PRZEGLĄD STATYSTYCZNY
R. LVI – ZESZYT 1 – 2009

ANNA PAJOR

BAYESIAN PORTFOLIO SELECTION WITH MSV MODELS*

1. INTRODUCTION

Markowitz mean–variance portfolio theory is the most popular approach to portfolio selection. In the Markowitz model the investor maximizes the expected return on the portfolio at a certain level of risk (measured by the standard deviation of return on the portfolio) or minimizes the risk at a certain level of expected return on the portfolio. The Markowitz theory (see [7]) assumes a known covariance matrix of asset returns. In this paper the covariance matrix of asset returns is modelled using multivariate stochastic volatility processes. An earlier Bayesian approach to the portfolio selection problem was considered by Winkler and Barry [16], Polson and Tew [12], and Soyer and Tanyeri [14]. In [16] Winkler and Barry consider a general multi-period model for portfolio selection. The investor chooses a portfolio to maximize the expected utility of his wealth at the end of a finite horizon. In their examples the data-generating process has a known variance and unknown mean. In [12] the minimum-variance portfolio is computed as a function of the predictive covariance matrix. But in the Polson and Tew model the predictive covariance matrix is time-invariant. In [14] Soyer and Tanyeri consider the multi-period portfolio selection problem from a Bayesian decision theoretic point of view. The multi-period optimal allocation problem is represented as a sequential decision problem (the investor maximizes the expected utility of his wealth at the end of a finite horizon problem). The work [1] introduces the dynamic factor models with stochastic volatility into dynamic one-period portfolio selection problem. A review of Bayesian works in portfolio management is provided in [13].

It is important to stress that in our previous work (see [10]) the one-period portfolio selection problem was considered. In this paper we consider the multi-period minimum conditional variance portfolio. In the optimization process we use the predictive distributions of future returns and the predictive conditional covariance matrices obtained from multivariate stochastic volatility models.

The other aim of the paper is to analyse and compare discrete-time Multivariate Stochastic Volatility (MSV) models from the point of view of their ability to select the optimal portfolio. The bivariate stochastic volatility models are used to describe the daily exchange rate of the euro against the Polish zloty and the daily exchange rate of the US dollar against the Polish zloty. Based on these two currencies we consider the Bayesian portfolio selection problem.

* Research supported by a grant from Cracow University of Economics.

In the next section we briefly present the optimal portfolio choice problem. Section 3 is devoted to the description of bivariate MSV specifications. In section 4 we present and discuss the empirical results. Some concluding remarks are presented in the last section.

2. PORTFOLIO SELECTION PROBLEM

Let $x_{j,t}$ denote the price of asset j (or the exchange rate as in our application) at time t for $j = 1,2, \dots, n$ and $t = 1, 2, \dots, T + s$. The vector of growth rates $y_t = (y_{1,t}, y_{2,t}, \dots, y_{n,t})'$, where $y_{j,t} = 100 \ln (x_{j,t}/x_{j,t-1})$, is modelled using the basic VAR(1) framework:

$$y_t - \delta = R(y_{t-1} - \delta) + \xi_t,\ t = 1,2,\dots,T,T+1,\dots,T+s, \tag{1}$$

where $\{\xi_t\}$ is a MSV process, T denotes the number of the observations used in estimation, and s is the forecast horizon, δ is a n-dimensional vector, R is a $n \times n$ matrix of parameters.

To introduce notation we denote by Θ_t the latent variable vector, by θ the parameter vector, and assume that

1) $\xi_t = \Sigma_t^{1/2} \varepsilon_t$, where $\{\varepsilon_t\} \sim iiN\,(0, I_n)$[1];

2) Σ_t is a function of the latent variables Θ_τ for $\tau \le t$, i.e. $\Sigma_t = \Sigma\,(\Theta_\tau; \tau \le t)$;

3) the vector ξ_t conditional on the σ-algebra $\sigma\,(\Theta_\tau; \tau \le t)$ is independent of $\sigma\,(\Theta_\tau; \tau > t)$.

Even thought these assumptions restrict the class of processes that we can consider, they significantly simplify the analysis and yield tractable formulae.

The s-period portfolio at time T is defined by a vector $w_{T|T+s} = (w_{1,T|T+s}, w_{2,T|T+s}, \dots, w_{n,T|T+s})'$, where $w_{i,T|T+s}$ is the fraction of wealth invested in asset i ($1 \le i \le n$). The return on the portfolio that places weight $w_{i,T|T+s}$ on asset i at time T is simply a weighted average of the returns on the individual assets. The weight applied to each return is the fraction of the portfolio invested in that asset:

$$R_{w,T|T+s} \approx \sum_{i=1}^{n} w_{i,T|T+s} z_{i,T|T+s}, \tag{2}$$

where $z_{i,T|T+s}$ is the rate of return on the asset i from the period T to $T + s$, i.e. $z_{i,T|T+s} = \sum_{t=T+1}^{T+s} y_{i,t}\ (i = 1,\dots,n)$. If $\Sigma_{T|T+s}$ is the matrix of conditional covariances of $z_{T|T+s}$, then the conditional variance of return on the portfolio is

$$Var\big(R_{w,T|T+s} \,\big|\, \psi_T, \Theta_T, \dots, \Theta_{T+s}\big) = V^2_{T|T+s}{}'\Sigma_{T|T+s} w_{T|T+s}, \tag{3}$$

where ψ_T is the σ-algebra generated by ε_τ Θ_τ for $\tau \le T$ i.e. ψT $= \sigma\,(\varepsilon_\tau, \Theta_\tau; \tau \le T)$.

[1] $\{\varepsilon_t\}$ is a sequence of independent and identically distributed normal random vectors with mean vector zero and covariance matrix I_n.

The vector of the rates of return at time $T + k$ ($k \leq s$) satisfies:

$$y_{T+k} - \delta = R^k (y_T - \delta) + \sum_{j=1}^{k} R^{k-j} \xi_{T+j}. \tag{4}$$

Let $G_{T+k} = \sigma(\Theta_\tau; \tau \leq T + k)$ be the σ-algebra generated by Θ_τ, $\tau \leq T + k$. That is, G_{T+k} contains all information about the latent variables up to time $T + k$ (is the entire path of the covariance matrix process). Based on equation (4) we have:

$$y_{T+k} \,|\, \psi_T, G_{T+k} \sim N\left(\delta + R^k (y_T - \delta), \sum_{j=1}^{k} R^{k-j} \Sigma_{T+j} (R^{k-j})'\right). \tag{5}$$

Under our assumptions:

$$z_{T|T+s} = s\delta + \sum_{j=1}^{s} R^j (y_T - \delta) + \sum_{j=1}^{s} \xi_{T+j} \sum_{i=0}^{s-j} R^i, \tag{6}$$

and

$$z_{T|T+s} \,|\, \psi_T, G_{T+s} \sim N\left(s\delta + \sum_{j=1}^{s} R^j (y_t - \delta), \Sigma_{T|T+s}\right), \tag{7}$$

where
$$\Sigma_{T|T+s} = \sum_{j=1}^{s} \left(\sum_{i=0}^{s-j} R^i\right) \Sigma_{T+j} \left(\sum_{i=0}^{s-j} R^i\right)'. \tag{8}$$

Consequently, the conditional variance of return on the portfolio is:

$$Var\left(R_{w,T|T+s} \,|\, \psi_T, \Theta_T, \ldots, \Theta_{T+s}\right) = w_{T|T+s}' \sum_{j=1}^{s} \left(\sum_{i=0}^{s-j} R^i\right) \Sigma_{t+j} \left(\sum_{i=0}^{s-j} R^i\right)' w_{T|T+s}.$$

The usual portfolio constrain gives $w_{1,T|T+s} + w_{2,T|T+s} + \ldots + w_{n,T|T+s} = 1$. We assume that short sales are allowed and $w_{i,T|T+s} < 0$ reflects a short selling. The standard approach assumes that the investor selects the portfolio with minimum variance (see [2]). Here we assume that the investor minimises the conditional variance of the portfolio. Then the problem for the investor reduces to solving the quadratic programming problem:

$$\min_{w_{T|T+s}} w_{T|T+s}' \Sigma_{T|T+s} w_{T|T+s} \text{ subject to } w_{1,T|T+s} + w_{2,T|T+s} + \ldots + w_{n,T|T+s} = 1.$$

In this way we obtain so-called the minimum conditional variance portfolio (the portfolio that has the lowest risk of any feasible portfolio):

$$w_{MV,T|T+s} = \frac{\Sigma_{T|T+s}^{-1} \iota}{\iota' \Sigma_{T|T+s}^{-1} \iota}, \tag{9}$$

which has a return:

$$R_{MV,T|T+s} = \frac{\iota' \sum_{T|T+s}^{-1} z_{T|T+s}}{\iota' \sum_{T|T+s}^{-1} \iota}, \tag{10}$$

and the conditional variance at time T:

$$Var\left(w_{MV,T|T+s}' y_T \,\middle|\, \psi_T, G_{T+s}\right) = V_{MV,T|T+s}^2 = \frac{1}{\iota' \sum_{T|T+s}^{-1} \iota}, \tag{11}$$

where ι is an $n \times 1$ vector of ones.

We see that the minimum conditional variance portfolio $w_{MV,T|t+s}$ and its conditional variance $V_{MV,T|T+s}^2$ depend on only the conditional covariance matrix $\Sigma_{T|T+s}$. Thus the choice of $w_{MV,T|T+s}$ is made at time T based on the predictive distribution for $\Sigma_{T|T+s}$ conditional on past data and information up to time T, ψ_T. In the Bayesian framework we have the predictive distribution of $\Sigma_{T|T+s}$ given by $p\,(\Sigma_{T|T+s}|\mathrm{y})$, which implies the predictive distribution of $w_{MV,T|t+s}$ and $V_{MV,T|T+s}^2$.

Now we consider a s-period portfolio selection problem where the investor wants to minimise the conditional variance of the portfolio with a given level of return $R_{w,T|T+s} \geq R^*_{p,T|T+s}$. This problem reduces to solving the quadratic programming problem:

$$\min_{w_{T|T+s}} w_{T|T+s}' \Sigma_{T|T+s} w_{T|T+s} \text{ subject to } \begin{cases} w_{T|T+s}' \iota = 1, \\ w_{T|T+s}' z_{T|T+s} \geq R^*_{p,T|T+s}, \end{cases}.$$

When $R_{w,T|T+s} = R^*_{p,T|T+s}$, the solution for the s-period portfolio is:

$$w_{MVR^*_p,T|T+s} = \frac{\left(\sum_{T|T+s}^{-1} z_{T|T+s} \iota' \sum_{T|T+s}^{-1} - \sum_{T|T+s}^{-1} \iota z_{T|T+s}' \sum_{T|T+s}^{-1}\right)\left(\iota R^*_{p,T|T+s} - z_{T|T+s}\right)}{\left(\iota' \sum_{T|T+s}^{-1} \iota\right)\left(z_{T|T+s}' \sum_{TT+s}^{-1} z_{T|T+s}\right) - \left(\iota' \sum_{T|T+s}^{-1} z_{T|T+s}\right)^2}. \tag{12}$$

Note that the classic portfolio choice scheme assumes the covariance matrix and expected returns at time T to be known. But the Bayesian approach to inference naturally leads to the posterior or predictive distributions of these quantities. Note that the minimum conditional variance portfolio $w_{MV,T|T+s}$, and the minimum conditional variance portfolio with a given level of return $w_{MVR^*_p,T|T+s}$ are random variables. The posterior (or predictive) distributions of $w_{MV,T|T+s}$, and $w_{MVR^*_p,T|T+s}$ or $V_{MV,T|T+s}$, and $V_{MVR^*_p,T|T+s}$ are induced by the distribution of $z_{T|T+s}$, and $\Sigma_{T|T+s}$. Thus the predictive distribution of the minimum variance portfolio $w_{MV,T|T+s}$ or $w_{MVR^*_p,T|T+s}$ can be used to provide an optimal portfolio. The optimal portfolio can be defined as the expected value (if there exists) of $w_{MV,T|T+s}$ or $w_{MVR^*_p,T|T+s}$. As the predictive mean (for $w_{MV,T|T+s}$ or $w_{MVR^*_p,T|T+s}$) may not exist, we can consider the predictive medians $w^{op}_{MV,T|T+s}$ or $w^{op}_{MVR^*_p,T|T+s}$ defined respectively by conditions:

$$\Pr\left\{w_{MV,i,T|T+s} \geq w^{op}_{MV,i,T|T+s} \,|\, y\right\} \leq 0.5 \text{ and } \Pr\left\{w_{MV,i,T|T+s} \leq w^{op}_{MV,i,T|T+s} \,|\, y\right\} \geq 0.5$$
$$\Pr\left\{w_{MVR^*_p,i,T|T+s} \geq w^{op}_{MVR^*_p,i,T|T+s} \,|\, y\right\} \leq 0.5 \text{ and}$$
$$\Pr\left\{w_{MVR^*_p,i,T|T+s} \leq w^{op}_{MVR^*_p,i,T|T+s} \,|\, y\right\} \geq 0.5, (i = 1, \ldots, n-1).$$

It is important to note that even for the simple case of $n = 2$ assets, there is no analytical solution for the optimal portfolio selection problem when we consider the MSV model. In this case one alternative approach is to use Monte Carlo methods to evaluate the quantiles of the posterior (or predictive) distributions of $w_{MV,T|T+s}$ and $w_{MVR^*_p,T|T+s}$, and then find the portfolio.

3. BIVARIATE STOCHASTIC VOLATILITY MODELS

The vector of growth rates $y_t = (y_{1,t}, y_{2,t})'$ is modelled here using the framework (1). In (1) δ is a 2-dimensional vector, R is a 2×2 matrix of parameters, and ξ_t is a bivariate SV process. By restricting to only bivariate time series, it is possible to estimate unparsimoniously parameterised MSV models. We assume that, conditionally on vector $\Theta_{t(i)}$ (consisting of model-specific latent variables) and the parameter vector θ_i, ξ_t follows a bivariate Gaussian distribution with mean vector $0_{[2\times1]}$ and covariance matrix Σ_t, i.e. $\xi_t \,|\, \Theta_{t(i)}, \theta_i \sim N(0_{[2\times1]}, \Sigma_t), t = 1, 2, \ldots, T+s$. Competing bivariate SV models are defined by imposing different structures on Σ_t. The three distinct elements of Σ_t can be described by one, two or three separate latent processes.

The elements of δ and R are common parameters. We assume for them the multivariate standard Normal prior $N(0, I_6)$, truncated by the restriction that all eigenvalues of R lie inside the unit circle. These parameters and the remaining (model-specific) parameters are a prior independent.

3.1. STOCHASTIC DISCOUNT FACTOR MODEL – SDF

The first specification considered here is the stochastic discount factor model (SDF) proposed in [6] by Jacquier, Polson and Rossi. The SDF process is defined as follows:

$$\begin{gathered} \xi_t = u_t \sqrt{h_t},\ \ln h_t = \phi \ln h_{t-1} + \sigma_h \eta_t, \\ \{u_t\} \sim iiN(0_{[2\times1]}, \Sigma),\ \{\eta_t\} \sim iiN(0,1),\ u_{j,t} \perp \eta_s,\ t,s \in Z, j = 1,2. \end{gathered} \tag{13}$$

Here $\{u_t\}$ is a sequence of independent and identically distributed normal random vectors with mean vector zero and constant covariance matrix Σ. Thus, we have $\xi_t \,|\, \Theta_{t(i)}, \theta_1 \sim N(0_{[2\times1]}, h_t\Sigma)$, where $\Theta_{t(1)} = h_t$. The conditional covariance matrix of ξ_t is time varying and stochastic, but all its elements have the same dynamics governed by h_t. Thus, the conditional correlation coefficient is time invariant.

In order to complete the Bayesian model, we have to specify a prior distribution on the parameter space. We assume the following prior structure:

$$p(\phi,\sigma_h^2,\ln h_0,\Sigma) = p(\phi)p(\sigma_h^2)p(\ln h_0)p(\Sigma),$$

where we use proper prior densities of the following distributions:

$$\phi \sim N(0,10^2)I_{(-1,1)}(\phi),\ \sigma_h^2 \sim IG(1,0.005),\ \ln h_0 \sim N(0,10^2),\ \Sigma \sim IW(2I,2,2).$$

The prior distribution for $(\phi,\sigma_h^2)'$ is the same as in the univariate SV model (see [9]). $I_{(-1,\ 1)}(.)$ denotes the indicator function of the interval (–1, 1). The symbol $IG(v_0, s_0)$ denotes the inverse Gamma distribution with mean $s_0/(v_0\text{-}1)$ and variance $s_0^2/\left[(v_0-1)^2(v_0-2)\right]$ (thus, here the prior mean for σ_h^2 does not exist, but σ_h^{-2} has a Gamma prior with mean 200 and standard deviation 200). The symbol $IW(B, d, 2)$ denotes the two-dimensional inverse Wishart distribution with d degrees of freedom and parameter matrix B (see [17]). $\ln h_0$ is treated as an additional parameter and estimated jointly with other parameters. The prior distributions used are relatively noninformative.

3.2. BASIC STOCHASTIC VOLATILITY MODEL – BSV

Next, we consider the basic stochastic volatility process (BSV), where $\xi_{1,t}$ and $\xi_{2,t}$ follow independent univariate SV processes $\xi_t \mid \Theta_{t(2)} \sim N(0_{[2\times 1]},\Sigma_t)$, where $\Sigma_t = Diag\,(h_{1,t}, h_{2,t})$. The conditional variance equations are

$$\ln h_{1,t} - \gamma_{11} = \phi_{11}(\ln h_{1,t-1} - \gamma_{11}) + \sigma_{11}\eta_{1,t}, \ln h_{2,t} - \gamma_{22} = \phi_{22}(\ln h_{2,t-1} - \gamma_{22}) + \sigma_{22}\eta_{2,t},$$

where $\eta_t = (\eta_{1,t},\eta_{2,t})'$, $\eta_t \sim iiN(0_{[2\times 1]},I_2)$, $\Theta_{t(2)} = (h_{1,t},h_{2,t})'$.

For the parameters we use the same specification of prior distribution as in the univariate SV model (see [9]), i.e. $\gamma_{jj} \sim N(0, 10^2)$, $\phi_{jj} \sim N(0, 10^2)\,I_{(-1,1)}(\phi_{jj})$, $\sigma_{jj}^2 \sim IG(1, 0.005)$, $\ln h_{j,0} \sim N(0, 10^2)$, $j = 1, 2$.

3.3. JSV MODEL

Now, we consider a SV process based on the spectral decomposition of the matrix Σ_t (see [11]). That is

$$\Sigma_t = P\ \Lambda_t\ P^{-1}, \tag{14}$$

where $\Lambda_t = Diag\,(\lambda_{1,t}, \lambda_{2,t})$ is the diagonal matrix consisting of all eigenvalues of Σ_t and P is the matrix consisting of the eigenvectors of Σ_t. For series $\{\ln\lambda_{j,t}\}$ (j = 1, 2), similarly as in the univariate SV process, we assume standard univariate autoregressive processes of order one, namely $\ln\lambda_{1,t} - \gamma_{11} = \phi_{11}(\ln\lambda_{1,t-1} - \gamma_{11}) + \sigma_{11}\eta_{1,t}$, $\ln\lambda_{2,t} - \gamma_{22} = \phi_{22}(\ln\lambda_{2,t-1} - \gamma_{22}) + \sigma_{22}\eta_{2,t}$,

where $\eta_t = (\eta_{1,t},\eta_{2,t})'$ and $\eta_t \sim iiN(0_{[2\times 1]},I_2)$, $\Theta_{t(3)} = (\lambda_{1,t},\lambda_{2,t})'$.

Log transformation for $\lambda_{j,t}$ is used to ensure the positiveness of Σ_t. If $|\phi_{jj}| < 1$ ($j = 1, 2$) then $\{\ln \lambda_{1,t}\}$ and $\{\ln \lambda_{2,t}\}$ are stationary and the JSV process is a white noise. In addition, P is an orthogonal matrix, i.e. $P'P = I_2$. The conditional covariance matrix of ξ_t can be written as (see [11]):

$$\Sigma_t = \begin{bmatrix} \lambda_{1,t} p_{11}^2 + \lambda_{2,t}(1 - p_{11}^2) & (\lambda_{1,t} - \lambda_{2,t}) p_{11}\sqrt{1 - p_{11}^2} \\ (\lambda_{1,t} - \lambda_{2,t}) p_{11}\sqrt{1 - p_{11}^2} & \lambda_{2,t} p_{11}^2 + \lambda_{1,t}(1 - p_{11}^2) \end{bmatrix}, p_{11} \in (0,1]. \tag{15}$$

Consequently, the conditional correlation coefficient is time-varying and stochastic if $p_{11} \neq 1$.

For the model-specific parameters we take the following prior distributions: $\gamma_{jj} \sim N(0, 10^2)$, $\phi_{jj} \sim N(0, 10^2) I_{(-1,1)}(\phi_{jj})$, $\sigma_{jj}^2 \sim IG(1, 0.005)$, $\ln \lambda_{j,0} \sim N(0, 10^2)$, $j = 1, 2$; $p_{11} \sim U(0, 1)$ (i.e. uniform over (0, 1)). Note that if $p_{11} = 1$, then we obtain the BSV model, but we formally exclude this value.

3.4. SJSV MODEL

In the JSV model the structure of the conditional covariance matrix is based on two separate latent variables. The next specification uses three separate latent processes (see [11]). In the definition of the JSV model we replace p_{11} by a process $p_{11,t}$ with value in (0,1]. Thus, we have: $w_t - \gamma_{21} = \phi_{21}(w_{t-1} - \gamma_{21}) + \sigma_{21}\eta_{21,t}$, $w_t = \ln[p_{11,t}/(1 - p_{11,t})]$, $\eta_t = (\eta_{11,t}, \eta_{22,t}, \eta_{21,t})'$, $\{\eta_t\} \sim iiN(0_{[3\times 1]}, I_3)$, $\Theta_{t(4)} = (\lambda_{1,t}, \lambda_{2,t}, p_{11,t})'$.

Now the number of the latent processes is equal to the number of distinct elements of the conditional covariance matrix. We assume the same prior distributions as previously.

3.5. TSV MODEL

The next specification (proposed by Tsay in [15], thus called TSV) uses the Cholesky decomposition of the conditional covariance matrix:

$$\Sigma_t = L_t G_t L_t', \tag{16}$$

where L_t is a lower triangular matrix with unitary diagonal elements, G_t is a diagonal matrix with positive diagonal elements:

$$L_t = \begin{bmatrix} 1 & 0 \\ q_{21,t} & 1 \end{bmatrix}, G_t = \begin{bmatrix} q_{11,t} & 0 \\ 0 & q_{22,t} \end{bmatrix}, \Sigma_t = \begin{bmatrix} \sigma_{11,t}^2 & \sigma_{12,t} \\ \sigma_{21,t} & \sigma_{22,t}^2 \end{bmatrix} = \begin{bmatrix} q_{11,t} & q_{11,t} q_{21,t} \\ q_{11,t} q_{21,t} & q_{11,t} q_{21,t}^2 + q_{22,t} \end{bmatrix}.$$

Series $\{q_{21,t}\}$, and $\{\ln q_{jj,t}\}$ ($j = 1, 2$), analogous to the univariate SV, are standard univariate autoregressive processes of order one, namely $q_{21,t} - \gamma_{21} = \phi_{21}(q_{21,t-1} - \gamma_{21}) + \sigma_{21}\eta_{21,t}$,

$$\ln q_{11,t} - \gamma_{11} = \phi_{11}\left(\ln q_{11,t-1} - \gamma_{11}\right) + \sigma_{11}\eta_{11,t}, \ln q_{22,t} - \gamma_{22} =$$
$$= \phi_{22}\left(\ln q_{22,t-1} - \gamma_{22}\right) + \sigma_{22}\eta_{22,t},$$

where $\eta_t = \left(\eta_{11,t}, \eta_{21,t}, \eta_{22,t}\right)'$ and $\{\eta_t\} \sim iiN(0_{[3\times 1]}, I_3), \Theta_{t(5)} = \left(q_{11,t}, q_{22,t}, q_{21,t}\right)'$.

We make similar assumptions about the prior distributions as previously. In particular: $\gamma_{ij} \sim N(0, 10^2)$, $\phi_{ij} \sim N(0, 10^2)I_{(-1,1)}(\phi_{ij})$, $\sigma_{ij}^2 \sim IG(1, 0.005)$, $\ln q_{ii,0} \sim N(0, 10^2)$ $i, j \in \{1, 2\}$, $i \geq j$, $q_{21,0} \sim N(0, 10^2)$.

3.6. BIVARIATE DCC-SDF MODEL

We consider also the DCC-SDF model proposed by Osiewalski and Pajor in [8]. The DCC–SDF model allows for different dynamics of each conditional variance or covariance (like DCC in [3]) and keeps just one latent process in the conditional covariance matrix in order to describe outliers (like SDF). The DCC-SDF process is defined as follows:

$$\begin{gathered}\xi_t = \Sigma_t^{1/2}\varepsilon_t\sqrt{h_t}, \ln h_t = \gamma + \phi(\ln h_{t-1} - \gamma) + \sigma_h\eta_t, \\ \{\varepsilon_t\} \sim iiN(0_{[2\times 1]}, I_2), \{\eta_t\} \sim iiN(0, 1), \varepsilon_{j,t} \perp \eta_s, t, s \in Z, j = 1, 2,\end{gathered} \tag{17}$$

where for the diagonal element of matrix Σ_t it is assumed

$$\sigma_{11,t}^2 = 1 + \alpha_1\xi_{1,t-1}^2 + \beta_1\sigma_{11,t-1}^2, \sigma_{11,t}^2 = a^2 + \alpha_2\xi_{2,t-1}^2 + \beta_2\sigma_{22,t-1}^2,$$

and for the off-diagonal element it is assumed $\sigma_{12,t} = \rho_{12,t}\sqrt{\sigma_{11,t}^2\sigma_{22,t}^2}$, where $\rho_{12,t}$ is the time-varying conditional correlation coefficient, modelled as $\rho_{12,t} = q_{12,t}\big/\sqrt{q_{11,t}q_{22,t}}$ with $q_{ij,t}$'s being entries of a symmetric positive definite matrix Q_t of the same order as the dimension of ξ_t. A simple specification for Q_t, considered by Engle in [3], assumes that

$$Q_t = (1 - b - c)S + b\xi_{t-1}^*\xi_{t-1}^{*'} + cQ_{t-1},$$

where b and c are nonnegative scalar parameters ($b + c < 1$), $\xi_{i,t}^* = \xi_{i,t}/\sigma_{ii,t}$, $i = 1, 2$. We keep Engle's basic structure and define S as a square matrix with ones on the diagonal and ρ_{12}, an unknown parameter from the interval (–1, 1). The initial condition for Q_t is $Q_0 = q_0 I_2$ with free $q_0 > 0$. We assume that a priori (b c)' is uniform over the unit simplex, $q_0 \sim Exp(1)$ (i.e. exponential with mean 1), $\gamma \sim N(0, 10^2)$, $a \sim Exp(1)$, $\rho_{12} \sim U([-1, 1])$ and $(\alpha_1, \alpha_2, \beta_1, \beta_2) \sim U([0, 1]^4)$, $\phi \sim N(0, 10^2)\, I_{(-1,1)}(\phi)$, $\sigma_h^2 \sim IG(1, 0{,}005)$, $\ln h_0 \sim N(0, 10^2)$. Analogous to [3], when the sum $b + c$ is equal to one, we have the integrated DCC-SDF model (IDCC-SDF).

4. EMPIRICAL RESULTS

We consider the daily exchange rate of the euro against the Polish zloty and the daily exchange rate of the US dollar against the Polish zloty from January 2, 2002 to June 29, 2007. The data were downloaded from the website of the National Bank of Poland. The dataset of the daily logarithmic growth (return) rates, y_t, consists of 1388 observations (for each series). As the first growth rates are used as initial conditions, thus T = 1387 remaining observations on y_t are modelled. The data are plotted in Figure 1. The return rates seem to be centred around zero, with changing volatility and the presence of outliers. The sample correlation (equals 0.598) indicates that the returns are positively correlated.

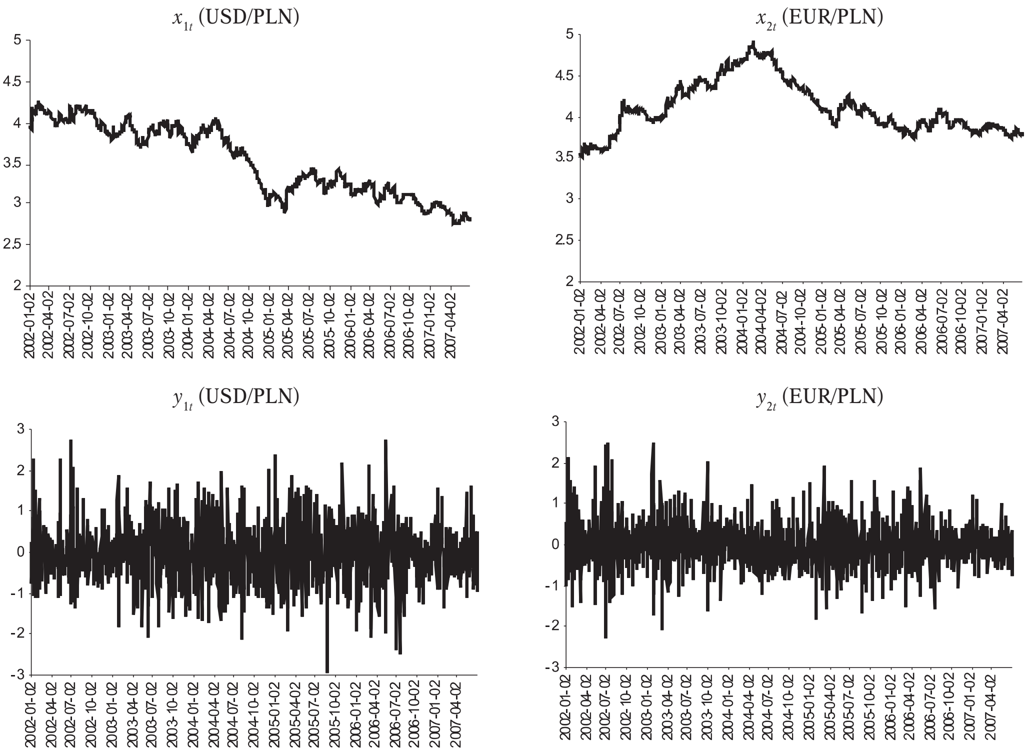

Figure 1. Daily exchange rates of the USD/PLN and the EUR/PLN (January 2, 2002 – June 29, 2007), and daily rates of return

Source: own elaboration.

4.1. BAYESIAN MODEL COMPARISON[2]

In Table 1 we rank the models by the increasing value of the decimal logarithm of the Bayes factor of VAR(1)-SJSV against the alternative models. Because in the

[2] All presented results were obtained with the use of the Gibbs sampler using 10^5 iterations after 5×10^4 burn-in Gibbs steps; see [4], [5], and [10] for details.

VAR(1)-TSV specification the conditional variances are not modelled in a symmetric way, we consider two cases: the VAR(1)-TSV_{USD_EUR} and VAR(1)-$TSV_{EUR\text{-}USD}$ models. These models differ in ordering of elements in y_t. In the VAR(1)-TSV_{USD_EUR} model $y_{1,t}$ denotes the daily growth rate of the PLN/USD, and $y_{2,t}$ is the daily growth rate of the PLN/EUR. In the VAR(1)-TSV_{EUR_USD} model the ordering of components in y_t is contrary to previous one.

Table 1

Logs of Bayes factors in favour of VAR(1)-SJSV model

Model	Number of latent processes	Number of parameters	Log_{10} ($B_{SJSV,i}$)	Rank
VAR(1)-SJSV	3	18	0	1
VAR(1)-TSV_{EUR_USD}	3	18	8.51	2
VAR(1)-TSV_{USD_EUR}	3	18	11.10	3
VAR(1)-JSV	2	15	19.60	4
VAR(1)-IDCC-SDF	1	18	32.00	5
VAR(1)-DCC-SDF	1	20	33.88	6
VAR(1)-SDF	1	12	50.20	7
VAR(1)-BMSV	2	14	158.51	8

Source: own calculations.

We see that for our data set the models with three latent processes describe the time-varying conditional covariance matrix much better than the models with one or two latent processes. The VAR(1)-SJSV model wins our model comparison, being about 8.5 orders of magnitude better than the VAR(1)-TSV_{EUR_USD} model. The VAR(1)-BMSV model with zero correlations is the worst (has the lowest marginal data density). The decimal log of the Bayes factor of the VAR(1)-BMSV model relative to the VAR(1)-SJSV model is 158.51. Assuming equal prior model probabilities, the VAR(1)-SDF model (with the constant conditional correlations) is about 108 orders of magnitude more probable a posterior than the VAR(1)-BMSV model, but about 18 orders of magnitude worse than the VAR(1)-IDCC-SDF model and about 50 orders of magnitude worse than the VAR(1)-SJSV model. The results indicate that the data strongly reject the assumption of zero or constant conditional correlation coefficient. Of course, our model comparison relies on the prior distributions for the parameters of the models, but these prior distributions are not very informative.

4.2. PORTFOLIO SELECTION WITH MSV MODELS

In this section we report the results of building the optimal portfolios using the MSV models. We consider the hypothetical portfolios, which consist of two currencies:

the US dollar and euro. We assume that there are no transaction costs and that we may reallocate zloty to long as well as to short positions across the currencies. Allocation decisions are made at time T based on the predictive distribution for y_{T+k} and Σ_{T+k} for $k = 1, \ldots, 60$.

In Figure 2 we show the quantiles of the predictive distributions of the minimum conditional variance portfolio $w_{MV,1,T|T+s}$ (the fraction of wealth invested in the US dollar). If the medians of the marginal predictive distributions are treated as point forecasts, in model with time-varying conditional correlation coefficient the optimum weights to invest in the PLN/USD are negative, indicating the short sale of the US dollar (the median of the marginal predictive distribution of $w_{MV,1,T|T+s}$ is equal to about -0.5). The short position on the US dollar is connected with corresponding long position on the euro.

We see that the predictive distributions are very widely dispersed and fat-tailed, thus leaving us with considerable uncertainty about the future returns of these portfolios. Surprisingly, in the VAR(1)-MSV models with one latent process the minimum conditional variance portfolios are estimated very precisely – the inter-quartile ranges are relatively small. Note that the predictive distribution of $w_{MV,1,T|T+s}$ produced by the VAR(1)-DCC-SDF model is located in areas of high predictive density obtained in the best model (i.e. VAR(1)-SJSV).

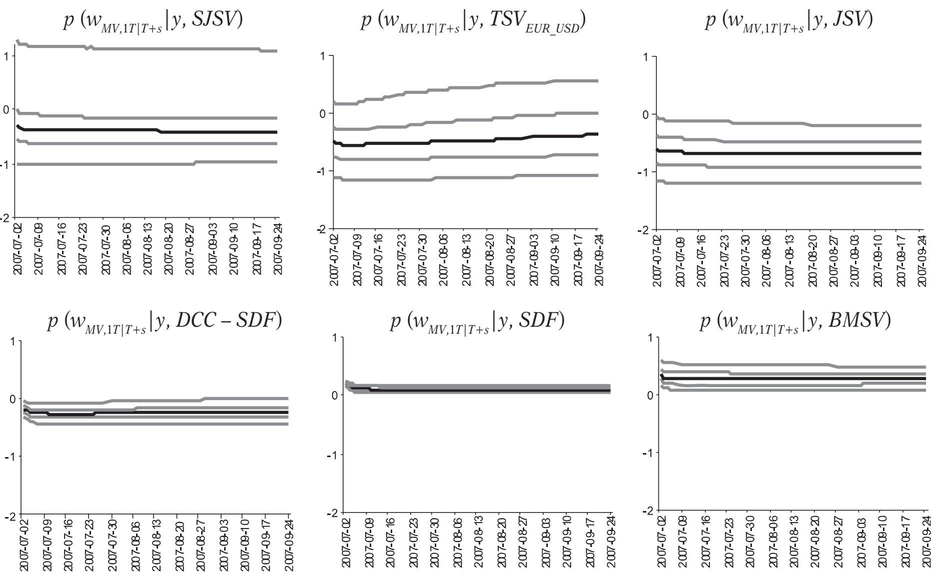

Figure 2. Quantiles of the predictive distributions of the minimum conditional variance portfolios (the fractions of wealth invested in the US dollar). The central black line represents the medians, the grey lines represent the quantiles of order 0.05, 0.25, 0.75, 0.95, respectively

Source: own elaboration.

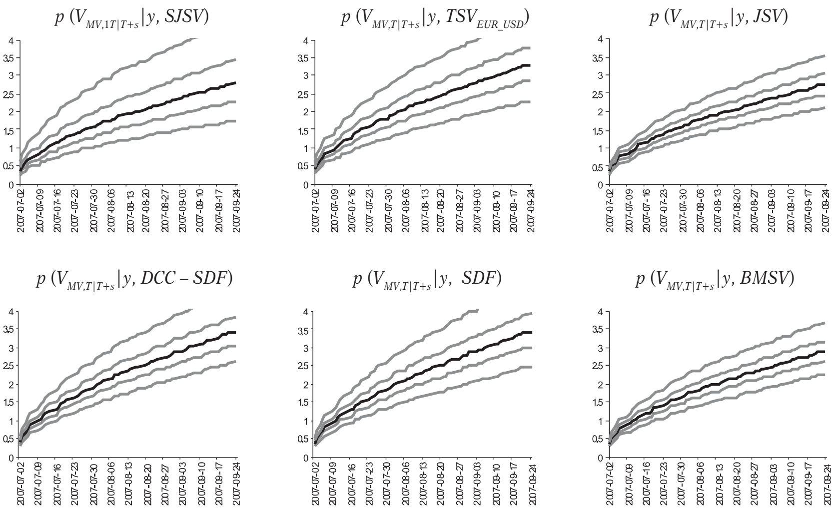

Figure 3. Quantiles of the predictive distributions of the conditional standard deviation of the minimum conditional variance portfolios. The central black line represents the medians, the grey lines represent the quantiles of order 0.05, 0.25, 0.75, 0.95, respectively

Source: own elaboration.

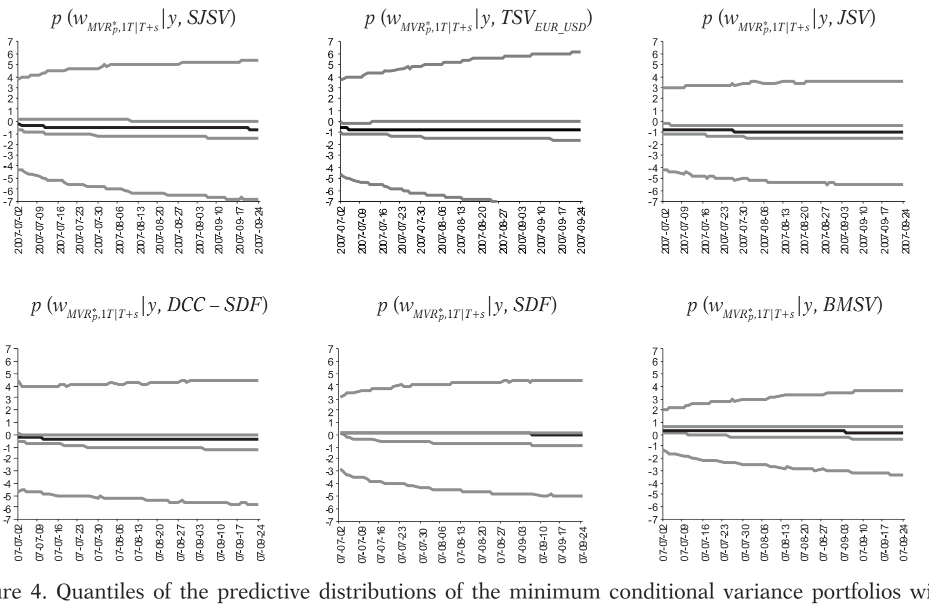

Figure 4. Quantiles of the predictive distributions of the minimum conditional variance portfolios with the return equals at least 5% on annual base (the fraction of wealth invested in the US dollar). The central black line represents the medians, the grey lines represent the quantiles of order 0.05, 0.25, 0.75, 0.95, respectively

Source: own elaboration.

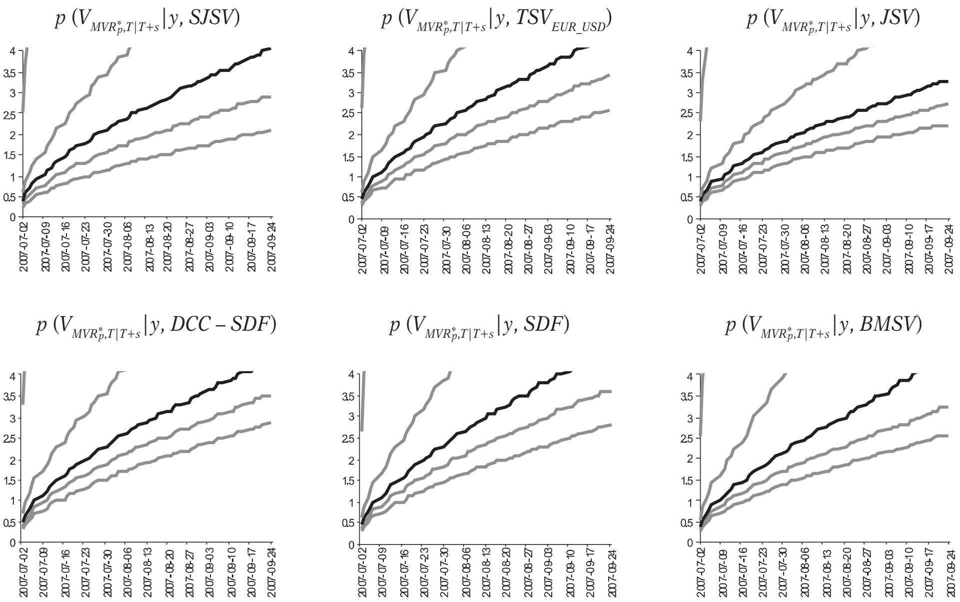

Figure 5. Quantiles of the predictive distributions of the conditional standard deviation of the minimum conditional variance portfolio with the return equals at least 5% on annual base. The central black line represents the medians, the grey lines represent the quantiles of order 0.05, 0.25, 0.75, 0.95, respectively

Source: own elaboration.

The predictive distributions related to the portfolio with bound on return are more diffuse – the inter-quartile ranges are higher (see Figure 4 and 5). Comparing the minimum conditional variance portfolio and the minimum conditional variance portfolio with the return equals at least 5%, we can see that the distributions of the forecasted value of $w_{MVR_p^*T|T+s}$ and $V_{MVR_p^*T|T+s}$ are more dispersed and have very thick tails. Thus uncertainty connected with the optimal portfolio with return at least 5% on annual base is huge. The quantiles of the conditional standard deviation of the optimal portfolios (see Figure 3 and 5) indicate increasing volatility with the forecast horizon.

Finally we use the medians of $w_{MVR_p^*,1,T|T+s}$ to construct hypothetical portfolios for $s = 1, 2, \ldots, 60$. Let $W_T = 10000$ PLN be the initial wealth of the investor at time T (on June 29, 2007). If we assume that there are no transaction costs and the investor uses the median of the predictive distribution of $w_{MVR_p^*T|T+s}$ (denoted by $w^{op}_{MVR_p^*,1,T|T+s}$) to construct optimal portfolio, then the investor's wealth at time $T + s$ is given by:

$$W_{MVR_p^*,T|T+s} = W_T\left[w^{op}_{MVR_p^*,1,T|T+s}\left(x_{1,T+s}/x_{1,T}\right) + w^{op}_{MVR_p^*,2,T|T+s}\left(x_{2,T+s}/x_{2,T}\right)\right],$$
$$s = 1,2,\ldots,60.$$

In Figure 6, we present the plot of $W_{MVR^*_p,T|T+s}$ for s = 1, 2, …, 60, and compare them with a bank deposit with the interest rate equal to 4.7% on annual base (the quotation of the 3-month Warsaw Interbank Offered Rate on June, 29 2007). Surprisingly, the best results we obtained in the VAR(1)-JSV model – at a 2-month horizon the average return of the optimal portfolios is equal to 0.098%, which represents annual return of 24.58%. In the VAR(1) – TSV models the average return of the optimal portfolios is equal to about 0.082% (i.e. 20.57% per annum), whereas in the VAR(1)-DCC-SDF and VAR(1)-IDCC-SDF models we have 0.041% and 0.053%, respectively. It is important to stress that the returns of the hypothetical investments are higher than of the bank deposit, indicating good forecasting properties of the MSV models with the time–varying conditional correlations. In VAR(1)-MSV model with zero or constant conditional correlation the average return of the portfolio is negative (in the VAR(1)-SDF and VAR(1)-BMSV models we obtained -0.001% and -0.019%, i.e. -0.25% and -4.72% per annum, respectively). Note that the average return of equally – weighted portfolio is equal to -0.047, i.e. -11.80% per annum.

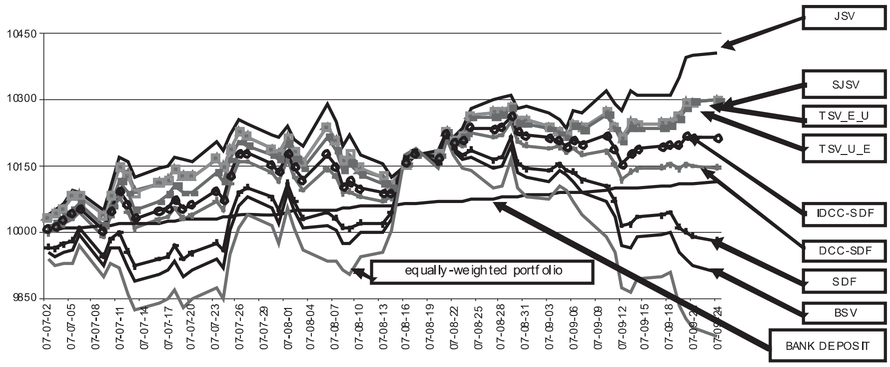

Figure 6. Wealth of the investor at time $T + s$ (the optimal portfolio is constructed on the medians of $w_{MVR^*_p T|T+s}$)

Source: own elaboration.

5. CONCLUSIONS

The paper proposes methods for optimal asset allocation under stochastic volatility. We apply the bivariate stochastic volatility models and the Bayesian approach to portfolio selection problem when the investor minimizes risk. The Bayesian approach leads to the predictive distributions of the returns and the conditional covariance matrix, which were passed on to the optimization procedure – construct optimal portfolio. The predictive distributions of the optimal portfolio are very spread and have heavy tails. But our experience leads us to believe that the VAR(1)-MSV models with time-varying conditional correlations are an appropriate tool for building the multi-period optimal minimum conditional variance portfolio. The multi-period optimal portfolios constructed in the VAR(1)-MSV models with time-varying conditional correlation offer a higher return than the bank deposit.

Cracow University of Economics

REFERENCES

[1] Aguilar O., West M., [2000], *Bayesian dynamic factor models and portfolio allocation*, „Journal of Business and Economic Statistics", Vol. 18, 338-357.
[2] Elton J.E., Gruber M.J., [1991], *Modern portfolio theory and investment analysis*, New York, John Wiley & Sons, Inc.
[3] Engle R.F., [2002], *Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models*, „Journal of Business and Economic Statistics", Vol. 20, 339-350.
[4] Gamerman D., [1997], *Markov Chain Monte Carlo. Stochastic Simulation for Bayesian Inference*, Champan and Hall, London.
[5] Jacquier E., Polson N., Rossi P., [1994], *Bayesian Analysis of Stochastic Volatility Models*, (*with Discussion*), „Journal of Business and Economic Statistics", Vol. 12, 371-417.
[6] Jacquier E., Polson N., Rossi P., [1995], *Model and Prior for Multivariate Stochastic Volatility Models*, technical report, University of Chicago, Graduate School of Business.
[7] Markowitz H.M., [1959], Portfolio Selection: Efficient Diversification of Investments, New York, John Wiely & Sons, Inc.
[8] Osiewalski J., Pajor A., [2007], *Flexibility and parsimony in multivariate financial modelling: a hybrid bivariate DCC–SV model*, [in:] Financial Markets: Principles of Modeling, Forecasting and Decision-Making, FindEcon Monograph Series: Advance in Financial Market Analysis, Number 3, red. W. Milo, P. Wdowiński, Łódź: Łódź University Press, 11-26.
[9] Pajor A., [2003], *Procesy zmienności stochastycznej w Bayesowskiej analizie finansowych szeregów czasowych* (*Stochastic Volatility Processes in Bayesian analysis of financial time series*), Cracow of University of Economics, Cracow.
[10] Pajor A., [2005], *Bayesian Analysis of Stochastic Volatility Model and Portfolio Allocation*, [in:] Issues in Modelling, Forecasting and Decision-Making in Financial Markets, Acta Universitatis Lodzensis – Folia Oeconomica, Vol. 192, 229-249.
[11] Pajor A., [2006], *Modelling the Conditional Covariance Matrix in Stochastic Volatility Models with Applications to the Main Exchange Rates in Poland*, Dynamic Econometric Models, Vol. 7, 170-177.
[12] Polson N.G., Tew B.N., [2000], *Bayesian Portfolio Selection: An Empirical Analysis of the S&P500 Index 1970-1996*, „Journal of Business and Economic Statistics", Vol. 18, No. 2, 164-173.
[13] Quintana J., Lourdes V., Aguilar O., Liu J., [2003], *Global gambling*, [in:] Bernardo J., Bayarri M., Berger J., Dawid A., Heckerman D., Smith A., and West M. (eds.), Bayesian Statistics, Vol. 7, 349-367, Oxford University Press.
[14] Soyer R., Tanyeri K., [2006], *Bayesian portfolio selection with multi-variate random variance models*, „European Journal of Operational Research", Vol. 171, 977-990.
[15] Tsay R.S., [2002], *Analysis of Financial Time Series. Financial Econometrics*, A Wiley-Interscience Publication, John Wiley & Sons, INC.
[16] Winkler R.L., Barry C.B., [1975], *A Bayesian model for portfolio selection and revision*, „Journal of Finance", Vol. 30, No. 1, 175-194.
[17] Zellner A., [1971], *An Introduction to Bayesian Inference in Econometrics*, J. Wiley, New York.

Praca wpłynęła do redakcji we wrześniu 2008 r.

BAYESOWSKI WYBÓR PORTFELA Z WYKORZYSTANIEM MODELI MSV

Streszczenie

W pracy porównano predyktywne własności wielowymiarowych modeli wariancji stochastycznej (ang. Multivariate Stochastic models, MSV) w kontekście budowy optymalnego portfela. Rozważono modele dwuwymiarowe o różnej strukturze macierzy warunkowych kowariancji. Uwzględniono specyfikacje procesu MSV o zerowych, stałych oraz zmiennych warunkowych współczynnikach korelacji. Modele te wykorzystano do konstrukcji wielookresowego portfela o minimalnym ryzyku (gdzie za miarę ryzyka przyjęto warunkową wariancję portfela) oraz portfela o minimalnym ryzyku, ale z zadaną przez inwestora dolną granicą stopy zwrotu. Jako przykład empiryczny przedstawiono sposób konstrukcji portfela walutowego złożonego ze złotowego kursu dolara amerykańskiego i euro.

Uzyskane wyniki wskazują na ogromną przydatność modeli MSV o zmiennych warunkowych korelacjach w wyborze optymalnego portfela walutowego.

Słowa kluczowe: analiza portfelowa, wielowymiarowe procesy wariancji stochastycznej, wnioskowanie bayesowskie

BAYESIAN PORTFOLIO SELECTION WITH MSV MODELS

Summary

In the paper we compare the predictive ability of discrete-time Multivariate Stochastic Volatility (MSV) models to optimal portfolio choice. We consider MSV models, which differ in the structure of the conditional covariance matrix (including the specifications with zero, constant and time-varying conditional correlations). Next, we construct the optimal portfolio under the assumption that the asset returns are described by the multivariate stochastic volatility models. We consider hypothetical portfolios, which consist of two currencies that were the most important for the Polish economy: the US dollar and euro. In the optimization process we use the predictive distributions of future returns and the predictive conditional covariance matrix obtained from the MSV models.

Key words: Portfolio analysis, Multivariate Stochastic Volatility models, Bayesian analysis