University of Innsbruck

Multivariate stochastic volatility via wishart processes - A continuation

Wolfgang Rinnergschwentner, Gottfried Tappeiner, Janette Walde

Working Papers in Economics and Statistics

2011-19

Research Platform eeecon Empirical and Experimental Economics

University of Innsbruck
http://eeecon.uibk.ac.at/

Multivariate Stochastic Volatility via Wishart Processes - A Continuation

Wolfgang RINNERGSCHWENTNER

Department of Statistics, University of Innsbruck, Universitaetsstrasse 15, A-6020 Innsbruck, Austria. E-mail: wolfgang.rinnergschwentner@uibk.ac.at

Gottfried TAPPEINER

Department of Economic Theory, Economic Policy and Economic History, University of Innsbruck, Universitaetsstrasse 15, A-6020 Innsbruck, Austria. E-mail: gottfried.tappeiner@uibk.ac.at

Janette WALDE

Corresponding author

Department of Statistics, University of Innsbruck, Universitaetsstrasse 15, A-6020 Innsbruck, Austria. E-mail: janette.walde@uibk.ac.at

Acknowledgements: The authors gratefully acknowledge helpful comments made by Stefan Lang and Alexander Ostermann.

Multivariate Stochastic Volatility via Wishart Processes - A Continuation

Abstract

This paper picks up on a model developed by Philipov and Glickman (2006) for modeling multivariate stochastic volatility via Wishart processes. MCMC simulation from the posterior distribution is employed to fit the model. However, erroneous mathematical transformations in the full conditionals cause false implementation of the approach. We adjust the model, upgrade the analysis and investigate the statistical properties of the estimators using an extensive Monte Carlo study. Employing a Gibbs sampler in combination with a Metropolis Hastings algorithm inference for the time-dependent covariance matrix is feasible with appropriate statistical properties.

Key Words: Bayesian time series; Stochastic covariance; Time-varying correlation; Markov Chain Monte Carlo

1 INTRODUCTION

Philipov and Glickman (2006) formulate a general model of multivariate stochastic volatility driven by Wishart processes. The model aims to provide a direct connection between univariate and multivariate models in order to increase flexibility in describing the behavior of stochastic covariances and improve volatility estimates and forecasts. Their model offers several advantages. Covariance matrices are used instead of vectors of log-variances, and therefore their framework is a multivariate extension of the scalar case (Philipov and Glickman 2006). Both variances and correlations evolve stochastically over time, and conditional volatility depends not only on past volatility, but also on past covariances. Additionally, restrictions imposed by previous stochastic volatility (SVOL) models can be tested in their framework, because many of the existing SVOL models can be regarded as special cases of their general setup.

This paper picks up on a deficiency in the work of Philipov and Glickman (2006). The derivation of the full conditionals in Philipov and Glickman (2006) is erroneous, and consequently the results of the simulation study therein are not valid. As the approach itself is promising and valuable for users, we adjust the model and upgrade the analysis:

1. Due to a mathematical error the derivations of the full conditionals are not correct. All formulae are corrected and the derivations are given in detail.

2. Although the expressions of the full conditionals become increasingly complex, we demonstrate how the approach can be appropriately implemented using Bayesian estimation methods.

3. Philipov and Glickman (2006) demonstrate the suitability of their approach using just one simulated data set. Although the approach is CPU time-consuming, we perform an extensive Monte Carlo study in order to demonstrate the appropriateness of the estimators and the validity of the approach.

The remainder of this article is structured as follows: Section 2 gives the model characteristics with all necessary formulae for the joint and conditional posterior distributions. The Monte Carlo study and its settings are described in Section 3. Thereafter, the results are discussed and the main conclusions summarized.

2 MODEL

The main characteristics of the model derived by Philipov and Glickman (2006) as well as the correctly deduced formulae for the full conditionals are given in this section. Detailed derivations are shown in Appendix A.

The paper by Philipov and Glickman (2006) contains several typographical errors. Appendix B lists the errors and their correction.

2.1 Philipov and Glickman Model

The proposed approach models the time development of k correlated normally distributed random variables. The dynamic covariance structure is driven by a stochastic process based on the Wishart distribution with ν degrees of freedom,

$$\mathbf{\Sigma}_t^{-1}|\nu, \mathbf{S}_{t-1} \sim \mathrm{Wish}(\nu, \mathbf{S}_{t-1}), \tag{1}$$

where the stochastic process of the observation units is given as $\mathbf{y}_t|\mathbf{\Sigma}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_t)$. Here $\mathbf{0}$ and $\mathbf{y}_t$ are $k \times 1$ vectors, $\mathbf{\Sigma}_t$ is the $k \times k$ covariance matrix, $\mathcal{N}(\cdot, \cdot)$ denotes the normal distribution, and the abbreviation 'Wish' the Wishart distribution. The matrix $\mathbf{S}_t$ is the time-dependent scale parameter and defined as $\mathbf{S}_t = 1/\nu \left(\mathbf{A}^{1/2}\right)\left(\boldsymbol{\Sigma}_t^{-1}\right)^d \left(\mathbf{A}^{1/2}\right)'$, where $\mathbf{A}$ is a positive definite symmetric parameter matrix, $\mathbf{A}^{1/2}$ is the lower triangular matrix of a Cholesky decomposition of $\mathbf{A}$ and d is a scale parameter. The matrix $\mathbf{A}$ shows the dependencies of each variance on the other variances and covariances. The scale parameter d can be regarded as a persistence parameter. For a more detailed interpretation of the parameters and their influence on the dynamic behavior of the covariance matrix see Philipov and Glickman (2006).

The conditional expected value for matrices following a Wishart distribution is given as $E\left(\boldsymbol{\Sigma}_t^{-1}|\mathbf{A}, \boldsymbol{\Sigma}_{t-1}, d\right) = \left(\mathbf{A}^{1/2}\right)\left(\boldsymbol{\Sigma}_{t-1}^{-1}\right)^d \left(\mathbf{A}^{1/2}\right)'$. The conditional distributions of the covariance matrices themselves follow the inverse-Wishart distribution. The conditional expectation is given as $E\left(\boldsymbol{\Sigma}_t|\mathbf{A}, \boldsymbol{\Sigma}_{t-1}, d\right) = \frac{\nu}{\nu-k-1}\left(\mathbf{A}^{-1/2}\right)' \boldsymbol{\Sigma}_{t-1}^d \left(\mathbf{A}^{-1/2}\right)$.

In order to estimate the parameters $\nu, d, \mathbf{A}$ and to obtain estimates for the latent variables $\boldsymbol{\Sigma}_t$ Bayesian methods are employed. After choosing appropriate prior distributions the joint posterior distribution can be derived. With the full conditionals the parameters can be estimated using MCMC methods.

The priors are fixed in the following way. For $\mathbf{A}^{-1}$ a Wishart distribution with γ_0 degrees of freedom and a $k \times k$ scale matrix $\mathbf{Q}_0$ is assumed. $\mathbf{Q}_0$ is the identity matrix and the degrees of freedom are $k+1$. The parameter d is drawn from a diffuse prior. A uniform distribution at the interval $[0; 1]$ is chosen. Philipov and Glickman (2006) choose for the parameter ν a gamma distribution as prior distribution. It is not really obvious why a gamma distribution should be chosen in our case for ν; a noninformative uniform distribution is chosen in this paper. In this process it must be remembered that only values greater than k are allowed. For the inverse covariance matrices a Wishart distribution is used as prior distribution (see Equation (1)). Given these assumptions for the prior distributions, the joint posterior distribution can be derived.

2.2 Full Conditionals

The expressions for the full conditionals for ν, d, $\mathbf{A}^{-1}$, as well as the expression for the acceptance ratio of $\mathbf{\Sigma}_t^{-1}$ in Philipov and Glickman (2006) are erroneous. For example, the expression of the acceptance ratio of $\mathbf{\Sigma}_t^{-1}$ contains the term $tr(\mathbf{S}_t^{-1}\mathbf{\Sigma}_{t+1}^{-1})$ (cf. p. 326), where $tr(\mathbf{A})$ is defined as the trace of matrix $\mathbf{A}$. This expression is simplified to:

$$tr(\mathbf{S}_t^{-1}\boldsymbol{\Sigma}_{t+1}^{-1}) = tr\left(\nu\left(\mathbf{A}^{1/2'}\right)^{-1}\left(\boldsymbol{\Sigma}_t^{-1}\right)^{-d}\left(\mathbf{A}^{1/2}\right)^{-1}\boldsymbol{\Sigma}_{t+1}^{-1}\right) = tr(\nu\mathbf{A}^{-1}\boldsymbol{\Sigma}_t^{d}\boldsymbol{\Sigma}_{t+1}^{-1}).$$

However, this simplification cannot be achieved with known mathematical methods (Horn and Johnson 1985).

Furthermore, in the derivation of the full conditional for $p\left(\mathbf{A}^{-1}|\cdot\right)$

$$\begin{aligned}
p\left(\mathbf{A}^{-1}|\cdot\right) &\propto \text{Wish}\left(\mathbf{A}^{-1}|\gamma_0, \mathbf{Q}_0\right)\prod_{t=1}^{T}\text{Wish}\left(\boldsymbol{\Sigma}_t^{-1}|\nu, \mathbf{S}_{t-1}\right) \\
&\propto \left|\mathbf{A}^{-1}\right|^{(\gamma_0-k-1)/2}\exp\left(-1/2\cdot\text{tr}\left(\mathbf{Q}_0^{-1}\mathbf{A}^{-1}\right)\right)\cdot \\
&\quad \prod_{t=1}^{T}\left|\mathbf{S}_{t-1}\right|^{-\nu/2}\exp\left(-1/2\cdot\text{tr}\left(\mathbf{S}_{t-1}^{-1}\boldsymbol{\Sigma}_t^{-1}\right)\right) \\
&\propto \left|\mathbf{A}^{-1}\right|^{(\nu T+\gamma_0-k-1)/2}\cdot\exp\Bigg(-1/2\cdot\text{tr}\Bigg(\mathbf{Q}_0^{-1}\mathbf{A}^{-1} + \\
&\quad \nu\sum_{t=1}^{T}\left(\mathbf{A}^{1/2'}\right)^{-1}\boldsymbol{\Sigma}_{t-1}^{d}\left(\mathbf{A}^{1/2}\right)^{-1}\boldsymbol{\Sigma}_t^{-1}\Bigg)\Bigg) \qquad (2)
\end{aligned}$$

the following expression is erroneously simplified:

$$\begin{aligned}
&\text{tr}\left(\mathbf{Q}_0^{-1}\mathbf{A}^{-1} + \nu\sum_{t=1}^{T}\left(\mathbf{A}^{1/2'}\right)^{-1}\boldsymbol{\Sigma}_{t-1}^{d}\left(\mathbf{A}^{1/2}\right)^{-1}\boldsymbol{\Sigma}_t^{-1}\right) = \\
&\text{tr}\left(\left(\mathbf{Q}_0^{-1} + \nu\sum_{t=1}^{T}\boldsymbol{\Sigma}_{t-1}^{d/2}\boldsymbol{\Sigma}_t^{-1}\boldsymbol{\Sigma}_{t-1}^{d/2}\right)\mathbf{A}^{-1}\right)
\end{aligned}$$

(Whether $\mathbf{\Sigma}_{t-1}^{d/2}$ results from a Cholesky or a spectral decomposition depends on whether the paper by Philipov and Glickman (2006) contains a typographical error or not.) This term also appears in the full conditionals for ν and d. Using expression (2) for the sampler for $\mathbf{A}^{-1}$ as well as the correct full conditionals for ν and d (cf. next section) the MCMC sampler becomes more complex, because the full conditionals no longer follow known distributions. For example, Philipov and Glickman (2006) draw the values for $\mathbf{A}^{-1}$ directly from a Wishart distribution and employ a simple Gibbs sampler. Now the Gibbs sampler must be combined with a Metropolis Hastings algorithm. The wrongly computed term can be found in all full conditionals.

Given the assumptions for the prior distributions the joint posterior distribution can be derived,

$$
\begin{aligned}
p(\mathbf{\Sigma}^{-1}, \mathbf{A}^{-1}, \nu, d|\mathbf{y}) \quad \propto \quad & \mathrm{Wish}(\mathbf{A}^{-1}|\gamma_0, \mathbf{Q}_0) \cdot \\
& \prod_{t=1}^{T} \mathrm{Wish}\left(\mathbf{\Sigma}_t^{-1}|\nu, \mathbf{S}_{t-1}\right) \cdot \mathcal{N}\left(\mathbf{y}_t|\mathbf{0}, \mathbf{\Sigma}_t\right) \qquad (3)
\end{aligned}
$$

For implementation of the MCMC sampler the full conditionals for the parameters have to be deduced and are summarized as follows (detailed derivations of the full conditionals are given in Appendix A).

$$
\begin{aligned}
p\left(\mathbf{\Sigma}_t^{-1}|\cdot\right) &\propto \text{Wish}\left(\mathbf{\Sigma}_t^{-1}|\nu, \widetilde{\mathbf{S}}_{t-1}\right) \cdot \left|\mathbf{\Sigma}_t^{-1}\right|^{(1-\nu d)/2} \cdot \exp\left(-\frac{1}{2}\, tr\left(\mathbf{S}_t^{-1}\mathbf{\Sigma}_{t+1}^{-1}\right)\right) \\
p\left(\mathbf{\Sigma}_T^{-1}|\cdot\right) &\propto \text{Wish}\left(\mathbf{\Sigma}_T^{-1}|\nu+1, \widetilde{\mathbf{S}}_{T-1}\right), \qquad (4)
\end{aligned}
$$

where $\widetilde{\mathbf{S}}_{t-1}$ is defined as $\widetilde{\mathbf{S}}_{t-1} = \left(\mathbf{S}_{t-1}^{-1} + \mathbf{y}_t\mathbf{y}_t'\right)^{-1}$.

$$
\begin{aligned}
p\left(\mathbf{A}^{-1}|\cdot\right) &\propto \left|\mathbf{A}^{-1}\right|^{(\gamma_0+\nu T-k-1)/2} \cdot \exp\left(-\frac{1}{2} tr\left(\mathbf{Q}_0^{-1}\mathbf{A}^{-1}\right)\right) \cdot \\
&\quad \prod_{t=1}^{T} \exp\left(-\frac{1}{2} tr\left(\mathbf{S}_{t-1}^{-1}\mathbf{\Sigma}_t^{-1}\right)\right) \qquad (5) \\
p\left(\nu|\cdot\right) &\propto \left(\frac{\left|\nu\mathbf{A}^{-1}\right|^{\nu/2}}{2^{\nu k/2}\prod_{i=1}^{k}\Gamma\left(\frac{\nu+1-i}{2}\right)}\right)^T \cdot \prod_{t=1}^{T} \left|\mathbf{\Sigma}_{t-1}\right|^{\nu d/2}\left|\mathbf{\Sigma}_t^{-1}\right|^{\nu/2} \cdot \\
&\quad \prod_{t=1}^{T} \exp\left(-\frac{1}{2} tr\left(\mathbf{S}_{t-1}^{-1}\mathbf{\Sigma}_t^{-1}\right)\right) \qquad (6) \\
p\left(d|\cdot\right) &\propto \prod_{t=1}^{T} \left|\mathbf{\Sigma}_{t-1}^{-1}\right|^{-\nu d/2} \exp\left(-\frac{1}{2}\, tr\left(\mathbf{S}_{t-1}^{-1}\mathbf{\Sigma}_t^{-1}\right)\right) \qquad (7)
\end{aligned}
$$

2.3 MCMC Sampler

In the following the superscript '*' denotes the proposed values and the superscript '$[m-1]$' the current state of the Markov chain.

2.3.1 Sampling $\mathbf{\Sigma}_t^{-1}$

To sample $\mathbf{\Sigma}_t^{-1}$ an independence chain Metropolis Hastings (MH) step is used. We use Wish $\left(\mathbf{\Sigma}_t^{-1}|\nu,\widetilde{\mathbf{S}}_{t-1}\right)$ as proposal density. For the acceptance ratio we obtain:

$$AR \;=\; \frac{\left|(\mathbf{\Sigma}_t^{*})^{-1}\right|^{(1-\nu d)/2}}{\left|\left(\mathbf{\Sigma}_t^{[m-1]}\right)^{-1}\right|^{(1-\nu d)/2}}\cdot\frac{\exp\left(-\frac{1}{2}tr\left(\mathbf{S}_t^{-1}\left(\mathbf{\Sigma}_t^{*}\right)\mathbf{\Sigma}_{t+1}^{-1}\right)\right)}{\exp\left(-\frac{1}{2}tr\left(\mathbf{S}_t^{-1}\left(\mathbf{\Sigma}_t^{[m-1]}\right)\mathbf{\Sigma}_{t+1}^{-1}\right)\right)},$$

where $\mathbf{S}_t^{-1}\left(\mathbf{\Sigma}_t^{*}\right)$ is defined as $\mathbf{S}_t^{-1}\left(\mathbf{\Sigma}_t^{*}\right) = \nu\left(\mathbf{A}^{1/2'}\right)^{-1}\left(\mathbf{\Sigma}_t^{*}\right)^{d}\left(\mathbf{A}^{1/2}\right)^{-1}$ and $\mathbf{S}_t^{-1}\left(\mathbf{\Sigma}_t^{[m-1]}\right)$ is given as $\mathbf{S}_t^{-1}\left(\mathbf{\Sigma}_t^{[m-1]}\right) = \nu\left(\mathbf{A}^{1/2'}\right)^{-1}\left(\mathbf{\Sigma}_t^{[m-1]}\right)^{d}\left(\mathbf{A}^{1/2}\right)^{-1}$. $\mathbf{\Sigma}_T^{-1}$ can be directly drawn from a Wishart distribution (cf. Equation (4)) with $\nu+1$ degrees of freedom and the scale matrix $\left(\mathbf{S}_{T-1}^{-1}+\mathbf{y}_T\mathbf{y}_T'\right)^{-1}$.

2.3.2 Sampling $\mathbf{A}^{-1}$

For the matrix $\mathbf{A}^{-1}$ a random-walk proposal is employed:

$$
\begin{aligned}
\left(\mathbf{A}^{-1}\right)^{*} = \left(\mathbf{A}^{-1}\right)^{[m-1]} \quad &+ \quad \boldsymbol{\Omega} \\
&\boldsymbol{\Omega} = \begin{pmatrix} \omega_{11} & \omega_{12} \\ \omega_{12} & \omega_{22} \end{pmatrix} \\
&(\omega_{11} \quad \omega_{12} \quad \omega_{22})' \sim \mathcal{N}\left(\mathbf{0}, \sigma^2_{\mathbf{A}^{-1}}\mathbf{I}\right), \qquad (8)
\end{aligned}
$$

where $\boldsymbol{\Omega}$ follows a symmetric matrix variate normal distribution. As the matrix $\boldsymbol{\Omega}$ has $k \cdot (k+1)/2$ different random elements, each one can be drawn from a $k \cdot (k+1)/2$-dimensional normal distribution (Gupta and Nagar 2000). The acceptance ratio is given as:

$$
\begin{aligned}
AR \quad = \quad & \frac{|\mathbf{A}^{*}|^{-(\gamma_0+\nu T-k-1)/2}}{|\mathbf{A}^{[m-1]}|^{-(\gamma_0+\nu T-k-1)/2}} \cdot \frac{exp\left(-\frac{1}{2}tr\left(\mathbf{Q}_0^{-1}(\mathbf{A}^{*})^{-1}\right)\right)}{exp\left(-\frac{1}{2}tr\left(\mathbf{Q}_0^{-1}(\mathbf{A}^{[m-1]})^{-1}\right)\right)} \cdot \\
& \frac{\exp\left(-\frac{1}{2}tr\left(\sum_{t=1}^{T}\mathbf{S}_{t-1}^{-1}\left(\mathbf{A}^{*}\right)\boldsymbol{\Sigma}_t^{-1}\right)\right)}{\exp\left(-\frac{1}{2}tr\left(\sum_{t=1}^{T}\mathbf{S}_{t-1}^{-1}\left(\mathbf{A}^{[m-1]}\right)\boldsymbol{\Sigma}_t^{-1}\right)\right)},
\end{aligned}
$$

where

$$
\begin{aligned}
\mathbf{S}_{t-1}^{-1}\left(\mathbf{A}^{*}\right) &= \nu\left(\left(\mathbf{A}^{*}\right)^{1/2'}\right)^{-1}\boldsymbol{\Sigma}_{t-1}^{d}\left(\left(\mathbf{A}^{*}\right)^{1/2}\right)^{-1} \\
\mathbf{S}_{t-1}^{-1}\left(\mathbf{A}^{[m-1]}\right) &= \nu\left(\left(\mathbf{A}^{[m-1]}\right)^{1/2'}\right)^{-1}\boldsymbol{\Sigma}_{t-1}^{d}\cdot\left(\left(\mathbf{A}^{[m-1]}\right)^{1/2}\right)^{-1}
\end{aligned}
$$

2.3.3 Sampling ν

The values for ν are also obtained with a random-walk proposal:

$$
\nu^{*} = \nu^{[m-1]} + \epsilon_{\nu}\,, \qquad \epsilon_{\nu} \sim \mathcal{N}\left(0, \sigma_{\nu}^{2}\right), \tag{9}
$$

where ϵ_{ν} is drawn from a normal distribution. The acceptance ratio can be derived as:

$$
\begin{aligned}
AR = & \left(\frac{\left|\nu^{*}\mathbf{A}^{-1}\right|^{\nu^{*}/2}}{2^{\nu^{*}k/2}\prod_{i=1}^{k}\Gamma\left(\frac{\nu^{*}+1-i}{2}\right)}\right)^{T}\cdot\left(\frac{\left|\nu^{[m-1]}\mathbf{A}^{-1}\right|^{\nu^{[m-1]}/2}}{2^{\nu^{[m-1]}k/2}\prod_{i=1}^{k}\Gamma\left(\frac{\nu^{[m-1]}+1-i}{2}\right)}\right)^{-T}\cdot \\
& \frac{\prod_{t=1}^{T}\left|\boldsymbol{\Sigma}_{t-1}^{-1}\right|^{-\nu^{*}d/2}\left|\boldsymbol{\Sigma}_{t}^{-1}\right|^{\nu^{*}/2}}{\prod_{t=1}^{T}\left|\boldsymbol{\Sigma}_{t-1}^{-1}\right|^{-\nu^{[m-1]}d/2}\left|\boldsymbol{\Sigma}_{t}^{-1}\right|^{\nu^{[m-1]}/2}}\cdot\frac{\prod_{t=1}^{T}exp\left(-\frac{1}{2}tr\left(\mathbf{S}_{t-1}^{-1}\left(\nu^{*}\right)\boldsymbol{\Sigma}_{t}^{-1}\right)\right)}{\prod_{t=1}^{T}exp\left(-\frac{1}{2}tr\left(\mathbf{S}_{t-1}^{-1}\left(\nu^{[m-1]}\right)\boldsymbol{\Sigma}_{t}^{-1}\right)\right)},
\end{aligned}
$$

whereby

$$
\begin{aligned}
\mathbf{S}_{t-1}^{-1}\left(\nu^*\right) &= \nu^* \left(\mathbf{A}^{1/2'}\right)^{-1} \boldsymbol{\Sigma}_{t-1}^{d} \left(\mathbf{A}^{1/2}\right)^{-1} \\
\mathbf{S}_{t-1}^{-1}\left(\nu^{[m-1]}\right) &= \nu^{[m-1]} \left(\mathbf{A}^{1/2'}\right)^{-1} \boldsymbol{\Sigma}_{t-1}^{d} \left(\mathbf{A}^{1/2}\right)^{-1}
\end{aligned}
$$

2.3.4 Sampling d

Finally, d is obtained using a random-walk proposal:

$$
d^* = d^{[m-1]} + \epsilon_d\,, \qquad \epsilon_d \sim \mathcal{N}\left(0, \sigma_d^2\right), \tag{10}
$$

where ϵ_d is drawn form a normal distribution. The acceptance ratio is:

$$
AR = \frac{\prod_{t=1}^{T} \left|\boldsymbol{\Sigma}_{t-1}\right|^{\nu d^*/2}}{\prod_{t=1}^{T} \left|\boldsymbol{\Sigma}_{t-1}\right|^{\nu d^{[m-1]}/2}} \cdot \frac{\prod_{t=1}^{T} exp\left(-\frac{1}{2} tr\left(\mathbf{S}_{t-1}^{-1}\left(d^*\right)\boldsymbol{\Sigma}_{t}^{-1}\right)\right)}{\prod_{t=1}^{T} exp\left(-\frac{1}{2} tr\left(\mathbf{S}_{t-1}^{-1}\left(d^{[m-1]}\right)\boldsymbol{\Sigma}_{t}^{-1}\right)\right)},
$$

and

$$\begin{aligned}
\mathbf{S}_{t-1}^{-1}\left(d^{*}\right) &= \nu\left(\mathbf{A}^{1/2'}\right)^{-1}\mathbf{\Sigma}_{t-1}^{d^{*}}\left(\mathbf{A}^{1/2}\right)^{-1} \\
\mathbf{S}_{t-1}^{-1}\left(d^{[m-1]}\right) &= \nu\left(\mathbf{A}^{1/2'}\right)^{-1}\mathbf{\Sigma}_{t-1}^{d^{[m-1]}}\left(\mathbf{A}^{1/2}\right)^{-1}
\end{aligned}$$

The procedure is implemented in MATLAB 7.7.0.471(R2008b).

3 MONTE CARLO STUDY

3.1 Simulation Design

The validity of our approach is analyzed using a Monte Carlo study. A two-dimensional model is used. The true values are chosen similar as in Philipov and Glickman (2006). For the parameters ν and d the same values are employed: $\nu = 20$ and $d = 0.7$. As only the value of $|\mathbf{A}|$ can be found in Philipov and Glickman (2006), the matrix $\mathbf{A}$ is fixed in order to have a similar value for the determinant, i.e. $\mathbf{A} = \binom{130\ 50}{50\ 130}$. The starting values for the data generating process (DGP), i.e. $\mathbf{\Xi}$, are chosen to guarantee appropriate values for the yields: $\mathbf{\Xi} = \binom{3\ 0.8}{0.8\ 3}$.

With these fixed values for the model parameters $\mathbf{A}$, ν, d and $\mathbf{\Xi}$, covariance matrices for 100 dates (i.e., $t = 1, ..., 100$) are drawn from the corresponding Wishart distribution. Using these covariance matrices yields $(\mathbf{y}_t)$ are simulated. We simulated 100 replicated data sets from the model. The yields of each data set are the observation units with which the model parameters as well as the covariance matrices are estimated. Note that with this procedure $\mathbf{\Xi}$ corresponds to the true $\mathbf{\Sigma}_0$, analogously to Philipov and Glickman (2006). However, after just a few dates (in our various experiments at maximum 10 dates) the starting values of the DGP are no longer relevant.

3.2 Sampling Approach

The employed sampling approach is conducted in two steps. In the first step the three parameters and the covariance matrices are sampled together. The starting values are chosen as follows: 5 for ν, 0.5 for d and the identity matrix for $\mathbf{A}$ and $\mathbf{\Sigma}_0$.

With the help of one data set the variances of the samplers for the parameters ν, d, and $\mathbf{A}$ are fixed. The variances of the samplers are set in order to obtain an acceptance rate of approximately 0.5.

More than a million simulation steps (1,080,000) are computed and 180,000 are dropped as burn-in phase. For the remaining values, dropping every 9,000th value gives a sample of uncorrelated draws, i.e. the effective sample size is 100.

Because the estimates of ν and $\mathbf{A}$ depend sensitively on the appropriateness of the values of the covariance matrix, values for ν and $\mathbf{A}$ are drawn again in a second step. The covariance matrices and d are held at the estimates in the first step. Accordingly, the estimates obtained in the first step for ν and $\mathbf{A}$ are employed as starting values. In the second simulation step 4,000 values for ν and $\mathbf{A}$ are drawn. The first 1,000 values correspond to the burn-in phase. Again, we obtain an effective sample size of 100 for ν and $\mathbf{A}$.

3.3 Evaluation Methods

In order to assess the quality of the samplers their convergence, the empirical autocorrelation function and all acceptance rates are investigated. Various statistical characteristics are used to evaluate the appropriateness of the estimates for $\mathbf{A}$, ν, and d. Analogously to Philipov and Glickman (2006), the findings for the parameter matrix $\mathbf{A}$ are condensed and scaled into the natural logarithm of the determinant of $\mathbf{A}$, i.e. $log|\mathbf{A}|$.

Additionally, for each estimate the highest posterior density interval (hpdi) is computed (Koop 2006).

Of course, a practitioner focuses on the quality of the covariance matrices' forecasts as these estimates are used further in economic models, i.e. the estimated conditional expectation of the inverse of the covariance matrices is investigated $E\left(\mathbf{\Sigma}_t^{-1}|\mathbf{A},\mathbf{\Sigma}_{t-1},d\right) = \left(\mathbf{A}^{1/2}\right)\left(\mathbf{\Sigma}_{t-1}^{-1}\right)^d\left(\mathbf{A}^{1/2}\right)'$. In order to calculate an estimate for the conditional mean, one can assume that $\mathbf{\Sigma}_{t-1}^{-1}$ is known then the estimate of the conditional mean is: $E\widehat{\left(\mathbf{\Sigma}_t^{-1}|\cdot\right)} = \left(\widehat{\mathbf{A}}^{1/2}\right)\left(\mathbf{\Sigma}_{t-1}^{-1}\right)^d\left(\widehat{\mathbf{A}}^{1/2}\right)'$, or when also respecting the latent nature of the covariance matrices the estimate of $\mathbf{\Sigma}_{t-1}^{-1}$ is plugged into: $E\widehat{\widehat{\left(\mathbf{\Sigma}_t^{-1}|\cdot\right)}} = \left(\widehat{\mathbf{A}}^{1/2}\right)\cdot\left(\widehat{\mathbf{\Sigma}}_{t-1}^{-1}\right)^d\left(\widehat{\mathbf{A}}^{1/2}\right)'$. The quality of the latter expression is of course more relevant for practical implementations.

Two approaches are employed to measure the quality of this statistic. On the one hand, the mean absolute percentage error ($MAPE$) is calculated as:

$$MAPE_{ij} = \frac{1}{T}\sum_{t=1}^{T}\frac{\left|(\mathbf{\Sigma}_t^{-1})_{ij} - E\widehat{(\mathbf{\Sigma}_t^{-1}|\cdot)}_{ij}\right|}{\left|(\mathbf{\Sigma}_t^{-1})_{ij}\right|}, \qquad i=1,2, j=1,2$$

for each data set. For the sake of completeness, the $MAPE$ is also computed for the condensed statistic $log|E(\mathbf{\Sigma}_t^{-1}|\cdot)|$. However, note that due to the nature of the log function this $MAPE$ will always be much smaller.

On the other hand, we calculated for each data set how many of the 100 true values lie in their correspondingly estimated 95% hpdi. This procedure was performed for each matrix element of the conditional mean as well as for the condensed statistic.

4 RESULTS

First, important characteristics are discussed using an example data set. The appropriateness of the samplers is analyzed for their convergence behavior, possible autocorrelation, and acceptance rate. The sampling results demonstrate early convergence of the samplers. The investigation of the corresponding autocorrelation shows that we work with a large effective sample size defined as the number of uncorrelated draws from the conditional distributions. The acceptance rates of the samplers are quite appropriate (cf. last column of Table 1).

> - - - Insert Table 1 - - - <

Regarding ν and d, the estimates of these parameters are close to their true values (cf. Table 1). However, the values of the matrix $\mathbf{A}$ are underestimated. Two questions now arise: Is this underestimation a bias? How does this statistical distortion influence the forecast for the covariance matrices?

Whether this underestimation is a systematic bias or not is answered below with the help of all Monte Carlo data sets. In order to answer the second question, the performance of the forecasts of the covariance matrices is analyzed. In addition to the estimates of the covariance matrices, the 95% highest posterior density intervals are computed. The results for each element of the explained variable $\mathbf{\Sigma}_t^{-1}$ are plotted in Figure 1 for $E\widehat{\left(\mathbf{\Sigma}_t^{-1}|\cdot\right)}$ and in Figure 3 for $E\widehat{\widehat{\left(\mathbf{\Sigma}_t^{-1}|\cdot\right)}}$. Figure 2 shows the equivalent plots for the condensed statistic $log|E(\mathbf{\Sigma}_t^{-1}|\mathbf{A}, \mathbf{\Sigma}_{t-1}, d)|$.

$>$ - - - Insert Figure 1 - - - $<$

The quality of the findings for $E\widehat{\left(\mathbf{\Sigma}_t^{-1}|\cdot\right)}$ is quite good, not only graphically but also as expressed by a small mean absolute percentage error for each element of $E\widehat{\left(\mathbf{\Sigma}_t^{-1}|\cdot\right)}$ ($MAPE$ for $E(\widehat{\mathbf{\Sigma}_t^{-1}|\cdot})_{11} = 0.02$, $MAPE$ for $E(\widehat{\mathbf{\Sigma}_t^{-1}|\cdot})_{12} = 0.08$, $MAPE$ for $E(\widehat{\mathbf{\Sigma}_t^{-1}|\cdot})_{22} = 0.11$). The same holds for the condensed statistic ($MAPE = 0.005$, cf. Figure 2 (a)).

> - - - Insert Figure 2 - - - <

For practical purposes it must be remembered that $\mathbf{\Sigma}_t$ is a latent variable and unknown to the user. Therefore, the estimate of $\mathbf{\Sigma}_t$ must also be plugged into the conditional expected value. Figure 3 shows the findings for $\widehat{E(\mathbf{\Sigma}_t^{-1}|\cdot)}$.

> - - - Insert Figure 3 - - - <

Although the fluctuations of the true values are not modeled very well using the posterior mean (posterior median) ($MAPE$ for $\widehat{E(\mathbf{\Sigma}_t^{-1}|\cdot)}_{11} = 0.23$, $MAPE$ for $\widehat{E(\mathbf{\Sigma}_t^{-1}|\cdot)}_{12} = 0.30$, $MAPE$ for $\widehat{E(\mathbf{\Sigma}_t^{-1}|\cdot)}_{22} = 0.33$), the estimated 95% hpdi almost always captures the true values. For the condensed statistic the hpdi contains the true values and the $MAPE$ is 2%. The findings demonstrate that the time dependence of the covariance matrices is appropriately modeled using the 95% hpdi (cf. Figure 2 (b)).

Whether the above findings can be generalized is now determined by analyzing the results for all data sets. The first three plots in Figure 4 show the estimates and their 95% hpdi of the parameters ν, d, and $log|\mathbf{A}|$ for each data set. The parameter d is estimated quite well. There is a high deviation of ν from its true value, and $log|\mathbf{A}|$ is slightly underestimated for almost all data sets.

$>$ - - - Insert Figure 4 - - - $<$

However, these statistical deficiencies do not seem to have a serious impact on the consequences for the estimates of $\mathbf{\Sigma}_t^{-1}$. Analyzing Figure 5 we see that a high fraction of the explained variable lies in the correspondingly estimated 95% hpdi. The findings are similar for each element of $\widehat{E(\mathbf{\Sigma}_t^{-1}|\cdot)}$ as well as for $log|\widehat{E(\mathbf{\Sigma}_t^{-1}|\cdot)}|$. Despite the small number of data sets, $\mathbf{\Sigma}_t^{-1}$ can be forecasted appropriately.

$>$ - - - Insert Figure 5 - - - $<$

Table 2 shows summary statistics for the 100 Monte Carlo data sets of the three parameters $\mathbf{A}$, ν, and d. The estimates are suitable as indicated by the mean and median and are not systematically biased as the true values of the parameters lie in their 95% hpdi. However, the 95% hpdi for ν and for each element of the matrix $\mathbf{A}$ is large. The estimates are highly volatile.

$>$ - - - Insert Table 2 - - - $<$

With regard to the forecasts for the explained variable, i.e. $\widehat{E(\mathbf{\Sigma}_t^{-1}|\cdot)}$, we again calculated how often the true value of the logarithm of the determinant of the inverse variance covariance matrix for each date and all data sets falls into the correspondingly estimated 95% hpdi. The findings for the condensed statistic are listed in the last column of Table 3. The results are quite satisfying and promising for practical applications. On average, for 94% of the 100 dates the condensed statistic of the covariance matrices lies in the estimated 95% hpdi. Consequently, the forecast for the covariance is appropriate and its time dependence is accurately estimated employing the 95% hpdi. However, regarding the $MAPE$ the forecasts for the covariance matrices are not estimated well via the mean of the sampling observations ($\overline{MAPE}$ for $E(\widehat{\boldsymbol{\Sigma}_t^{-1}|\cdot})_{11} = 0.38$, $\overline{MAPE}$ for $E(\widehat{\boldsymbol{\Sigma}_t^{-1}|\cdot})_{12} = 0.47$, $\overline{MAPE}$ for $E(\widehat{\boldsymbol{\Sigma}_t^{-1}|\cdot})_{22} = 0.38$, where $\overline{MAPE}$ denotes the mean of the $MAPE$s over all data sets). For the condensed statistic we have: $\overline{MAPE} = 0.02$.

> - - - Insert Table 3 - - - <

5 CONCLUSIONS

Assuming that covariance matrices follow a stochastic process instead of having a deterministic structure gives more flexibility in modeling stylized facts and observed features of e.g. volatilities of assets. Possible applications in economics are portfolio management, risk management or pricing derivatives.

Philipov and Glickman (2006) show in their paper an interesting approach modeling multivariate stochastic volatility via Wishart processes. We pick up on their model and upgrade the theoretical derivations. Focus is put on the practical implementation of the approach using Bayesian estimation methods and on the quality of the estimators. The latter is analyzed profoundly using a Monte Carlo study.

Our analysis shows that the model parameters are estimated unbiased, but the corresponding 95% hpdi can be quite large. Whether the quality of the estimate of $\mathbf{\Sigma}_t^{-1}$ is sufficient for practical applications cannot be definitively answered. When employing the 95% hpdi the true values of $\mathbf{\Sigma}_t^{-1}$ are appropriately modeled. Estimating $\mathbf{\Sigma}_t^{-1}$ merely via the posterior mean of the sampling observations gives a large mean absolute percentage error that must be considered for applications.

APPENDIX A: DERIVATIONS

A.1 Joint Posterior Distribution

$$
\begin{aligned}
p\left(\boldsymbol{\Sigma}^{-1}, \mathbf{A}^{-1}, \nu, d|\mathbf{y}\right) \;\; \propto \;\; & p\left(\mathbf{A}^{-1}\right) \cdot p(\nu) \cdot p(d) \cdot \prod_{t=1}^{T} p\left(\boldsymbol{\Sigma}_t^{-1}\right) \cdot p\left(\mathbf{y}_t|\boldsymbol{\Sigma}_t^{-1}, \mathbf{A}^{-1}, \nu, d\right) \\
\propto \;\; & \left(2^{\gamma_0 k/2} \pi^{k(k-1)/4} \prod_{i=1}^{k} \Gamma\left(\frac{\gamma_0+1-i}{2}\right)\right)^{-1} \cdot |\mathbf{Q}_0|^{-\gamma_0/2} \cdot \\
& \left|\mathbf{A}^{-1}\right|^{(\gamma_0-k-1)/2} \exp\left(-\frac{1}{2} tr\left(\mathbf{Q}_0^{-1}\mathbf{A}^{-1}\right)\right) \cdot \\
& \prod_{t=1}^{T} \left[\left(2^{\nu k/2} \pi^{k(k-1)/4} \prod_{i=1}^{k} \Gamma\left(\frac{\nu+1-i}{2}\right)\right)^{-1} \cdot |\mathbf{S}_{t-1}|^{-\nu/2} \cdot \right. \\
& \left|\boldsymbol{\Sigma}_t^{-1}\right|^{(\nu-k-1)/2} \cdot \exp\left(-\frac{1}{2} tr\left(\mathbf{S}_{t-1}^{-1}\boldsymbol{\Sigma}_t^{-1}\right)\right) \cdot (2\pi)^{-k/2} \cdot \\
& \left. \left|\boldsymbol{\Sigma}_t^{-1}\right|^{1/2} \cdot \exp\left(-\frac{1}{2}\mathbf{y}_t' \boldsymbol{\Sigma}_t^{-1} \mathbf{y}_t\right)\right]
\end{aligned}
$$

A.2 Full Conditional of d

$$
\begin{aligned}
p\left(d|\cdot\right) \;\; \propto \;\; & \prod_{t=1}^{T} \left[\left(|\mathbf{S}_{t-1}|^{-\nu/2} \cdot \exp\left(-\frac{1}{2} tr\left(\mathbf{S}_{t-1}^{-1}\boldsymbol{\Sigma}_t^{-1}\right)\right)\right)\right] \\
\propto \;\; & \prod_{t=1}^{T} \left[\left|\boldsymbol{\Sigma}_{t-1}^{-1}\right|^{-\nu d/2} \cdot \exp\left(-\frac{1}{2}\, tr\left(\mathbf{S}_{t-1}^{-1}\boldsymbol{\Sigma}_t^{-1}\right)\right)\right]
\end{aligned}
$$

A.3 Full Conditional of ν

$$
\begin{aligned}
p(\nu|\cdot) &\propto \prod_{t=1}^{T}\left(2^{\nu k/2}\pi^{k(k-1)/4}\prod_{i=1}^{k}\Gamma\left(\frac{\nu+1-i}{2}\right)\right)^{-1}\cdot|\mathbf{S}_{t-1}|^{-\nu/2} \\
&\quad \left|\boldsymbol{\Sigma}_t^{-1}\right|^{(\nu-k-1)/2}\exp\left(-\frac{1}{2}tr\left(\mathbf{S}_{t-1}^{-1}\boldsymbol{\Sigma}_t^{-1}\right)\right) \\
&\propto \left(\frac{\left|\nu\mathbf{A}^{-1}\right|^{\nu/2}}{2^{\nu k/2}\prod_{i=1}^{k}\Gamma\left(\frac{\nu+1-i}{2}\right)}\right)^{T}\prod_{t=1}^{T}|\boldsymbol{\Sigma}_{t-1}|^{\nu d/2} \\
&\quad \prod_{t=1}^{T}\left|\boldsymbol{\Sigma}_t^{-1}\right|^{\nu/2}\cdot\exp\left(-\frac{1}{2}\sum_{t=1}^{T}tr\left(\mathbf{S}_{t-1}^{-1}\boldsymbol{\Sigma}_t^{-1}\right)\right)
\end{aligned}
$$

A.4 Full Conditional of $\boldsymbol{\Sigma}_t^{-1}$

$$
\begin{aligned}
p\left(\boldsymbol{\Sigma}_t^{-1}|\cdot\right) &\propto \left|\boldsymbol{\Sigma}_t^{-1}\right|^{(\nu-k-1)/2}\cdot\exp\left(-\frac{1}{2}\,tr\left(\mathbf{S}_{t-1}^{-1}\boldsymbol{\Sigma}_t^{-1}\right)\right)\cdot|\boldsymbol{\Sigma}_t|^{-\frac{1}{2}}\cdot \\
&\quad \exp\left(-\frac{1}{2}\,tr\left(\mathbf{y}_t'\boldsymbol{\Sigma}_t^{-1}\mathbf{y}_t\right)\right)\cdot|\mathbf{S}_t|^{-\nu/2}\exp\left(-\frac{1}{2}\,tr\left(\mathbf{S}_t^{-1}\boldsymbol{\Sigma}_{t+1}^{-1}\right)\right) \\
&\propto \left|\boldsymbol{\Sigma}_t^{-1}\right|^{(\nu-k-1)/2}\cdot\exp\left(-\frac{1}{2}\,tr\left(\left(\mathbf{S}_{t-1}^{-1}+\mathbf{y}_t\mathbf{y}_t'\right)\boldsymbol{\Sigma}_t^{-1}\right)\right)\cdot \\
&\quad \left|\boldsymbol{\Sigma}_t^{-1}\right|^{(1-\nu d)/2}\cdot\exp\left(-\frac{1}{2}\,tr\left(\mathbf{S}_t^{-1}\boldsymbol{\Sigma}_{t+1}^{-1}\right)\right) \\
&\propto \mathrm{Wish}\left(\boldsymbol{\Sigma}_t^{-1}|\nu,\widetilde{\mathbf{S}}_{t-1}\right)\cdot \\
&\quad \left|\boldsymbol{\Sigma}_t^{-1}\right|^{(1-\nu d)/2}\cdot\exp\left(-\frac{1}{2}\,tr\left(\mathbf{S}_t^{-1}\boldsymbol{\Sigma}_{t+1}^{-1}\right)\right),
\end{aligned}
$$

where $\widetilde{\mathbf{S}}_{t-1}=\left(\mathbf{S}_{t-1}^{-1}+\mathbf{y}_t\mathbf{y}_t'\right)^{-1}$.

A.5 Full Conditional of $\boldsymbol{\Sigma}_T^{-1}$

$$
\begin{aligned}
p\left(\boldsymbol{\Sigma}_T^{-1}|\cdot\right) &\propto \left|\boldsymbol{\Sigma}_T^{-1}\right|^{(\nu-k-1)/2} \cdot \exp\left(-\frac{1}{2}tr\left(\mathbf{S}_{T-1}^{-1}\boldsymbol{\Sigma}_T^{-1}\right)\right) \cdot \\
&\quad \left|\boldsymbol{\Sigma}_T\right|^{-1/2} \cdot \exp\left(-\frac{1}{2}\mathbf{y}_T'\boldsymbol{\Sigma}_T^{-1}\mathbf{y}_T\right) \\
&\propto \left|\boldsymbol{\Sigma}_T^{-1}\right|^{(\nu+1-k-1)/2} \exp\left(-\frac{1}{2}tr\left(\left(\mathbf{S}_{T-1}^{-1}+\mathbf{y}_T\mathbf{y}_T'\right)\boldsymbol{\Sigma}_T^{-1}\right)\right) \\
&\propto \text{Wish}\left(\boldsymbol{\Sigma}_T^{-1}\left|\nu+1, \left(\mathbf{S}_{T-1}^{-1}+\mathbf{y}_T\mathbf{y}_T'\right)^{-1}\right.\right)
\end{aligned}
$$

A.6 Full Conditional of $\mathbf{A}^{-1}$

$$
\begin{aligned}
p\left(\mathbf{A}^{-1}|\cdot\right) &\propto \left|\mathbf{A}^{-1}\right|^{(\gamma_0-k-1)/2} \cdot \exp\left(-\frac{1}{2}tr\left(\mathbf{Q}_0^{-1}\mathbf{A}^{-1}\right)\right) \cdot \\
&\quad \prod_{t=1}^{T}\left[\left|\mathbf{S}_{t-1}\right|^{-\nu/2} \cdot \exp\left(-\frac{1}{2}\,tr\left(\mathbf{S}_{t-1}^{-1}\boldsymbol{\Sigma}_t^{-1}\right)\right)\right] \\
&\propto \left|\mathbf{A}^{-1}\right|^{(\gamma_0+\nu T-k-1)/2} \cdot \exp\left(-\frac{1}{2}\,tr\left(\mathbf{Q}_0^{-1}\mathbf{A}^{-1}\right)\right) \cdot \\
&\quad \prod_{t=1}^{T}\exp\left(-\frac{1}{2}\,tr\left(\mathbf{S}_{t-1}^{-1}\boldsymbol{\Sigma}_t^{-1}\right)\right)
\end{aligned}
$$

APPENDIX B: TYPOGRAPHICAL ERRORS

Pages 315 - 317

$E\left(\boldsymbol{\Sigma}_t \mid \mathbf{A},\boldsymbol{\Sigma}_{t-1}\right)=\frac{\nu}{\nu-k-1}\left(\mathbf{A}^{-1/2}\right)\boldsymbol{\Sigma}_{t-1}^{d}\left(\mathbf{A}^{-1/2}\right)'$
$E\left(\boldsymbol{\Sigma}_t \mid \mathbf{A},\boldsymbol{\Sigma}_{t-1}\right)=\frac{\nu}{\nu-k-1}\left(\mathbf{A}^{1/2\prime}\right)^{-1}\boldsymbol{\Sigma}_{t-1}^{-d}\left(\mathbf{A}^{1/2}\right)^{-1}$
...Cholesky decomposition of $\left(\boldsymbol{\Sigma}_t^{-1}\right)^{d}$...
...Cholesky decomposition of $\left(\boldsymbol{\Sigma}_{t-1}^{-1}\right)^{d}$...
$\Gamma\left(\frac{\nu+j-1}{2}\right)$
$\Gamma\left(\frac{\nu+1-j}{2}\right)$
$p\left(\boldsymbol{\Sigma}^{-1},\mathbf{A},\nu,d \mid \mathbf{y}\right)\propto \mathrm{Wish}\left(\gamma_0,\mathbf{Q}_0\right)\mathrm{Gam}(\nu)\ldots$
$p\left(\boldsymbol{\Sigma}^{-1},\mathbf{A}^{-1},\nu,d \mid \mathbf{y}\right)\propto \mathrm{Wish}\left(\mathbf{A}^{-1} \mid \gamma_0,\mathbf{Q}_0\right)\mathrm{Gam}(\nu-k \mid \alpha,\beta)\ldots$

Note: The gray lines show the erroneous term; the white lines demonstrate the corresponding correct term.

Page 326

draw the last $\boldsymbol{\Sigma}$ directly
draw the last $\boldsymbol{\Sigma}^{-1}$ directly
$\widetilde{\mathbf{S}}_{t-1} = \mathbf{S}_{t-1}^{-1} + \mathbf{y}_t\mathbf{y}_t'$
$\widetilde{\mathbf{S}}_{t-1} = \left(\mathbf{S}_{t-1}^{-1} + \mathbf{y}_t\mathbf{y}_t'\right)^{-1}$
$\widetilde{\nu} = \nu(1-d)+1$
$\widetilde{\nu} = \nu$
$\widetilde{\nu} = \nu(1-d)+1$
$\widetilde{\nu} = \nu+1$
$\text{Wish}\left(\boldsymbol{\Sigma}_t \vert \nu, \widetilde{\mathbf{S}}_{t-1}\right)$
$\text{Wish}\left(\boldsymbol{\Sigma}_t^{-1} \vert \nu, \widetilde{\mathbf{S}}_{t-1}\right)$
$\text{AR=}\ldots\exp\left(-\frac{1}{2}tr\left(\nu\,\mathbf{A}^{-1}(\boldsymbol{\Sigma}_t^{*})^{-d}\boldsymbol{\Sigma}_{t+1}^{-1}\right)\right)\,/$ $\exp\left(-\frac{1}{2}tr\left(\nu\,\mathbf{A}^{-1}(\boldsymbol{\Sigma}_t^{[m-1]})^{-d}\boldsymbol{\Sigma}_{t+1}^{-1}\right)\right)$
$\text{AR=}\ldots\exp\left(-\frac{1}{2}tr\left(\nu\,\left(\mathbf{A}^{1/2'}\right)^{-1}(\boldsymbol{\Sigma}_t^{*})^{d}\left(\mathbf{A}^{1/2}\right)^{-1}\boldsymbol{\Sigma}_{t+1}^{-1}\right)\right)\,/$ $\exp\left(-\frac{1}{2}tr\left(\nu\,\left(\mathbf{A}^{1/2'}\right)^{-1}(\boldsymbol{\Sigma}_t^{[m-1]})^{d}\left(\mathbf{A}^{1/2}\right)^{-1}\boldsymbol{\Sigma}_{t+1}^{-1}\right)\right)$

Page 327

$\mathbf{Q}^{-1} = \sum_{t=1}^{T}\left(\boldsymbol{\Sigma}_t^{-1}\right)^{-d/2}\boldsymbol{\Sigma}_t^{-1}\left(\boldsymbol{\Sigma}_{t-1}^{-1}\right)^{-d/2}$
$\mathbf{Q}^{-1} = \sum_{t=1}^{T}\left(\boldsymbol{\Sigma}_{t-1}^{-1}\right)^{-d/2}\boldsymbol{\Sigma}_t^{-1}\left(\boldsymbol{\Sigma}_{t-1}^{-1}\right)^{-d/2}$
$p(d \vert rest) \propto p(d)\prod_{t=1}^{T}\left[\left\lvert\boldsymbol{\Sigma}_t^{-1}\right\rvert^{-d\nu/2}\exp\left(-\frac{1}{2}tr\left(\mathbf{S}_t^{-1}\left(\boldsymbol{\Sigma}_{t-1}^{-1}\right)^{-d}\right)\right)\right]$
$p(d \vert rest) \propto \prod_{t=1}^{T}\left[\left\lvert\boldsymbol{\Sigma}_{t-1}^{-1}\right\rvert^{-d\nu/2}\exp\left(-\frac{1}{2}tr\left(\mathbf{S}_{t-1}^{-1}\boldsymbol{\Sigma}_t^{-1}\right)\right)\right]$
$\mathbf{A}^{-1}\mathbf{Q}(d)^{-1}$
$\sum_{t=1}^{T}\mathbf{A}^{-1/2'}\boldsymbol{\Sigma}_{t-1}^{d}\mathbf{A}^{-1/2}\boldsymbol{\Sigma}_t^{-1}$
$\psi = -\frac{\nu}{2}\sum_{t=1}^{T}\ln\left\lvert\boldsymbol{\Sigma}_t^{-1}\right\rvert$
$\psi = -\frac{\nu}{2}\sum_{t=1}^{T}\ln\left\lvert\boldsymbol{\Sigma}_{t-1}^{-1}\right\rvert$
$p(\nu \vert rest) \propto \ldots\left(\frac{\left\lvert\nu\mathbf{A}^{-1}\right\rvert^{\nu/2}}{2^{\nu k}\prod_{j=1}^{k}\Gamma\left(\frac{\nu+j-1}{2}\right)}\right)^{T}$
$p(\nu \vert rest) \propto \ldots\left(\frac{\left\lvert\nu\mathbf{A}^{-1}\right\rvert^{\nu/2}}{2^{\nu k/2}\prod_{j=1}^{k}\Gamma\left(\frac{\nu+1-j}{2}\right)}\right)^{T}$
$\mathbf{A}^{-1}\mathbf{Q}_t^{-1}$
$\mathbf{A}^{-1/2'}\boldsymbol{\Sigma}_{t-1}^{d}\mathbf{A}^{-1/2}\boldsymbol{\Sigma}_t^{-1}$

Note: The gray lines show the erroneous term; the white lines demonstrate the corresponding correct term.

REFERENCES

GUPTA, A., AND D. NAGAR (2000): *Matrix variate Distributions.* Chapman and Hall/CRC, Florida.

HORN, R., AND C. JOHNSON (1985): *Matrix Analysis.* Cambridge University Press, Cambridge, Massachusetts.

KOOP, G. (2006): *Bayesian Econometrics.* Wiley, West Sussex.

PHILIPOV, A., AND M. GLICKMAN (2006): "Multivariate Stochastic Volatility via Wishart Processes," *Journal of Business and Economic Statistics*, 24(3), 313–328.

Table 1: Descriptive statistics of the sampling observations for ν, d, the elements of $\mathbf{A}$, and $log|\mathbf{A}|$ for the example data set.

	True Value	*Posterior Mean*	*Posterior Median*	*Standard Deviation*	*Acceptance Rate*
ν	20	19.59	19.41	1.51	0.50
d	0.7	0.71	0.71	0.02	0.43
$\mathbf{A}_{11}$	130	98.96	98.96	3.06	0.40
$\mathbf{A}_{12}$	50	17.70	17.92	4.06	
$\mathbf{A}_{22}$	130	116.33	116.33	3.00	
$log\|\mathbf{A}\|$	9.58	9.32	9.19	0.04	

Table 2: Statistics of the sampling observations for each element of the matrix $\mathbf{A}$ ($\mathbf{A}_{11}, \mathbf{A}_{12}, \mathbf{A}_{22}$), $log|\mathbf{A}|$, ν, and d for all Monte Carlo data sets. 'AR' denotes the average acceptance rate for all 100 data sets. The 95% hpdi are given in the last row.

	$\mathbf{A}_{11}$	$\mathbf{A}_{12}$	$\mathbf{A}_{22}$	$log\|\mathbf{A}\|$	ν	d
true value	130	50	130	9.58	20	0.7
observations	100	100	100	100	100	100
mean	91.95	33.67	92.58	8.74	28.43	0.73
median	83.25	32.64	80.77	8.64	26.05	0.74
standard deviation	37.74	22.71	39.75	0.70	15.61	0.02
minimum	28.77	-10.58	30.18	6.76	3.89	0.63
25th percentile	68.18	18.50	67.79	8.18	15.39	0.72
75th percentile	105.84	42.64	106.27	9.25	37.29	0.74
maximum	254.94	154.01	311.18	11.16	85.38	0.78
AR	0.44	0.44	0.44	0.44	0.57	0.38
95% hpdi	[48.47; 173.09]	[-1.78; 84.42]	[30.18; 159.63]	[7.65; 10.19]	[3.89; 57.14]	[0.68; 0.76]

Table 3: Analysis of the forecast for $\mathbf{\Sigma}_t$. Statistical characteristics of the fraction of the true values lying in their estimated 95% hpdi for each of the elements of $\widehat{E(\widehat{\mathbf{\Sigma}_t^{-1}}|\cdot)}$ and for the condensed statistic $log|\widehat{E(\widehat{\mathbf{\Sigma}_t^{-1}}|\cdot)}|$ are given.

	$\widehat{E(\widehat{\mathbf{\Sigma}_t^{-1}}|\cdot)}_{ij}$			$log|\widehat{E(\widehat{\mathbf{\Sigma}_t^{-1}}|\cdot)}|$
ij	11	12	22	
observations	100×100	100×100	100×100	100×100
mean	0.92	0.91	0.93	0.94
median	0.94	0.94	0.96	0.95
standard deviation	0.10	0.12	0.09	0.06
minimum	0.41	0.32	0.53	0.69
25th percentile	0.92	0.90	0.92	0.91
75th percentile	0.98	0.97	0.98	0.99
maximum	1	1	1	1

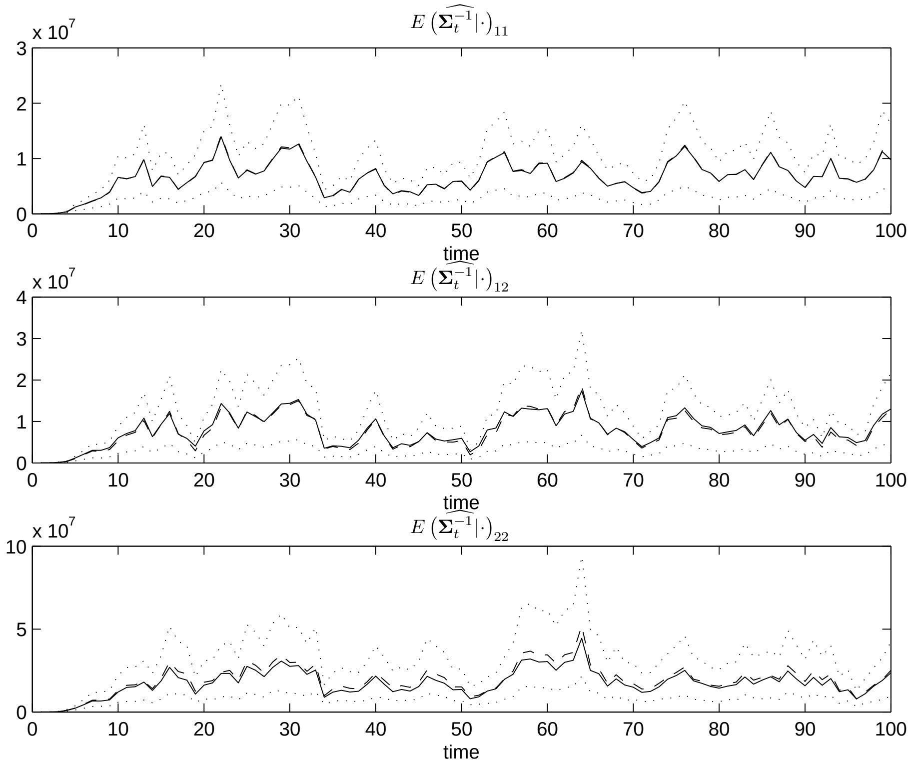

Figure 1: Plot of the elements of $E\left(\widehat{\boldsymbol{\Sigma}_t^{-1}|\cdot}\right)$ for the example data set, where the dotted line is the 95% hpdi, the solid line denotes the true values, and the dashed line corresponds to the mean of the sampling observations.

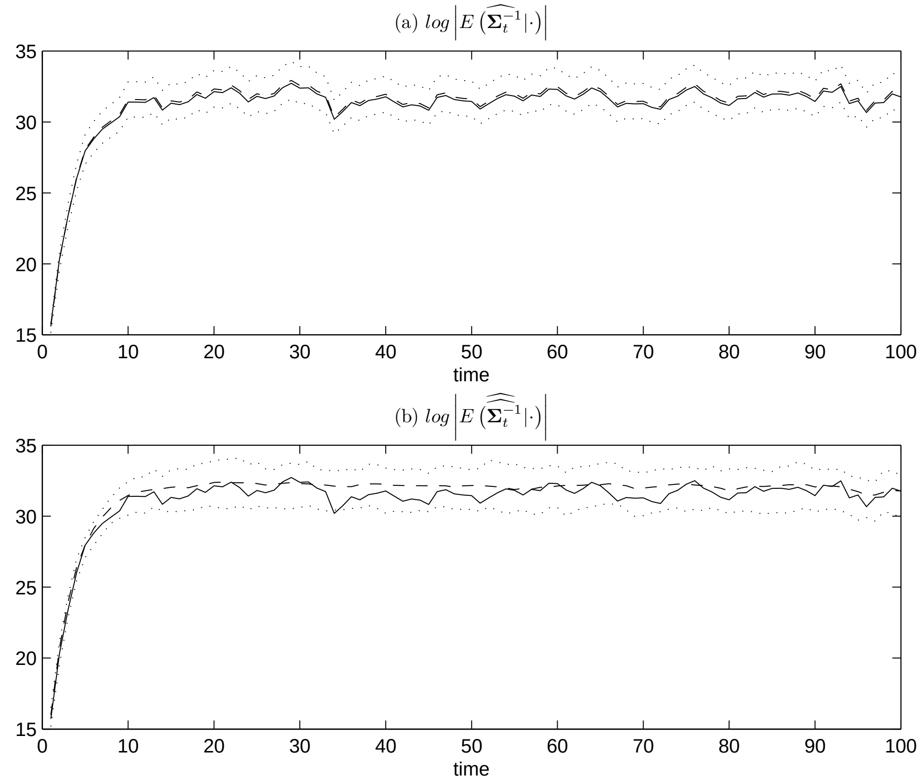

Figure 2: Plot of the estimates of $log|E\left(\widehat{\boldsymbol{\Sigma}_t^{-1}}|\cdot\right)|$ (- - -), $log|E\left(\widehat{\widehat{\boldsymbol{\Sigma}_t^{-1}}}|\cdot\right)|$ (- - -), their 95% hpdi ($\cdots$), and their true values (—) for the example data set.

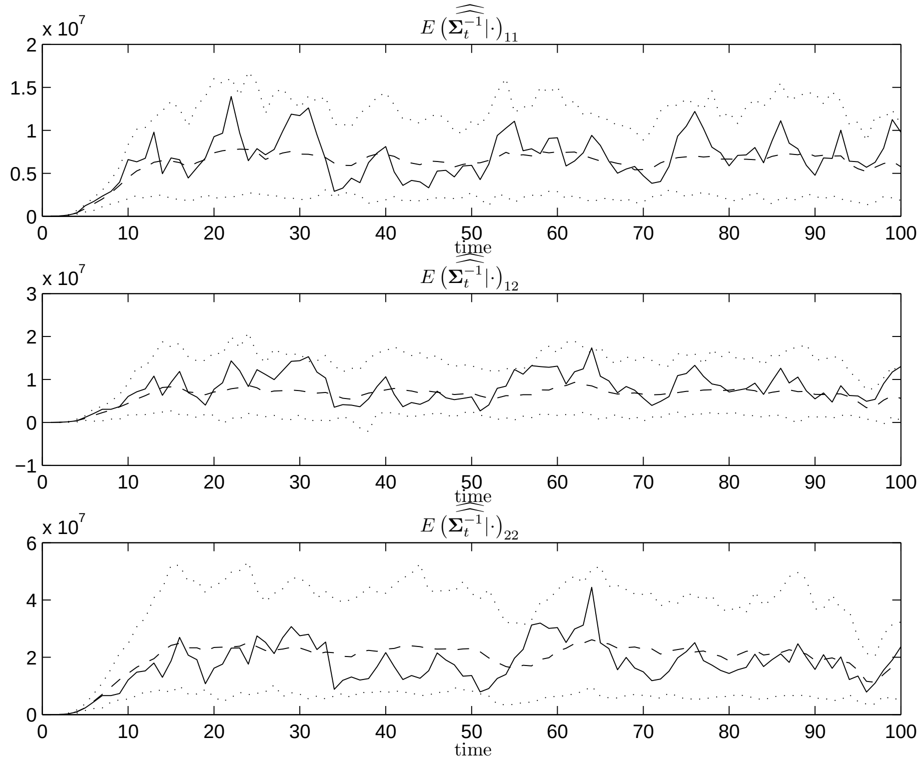

Figure 3: Plot of the elements of $E\left(\widehat{\widehat{\boldsymbol{\Sigma}_t^{-1}}}|\cdot\right)$ for the example data set, where the dotted line is the 95% hpdi, the solid line denotes the true values, and the dashed line corresponds to the mean of the sampling observations.

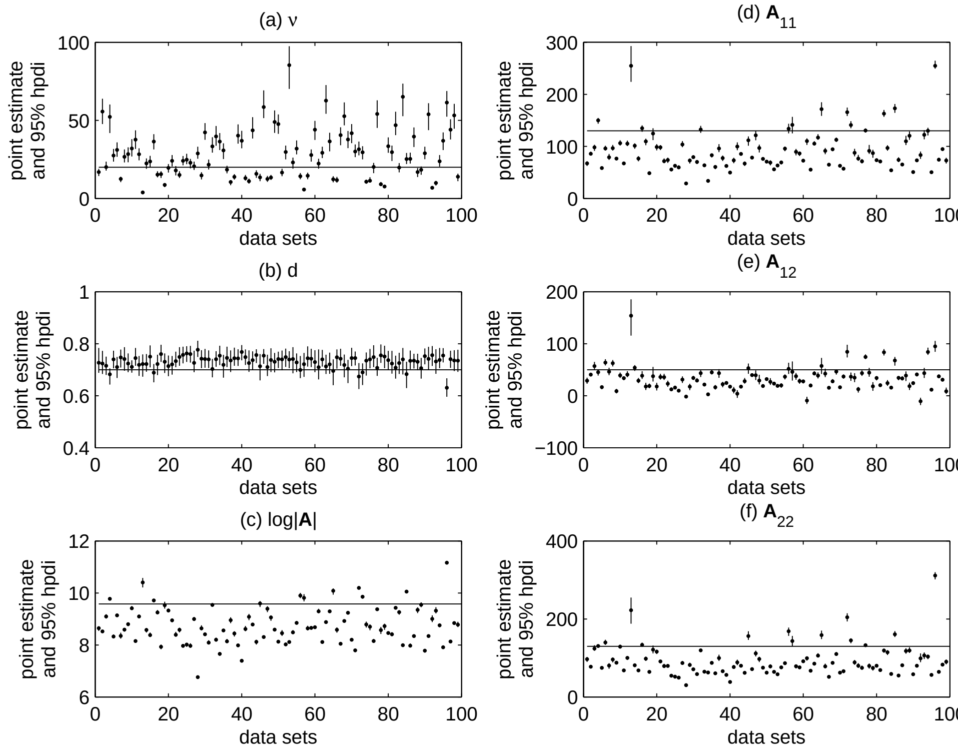

Figure 4: Estimation results (- - -) for ν, d, $log|\mathbf{A}|$, $\mathbf{A}_{11}$, $\mathbf{A}_{12}$, and $\mathbf{A}_{22}$ for each data set. True values are denoted by the horizontal line, and the corresponding 95% hpdi by the vertical lines.

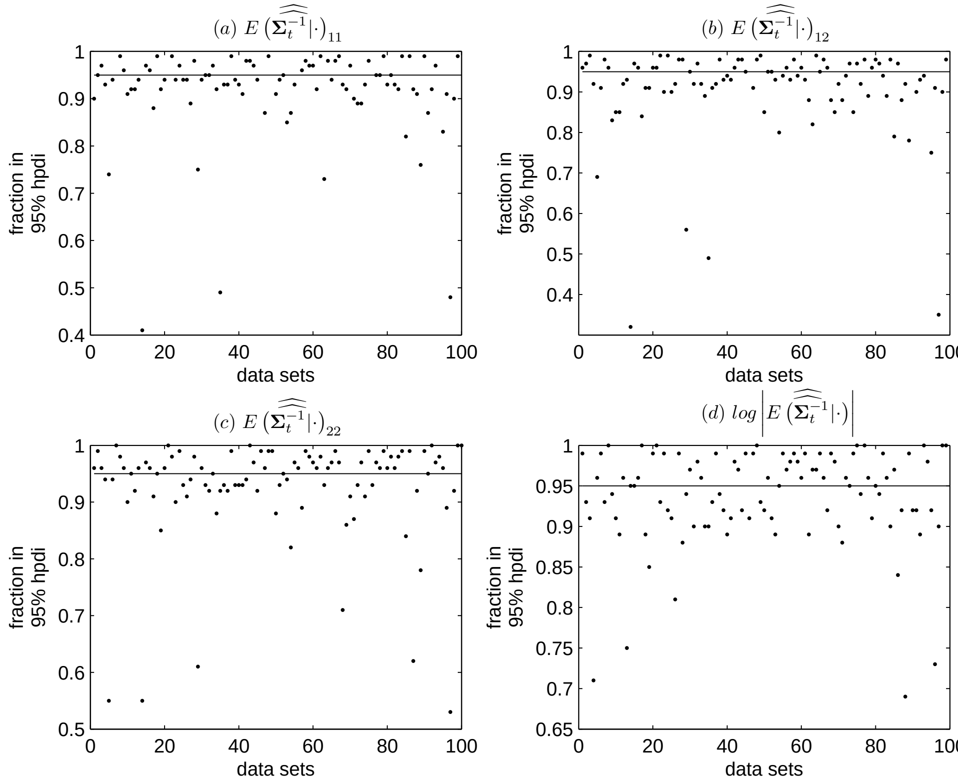

Figure 5: The fraction of the true values of the conditional expected value of the explained variable lying in the correspondingly estimated 95% hpdi is plotted. The first three graphs show the findings for the individual elements of $\widehat{E(\mathbf{\Sigma}_t^{-1}|\cdot)}$; the last plot gives the results for $log|\widehat{E(\mathbf{\Sigma}_t^{-1}|\cdot)}|$.

University of Innsbruck - Working Papers in Economics and Statistics
Recent Papers can be accessed on the following webpage:

http://eeecon.uibk.ac.at/wopec/

2011-19 **Wolfgang Rinnergschwentner, Gottfried Tappeiner, Janette Walde:** Multivariate stochastic volatility via wishart processes - A continuation

2011-18 **Jan Verbesselt, Achim Zeileis, Martin Herold:** Near Real-Time Disturbance Detection in Terrestrial Ecosystems Using Satellite Image Time Series: Drought Detection in Somalia

2011-17 **Stefan Borsky, Andrea Leiter, Michael Pfaffermayr:** Does going green pay off? The effect of an international environmental agreement on tropical timber trade

2011-16 **Pavlo Blavatskyy:** Stronger Utility

2011-15 **Anita Gantner, Wolfgang Höchtl, Rupert Sausgruber:** The pivotal mechanism revisited: Some evidence on group manipulation

2011-14 **David J. Cooper, Matthias Sutter:** Role selection and team performance

2011-13 **Wolfgang Höchtl, Rupert Sausgruber, Jean-Robert Tyran:** Inequality aversion and voting on redistribution

2011-12 **Thomas Windberger, Achim Zeileis:** Structural breaks in inflation dynamics within the European Monetary Union

2011-11 **Loukas Balafoutas, Adrian Beck, Rudolf Kerschbamer, Matthias Sutter:** What drives taxi drivers? A field experiment on fraud in a market for credence goods

2011-10 **Stefan Borsky, Paul A. Raschky:** A spatial econometric analysis of compliance with an international environmental agreement on open access resources

2011-09 **Edgar C. Merkle, Achim Zeileis:** Generalized measurement invariance tests with application to factor analysis

2011-08 **Michael Kirchler, Jürgen Huber, Thomas Stöckl:** Thar she bursts - reducing confusion reduces bubbles *modified version forthcoming in American Economic Review*

2011-07 **Ernst Fehr, Daniela Rützler, Matthias Sutter:** The development of egalitarianism, altruism, spite and parochialism in childhood and adolescence

2011-06 **Octavio Fernández-Amador, Martin Gächter, Martin Larch, Georg Peter:** Monetary policy and its impact on stock market liquidity: Evidence from the euro zone

2011-05 **Martin Gächter, Peter Schwazer, Engelbert Theurl:** Entry and exit of physicians in a two-tiered public/private health care system

2011-04 **Loukas Balafoutas, Rudolf Kerschbamer, Matthias Sutter:** Distributional preferences and competitive behavior *forthcoming in Journal of Economic Behavior and Organization*

2011-03 **Francesco Feri, Alessandro Innocenti, Paolo Pin:** Psychological pressure in competitive environments: Evidence from a randomized natural experiment: Comment

2011-02 **Christian Kleiber, Achim Zeileis:** Reproducible Econometric Simulations

2011-01 **Carolin Strobl, Julia Kopf, Achim Zeileis:** A new method for detecting differential item functioning in the Rasch model

2010-29 **Matthias Sutter, Martin G. Kocher, Daniela Rützler and Stefan T. Trautmann:** Impatience and uncertainty: Experimental decisions predict adolescents' field behavior

2010-28 **Peter Martinsson, Katarina Nordblom, Daniela Rützler and Matthias Sutter:** Social preferences during childhood and the role of gender and age - An experiment in Austria and Sweden *Revised version forthcoming in Economics Letters*

2010-27 **Francesco Feri and Anita Gantner:** Baragining or searching for a better price? - An experimental study. *Revised version accepted for publication in Games and Economic Behavior*

2010-26 **Loukas Balafoutas, Martin G. Kocher, Louis Putterman and Matthias Sutter:** Equality, equity and incentives: An experiment

2010-25 **Jesús Crespo-Cuaresma and Octavio Fernández Amador:** Business cycle convergence in EMU: A second look at the second moment

2010-24 **Lorenz Goette, David Huffman, Stephan Meier and Matthias Sutter:** Group membership, competition and altruistic versus antisocial punishment: Evidence from randomly assigned army groups

2010-23 **Martin Gächter and Engelbert Theurl:** Health status convergence at the local level: Empirical evidence from Austria *(revised Version March 2011)*

2010-22 **Jesús Crespo-Cuaresma and Octavio Fernández Amador:** Buiness cycle convergence in the EMU: A first look at the second moment

2010-21 **Octavio Fernández-Amador, Josef Baumgartner and Jesús Crespo-Cuaresma:** Milking the prices: The role of asymmetries in the price transmission mechanism for milk products in Austria

2010-20 **Fredrik Carlsson, Haoran He, Peter Martinsson, Ping Qin and Matthias Sutter:** Household decision making in rural China: Using experiments to estimate the influences of spouses

2010-19 **Wolfgang Brunauer, Stefan Lang and Nikolaus Umlauf:** Modeling house prices using multilevel structured additive regression

2010-18 **Martin Gächter and Engelbert Theurl:** Socioeconomic environment and mortality: A two-level decomposition by sex and cause of death

2010-17 **Boris Maciejovsky, Matthias Sutter, David V. Budescu and Patrick Bernau:** Teams make you smarter: Learning and knowledge transfer in auctions and markets by teams and individuals

2010-16 **Martin Gächter, Peter Schwazer and Engelbert Theurl:** Stronger sex but earlier death: A multi-level socioeconomic analysis of gender differences in mortality in Austria

2010-15 **Simon Czermak, Francesco Feri, Daniela Rützler and Matthias Sutter:** Strategic sophistication of adolescents - Evidence from experimental normal-form games

2010-14 **Matthias Sutter and Daniela Rützler:** Gender differences in competition emerge early in live

2010-13 **Matthias Sutter, Francesco Feri, Martin G. Kocher, Peter Martinsson, Katarina Nordblom and Daniela Rützler:** Social preferences in childhood and adolescence - A large-scale experiment

2010-12 **Loukas Balafoutas and Matthias Sutter:** Gender, competition and the efficiency of policy interventions

2010-11 **Alexander Strasak, Nikolaus Umlauf, Ruth Pfeifer and Stefan Lang:** Comparing penalized splines and fractional polynomials for flexible modeling of the effects of continuous predictor variables

2010-10 **Wolfgang A. Brunauer, Sebastian Keiler and Stefan Lang:** Trading strategies and trading profits in experimental asset markets with cumulative information

2010-09 **Thomas Stöckl and Michael Kirchler:** Trading strategies and trading profits in experimental asset markets with cumulative information

2010-08 **Martin G. Kocher, Marc V. Lenz and Matthias Sutter:** Psychological pressure in competitive environments: Evidence from a randomized natural experiment: Comment

2010-07 **Michael Hanke and Michael Kirchler:** Football Championships and Jersey sponsors' stock prices: An empirical investigation

2010-06 **Adrian Beck, Rudolf Kerschbamer, Jianying Qiu and Matthias Sutter:** Guilt from promise-breaking and trust in markets for expert services - Theory and experiment

2010-05 **Martin Gächter, David A. Savage and Benno Torgler:** Retaining the thin blue line: What shapes workers' intentions not to quit the current work environment

2010-04 **Martin Gächter, David A. Savage and Benno Torgler:** The relationship between stress, strain and social capital

2010-03 **Paul A. Raschky, Reimund Schwarze, Manijeh Schwindt and Ferdinand Zahn:** Uncertainty of governmental relief and the crowding out of insurance

2010-02 **Matthias Sutter, Simon Czermak and Francesco Feri:** Strategic sophistication of individuals and teams in experimental normal-form games

2010-01 **Stefan Lang and Nikolaus Umlauf:** Applications of multilevel structured additive regression models to insurance data

University of Innsbruck

Working Papers in Economics and Statistics

2011-19

Wolfgang Rinnergschwentner, Gottfried Tappeiner, Janette Walde

Multivariate stochastic volatility via wishart processes - A continuation

Abstract
This paper picks up on a model developed by Philipov and Glickman (2006) for modeling multivariate stochastic volatility via Wishart processes. MCMC simulation from the posterior distribution is employed to fit the model. However, erroneous mathematical transformations in the full conditionals cause false implementation of the approach. We adjust the model, upgrade the analysis and investigate the statistical properties of the estimators using an extensive Monte Carlo study. Employing a Gibbs sampler in combination with a Metropolis Hastings algorithm inference for the time-dependent covariance matrix is feasible with appropriate statistical properties.

ISSN 1993-4378 (Print)
ISSN 1993-6885 (Online)