graph 4. p. 34). This evidence is broadly consistent with data on currency futures traded on the Chicago Mercantile Exchange (the most important market for forex futures), pointing out a significant increase in net non-commercial short open positions denominated in Yen since the beginning of 2006 (Galati et al. (2007), Graph 7 p. 39)².

As one of the cross rates considered in the analysis includes the Euro, the validity of the 'Lucas Critique' (Lucas (1976)) must be seriously taken into account in the present context. The beginning of the European Monetary Union (EMU), with the introduction of the new European currency in 1999, represents a significant policy regime shift, changing the structure of the economic system under investigation, and thereby potentially affecting the coefficients of the estimated equations. For this reason, instead of resorting to an artificial "synthetic Euro" series for the pre-EMU period, we prefer to include in our analysis only the historical Euro data, collected since the official beginning of EMU. The sample period is from 01-01-1999 to 03-11-2011, yielding a total of 3351 daily observations.

The selection of this sample period is consistent with various empirical contributions highlighting the existence of important structural breaks associated with the creation of the EMU. Using a Markov-switching GARCH dispersion model, Wilfling (2009) documents statistically significant volatility regime-switching (from a

¹The Japanese currency has been one of the most widely used "funding" currencies in these operations (together with the Swiss Franc), mainly as a consequence of the "near-zero interest rate" policy implemented by the Bank of Japan since the late 1990's. The literature points out that the Yen exhibited the higher degree of skewness among advanced countries' currencies, attributing this to large periodic yen appreciations induced by the unwinding of speculative carry trade operations (Brunnermeier et al. (2009)). An often quoted example, in this regard, is the dramatic Yen appreciation against the US Dollar occurred on October 7th and 8th of 1998.

²Quite interestingly, the importance of the Japanese currency in the build-up of carry trades operations is supported not only by different data sources, but also by the predictions of some recent theoretical models of speculative exchange rate dynamics. Nirei and Sushko (2010) outline a stochastic game model of strategic "carry traders" facing the common risk of a sudden depreciation of the "high yield", "target" currency. Aggregate uncertainty about this event makes carry traders' actions strategic complements, generating herd behavior, and leading to endogenous episodes of "explosive" carry unwinding. One important result of this paper is that the intra-day log-return series of the YEN/USD exchange rate (1999-2007) is consistent with some basic predictions from this model. More specifically, Yen appreciation jumps exhibit dependence and extreme variability, whereas Yen depreciation jumps appear to be white noise

”‘high’” to a ”‘low’” volatility regime) for all 11 EMU currencies in the run-up to European monetary unification. Further evidence strongly supporting the empirical relevance of the ‘Lucas Critique’ in the present context is provided in Van Bergeijk and Berk (2001), where a significant structural break in the term structure equation of one ”core” Euro area country (Germany) is detected at the end of 1998.

We consider the percentage log-differences of the exchange rates and denote them as $y_{1,t}$, $y_{2,t}$ and $y_{3,t}$ for Euro, Yen and Pound, respectively (see left column of Figure 8). The main descriptive statistics for the whole sample are in Table 5. The mean and median returns for the three rates are not significantly different from zero, while their standard deviations are approximately equal to 0.006. EUR/USD and YEN/USD series denote a slight leftward skewness, while the GBP/USD series is not significantly skewed. On the other hand, the three series have a high level of leptokurtosis. This departure from normality is further detected by the Jarque-Bera (JB) statistics, which motivates the use of models accounting for asymmetry and heavy-tails (such as stochastic volatility models). The Ljung-Box statistics for the log-returns at the i -th lag, $LBQ_1(i)$, show the presence of some serial correlation, particularly for the GBP/USD series. The Ljung-Box statistics for the squared log-returns, $LBQ_2(i)$, are instead all strongly significant, pointing out the existence of a dynamic structure in the conditional volatility. The variations in the volatility are also evident in the squared return series (see right column of Figure 8).

From both the left and right columns of Figure 8, it is possible to see a structural break in the volatility from the second half of 2007. This sudden change of volatility can be clearly ascribed to the effects of the recent financial crisis of 2007-2009 as shown, among others, by the empirical analysis carried out in Hui et al. (2009)³. In order to account for the effects of this financial crisis on exchange rate returns, we performed a test on squared recursive residuals which highlights a significant change in volatility for the three series at mid August 2007. A test on squared recursive residuals highlights a significant change in volatility for the three series in mid-August 2007. The dashed vertical bar in each graph indicates the date of this break and the corresponding sub-periods, 01/01/1999-14/08/2007, 15/08/2007-31/12/2009 and 01/01/2010-03/11/2011. In each sub-sample we detect the presence of outliers, defined as the observations exceeding in absolute value three times the standard deviation. The horizontal gray lines show the confidence bands approximately at the 1% significance level used to detect the outliers. The presence of outliers motivate the use of heavy-

³These authors address the dynamics of some major cross rates against the US Dollar during this crisis splitting the sample period into two sub-periods corresponding, respectively, to the earlier crisis stage (09/09/2007-12/09/2008) and to the subsequent phase following the Lehman Brothers default (15/09/2008-31/03/2009). Quite interestingly, they document a significant increase in the daily volatility of the EU/USD, GBP/USD, and YEN/USD exchange rates during the latter sub-period, and relate the above process to the severe liquidity problems associated with the deleveraging of the US financial sector and to the surge of carry-trade incentives after the Lehman default. The empirical estimates reveal a significant impact of the market-wide liquidity risk (measured by the spread between the LIBOR and the overnight index swap rate) on the movements of the Euro and the Pound against the US Dollar, whereas carry-trade incentives turn out to be more important in the case of the YEN/USD rate (see Hui et al. (2009), Tables 2a, 2b and 2d, pages 16-19).

Figure 7: EUR-USD, YEN-USD and GBP-USD exchange rates for the period 01/01/1999-03/11/2011 at a daily frequency. The vertical dashed lines correspond to dates of beginning of the 2007 financial crisis (15/08/2007) and the beginning of the Greek's debt crisis (31/12/2009).

Figure 8: Log-differences (left column) and squared log-differences (right column) of EUR-USD, YEN-USD and GBP-USD exchange rates for the period 01/01/1999-03/11/2011 at a daily frequency. The vertical dashed lines correspond to dates of beginning of the 2007 financial crisis (15/08/2007) and the beginning of the Greek's debt crisis (31/12/2009).

	01/01/1999-03/11/2011			01/01/1999-14/08/2007			15/08/2007-31/12/2009			31/12/2009-03/11/2011		
	EUR/USD	YEN/USD	GPB/USD	EUR/USD	YEN/USD	GPB/USD	EUR/USD	YEN/USD	GPB/USD	EUR/USD	YEN/USD	GPB/USD
Mean	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Median	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Max	0.026	0.035	0.034	0.024	0.019	0.034	0.026	0.035	0.028	0.020	0.017	0.030
Min	-0.039	-0.052	-0.049	-0.023	-0.021	-0.031	-0.039	-0.052	-0.049	-0.025	-0.014	-0.023
St.Dev.	0.006	0.006	0.007	0.006	0.005	0.006	0.008	0.009	0.009	0.007	0.006	0.006
Skewness	-0.117	0.052	-0.246	0.028	0.057	-0.121	-0.440	-0.046	-0.542	0.011	0.096	0.480
Kurtosis	4.730	7.295	6.383	3.777	3.697	5.205	6.174	7.045	6.216	3.406	2.722	6.196
JB	425.275*	2576.480*	1630.965*	56.804*	46.791*	460.990*	280.802*	423.579*	298.135*	3.297	2.272	222.282*
$LBQ_1(1)$	0.096	1.890	9.554*	3.367	0.129	5.655*	1.170	2.314	3.750	4.945*	0.244	0.315
$LBQ_1(20)$	21.221	46.634*	30.409	20.373	18.725	25.070	27.136	44.271*	35.778*	22.585	21.320	24.709
$LBQ_2(1)$	0.372	32.817*	28.347*	0.844	0.836	17.726*	0.372	2.904	3.862*	0.300	0.074	0.106
$LBQ_2(20)$	470.521*	1139.235*	211.756*	130.340*	115.650*	97.744*	154.621*	140.255*	49.013*	25.353*	38.333*	7.318

Table 5: Descriptive statistics of daily spot exchange rates in log-differences for the whole sample period 01/01/1999-11/06/2009 (3350 observations), and for the subperiods 01/01/1999-14/08/2007 (2249), 15/08/2007-31/12/2009 (621 observations) and 01/01/2010-03/11/2011 (480 observations). For the Jarque-Bera (JB) and the Ljung-Box (LBQ) statistics the symbol ”*” indicates that the null hypotheses of normality and absence of autocorrelation are rejected at the 5% significance level.

tails innovations (such as the Student-t) for the log-return.

Due to the presence of structural breaks, we compute the normality and autocorrelation test statistics along the above quoted sub-samples. In line with the above analysis, Table 5(b) shows an increase in the standard deviations during the latter sub-period, particularly for the GBP/USD series. Although the Jarque-Bera tests suggest the existence of significant departure from the normality in both sub-periods, the latter second sub-period is characterized by a higher level of kurtosis for all the series.

In order to test the significance of the autocorrelation structure, we take into account the volatility break and remove the observations exceeding three times the estimated standard deviation in each sub-sample. After removing the outliers the Ljung-Box statistics on the log-differences (see Table 5(b)) allow us to conclude that, differently from the former sub-period, the series exhibit some significant serial correlation in the latter. The presence of autocorrelation call for the use of autoregressive models. The Ljung-Box statistics on the squared log-returns reiterate the presence of a time-varying conditional volatility. This fact, together with the excess of kurtosis, call for the use of stochastic volatility models.

6.2 Joint constant correlation test

In order to check the significance of the time variations in the correlation structure, we follow the procedure suggested in Billio and Caporin (2005). First, we consider a VAR model for the mean and CCC-GARCH(1,1) for the variance, and obtain the residuals, ε_t , $t = 1, \dots, T$. At each date t , we estimate the correlation matrix $R_t = \{\rho_{ij,t}\}_{i,j}$, $i, j = 1, \dots, 3$, on the non-overlapping subsamples, $(\varepsilon_{t-\tau+1}, \dots, \varepsilon_t)$ and $(\varepsilon_{t+1}, \dots, \varepsilon_{t+\tau})$, to obtain the estimates $\hat{R}_{1t}$ and $\hat{R}_{2t}$, respectively.

Let $\phi(\mathbf{x}) = (0.5 \log((1 + x_1)/(1 - x_1)), \dots, 0.5 \log((1 + x_k)/(1 - x_k)))'$ be the multivariate Fisher z-transform, which goes from $\mathbb{R}^k$ to $[0, 1]^k$, and define the transformed vector of correlations, $\boldsymbol{\varsigma}_t = \phi(\text{vech}(R_t))$, $t = 1, \dots, T$. Then its empirical counterpart $\hat{\boldsymbol{\varsigma}}_t = \phi(\text{vech}(\hat{R}_t))$ follows (see Rao (1979)) the k -variate normal distribution $\mathcal{N}_k(\boldsymbol{\varsigma}_t, V\tau^{-1})$, where the variance-covariance matrix, V , is defined as follows: $\mathbb{V}(\hat{\varsigma}_{ij,t}) = 1$ and

$$\begin{aligned} \mathbb{Cov}(\hat{\varsigma}_{ij,t}, \hat{\varsigma}_{kl,t}) &= \\ &= [\rho_{ik,t}\rho_{jl,t} + \rho_{il,t}\rho_{jk,t} - \rho_{kl,t}(\rho_{ik,t}\rho_{jk,t} + \rho_{il,t}\rho_{jl,t}) - \rho_{ij,t}(\rho_{ik,t}\rho_{il,t} + \rho_{jk,t}\rho_{jl,t}) \\ &\quad + 0.5\rho_{ij,t}\rho_{kl,t}(\rho_{ik,t}^2 + \rho_{il,t}^2 + \rho_{jk,t}^2 + \rho_{jl,t}^2)] [(1 - \rho_{ij,t}^2)(1 - \rho_{kl,t}^2)]^{-1}. \end{aligned} \quad (63)$$

Let $\mathbf{i}_k$ be a k -dimensional vector of linear combinators, then the test statistics

$$D_t = \frac{\hat{\boldsymbol{\varsigma}}_{2t}'\mathbf{i}_k - \hat{\boldsymbol{\varsigma}}_{1t}'\mathbf{i}_k}{\sqrt{\mathbf{i}_k'V_{1,t}\mathbf{i}_k\tau^{-1} + \mathbf{i}_k'V_{2,t}\mathbf{i}_k\tau^{-1}}} \quad (64)$$

follows under the null hypothesis of equal correlation: $\mathcal{H}_0 : (\boldsymbol{\varsigma}_{2t}'\mathbf{i}_k - \boldsymbol{\varsigma}_{1t}'\mathbf{i}_k) = 0$, the normal distribution: $\mathcal{N}(0, \mathbf{i}_k'V_{1,t}\mathbf{i}_k\tau^{-1} + \mathbf{i}_k'V_{2,t}\mathbf{i}_k\tau^{-1})$.

Figure 9: Left column: correlations $\rho_{12,t}$, $\rho_{13,t}$ and $\rho_{23,t}$ between (EUR-USD,YEN-USD), (EUR-USD,GBP-USD) and (GBP-USD,YEN-USD), respectively, estimated sequentially over the period 01/01/1999-03/11/2011 by using a window size of $\tau = 60$ observations. Right column: constant correlation test ($e_{12,t}$, $e_{13,t}$ and $e_{23,t}$) for each correlation. The horizontal grey lines represent the confidence bands for the joint test of significance at the 1% and 5% significance levels. The vertical dashed lines correspond to dates of beginning of the 2007 financial crisis (15/08/2007) and the beginning of the Greek's debt crisis (31/12/2009).

Focusing on the left column plots reported in Fig. 9, we observe that the correlation pattern between EU/USD and GBP/USD volatilities ($\rho_{13,t}$) always oscillates around positive values, albeit inside a relatively large interval (0.4-0.8). The remaining two correlations exhibit instead a different pattern. Although positive for most of the time, these correlations display negative values both at the beginning of the sample and during the more recent period of financial crisis. More specifically, the largest negative values of $\rho_{12,t}$ and $\rho_{23,t}$ are recorded around observation 2500 (1st of August 2008), i.e. one of the more acute phases of the financial crisis. Since both $\rho_{12,t}$ and $\rho_{23,t}$ involve correlations with the YEN/USD, this captures a tendency for this exchange rate to invert its correlation structure from positive to negative under periods of particular financial stress. This, in turn, is likely to reflect the peculiar speculative nature of the Japanese Yen mentioned above, acting as a safe-heaven anchor for the funding of carry trade operations during particularly turbulent periods.

Turning to the constant correlation test (Fig. 9, right columns plots) it is apparent that, for all exchange rate pairs ($e_{12,t}$, $e_{13,t}$, $e_{23,t}$), this statistics exceeds quite often the confidence bands over the whole sample. This preliminary evidence supports therefore the existence of significant changes (at both the 1% and 5% significance

Figure 10: Pairwise joint constant correlation test for $(\rho_{12,t}, \rho_{23,t})$ (first row), $(\rho_{23,t}, \rho_{13,t})$ (second row) and $(\rho_{12,t}, \rho_{13,t})$ (third row). Joint test of constant correlation for $(\rho_{12,t}, \rho_{23,t}, \rho_{13,t})$ (fourth row). In all the rows the vertical grey lines represents the confidence bands for the joint test of significance at the 1% and 5% significance levels. The vertical dashed lines correspond to dates of beginning of the 2007 financial crisis (15/08/2007) and the beginning of the Greek's debt crisis (31/12/2009).

levels) in the correlations. The time-varying correlations motivates the application of our proposed models to exchange rates data. Focusing on constant correlation tests involving the YEN/USD exchange rate ($e_{12,t}$, $e_{23,t}$) two distinctive features stand out. First, the overall pattern of this statistic is different during the financial crisis (see vertical bars in the plots), with sharp upward and downward movements which are on average larger than those recorded during the pre-crisis period. Second, this amplified pattern characterizes not only the crisis period, but also the months immediately preceding the financial turmoil. This evidence reiterates the speculative nature of the YEN/USD exchange rate underlined in the previous discussion. Moreover, the abrupt movements of constant correlation tests involving the YEN/USD rate on the verge of the financial crisis, suggest the existence of early warning signals of the upcoming turmoil originating from the foreign exchange market.

Figure 10 reports plots of pairwise joint constant correlation tests (first three rows) and of the joint test of constant correlation for all cross-rates against the US Dollar (fourth row). Overall, these plots display very similar patterns, and confirm that a dynamic structure in co-movements is a pervasive empirical regularity characterizing both the former part of the sample and the more recent phases of financial turmoil. All plots appearing in Figure 10 reiterate the anticipatory nature of these test statistics with respect to the 2007-2008 financial crisis, as witnessed by the sharp upward and

downward movements occurring just before August 2007 (the first vertical dashed line in all figures). As discussed above, the presence of the Japanese Yen among the currencies considered is likely to play a prominent role in this regard. Focusing on the joint test for constant correlation (Figure 4, fourth row) we find that, in line with most of the evidence documented in Figure 9, rejections of the null hypothesis tend to be stronger during the second half of the sample. More specifically, we document three significant rejections in the phase immediately preceding the financial turmoil, two significant rejections during the former crisis period (2007-2009), and one significant rejection during the period following the Greek's debt crisis (2010-2011).

6.3 Results of the MS-DC-MSV3 model

We apply our Bayesian MS-DC-MSV3 and the proposed MCMC algorithm (see Section 4) for posterior approximation. We assume the prior hyperparameters discussed in Section 5, run the MCMC chain for 10,000 iterations and drop the initial sample of 5,000 iterations. Figure 11 plots the posterior mean and the 95% credibility intervals for the regression coefficients θ_1 and the log-volatility parameters θ_2 . Vertical bars indicate the different subset of parameters within each vector.

The left graph of Figure 11 plots the posterior mean (dark line) and the credibility intervals (gray area) for the regression intercepts $\mathbf{a}_{00}$ and $\mathbf{a}_{00} + \mathbf{a}_{01}$ (first and second part of the graph). One can find a substantial evidence of negative returns associated with the low-volatility regimes and similar substantial evidence of positive returns in high-volatility regimes. The result is in line with previous studies in the literature on Markov-switching models for exchange rates (see e.g. Dueker and Neely (2007)).

The credibility intervals for the autoregressive coefficients A_{10} and $A_{10} + A_{11}$ (third and fourth block, left graph of Figure 11) show a strong evidence in favour of the absence of serial correlation in the daily exchange rate returns. These results support the restrictions imposed by Kim et al. (1998) and are in line with those obtained by Loddo et al. (2011), which analyse the same rates on a shorter time period (2001-2005) assuming a VAR model with one lag. We arrive at the same conclusion with a Markov-switching VAR model and extend their results, by documenting no substantial differences, in the returns autocorrelation, between the high- and low-volatility regimes.

The right graph of Figure 11 plots the posterior mean (dark line) and the credibility intervals (gray area) for the intercept $\mathbf{b}_{00}$ and $\mathbf{b}_{00} + \mathbf{b}_{01}$ (blocks one and two of parameters) and for the autoregressive coefficients B_{10} and $B_{10} + B_{11}$ (block third and four) of the log-volatility process. We find evidence of increase of the volatility level during the high-volatility regime, and significant volatility persistence during both high and low volatility phases.

Focusing on the ν parameter, corresponding to the estimated degrees of freedom of the Wishart distribution (see Eq. 19), its posterior density is highly concentrated on relatively high values (30.862). Thus, in line with Asai and McAleer (2009), the variability in correlation dynamics appears, to a large extent, to be driven by the underlying Wishart process.

Figure 11: Posterior means (solid lines) and 95% credibility intervals (gray areas) for the components of the parameter vectors θ_1 and θ_2 . Vertical bars denote the different blocks of parameters: $\mathbf{a}_{00}, \mathbf{a}_{00} + \mathbf{a}_{01}, \mathbf{A}_{10}$ and $\mathbf{A}_{10} + \mathbf{A}_{11}$ (left chart) and $\mathbf{b}_{00}, \mathbf{b}_{00} + \mathbf{b}_{01}, \mathbf{B}_{10}, \mathbf{B}_{10} + \mathbf{B}_{11}$ and Σ_η (right chart).

Figure 12: Estimated posterior densities (solid lines) for the components of the parameter vector θ_3 .

As explained in Section 3, the d parameter crucially affects the time-varying correlation dynamics in our model. More specifically, this parameter drives the dependence of correlations at time t on past correlation values (see Eq. 20). Values of d largely far from zero (0.824), such as those indicated in the estimated posterior density in Figure 12, provide therefore strong evidence supporting the existence of significant dynamic correlations.

The estimated posterior density for λ_1 is strictly positive and strongly concentrated around a value of 0.296. Since λ_1 affects the long-run correlation structure (see Eq. 21), a strictly positive value for this parameter points out a significant "shrinking effect", namely a tendency for $\bar{Q}_t$ to move towards the possibly non-zero correlations structure in both high and low volatility states. This result is in line with the stylized facts about financial contagion effects reported in Corsetti et al. (2011) (see also discussion later on in this section).

As regards to the configuration of the long-run correlation parameters during the regimes of low ($\bar{D}_0$) and high ($\bar{D}_1$) volatility, there is a strong evidence in favour of correlation switches in the long-run. More specifically, one common feature for all exchange rates pairs is that, in the low-volatility regime, these long-run correlations are always negative (see $\bar{D}_{0,12}$, $\bar{D}_{0,13}$ and $\bar{D}_{0,23}$ in Figure 12 and in Table 6). The underlying intuition for this evidence is that, during market periods of low volatility, there are no contagion effects from each nominal exchange rate to the others. This result is reversed during the high-volatility regime. Moreover, the long-run correlation coefficients are now positive for the exchange rate pairs including the Japanese Yen (i.e. both $\bar{D}_{1,12}$ and $\bar{D}_{1,23}$ are greater than zero). The long-run correlation coefficient for the high-volatility regime remains negative focusing on the the Euro/US\$ and Pound/US\$ pair ($\bar{D}_{1,13} < 0$). Overall, this latter evidence points out the existence of significant contagion effects when the market switches from low-volatility to turbulent periods. At the same time, however, our results suggest that such contagion effects arise only for exchange rate pairs including the Japanese currency, reiterating the peculiar speculative nature of the Yen already discussed in Section 6.2.

Figure 13 plots the posterior means of conditional volatilities (black lines) and the alternative volatility regimes (red lines) for the three nominal exchange rates against the US Dollar (Euro/Dollar, Yen/Dollar, and Pound/Dollar respectively). All these bilateral exchange rates exhibit substantial evidence of time-varying volatility along the sample, as well as frequent simultaneous switches between low and high volatility regimes. Focusing on the last part of the sample, and more specifically around observation 2500 (1st August 2008), a relevant volatility increase is apparent for all exchange rates. Since simultaneous volatility increases in different asset returns are commonly recognized as one of the main stylized fact characterizing financial contagion, this evidence indicates that the recent US sub-prime crisis significantly affected the foreign exchange markets of major cross rates against the US Dollar. Note, moreover, that this volatility increase is in line with the empirical evidence discussed in Corsetti et al. (2011), where a sharp increase in the volatility of stock market returns is detected during the years 2008-2009 (see ibidem, Figure 2).

The volatility transmission is not the only contagion mechanism reported in the

θ_3	Mean	SE	NSE	95% CI	CD
ν	30.862	1.154	2.637	[29.000,33.000]	0.252
d	0.824	0.011	0.026	[0.806,0.842]	0.005
λ_1	0.296	0.016	0.016	[0.267,0.323]	0.442
$\bar{D}_{0,11}$	1.214	0.024	0.058	[1.178,1.267]	0.330
$\bar{D}_{0,12}$	-0.064	0.007	0.012	[-0.082,-0.052]	0.428
$\bar{D}_{0,13}$	-0.141	0.009	0.013	[-0.160,-0.128]	0.271
$\bar{D}_{0,22}$	1.137	0.010	0.021	[1.120,1.157]	0.120
$\bar{D}_{0,23}$	-0.036	0.006	0.010	[-0.047,-0.025]	0.361
$\bar{D}_{0,33}$	1.191	0.021	0.046	[1.165,1.245]	0.356
$\bar{D}_{1,11}$	1.156	0.042	0.089	[1.079,1.222]	0.472
$\bar{D}_{1,12}$	0.000	0.020	0.038	[-0.035,0.040]	0.143
$\bar{D}_{1,13}$	-0.101	0.034	0.070	[-0.155,-0.041]	0.168
$\bar{D}_{1,22}$	1.107	0.030	0.059	[1.054,1.170]	0.232
$\bar{D}_{1,23}$	0.047	0.018	0.027	[0.014,0.087]	0.120
$\bar{D}_{1,33}$	1.218	0.027	0.053	[1.161,1.265]	0.084

θ_4	Mean	SE	NSE	95% CI	CD
$p_{1,11}$	0.964	0.007	0.017	[0.952,0.978]	0.491
$p_{1,22}$	0.932	0.011	0.026	[0.914,0.953]	0.305
$p_{2,11}$	0.977	0.008	0.016	[0.956,0.989]	0.427
$p_{2,22}$	0.877	0.043	0.100	[0.793,0.971]	0.282

Table 6: Estimates of the parameters θ_3 (top) and θ_4 (bottom) of the MS-DC-MSV3 model, on the daily exchange rate log-returns. We draw 10,000 sample from the posterior distribution of the parameters using MCMC, discard 5,000 draws and compute the posterior mean, the standard error (SE), the numerical standard error (NSE), the 95% credibility interval (CI) and p-value of the convergence diagnostic statistics (CD).

literature. In fact, also correlation changes may provide evidence of contagion effects (e.g., see Forbes and Rigobon (2002)). Figure 14 plots the conditional correlations (black lines) between alternative exchange rate pairs against the US Dollar, namely: Euro and Yen ($\rho_{12,t}$), Euro and Pound ($\rho_{13,t}$), Yen and Pound ($\rho_{23,t}$). The same figure indicates the alternative volatility regimes associated with estimated conditional correlations (stepwise red lines), namely: low volatility ($s_t = 0$), high volatility ($s_t = 1$), low volatility with shrinking towards zero correlation ($s_t = 2$), and high volatility with shrinking towards zero correlation ($s_t = 3$). According to Forbes and Rigobon (2002), contagion may be defined as a significant increase in cross-market correlation between asset returns in the aftermath of an (exogenous) crisis event.

More recently, however, the above definition has been widely criticized in the literature relying on standard factor models in order to evaluate the empirical evidence on financial contagion stemming from correlation analysis (Corsetti et al. (2005, 2011)). The main point of this critique is that, for given factor loadings, correlation between asset returns depends on the variance of the common factor affecting returns relative to the variance of idiosyncratic asset noise. Correlation between returns will therefore rise, during a crisis, only to the extent that movements in the common factor are relatively large. In a more general perspective, however, correlation may increase or decrease during a crisis. The empirical evidence for various financial markets is broadly consistent with the above remarks: although correlations of returns often increase during periods of financial turbulence, there are indeed many crisis episodes in which correlations actually fall at the outbreak of a crisis (see Corsetti et al. (2001) and Corsetti et al. (2011), Fig. 4).

Corsetti et al. (2005) outline an innovative econometric framework to test the null hypothesis of interdependence (defined as a variation in correlation consistent with the data-generating process) against the alternative hypothesis of contagion (defined as a structural break in the international transmission of shocks). The central idea of this framework is to assess whether the correlation of asset returns observed during a crisis period is significantly larger (or smaller) than an adjusted-correlation measure assuming only interdependence.

Although the econometric framework of the present paper differs from that outlined in Corsetti et al. (2005, 2011) (since we do not rely on a factor model to explain the dynamics of asset returns), the approach taken by these authors provides some useful insights to evaluate the empirical evidence reproduced in Figure 14.

More specifically, the following two issues deserve, in our opinion, particular attention. First, what is the overall pattern of cross-market correlations and, most importantly, if this pattern exhibits significant changes during periods of financial turbulence. The previous discussion suggests that cross correlation patterns should tend to be more erratic and unstable in the presence of financial contagion. Second, what is the frequency of the state $s_t = 3$ along the sample and if this frequency exhibits significant changes in some specific sub-periods. The state $s_t = 3$ corresponds to a high volatility regime in the presence of shrinking to zero correlation. On the other hand, the econometric framework outlined in Corsetti et al. (2005, 2011) relies on identifying large correlation discrepancies relative to a given benchmark. We expect,

Figure 13: Posterior means (solid lines, left axes) and 95% credibility regions (gray areas, left axes) of the log-volatility $\mathbf{h}_t$. Each figure includes $\hat{s}_{1,t}$ and each figure includes $\hat{s}_t = \hat{s}_{1,t} + 2\hat{s}_{2,t}$ (stepwise, right axes). The vertical dashed lines correspond to dates of beginning of the 2007 financial crisis (15/08/2007) and the beginning of the Greek's debt crisis (31/12/2009).

Figure 14: Posterior means (solid lines, left axes) and 95% credibility regions (gray areas, left axes) of the correlation Ω_t . Each figure includes $\hat{s}_t = \hat{s}_{1,t} + 2\hat{s}_{2,t}$ (stepwise, right axes). The vertical dashed lines correspond to dates of beginning of the 2007 financial crisis (15/08/2007) and the beginning of the Greek's debt crisis (31/12/2009).

therefore, the frequency of $s_t = 3$ to be higher in periods characterized by financial contagion, since in these periods there will be a more frequent correction of excessive correlation movements induced by a stronger transmission mechanism of international shocks.

The dynamic pattern of conditional correlations in Figure 14 reveals a clear difference between $\rho_{13,t}$ and the remaining correlation parameters relative to the Japanese currency against the Euro ($\rho_{12,t}$) or against the Pound ($\rho_{23,t}$).

While $\rho_{13,t}$ displays a rather smooth path, with values (almost) always positive along the whole sample, the remaining correlations exhibit more variability, with more frequent switches between positive and negative values. This evidence reiterates the peculiar speculative nature of the Yen already discussed in previous sections and suggests that, in the context of our sample, contagion episodes are more likely to occur in bilateral correlations involving the Japanese currency.

Focusing on $\rho_{12,t}$ and $\rho_{23,t}$, a structural break is clearly apparent in the last part of the sample. Since mid-2007, these conditional correlations show a much higher degree of instability and increased clusterings of positive and negative spikes with respect to earlier periods.

Overall, this evidence points out that recent episodes of turmoil on the US sub-prime market and on the European sovereign bonds market produced a relevant discontinuity in the international transmission of financial shocks. More specifically, our DCC model captures significant contagion effects from these financial crises to the foreign exchange market, as witnessed by the more erratic and unstable paths of $\rho_{12,t}$ and $\rho_{23,t}$ towards the end of the sample.

Focusing again on the last part of the sample we finally observe, for all bilateral exchange rate pairs, a notable increase in the frequency of $s_t = 3$, namely of the high-volatility regime in the presence of shrinking to zero correlation, with respect to earlier periods. Since, as discussed before, the occurrence of $s_t = 3$ is expected to be more likely in periods of financial contagion, this finding corroborates the main evidence on contagion effects obtained through the estimates of dynamic conditional correlations.

7 Conclusion

We propose some new Markov-switching stochastic correlation models (MS-DC-MVS). We follow a Bayesian inference approach to parameter and latent variable estimation and develop a MCMC algorithm for posterior approximation. The results on simulated data show us that the proposed algorithm is efficient. The MS-DC-MVS allow us to study the relationships between exchange rates data, to disentangle the correlation regimes from the volatility regimes and to analyse contagion effects on currency markets. We found evidence of coexistence of different kinds of volatility and correlation regimes. This class of models can be useful to study whether similar or not similar relationships are existing on financial markets.

acknowledgement

This research is supported by PRIN 2010-11 grant by the Italian Ministry of Education, University and Research (MIUR) and by the European Union, Seventh Framework Programme FP7/2007-2013 under grant agreement SYRTO-SSH-2012-320270.

References

- Aas, K. and Haff, I. (2006). The Generalized Hyperbolic Skew Student's t-Distribution. *Journal of Financial Econometrics*, 4(2):275–309.
- Aguilar, O. and West, M. (2000). Bayesian Dynamic Factor Models and Portfolio Allocation. *Journal of Business and Economic Statistics*, 18:338–357.
- Amisano, G. and Casarin, R. (2007). Particle Filters for Markov-Switching Stochastic Correlation Models. In *Proceedings SIS, 2007, Risk and Prediction*. CLEUP, Padova.
- Asai, M. and McAleer, M. (2005). Dynamic Correlations in Symmetric Multivariate SV Models. In Zerger, A. and Argent, R. M., editors, *MODSIM 2005 International Congress on Modelling and Simulation*. Modelling and Simulation Society of Australia and New Zealand.
- Asai, M. and McAleer, M. (2009). The Structure of Dynamic Correlations in Multivariate Stochastic Volatility Models. *Journal of Econometrics*, 150(2):182–192.
- Asai, M., McAleer, M., and Yu, J. (2006). Multivariate Stochastic Volatility: A Review. *Econometric Reviews*, (2–3):145–175.
- Azzalini, A. and Dalla Valle, A. (1996). The Multivariate Skew-Normal Distribution. *Biometrika*, 83:715–726.
- Bauwens, L. and Laurent, S. (2002). A New Class of Multivariate Skew Densities, with Application to garch Models. Technical report N.2002-20, CORE, University of Lovain.
- Bauwens, L., Laurent, S., and Rombouts, J. (2006). Multivariate GARCH Models: A Survey. *Journal of Applied Econometrics*, 21(1):79–109.
- Billio, M. and Caporin, M. (2005). Multivariate Markov Switching Dynamic Conditional Correlation GARCH Representations for Contagion Analysis. *Statistical Methods and Applications*, 14(2):145–161.
- Billio, M. and Caporin, M. (2009). A Generalized Dynamic Conditional Correlation Model for Portfolio Risk Evaluation. *Mathematics and Computers in Simulation*, 79(8):2566–2578.

- Billio, M. and Casarin, R. (2011). Beta Autoregressive Transition Markov-Switching Models for Business Cycle Analysis. *Studies in Nonlinear Dynamics and Econometrics*, 15(4):1–30.
- Billio, M., Casarin, R., Ravazzolo, F., and van Dijk, H. (2012). Combination Schemes for Turning Point Predictions. *Quarterly Reviews of Economics and Finance*, 52:402–412.
- Billio, M., Monfort, A., and Robert, C. (1999). Bayesian Estimation of Switching ARMA Models. *Journal of Econometrics*, 93:229–255.
- Bollerslev, T. (1987). A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return. *Review of Economics and Statistics*, 69(3):542–547.
- Bollerslev, T. (1990). Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Approach. *Review of Economics and Statistics*, 72(3):498–505.
- Bollerslev, T., Engle, R., and Wooldridge, J. (1988). A Capital Asset Pricing Model with Time Varying Covariances. *Journal of Political Economy*, 96(1):116–131.
- Branco, M. and Dey, D. (2001). A General Class of Multivariate Skew Elliptical Distributions. *Journal of Multivariate Analysis*, 79:99–113.
- Brunnermeier, M., Nagel, S., and Pedersen, L. (2009). Carry Trades and Currency Crashes. *NBER Macroeconomics Annual*, 23:313–348.
- Calvet, L., Fisher, A., and Thompson, S. (2006). Volatility Comovement: A Multifrequency Approach. *Journal of Econometrics*, 131(1-2):179–215.
- Cappiello, L., Engle, R., and Sheppard, K. (2006). Asymmetric Dynamics in the Correlations of Global Equity and Bond Returns. *Journal of Financial Econometrics*, 4:537–572.
- Casarin, R. and Sartore, D. (2007). Matrix-State Particle Filter for Wishart Stochastic Volatility Process. In *Proceedings SIS, 2007, Risk and Prediction*. CLEUP, Padova.
- Chan, D., Kohn, R., and Kirby, C. (2005). Multivariate Stochastic Volatility with Leverage. *Econometric Reviews*, 25(2-3):245–274.
- Chiang, T., Jeon, B., and Li, H. (2007). Dynamic Correlation Analysis of Financial Contagion: Evidence from Asian Markets. *Journal of International Money and Finance*, 26(7):1206–1228.
- Chib, S. and Greenberg, E. (1995). Understanding the Metropolis-Hastings Algorithm. *The American Statistician*, 49:327–335.

- Chib, S., Nardari, F., and Shephard, N. (2006). Analysis of High-Dimensional Multivariate Stochastic Volatility Models. *Journal of Econometrics*, 134(2):341–371.
- Clark, T. E. (2011). Real-Time Density Forecasts From Bayesian Vector Autoregressions With Stochastic Volatility. *Journal of Business and Economic Statistics*, 29(3):327–341.
- Clark, T. E. and Ravazzolo, F. (2012). The Macroeconomic Forecasting Performance of Autoregressive Models with Alternative Specifications of Time-Varying Volatility. Working Paper N. 12-18, Federal Reserve Bank of Cleveland.
- Corsetti, G., Pericoli, M., and Sbracia, M. (2001). Correlation Analysis of Financial Contagion: What One Should Know Before Running a Test. Technical report, Temi di discussione N. 408, Banca d’Italia.
- Corsetti, G., Pericoli, M., and Sbracia, M. (2005). Some Contagion, some Interdependence: More Pitfalls in Tests of Financial Contagion. *Journal of International Money and Finance*, 24(8):1177 – 1199.
- Corsetti, G., Pericoli, M., and Sbracia, M. (2011). *Financial Contagion: The Viral Threat to the Wealth of Nations*, chapter Correlation Analysis of Financial Contagion, pages 11–20. John Wiley & Sons.
- Creal, D., Koopman, S. J., and Lucas, A. (2011). A Dynamic Heavy-Tailed Model for Time-Varying Volatilities and Correlation. *Journal of Business and Economic Statistics*, 29(4):552–563.
- Dueker, M. and Neely, C. J. (2007). Can Markov Switching Models Predict Excess Foreign Exchange Returns? *Journal of Banking and Finance*, 31(2):279–296.
- Dungey, M., Fry, R., and Martin, V. (2004). Currency Market Contagion in the Asia-Pacific Region. *Australian Economic Papers*, 43(4):379–395.
- Engle, R. (2002). Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models. *Journal of Business and Economic Statistics*, 20(3):339–350.
- Engle, R. F. and Kroner, K. F. (1995). Multivariate Simultaneous Generalized ARCH. *Econometric Theory*, 11(1):122–150.
- Ferreira, J. and Steel, M. (2007). Bayesian Multivariate Regression Analysis with a New Class of Skewed Distributions. *Statistica Sinica*, 17:505–529.
- Forbes, K. and Rigobon, R. (2002). No Contagion, Only Interdependence: Measuring Stock Market Co-Movements. *Journal of Finance*, 57(5):2223–2261.
- Frühwirth-Schnatter, S. (2006). *Mixture and Markov-switching Models*. Springer-Verlag, New York.

- Galati, G., Heath, A., and McGuire, P. (2007). Evidence of Carry Trade Activity. *BIS Quarterly Review*, September:27–41.
- Galeano, P. and Ausín, M. C. (2010). The Gaussian Mixture Dynamic Conditional Correlation Model: Parameter Estimation, Value at Risk Calculation, and Portfolio Selection. *Journal of Business and Economic Statistics*, 28(4):559–571.
- Geweke, J. (1992). Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments. In Bernardo, J. M., Berger, J. O., Dawid, A. P., and Smith, A. F. M., editors, *Bayesian Statistics 4*, pages 169–193. Oxford University Press, Oxford.
- Gourieroux, C., Jasiak, J., and Sufana, R. (2004). The Wishart Autoregressive Process of Multivariate Stochastic Volatility. Working Paper, CREST.
- Harvey, A., Ruiz, E., and Shephard, N. (1994). Multivariate Stochastic Variance Models. *Review of Economic Studies*, 61:247–264.
- Horen, N., Jager, H., and Klaassen, F. (2006). Foreign Exchange Market Contagion in the Asian Crisis: A Regression Based Approach. *Review of World Economics*, 142(2):374–401.
- Hui, C., Genberg, H., and Chung, T. (2009). Liquidity, Risk Appetite and Exchange Rate Movements During the Financial Crisis of 2007-2009. Hong Kong Monetary Authority, Working Paper 11/2009.
- Jacquier, E., Polson, N., and Rossi, P. (1994). Bayesian Analysis of Stochastic Volatility Models. *Journal of Business and Economic Statistics*, 12:281–300.
- Jacquier, E., Polson, N., and Rossi, P. (2004). Bayesian Analysis of Stochastic Volatility Models with Fat-Tails and Correlated Errors. *Journal of Econometrics*, 122:185–212.
- Kim, S., Shephard, N., and Chib, S. (1998). Stochastic Volatility: Likelihood Inference and Comparison with Arach Models. *Review of Economic Studies*, 65:361–394.
- Kraft, D. and Engle, R. (1982). Autoregressive Conditional Heteroscedasticity in Multiple Time Series. Mimeo, Dept.of Economics, University of California, San Diego.
- Krolzig, H.-M. (1997). *Markov Switching Vector Autoregressions. Modelling, Statistical Inference and Application to Business Cycle Analysis*. Springer-Verlag, Berlin.
- Lenk, P. J. and DeSarbo, W. S. (2000). Bayesian Inference for Finite Mixtures of Generalized Linear Models with Random Effects. *Psychometrika*, 65:93–119.
- Liesenfeld, R. and Richard, J. (2003). Univariate and Multivariate Stochastic Volatility Models: Estimation and Diagnostics. *Journal of Empirical Finance*, 10:505–531.

- Loddo, A., Ni, S., and Sun, D. (2011). Selection of Multivariate Stochastic Volatility Models via Bayesian Stochastic Search. *Journal of Business and Economic Statistics*, 29(3):342–355.
- Lucas, R. (1976). Econometric Policy Evaluation: A Critique. In Lucas, R., editor, *Studies in Business-Cycle Theory*. MIT Press, Cambridge (MA).
- Mosconi, R. and Seri, R. (2006). Non Causality in Bivariate Binary Time Series. *Journal of Econometrics*, 132:379–407.
- Nirei, M. and Sushko, V. (2010). Speculative Dynamics and Currency Crash Risk. Mimeo, University of California, Santa Cruz, November.
- Pelletier, D. (2006). Regime Switching for Dynamic Correlations. *Journal of Econometrics*, 131:445–473.
- Philipov, A. and Glickman, M. (2006). Multivariate Stochastic Volatility via Wishart Processes. *Journal of Economics and Business Statistics*, 24:313–328.
- Pérez-Rodríguez, J. (2006). The Euro and Other Major Currencies Floating Against the U.S. Dollar. *Atlantic Economic Journal*, 34(4):367–384.
- Robert, C. P. (2001). *The Bayesian Choice, 2nd ed.* Springer-Verlag, New York.
- Sahu, S., Dey, D., and Branco, D. (2003). A New Class of Multivariate Skew Distributions with Applications to Bayesian Regression Models. *Canadian Journal of Statistics*, 31:129–150.
- Silvennoinen, A. and Teräsvirta, T. (2008). Multivariate GARCH Models. In Andersen, T., Davis, R., Kreiss, J.-P., and Mikosch, T., editors, *Handbook of Financial Time Series*. Springer-Verlag, New York.
- So, M. and Choi, C. (2008). A Multivariate Threshold Stochastic Volatility Model. *Mathematics and Computers in Simulation*, 79:306–317.
- So, M., Li, W., and Lam, K. (2002). A Threshold Stochastic Volatility Model. *Journal of Forecasting*, 21:473–500.
- So, M. P., Lam, K., and Li, W. (1998). A Stochastic Volatility Model with Markov Switching. *Journal of Business and Economic Statistics*, 18:429–446.
- Tai, S. (2007). Market Integration and Contagion: Evidence from Asian Emerging Stock and Foreign Exchange Markets. *Emerging Markets Review*, 8(4):264–283.
- Tanner, M. (1993). *Tools for Statistical Inference (Lecture Notes in Statistics 67)*. Springer-Verlag, New York.
- Tanner, M. and Wong, W. (1987). The Calculation of Posterior Distributions by Data Augmentation. *Journal of the American Statistical Association*, 82:528–550.

- Taylor, S. (1986). *Modelling Financial Time Series*. Chichester, Wiley.
- Taylor, S. (1994). Modelling Stochastic Volatility. *Mathematical Finance*, 4:183–204.
- Van Bergeijk, P. and Berk, J. (2001). European Monetary Union, the Term Structure, and the Lucas Critique. *Kyklos*, (4):547–556.
- Wilfling, B. (2009). Volatility Regime-Switching in European Exchange Rates Prior to Monetary Unification. *Journal of International Money and Finance*, 3(2):240–270.
- Yu, J. and Meyer, R. (2006). Multivariate Stochastic Volatility Models: Bayesian Estimation and Model Comparison. *Econometric Reviews*, (2-3):361–384.