

Anna Pajor* and Justyna Wróblewska

VEC-MSF models in Bayesian analysis of short- and long-run relationships

DOI 10.1515/snde-2016-0004

Abstract: The paper aims at developing new Bayesian Vector Error Correction – Stochastic Volatility (VEC-SV) models, which combine the VEC representation of a VAR structure with stochastic volatility, represented by either the multiplicative stochastic factor (MSF) process or the MSF-SBEKK specification. Appropriate numerical methods (MCMC-based algorithms) are adapted for estimation and comparison of these type of models. Based on data coming from the Polish economy (time series of unemployment, inflation, interest rates, and of PLN/EUR, PLN/USD and EUR/USD exchange rates) it is shown that the models and numerical methods proposed in our study work well in simultaneous modelling of volatility and long-run relationships.

Keywords: Bayesian analysis; cointegration; Markov Chain Monte Carlo methods; stochastic volatility.

JEL Classification: C11; C32.

1 Introduction

A frequently used model structure for macroeconomic time series is the vector autoregression with the error correction mechanism and with a constant conditional covariance matrix. In turn, conditional heteroscedasticity processes serve as typical models for the volatility of financial time series. However, some research has shown that the variability of macroeconomic time series (measured, for example, by the conditional covariance matrix) can vary over time, and also that on financial markets there exist long-term relationships. Therefore, it appears essential to formulate such models in which the possible presence of long-run relationships and time-variable volatility are simultaneously taken into account. A fast growing econometrics literature (see, e.g. Cogley and Sargent 2005; Primiceri 2005; Sims and Zha 2006; Koop, León-González, and Strachan 2009; Koop and Potter 2011; Koop 2012; Belmonte and Koop 2014) has mainly focused on two particular extensions of a basic vector autoregressive (VAR) model: specifications with time-varying parameters and the ones with time-varying conditional covariances. Our particular attention is paid to the vector error correction (VEC) representation of VAR with the foregoing extensions (see, e.g. Seo 2007; Herwartz and Lütkepohl 2011; Koop, León-González, and Strachan 2011; Cavaliere et al. 2015). However, due to the lack of identification of cointegrating vectors (only the space spanned by the cointegrating vectors can be identified), the literature on the VEC models is much more limited than that on VAR. Also, studies on the Bayesian inference within the corresponding VEC structures are quite rare. Therefore, it is desirable to fill this gap in the literature and to develop Bayesian methods for VEC models allowing for stochastic volatility (SV).

The paper aims at developing new Bayesian Vector Error Correction – Stochastic Volatility (VEC-SV) models, which combine the VEC representation of a VAR structure with stochastic volatility, represented by either the multiplicative stochastic factor (MSF) process or the MSF-SBEKK specification. The MSF structure is one of the simplest amongst the Multivariate Stochastic Volatility (MSV) models (since it features only one latent process, and the conditional covariance matrix does not depend on the past information set), and it can be used to describe time-varying volatility in a parsimonious manner. It is worth mentioning that this SV specification has been used in constructing hybrid models combining multivariate stochastic volatility and

*Corresponding author: Anna Pajor, Department of Mathematics, Cracow University of Economics, Kraków, Poland, e-mail: eopajor@cyf-kr.edu.pl

Justyna Wróblewska: Department of Econometrics and Operations Research, Cracow University of Economics, Kraków, Poland

multivariate generalized conditional heteroscedasticity structures (MSF-SBEKK, among others), proposed by (Osiewalski 2009; Osiewalski and Pajor 2009), and further developed and extended by (Osiewalski and Pajor 2010; Pajor and Osiewalski 2012; Osiewalski and Osiewalski 2012).

Another objective of this paper is to develop Bayesian methodology for estimation and comparison of the models in question. Bayesian statistical analysis in the class of proposed models requires using some numerical methods, in particular Markov Chain Monte Carlo (MCMC) algorithms. Therefore, one of objectives of the research is to adapt relevant numerical methods to allow the estimation, prediction and comparison of such model structures. In particular, we employ the Gibbs sampler and the Metropolis-Hastings procedure.

The paper is organized as follows. In Section 2 the basic framework is set and the Bayesian models are formulated. In Section 3 we elaborate on the numerical methods. Finally, Section 4 is devoted to an empirical study, in which we use a standard set of the Polish macroeconomics variables (including the unemployment, inflation and interest rates), and show that relaxing the usual assumption of conditional homoscedasticity by introducing the MSF or the MSF-SBEKK structures is empirically valid. The resulting VEC-MSF and VEC-MSF-SBEKK models are also used for the analysis of relationships and volatility of selected main official Polish exchange rates.

2 Bayesian VEC-SV models

Consider a linear n -variate and k -order vector autoregressive [VAR(k)] process with deterministic terms and a stochastic volatility (SV) structure, represented in the form of a vector error correction (VEC) process, i.e. Vector Error Correction with Stochastic Volatility (VEC-SV) process:

$$\Delta x_t = \tilde{\Pi}x_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta x_{t-i} + \Phi D_t + \varepsilon_t, \quad t = 1, 2, \dots, T, \quad (1)$$

$$\varepsilon_t | \psi_{t-1}, D_t, q_t, \theta \sim N(0, \Sigma_t), \quad (2)$$

where x_t is an $n \times 1$ random vector, $\tilde{\Pi}$ and Γ_i are $n \times n$ matrices of real coefficients ($i = 1, \dots, k-1$), matrix D_t is comprised of deterministic variables such as the constant and seasonal centred dummies, Φ is a parameter matrix, q_t is a latent variable, θ is a vector of parameters, ψ_{t-1} denotes the past of the process $\{x_t\}$ up to time $t-1$, and $\Sigma_t = \Sigma(q_t, \psi_{t-1})$. Moreover, $\tilde{\Pi} = \alpha \tilde{\beta}'$, with α and $\tilde{\beta}'$ being some $n \times r$ matrices, where $r < n$ is the number of cointegration relationships (if they exist). The initial conditions $x_{-k+1}, x_{-k+2}, \dots, x_0$ are assumed to be known.

2.1 MSF and MSF-SBEKK structures in the VEC model

In this paper we consider two alternative stochastic volatility structures for matrix Σ_t : multiplicative stochastic factor (MSF) and hybrid MSF-SBEKK (type I; see Osiewalski and Pajor 2009).

The multiplicative stochastic factor structure for matrix Σ_t is as follows:

$$\Sigma_t = q_t \Sigma, \quad (3)$$

with $\ln q_t = \phi \ln q_{t-1} + \sigma_q \eta_t$, $\{\eta_t\} \sim iin(0, 1)$.

Henceforth, this specification will be referred to as the VEC-MSF (VEC with Multiplicative Stochastic Factor) process. Although the VEC-MSF process features non-zero time-varying conditional covariances, the conditional correlations remain constant over time. Such a result is attributable to the fact that the very same q_t factor drives the dynamics of each element of Σ_t . The idea of the MSF structure dates back to (Quintana and West 1987), who specified the scaling factors to be some known constants. Under a different name of the stochastic discount factor (SDF) process, the MSF specification was then revived (though only conceptually, without being used in any application) by (Jacquier, Polson, and Rossi 1995) in modelling financial time series volatility. Later, the process was employed, e.g. by (Osiewalski and Pajor 2009; Pajor 2011; Pajor and Osiewalski 2012; Osiewalski and Osiewalski 2013).

As an alternative MSV specification we employ the MSF-SBEKK structure proposed by (Osiewalski 2009). Following (Osiewalski and Pajor 2009), for matrix Σ_t we assume the so-called type I hybrid MSF-SBEKK process:

$$\Sigma_t = q_t \tilde{\Sigma}_t, \quad (4)$$

$$\tilde{\Sigma}_t = (1-a-b)\Sigma + b(\varepsilon_{t-1}\varepsilon_{t-1}') + a\tilde{\Sigma}_{t-1}, \quad (5)$$

with $\ln q_t = \phi \ln q_{t-1} + \sigma_q \eta_t$, $\{\eta_t\} \sim iin(0, 1)$, $a, b \in \mathbb{R}$.

Matrix $\tilde{\Sigma}_t$ is square, of order n and follows the scalar BEKK(1,1) structure. The specification resulting from introducing MSF-SBEKK error term (ε_t) into the VEC model is further referred to as VEC-MSF-SBEKK. The presence of the scalar BEKK(1,1) structure in the conditional covariance matrix allows us to model time-varying conditional correlations without introducing more latent processes. In the hybrid MSF-SBEKK model two simple basic structures are nested. In the limiting case of $\sigma_g \rightarrow 0$ and $\phi = 0$ we obtain the VEC-SBEKK process, while setting $b=0$ and $a=0$ leads to the VEC-MSF case. As regards the initial conditions for $\tilde{\Sigma}_t$, we assume $\varepsilon_0 = 0$, $\tilde{\Sigma}_0 = s_{0,\Sigma} I_n$, where $s_{0,\Sigma} > 0$ and I_n denotes the identity matrix of size n .

Equation (1) can be decomposed and written as:

$$\Delta x_t = \alpha[\tilde{\beta}', \Phi_1'] \begin{bmatrix} x_{t-1} \\ D_t^{(1)} \end{bmatrix} + \sum_{i=1}^{k-1} \Gamma_i \Delta x_{t-i} + \Phi_2 D_t^{(2)} + \varepsilon_t = \alpha\beta' z_{1,t} + \Gamma' z_{2,t} + \Gamma_s' z_{3,t} + \varepsilon_t, \quad (6)$$

where $\beta' = [\tilde{\beta}', \Phi_1']$, $z_{1,t} = [x_{t-1}, D_t^{(1)}]'$, $z_{2,t} = (\Delta x_{t-1}', \Delta x_{t-2}', \dots, \Delta x_{t-k+1}')'$, $z_{3,t} = D_t^{(2)}$, $\Gamma = [\Gamma_1, \Gamma_2, \dots, \Gamma_{k-1}]'$, $\Gamma_s = \Phi_2'$, and $\alpha \Phi_1' D_t^{(1)} + \Phi_2 D_t^{(2)} = \Phi D_t$.

In order to simplify the notation let us write the basic model (6) in a matrix form:

$$Z_0 = Z_1 \Pi' + Z_2 \Gamma + Z_3 \Gamma_s + E, \quad (7)$$

where $\Pi = \alpha\beta'$, $Z_0 = [\Delta x_1, \Delta x_2, \dots, \Delta x_T]' = [z_{0,1}, z_{0,2}, \dots, z_{0,T}]'$, $Z_1 = [z_{1,1}, z_{1,2}, \dots, z_{1,T}]'$, $Z_2 = [z_{2,1}, z_{2,2}, \dots, z_{2,T}]'$, $Z_3 = [z_{3,1}, z_{3,2}, \dots, z_{3,T}]'$, $E = [\varepsilon_1, \varepsilon_2, \dots, \varepsilon_T]'$.

The conditional distribution of x_t (given the past of the process, ψ_{t-1} , the deterministic variables, D_t , the parameters and the latent variable q_t) is n -variate Normal with the mean $\mu_t = x_{t-1} + \tilde{\Pi}x_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta x_{t-i} + \Phi D_t$ and the covariance matrix Σ_t :

$$p(x_t | \psi_{t-1}, D_t, \alpha, \beta, \Gamma, \Gamma_s, q_t, \Sigma, \theta_\Sigma, q_{0,\Sigma}) = f_{N,n}(x_t | \mu_t, \Sigma_t), \quad (8)$$

where θ_Σ and $q_{0,\Sigma}$ are the vectors of the stochastic volatility parameters: in the VEC-MSF model $\theta_\Sigma = (\phi, \sigma_q^2)'$, and $q_{0,\Sigma} = \ln q_0$, whereas in the VEC-MSF-SBEKK model we have $\theta_\Sigma = (\phi, \sigma_q^2, a, b)'$ and $q_{0,\Sigma} = (\ln q_0, s_{0,\Sigma})'$. The vector $q_{0,\Sigma}$ is treated as an additional vector of parameters and is estimated jointly with other parameters. The density of the data (given the parameters) is the mixture (over $q = (q_1, q_2, \dots, q_T)'$) distribution:

$$p(x | D, \alpha, \beta, \Gamma, \Gamma_s, \Sigma, \theta_\Sigma, q_{0,\Sigma}) = \int p(x | D, \alpha, \beta, \Gamma, \Gamma_s, \Sigma, \theta_\Sigma, q_{0,\Sigma}, q) p(q | \theta_\Sigma, q_{0,\Sigma}) dq, \quad (9)$$

where $x = [x_1, x_2, \dots, x_T]'$ denotes the full data set and $D = [D_1, D_2, \dots, D_T]$. The two densities on the right hand side of (9) are given as:

$$p(x | D, \alpha, \beta, \Gamma_1, \Gamma_s, \Sigma, \theta_\Sigma, q_{0,\Sigma}, q) = \prod_{t=1}^T f_{N,n}(x_t | \mu_t, \Sigma_t),$$

and

$$p(q | \theta_\Sigma, q_{0,\Sigma}) = \prod_{t=1}^T q_t^{-1} f_{N,1}(\ln q_t | \phi \ln q_{t-1}, \sigma_q^2).$$

2.2 The prior distribution and Bayesian VEC-MSF model

We are interested in making inference about both the parameter vector $\theta = (\text{vec}\alpha', \text{vec}\beta', \text{vec}\Gamma', \text{vec}\Gamma_s', \text{vech}\Sigma', \theta_\Sigma', q_{0,\Sigma})'$ and the latent variables vector q . Within the Bayesian approach, parameters and latent variables are treated as random variables. The joint density of the vector of observations, latent variables and parameters, which determines the Bayesian model, can be written as follows:

$$\begin{aligned} p(x, q, \theta | D) &= p(x | D, q, \theta) p(q | \theta) p(\theta) = \\ &= p(\theta) \left[\prod_{t=1}^T f_{N,n}(x_t | \mu_t, \Sigma_t) \right] \left[\prod_{t=1}^T q_t^{-1} f_{N,1}(\ln q_t | \phi \ln q_{t-1}, \sigma_q^2) \right], \end{aligned} \quad (10)$$

where $p(\theta)$ denotes the density of prior distribution. This density is also conditioned by certain initial observations, which are omitted from our notation. To complete the Bayesian models, we specify the prior distribution of the parameter vector θ :

$$p(\theta) = p(\alpha | \beta) p(\beta) p(\Gamma) p(\Gamma_s) p(\Sigma) p(\theta_\Sigma) p(q_{0,\Sigma}) I_{[0,1]}(|\lambda|_{\max}), \quad (11)$$

where $I_{[a,b]}(\cdot)$ denotes the indicator function of the interval $[a, b]$ and λ is the vector of the eigenvalues of the companion matrix, that is the matrix of the form:

$$\begin{bmatrix} \tilde{\Pi} + I_n + \Gamma_1 & \Gamma_2 - \Gamma_1 & \cdots & \Gamma_{k-1} - \Gamma_{k-2} & -\Gamma_{k-1} \\ I_n & 0 & \cdots & 0 & 0 \\ 0 & I_n & & 0 & 0 \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I_n & 0 \end{bmatrix},$$

which makes it possible to write the analysed process in the VAR(1) form.

According to (11), we assume that certain blocks of the parameters are *a priori* independent, and that the stability condition is imposed on the parameters of the VEC process. Furthermore, for matrix Π the following parameterisations are used (as proposed by Koop, León-González, and Strachan 2010):

$$\alpha\beta' = (\alpha M_\Pi)(\alpha M_\Pi^{-1})' \equiv AB',$$

where M_Π is an $r \times r$ symmetric positive-definite matrix, A and B are unrestricted matrices. Moreover, $\alpha = A(B'B)^{-\frac{1}{2}}$, and $\beta = B(B'B)^{-\frac{1}{2}}$, so β has orthonormal columns, and it is an element of the Stiefel manifold $V_{r,m}$ (represented by the matrix space of $m \times r$ matrices with orthonormal columns). The data inform only about the cointegration space, which is the element of the Grassmann manifold $G_{r,m-r}$, i.e. the space of r -dimensional hyperplanes in $\mathbb{R}^m$. In order to take into account the many-to-one relationship between the Stiefel and the Grassmann manifolds, we normalise the columns of β to have positive first elements (with the help of a diagonal matrix whose elements are equal to either 1 or -1).

Next, we state the prior distributions (see Koop, León-González, and Strachan 2010 for the discussion).

- For matrix B the matrix normal distribution is used:

$p(B | \tau, r) = f_{mN}(B | 0, I_r, P_\tau)$, which leads us to the matrix angular central Gaussian (MACG) distribution for β : $p(\beta) = f_{MACG}(\beta | P_\tau)$ (see e.g. Chikuse 2002).

The prior information for the space spanned by β may be incorporated into the model via matrix P_τ , which is constructed as follows:

$$P_\tau = H_B H_B' + \tau H_B^\perp H_B^{\perp'},$$

where H_B is a matrix with orthonormal columns containing prior information about the cointegration space and $H_B^\perp$ represents its orthogonal complement. If we assume that parameter matrix P_τ in the MACG

distribution is an identity matrix, we obtain the uniform distribution over the Stiefel manifold $V_{r,m}$ and also the uniform distribution over the Grassmann manifold $G_{r,m-r}$. Additionally, it is worth recalling that unless the prior information about the cointegration space of the highest considered rank is available, the researcher willing to perform Bayesian model selection (including inference about the cointegration rank) should use the uniform prior over the Grassmann manifold.

- On matrix A we also impose the matrix normal distribution:

$$p(A | \mu_A, G, \nu, r) = f_{mN}(A | \mu_A, \nu I_r, G).$$

Parameters ν and τ control the degrees of informativeness of the distributions stated above, and may be either set arbitrarily by the researcher or estimated. In the latter case, usually inverse gamma prior distributions are used: $p(\nu | s_\nu, n_\nu) = f_{IG}(\nu | s_\nu, n_\nu)$ and $p(\tau | s_\tau, n_\tau) = f_{IG}(\tau | s_\tau, n_\tau)$ (see e.g. Koop, León-González, and Strachan 2010). Such values of s_τ and n_τ should be set that almost all the resulting prior probability is allocated close to zero. For τ close to zero one imposes most of the prior probability to the spaces close to those spanned by H_B , whereas for τ equal 1 one gets noninformative priors for the estimated spaces (see Koop, León-González, and Strachan 2010 for the discussion).

The priors for the remaining parameters are as follows:

- $p(\Gamma | \mu_\Gamma, H, h) = f_{mN}(\Gamma | \mu_\Gamma, H, hI_l)$, where $l = n(k-1)$, H – an $n \times n$ positive-definite symmetric matrix,
- $p(\Gamma_s | \mu_{\Gamma_s}, H_s, h_s) = f_{mN}(\Gamma_s | \mu_{\Gamma_s}, H_s, h_s I_{l_s})$, where H_s – an $n \times n$ positive-definite symmetric matrix, l_s – the number of deterministic terms in D_t ,
where $p(h | n_h, s_h) = f_{IG}(h | n_h, s_h)$ and $p(h_s | n_{hs}, s_{hs}) = f_{IG}(h_s | n_{hs}, s_{hs})$.
- $p(\phi) \propto f_{N,1}(\phi | \mu_\phi, \Omega_\phi) I_{(-1,1)}(\phi)$, – the normal density with mean μ_ϕ and variance Ω_ϕ , truncated by the restriction $|\phi| < 1$;
- $p(\sigma_a^2) = f_{IG}(\sigma_a^2 | a_\sigma, b_\sigma)$ – the density of the inverse gamma distribution with the mean $b_\sigma/(a_\sigma - 1)$ and variance equal to $b_\sigma^2 / [(a_\sigma - 1)^2(a_\sigma - 2)]$;
- $p(\Sigma) = f_{IW}(\Sigma | \Omega_\Sigma, \mu_\Sigma, n)$ – the inverted Wishart distribution with mean $\Omega_\Sigma/(\mu_\Sigma - n - 1)$, $\mu_\Sigma > n + 1$;
- $p(a, b) \propto I_{(0,1)}(a + b)$ – the uniform distribution over the unit simplex;
- $p(\ln q_0) = f_{N,1}(\ln q_0 | \mu_q, \Omega_q)$;
- $p(s_{0,\Sigma}) = f_{Exp}(s_{0,\Sigma} | \mu_{q,\Sigma})$ – the exponential distribution with mean $1/\mu_q$.

Finally, for θ_Σ we use the same priors as in (Osiewalski and Pajor 2009; Pajor and Osiewalski 2012).

3 MCMC methods for Bayesian estimation and model comparison

The joint posterior distribution, represented by $p(q, \theta | x)$, is too complicated to obtain a closed-form expression for any of its characteristics. To estimate these, we resort to MCMC methods, which allow one to generate a (pseudo-) random sample from the considered distribution.

3.1 Gibbs and Metropolis-Hastings algorithms

In order to draw from the joint posterior distribution of unknown parameters and latent variables we use a MCMC method, specifically, the Metropolis and Hastings algorithm within the Gibbs sampler. The procedure generates drawings from a joint distribution (as a stationary distribution) by sequentially sampling from the full conditional distributions (see, e.g. Gamerman 1997). Thus, we construct a Gibbs sampler with limiting distribution equal to the joint posterior distribution $p(q, \theta | x)$. The Gibbs sampler for the univariate stochastic volatility models, presented by (Jacquier, Polson, and Rossi 1994), can be easily generalised to the VEC-SV models.

In the VEC-MSF model one has to draw the $[T + nr + mr + l_s n + n^2(k-1) + n(n+1)/2]$ -dimensional vector $(\theta', q')'$, which is split up into $T + 12$ separate Gibbs steps, 11 of which are quite straightforward, while the remaining one poses more of a challenge.

Before presenting the conditional posterior distributions for the model parameters, let us introduce some additional notation:

$$W_i = \begin{bmatrix} w_{i,1} \\ w_{i,2} \\ \vdots \\ w_{i,T} \end{bmatrix} \text{ with } w_{i,t} = \Sigma_t^{-1/2} z'_{i,t} \text{ for } i = 0, 1, 2, 3,$$

where $\Sigma_t^{-1/2}$ denotes the matrix inverse of the Cholesky factor of Σ_t .

The equation for the observation t may be written as follows:

$$w_{0,t} = w_{1,t} \text{vec}(BA') + w_{2,t} \text{vec}(\Gamma) + w_{3,t} \text{vec}(\Gamma_s) + u_t, \quad \{u_t\} \sim iN(0, I_n).$$

We can now move on to presenting the conditional distributions of the VEC form parameters. They are either Normal (for the vectorisations of the matrix parameters) or inverse Gamma for the scalar ones:

$$- \quad p(\text{vec}(A') | \cdot, x) = f_N(\text{vec}(A') | \bar{\mu}_{A'}, \bar{\Omega}_{A'}), \quad (12)$$

where

$$\bar{\Omega}_{A'} = \left[\left(G^{-1} \otimes \frac{1}{\nu} I_r \right) + (I_n \otimes B') W_1' W_1 (I_n \otimes B) \right]^{-1},$$

$$\bar{\mu}_{A'} = \bar{\Omega}_{A'} \left[\left(G^{-1} \otimes \frac{1}{\nu} I_r \right) \text{vec}(\mu'_A) + (I_n \otimes B') W_1' [W_0 - W_2 \text{vec}(\Gamma) - W_3 \text{vec}(\Gamma_s)] \right],$$

$$- \quad p(\text{vec}(B) | \cdot, x) = f_N(\text{vec}(B) | \bar{\mu}_B, \bar{\Omega}_B), \quad (13)$$

where

$$\bar{\Omega}_B = [(mI_r \otimes P_{1/r}) + (A' \otimes I_m) W_1' W_1 (A \otimes I_m)]^{-1},$$

$$\bar{\mu}_B = \bar{\Omega}_B (A' \otimes I_m) W_1' [W_0 - W_2 \text{vec}(\Gamma) - W_3 \text{vec}(\Gamma_s)],$$

$$- \quad p(\text{vec}(\Gamma) | \cdot, x) = f_N(\text{vec}(\Gamma) | \bar{\mu}_\Gamma, \bar{\Omega}_\Gamma), \quad (14)$$

$$\text{where } \bar{\Omega}_\Gamma = \left[\left(H^{-1} \otimes \frac{1}{h} I_l \right) + W_2' W_2 \right]^{-1},$$

$$\bar{\mu}_\Gamma = \bar{\Omega}_\Gamma \left\{ \left(H^{-1} \otimes \frac{1}{h} I_l \right) \text{vec}(\mu_\Gamma) + W_2' [W_0 - W_1 \text{vec}(BA') - W_3 \text{vec}(\Gamma_s)] \right\},$$

$$- \quad p(\text{vec}(\Gamma_s) | \cdot, x) = f_N(\text{vec}(\Gamma_s) | \bar{\mu}_{\Gamma_s}, \bar{\Omega}_{\Gamma_s}), \quad (15)$$

$$\text{where } \bar{\Omega}_{\Gamma_s} = \left[\left(H_s^{-1} \otimes \frac{1}{h_s} I_{ls} \right) + W_3' W_3 \right]^{-1},$$

$$\bar{\mu}_{\Gamma_s} = \bar{\Omega}_{\Gamma_s} \left\{ \left(H_s^{-1} \otimes \frac{1}{h_s} I_{ls} \right) \text{vec}(\mu_{\Gamma_s}) + W_3' [W_0 - W_1 \text{vec}(BA') - W_2 \text{vec}(\Gamma)] \right\},$$

$$- \quad p(v | \cdot, x) = f_{IG} \left(v | n_v + \frac{nr}{2}, s_v + \frac{1}{2} \text{tr}[(A - \mu_A)' G^{-1} (A - \mu_A)] \right), \quad (16)$$

$$- \quad p(h | \cdot, x) = f_{IG} \left(h | n_h + \frac{nl}{2}, s_h + \frac{1}{2} \text{tr}[H^{-1}(\Gamma - \mu_\Gamma)'(\Gamma - \mu_\Gamma)] \right), \quad (17)$$

$$- \quad p(h_s | \cdot, x) = f_{IG} \left(h_s | n_{hs} + \frac{nl_s}{2}, s_{hs} + \frac{1}{2} \text{tr}[H_s^{-1}(\Gamma_s - \mu_{\Gamma_s})'(\Gamma_s - \mu_{\Gamma_s})] \right), \quad (18)$$

- additionally, in the case of estimated τ with an inverse gamma prior imposed $[\tau \sim IG(n_\tau, s_\tau)]$, the full conditional posterior distribution is of the form:

$$p(\tau | \cdot, x) \propto |P_\tau|^{-r/2} \tau^{-n_\tau-1} \exp \left\{ -\frac{1}{\tau} \left[s_\tau + \frac{1}{2} \text{tr}(mB'H_B^\perp H_B^{\perp'} B) \right] \right\}. \quad (19)$$

Distribution (19) does not belong to any standard distribution families. To obtain a sample from it, a Metropolis-Hastings algorithm can be employed with the proposed values drawn from the inverse gamma distribution: $IG\left(n_\tau + \frac{mr}{2}, s_\tau + \frac{1}{2} \text{tr}(mB'H_B^\perp H_B^{\perp'} B)\right)$. For a discussion about the meaning of τ and further recommendations for its estimation see e.g. (Koop, León-González, and Strachan 2010).

The conditional posteriors of the MSF parameters are defined by the following inverse gamma, truncated normal, inverted Wishart, and normal distributions:

$$- \quad p(\sigma_q^2 | \cdot, x) = f_{IG}(\sigma_q^2 | \mu_\sigma + T/2, s_\sigma), \quad (20)$$

$$\text{with } s_\sigma = b_\sigma + 0.5 \sum_{t=1}^T (\ln q_t - \phi \ln q_{t-1})^2;$$

$$- \quad p(\phi | \cdot, x) \propto f_{N,1}(\phi | \alpha_*, \sigma_q^2 A_*^{-1}) I_{(-1, 1)}(\phi), \quad (21)$$

where

$$\alpha_* = A_*^{-1}(W'Q + \sigma_q^2 \Omega_{\phi} \mu_\phi), \quad A_{*ij} = W'W + \sigma_q^2 \Omega_{\phi},$$

$$W = (\ln q_0, \dots, \ln q_{T-1})', \quad Q = (\ln q_1, \dots, \ln q_T)';$$

$$- \quad p(\Sigma | \cdot, x) \propto f_{IW}(\Sigma | \Omega_\Sigma + TS_0, \mu_\Sigma + T), \quad (22)$$

$$\text{where } S_0 = T^{-1} \sum_{t=1}^T q_t^{-1} (x_t - \mu_t)(x_t - \mu_t)';$$

$$- \quad p(\ln q_0 | \cdot, x) = f_{N,1}(\ln q_0 | b_2 b_1^{-1}, b_1^{-1}), \quad (23)$$

$$\text{where } b_1 = \phi^2 / \sigma_q^2 + 1 / \Omega_q, \quad b_2 = \phi \ln q_1 / \sigma_q^2 + \mu_q / \Omega_q.$$

Sampling from all these conditional distributions is straightforward (also in the case of the truncated ones, for which rejection sampling is applied).

The conditional posterior distributions of the unobserved variables in the VEC-MSF model are the following:

$$- \quad p(q_t | \cdot, x) \propto q_t^{-(n+2)/2} \exp[-(2q_t)^{-1}(x_t - \mu_t)' \Sigma^{-1}(x_t - \mu_t)] \times \\ \times \exp[-(2\sigma_{*q}^2)^{-1}(\ln q_t - s_{*t})^2], \quad (24)$$

$$\text{where for } 1 \leq t \leq T-1: \quad s_{*t} = [\phi(\ln q_{t-1} + \ln q_{t+1})](1 + \phi^2)^{-1}, \quad \sigma_{*q}^2 = \sigma_q^2(1 + \phi^2)^{-1}, \quad \text{and for } t=T: \quad s_{*T} = \phi \ln q_{T-1}, \\ \sigma_{*q}^2 = \sigma_q^2.$$

Density kernel (24) is not standard, so in order to sample from $p(q_t | \cdot, x)$ we use an acceptance-rejection Metropolis-Hastings algorithm with the candidate generating distribution obtained by approximating the log-normal kernel in (24) by the kernel of the inverse gamma distribution with the same mean and variance as this log-normal distribution (see Jacquier, Polson, and Rossi, 1994). This yields an inverse Gamma distribution

from which direct draws are easily generated (the product of two inverse Gamma distributions is still an inverse Gamma distribution). The candidate generating density is:

$$q(q_t | \cdot) \propto q_t^{-(\varphi+1)} \exp(-\theta_t^q q_t^{-1}), \quad (25)$$

$$\text{where } \varphi = \frac{(1-2e^{\sigma_q^2})}{(1-e^{\sigma_q^2})} + \frac{n}{2}, \quad \theta_t^q = \left(\varphi - \frac{n}{2} - 1\right) \exp\left(s_{*t} + \frac{1}{2}\sigma_q^2\right) + (x_t - \mu_t)' \Sigma^{-1} (x_t - \mu_t) / 2.$$

Let us now describe the Gibbs sampler used in the VEC-MSF-SBEKK model. The numerical procedure is much more computationally demanding and time-consuming. Because matrix $\tilde{\Sigma}_t$ in the SBEKK structure is a function of the past information set, the full conditional posterior distribution for the VEC structure is very complicated:

$$\begin{aligned} p(\text{vec}A', \text{vec}B', \text{vec}\Gamma', \text{vec}\Gamma_s', v, h, h_s, \tau | \cdot, x) \propto \\ \propto p(\text{vec}A', \text{vec}B', \text{vec}\Gamma', \text{vec}\Gamma_s', v, h, h_s, \tau) \prod_{t=1}^T f_{N,n}(x_t | \mu_t, q_t \tilde{\Sigma}_t). \end{aligned} \quad (26)$$

Drawing from the above full conditional posterior is riddled with serious numerical obstacles, unless its dimension (n) is small. In order to sample from (26) we implement the random walk Metropolis-Hastings algorithm with normal and truncated normal distributions centred at the previous values of the chain. The covariance matrix of the proposal distribution is a form of approximation of a specific part of the posterior covariance matrix, and it is determined by initial draws. For data sets used in our empirical illustrations the acceptance rate amounts to about 27%.

The conditional posterior distributions of ϕ , σ_q^2 and $\ln q_0$ in the MSF-SBEKK structure are also truncated normal, inverse gamma and normal, respectively, as in (19), (21) and (23). On the other hand, the conditional posterior of the parameters featured by the SBEKK part of the model is very complicated:

$$- \quad p(a, b, s_{0,\Sigma}, \Sigma | \cdot, x) \propto p(a, b, s_{0,\Sigma}, \Sigma) \prod_{t=1}^T f_{N,n}(x_t | \mu_t, \Sigma_t). \quad (27)$$

Parameters a , b , $s_{0,\Sigma}$, and elements of matrix Σ can be sampled using the Metropolis-Hastings steps within the Gibbs sampler. We implement the random walk Metropolis-Hastings algorithm with truncated Student's t distribution (with 3 degrees of freedom) centred at the previous values of the chain (similarly to Osiewalski and Pajor 2009). The covariance matrix of the latter is determined by initial draws of the algorithm. For large n sampling from this full conditional posterior distribution may be numerically more demanding.

The conditional posterior distributions of the unobserved variables in the VEC-MSF-SBEKK model are of the following form:

$$- \quad p(q_t | \cdot, x) \propto q_t^{-1} f_{N,1}(\ln q_t | \phi \ln q_{t-1}, \sigma_q^2) f_{N,1}(\ln q_{t+1} | \phi \ln q_t, \sigma_q^2) f_{N,n}(x_t | \mu_t, q_t \tilde{\Sigma}_t), \quad (28)$$

for $1 \leq t \leq T-1$,

$$- \quad p(q_T | \cdot, x) \propto q_T^{-1} f_{N,1}(\ln q_T | \phi \ln q_{T-1}, \sigma_q^2) f_{N,n}(x_T | \mu_T, q_T \tilde{\Sigma}_T). \quad (29)$$

In order to sample from (28) and (29) we also use the Metropolis-Hastings algorithm. Following (Osiewalski and Pajor 2009), the candidate generating density is an inverse Gamma density:

$$q(q_t | \cdot) \propto q_t^{-(\varphi+1)} \exp(-\theta_t^{q,B} q_t^{-1}), \quad (30)$$

$$\text{where } \varphi = \frac{(1-2e^{\sigma_q^2})}{(1-e^{\sigma_q^2})} + \frac{n}{2}, \quad \theta_t^{q,B} = \left(\varphi - \frac{n}{2} - 1\right) \exp\left(s_{*t} + \frac{1}{2}\sigma_q^2\right) + \frac{1}{2}(x_t - \mu_t)' \tilde{\Sigma}_t^{-1} (x_t - \mu_t).$$

We stress that the Metropolis-Hastings steps are implemented within the Gibbs sampler. Thus we create Metropolis subchains within the Markov chain constructed by the Gibbs sampler to sample from the non-standard distributions.

To summarise, Bayesian VEC-MSF or VEC-MSF-SBEKK models are analysed, using Gibbs sampling as a tool for simulating samples from the posterior distributions. In the case of the VEC-MSF-SBEKK structure the numerics is far more demanding even for small n . On the other hand, the forms of the full conditional posterior distributions in the VEC-MSF model make steps of the Gibbs sampler easy even for large n . However, the MSF specification is very restrictive since it assumes the same dynamics for all entries of Σ_t . This assumption is a price to be paid for the ease of numerical calculations.

3.2 Harmonic mean estimator with Lenk's correction

The marginal data densities are essential in the formal Bayesian model selection and model averaging. (Newton and Raftery 1994) proposed a simple (and hence popular) method of estimation of the marginal data density based on the harmonic mean. The harmonic mean estimator (HME) is given as:

$$\hat{p}_{HME}(x) = \left[\frac{1}{m} \sum_{p=1}^m \frac{1}{p(x|\theta_{(p)}, q_{(p)})} \right]^{-1} \quad (31)$$

where $\{\theta_{(p)}, q_{(p)}\}_{p=1}^m$ are the draws generated from the posterior distribution of the parameters and latent variables, using the MCMC methods.

Even though the HME is consistent (see Newton and Raftery 1994), the computed HME suffers from some serious shortcomings, for example it overestimates the marginal data density. In the paper (Lenk 2009) the source of this “simulation pseudo-bias” of the HME is identified, and several methods of estimating the “bias” adjustment factor are proposed. The adjusted HME for the marginal data density, proposed by (Lenk 2009), is given by the formula:

$$\hat{p}_{AHME}(x) = \hat{P}(C) \left[\frac{1}{m} \sum_{p=1}^m \frac{1}{p(x|\theta_{(p)}, q_{(p)})} \right]^{-1}, \quad (32)$$

where $\hat{P}(C)$ is an evaluation of the prior probability of subset $C \subseteq \Theta$ [i.e. $P(C)$], where Θ denotes the space of parameters θ and latent variables q . In Equation (32) $\{\theta_{(p)}, q_{(p)}\}_{p=1}^m$ are drawn from the posterior distribution of the vector $(\theta', q')'$ restricted to the subset C . Following (Lenk 2009) we assume that $C = \{(\theta', q')': p(x|\theta, q) \geq L\}$, where $L = \min_{(\theta', q')' \in \{(\theta_{(p)}, q_{(p)})\}_{p=1}^m} p(x|\theta, q)$. $P(C)$ is approximated using importance sampling, drawing from

the truncated [to the support of $p(\theta)$] multivariate normal distribution for all parameters and the prior distribution for the latent variables. Starting from the identity: $P(C) = \int_{\Theta} p(q|\theta) p(\theta) I_C(\theta, q) d\theta dq$, we obtain:

$$P(C) = \int_{\Theta} \frac{p(\theta)}{s(\theta)} I_C(\theta, q) s(\theta) p(q|\theta) d\theta dq, \text{ where } p(\theta) \text{ is the probability density function of the prior distribution}$$

of parameters, $p(q|\theta)$ is the density function of the prior distribution for the latent variables (given the parameters), $s(\theta)$ is the probability density function of the truncated multivariate normal distribution. The last identity yields the importance sampling estimator of $P(C)$ with the importance function of the form: $s(\theta) p(q|\theta)$. The parameters of the truncated normal distribution (the mean and covariances) are obtained basing on the MCMC draws from the posterior distribution.

The estimate of $P(C)$ is:

$$\hat{P}(C) = \frac{1}{J} \sum_{j=1}^J \frac{p(\theta_{(j)}^s) I_C(\theta_{(j)}^s, q_{(j)}^s)}{s(\theta_{(j)}^s)}, \quad (33)$$

where $\{\theta_{(j)}^s, q_{(j)}^s\}_{j=1}^J$ are drawn from the importance sampling distribution defined by the probability density function $p(q|\theta) s(\theta)$. The normalising constants of probability density functions $s(\theta)$ and $p(\theta)$ are not known, thus they are calculated, using the direct sampling method.

4 Empirical illustrations

Data used in the paper come from the Polish economy. In the first example we analyse financial data of daily quotations on three major exchange rates: PLN/EUR, PLN/USD and EUR/USD. The second one focuses on macroeconomic data of unemployment, inflation and interest rates.

4.1 Long-run relationships among exchange rates

We start by modelling the exchange rates (where, on the ground of lack of arbitrage opportunities, the cointegrating vector may be assumed to be known *a priori*) to show that our model and numerical methods are suitable for modelling a long-run relationship in financial data. Let us consider two average daily main Polish official exchange rates: the zloty (PLN) values of the US dollar, and the zloty (PLN) values of the euro, over the period from January 3, 2005 to September 30, 2011 (downloaded from the website of the National Bank of Poland). The dataset of the daily logarithmic growth rates (expressed in percentage terms) consists of 1662 observations (for each series). The first observation is sacrificed as an initial condition, thus $T = 1661$. The Polish official exchange rates are linked to the exchange rates quoted on Forex, the international currencies market. Thus, while building time-series models for the two Polish exchange rates [EUR/PLN ($x_{2,t}$) and USD/PLN ($x_{3,t}$)], we introduce some extra variable from the international market – the euro value of the US dollar, EUR/USD ($x_{1,t}$, downloaded from <http://stoq.com>).

We use the same dataset as in (Pajor 2011), where exogeneity in models with latent variables was considered, and the cointegrating vector was assumed to be known. The relationship: $(\text{EUR/PLN})/(\text{USD/PLN}) \approx \text{EUR/USD}$ was introduced by assuming that this relation (in logarithmic terms) is a cointegration equation in the sense of (Engle and Granger 1987). This yielded the cointegrating vector $(1, -1, 1)$ and the long-run (equilibrium) relationship: $\ln x_{2,t} - \ln x_{3,t} = \ln x_{1,t}$. The assumption that $\ln x_{1,t}$, $\ln x_{2,t}$ and $\ln x_{3,t}$ are cointegrated has been checked informally using the Phillips and Perron test applied to the series $ECM_t = 100(\ln x_{1,t} - \ln x_{2,t} + \ln x_{3,t})$. In this paper the cointegrating vector is formally estimated.

The Polish official exchange rates with the addition of certain variables related to the foreign exchange market were previously modelled by (Osiewalski and Pipień 2004) within the VEC-GARCH structure, and under assumption that the cointegrating vector is known.

The data are plotted in Figure 1, depicting the graphs of logarithms (multiplied by 100) of exchange rates in terms of their levels and of their increments. It can be seen from the graphs that the growth rates are centred around zero, and feature time-varying volatility and outliers. The series are more volatile in the period of the financial crisis 2008–2009. Moreover, the official exchange rates of the National Bank of Poland appear more volatile than the EUR/USD exchange rates.

A visual inspection of the data reveals that they do not satisfy the constant conditional variance assumption, and therefore the conditional covariance matrix of the respective VAR models need to be stochastic. Summary statistics for the analysed series are shown in Table 1. The arithmetic means of the growth rates of the official Polish exchange rates are positive, but the standard deviations are relatively high. A much higher than 3 value of the kurtosis pronounced leptokurtosis of the empirical distribution. As expected, the EUR/PLN and USD/PLN exchange rates are positively correlated.

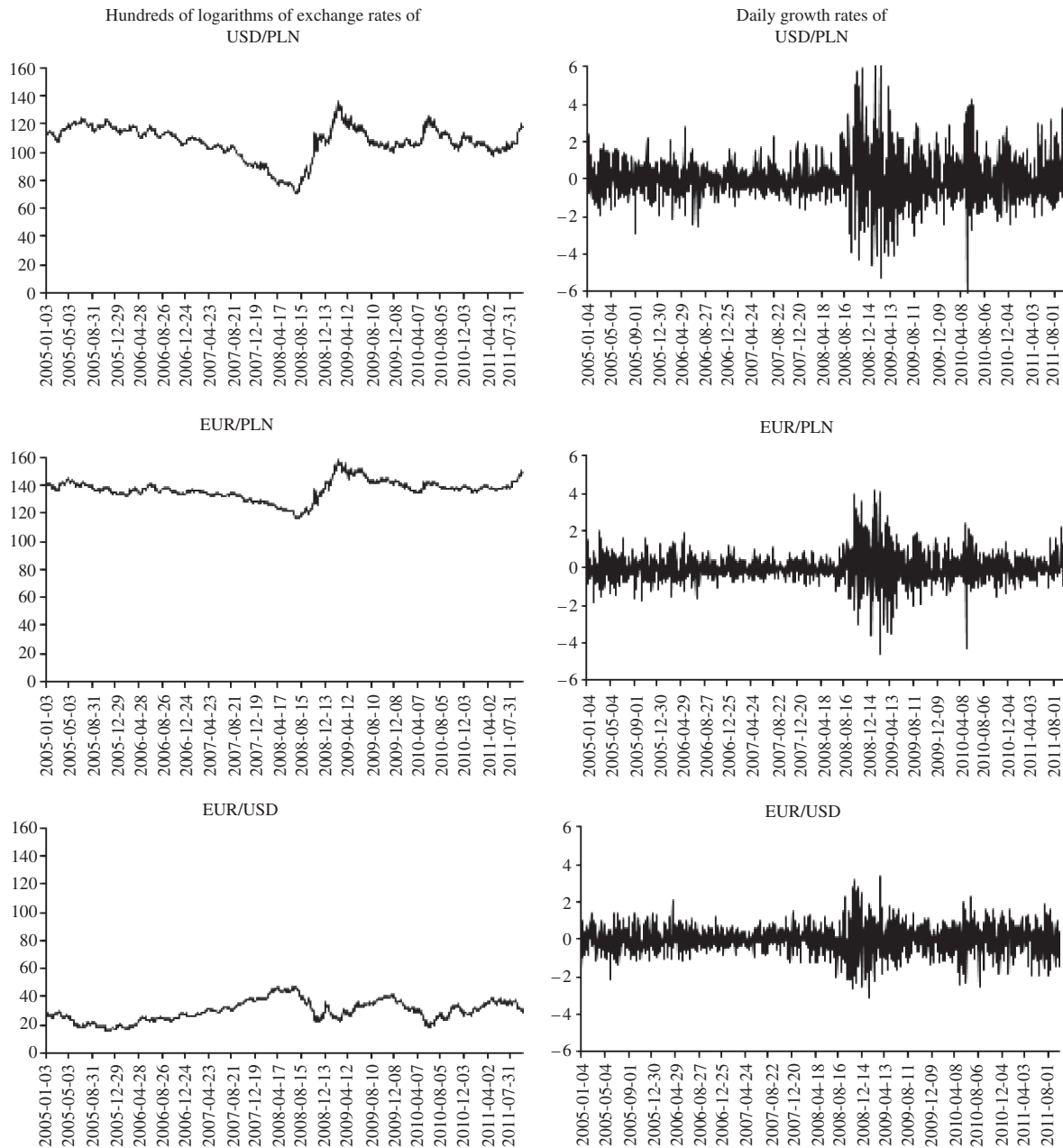


Figure 1: Exchange rates of USD/PLN, EUR/PLN and EUR/USD (January 3, 2005 – September 30, 2011).

Due to the typical properties of daily growth rates of exchange rates, we consider the VEC-MSF and the VEC-MSF-SBEKK models with one cointegrating vector ($r=1$), containing an unrestricted vector of constants, and the number of lags equal to 2 ($k=2$).¹ We set the following values of hyperparameters:

¹ The Lindley-type test indicates that such an autoregressive lag order is high enough to capture the autocorrelation of the series. We set the null and alternative hypotheses as: $H_0: \Gamma_2 = 0$ and $H_1: \Gamma_2 \neq 0$, where $\text{vec}(\Gamma_2) = \theta_0$ is a 9×1 vector. Then the quadratic form: $F(\theta_0) = [\theta_0 - E(\theta_0 | x)]' V^{-1}(\theta_0 | x) [\theta_0 - E(\theta_0 | x)]$, where $E(\theta_0 | x)$ and $V(\theta_0 | x)$ are, respectively, the vector of posterior means of θ_0 and the posterior covariance matrix, is a *posteriori* approximately ($T \rightarrow \infty$) χ^2 distributed with 9 degrees of freedom. In the VEC-MSF-SBEKK model we obtain $F(0) = 4.59$, and $p(F(\theta_0) > F(0) | x) \approx 0.86$. Thus, the assumption that $\Gamma_2 = 0$ is not rejected by the data.

Table 1: Descriptive statistics of analyzed exchange rates ($x_{i,t}$) and of daily growth rates ($y_{i,t}$).

	$y_{i,t}$			$x_{i,t}$		
	USD/PLN	EUR/PLN	EUR/USD	USD/PLN	EUR/PLN	EUR/USD
Mean	0.0047	0.0050	-0.0003	2.9319	3.9343	1.3513
SD	0.0280	0.0182	0.0172	0.0081	0.0067	0.0025
Skewness	0.4802	0.2530	-0.0113	-0.6198	-0.0419	0.3538
Kurtosis	7.6101	8.8404	4.9594	3.3365	3.8930	2.4680
Minimum	-6.7485	-4.5895	-3.1163	2.0220	3.2026	1.1670
Maximum	6.2677	4.1467	3.4157	3.8978	4.8999	1.5995
Empirical correlations						
USD/PLN	1	0.7502	-0.3082	1	0.8041	-0.8452
EUR/PLN		1	-0.1313		1	-0.3742
EUR/USD			1			1

- for matrix B : $P_\tau = I_3$ (the uniform distribution over the Stiefel manifold), $\tau = 1$;
- for matrix A : $G = I_3$, $\mu_A = 0_{3 \times 1}$, $s_v = 2$, $n_v = 3$;
- for matrix Γ : $\mu_\Gamma = 0_{3 \times 3}$, $H = I_3$, $s_h = 2$, $n_h = 3$;
- for matrix Γ_s : $\mu_{\Gamma_s} = 0_{3 \times 1}$, $H_s = I_3$, $s_{hs} = 2$, $n_{hs} = 3$;
- for MSF parameters: $\mu_\phi = 0$, $\Omega_\phi = 100$, $a_\sigma = 1$, $b_\sigma = 0.005$, $\mu_q = 0$, $\Omega_q = 100$, $\mu_\Sigma = 5$, $\Omega_\Sigma = I_3$,
- for SBEKK parameters: $\mu_{q,\Sigma} = 1$.

As mentioned above, the relationship: (EUR/PLN)/(USD/PLN) $\approx$ EUR/USD indicates that the posterior cointegrating space can be spanned by the vector (1, -1, 1). One of the objectives of the study is to check whether formal Bayesian estimation of the cointegrating space leads to the same cointegrating vector, which was a priori assumed in our previous research.

The posterior means and standard deviations (in parentheses) of the cointegrating vector as well as the adjustment coefficients are presented bellow. As suggested by (Villani 2006), the point estimates of the cointegrating vectors are obtained with the use of a loss function based on the projective Frobenius distance $\hat{\beta} = \arg \min_{\tilde{\beta} \in G_{r,m-r}} E[l(\beta, \tilde{\beta})]$, where $l(\beta, \tilde{\beta}) = 2(r - \text{tr}(\beta\beta'\tilde{\beta}\tilde{\beta}'))$. This function is minimised at $\hat{\beta} = (v_1, v_2, \dots, v_r)$,

where v_i ($i = 1, 2, \dots, r$) are the eigenvectors corresponding to the r largest eigenvalues of the matrix $E(\beta\beta')$. The

measures of cointegration space variation are calculated as $\tau_{sp(\beta)} = \sqrt{\frac{r - \sum_{i=1}^r \lambda_i}{r(m-r)/m}}$, where λ_i is the i -th largest

eigenvalue of $E(\beta\beta')$. This measure takes values from 0 to 1. The lower is the value, the tighter is the posterior distribution (see Villani 2006). To make comparison with the theoretical cointegrating vector more readable, the point estimates of the cointegrating and the adjustment coefficient vectors are normalised in such a manner that the first element of β equals 1. Note that in the case of one cointegrating vector such a procedure does not alter the cointegration space. We obtained the following posterior means (and standard deviations) of β and α :

$$E(\beta | x, MSF) = [1.000 \quad -0.997 \quad 0.999], \tau_{sp(\beta)} = 0.0007,$$

$$E(\alpha | x, MSF) = \begin{bmatrix} 0.080 & 0.017 & -0.961 \\ (0.097) & (0.094) & (0.116) \end{bmatrix},$$

$$E(\beta | x, MSF - SBEKK) = [1.000 \quad -0.998 \quad 1.000], \tau_{sp(\beta)} = 0.0006,$$

$$E(\alpha | x, MSF - SBEKK) = \begin{bmatrix} 0.054 & 0.314 & -0.624 \\ (0.081) & (0.070) & (0.090) \end{bmatrix}.$$

Overall, the results confirm that exchange rates of USD/PLN, EUR/PLN and EUR/USD are cointegrated with coefficients (1, -1, 1).

In the VEC-MSF model the posterior means and standard deviations (in parentheses) of the VAR parameters are:

$$\begin{bmatrix} \Delta x_{t,1} \\ \Delta x_{t,2} \\ \Delta x_{t,3} \end{bmatrix} = \begin{bmatrix} 0.088 & 0.020 & -1.602 \\ (0.119) & (0.107) & (0.624) \end{bmatrix} \begin{bmatrix} 1.000 \\ -0.997 \text{ (0.0007)} \\ 0.999 \end{bmatrix} \begin{bmatrix} x_{t-1,1} \\ x_{t-1,2} \\ x_{t-1,3} \end{bmatrix} + \begin{bmatrix} -0.076 & -0.040 & -0.013 \\ (0.050) & (0.041) & (0.030) \\ -0.115 & -0.072 & 0.020 \\ (0.047) & (0.039) & (0.028) \\ -0.135 & -0.042 & 0.006 \\ (0.060) & (0.048) & (0.035) \end{bmatrix} \begin{bmatrix} \Delta x_{t-1,1} \\ \Delta x_{t-1,2} \\ \Delta x_{t-1,3} \end{bmatrix} + \begin{bmatrix} 0.003 \\ (0.028) \\ -0.032 \\ (0.022) \\ 0.206 \\ (0.201) \end{bmatrix} + \hat{\varepsilon}_t,$$

and

$$\ln q_t = 0.987 \ln q_{t-1} + \sqrt{0.0207} \eta_t + \hat{\eta}_t, \\ (0.0063) \quad (0.0063)$$

$$E(\ln q_0 | x, MSF) = 0.055, \quad D(\ln q_0 | x, MSF) = 0.488,$$

$$E(\Sigma | x, MSF) = \begin{bmatrix} 0.552 & -0.041 & -0.230 \\ (0.223) & (0.022) & (0.095) \\ -0.041 & 0.503 & 0.534 \\ (0.022) & (0.204) & (0.217) \\ -0.230 & 0.534 & 0.780 \\ (0.095) & (0.217) & (0.315) \end{bmatrix},$$

where $(\hat{\varepsilon}_t', \hat{\eta}_t)'$ is the vector of relevant residuals.

In turn, in the VEC-MSF-SBEKK model the posterior means and standard deviations of the VAR parameters are:

$$\begin{bmatrix} \Delta x_{t,1} \\ \Delta x_{t,2} \\ \Delta x_{t,3} \end{bmatrix} = \begin{bmatrix} 0.054 & 0.314 & -0.624 \\ (0.081) & (0.070) & (0.090) \end{bmatrix} \begin{bmatrix} 1.000 \\ -0.998 \text{ (0.0006)} \\ 1.000 \end{bmatrix} \begin{bmatrix} x_{t-1,1} \\ x_{t-1,2} \\ x_{t-1,3} \end{bmatrix} + \begin{bmatrix} -0.080 & -0.022 & -0.027 \\ (0.042) & (0.040) & (0.030) \\ -0.295 & -0.049 & 0.014 \\ (0.033) & (0.028) & (0.017) \\ -0.312 & -0.021 & 0.006 \\ (0.044) & (0.043) & (0.030) \end{bmatrix} \begin{bmatrix} \Delta x_{t-1,1} \\ \Delta x_{t-1,2} \\ \Delta x_{t-1,3} \end{bmatrix} + \begin{bmatrix} 0.014 \\ (0.022) \\ -0.100 \\ (0.064) \\ 0.109 \\ (0.121) \end{bmatrix} + \hat{\varepsilon}_t,$$

and

$$\ln q_t = 0.664 \ln q_{t-1} + \sqrt{0.162} \eta_t + \hat{\eta}_t, \\ (0.086) \quad (0.043)$$

$$E(\ln q_0 | x, MSF - SBEKK) = 5.760, \quad D(\ln q_0 | MSF - SBEKK) = 1.753,$$

$$E(a | x, MSF - SBEKK) = 0.029, \quad D(a | MSF - SBEKK) = 0.003,$$

$$E(b | x, MSF - SBEKK) = 0.969, \quad D(b | MSF - SBEKK) = 0.004,$$

$$E(q_{0,\Sigma} | x, MSF - SBEKK) = 0.044, \quad D(q_{0,\Sigma} | x, MSF - SBEKK) = 0.026,$$

$$E(\Sigma | x, MSF - SBEKK) = \begin{matrix} \begin{matrix} 0.622 & -0.018 & -0.186 \\ (0.460) & (0.134) & (0.271) \end{matrix} \\ \begin{matrix} -0.018 & 0.219 & 0.138 \\ (0.134) & (0.099) & (0.124) \end{matrix} \\ \begin{matrix} -0.186 & 0.138 & 0.414 \\ (0.271) & (0.124) & (0.225) \end{matrix} \end{matrix}.$$

Similar characteristics of posterior distributions were obtained in the VEC-SV models with the known cointegration relation (therefore they are not presented here), with the marginal posterior distributions being only less dispersed. The inference about the individual volatility of each time series (measured by the conditional standard deviations) is also very similar in both considered VEC-SV models: with either the known or estimated cointegrating vector. The time plots of conditional standard deviations (for each $t=1, 2, \dots, T$; $T=1661$) are presented in Figure 2, where the upper line represents the posterior mean plus two standard deviations, and the lower one – the posterior mean minus two standard deviations. The models produce volatility peaks at the same time moments. Note that the dynamic pattern of the volatility for all returns is the same in the VEC-MSF model, which is due to the fact that the dynamics of volatilities is governed by only one latent process. However, in spite of the simplicity of the structure of the conditional covariance matrix, we allow the variances of multivariate returns to vary over time. The plots of the main posterior characteristics of conditional correlations are presented in Figure 3. One can easily notice that the constant conditional correlation hypothesis (represented by the MSF specification) is not supported by the data, which is consistent with the findings presented by (Osiewalski and Osiewalski 2013). Due to numerical problems in calculating the marginal likelihoods, we do not present formal Bayesian model comparison of the MSF and MSF-SBEKK specifications. The use of the Lindley-type test (based on the highest posterior density region, see, e.g. Box and Tiao 1973; Bauwens, Lubrano, and Richard 1999) would lead to rejection of the hypothesis. However, the assumption of constant conditional correlations does not influence the inference about the cointegration relationship. The distance measures between the estimated cointegrated spaces and the theoretical one [i.e. spanned by the vector $(1, -1, 1)^2$] obtained in the VEC-MSF and the VEC-MSF-SBEKK models are very close to zero and almost the same (although they are slightly lower in the VEC-MSF models). In the set of the VEC-MSF-SBEKK models the distance measure takes the highest value in the model with five lags ($5.5 \cdot 10^{-6}$) and they fall to $3.3 \cdot 10^{-6}$ in the model with two lags.

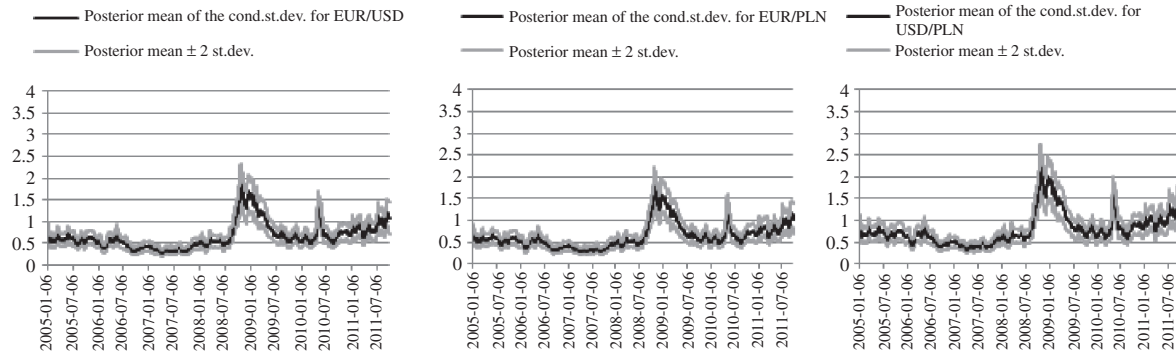
Our application of the VEC-SV models to exchange rates yields results, which are consistent with the theory. The cointegrating space appears indeed to be spanned by the vector $(1, -1, 1)$. This illustration shows that our methods can accurately estimate the long-run relationship in the VEC model with a time-variable conditional covariance matrix.

4.2 A small model for monetary policy

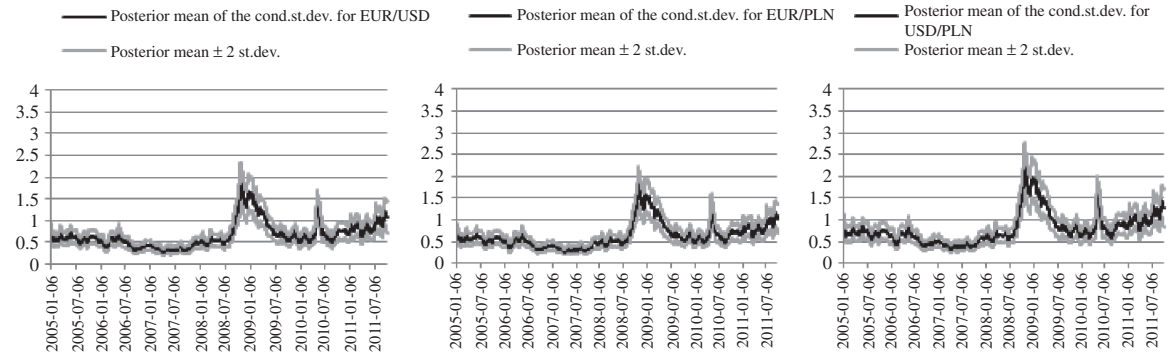
In the second example the methods developed in the paper are applied to a system with three Polish macroeconomic variables. Before presenting the results, it is worth emphasizing that the papers cited in the introduction pertain predominantly to the US economy. To the best of our knowledge, similar analyses for the

² Following (Larsson and Villani 2001) we based this measure on the Frobenius matrix norm. We thank the Referee for this suggestion.

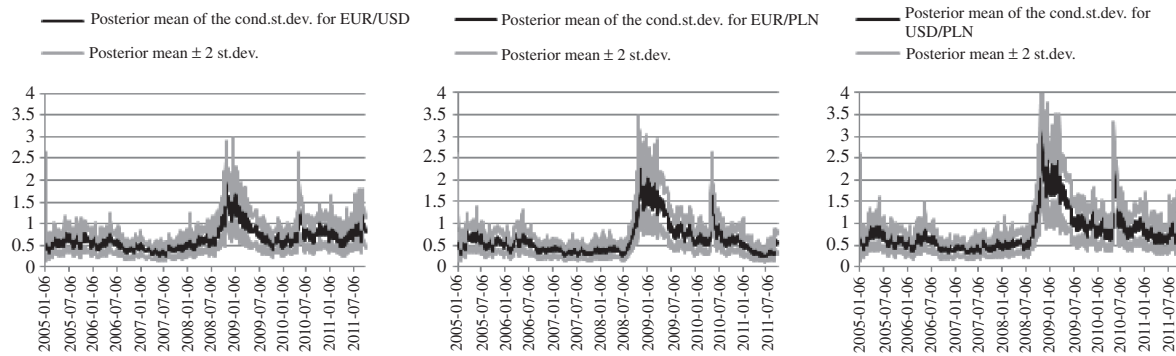
VEC-MSF model, $r = 1$, cointegrating vector known



VEC-MSF model, $r = 1$, cointegrating vector estimated



VEC-MSF-SBEKK model, $r = 1$, cointegrating vector known



VEC-MSF-SBEKK model, $r = 1$, cointegrating vector estimated

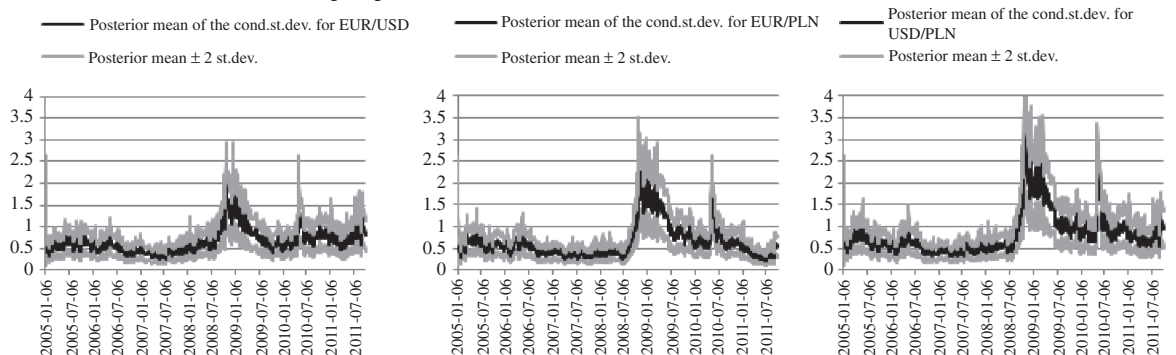
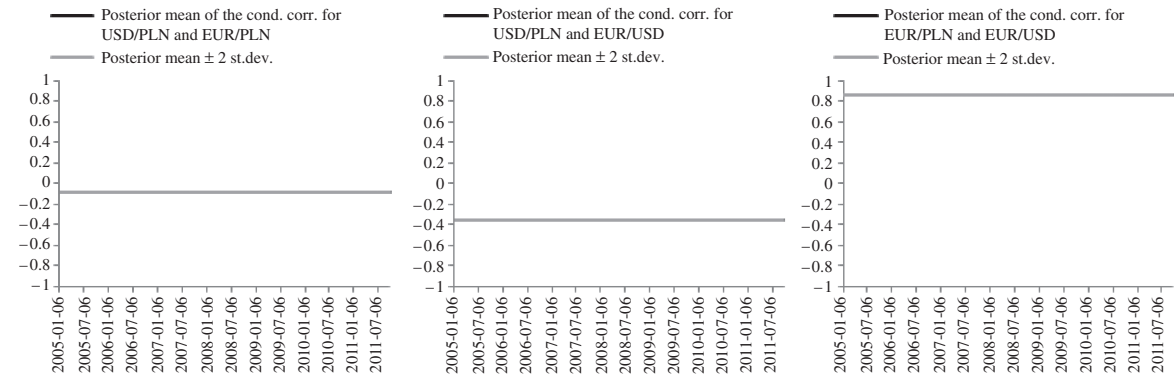


Figure 2: Conditional standard deviations (posterior mean ± 2 standard deviations). The results are obtained in the VEC-MSF and VEC-MSF-SBEKK models with $r = 1$.

VEC-MSF model, $r = 1$, cointegrating vector estimated



VEC-MSF-SBEKK model, $r = 1$, cointegrating vector estimated

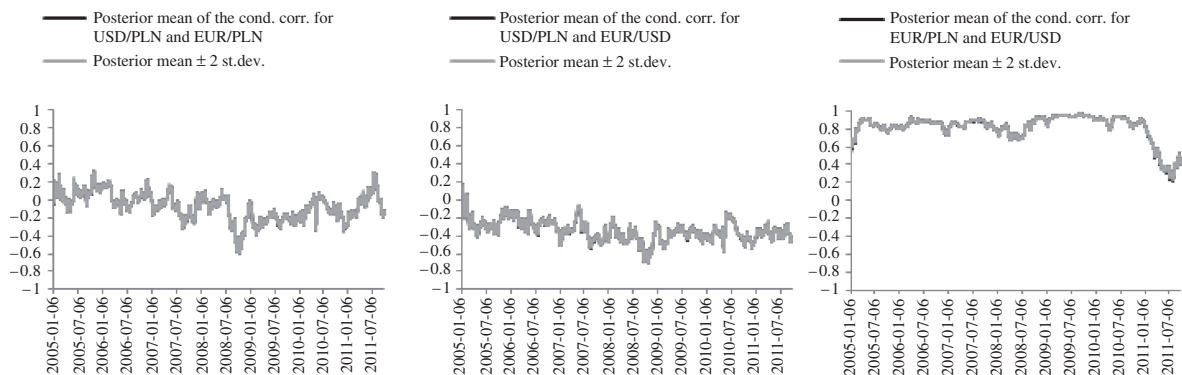


Figure 3: Conditional correlations (posterior mean $\pm$ 2 standard deviations). The results are obtained in the VEC-MSF and VEC-MSF-SBEKK models with $r=1$ (cointegrating vector is estimated).

Polish economy, with the use of the (Bayesian) VEC-SV models, have not been carried out as yet. Therefore, the paper aims to fill this gap.

In this study a so called small model of monetary policy is considered. The model is based on three major macroeconomic variables: inflation of consumer prices, Δp_t , unemployment rate, U_t , and short-term interest rates, r_t . The latter can be regarded as a proxy for the monetary policy.

The seasonally unadjusted data cover the period from 1995Q1 to 2012Q4. The seasonality of the analysed series is modelled in a deterministic manner, *via* zero-mean seasonal dummies. The first 3 years are used as a training sample to determine the hyperparameters of the prior distribution for the cointegration space (the Johansen procedure in the model with two lags and r cointegrating vectors β has been used, which further have been utilized to calculate $P_\tau = \beta\beta' + 0.5\beta_\perp\beta_\perp'$). The next five quarters are sacrificed as initial conditions. Eventually, there are 55 modelled observations (see Figure 4).

As the analysed time series are rather short, it is very important to test whether the variances and covariances are constant or not. A visual inspection of the time paths of the series suggests that one could expect a time-varying covariance structure. The empirical validity of this presumption is to be formally tested *via* Bayesian model comparison. The set of competing models include 72 non-nested specifications. Along with the number of latent processes driving covariances ($l=0$ in the a VEC model, $l=1$ in the VEC-MSF and VEC-MSF-SBEKK forms), the models can differ in the VAR order ($k \in \{2, 3, 4, 5\}$), the type of incorporated deterministic term ($d=4$ denotes a restricted constant, $d=3$ – an unrestricted one) and the cointegration rank (r can be equal to 1 or 2). Additionally, we consider VAR models for the first differences of the analysed processes ($r=0$) and VAR models for the levels of these processes ($r=3$). We assume that the prior probability of each of the specifications compared is equal to 0.0139, which is the reciprocal of the total number of models considered.

Table 2 summarizes the results of the model comparison. In all three considered cases of the covariance structure (homoscedastic, MSF and MSF-SBEKK), the model that ranks first is the one that features two lags,

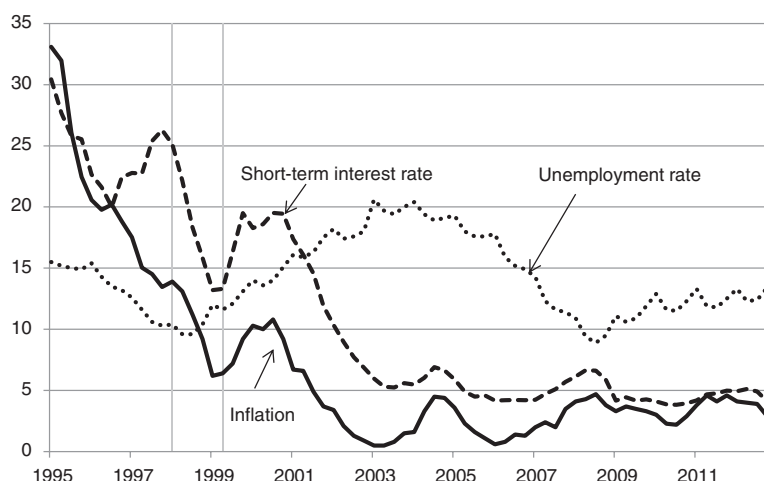


Figure 4: The analysed macro-data series (1995Q1 – 2012Q4).

Table 2: Natural logarithm of estimates of marginal likelihoods, numerical standard errors and ranks.

k, d, r	VEC			VEC-MSF			VEC-MSF-SBEKK		
	$\ln p(x)$	NSE	Rank (under all models)	$\ln p(x)$	NSE	Rank (under all models)	$\ln p(x)$	NSE	Rank (under all models)
2,3,0	-196.94	0.46	24 (23)	-182.06	0.56	24 (5)	-177.58	0.32	24 (3)
2,3,1	-198.69	0.50	4 (26)	-185.61	0.50	3 (6)	-181.39	0.50	3 (4)
2,3,2	-203.16	0.98	7 (37)	-187.89	0.75	5 (8)	-202.69	1.17	12 (36)
2,3,3	-214.66	1.03	19 (54)	-204.34	0.89	20 (40)	-225.80	1.30	17 (63)
2,4,1	-191.89	0.49	1 (15)	-177.38	0.42	1 (2)	-174.29	0.87	1 (1)
2,4,2	-196.88	0.60	2 (22)	-190.41	1.04	7 (13)	-202.45	1.03	11 (34)
3,3,0	-204.55	1.03	3 (41)	-190.55	0.87	2 (14)	-187.38	0.60	2 (7)
3,3,1	-206.53	0.36	12 (45)	-196.56	0.85	12 (21)	-201.64	1.02	10 (32)
3,3,2	-211.16	0.84	16 (49)	-200.30	1.06	15 (30)	-223.44	1.10	15 (61)
3,3,3	-220.83	1.09	23 (60)	-212.93	0.88	22 (50)	-238.38	1.16	21 (67)
3,4,1	-199.04	0.52	5 (27)	-187.59	0.51	4 (8)	-189.79	0.64	6 (12)
3,4,2	-206.57	0.74	13 (46)	-193.21	0.70	10 (18)	-231.97	0.68	19 (65)
4,3,0	-210.72	0.92	9 (48)	-197.66	1.03	8 (24)	-192.89	0.64	4 (17)
4,3,1	-213.70	0.67	18 (52)	-201.14	1.05	16 (31)	-216.74	0.79	14 (57)
4,3,2	-216.67	0.66	21 (56)	-206.40	1.23	21 (44)	-247.65	1.37	22 (68)
4,3,3	-219.61	1.40	22 (59)	-219.12	0.96	23 (58)	-237.05	1.00	20 (66)
4,4,1	-203.21	0.81	8 (38)	-192.00	0.80	9 (16)	-193.77	0.55	8 (20)
4,4,2	-215.67	0.89	20 (55)	-199.62	0.73	14 (28)	-225.09	1.05	16 (62)
5,3,0	-205.24	1.15	15 (42)	-193.25	1.04	13 (19)	-188.09	1.05	7 (9)
5,3,1	-209.38	0.81	14 (47)	-202.39	1.03	17 (33)	-213.94	0.88	13 (53)
5,3,2	-212.96	0.75	17 (51)	-204.08	0.75	19 (39)	-253.58	1.18	23 (69)
5,4,1	-199.91	1.03	6 (29)	-188.61	0.60	6 (11)	-197.88	0.79	9 (25)
5,4,2	-205.82	1.25	11 (43)	-202.46	1.11	18 (35)	-231.00	1.11	18 (64)

k , VAR order; d , deterministic term; r , cointegration rank; $r=3$ – VAR for levels, $r=0$ – VAR for first differences. The P_r matrix is determined with the help of the Johansen procedure employed in the model with r cointegrating relations.

one cointegrating relation and a constant restricted to the cointegration space. Generally, among the models with the highest posterior probability are almost exclusively those with either no or only one cointegrating relation, a time-varying covariance matrix and a short lag structure. Note that the number of lags seems to be less important than the cointegration rank. Due to high values of numerical standard errors of HME the ranking positions of some models cannot be clearly identified, but even if we take into account the error for

Table 3: The most probable models.

VAR order (k)	Deterministic term (d)	Cointegration rank (r)	Model (I)	$P(M_{(k,d,r,0)} x)$
2	4	1	VEC-MSF-SBEKK	0.92261
2	4	1	VEC-MSF	0.04172
2	3	0	VEC-MSF-SBEKK	0.03435
2	3	1	VEC-MSF-SBEKK	0.00076
2	3	0	VEC-MSF	0.00039
2	3	1	VEC-MSF	0.00001
3	3	0	VEC-MSF-SBEKK	0.000002
3	4	1	VEC-MSF	0.000002
2	3	2	VEC-MSF	0.000001

The P_{τ} matrix is determined with the help of the Johansen procedure employed in the model with r cointegrating relations.

the VEC-MSF-SBEKK model with $k=2$, $d=4$ and $r=1$, there is no doubt that the marginal data density for this specification is the highest. We can also observe high positive correlations between ordering of the models in all considered cases of the covariance structure. Table 3 presents models with the highest posterior probability. Among them there are only the VEC-MSF-SBEKK and VEC-MSF specifications, so there is a strong evidence that the covariance matrix is time-varying. Only the first three models obtained higher posterior probability than the assumed prior probability, and almost all posterior probability is gathered by the VEC-MSF-SBEKK

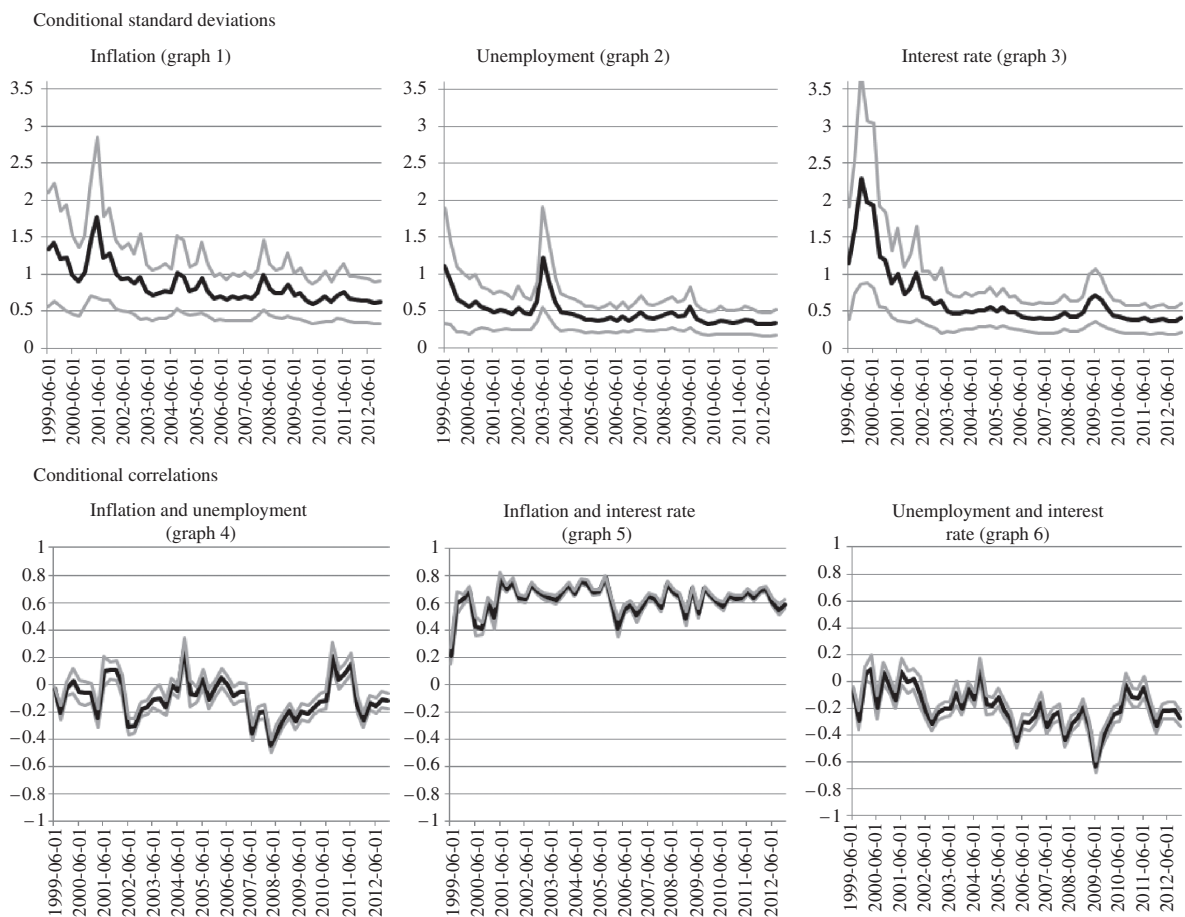


Figure 5: Conditional standard deviations and correlations (posterior mean $\pm$ 2 standard deviations). The results are obtained in the VEC-MSF-SBEKK model with $r=2$, $k=1$, $d=4$.

model with two lags in the VAR form, one cointegration relation and a constant restricted to the cointegration space. The second one is the model with the same VEC structure but with the MSF covariance structure, and the third specification is the VAR-MSF-SBEKK model for the first differences.

Figure 5 displays the time plots of conditional standard deviation and the cross correlations of the innovation processes in each of the three equations of the model with the highest posterior probability. Black lines represent posterior means, and the grey ones – the means plus/minus two posterior standard deviations. The posterior means of the conditional standard deviations lead us to the conclusion that all the error terms are heteroscedastic, but the areas between grey lines representing the uncertainty suggest that the Lindley-type test would probably indicate that the error term in the inflation equation could be homoscedastic. In the unemployment equation we can observe a high pick of volatility in year 2003, when the unemployment rate was the highest. The volatility of the interest rate shocks was clearly higher at the beginning of the analysed sample (in years 1999–2002), when the interest rates were high.

The posteriors of conditional correlations are quite tight, and the presented figures clearly indicate that these correlations keep changing during the analysed period. It can be seen that all the conditional correlations of error terms for interest rate and for inflation are positive and relatively high (Figure 5, graph 5). The conditional correlations between the error terms for the inflation and unemployment as well as for the unemployment and interest rate change their signs (Figure 5, graphs 4 and 6). Also, they share similar patterns, which probably results from positive correlations of the error terms in the equations of the interest rate and inflation.

As has been mentioned before, the analysed series is rather short, therefore we have decided to perform a sensitivity analysis of the Bayesian model comparison for two different prior specifications for the cointegration space. Contrary to the previous approach of setting the P_τ matrix, now it is determined with the

Table 4: Natural logarithm of estimates of marginal likelihoods, numerical standard errors and ranks.

k, d, r	VEC			VEC-MSF			VEC-MSF-SBEKK		
	$\ln p(x)$	NSE	Rank (under all models)	$\ln p(x)$	NSE	Rank (under all models)	$\ln p(x)$	NSE	Rank (under all models)
2,3,0	-196.94	0.46	3 (20)	-182.06	0.56	1 (2)	-177.58	0.32	1 (1)
2,3,1	-199.18	0.81	4 (23)	-188.74	0.56	4 (8)	-196.39	0.35	7 (17)
2,3,2	-203.16	0.98	5 (30)	-187.89	0.75	3 (6)	-202.69	1.17	10 (29)
2,3,3	-214.66	1.03	17 (50)	-204.34	0.89	18 (33)	-225.80	1.30	17 (63)
2,4,1	-195.25	1.09	1 (15)	-184.21	0.79	2 (3)	-187.70	0.88	3 (5)
2,4,2	-196.88	0.60	2 (19)	-190.41	1.04	5 (9)	-202.45	1.03	9 (27)
3,3,0	-204.55	1.03	6 (34)	-190.55	0.87	6 (10)	-187.38	0.60	2 (4)
3,3,1	-209.50	0.83	13 (45)	-198.11	0.74	12 (22)	-206.59	0.68	11 (40)
3,3,2	-211.16	0.84	15 (47)	-200.30	1.06	14 (25)	-223.44	1.10	14 (60)
3,3,3	-220.83	1.09	23 (58)	-212.93	0.88	22 (48)	-238.38	1.16	21 (67)
3,4,1	-208.46	0.72	11 (43)	-195.70	0.49	9 (16)	-193.88	0.80	6 (14)
3,4,2	-206.57	0.74	9 (39)	-193.21	0.70	7 (12)	-231.97	0.68	19 (65)
4,3,0	-210.72	0.92	14 (46)	-197.66	1.03	11 (21)	-192.89	0.64	5 (11)
4,3,1	-215.76	0.82	19 (52)	-207.44	1.06	21 (41)	-222.85	0.88	13 (59)
4,3,2	-216.67	0.66	21 (54)	-206.40	1.23	20 (38)	-247.65	1.37	22 (68)
4,3,3	-219.61	1.40	22 (57)	-219.12	0.96	23 (56)	-237.05	1.00	20 (66)
4,4,1	-216.63	1.11	20 (53)	-203.45	1.14	16 (31)	-217.02	1.04	12 (55)
4,4,2	-215.67	0.89	18 (51)	-199.62	0.73	13 (24)	-225.09	1.05	16 (62)
5,3,0	-205.24	1.15	7 (35)	-193.25	1.04	8 (13)	-188.09	1.05	4 (7)
5,3,1	-208.91	0.83	12 (44)	-205.88	1.16	19 (37)	-224.54	0.79	15 (61)
5,3,2	-212.96	0.75	16 (49)	-204.08	0.75	17 (32)	-253.58	1.18	23 (69)
5,4,1	-207.99	0.77	10 (42)	-196.73	1.03	10 (18)	-202.04	0.74	8 (26)
5,4,2	-205.82	1.25	8 (36)	-202.46	1.11	15 (28)	-231.00	1.11	18 (64)

k , VAR order; d , deterministic term; r , cointegration rank; $r=3$ – VAR for levels, $r=0$ – VAR for first differences. The P_τ matrix is determined with the help of the Johansen procedure employed in the model with $n-1$ cointegrating relations.

Table 5: The most probable models.

VAR order (k)	Deterministic term (d)	Cointegration rank (r)	Model (I)	$P(M_{(k,d,r,0)} \mathbf{x})$
2	3	0	VEC-MSF-SBEKK	0.98737
2	3	0	VEC-MSF	0.01116
2	4	1	VEC-MSF	0.00130
3	3	0	VEC-MSF-SBEKK	0.00005
2	4	1	VEC-MSF-SBEKK	0.00004
2	3	2	VEC-MSF	0.00003
5	3	0	VEC-MSF-SBEKK	0.00003
2	3	1	VEC-MSF	0.000014
2	4	2	VEC-MSF	0.000003
3	3	0	VEC-MSF	0.000002

The P_r matrix is determined with the help of the Johansen procedure employed in the model with $n - 1$ cointegrating relations.

help of the Johansen procedure employed in the model with $n - 1$ (rather than r) cointegrating relations. Tables 4 and 5 summarize the obtained results. The ranking of the most probable models with the homoscedastic covariance structure remains almost unchanged, whereas in the two other cases under study (i.e. VEC-MSF and VEC-MSF-SBEKK) we can observe some major differences in their performance. It emerges now that the models with two lags and no cointegrating relations are superior. As for the covariance structure we still have strong evidence for its time-variability, and almost all posterior model probability is gained by the MSF-SBEKK specification. Summing up, our sensitivity analysis shows that, as it might have been expected, in such richly parameterised models and with samples of so small sizes the prior assumptions affect the posterior results quite considerably.

5 Concluding remarks

The paper is focused on the construction and estimation of Bayesian Vector Error Correction -Stochastic Volatility (VEC-SV) models. As regards the stochastic volatility part of the specification we start with the MSF structure, which is the simplest among the MSV structures, thereby allowing a parsimonious way of modelling time-variability of volatility. Such a framework enables us to capture the long-run relationships among processes, and also to formally examine the presence of time-variation in the conditional covariance matrix.

The main conclusion of this paper is that both the heteroskedasticity and the long-run relationships have to be taken into account simultaneously in macroeconomic models as well as in financial analyses. Our application of the VEC-MSF and VEC-MSF-SBEKK models to exchange rates demonstrates that our methods can accurately estimate the long-run relationships in the VEC model for daily market prices. The results of Bayesian model comparison performed in the macroeconomic model example indicate that the ranking of different VEC structures (covering various numbers of lags and cointegrating relations, and two types of incorporated deterministic terms) is rather stable across different covariance matrix specifications.

The MSF structure allows conditional variances of the joint distribution of variables to change over time, but conditional correlations remain constant. In the MSF-SBEKK structure the conditional covariance matrix depends on the past of the process, as in the MGARCH models, which can lead to serious numerical obstacles. Thus, we deem that further research should rather include an MSV specification with stochastic conditional correlations based on more than one latent process.

On the other hand, time-variability of shocks has a potential effect on the propagation mechanism of the innovations (see, e.g. Koop, León-González, and Strachan 2009), so allowing parameters in the conditional mean of the process to be time-varying appear one of the most wanted generalisations.

Acknowledgments: We would like to thank the anonymous referees for very useful comments and suggestions that helped us to prepare the final version of our paper, and Łukasz Kwiatkowski for language verification of the manuscript. Research was partly financed from the funds granted to the Faculty of Management at Cracow University of Economics, within the framework of the subsidy for the maintenance of research potential. Research was partly financed from the funds granted to the Faculty of Finance and Law at Cracow University of Economics, within the framework of the subsidy for the maintenance of research potential.

References

- Bauwens, Luc, Michel Lubrano, and Jean-François Richard. 1999. *Bayesian Inference in Dynamic Econometric Models*. Oxford University Press, New York.
- Belmonte, Miguel, and Gary Koop. 2014. "Model Switching and Model Averaging in Time-Varying Parameter Regression Models." In *Bayesian Model Comparison*, edited by Ivan Jeliazkov and Dale J. Poirier, *Advances in Econometrics*, Vol. 34, 45–69. Bingley: Emerald Group Publishing Limited. Accessed June 12, 2015. doi: 10.1108/S0731-905320140000034004.
- Box, George E.P., and George C. Tiao. 1973. *Bayesian Inference in Statistical Analysis*. Massachusetts: Addison-Wesley Publishing Company.
- Cavaliere, Giuseppe, Luca De Angelis, Anders Rahbek, and A.M. Robert Taylor. 2015. "A Comparison of Sequential and Information-based Methods for Determining the Co-integration Rank in Heteroskedastic VAR Models." *Oxford Bulletin of Economics and Statistics* 77: 106–128.
- Chikuse, Yasuko. 2002. "Statistics on Special Manifolds." *Lecture Notes in Statistics* 174. Springer-Verlag, New York.
- Cogley, Timothy, and Thomas J. Sargent. 2005. "Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S." *Review of Economic Dynamics* 8: 262–302.
- Engle, Robert F., and C.W.J. Granger. 1987. "Co-integration and Error Correction: Representation, Estimation and Testing." *Econometrica* 55: 251–276.
- Gamerman, Dani. 1997. *Markov Chain Monte Carlo. Statistic Simulation for Bayesian Inference*. London: Chapman and Hall.
- Herwartz, Helmut, and Helmut Lütkepohl. 2011. "Generalized Least Squares Estimation for Cointegration Parameters Under Conditional Heteroskedasticity." *Journal of Time Series Analysis* 32 (3): 281–291.
- Jacquier, Eric, Nicholas G. Polson, and Peter E. Rossi. 1994. "Bayesian Analysis of Stochastic Volatility Models." *Journal of Business and Economic Statistics* 12 (4): 371–389.
- Jacquier, Eric, Nicholas G. Polson, and Peter E. Rossi. 1995. "Priors and Models for Multivariate Stochastic Volatility Models." CIRANO Working Paper No. 95s-18, Montréal. Accessed September 19, 2016. <http://www.cirano.qc.ca/files/publications/95s-18.pdf>.
- Koop, Gary. 2012. "Using VARs and TVP-VARs with Many Macroeconomic Variables." *Central European Journal of Economic Modelling and Econometrics* 4 (3): 43–167.
- Koop, Gary, Roberto León-González, and Rodney W. Strachan. 2009. "On the Evolution of Monetary Policy." *Journal of Economic Dynamics and Control* 33: 997–1017.
- Koop, Gary, Roberto León-González, and Rodney W. Strachan. 2010. "Efficient Posterior Simulation for Cointegrated Models with Priors on the Cointegration Space." *Econometric Reviews* 29: 224–242.
- Koop, Gary, Roberto León-González, and Rodney W. Strachan. 2011. "Bayesian Inference in the Time Varying Cointegration Model." *Journal of Econometrics* 165: 210–220.
- Koop, Gary, and Simon M. Potter. 2011. "Time Varying VARs with Inequality Restrictions." *Journal of Economic Dynamics and Control* 35: 1126–1138.
- Larsson, Rolf, and Mattias Villani. 2001. "A Distance Measure between Cointegration Spaces." *Economics Letters* 70: 21–27.
- Lenk, Peter. 2009. "Simulation Pseudo-bias Correction to the Harmonic Mean Estimator of Integrated Likelihoods." *Journal of Computational and Graphical Statistics* 18: 941–960.
- Newton, Michael A., and Adrian E. Raftery. 1994. "Approximate Bayesian Inference by the Weighted Likelihood Bootstrap." *Journal of the Royal Statistical Society B* 56 (1): 3–48.
- Osiewalski, Jacek. 2009. "New Hybrid Models of Multivariate Volatility (a Bayesian Perspective)." *Przegląd Statystyczny (Statistical Review)* 56 (1): 15–22.
- Osiewalski, Jacek, and Anna Pajor. 2009. "Bayesian Analysis for Hybrid MSF–SBEKK Models of Multivariate Volatility." *Central European Journal of Economic Modelling and Econometrics* 1 (2): 179–202.
- Osiewalski, Jacek, and Anna Pajor. 2010. "Bayesian Value-at-Risk for a Portfolio: Multi- and Univariate Approaches using MSF–SBEKK Models." *Central European Journal of Economic Modelling and Econometrics* 2 (4): 253–277.
- Osiewalski, Krzysztof, and Jacek Osiewalski. 2012. "Missing Observations in Daily Returns –Bayesian Inference within the MSF–SBEKK Model." *Central European Journal of Economic Modelling and Econometrics* 4 (3): 169–197.

- Osiewalski, Krzysztof, and Jacek Osiewalski. 2013. "A Long-Run Relationship between Daily Prices on Two Markets: The Bayesian VAR(2)–MSF-SBEKK Model." *Central European Journal of Economic Modelling and Econometrics* 5 (1): 65–83.
- Osiewalski, Jacek, and Mateusz Pipień. 2004. "Bayesian comparison of bivariate GARCH processes. The Role of the Conditional Mean Specification". In *New Directions in Macromodelling*, edited by Aleksander Welfe, 173–196, Amsterdam: Elsevier.
- Pajor, Anna. 2011. "A Bayesian Analysis of Exogeneity in Models with Latent Variables." *Central European Journal of Economic Modelling and Econometrics* 3 (2): 49–73.
- Pajor, Anna, and Jacek Osiewalski. 2012. "Bayesian Value-at-Risk and Expected Shortfall for a Large Portfolio (Multi- and Univariate Approaches)." *Acta Physica Polonica A* 121, (2-B): B-101-B-109.
- Primiceri, Giorgio E. 2005. "Time Varying Structural Vector Autoregressions and Monetary Policy." *Review of Economic Studies* 72: 821–852.
- Quintana, Jose M., and Michael West. 1987. "An analysis of international exchange rates using multivariate DLMS." *The Statistician* 36: 275–281.
- Seo, Byeongseon. 2007. "Asymptotic Distribution of the Cointegrating Vector Estimator in Error Correction Models with Conditional Heteroskedasticity." *Journal of Econometrics* 137: 68–111.
- Sims, Christopher A., and Tao Zha. 2006. "Were there Regime Switches in Macroeconomic Policy?" *American Economic Review* 96: 54–81.
- Villani, Mattias. 2006. "Bayesian Point Estimation of the Cointegration Space" *Journal of Econometrics* 134: 645–664.

Supplemental Material: The online version of this article (DOI: 10.1515/snde-2016-0004) offers supplementary material, available to authorized users.