Multivariate Stochastic Volatility

William J. McCausland *Université de Montréal, CIREQ and CIRANO

Shirley Miller †Université de Montréal

Denis Pelletier ‡North Carolina State University

Current version: February 20, 2012

Abstract

We propose a new technique for the analysis of multivariate stochastic volatility models, based on efficient draws of volatility from its conditional posterior distribution. It applies to models with several kinds of cross-sectional dependence. Full VAR coefficient and covariance matrices give cross-sectional volatility dependence. Mean factor structure allows conditional correlations, given states, to vary in time. The conditional return distribution features Student's t marginals, with asset-specific degrees of freedom, and copulas describing cross-sectional dependence. We draw volatility as a block in the time dimension and one-at-a-time in the cross-section. Following McCausland (2012), we use close approximations of the conditional posterior distributions of volatility blocks as Metropolis-Hastings proposal distributions. We illustrate using daily return data for ten currencies. We report results for univariate stochastic volatility models and two multivariate models.

1 Introduction

Multivariate volatility models are a powerful inferential tool. By featuring different kinds of dynamic cross-sectional dependence among multiple asset returns, they can capture many different stylized facts.

It is well known that asset return volatility varies over time, changing in response to news and revised expectations of future performance. It tends to cluster, so that large price changes tend to be followed by other large changes. Volatility is not independent across markets and assets, and this cross-sectional dependence is time-varying. Cross-sectional

*Mailing address: Département de sciences économiques, C.P. 6128, succursale Centre-ville, Montréal QC H3C 3J7, Canada. e-mail: william.j.mccausland@umontreal.ca. Web site: www.cirano.qc.ca/∼mccauslw.

†e-mail: shirley.miller.lira@umontreal.ca.

‡Mailing address: Department of Economics, Campus Box 8110, North Carolina State University, Raleigh, 27695-8110, USA. e-mail: denis_pelletier@ncsu.edu. Web site: http://www4.ncsu.edu/∼dpellet.

correlations increase substantially in periods of high market volatility, especially in bear markets. The distribution of returns is heavy tailed compared with the normal distribution, even when one conditions on current market conditions. There is an asymmetric relation between price and volatility changes known as the "leverage effect": increases in volatility are associated more with large decreases in price than with large increases.

Multivariate volatility models that can capture these stylized facts are in high demand in finance given their many important applications, especially in modern portfolio management. Learning about the joint distribution of asset returns is a key element for the construction, diversification, evaluation and hedging of portfolios. Accurate estimation of the covariance matrix of multiple asset returns allows the investor to timely identify opportunities or risks associated with a particular portfolio. It is important to track changes in correlations to assess the risk of a portfolio, especially during periods of market stress. Financial crises usually have a strong impact on correlation. Markets tend to behave as one during big crashes — as the risk of some assets increases, investors wish to sell other risky investments. The result is more highly correlated returns. The pessimistic conclusion is that diversification is least effective at reducing risk at the very times when risk is highest. An awareness of this fact avoids false optimism.

As with univariate volatility models, there are two main types of multivariate volatility models: observation-driven and parameter-driven. In observation-driven models, volatility is a deterministic function of observed variables, which allows straightforward evaluation of the likelihood function. This advantage has made the observation-driven GARCH model and its extensions very popular for univariate problems.

In parameter-driven volatility models, known as stochastic volatility (SV) models, volatility is a latent stochastic process. Jacquier, Polson, and Rossi (1994) and Geweke (1994) give evidence suggesting that SV models are more realistic. They are also more natural discrete time representations of the continuous time models upon which much of modern finance theory is based. Unfortunately, computation of the likelihood function, which amounts to integrating out latent states, is difficult. However, since the introduction of Bayesian Markov chain Monte Carlo (MCMC) methods by Jacquier, Polson, and Rossi (1994) for univariate SV models, inference for these models has become much more feasible. These methods require the evaluation of the joint density of returns, states and parameters, which is straightforward. In addition, simulation methods for Bayesian inference make exact finite sample inference possible.

This paper focuses on multivariate stochastic volatility (MSV) models, which are parameter-driven. For a literature review of multivariate GARCH type models, which are observation-driven, see Bauwens, Laurent, and Rombouts (2006). We propose new MCMC methods for Bayesian analysis of MSV models, based on efficient draws of volatility from its conditional posterior distribution.

There are many different types of MSV models. In Section 2, we describe a MSV model that encompasses several special cases of interest and compare it to other models. Two difficulties arise when we extend volatility models to the multivariate case. First, the conditional variance of returns given states must be positive definite at every point in time. Second, there is a severe trade-off between parsimony and flexibility. As the number of assets increases, the number of potential parameters increases quickly, leading to a danger of overfitting. Reining in the number of parameters forces the modeler to make choices, and much of the difference between MSV model specifications reflects a choice about how to do this. This has implications on which stylized facts can be captured by the model.

We show that our estimation approach is quite flexible and we do not rely much on any special structure for the MSV model considered. It applies to models with several kinds of cross-sectional dependence. We can specify full first order VAR coefficient and covariance matrices for the evolution of volatilities. We can include a mean factor structure, which allows conditional return correlations, given asset and factor volatilities, to vary over time, and for these correlations to covary with variances. We can also model cross-sectional conditional return dependence given latent asset and factor volatilities using copulas. Copulas allow one to represent a multivariate distribution in a very flexible way by decoupling the choice of marginal distributions — which can be different from each other — from the choice of the dependence structure. Copulas have been used in multivariate GARCH-type models, but to our knowledge, this is the first study to introduce copulas in MSV models.

We introduce a new prior for correlation matrices, which we use in the context of Gaussian copulas. It is based on a geometric interpretation of correlation coefficients. The prior is a first step towards a model for time varying correlations where assets are exchangeable, avoiding a problem with models based on the Cholesky decomposition – their predictions are not invariant to the arbitrary choice of how to order assets.

We allow heavy-tailed returns. In our applications, we use Student's t marginals, but this is not an essential choice, and we don't rely on data augmentation to obtain conditional Gaussianity, unlike with many models using Student's t distributions. In general, we allow the marginal distribution to vary by asset, which in our applications translates to asset-specific degrees of freedom parameters. We also depart from the usual assumption of Gaussian factors and allow Student's t factors.

Different MCMC methods have been proposed for inference in MSV models and sometimes they are quite model specific. The estimation technique proposed by Chib, Nardari, and Shephard (2006) (CNS) is one of the most popular, especially when analyzing a large number of asset returns. The CNS model includes factors in mean, heavy tailed errors for returns, and jumps. Factor volatilities and the volatilities of the idiosyncratic components of returns are conditionally independent given parameters. Factors are Gaussian.

An important feature of the CNS procedure is sampling the factor loading matrix and the latent factors as a single block. This is more numerically efficient than using separate blocks to draw factor loadings and factors. The procedure exploits the conditional independence of volatilities to draw all volatilities and some associated parameters as a single block, using the procedure proposed by Kim, Shephard, and Chib (1998) (KSC) for univariate SV models.

The procedure in Kim, Shephard, and Chib (1998) is an example of the auxiliary mixture approach to inference in state space models, whereby non-linear or non-Gaussian state space models are first transformed into linear models and then the distribution of the transformed error is approximated by a mixture of Gaussian distributions. The mixture can be dealt with using data augmentation — adding mixture component indicators yields a linear Gaussian model when one conditions on them. The transformation is model specific, but many other models have yielded to this approach. Some relevant articles are Chib, Nardari, and Shephard (2002) and Omori, Chib, Shephard, and Nakajima (2007) for other univariate SV models; Stroud, Müller, and Polson (2003) for Gaussian, but non-linear, state space models with state dependant variances; Frühwirth-Schnatter and Wagner (2006) for state space models with Poisson counts; and Frühwirth-Schnatter and Früwirth (2007) for logit and multinomial logit models.

The approximation of the transformed error distribution as a mixture can be corrected by reweighting, as in Kim, Shephard, and Chib (1998) or by an additional Metropolis accept-reject, as in Stroud, Müller, and Polson (2003), Frühwirth-Schnatter and Wagner (2006) and Frühwirth-Schnatter and Früwirth (2007).

CNS draw log volatilities, component indicators and some parameters based on the approximate transformed model. These Metropolis-Hastings updates preserve an approximate posterior distribution implied by the approximate model. All other updates of unknown quantities preserve the exact posterior distribution. Thus the stationary distribution of a sweep through all the blocks is neither the exact nor the approximate posterior distribution. We cannot expect the method to be simulation consistent.

McCausland (2012) proposed an alternative procedure to draw all latent states in univariate state space models as a block, preserving their exact conditional posterior distribution. This *HESSIAN* method is fast and numerically efficient and does not require data augmentation. It can be used to draw joint blocks of states and parameters. It is based on a non-Gaussian proposal distribution that captures some of the departure from Gaussianity of the conditional posterior distribution of the states. The HESSIAN method uses routines to compute derivatives of the log measurement density at a point, but is not otherwise model specific.

While the HESSIAN method is only for univariate states, we can apply it to draw volatilities as a single block in the time dimension but one-at-a-time in the cross-section dimension. We will see that the conditional distribution of one state sequence, given the others, parameters and data, can be seen as the conditional posterior distribution of states in a univariate state space model. So, following McCausland (2012), we obtain very close approximations to these conditional posterior distributions, which we use as proposal distributions. We are also able to draw a single volatility sequence, together with some of its associated parameters, as a single block. Because of strong dependence between volatilities and these parameters, the result is higher numerical efficiency.

To apply the HESSIAN method in this way, we require only that the multivariate state sequence be a Gaussian first-order vector autoregressive process and that the conditional distribution of the observed vector depend only on the contemporaneous state vector. This requirement is satisfied for a wide variety of state space models, including but not limited to multivariate stochastic volatility models.

In Section 3, we describe in detail our methods for posterior simulation. In Section 4, we validate the correctness of our proposed algorithm using a test of program correctness similar to that proposed by Geweke (2004). In Section 5, we present an empirical application using a data set of daily returns of foreign exchange rates and compare the results of different specifications of the MSV model with the results for univariate SV models. Finally, in Section 6, we conclude and outline some possible extensions.

2 The Model

This section describes the most general discrete-time MSV model considered in this paper, and identifies some special cases of interest. We compare it to other specifications in the literature. We also describe prior distributions used in our empirical applications. Table 1 describes all of the model's variables. The notation is similar to that in Chib, Nardari, and Shephard (2006).

There are p observed return sequences, q factors and $m = p + q$ latent log volatility states. The conditional distribution of the factor vector $f_t = (f_{1t}, \ldots, f_{qt})$ and the return vector $r_t = (r_{1t}, \ldots, r_{pt})$ given the contemporaneous state vector α_t is given by

$$r_t = Bf_t + V_t^{1/2}\epsilon_{1t}, \quad f_t = D_t^{1/2}\epsilon_{2t},$$

or alternatively

$$y_t = \begin{bmatrix} r_t \\ f_t \end{bmatrix} = \begin{bmatrix} V_t^{1/2} & BD_t^{1/2} \\ 0 & D_t^{1/2} \end{bmatrix} \epsilon_t, \tag{1}$$

where B is a $p \times q$ factor loading matrix, $V_t = \mathrm{diag}(\exp(\alpha_{1t}), \ldots, \exp(\alpha_{pt}))$ and $D_t = \mathrm{diag}(\exp(\alpha_{p+1,t}), \ldots, \exp(\alpha_{p+q,t}))$ are matrices of idiosyncratic and factor volatilities, and $\epsilon_t = (\epsilon_{1t}^\top, \epsilon_{2t}^\top)^\top$ is an vector of innovations, specified below, in terms of parameters ν and R.

Given parameters $\bar{\alpha}$, A and Σ, the state is a Gaussian first order vector autoregression, given by

$$\alpha_1 \sim N(\bar{\alpha}, \Sigma_0), \qquad \alpha_{t+1}|\alpha_t \sim N((I - A)\bar{\alpha} + A\alpha_t, \Sigma), \tag{2}$$

where the derived parameter Σ_0 is chosen to make the state sequence stationary:

$$\mathrm{vec}\,\Sigma_0 = (I_{m^2} - A \otimes A)^{-1}\mathrm{vec}\,\Sigma.$$

See (Hamilton, 1994, p.265) for details on computing the marginal variance Σ_0.

Symbol	dimensions	description
$\bar{\alpha}$	$m \times 1$	mean of state α_t
A	$m \times m$	coefficient matrix for α_t
Σ	$m \times m$	variance of state innovation
B	$p \times q$	factor loading matrix
ν	$m \times 1$	vector of degrees of freedom parameters
R	$m \times m$	Gaussian copula parameter
ϵ_t	$m \times 1$	period t return/factor innovation
α_t	$m \times 1$	period t state
r_t	$p \times 1$	period t return vector
f_t	$q \times 1$	period t factor
y_t	$m \times 1$	$(r_t^\top, f_t^\top)^\top$

Table 1: Table of symbols

We assume the conditional independence relationships implied by the following joint density decomposition:

$$\begin{aligned}\pi(\bar{\alpha}, A, \Sigma, \nu, B, R, \alpha, f, r) = \pi(\bar{\alpha}, A, \Sigma, \nu, B)\pi(R) \\ \cdot \pi(\alpha_1|\bar{\alpha}, A, \Sigma) \prod_{t=1}^{n-1} \pi(\alpha_{t+1}|\alpha_t, \bar{\alpha}, A, \Sigma) \\ \cdot \prod_{t=1}^{n} \left[\pi(f_t|\alpha_t)\pi(r_t|B, R, f_t, \alpha_t)\right].\end{aligned}$$

We specify the distribution of $\epsilon_t = (\epsilon_{1t}, \ldots, \epsilon_{mt})$ by providing marginal distributions, which may differ, and a copula function describing dependence. See Patton (2009) for an overview of the application of copulas in the modelling of financial time series and Kolev, dos Anjos, and de M. Mendez (2006) for a survey and contributions to copula theory.

The marginal distribution of ϵ_{it} is given by the cumulative distribution function (cdf) $F_\epsilon(\epsilon_{it}|\theta_i)$. Let $\pi(\epsilon_{it}|\theta_i)$ be its density function. Sklar (1959) provides a theorem on the relationship between marginal distributions, joint distributions and a *copula function*. It states that if $F(\epsilon_1, \ldots, \epsilon_m)$ is an m-dimensional cdf with marginals $F_1(\epsilon_1), \ldots, F_m(\epsilon_m)$, then there exists a unique copula function C such that F that can be written as:

$$F(\epsilon_1, \ldots, \epsilon_m) = C(F_1(\epsilon_1), \ldots, F_m(\epsilon_m)).$$

A copula function is a cdf on $[0,1]^m$ with marginal distributions that are uniform on $[0,1]$. Conversely, if $\epsilon = (\epsilon_1, \ldots, \epsilon_m)$ is a random vector with cdf F and continuous marginal cdfs F_i, $i = 1, \ldots, m$, then the copula of ϵ, denoted C, is the cdf of $(u_1, \ldots, u_m)$, where u_i is the probability integral transform of ϵ_i: $u_i = F_i(\epsilon_i)$. The distribution of the u_i is uniform on $[0,1]$. Thus

$$C(u_1, \ldots, u_m) = F(F_1^{-1}(u_1), \ldots, F_m^{-1}(u_m)).$$

In this paper, we assume Student's t marginals with asset-specific degrees of freedom. This allows for fat tails. However, the Student-t cdf could be replaced by another one and most of the derivations presented below would still hold. We choose a Gaussian copula with variance matrix

$$R = \begin{bmatrix} R_{11} & 0 \\ 0 & I_q \end{bmatrix},$$

where R_{11}, and thus R, are correlation matrices. One could replace the Gaussian copula with another, and the derivations below could be modified accordingly. However, there would be a computational cost. We take advantage of the fact that the derivatives of a log Gaussian density are non-zero only up to second order.

We denote the Gaussian copula with correlation matrix R as C_R:

$$C_R(u_1, \ldots, u_m) = \Phi_R(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_m)).$$

Here Φ denotes the standard univariate Gaussian cdf and ϕ, its density. Φ_R and ϕ_R denote the cdf and density of the m-variate Gaussian distribution with mean zero and covariance R. Then the multivariate density of ϵ_t is the product of the Gaussian copula density and the Student-t marginal density functions:

$$\pi_\epsilon(\epsilon_t|\theta) = c_R(F_\epsilon(\epsilon_{1t}|\theta_1), \ldots, F_\epsilon(\epsilon_{mt}|\theta_m)) \prod_{i=1}^{m} \pi(\epsilon_{it}|\theta_i), \tag{3}$$

where

$$c_R(u_1, \ldots, u_m) = \frac{\partial^{(m)} C_R(u_1, \ldots, u_m)}{\partial u_1 \cdots \partial u_m} = \frac{\phi_R(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_m))}{\prod_{i=1}^{m} \phi(\Phi^{-1}(u_i))}.$$

Letting $x_i \equiv \Phi^{-1}(u_i)$, $i = 1, \ldots, m$ and $x \equiv (x_1, \ldots, x_m)$, we can write

$$\log c_R(u_1, \ldots, u_m) = -\frac{1}{2}(\log |R| + \log(2\pi) + x^\top (R^{-1} - I))x. \tag{4}$$

We use the notation π_ϵ here instead of the generic π to clarify that it is the density function of ϵ_t. We can now write the conditional density of y_t given α_t, B, ν and R as

$$\pi(y_t|\alpha_t, B, \nu, R) = \pi_\epsilon\left(\begin{bmatrix} V_t^{-1/2}(r_t - Bf_t) \\ D_t^{-1/2} f_t \end{bmatrix} \middle| \nu, R\right) \prod_{i=1}^{m} \exp(-\alpha_{it}/2). \tag{5}$$

2.1 Alternative MSV models

As mentioned before, different MSV model specifications reflect, to a large extent, different restrictions chosen by the modeller to balance flexibility and parsimony. In our model, we can impose restrictions on the parameters governing the marginal distribution of volatility in (2), the parameters governing the conditional distribution of returns and factors given volatility, equation (1), or both. These choices have different implications for the stylized facts that a MSV model can capture.

First let us consider restrictions on the marginal distribution of volatilities. At one extreme, giving the most flexibility for volatility dynamics, we can specify A and Σ in equation (2) as full matrices. At another extreme, we can impose prior independence among log volatilities by specifying diagonal matrices for A and Σ. This can be much less computationally demanding, which makes it especially attractive when the number of volatilities to estimate is large. Several intermediate possibilities are possible, including the relatively parsimonious specification in Section 2.2, where A and Σ are not diagonal, but have $O(m)$ free elements.

We now consider cross-sectional dependence arising from the conditional distribution of returns given parameters and volatilities, marginal of latent factors. For comparison purposes, it will be helpful to write out the conditional variance of returns given returns and factor volatilities:

$$\mathrm{Var}[r_t|\alpha_t] = V_t^{1/2} R_{11} V_t^{1/2} + B D_t B^\top. \tag{6}$$

In the case where we have no factors, $q = 0$, then the second term disappears. The conditional variance varies in time, but the conditional correlation R_{11} is constant. Models with constant correlations have been studied by Harvey, Ruiz, and Shephard (1994), Danielsson (1998), Smith and Pitts (2006) and So, Li, and Lam. (1997). Other authors, including Yu and Meyer (2006), Philipov and Glickman (2006), Gourieroux (2006), Gourieroux, Jasiak, and Sufana (2004), Carvalho and West (2006) and Asai and McAleer (2009), consider models in which the return innovation correlation is time-varying, which is more realistic. However, as the number of assets increases, the estimation of a separate time varying correlation matrix becomes very challenging. Furthermore, when the dynamics of correlation and volatility are modelled separately, it is difficult to capture the empirical regularity that correlation and volatility covary.

Introducing latent factors in mean is another way to introduce time-varying correlations. Factors in mean models exploit the idea that co-movements of asset returns are driven by a small number of common underlying variables, called factors. The factors are typically modelled as univariate SV processes. Usually, factor MSV models give R_{11} as the identity matrix, in which case $\mathrm{Var}(r_t|\alpha_t) = V_t + B D_t B^\top$. The main attractions of mean factor models is that they are parsimonious, they lead to time varying conditional correlations and they have a natural link with the arbitrage pricing theory (APT), an influential theory of asset pricing. APT holds that the expected return of a financial asset can be modelled as a linear function of various factors. In addition, the mean factor structure allows the conditional correlations and conditional variances to covary. This is an important feature for portfolio analysis, especially when there are turbulent periods. See Longin and Solnik (2001) and Ang and Chen (2002) for empirical studies showing the positive correlation of the conditional variances and conditional correlations. Given all these characteristics, factor MSV models have become very popular in the literature, and different versions have been proposed. The basic model assumed normal returns, a constant factor loading matrix and zero factor mean. See, for example, Jacquier, Polson, and Rossi (1995), Pitt and Shephard (1999) and Aguilar and West (2000). Other studies proposed some extensions to the basic structure such as jumps in the return equation and heavy-tailed returns (Chib, Nardari, and Shephard (2006)), time varying factor loading matrices and regime-switching factors (Lopes and Carvalho (2007)) or first-order autoregressive factors (Han (2006)). See Chib, Omori, and Asai (2009) for a brief description and comparison of the different types of MSV models mentioned. Allowing for heavy tails in the distributions of returns is desirable because empirical evidence has shown that returns present higher conditional kurtosis than a Gaussian distribution does.

If we compare these models to the one described at the beginning of this section, we notice that the MSV model specification that we work with is fairly general and incorporates some other specifications as special cases. In its most general version, without parameter restrictions, the model allows for cross-sectional volatility dependence. It allows time-varying conditional correlations through the specification of a mean factor structure. It also incorporates cross-sectional conditional return dependence through copulas. The conditional variance matrix of returns in equation (6) is time-varying. The conditional correlation matrix is also time varying, and covaries with the conditional variances.

We can impose some parameter restrictions and obtain some interesting special cases:

- Independent states in cross section: A and Σ diagonal.
- Conditionally independent returns given factors and states: R diagonal.
- No factors: $q = 0$. In this case, the conditional variance-covariance matrix of returns is given by $\mathrm{Var}(r_t|\alpha_t) = V_t^{1/2} R_{11} V_t^{1/2}$ which is still time-varying but the conditional correlation matrix will be R_{11} which is constant.

2.2 Prior Distributions

2.2.1 Prior for $\bar{\alpha}$, A, Σ, ν, and B

We now describe a prior for a low dimensional specification of $\bar{\alpha}$, A, Σ, ν, and B.

We parameterize A and Σ in the following parsimonious way:

$$\Sigma = (\mathrm{diag}(\sigma))^2 + \begin{bmatrix} \beta\beta^T & 0 \\ 0 & 0 \end{bmatrix}, \quad A = \mathrm{diag}(\lambda) + \begin{bmatrix} (1/p)\delta\iota_p^\top & 0 \\ 0 & 0 \end{bmatrix}.$$

where σ and λ are $m \times 1$ vectors, β and δ are $p \times 1$ vectors and ι_p is the $p \times 1$ vector of ones.

We organize the parameters associated with each series i (a return for $i = 1, \ldots, p$ or a factor for $i = p+1, \ldots, m$) as

$$\theta_i = \begin{cases} (\bar{\alpha}_i, \tanh^{-1}(\lambda_i), \tanh^{-1}(\lambda_i + \delta_i), \log\sigma_i, \beta_i/\sigma_i, \log\nu_i, B_{i1}, \ldots, B_{iq})^\top, & 1 \le i \le p, \\ (\tanh^{-1}(\lambda_i), \log\sigma_i, \log\nu_i)^\top, & p+1 \le i \le m, \end{cases}$$

and organize the vector of all these parameters as $\theta = (\theta_1^\top, \dots, \theta_m^\top)^\top$.

We suppose that the θ_i are *a priori* independent, multivariate normal, and that the parameters have the prior means and variances given in Table 2. For each $i = 1, \dots, m$,

Parameter	mean	variance
$\bar{\alpha}_i$	-11.0	$2^2 = 4$
$\tanh^{-1}(\lambda_i)$	2.1	$(0.25)^2 = 0.0625$
$\tanh^{-1}(\lambda_i + \delta_i)$	2.3	$(0.25)^2 = 0.0625$
$\log \sigma_i$	-2.0	$(0.5)^2 = 0.25$
β_i/σ_i	0.0	$(0.5)^2 = 0.25$
$\log \nu_i$	3.0	$(0.5)^2 = 0.25$

Table 2: Parameter means and variances of prior distributions.

the correlation coefficient between σ_i and $\tanh^{-1}(\lambda_i)$ is -0.8. All other correlations are zero. The prior probability that the A matrix is such that α is not stationary is close enough to zero that we have not seen an example in prior simulations.

2.2.2 Prior for R

We can interpret the correlations in the $p \times p$ correlation matrix R as the cosines of angles between vectors in $\mathbb{R}^l$, where $l \geq p$. There are p vectors, one for each asset, and the $\binom{p}{2}$ angles between distinct vectors give the various correlations.

We reparameterize the information in R. The new parameter is an $p \times l$ matrix V whose rank is p and whose rows have unit Euclidean length. The rows of V give p points on the surface of the unit l-dimensional hypersphere centred at the origin. In putting a prior on V, we induce a prior on $R = VV^\top$. It is easy to see that $VV^\top$ is a $p \times p$ symmetric positive definite matrix with unit diagonal elements. In other words, it is a full rank correlation matrix. Conversely, for any full correlation matrix R and any $l \geq p$, there is an $p \times l$ real matrix V with rows of unit length and rank p such that $VV^\top = R$: take the Cholesky decomposition $R = LL^\top$ and let $V = [L \quad 0_{p,l-p}]$.

We choose a prior such that the rows v_i of V are independent and identically distributed. This ensures that the prior does not depend on how the assets are ordered. We could relax this to exchangeable v_i and retain this advantage. This kind of invariance is difficult to achieve if one specifies a prior on the Cholesky decomposition of the correlation matrix. A disadvantage of the V parameterization is that the number of non-zero elements of V is lp, while the number of non-zero elements of the Cholesky factor is $p(p+1)/2$. Another issue is that V is not identified. However, since $VV^\top$ is identified, this is not a problem.

We will call the vector $(1, 0, ..., 0)$ in R^l the north pole of the hypersphere. Let $\zeta_i \equiv \cos^{-1}(V_{i1})$, the angle between v_i and the north pole. We specify a marginal density $\pi(\zeta_i)$ and let the conditional distribution $v_i|\zeta_i$ be uniform on the set of points on the surface of the unit hyperphere at an angle of ζ_i from the north pole. This set is the surface of an $(l-1)$ dimensional hypersphere of radius $\sin\zeta_i$.

This gives the following density for v_i on the unit l-dimensional hypersphere:

$$\pi(v_i) = \pi(\zeta_i) 2 \frac{\pi^{(l-1)/2}}{\Gamma(\frac{l-1}{2})} \sin^{l-2}\zeta_i.$$

In our applications, we use $\zeta_i/\pi \sim \mathrm{Be}(4,4)$.

3 Posterior inference using MCMC

We use MCMC methods to simulate the posterior distribution, with density $\pi(\bar{\alpha}, A, \Sigma, \nu, B, R, \alpha, f|r)$. We use a multi-block Gibbs sampler. The result is an ergodic chain whose stationary distribution is the target distribution. The sequence of steps in a single sweep through the blocks is

1. For $i = 1, \ldots, m$, update (θ_i, α_i) as described in 3.1, preserving the conditional posterior distribution $\theta_i, \alpha_i | \theta_{-i}, \alpha_{-i}, R, B_{-i}, f, r$, where α_{-i} is the vector of all state sequences except the i'th and θ_{-i} is the vector of all parameter values in θ except those in θ_i.

2. Update (B, f) as described in 3.2, preserving the conditional distribution $B, f|\theta, \alpha, R, r$.

3. Update f as described in 3.3, preserving the conditional distribution $f|\theta, \alpha, R, B, r$.

4. Update R as described in 3.4, preserving the conditional distribution $R|\theta, \alpha, B, f, r$.

In the following subsections, we describe each of these steps.

3.1 Draw of θ_i, α_i

We draw (θ_i, α_i) as a single Metropolis-Hastings block. Drawing a volatility sequence together with its associated parameters in one block is more efficient than drawing them separately because of their posterior dependence.

Our proposal of (θ_i, α_i) consists of a random walk proposal of θ_i^* followed by a (conditional) independence proposal of α_i^* given θ_i^*. This gives a joint proposal that we accept or reject as a unit. The acceptance probability is given by

$$\min\left(1, \frac{\pi(\theta_i^*)\pi(\alpha_i^*|\theta_i^*, \theta_{-i}, \alpha_{-i})\pi(y_t|\alpha_i^*, \alpha_{-i}, \theta_i^*, \theta_{-i}, R)}{\pi(\theta_i)\pi(\alpha_i|\theta, \alpha_{-i})\pi(y_t|\alpha, \theta, R)} \cdot \frac{g(\alpha_i^*|\theta_i^*, \theta_{-i}, \alpha_{-i}, R)}{g(\alpha_i|\theta, \alpha_{-i}, R)}\right),$$

where $g(\alpha_i^*|\theta_i^*, \theta_{-i}, R)$ is an independence (it does not depend on α_i) conditional proposal density for α_i^* given θ_i^*.

A key issue for independence proposals is the specification of the proposal density. To obtain high numerical efficiency for the draw of a vector with thousands of observations, we need an extremely close approximation. We will see that the conditional posterior distribution of α_i has the form of the target distributions approximated in McCausland (2012). These approximations are very close, and we will exploit them here.

3.1.1 Draw of $\theta_i^* | \theta_i, \alpha_{-i}^*, \omega$

We use a random walk Metropolis proposal for θ_i^*. The random walk $(\theta_i^* - \theta_i)$ is Gaussian with mean zero and covariance matrix Ξ. We obtain Ξ using an adaptive random walk Metropolis algorithm, described in Vihola (2011), during a burn-in period — the random walk proposal variance is adjusted after each draw to track a target acceptance probability. We use the final value of Ξ at the end of the burn-in period as the proposal covariance matrix for all future draws. Thus our posterior simulator is a true Markov chain after the burn-in period and so standard MCMC theory applies to the retained posterior sample.

3.1.2 Draw of $\alpha_i^* | \theta_i^*, \omega$

We now discuss the draw of the conditional proposal $\alpha_i^* | \theta_i^*, \theta_{-i}, \alpha_{-i}, R$ using the HESSIAN method in McCausland (2012).

The HESSIAN method is for simulation smoothing in state space models with univariate Gaussian states and observable vectors that are not necessarily Gaussian. It involves a direct independence Metropolis-Hastings update of the entire sequence of states as a single block. The proposal is a much closer approximation of the target distribution than is any multivariate Gaussian approximation. The result is a Metropolis-Hastings update that is not only tractable, but very numerically efficient. One can also update states jointly with parameters by constructing joint proposal distributions, as we do here.

Drawing states as a block is much more efficient than one-at-a-time draws in the usual case where the posterior autocorrelation of states is high. Adding parameters to the block leads to even higher numerical efficiency when there is strong posterior dependence between parameters and states. The HESSIAN method does not require data augmentation or model transformations, unlike auxiliary mixture sampling methods, where the model is transformed and augmented so that conditioning on auxiliary variables yields a linear Gaussian state space model. The auxiliary mixture approach has been used for univariate state space models by Omori, Chib, Shephard, and Nakajima (2007) and Kim, Shephard, and Chib (1998) . Approximating distributions of the transformed model by mixtures of Gaussian random variables results in slightly incorrect posterior draws. In some cases, this is corrected using reweighting or an additional accept/reject step. We have seen that in Chib, Nardari, and Shephard (2006), some blocks update the true posterior and some blocks update the approximate (mixture approximation) posterior. The stationary distribution is neither the approximate distribution nor the true distribution, and it is not clear to us how one could compensate for the error. Draws from the HESSIAN approximate distribution are exact, in the sense that draws of α_i^* are consistent with the evaluation of the proposal density used to compute the Metropolis-Hastings acceptance probability.

The HESSIAN method uses an approximation $g(\alpha|y)$ of $\pi(\alpha|y)$ for univariate models in which $\alpha \sim N(\bar{\Omega}^{-1}\bar{c}, \bar{\Omega})$, with $\bar{\Omega}$ tridiagonal and $\pi(y|\alpha) = \prod_{t=1}^{n} \pi(y_t|\alpha_t)$. One needs to specify $\bar{\Omega}$, the precision, and $\bar{c}$, the co-vector, and provide routines to compute the first five derivatives of $\log \pi(y_t|\alpha_t)$. The approximation $g(\alpha|y)$ is so close to $\pi(\alpha|y)$ that we can use it as a proposal distribution to update the entire sequence $\alpha = (\alpha_1, \ldots, \alpha_n)$ as a block.

Here states are multivariate, but we can draw state sequences one at a time in the cross-sectional dimension, using approximations of the conditional distribution of each state sequence α_i given the rest of the states (α_{-i}), parameters and data. The conditional density we need to approximate is

$$\pi(\alpha_i|\alpha_{-i}, y) \propto \pi(\alpha_i|\alpha_{-i}) \prod_{t=1}^{n} \pi(y_t|\alpha_t).$$

In Appendix A, we show that $\alpha_i|\alpha_{-i} \sim N((\bar{\Omega}^{(i)})^{-1}\bar{c}^{(i)}, \bar{\Omega}^{(i)})$, where the co-vector $\bar{c}^{(i)}$ is a $n \times 1$ vector and the precision $\bar{\Omega}^{(i)}$ is a tridiagonal $n \times n$ matrix, as required by the HESSIAN method. We also describe there how to compute the elements of $\bar{\Omega}^{(i)}$ and $\bar{c}^{(i)}$ in terms of the elements of $\bar{\Omega}$ and $\bar{c}$.

We just need to compute five derivatives of $\log \pi(y_t|\alpha_{it}, \alpha_{-i,t})$ with respect to α_{it}. We do not need to write down the complete analytical expressions of these derivatives, we just need to evaluate them at a point. To do this, we use automatic routines to combine derivatives of primitive functions according to Faa di Bruno's rule, which is a generalization of the chain rule to higher derivatives. It allows us to take two vectors of derivative values and call a function that returns a vector of derivatives of a composite function. Appendix B describes the Faa di Bruno formula and how we use it to evaluate five derivatives of $\log \pi(y_t|\alpha_{it}, \alpha_{-i,t})$.

3.2 Draw of (B, f)

In this block, we update B and f simultaneously in a way that preserves the posterior distribution of B and f given everything else but does not change the value of the matrix-vector products Bf_t. Adding this block improves the posterior mixing of the poorly identified scale of the B matrix. At the same time, it is fairly cheap computationally, because the Bf_t do not change.

We first draw a random $q \times q$ matrix Λ. The diagonal elements are iid, with $n\Lambda_{ii} \sim \chi^2(n)$, and the non-diagonal elements are zero. With probability 1/2, we form proposals $B^* = B\Lambda$, $f_t^* = \Lambda^{-1} f_t$, $t = 1, \ldots, n$ and with complementary probability, we form $B^* = B\Lambda^{-1}$, $f_t^* = \Lambda f_t$, $t = 1, \ldots, n$. In the first case, we accept with probability

$$\min\left(1, |\Lambda|^{-(n-p)} \frac{\pi(B^*) \prod_{t=1}^{n} \pi(f_t^*|\alpha, \nu)}{\pi(B) \prod_{t=1}^{n} \pi(f_t|\alpha, \nu)}\right),$$

and in the second case, we accept with probability

$$\min\left(1, |\Lambda|^{(n-p)} \frac{\pi(B^*) \prod_{t=1}^n \pi(f_t^*|\alpha,\nu)}{\pi(B) \prod_{t=1}^n \pi(f_t|\alpha,\nu)}\right).$$

The factors $|\Lambda|^{-(n-p)}$ and $|\Lambda|^{(n-p)}$ are products of the Jacobian matrices for the multiplicative transformations of n vectors f_t and p rows of B.

3.3 Draw of f

We draw each f_t from its conditional posterior distribution using a random walk proposal. Because the random walk involves only two function evaluations, it is quite cheap computationally. We use a proposal variance matrix $(2.38)^2(B^\top V_t^{-1}B + D_t^{-1})^{-1}$. The matrix $(B^\top V_t^{-1}B + D_t^{-1})^{-1}$ is a crude but cheap approximation of the conditional posterior variance of f_t, obtained by setting $\nu_i = \infty$, $i = 1, \ldots, m$, and $R = I$. The scaling factor $(2.38)^2$ comes from Gelman, Roberts, and Gilks (1996), and it is optimal when the target distribution is univariate Gaussian.

3.4 Draw of R

We draw the rows of V one-at-a-time. We use a random walk M-H proposal to update the row vector v_i. It is a random walk on the l-dimensional unit hypersphere: the direction of the walk is uniform and the angle of the walk has some arbitrary distribution. Let d be the direction vector, normalized so that it has unit length. To draw the proposal v_i^*:

1. Draw the angle ζ_i between the proposal v_i^* and the current state. We use $\zeta_i/\pi \sim \mathrm{Be}(1, 199)$.

2. Draw the direction d from the uniform distribution on the unit l-dimensional hypersphere[1].

3. Compute $d_\perp$, the projection of d onto the hyperplane perpendicular to v_i:

$$d_\perp = d - \frac{v_i d}{||v_i||^2} v_i$$

4. Compute:

$$v_i^* = \cos\zeta_i \cdot v_i + \sin\zeta_i \cdot \frac{d_\perp}{||d_\perp||}$$

[1] We can draw from a uniform distribution on a unit hypersphere by drawing a spherically symmetric normal random vector of the same dimension, and dividing it by its length.

5. Accept with probability

$$\min\left(1, \frac{\pi(f,r|\alpha,\theta,B,R^*,f)\pi(\zeta_i^*)}{\pi(f,r|\alpha,\theta,B,R,f)\pi(\zeta_i)}\frac{\zeta_i}{\zeta_i^*}\right).$$

4 Getting it Right

Here we perform a computational experiment with artificial data to put the implementation of our methods to the test. We use a simulation strategy similar to that proposed by Geweke (2004) for testing the correctness of posterior simulators and detecting any analytical and coding errors there may be. This procedure replaces the common exercise of generating a single artificial data set using known values of the parameters, applying a simulation method to these data and verifying that the "true value" falls in a region of high posterior probability.

Like the approach of Geweke (2004), our approach is based on the simulation of the joint distribution of parameters, states, factors and data. We use a single simulator, a Gibbs sampler that alternates between updates of the posterior distribution, described in the previous section, and draws of returns given parameters, states and factors, described in Appendix C. If the simulator works correctly, then the marginal distribution of the parameters must agree with the specified prior distributions. We can test a wide range of implications of this condition.

This formal approach is a more stringent way to verify the correctness of posterior simulators, as not all errors lead to obviously incorrect results. Reasonable but incorrect results are worse than obvious errors, because they can mislead. The test applied here can discriminate much more effectively between correct code and alternatives with minor coding errors. Also, simulation results often provide clues to the source of any errors.

Here in detail is how we generate a sample from the joint distribution of $\bar{\alpha}$, A, Σ, B, ν, R, α, f and r. The first draw $(\bar{\alpha}^{(1)}, A^{(1)}, \Sigma^{(1)}, B^{(1)}, \nu^{(1)}, R^{(1)}, \alpha^{(1)}, f^{(1)}, r^{(1)})$ comes directly from the model. See Appendix C for a description of how to draw from $\pi(r|\bar{\alpha}, A, \Sigma, B, \nu, R, \alpha, f)$. Then, we draw subsequent values by iterating the following Gibbs blocks:

1. For $i = 1, \ldots, m$, update θ_i, α_i as described in Section 3.1.
2. For $t = 1, \ldots, n$, update f_t as described in Section 3.3.
3. Update B and $(f_1, \ldots, f_n)$ as described in Section 3.2.
4. Update R as described in Section 3.4.
5. Update r as described in Appendix C.

We obtain a sample $\{\theta_i^{(j)}\}_{j=1}^J$ of size $J = 10^8$ for $i = 1, \ldots, m$. We construct, for $i = 1, \ldots, m$ and $j = 1, \ldots, J$ the vectors

$$z^{(i,j)} \equiv L_i^{-1}(\theta_i^{(j)} - \mu_i),$$

where μ_i is the prior mean and L_i is the lower Cholesky factor of the prior variance of θ_i. If the $\theta_i^{(j)}$ are truly multivariate Gaussian with variance $L_i L_i^\top$, the elements of $z^{(i,j)}$ are iid $N(0,1)$. The vectors $z^{(i,j)}$ have length $K_i = 6+q$ for $i = 1,\ldots,p$ and length $K_i = 3$ for $i = p+1,\ldots,m$. Since the $z^{(i,j)}$, $i = 1,\ldots,m$, are independent, we have $\sum_{i=1}^m z_i^\top z_i \sim \chi^2((6+q)p+3q)$.

We construct the following sample frequencies for quantiles $Q = 0.1, 0.3, 0.5, 0.7, 0.9$, return and factor indices $i = 1,\ldots,m$, and parameter indices $k = 1,\ldots,K_i$

$$\hat{I}_{ik}^{(Q)} = \frac{1}{J}\sum_{j=1}^{J} 1\left(z_k^{(i,j)} \leq \Phi^{-1}(Q)\right),$$

as well as the sample frequencies

$$\hat{I}_{0k}^{(Q)} = \frac{1}{J}\sum_{j=1}^{J} 1\left(\sum_{i=1}^{m}(z^{(i,j)})^\top z^{(i,j)} \leq F^{-1}(Q)\right),$$

where F is the cdf of the χ^2 distribution with $(6+q)p+3q$ degrees of freedom.

Standard results for laws of large numbers and central limit theorems for ergodic chains apply, so we should observe sample frequencies close to Q. Table 3 shows the sample frequencies $\hat{I}_{ik}^{(Q)}$ and their estimated numerical errors $s_{ik}^{(Q)}$, obtained using the method of batch means. We observe that for all cases, the sample frequencies are very similar to their respectively Q values. This fails to cast doubt on the correctness of the implementation of the proposed algorithm.

5 Empirical Results

In this section we apply our methods to historical exchange rate data. We describe the data and report estimation results for various models.

5.1 Data

We analyze daily returns of 10 currencies relative to the US dollar: the Swiss Franc (CHF), Euro (EUR), Australian Dollar (AUD), New Zealand Dollar (NZD), Mexican Peso (MXN), Brazil Real (BRL), British Pound (GBP), Canadian Dollar (CAD), Japanese Yen (JPY) and Singapore Dollar (SGD). The exchange rates are the noon spot rate obtained from the Federal Reserve Bank of New York. The sample covers the period from January 5, 1999 to December 31, 2008. We compute the log returns of the exchange rates and remove returns for those days when one or more of the markets was closed, giving 2503 observations for each return series.

Table 4 presents some descriptive statistics. All series present excess kurtosis, but the magnitude varies from one currency to another, from around 2 for the Euro to about 27

i	k	$\hat{I}_{i,k}^{(0.1)}$	$s_{i,k}^{(0.1)}$	$\hat{I}_{i,k}^{(0.3)}$	$s_{i,k}^{(0.3)}$	$\hat{I}_{i,k}^{(0.5)}$	$s_{i,k}^{(0.5)}$	$\hat{I}_{i,k}^{(0.7)}$	$s_{i,k}^{(0.7}$	$\hat{I}_{i,k}^{(0.9)}$	$s_{i,k}^{(0.9)}$
1	1	0.1004	0.00068	0.3003	0.00084	0.5001	0.00077	0.6997	0.00063	0.8997	0.00036
1	2	0.0997	0.00018	0.2994	0.00030	0.4994	0.00034	0.6996	0.00029	0.8998	0.00018
1	3	0.1000	0.00015	0.3000	0.00025	0.4998	0.00026	0.6997	0.00024	0.8996	0.00017
1	4	0.0997	0.00016	0.2996	0.00027	0.4998	0.00028	0.6998	0.00026	0.8998	0.00016
1	5	0.1001	0.00016	0.3002	0.00027	0.5001	0.00027	0.7000	0.00027	0.9001	0.00018
1	6	0.1002	0.00015	0.3001	0.00024	0.5001	0.00027	0.7003	0.00028	0.9005	0.00017
1	7	0.0993	0.00052	0.2993	0.00098	0.4998	0.00110	0.7000	0.00090	0.9002	0.00043
1	8	0.1002	0.00050	0.3008	0.00096	0.5010	0.00110	0.7003	0.00092	0.9004	0.00048
2	1	0.1008	0.00089	0.3005	0.00100	0.5007	0.00098	0.7004	0.00078	0.9002	0.00041
2	2	0.0999	0.00018	0.3003	0.00034	0.5001	0.00033	0.7001	0.00031	0.9003	0.00016
2	3	0.0999	0.00016	0.2996	0.00032	0.4995	0.00037	0.6999	0.00033	0.9000	0.00020
2	4	0.1000	0.00019	0.3001	0.00034	0.5005	0.00029	0.7004	0.00025	0.9003	0.00014
2	5	0.0999	0.00016	0.2996	0.00033	0.4999	0.00038	0.7002	0.00029	0.8999	0.00017
2	6	0.1001	0.00017	0.3002	0.00033	0.5001	0.00033	0.7002	0.00024	0.9002	0.00015
2	7	0.1001	0.00043	0.3001	0.00085	0.4997	0.00106	0.6997	0.00084	0.9001	0.00045
2	8	0.0990	0.00043	0.2980	0.00088	0.4974	0.00116	0.6983	0.00091	0.8994	0.00050
3	1	0.0999	0.00009	0.3000	0.00015	0.4999	0.00016	0.6999	0.00013	0.8999	0.00009
3	2	0.1000	0.00010	0.3000	0.00015	0.4999	0.00017	0.7000	0.00015	0.9001	0.00011
3	3	0.1000	0.00009	0.3000	0.00016	0.5000	0.00018	0.6999	0.00016	0.8999	0.00009
4	1	0.1000	0.00008	0.2999	0.00015	0.4998	0.00015	0.6998	0.00015	0.8999	0.00009
4	2	0.1000	0.00010	0.3001	0.00013	0.4999	0.00015	0.7000	0.00013	0.9001	0.00009
4	3	0.0998	0.00009	0.2999	0.00014	0.5000	0.00015	0.7001	0.00015	0.9000	0.00009

Table 3: “Getting it right” sample quantiles

	Mean	Std. Dev.	Skewness	Excess Kurtosis
CHF	-2.56	10.81	-0.30	2.45
EUR	1.45	9.98	0.10	1.90
AUD	1.27	13.45	-0.88	16.66
NZD	0.83	13.40	-0.60	5.75
MXN	3.49	9.48	1.31	26.91
BRL	6.67	19.33	0.45	14.13
GBP	-1.19	8.97	-0.29	5.04
CAD	-2.22	8.80	-0.20	9.36
JPY	-2.05	10.35	-0.36	2.70
SGD	-1.44	4.75	-0.19	4.44

Table 4: Descriptive statistics of data: annualized mean, annualized standard deviation, skewness and excess kurtosis. The sample period is from January 5, 1999 to December 31, 2008.

for the Mexican Peso. Sample volatility varies a lot across currencies, with the Brazilian Real, and the Australian and New Zealand Dollars being the most volatile currencies. Although the sample statistics differ substantially across currencies, we can also observe some commonalities in Figure 1. This shows time plots of the 10 return series and we notice that all returns exhibit their most volatile episodes at the end of the sample, which corresponds to the financial crisis of 2008.

In Table 5 we show the sample correlation matrix for the entire period. Correlation coefficients vary from -0.9 to 0.8. The strongest negative correlation is for the pair (EUR,CHF) and the strongest positive correlation is for the pair (AUD,NZD). The MXN and BRL are the least correlated with the rest of currencies.

	CHF	EUR	AUD	NZD	MXN	BRL	GBP	CAD	JPY	SGD
CHF	1.000	-0.902	-0.372	-0.370	-0.052	0.010	-0.609	0.305	0.382	0.423
EUR	-0.902	1.000	0.516	0.499	-0.067	-0.123	0.692	-0.409	-0.266	-0.466
AUD	-0.372	0.516	1.000	0.822	-0.362	-0.319	0.500	-0.568	-0.018	-0.446
NZD	-0.370	0.499	0.822	1.000	-0.267	-0.253	0.487	-0.485	-0.032	-0.429
MXN	-0.052	-0.067	-0.362	-0.267	1.000	0.479	-0.152	0.296	-0.162	0.162
BRL	0.010	-0.123	-0.319	-0.253	0.479	1.000	-0.160	0.244	-0.087	0.213
GBP	-0.609	0.692	0.500	0.487	-0.152	-0.160	1.000	-0.392	-0.162	-0.399
CAD	0.305	-0.409	-0.568	-0.485	0.296	0.244	-0.392	1.000	0.005	0.348
JPY	0.382	-0.266	-0.018	-0.032	-0.162	-0.087	-0.162	0.005	1.000	0.362
SGD	0.423	-0.466	-0.446	-0.429	0.162	0.213	-0.399	0.348	0.362	1.000

Table 5: Sample daily correlation for the period from January 1999 to December 2008 (2503 observations)

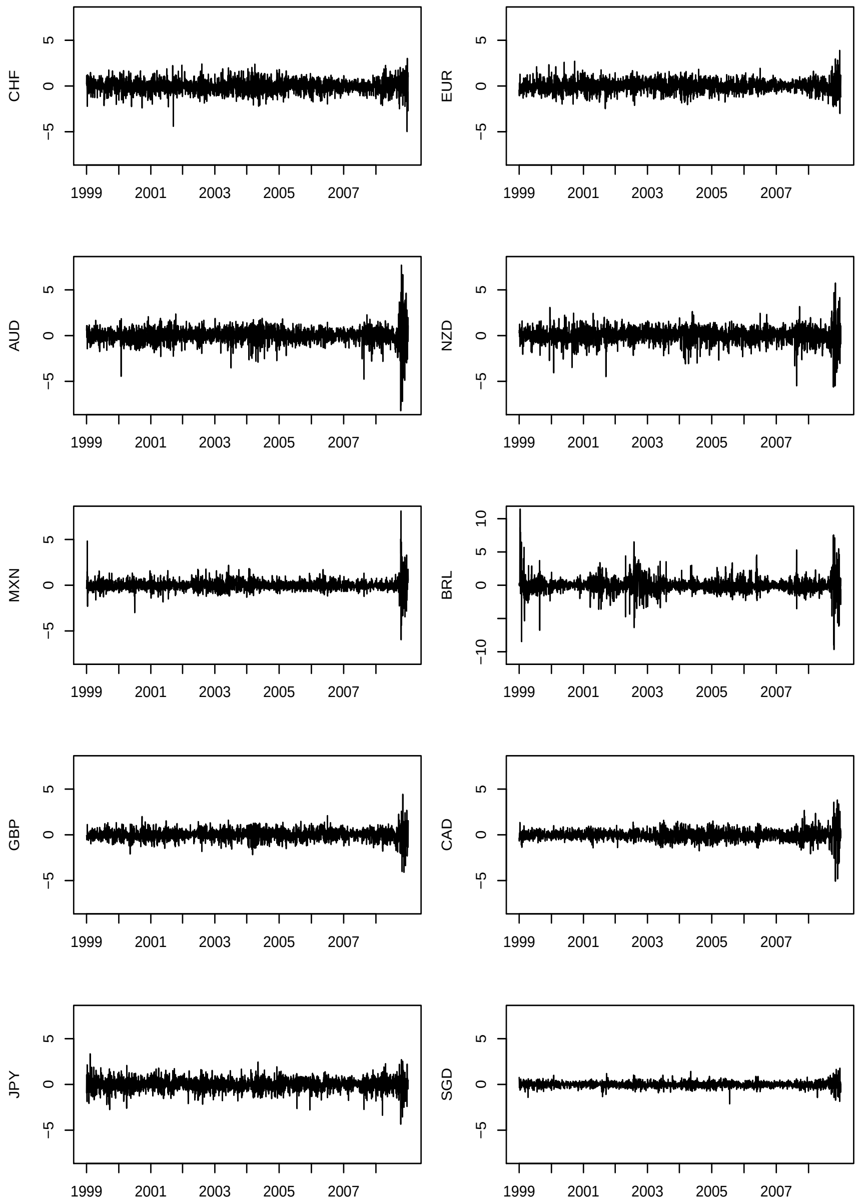

Figure 1: Time plots of daily returns series (in percentage). The sample period is from January 5, 1999 to December 31, 2008.

5.2 Estimation Results

We estimate three models: a model with independent currencies, each governed by a univariate SV model with Student's t innovations (SVt), a MSV model with no factors (MSV-q0) and a MSV model with one factor (MSV-q1). We use comparable priors in the three models and compare the posterior distribution of parameters, volatilities and correlations across models.

Figures 2 and 3 show posterior densities of the parameters of the volatility equation across currencies and models. These are computed in R using the default kernel density estimation method[2]. The solid line corresponds to the univariate SVt model, the dashed line to the MSV-q0 model and the dotted line to the MSV-q1 model. Tables 5 through 10 in Appendix D give posterior parameter means, standard deviations, numerical standard errors (NSE) for the mean, and relative numerical efficiency (RNE) for the mean. The NSE and RNE are are computed using the R library coda, using a time series method based on an estimate of the spectral density at 0.

For the SVt model, the A and Σ matrices are diagonal, so that $\bar{\alpha}_i, A_{ii}$ and σ_{ii} are the parameters of the i'th univariate SV model, $i = 1, \ldots, 10$. For the MSV-q0 and MSV-q1 models there are non-zero off-diagonal elements. In the SVt and MSV-q0 models, the α_{ti}, governed by the $\bar{\alpha}$, A and Σ matrices, are the only source of volatility, while in the MSV-q1 model they give the idiosyncratic volatility, the part of volatility not attributable to the common factor.

We observe that the posterior density of $\bar{\alpha}_i$ and A_{ii} for the MSV models is shifted left compared with the univariate SVt models for all the currencies except MXN and BRL, for which the three posterior densities of $\bar{\alpha}_i$ are very similar[3]. At the same time, the posterior densities of σ_{ii} are shifted right, relative to the univariate models[4]. With respect to the parameter ν_i, for half of the currencies the posterior distribution looks very similar, while for the other half there are some differences, but without a clear pattern.

Passing from the univariate SVt models to the multivariate MSV-q0 model, we obtain in most cases a lower mean, lower persistence and higher volatility of idiosyncratic volatility. The MSV-q0 model allows returns to be conditionally correlated but still with currency-specific degrees of freedom. We see that the posterior mean of the degrees of freedom parameter varies from one currency to another in line with what we observed in the descriptive statistics.

In the MSV-q1 model, there is both idiosyncratic and factor volatility. Figure 4 show a plot of the factor volatility and Table 11 presents the posterior parameter distribution statistics of the factor volatility equation. In our model the Bf_t are identified but not B and

[2]The default algorithm disperses the mass of the empirical distribution function over a regular grid of at least 512 points and then uses the fast Fourier transform to convolve this approximation with a discretized version of the kernel and then uses linear approximation to evaluate the density at the specified points.

[3]A_{ii} denotes the AR(1) coefficient in the volatility equation (2).

[4]In the case of the MSV models, σ_{ii} represents the square root of the diagonal elements of the variance matrix Σ correspondent to the volatility equation. It measures the volatility of volatility.

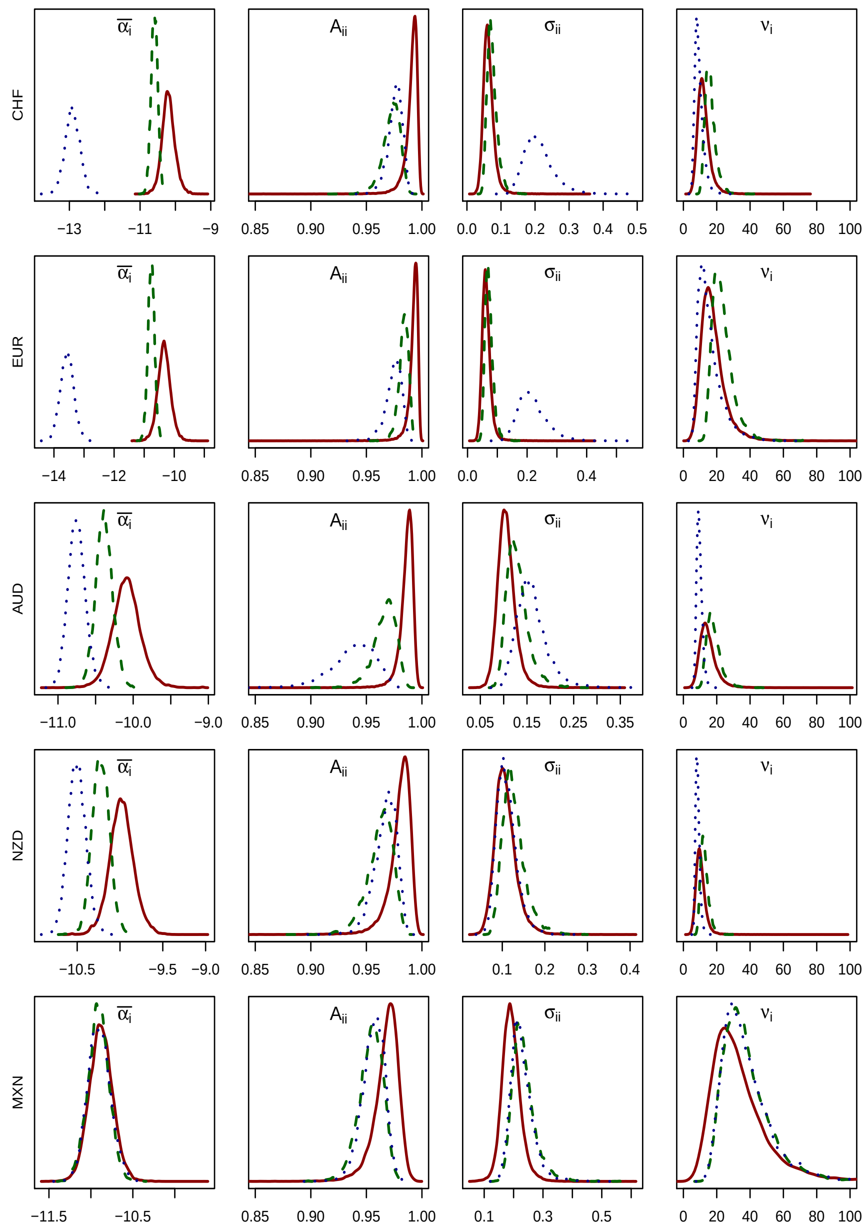

Figure 2: Comparison of posterior parameter distributions. The solid line corresponds to the SVt model, the dashed line to the MSV-q0 model and the dotted line to the MSV-q1 model. Posterior densities are based on 45,000 draws, after discarding 6,000. Densities are computed in R using the default kernel density estimation method.

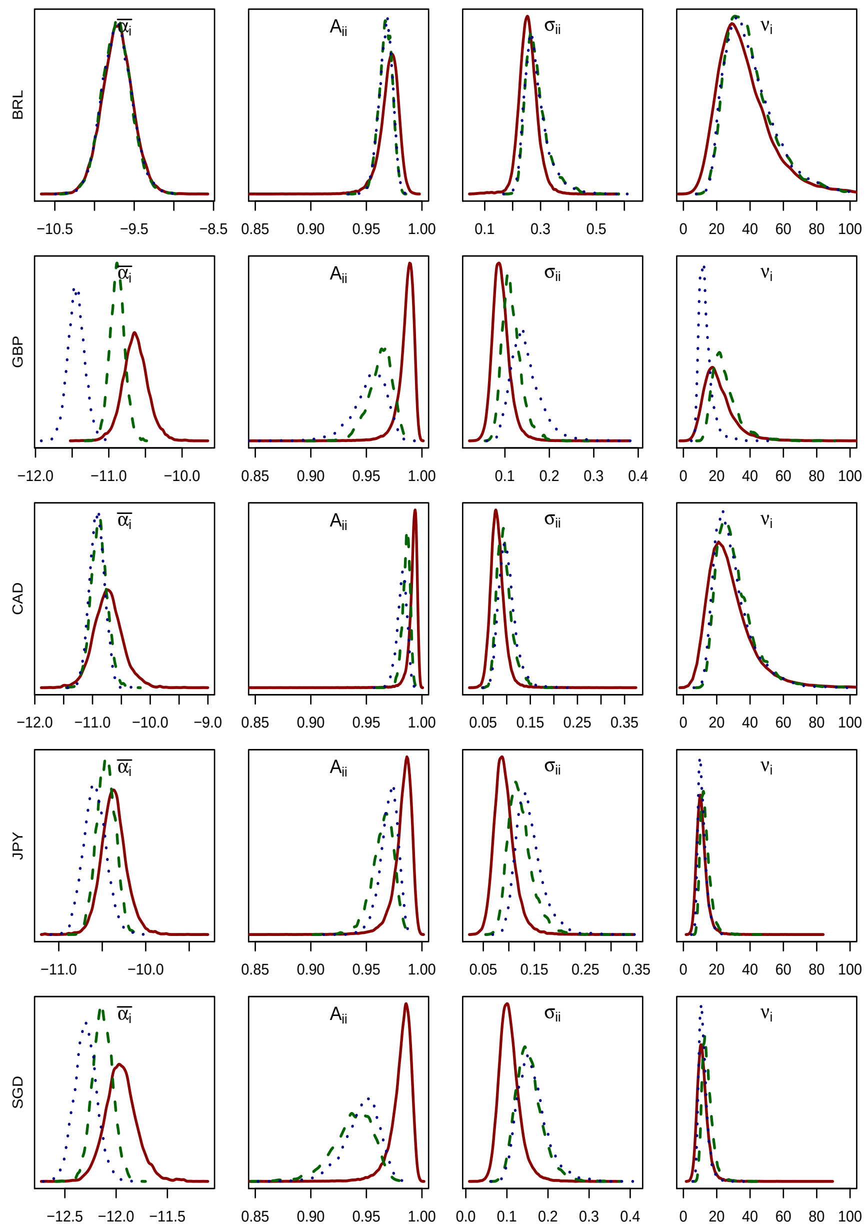

Figure 3: Comparison of posterior parameter distributions. The solid line correspond to the SVt model, the dashed line to the MSV-q0 model and the dotted line to the MSV-q1 model. Posterior densities are based on 45,000 draws, after discarding 6,000. Densities are computed in R using the default kernel density estimation method.

the f_t separately. The posterior distribution of B is thus quite sensitive to the priors for B and the parameters of the factors[5]. We set $\bar{\alpha}_{11} = 0$ to normalize the variance of the factor to one. Other normalization strategies are possible. Note that there are only two parameters to estimate for the factor volatility equation: $A_{11,11}$, the persistence parameter, and ν_{11} the factor volatility's degree of freedom. The posterior mean of $A_{11,11}$ is 0.99, indicating that the factor volatility is more persistent than the idiosyncratic volatilities. The posterior mean of ν_{11} is around 21, which suggest the conditional factor distribution is not much more fat-tailed than a Gaussian distribution.

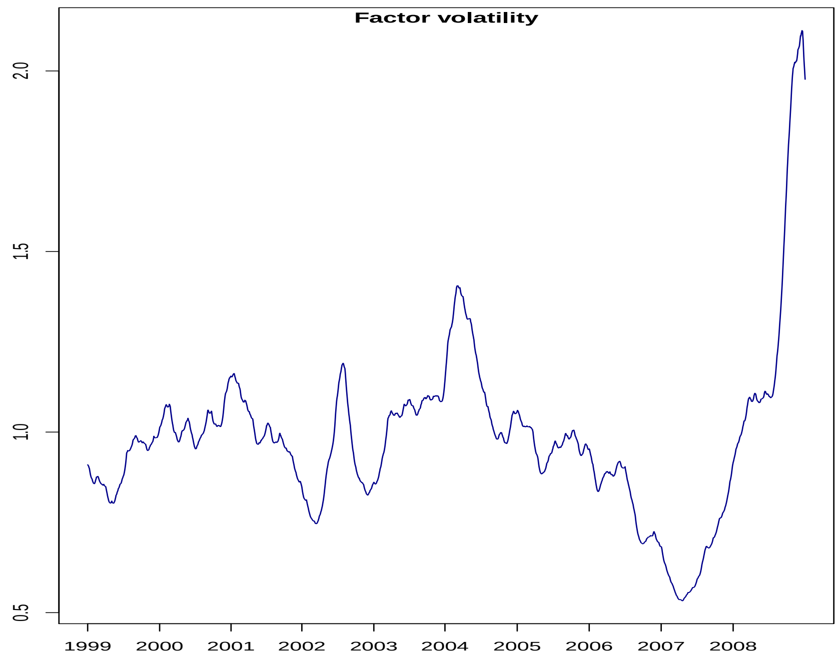

Figure 4: Time series plot of the posterior mean of the factor volatility of MSV-q1 model.

We calculate the time varying decomposition of variance into factor and idiosyncratic components and plot them in Figure 5. The solid line correspond to the factor component and the dashed line to the idiosyncratic component. We see that the factor is capturing most of the co-movement among the CHF, EUR and GBP currencies. The factor volatility contribution for CHF and EUR currencies is more than 80 percent for most of the period and for the GBP is slightly greater than 50 percent. For the rest of currencies, the idiosyncratic contribution is higher than the factor, specially for the case of BRL and MXN currencies where the factor contribution is close to zero. This is consistent with the low

[5]Table 12 present the posterior parameter of the elements of the B matrix. The relative numerical efficiency of these parameters are low but the efficiency is improved for the Bf_t.

correlation between these two currencies and the rest. Thus, this suggest that the factor can be identified as an "European" factor, as the CHF, EUR and GBP currencies are the three European currencies in our sample and the factor seems to capture the shocks that affect this region.

These results also explain why we see a big move to the left in the posterior distribution of the $\bar{\alpha}_i$ (mean idiosyncratic log volatility) for CHF, EUR and GBP. In Figure 6 we present the time series plot of the annualized total volatility for the 10 currencies analyzed obtained with MSV-q0 and MSV-q1. We see that estimates are similar across models except for the three currencies with the higher factor contributions, where we notice that for the MSV-q1 model the idiosyncratic volatility has higher mean and is more persistent, compared with the MSV-q0 model.

We now analyze estimates of correlations between currencies, across models. As we have discussed in previous sections, the MSV-q0 model, with no factors, implies a time varying variance matrix, but a time invariant correlation matrix; while the MSV model with factors implies that both the variance and the correlation matrices of returns are time-varying. Tables 14 and 15 in Appendix D.4 show the posterior mean of the R_{11} matrix for MSV-q0 and MSV-q1, respectively. In the case of MSV-q0, R_{11} is the conditional correlation matrix of the returns, $\text{Corr}(r_t|\alpha_t)$. For the MSV-q1 model, we show in Table 16 the average across the time dimension of the posterior mean of the corresponding $\text{Corr}(r_t|\alpha_t)$ matrix. If we compare these results with those of MSV-q0 and the sample correlation matrix showed in Table 5 we can see that the estimate of the correlation matrix for the MSV-q1 model is closer to the corresponding sample correlation matrix. The estimated conditional correlation matrix for the MSV-q0 model agrees with respect to sign but the magnitudes of the correlation estimates are much smaller.

6 Conclusions

We have introduced a new approach for estimating multivariate stochastic volatility models. This approach uses a numerically efficient method to draw volatilities as a block in the time dimension and one-at-a-time in the cross sectional dimension. The proposed algorithm is flexible, allowing different specifications and types of dependence. We can model time-varying conditional correlation matrices by incorporating factors in the return equation, where the factors are independent SV processes with Student's t innovations. Furthermore, we can incorporate copulas to allow conditional return dependence given volatility, allowing different Student's t marginals to capture return heterogeneity. We have tested the correctness of our implementation of the proposed method using procedures similar to those suggested by Geweke (1994).

We apply the proposed method to an exchange rate data set and compare posterior distributions of parameters and volatility with those obtained with univariate SV models with Student's t innovations. We estimate two multivariate models, one in which we do

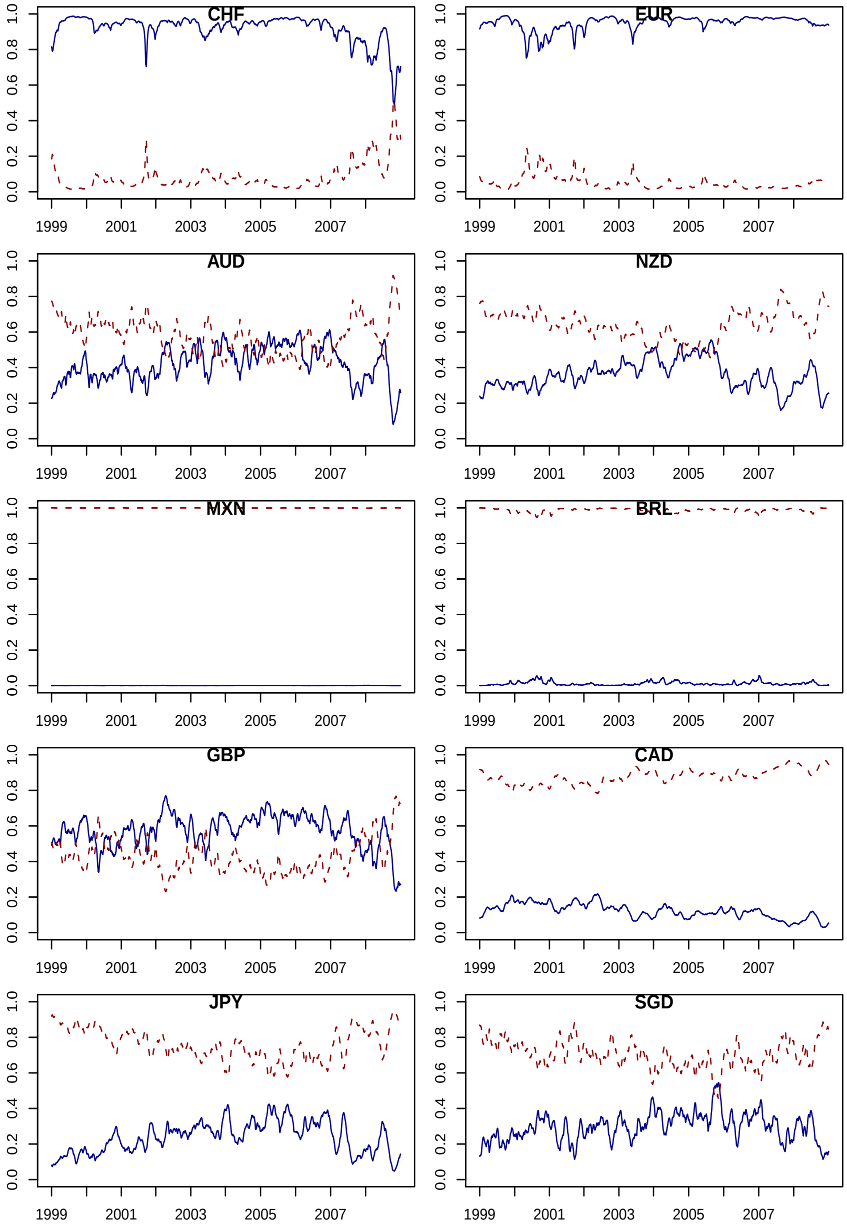

Figure 5: Time varying proportion of conditional variance of returns ($\mathrm{Var}(r_t|\alpha_t) = V_t^{1/2} R_{11} V_t^{1/2} + BD_t B^\top$) explained by the idiosyncratic and factor components in MSV-q1 model. The solid line corresponds to the factor contribution and the dashed line corresponds to the idiosyncratic contribution.

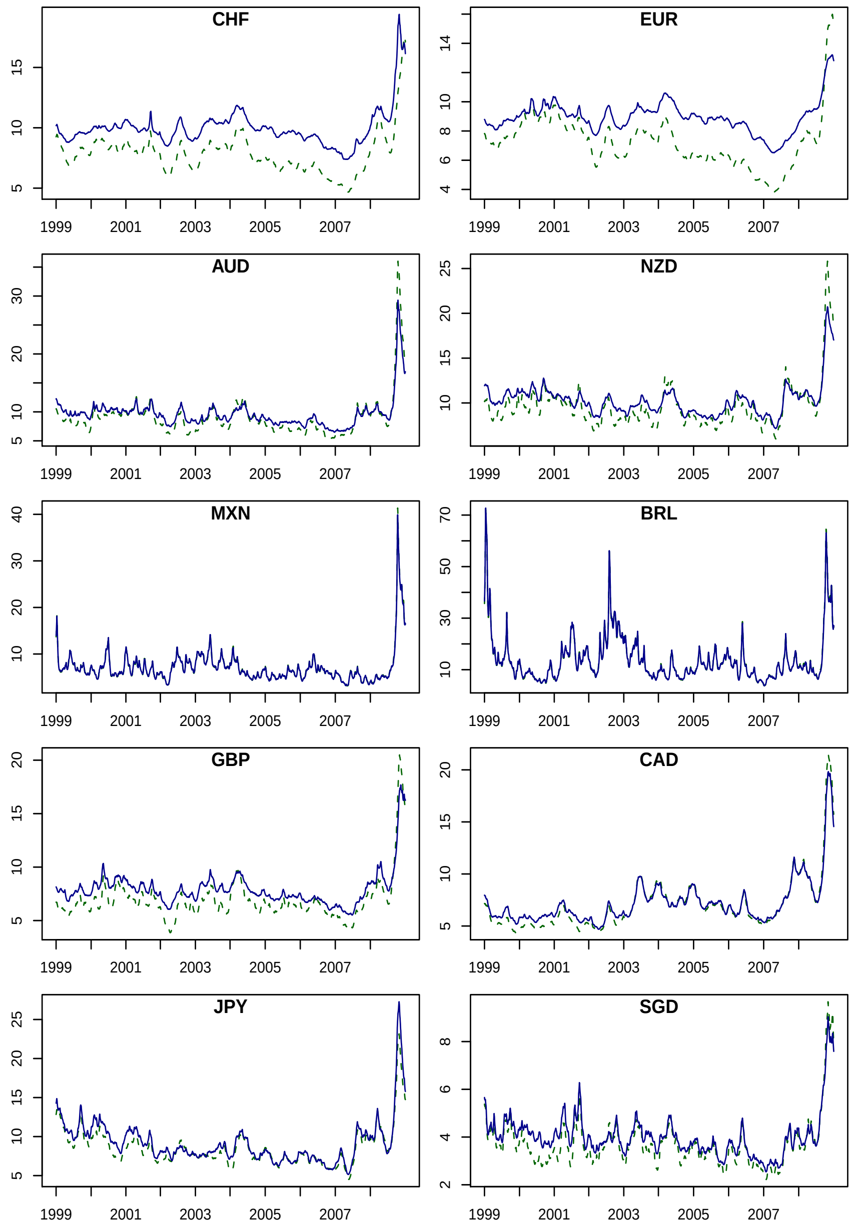

Figure 6: Comparison of annualized posterior mean of volatility. The dashed line corresponds to MSV-q0 and the solid line to MSV-q1.

not include factors and another in which we introduce one factor. We find that for most of the currencies, the multivariate approach with no factors gives a lower mean, lower persistence and higher volatility of volatility than the univariate model. The factor in the factor multivariate model seems to be a kind of "European" factor, as it is mainly capturing co-movement of three European currencies. The factor volatility is more persistent than the idiosyncratic volatilities. It would be interesting to introduce additional factors to see if we can capture other co-movements.

Applying the HESSIAN method one-at-a-time in the cross section only requires that the multivariate state sequence be a Gaussian first-order vector autoregressive process and that the conditional distribution of the observed vector depend only on the contemporaneous state vector. This requirement is satisfied for a wide variety of state space models, including but not limited to multivariate stochastic volatility models.

Using the HESSIAN method overcomes two disadvantages of the auxiliary mixture approach. First, it is less model specific — it does not require the researcher to find a suitable transformation for the model at hand. Second, it is exact — we do not need to correct for mixture approximation, using reweighting or additional Metropolis-Hastings steps, or settle for simulators that are not simulation consistent.

We hope to extend this work to compute marginal likelihoods and to compare the results from different specifications. Also, we hope to extend the model to incorporate leverage effects.

A Computing $\bar{\Omega}^{(i)}$ and $\bar{c}^{(i)}$

We show here how to compute $\bar{\Omega}^{(i)}$ and $\bar{c}^{(i)}$, the conditional precision and covector of the conditionally Gaussian distribution $\alpha_i|\alpha_{-i}$. We start by defining $\bar{\Omega}$ and $\bar{c}$, the prior precision and covector of α. The precision $\bar{\Omega}$ is a $nm \times nm$ block band-diagonal matrix. We will use the notation $\bar{\Omega}_{st}$, $s, t = 1, \ldots, n$, to denote the $m \times m$ submatrix starting at row $(s-1)m+1$ and column $(t-1)m+1$. The non-zero submatrices are the diagonal blocks $\bar{\Omega}_{tt}$ and the off-diagonal blocks $\bar{\Omega}_{t,t+1}$ and $\bar{\Omega}_{t-1,t}$, given by

$$\begin{aligned}
\bar{\Omega}_{tt} &= \Sigma^{-1} + A^\top \Sigma^{-1} A, \quad t = 2, \ldots, n-1, \\
\bar{\Omega}_{11} &= \Sigma_0^{-1} + A^\top \Sigma^{-1} A, \\
\bar{\Omega}_{nn} &= \Sigma^{-1}, \\
\bar{\Omega}_{t,t+1} &= -A^\top \Sigma^{-1}, \quad t = 1, \ldots, n-1, \\
\bar{\Omega}_{t-1,t} &= -\Sigma^{-1} A, \quad t = 1, \ldots, n-1.
\end{aligned} \tag{7}$$

The co-vector is a $nm \times 1$ vector stacking n $m \times 1$ subvectors $\bar{c}_t$, given by:

$$\begin{aligned}
\bar{c}_t &= \Sigma^{-1}(I-A)\bar{\alpha} - A^\top \Sigma^{-1}(I-A)\bar{\alpha}, \quad t = 2, \ldots, n-1 \\
\bar{c}_1 &= \Sigma_0^{-1}\bar{\alpha} - A^\top \Sigma^{-1}(I-A)\bar{\alpha}, \\
\bar{c}_n &= \Sigma^{-1}(I-A)\bar{\alpha}.
\end{aligned} \tag{8}$$

We now derive the $n \times n$ precision $\bar{\Omega}^{(i)}$ and $n \times 1$ co-vector $\bar{c}^{(i)}$ of the conditional distribution $\alpha_i|\alpha_{-i}$. We know that the conditional density $\pi(\alpha_i|\alpha_{-i})$ is proportional to the joint density $\pi(\alpha)$. Matching coefficients of the first- and second-order monomial terms of $\log \pi(\alpha_i|\alpha_{-i})$ gives the non-zero elements

$$\bar{\Omega}^{(i)}_{tt} = (\bar{\Omega}_{tt})_{ii}, \quad \bar{\Omega}^{(i)}_{t,t+1} = \bar{\Omega}^{(i)}_{t+1,t} = (\bar{\Omega}_{t,t+1})_{ii}.$$

$$\bar{c}^{(i)}_t = (\bar{c}_t)_i - \sum_{j \neq i} \left[(\bar{\Omega}_{tt})_{ji}\alpha_{tj} + (\bar{\Omega}_{t,t+1})_{ji}\alpha_{t+1,j} + (\bar{\Omega}_{t-1,t})_{ji}\alpha_{t-1,j} \right].$$

B Computing $\log \pi(y_t|\alpha_t, \nu, B, R)$ and its derivatives with respect to α_{it}

Using equations (3), (4), and (5), we can write $\log \pi(y_t|\alpha_t, B, \nu, R)$ in the following way:

$$\log \pi(y_t|\alpha_t, \nu, B, R) = -\frac{1}{2}\left\{ \log|R| + \log 2\pi + x_t^\top (R^{-1} - I)x_t + \sum_{i=1}^m \left[\alpha_{it} + (\nu_i + 1)\log\left(1 + \frac{\epsilon^2_{it}}{\nu_i}\right)\right]\right\}$$
$$+ \sum_{i=1}^m \left[\log\Gamma\left(\frac{\nu_i+1}{2}\right) - \log\Gamma\left(\frac{\nu_i}{2}\right) - \frac{1}{2}\log(\nu_i\pi)\right],$$

where $x_t = (x_{1t}, \ldots, x_{mt})$ and for $i = 1, \ldots, m$,

$$x_{it} = \Phi^{-1}(u_{it}), \quad u_{it} = F_\epsilon(\epsilon_{it}|\nu_i)),$$

$$\epsilon_{it} = \begin{cases} \exp(-\alpha_{it}/2)(r_{it} - \sum_{j=1}^q B_{ij} f_{jt}), & i = 1, \ldots, p, \\ \exp(-\alpha_{it}/2) f_{i-p,t}, & i = p+1, \ldots, m. \end{cases}$$

We can evaluate $\log \pi(y_t|\alpha_t, B, \nu, R)$ as a function of α_{it} bottom up, evaluating the ϵ_{it} at α_{it}, then the u_{it} at ϵ_{it}, then the x_{it} at u_{it} then $\log \pi(y_t|\alpha_t, B, \nu, R)$ at ϵ_t and x_t.

We require five derivatives of $\log \pi(y_t|\alpha_t, B, \nu, R)$ with respect to α_{it}, evaluated at α_{it}. Because it is a multi-level compound function of the α_{it}, computing these derivatives in closed form would be extremely tedious and prone to error. Fortunately, we do not need to. Instead, we compute any values we need, bottom up, using Faà di Bruno's formula (B.5 below) at each step to compute derivatives of a compound function by combining derivatives of its component functions.

We proceed using the following steps.

1. Compute five derivatives of $\psi(\alpha_{it}) \equiv \log \pi_\epsilon(e^{-\alpha_{it}/2}\eta_{it}|\theta_i)$ with respect to α_{it} at α_{it}, as described in B.1.

2. Compute five derivatives of $x^\top(R^{-1} - I)x$ with respect to x_{it} at x_{it}, as described in B.2.

3. Compute five derivatives of x_{it} with respect to u_{it} at u_{it}, as described in B.3.

4. Compute five derivatives of u_{it} with respect to α_{it} at α_{it}, as described in B.4.

5. Use the Faà di Bruno formula, described in B.5, to compute five derivatives of x_{it} with respect to α_{it} at α_{it}. Inputs are the derivatives of x_{it} with respect to u_{it} at step 3 and the derivatives of u_{it} with respect to α_{it} at step 4.

6. Use the Faà di Bruno formula to compute five derivatives of $x^\top(R^{-1} - I)x$ with respect to α_{it} at α_{it}. Inputs are the derivatives of $x^\top(R^{-1} - I)x$ with respect to x_{it} at step 2 and the derivatives of x_{it} with respect to α_{it} at step 5.

7. Compute five derivatives of $\log \pi(y_t|\alpha_t, \theta, B, R)$ with respect to α_{it} at α_{it} directly using the derivatives at steps 1 and 6.

For convenience, we define

$$\eta_t = \begin{bmatrix} \eta_{1t} \\ \vdots \\ \eta_{mt} \end{bmatrix} = \begin{bmatrix} r_t - Bf_t \\ f_t \end{bmatrix},$$

B.1 Derivatives of $\psi(\alpha_{it})$ with respect to α_{it}

For the special case of Student's t F,

$$\pi_\epsilon(e^{-\alpha_{it}/2}\eta_{it}|v_i) = \frac{\Gamma(\frac{\nu_i+1}{2})}{\sqrt{\nu_i\pi}\Gamma(\frac{\nu_i}{2})}\left(1 + \frac{e^{-\alpha_{it}}\eta_{it}^2}{\nu_i}\right)^{-\frac{\nu_i+1}{2}}$$

$$\psi(\alpha_{it}) = \log\left[\frac{\Gamma(\frac{\nu_i+1}{2})}{\sqrt{\nu_i\pi}\Gamma(\frac{\nu_i}{2})}\right] - \frac{\nu_i+1}{2}\log(1+s_{it})$$

where $s_{it} \equiv e^{-\alpha_{it}}\eta_{it}^2/\nu_i$. Noting that $\partial s_{it}/\partial\alpha_i = -s_{it}$, we compute

$$\psi'(\alpha_{it}) = \frac{\nu_i+1}{2}\frac{s_{it}}{1+s_{it}}, \quad \psi''(\alpha_{it}) = -\frac{\nu_i+1}{2}\frac{s_{it}}{(1+s_{it}^2)},$$

$$\psi'''(\alpha_{it}) = \frac{\nu_i+1}{2}\frac{s_{it}(1-s_{it})}{(1+s_{it})^3}, \quad \psi^{(4)}(\alpha_{it}) = -\frac{\nu_i+1}{2}\frac{s_{it}(1-4s_{it}+s_{it}^2)}{(1+s_{it})^4},$$

$$\psi^{(5)}(\alpha_{it}) = \frac{\nu_i+1}{2}\frac{s_{it}(1-11s_{it}+11s_{it}^2-s_{it}^3)}{(1+s_{it})^5}.$$

B.2 Derivatives of $x^\top(I-R^{-1})x$ with respect to x_{it}

In this section we show how to compute partial derivatives of $\log c(u_1,\dots,u_m)$ with respect to the u_i. We can write

$$\begin{aligned}\log c_R(u_1,\dots,u_m) &= \log\phi_R(\Phi^{-1}(u_1),\dots,\Phi^{-1}(u_m)) - \sum_{i=1}^m \log\phi(\Phi^{-1}(u_i))\\ &= \frac{1}{2}|H| + \frac{1}{2}x^\top(I-R^{-1})x,\end{aligned}$$

where $x=(x_1,\dots,x_m)=(\Phi^{-1}(u_1),\dots,\Phi^{-1}(u_m))$.

The gradient and Hessian of $\log(c_R)$ with respect to u are

$$\frac{\partial \log c(u)}{\partial x} = (I-R^{-1})x,$$

$$\frac{\partial \log c(u)}{\partial x \partial x^\top} = I - R^{-1}.$$

All third order partial derivatives and higher are zero.

B.3 Derivatives of x_{it} with respect to u_{it}

We now use the relationship $\Phi(x_i)=u_i$ to compute derivatives of x_i with respect to u_i. Differentiating with respect to u_i gives $\phi(x_i)\frac{\partial x_i}{\partial u_i}=1$, and thus

$$\frac{\partial x_i}{\partial u_i} = \frac{1}{\phi(x_i)}.$$

Taking further derivatives gives

$$\begin{aligned}\frac{\partial^2 x_i}{\partial u_i} &= 2\pi e^{x_i^2} x_i,\\ \frac{\partial^3 x_i}{\partial u_i} &= (2\pi)^{3/2} e^{3x_i^2/2}(2x_i^2+1),\\ \frac{\partial^4 x_i}{\partial u_i} &= (2\pi)^2 e^{2x_i^2}(6x_i^3+7x_i),\\ \frac{\partial^5 x_i}{\partial u_i} &= (2\pi)^{5/2} e^{5x_i^2/2}(24x_i^4+46x_i^2+7).\end{aligned}$$

B.4 Derivatives of $F_\epsilon(e^{-\alpha_{it}/2}\eta_{it}|\theta_i)$

We describe here how to compute five derivatives of $F_\epsilon(e^{-\alpha_{it}/2}\eta_{it}|\theta_i)$ with respect to α_{it}. We write down the derivatives in terms of $\psi(\alpha_{it}) \equiv \log \pi_\epsilon(e^{-\alpha_{it}/2}\eta_{it}|\theta_i)$:

$$\frac{\partial F_\epsilon(e^{-\alpha_{it}/2}\eta_{it}|\theta_i)}{\partial \alpha_{it}} = \pi_\epsilon(e^{-\alpha_{it}/2}\eta_{it}|\theta_i)\left(-\frac{1}{2}e^{-\alpha_{it}/2}\eta_{it}\right) = -\frac{\eta_{it}}{2}e^{-0.5\alpha_{it}+\psi(\alpha_{it})}$$

Then

$$\frac{\partial^2 F_\epsilon(e^{-\alpha_{it}/2}\eta_{it}|\theta_i)}{\partial \alpha_{it}^2} = -\frac{\eta_{it}}{2}e^{-0.5\alpha_{it}+\psi(\alpha_{it})}[-0.5+\psi'(\alpha_{it})]$$

$$\frac{\partial^3 F_\epsilon(e^{-\alpha_{it}/2}\eta_{it}|\theta_i)}{\partial \alpha_{it}^3} = -\frac{\eta_{it}}{2}e^{-0.5\alpha_{it}+\psi(\alpha_{it})}\left[\psi''(\alpha_{it})+(-0.5+\psi'(\alpha_{it})^2\right]$$

$$\frac{\partial^4 F_\epsilon(e^{-\alpha_{it}/2}\eta_{it}|\theta_i)}{\partial \alpha_{it}^4} = -\frac{\eta_{it}}{2}e^{-0.5\alpha_{it}+\psi(\alpha_{it})}\left[\psi'''(\alpha_{it})+3(-0.5+\psi'(\alpha_{it}))\psi''(\alpha_{it})+(-0.5+\psi'(\alpha_{it}))^3\right]$$

$$\begin{aligned}\frac{\partial^5 F_\epsilon(e^{-\alpha_{it}/2}\eta_{it}|\theta_i)}{\partial \alpha_{it}^5} = -\frac{\eta_{it}}{2}e^{\psi(\alpha_{it})}\Big[&\psi^{(4)}(\alpha_{it}) + 4(-0.5+\psi'(\alpha_{it}))\psi'''(\alpha_{it})+3(\psi''(\alpha_{it}))^2\\ &+6(-0.5+\psi'(\alpha_{it}))^2\psi''(\alpha_{it})+(-0.5+\psi'(\alpha_{it}))^4\Big]\end{aligned}$$

B.5 Faà di Bruno Formula

The Faà di Bruno Formula combines the derivatives of primitive functions to obtain the derivatives of composite functions. We can use it to evaluate exact multiple derivatives of compound functions at a point without needing to write out the derivatives of the compound function in closed form.

For the composite function $h = f \circ g$, the Faà di Bruno formula gives

$$\begin{aligned}h' &= f'g',\\ h'' &= f'g''+f''(g')^2,\\ h''' &= f'g'''+3f''g'g''+f'''(g')^3,\\ h^{(4)} &= f'g^{(4)}+4f''g'g'''+3f''(g'')^2+6f'''(g')^2g''+f^{(4)}(g')^4,\\ h^{(5)} &= f'g^{(5)}+5f''g'g^{(4)}+10f''g''g'''+15f'''(g'')^2g'+10f'''g'''(g')^2+10f^{(4)}g''(g')^3+f^{(5)}(g')^5.\end{aligned}$$

If $f^{(j)}=0$ for $j>2$, the third and higher derivatives simplify to

$$\begin{aligned}h''' &= f'g'''+3f''g'g'',\\ h^{(4)} &= f'g^{(4)}+4f''g'g'''+3f''(g'')^2,\\ h^{(5)} &= f'g^{(5)}+5f''g'g^{(4)}+10f''g''g'''.\end{aligned}$$

C Sampling $r|\alpha, \theta, f, B, R$

We draw r from $\pi(r|\alpha, \theta, f, B, R)$ using the following steps:

1. Compute the Cholesky decomposition $R = LL^\top$ of the correlation matrix R.
2. For each $t = 1, ..., n$:
   (a) Draw $z \sim N(0, I_m)$.
   (b) Set $g = Lz$
   (c) Compute the integral probability transform $u_i = \Phi(g_i)$, $i = 1, ..., m$, where Φ is the standard univariate Gaussian cdf.
   (d) Transform each of the u_i to a Student's t with ν_i degree of freedom: $t_i = F^{-1}(u_i)$, where F^{-1} is the inverse cdf of a Student's t distribution with ν_i degrees of freedom.
   (e) Scale each of the t_i random variables to form $\epsilon_{ti} = t_i \exp(0.5\alpha_{ti})$.
   (f) Form $r_t = Bf_t + \epsilon_t$.

D Tables of results

D.1 Posterior parameter's for univariate SV-t models

Parameters	Mean	Std	NSE	RNE
CHF				
$\bar{\alpha}_i$	-10.195	0.189	1.900e-03	2.048e-01
A_{ii}	0.993	0.003	0.000e+00	4.056e-01
σ_{ii}	0.061	0.010	1.000e-04	5.362e-01
ν_i	12.530	3.062	2.110e-02	4.227e-01
σ_α	0.491	0.139	7.0000e-04	7.0480e-01
EUR				
$\bar{\alpha}_i$	-10.333	0.218	1.600e-03	3.843e-01
A_{ii}	0.994	0.002	0.000e+00	5.048e-01
σ_{ii}	0.061	0.010	1.000e-04	4.387e-01
ν_i	18.507	6.252	3.760e-02	5.538e-01
σ_α	0.540	0.143	6.0000e-04	1.0815e+00
AUD				
$\bar{\alpha}_i$	-10.089	0.190	1.400e-03	3.465e-01
A_{ii}	0.988	0.004	0.000e+00	3.944e-01
σ_{ii}	0.104	0.013	1.000e-04	3.839e-01
ν_i	15.225	4.279	3.350e-02	3.261e-01
σ_α	0.657	0.133	7.0000e-04	8.1990e-01
NZD				
$\bar{\alpha}_i$	-9.980	0.144	1.100e-03	3.473e-01
A_{ii}	0.983	0.006	0.000e+00	5.296e-01
σ_{ii}	0.104	0.018	1.000e-04	6.050e-01
ν_i	10.547	2.294	1.410e-02	5.258e-01
σ_α	0.557	0.131	6.0000e-04	9.1270e-01
MXN				
$\bar{\alpha}_i$	-10.880	0.142	1.000e-03	3.906e-01
A_{ii}	0.971	0.008	0.000e+00	5.010e-01
σ_{ii}	0.188	0.023	1.000e-04	5.632e-01
ν_i	33.955	13.626	7.180e-02	7.197e-01
σ_α	0.776	0.129	6.0000e-04	8.4840e-01

Table 6: Posterior statistics of parameters of univariate SV models with student-t errors. First column shows the posterior mean, second column show posterior standard deviation, the third colum show the numerical standard error of the mean based on an estimate of the spectral density at 0, and the last column gives the relative numerical efficiency. σ_α represents the unconditional standard deviation of α_i. Estimations are based on 50,000 draws after discarding 1,000.

Parameters	Mean	Std	NSE	RNE
BRL				
$\bar{\alpha_i}$	-9.711	0.194	1.200e-03	5.672e-01
A_{ii}	0.973	0.006	0.000e+00	6.677e-01
σ_{ii}	0.253	0.023	1.000e-04	7.442e-01
ν_i	37.990	15.446	1.012e-01	4.661e-01
σ_α	1.073	0.148	8.0000e-04	7.2330e-01
GBP				
$\bar{\alpha_i}$	-10.637	0.174	1.500e-03	2.615e-01
A_{ii}	0.989	0.004	0.000e+00	4.604e-01
σ_{ii}	0.088	0.013	1.000e-04	5.834e-01
ν_i	22.118	8.072	5.420e-02	4.429e-01
σ_α	0.566	0.135	6.0000e-04	8.6320e-01
CAD				
$\bar{\alpha_i}$	-10.729	0.255	2.100e-03	3.050e-01
A_{ii}	0.993	0.002	0.000e+00	4.161e-01
σ_{ii}	0.078	0.010	1.000e-04	4.885e-01
ν_i	28.675	11.221	6.720e-02	5.579e-01
σ_α	0.662	0.150	6.0000e-04	1.1276e+00
JPY				
$\bar{\alpha_i}$	-10.369	0.148	1.100e-03	3.370e-01
A_{ii}	0.986	0.005	0.000e+00	2.992e-01
σ_{ii}	0.087	0.014	1.000e-04	3.502e-01
ν_i	11.220	2.528	1.580e-02	5.121e-01
σ_α	0.504	0.116	6.0000e-04	7.1950e-01
SGD				
$\bar{\alpha_i}$	-11.953	0.154	1.200e-03	3.081e-01
A_{ii}	0.984	0.006	0.000e+00	3.795e-01
σ_{ii}	0.102	0.017	1.000e-04	4.366e-01
ν_i	11.814	2.681	1.560e-02	5.915e-01
σ_α	0.568	0.136	7.0000e-04	6.7830e-01

Table 7: Posterior statistics of parameters of univariate SV models with student-t errors. First column shows the posterior mean, second column show posterior standard deviation, the third colum show the numerical standard error of the mean based on an estimate of the spectral density at 0, and the last column gives the relative numerical efficiency. σ_α represents the unconditional standard deviation of α_i. Estimations are based on 50,000 draws after discarding 1,000.

D.2 Posterior parameter's for MSV-q0 model

Parameters	Mean	Std	NSE	RNE
CHF				
$\bar{\alpha}_i$	-10.594	0.095	7.1266e-03	4.4129e-03
A_{ii}	0.97	0.008	2.3179e-04	2.8595e-02
A_{ij}	0.002	0.001	2.3891e-05	2.7707e-02
σ_{ii}	0.0746	0.016	3.6917e-04	4.4277e-02
ν_i	16.19	3.216	7.7741e-02	4.2774e-02
σ_α	0.324	0.060	1.3822e-03	4.7283e-02
EUR				
$\bar{\alpha}_i$	-10.764	0.112	7.1074e-03	6.2172e-03
A_{ii}	0.98	0.004	1.0760e-04	4.2351e-02
A_{ij}	0.001	0.000	1.1393e-05	4.2358e-02
σ_{ii}	0.0722	0.014	2.9354e-04	5.4371e-02
ν_i	23.50	6.342	1.5043e-01	4.4441e-02
σ_α	0.404	0.070	1.5928e-03	4.7762e-02
AUD				
$\bar{\alpha}_i$	-10.387	0.114	8.1451e-03	4.8991e-03
A_{ii}	0.97	0.010	2.3804e-04	4.6442e-02
A_{ij}	0.003	0.001	2.6327e-05	4.8545e-02
σ_{ii}	0.1294	0.024	4.3514e-04	7.3557e-02
ν_i	18.15	4.237	8.6456e-02	6.0037e-02
σ_α	0.513	0.084	2.0306e-03	4.2290e-02
NZD				
$\bar{\alpha}_i$	-10.226	0.101	6.9267e-03	5.2884e-03
A_{ii}	0.96	0.012	2.4565e-04	5.6291e-02
A_{ij}	0.003	0.001	2.7232e-05	4.8866e-02
σ_{ii}	0.1268	0.026	6.1017e-04	4.7135e-02
ν_i	12.78	2.444	4.6843e-02	6.8060e-02
σ_α	0.472	0.082	1.8883e-03	4.6627e-02
MXN				
$\bar{\alpha}_i$	-10.911	0.127	8.3612e-03	5.7422e-03
A_{ii}	0.95	0.011	2.3910e-04	5.0846e-02
A_{ij}	0.004	0.001	2.6614e-05	5.4285e-02
σ_{ii}	0.2353	0.041	1.0775e-03	3.6274e-02
ν_i	38.12	14.615	3.3147e-01	4.8600e-02
σ_α	0.792	0.118	3.2518e-03	3.2956e-02

Table 8: Posterior statistics of parameters of log volatility equation in the MSV model with q=0. First column shows the posterior mean, second column show posterior standard deviation, the third colum show the numerical standard error of the mean based on an estimate of the spectral density at 0, and the fourth column gives the relative numerical efficiency. σ_α represents the unconditional standard deviation of α_i. Estimations are based on 45,000 draws after discarding 6,000.

Parameters	Mean	Std	NSE	RNE
BRL				
$\bar{\alpha_i}$	-9.721	0.186	1.3186e-02	4.9542e-03
A_{ii}	0.97	0.006	1.4302e-04	4.5711e-02
A_{ij}	0.002	0.001	1.6865e-05	6.1487e-02
σ_{ii}	0.2869	0.044	7.7500e-04	7.9142e-02
ν_i	40.69	16.323	3.7063e-01	4.8490e-02
σ_α	1.137	0.166	2.9253e-03	8.0389e-02
GBP				
$\bar{\alpha_i}$	-10.884	0.101	8.1928e-03	3.7836e-03
A_{ii}	0.96	0.011	3.3021e-04	2.7418e-02
A_{ij}	0.003	0.001	3.4675e-05	2.8054e-02
σ_{ii}	0.1158	0.022	4.6367e-04	5.6518e-02
ν_i	25.25	8.309	2.1649e-01	3.6826e-02
σ_α	0.433	0.072	1.5308e-03	5.5078e-02
CAD				
$\bar{\alpha_i}$	-10.884	0.142	8.3868e-03	7.1511e-03
A_{ii}	0.99	0.004	7.2136e-05	5.9020e-02
A_{ij}	0.001	0.000	8.9339e-06	5.2327e-02
σ_{ii}	0.0960	0.017	3.2542e-04	6.9751e-02
ν_i	31.07	11.796	2.4744e-01	5.6812e-02
σ_α	0.578	0.094	1.9962e-03	5.5770e-02
JPY				
$\bar{\alpha_i}$	-10.456	0.106	7.2345e-03	5.3986e-03
A_{ii}	0.96	0.010	2.4814e-04	4.2152e-02
A_{ij}	0.003	0.001	2.7392e-05	4.0232e-02
σ_{ii}	0.1227	0.024	4.1054e-04	8.7760e-02
ν_i	13.21	3.156	1.0109e-01	2.4375e-02
σ_α	0.473	0.082	1.5759e-03	6.7970e-02
SGD				
$\bar{\alpha_i}$	-12.132	0.097	7.7733e-03	3.8583e-03
A_{ii}	0.94	0.018	4.3584e-04	4.4761e-02
A_{ij}	0.006	0.002	4.8223e-05	4.1560e-02
σ_{ii}	0.1550	0.033	7.7144e-04	4.6245e-02
ν_i	14.41	3.319	7.9066e-02	4.4053e-02
σ_α	0.450	0.074	1.5790e-03	5.5276e-02

Table 9: Posterior statistics of parameters of log volatility equation in the MSV model with q=0. First column shows the posterior mean, second column show posterior standard deviation, the third colum show the numerical standard error of the mean based on an estimate of the spectral density at 0, and the fourth column gives the relative numerical efficiency. σ_α represents the unconditional standard deviation of α_i. Estimations are based on 45,000 draws after discarding 6,000.

D.3 Posterior parameter's for MSV-q1 model

Parameters	Mean	Std	NSE	RNE
CHF				
$\bar{\alpha_i}$	-12.913	0.210	1.1413e-02	8.5047e-03
A_{ii}	0.98	0.007	1.9454e-04	2.8507e-02
A_{ij}	0.001	0.001	2.1978e-05	2.5953e-02
σ_{ii}	0.211	0.043	1.4328e-03	2.2969e-02
ν_i	9.423	2.724	1.0473e-01	1.6920e-02
σ_α	0.974	0.165	5.8899e-03	1.9739e-02
EUR				
$\bar{\alpha_i}$	-13.565	0.222	1.4209e-02	6.0833e-03
A_{ii}	0.97	0.007	2.7650e-04	1.7289e-02
A_{ij}	0.001	0.001	3.3092e-05	1.3731e-02
σ_{ii}	0.218	0.048	1.7203e-03	1.9101e-02
ν_i	15.823	7.375	3.2775e-01	1.2660e-02
σ_α	0.977	0.163	6.6350e-03	1.5145e-02
AUD				
$\bar{\alpha_i}$	-10.751	0.116	8.6855e-03	4.4563e-03
A_{ii}	0.94	0.019	5.6233e-04	2.9813e-02
A_{ij}	0.005	0.002	5.5356e-05	2.9620e-02
σ_{ii}	0.161	0.034	8.8369e-04	3.6101e-02
ν_i	9.641	1.708	3.8678e-02	4.8760e-02
σ_α	0.474	0.079	1.9784e-03	3.9497e-02
NZD				
$\bar{\alpha_i}$	-10.503	0.107	7.1830e-03	5.5007e-03
A_{ii}	0.97	0.010	2.6698e-04	3.7050e-02
A_{ij}	0.002	0.001	2.3473e-05	3.5510e-02
σ_{ii}	0.110	0.024	5.4657e-04	4.7496e-02
ν_i	8.528	1.248	3.1835e-02	3.8428e-02
σ_α	0.439	0.076	2.1354e-03	3.1525e-02
MXN				
$\bar{\alpha_i}$	-10.902	0.136	8.8611e-03	5.8568e-03
A_{ii}	0.96	0.010	2.8022e-04	3.2519e-02
A_{ij}	0.003	0.001	2.6870e-05	3.8055e-02
σ_{ii}	0.229	0.039	8.4206e-04	5.4197e-02
ν_i	37.830	14.900	4.0716e-01	3.3482e-02
σ_α	0.788	0.117	2.5907e-03	5.1014e-02

Table 10: Posterior statistics of parameters of log volatility equation in the MSV model with q=1. First column shows the posterior mean, second column show posterior standard deviation, the third colum show the numerical standard error of the mean based on an estimate of the spectral density at 0, and the fourth column gives the relative numerical efficiency. σ_α represents the unconditional standard deviation of α_i and Kurtosis is the unconditional excess kurtosis. Estimations are based on 45,000 draws after discarding 6,000.

Parameters	Mean	Std	NSE	RNE
BRL				
$\bar{\alpha_i}$	-9.715	0.190	1.2265e-02	6.0073e-03
A_{ii}	0.97	0.006	1.3283e-04	5.4991e-02
A_{ij}	0.002	0.001	2.0593e-05	3.3084e-02
σ_{ii}	0.286	0.046	1.0295e-03	4.9610e-02
ν_i	40.773	16.229	4.3679e-01	3.4516e-02
σ_α	1.136	0.170	3.9321e-03	4.6466e-02
GBP				
$\bar{\alpha_i}$	-11.442	0.116	8.6300e-03	4.4953e-03
A_{ii}	0.95	0.016	4.2810e-04	3.3885e-02
A_{ij}	0.004	0.001	4.1659e-05	3.1862e-02
σ_{ii}	0.147	0.034	6.8056e-04	6.3592e-02
ν_i	13.247	4.334	1.4878e-01	2.1219e-02
σ_α	0.491	0.085	1.8577e-03	5.2150e-02
CAD				
$\bar{\alpha_i}$	-10.919	0.129	7.0184e-03	8.3930e-03
A_{ii}	0.98	0.005	1.2659e-04	3.1920e-02
A_{ij}	0.001	0.000	1.1682e-05	3.2690e-02
σ_{ii}	0.101	0.019	4.3968e-04	4.5543e-02
ν_i	29.651	11.388	2.8737e-01	3.9266e-02
σ_α	0.540	0.087	1.8511e-03	5.5488e-02
JPY				
$\bar{\alpha_i}$	-10.584	0.131	8.0320e-03	6.7004e-03
A_{ii}	0.97	0.008	2.1608e-04	3.5409e-02
A_{ij}	0.002	0.001	2.1152e-05	3.5898e-02
σ_{ii}	0.141	0.027	6.0917e-04	4.9221e-02
ν_i	11.341	2.611	6.5062e-02	4.0255e-02
σ_α	0.588	0.097	2.3360e-03	4.2833e-02
SGD				
$\bar{\alpha_i}$	-12.297	0.109	7.8026e-03	4.9218e-03
A_{ii}	0.95	0.016	5.0045e-04	2.5373e-02
A_{ij}	0.004	0.001	4.3480e-05	2.9719e-02
σ_{ii}	0.163	0.037	1.1153e-03	2.7135e-02
ν_i	11.883	2.633	5.5718e-02	5.5822e-02
σ_α	0.504	0.085	2.4041e-03	3.1497e-02

Table 11: Posterior statistics of parameters of log volatility equation in the MSV model with q=1. First column shows the posterior mean, second column show posterior standard deviation, the third colum show the numerical standard error of the mean based on an estimate of the spectral density at 0, and the fourth column gives the relative numerical efficiency. σ_α represents the unconditional standard deviation of α_i and Kurtosis is the unconditional excess kurtosis. Estimations are based on 45,000 draws after discarding 6,000.

	Mean	Std	NSE	RNE
$AR(1)$	0.9903	0.003	6.2776e-05	4.453e-02
ν	21.25	7.831	2.0041e-01	3.8173e-02

Table 12: Posterior statistics of parameters of the factor volatility for MSV model with q=1. For the factor volatility, we set $\bar{\alpha} = 0, \sigma = 1.0$. Estimations are based on 45,000 draws after discarding 6,000.

	Mean	Std	NSE	RNE
B_1	-6.0683e-03	4.0507e-04	3.4856e-05	3.3765e-03
B_2	5.5567e-03	3.7097e-04	3.2516e-05	3.2540e-03
B_3	3.8570e-03	2.6975e-04	2.2083e-05	3.7304e-03
B_4	3.8550e-03	2.7787e-04	2.1401e-05	4.2145e-03
B_5	-8.6580e-05	8.8846e-05	2.0822e-06	4.5518e-02
B_6	-7.0001e-04	1.4292e-04	6.0137e-06	1.4121e-02
B_7	3.7751e-03	2.6207e-04	2.1915e-05	3.5753e-03
B_8	-1.5098e-03	1.3428e-04	8.2632e-06	6.6019e-03
B_9	-2.6726e-03	2.1028e-04	1.4747e-05	5.0834e-03
B_{10}	-1.3541e-03	1.0205e-04	7.4755e-06	4.6590e-03

Table 13: Posterior statistics of loading factor matrix B for MSV model with q=1. First column shows the posterior mean, second column show posterior standard deviation, the third colum show the numerical standard error of the mean based on an estimate of the spectral density at 0, and the fourth column gives the relative numerical efficiency. Estimations are based on 45,000 draws after discarding 6,000.

D.4 Correlation matrices

	CHF	EUR	AUD	NZD	MXN	BRL	GBP	CAD	JPY	SGD
CHF	1.000	-0.354	-0.148	-0.133	-0.035	-0.007	-0.245	0.092	0.163	0.151
EUR	-0.354	1.000	0.186	0.169	-0.003	-0.037	0.262	-0.120	-0.130	-0.166
AUD	-0.148	0.186	1.000	0.313	-0.080	-0.089	0.163	-0.186	-0.083	-0.161
NZD	-0.133	0.169	0.313	1.000	-0.060	-0.070	0.156	-0.152	-0.057	-0.147
MXN	-0.035	-0.003	-0.080	-0.060	1.000	0.204	-0.016	0.081	-0.042	0.049
BRL	-0.007	-0.037	-0.089	-0.070	0.204	1.000	-0.038	0.076	-0.008	0.077
GBP	-0.245	0.262	0.163	0.156	-0.016	-0.038	1.000	-0.095	-0.105	-0.131
CAD	0.092	-0.120	-0.186	-0.152	0.081	0.076	-0.095	1.000	0.036	0.103
JPY	0.163	-0.130	-0.083	-0.057	-0.042	-0.008	-0.105	0.036	1.000	0.197
SGD	0.151	-0.166	-0.161	-0.147	0.049	0.077	-0.131	0.103	0.197	1.000

Table 14: Posterior mean of correlation matrix R_{11} for MSV-q0. Estimation is based on 45,000 draws after discarding 6,000.

	CHF	EUR	AUD	NZD	MXN	BRL	GBP	CAD	JPY	SGD
CHF	1.000	-0.055	0.134	0.126	-0.095	-0.101	0.034	-0.082	0.054	-0.056
EUR	-0.055	1.000	-0.045	-0.047	0.008	-0.004	-0.027	0.025	0.075	0.055
AUD	0.134	-0.045	1.000	0.308	-0.101	-0.093	0.051	-0.160	-0.010	-0.100
NZD	0.126	-0.047	0.308	1.000	-0.074	-0.069	0.058	-0.118	0.013	-0.090
MXN	-0.095	0.008	-0.101	-0.074	1.000	0.204	-0.032	0.091	-0.042	0.057
BRL	-0.101	-0.004	-0.093	-0.069	0.204	1.000	-0.030	0.077	-0.016	0.078
GBP	0.034	-0.027	0.051	0.058	-0.032	-0.030	1.000	-0.020	-0.011	-0.021
CAD	-0.082	0.025	-0.160	-0.118	0.091	0.077	-0.020	1.000	-0.010	0.063
JPY	0.054	0.075	-0.010	0.013	-0.042	-0.016	-0.011	-0.010	1.000	0.174
SGD	-0.056	0.055	-0.100	-0.090	0.057	0.078	-0.021	0.063	0.174	1.000

Table 15: Posterior mean of correlation matrix R_{11} for MSV-q1. Estimation is based on 45,000 draws after discarding 6,000.

	CHF	EUR	AUD	NZD	MXN	BRL	GBP	CAD	JPY	SGD
CHF	1.000	-0.937	-0.593	-0.546	-0.005	0.071	-0.721	0.311	0.475	0.508
EUR	-0.937	1.000	0.621	0.573	-0.019	-0.099	0.734	-0.330	-0.455	-0.518
AUD	-0.593	0.621	1.000	0.575	-0.090	-0.137	0.517	-0.336	-0.322	-0.417
NZD	-0.546	0.573	0.575	1.000	-0.072	-0.116	0.484	-0.294	-0.282	-0.385
MXN	-0.005	-0.019	-0.090	-0.072	1.000	0.205	-0.037	0.092	-0.027	0.060
BRL	0.071	-0.099	-0.137	-0.116	0.205	1.000	-0.096	0.107	0.036	0.121
GBP	-0.721	0.734	0.517	0.484	-0.037	-0.096	1.000	-0.273	-0.373	-0.423
CAD	0.311	-0.330	-0.336	-0.294	0.092	0.107	-0.273	1.000	0.157	0.236
JPY	0.475	-0.455	-0.322	-0.282	-0.027	0.036	-0.373	0.157	1.000	0.392
SGD	0.508	-0.518	-0.417	-0.385	0.060	0.121	-0.423	0.236	0.392	1.000

Table 16: Average of posterior mean of conditional correlation matrix $Corr(r_t|\alpha_t)$ for MSV-q1. Estimation is based on $45,000$ draws after discarding $6,000$.

References

Aguilar, O., and West, M. (2000). 'Bayesian dynamic factor models and variance matrix discounting for portfolio allocation', *Journal of Business and Economic Statistics*, 18: 338–357.

Ang, A., and Chen, J. (2002). 'Asymmetric correlations of equity portfolios', *Journal of Financial Econometrics*, 63: 443–494.

Asai, and McAleer (2009). 'The structure of dynamic correlations in multivariate stochastic volatility models', *Journal of Econometrics*, 150: 182–192.

Bauwens, L., Laurent, S., and Rombouts, J. V. K. (2006). 'Multivariate GARCH : A survey', *Journal of Applied Econometrics*, 21: 79–109.

Carvalho, C. M., and West, M. (2006). 'Dynamic matrix-variate graphical models', *Bayesian Analysis*, 1: 1–29.

Chib, S., Nardari, F., and Shephard, N. (2002). 'Markov Chain Monte Carlo Methods for Stochastic Volatility Models', *Journal of Econometrics*, 108: 281–316.

——— (2006). 'Analysis of High Dimensional Multivariate Stochastic Volatility Models', *Journal of Econometrics*, 134: 341–371.

Chib, S., Omori, Y., and Asai, M. (2009). 'Multivariate Stochastic Volatility.', in T. G. Andersen, R. A. Davis, J.-P. Kreib, and T. Mikosch (eds.), *Handbook of Financial Time Series*, pp. 365–400. Springer.

Danielsson, J. (1998). 'Multivariate stochastic volatility models: Estimation and comparison with VGARCH models.', *Journal of Empirical Finance*, 5: 155–173.

Frühwirth-Schnatter, S., and Früwirth, R. (2007). 'Auxiliary mixture sampling with applications to logistic models', *Computational Statistics and Data Analysis*, 51: 3509–3528.

Frühwirth-Schnatter, S., Früwirth, R., Held, L., and Rue, H. (2009). 'Improved auxiliary mixture sampling for hierarchical models of non-Gaussian data', 19: 479–492.

Frühwirth-Schnatter, S., and Wagner, H. (2006). 'Auxiliary mixture sampling for parameter-driven models of time series of counts with applications to state space modelling', *Biometrika*, 93: 827–841.

Gelman, A., Roberts, G. O., and Gilks, W. R. (1996). 'Efficient Metropolis Jumping Rules', pp. 599–607.

Geweke, J. (1994). 'Bayesian Comparison of Econometric Models', Working Paper, Federal Reserve Bank of Minneapolis Research Department.

Geweke, J. (2004). 'Getting it Right: Joint Distribution Tests of Posterior Simulators', *Journal of the American Statistical Association*, 99: 799–804.

Gourieroux, C. (2006). 'Continous time Wishart process for stochastic risk.', *Econometric Reviews*, 25: 177–217.

Gourieroux, C., Jasiak, J., and Sufana, R. (2004). 'The Wishart autorregresive process of multivariate stochastic volatility.', Discussion Paper, University of Toronto.

Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press, Princeton, NJ.

Han (2006). 'The economics value of volatility modelling: Asset allocation with a high dimensional dynamic latent factor multivariate stochastic volatility model.', *Review of Financial Studies*, 19: 237–271.

Harvey, A. C., Ruiz, E., and Shephard, N. (1994). 'Multivariate stochastic variance models', *Review of Economic Studies*, 61: 247–264.

Jacquier, E., Polson, N., and Rossi, P. (1994). 'Bayesian Analysis of Stochastic Volatility Models', *Journal of Business and Economic Statistics*, 12(4): 371–388.

Jacquier, E., Polson, N., and Rossi, P. (1995). 'Models and prior distributions for multivariate stochastic volatility.', Working paper, Graduate School of Business, University of Chicago.

Kim, S., Shephard, N., and Chib, S. (1998). 'Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models', *Review of Economic Studies*, 65(3): 361–393.

Kolev, N., dos Anjos, U., and de M. Mendez, B. V. (2006). 'Copulas: a review and recent developments.', *Stochastic Models*, 22: 617–660.

Longin, F., and Solnik, B. (2001). 'Extreme correlations of international equity markets.', *Journal of Finance*, 56: 649–676.

Lopes, H. F., and Carvalho, C. M. (2007). 'Factor stochastic volatility with time varying loadings and Markov switching regimes.', *Journal of Statistical Planning and Inference*, 137: 3082–3091.

McCausland, W. J. (2012). 'The HESSIAN Method: Highly Efficient State Smoothing, in a Nutshell', Forthcoming, Journal of Econometrics.

Omori, Y., Chib, S., Shephard, N., and Nakajima, J. (2007). 'Stochastic volatility with leverage: fast and efficient likelihood inference', *Journal of Econometrics*, 140: 425–449.

Patton, A. J. (2009). 'Copula-Based Models for Financial Time Series.', in T. G. Andersen, R. A. Davis, J.-P. Kreib, and T. Mikosch (eds.), *Handbook of Financial Time Series*, pp. 767–784. Springer.

Philipov, A., and Glickman, M. E. (2006). 'Multivariate stochastic volatility via Wishart processes', *Journal of Business and Economic Statistics*, 24: 313–328.

Pitt, M. K., and Shephard, N. (1999). 'Time varying covariances: a factor stochastic volatility approach with discussion.', in J. Bernardo, J. Berger, A. Dawid, and A. Smith (eds.), *Bayesian Statistics*, pp. 547–570. Oxford University Press, Oxford.

Sklar, A. (1959). 'Fonctions de répartition à n dimensions et leurs marges.', *Publications de l'Institut Statistique de l'Université de Paris*, 8: 229–231.

Smith, M., and Pitts, A. (2006). 'Foreign exchange intervention by the Bank of Japan: Bayesian analyisis using a bivariate stochastic volatility model.', *Econometric Reviews*, 25: 425–451.

So, M. K. P., Li, W. K., and Lam., K. (1997). 'Multivariate modelling of the autorregresive random variance process.', *Journal of Time Series Analysis*, 18: 429–446.

Stroud, J. R., Müller, P., and Polson, N. G. (2003). 'Nonlinear State-Space Models With State-Dependent Variances', *Journal of the American Statistical Association*, 98: 377–386.

Vihola, M. (2011). 'Robust Adaptive Metropolis Algorithm with Coerced Acceptance Rate', *Statistics and Computing*.

Yu, J., and Meyer, R. (2006). 'Multivariate stochastic volatility models: Bayesian estimation and model comparison.', *Econometric Reviews*, 25: 361–384.