Portfolio construction with Bayesian GARCH forecasts

Wolfgang Polasek and Momtchil Pojarliev

Institute of Statistics and Econometrics
University of Basel
Holbeinstrasse 12
CH-4051 Basel
email: Momtchil.Pojarliev@unibas.ch
email: Wolfgang.Polasek@unibas.ch

September 8, 2001

Abstract

Mean-variance portfolio construction depends on forecasts of the variance of the assets in the portfolio. Recent studies have shown that the portfolio weights are very sensitive to changes in the one step ahead variance forecasts. Thus, we propose a Bayesian approach where prior information could be included in the model to reduce the variability of the optimal portfolio weights. We compare the performance of a global regional portfolio (North America, Europe and the Pacific) based on variance forecasts by a Bayesian VAR-GARCH in mean model with the benchmark (MSCI World index).

Keywords: Portfolio construction, multivariate ARCH in mean models, volatility forecasts, global minimum variance portfolio

1 Introduction

Previous research [3], [5] and [9] has found that multivariate time series forecasts can be used for active portfolio management. The information gained by using multivariate models can be successfully transformed into higher portfolio returns by quantitative portfolio strategies when the right combination of volatility modeling and portfolio strategy is found. Polasek and Ren [8] have shown that a multivariate ARCH in mean (ARCH-M) model can provide better forecasts for stock indices than traditional VAR and VAR-GARCH models. Pojarliev and Polasek (PP henceforth) [5] and [6] have shown that optimized global minimum variance (GMV) portfolio weights are very sensitive to changes in the forecasts of the variance-covariance matrix of the assets in the portfolio and that small changes in the forecasts lead to large active position. Thus, we

will investigate a Bayesian approach where prior information can be included in the time series analysis and we use the predictive density of the covariance matrix in order to reduce the variability of the optimal portfolio weights.
The paper is organized as follows: Section 2 describes how the information obtained by the volatility forecasts can be used for portfolio construction. In section 3 we compare our portfolio (which consists of the MSCI North America, MSCI Europe and MSCI Pacific indices) with the benchmark (MSCI World index) where we choose the last 56 months as back-testing horizon (February 1995 until September 1999). In the last section we summarize our results. The multivariate ARCH in mean model, which is used to forecast the variance matrix of the monthly returns of the three MSCI indices from February 1990 until September 1999 is described in the Appendix A.

2 Portfolio construction

Quantitative models of portfolio analysis use statistical and mathematical tools of empirical finance to construct optimal portfolios (see [1]). Beating benchmarks using bottom up strategy is a process divided into two parts: stock selection (the ranking of stocks from best to worst using fundamental analysis) and portfolio construction (computing the weights of the assets in the portfolio). We focus on the second part of this problem: the portfolio selection. Assuming a portfolio which consists of the 3 MSCI indices, we have to compute the weights of the portfolio for each index. The differences between the portfolio weights and the weights in the MSCI World index (the benchmark) are our active portfolio weights. The weights of the *global minimum variance* (GMV) portfolio depends only on the precision matrix $\mathbf{H}_t^{-1}$ (see [1] and [5]). Active portfolio management is based on forecasts. The difference between our forecasts and the consensus will results in the active weights (the active weights are the differences between th benchmark weights and the portfolio weights). Thus, forecasting the variance-covariance matrix, we can compute the optimal weights of a GMV portfolio. We are using the Bayesian VAR(2)-GARCH(1,1)-M(2) model, described in the Appendix A to forecast the variance matrix of the monthly returns of the MSCI North America, MSCI Europe and MSCI Pacific indices from February 1990 until September 1999 (116 observations). As in [9] we assume that a rational investor maximizes expected utility over the predictive distribution of future returns. Let I_t be information observed until time t, then we determine the predictive distribution for time *t+1* which is summarized by the predictive mean $\mu_{t+1} = E(r_{t+1}|I_t)$ and the predictive variance $\hat{\mathbf{H}}_{t+1} = Var(r_{t+1}|I_t)$.
The Bayesian GMV portfolio weights are found from the optimization problem

$$min \mathbf{w}'\mathbf{H}\mathbf{w}$$

subject to

$$\mathbf{w}'\iota = 1,$$

and are computed using the following steps:

1. Using the marginal likelihood criterion for the model selection, we estimate a Bayesian VAR(2)-GARCH(1,1)-M(2) model (see Appendix A) and obtain the MCMC output θ_{tj}, where $j = 1, ..., 105$ and $t = 1, ..., 116$.

2. We compute the 'mean' forecasted variance matrix from the predictive distributions

$$\hat{\mathbf{H}}_{t+1} = \frac{1}{105} \sum \mathbf{H^{(j)}}_{t+1}$$

   where $\mathbf{H}^{(j)}_{t+1}$ is the predicted covariance matrix for time *t+1* according to the variance equation of the VAR-GARCH-M model (see Appendix A) for the j-th sample.

3. We compute the GMV weights vector for the portfolio

$$\mathbf{w}_t = \frac{\hat{\mathbf{H}}^{-1}_{t+1} \iota}{\iota' \hat{\mathbf{H}}_{t+1} \iota}. \quad (1)$$

   where ι is a vector of ones and $\mathbf{w}_t$ is the weights vector.

Figure 1 compares the weights of the Bayesian portfolio and "classical" portfolio computed as in PP [6] from February 1995 until September 1999. The weights for the Pacific index are not shown since the weights sum to 100%. In contrast to the classical portfolio weights, the weights of the Bayesian portfolio are stable. This will reduce the transactions cost.

3 Portfolio comparisons

We perform the back-testing with the estimated weights by comparing the portfolio with the MSCI World index by using the same criteria as PP (2000) [6].

(a) annualized returns (in percent)

(b) annualized standard deviation (in percent)

(c) cumulative return for 56 months, 3 years and year to date (in percent)

(d) *Sharpe* ratio

(e) *success* rate

(f) tracking error

The *Sharpe* ratio is defined as the expected excess return of portfolio P divided by the risk of portfolio P:

$$S_p = \frac{r_p - r_{riskfree}}{\sigma_p},$$

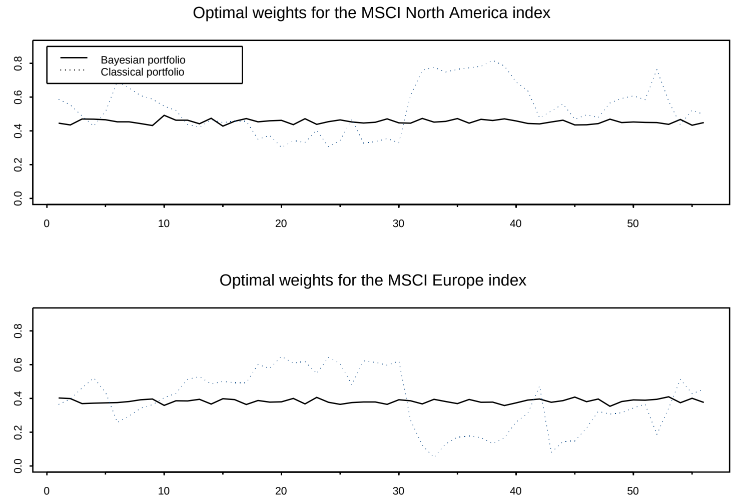

Figure 1: The portfolios weights from February 1995 until September 1999. The weights of the classical portfolio are computed as in Pojarliev and Polasek, (2000) [6].

where the risk free rate $r_{riskfree}$ is assumed to be 3% per year.
The *success* rate is the percentage of times (in months) in which the portfolio returns beat the benchmark returns.
The tracking error T_p is defined as the standard deviation (SD) of the active returns r_a, where the active returns are the difference between the benchmark returns and the portfolio returns, i.e.

$$T_p = \sigma(r_a) \quad with \quad r_a = r_m - r_p$$

with r_m the returns of the MSCI World index and r_p the portfolio returns.
Table 1 shows the performance of the portfolio for the last 56 months. Figure 2 plots the cumulative returns of the Bayesian portfolio and the benchmark for five years (02/1995-09/1999), three years (10/96-09/99) and year to date (01/1999-09/1999).
Note that the portfolio dominates on the benchmark as it has larger mean returns and smaller standard deviation for the whole back-testing period. The portfolio performs better than the MSCI World index for all sub-periods (see figure 2). The *success* rate is larger than 50% and shows that active portfolio management with Bayesian methods can beat the benchmark.
At each forecasting horizon we have simulated the predictive distribution for the covariance matrix. For each simulated variance matrix (sample size of 105) we compute the GMV weights. Figure 3 plots the distribution of the portfolio weights for the three indices for September 1999. Using the weights distribu-

	Portfolio	benchmark
Annualized returns	16.09%	14.95%
Annualized SD	12.67 %	12.91%
Returns		
56 months (02.95-09.99)	75.20%	69.51%
3 years (10.96-09.99)	46.17%	43.69%
year to date	5.84%	5.82%
Sharpe ratio	1.03	0.92
success rate	60.7%	
tracking error	0.35%	

Table 1: Performance of the Bayesian GMV portfolio and the benchmark (MSCI World index) from February 1995 until September 1999. The returns are calculated as $\sum_{t=1}^{n} r_t$, with n equals to the number of months in the evaluation period (56, 36 and 9 respectively). The annualized returns are computed as the mean portfolio returns multiplied by 12, the annualized SD as the standard deviation of the portfolio returns multiplied by $\sqrt{12}$. The *Sharpe* ratio is computed as the ratio of the annualized excess return (the annualized return minus 3%) and the annualized SD. The *success* rate denotes the number of months in percentage in which the portfolio returns are larger than the returns of the benchmark. The tracking error is computed as the annualized SD of the active returns.

tions for the first nine months of 1999 and the returns of the three MSCI indices from December 1998 until August 1999 we plot a simulated cumulative portfolio returns distribution of $N = 500$ for the year to date (01/99-09/99) in figure 4. The benchmark return for this period is 5.82%. According to the simulated distribution the mean returns of our portfolio is 6.9%. We compute the probability to beat the benchmark in this period (67.6%). The range of the distribution is from -0.039% to 15.67%. The maximum can be interpreted as "portfolio potential". Using the simulated returns distribution we compute the 99%-VaR for the 9 months sub-period 1.51%. This shows that with probability of 99%, the cumulative returns of the Bayesian portfolio will be larger than 1.51% in the first 9 months of 1999.

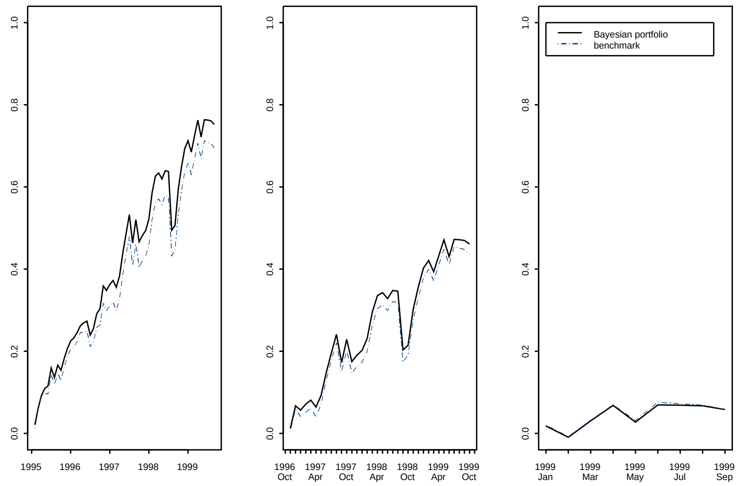

Figure 2: Cumulative returns of the Bayesian portfolio and the benchmark (MSCI World index) for the tree sub-periods (see table 1).

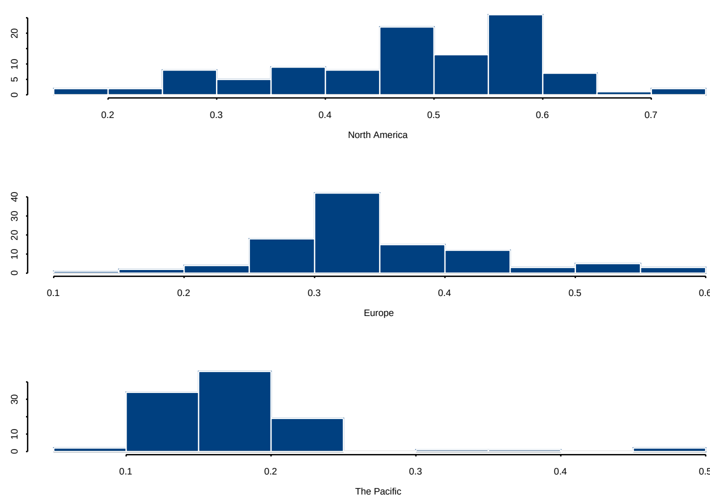

Figure 3: Estimated weights distribution of the portfolio for September 99.

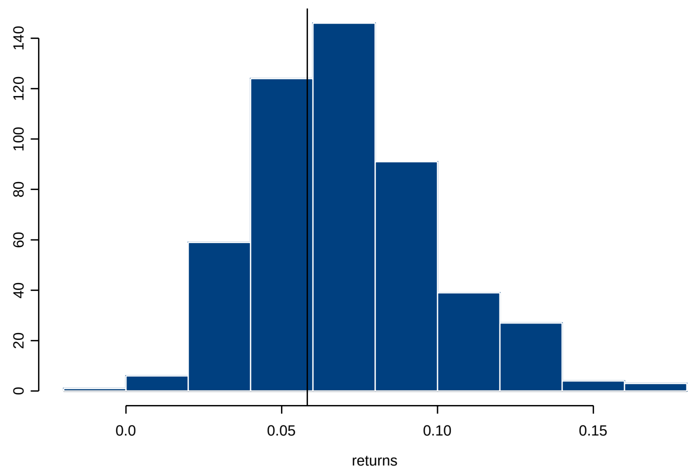

Figure 4: Estimated distribution of the cumulative returns of portfolio 3 from January 99 until September 99. The vertical line indicates the cumulative return of the MSCI World index for this period (5.82%).

4 Conclusion

In this paper we use a Bayesian multivariate ARCH in mean model for forecasting the variance matrix and therefore estimating the optimal weights of portfolio, which consists of the MSCI North America, MSCI Europe and MSCI Pacific indices. Comparing the portfolio performance with the benchmark shows that the information obtained by the multivariate time series forecasts can be used to beat the market. Furthermore the portfolio weights are stable over time in contrast to the "classical" approach for portfolio construction as in PP [6]. Using the Bayesian approach in portfolio decisions can avoid large changes in the weights and save transactions costs.

References

[1] J. V. Campbell, A. W. Lo, and A. C. MacKinlay, *The Econometrics of Financial Markets*, Princeton University Press, 1997.

[2] R. F. Engle, D. M. Lilien, and R. P. Robins, *Estimating Time-Varying Risk Premia in the Term Structure*, Econometrica (1987), 391–407.

[3] A. P. Frost and E. J. Savarino, *An Empirical Bayesian Approach to Efficient Portfolio Selection*, Journal of Financial and Quantitative Analysis (1986), 293–305.

[4] B. R. Litterman, *A Statistical Approach to Economic Forecasting*, Journal of Business and Economics Statistics (1986), no. 4, 1–24.

[5] M. Pojarliev and W. Polasek, *Applying Multivariate Time Series Forecasts for Active Portfolio Management*, University of Basel, http://www.unibas.ch/iso, 2000.

[6] ______, *Portfolio Construction Using Multivariate Time Series Forecasts*, University of Basel, http://www.unibas.ch/iso, 2000.

[7] W. Polasek, *The BASEL Package: User Guide, PC Version 1.0.*, University of Basel, http://www.unibas.ch/iso, 2000.

[8] W. Polasek and L. Ren, *Volatility Analysis during the Asia Crisis: A Multivariate GARCH-M Model for Stock Returns in the US, Germany and Japan*, 2001, pp. 93–107.

[9] N. G. Polson and B. V. Tew, *Bayesian Portfolio Selection: An Empirical Analysis of the SP 500 Index 1970-1996*, Journal of Business & Economic Statistics **18** (1990), 164–173.

A Appendix: The VAR-GARCH-M model

We describe the extension of univariate ARCH-in-mean models to multivariate VAR-GARCH-M processes from a Bayesian point of view [8]. The models are estimated by MCMC methods and model selection is done using the marginal likelihood criterion. For the estimation we have used the BASEL package [7]. To describe the interactions of returns and conditional variances in a VAR model the univariate ARCH-M model of Engle et al. [2] is extended to the multivariate case. Thus, a VAR(k) model of dimension M , i.e. the VAR(k)-GARCH(p,q)-M(r) model, is defined as

$$y_t^l \quad = \quad \beta_0^l + \sum_{m=1}^{M} \sum_{i=1}^{k} \beta_i^{lm} y_{t-i}^m + \sum_{m=1}^{M} \sum_{i=1}^{r} \psi_i^{lm} h_{t-i}^m + u_t^l \tag{2}$$

with heteroskedastic errors $u_t^l \sim N[0, h_t^l], \quad l = 1, \ldots, M$. The conditional variance is parameterized as

$$h_t^l = \alpha_0^l + \sum_{m=1}^{M} (\sum_{i=1}^{p} \alpha_i^{lm} h_{t-i}^m + \sum_{i=1}^{q} \phi_i^{lm} u_{m,t-i}^2), \tag{3}$$

with all coefficients being positive: $\alpha_0^{lm} > 0$, $\alpha_i^{lm} \geq 0$, $\phi_i^{lm} \geq 0$ and $m, l = 1, \ldots, M$. The mean equation (2) can be written as

$$\mathbf{y}_t = \beta_0 + \sum_{i=1}^{k} \beta_i \mathbf{y}_{t-i} + \sum_{i=1}^{r} \mathbf{\Psi}_i vech \mathbf{H}_{t-i} + \mathbf{u}_t = \mu_t + \mathbf{u}_t, \tag{4}$$

where $\mathbf{y}_t = (y_{t1}, \ldots, y_{tM})'$ is an $M \times 1$ vector of observed time series at time t, β_i $(i = 1, \ldots, k)$ and $\mathbf{\Psi}_i$ $(i = 1, \ldots, r)$ are fixed $M \times M$ coefficient matrices, $\beta_0 = (\beta_{10}, \ldots, \beta_{M0})'$ is a fixed $M \times 1$ vector of intercept terms, $\mu_t = (\mu_t^1, \ldots, \mu_t^M)'$ is the $M \times 1$ vector of conditional means and $\mathbf{u}_t = (u_{t1}, \ldots, u_{tM})^T$ is an $M \times 1$ vector of error terms. $\mathbf{H}_t$ is the conditional covariance matrix of the M dimensional observation at time t and $vech\mathbf{H}_t$ is the vectorization of the lower half of the covariance matrix.
The above model is rewritten as a multivariate regression system

$$\mathbf{Y} = \mathbf{BX} + \mathbf{\Psi}\tilde{\mathbf{H}} + \mathbf{U}, \tag{5}$$

with $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_T]_{(M \times T)}$ and $\mathbf{U} = [\mathbf{u}_1, \ldots, \mathbf{u}_T]_{(M \times T)}$, where the coefficient matrices are defined as

$$\mathbf{B} = [\beta_0, \beta_1, \ldots, \beta_k]_{(M \times (\tilde{M}k+1))}, \quad \mathbf{\Psi} = [\mathbf{\Psi}_1, \ldots, \mathbf{\Psi}_r]_{(M \times \tilde{M}r)}.$$

The regressor matrices are partitioned in transposed form as

$$\mathbf{X} = [\mathbf{x}_0, \ldots, \mathbf{x}_{T-1}]_{((1+\tilde{M}k) \times T)}, \quad \tilde{\mathbf{H}} = [\tilde{\mathbf{h}}_0, \ldots, \tilde{\mathbf{h}}_{T-1}]_{(\tilde{M}r \times T)}$$

with the columns defined with $\tilde{M} = M(M+1)/2$ as

$$\mathbf{x}_t = \begin{pmatrix} 1 \\ \mathbf{y}_t \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{y}_{t-k+1} \end{pmatrix}, \quad \tilde{\mathbf{h}}_t = \begin{pmatrix} vech\mathbf{H}_t \\ \cdot \\ \cdot \\ \cdot \\ vech\mathbf{H}_{t-r+1} \end{pmatrix}.$$

We now show that the conditional structure of the proposed VARCH-M model makes the MCMC and the Gibbs sampler convenient to apply in blocks of the parameters.
The Bayesian VAR(k)-GARCH(p,q)-M(r) model is then given by

$$\mathbf{Y} \sim N_{T\times M}[\mathbf{B}\mathbf{X} + \mathbf{\Psi}\tilde{\mathbf{H}}, diag(\mathbf{H}_1, \ldots, \mathbf{H}_T)], \tag{6}$$

$$vech\mathbf{H}_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i vech(u_{t-i}u'_{t-i}) + \sum_{j=1}^{p} \mathbf{\Phi}_j vech\mathbf{H}_{t-j},$$

and the prior distributions are chosen from the families of normal distributions, hence

$$\mathbf{B} \sim N_{M\times(1+\tilde{M}k)}[\mathbf{B}_*, \mathbf{\Sigma}_{B_*} \otimes \mathbf{I}_M]. \tag{7}$$

Stationarity condition is given be Engle and Koenker (1996)

$$det|\sum_{i=1}^{q} \alpha_i + \sum_{i=1}^{p} \phi_i| > 0, \tag{8}$$

$$\mathbf{\Psi} \sim N_{M\times\tilde{M}r}[\mathbf{\Psi}_*, \mathbf{\Sigma}_{\Psi_*} \otimes \mathbf{I}_M],$$

where all of the hyper-parameters (which are denoted with a star) are known a priori. The joint distribution for the data $\mathbf{Y}$ and the parameters $\theta = (\mathbf{B}, \mathbf{\Psi}, \mathbf{A}, \mathbf{\Phi})$ is with $\mathbf{A} = (\alpha_0, \alpha_1, \ldots, \alpha_q)$ and $\mathbf{\Phi} = (\phi_0, \phi_1, \ldots, \phi_p)$

$$\begin{aligned} p(\theta, \mathbf{Y}) &= N[\mathbf{Y}|\mathbf{B}\mathbf{X} + \mathbf{\Psi}\tilde{\mathbf{H}}, diag(\mathbf{H}_1, \ldots, \mathbf{H}_T)] \\ &\cdot N[\mathbf{B}|\mathbf{B}_*, \mathbf{\Sigma}_{B_*} \otimes \mathbf{I}_M] \cdot N[\mathbf{\Psi}|\mathbf{\Psi}_*, \mathbf{\Sigma}_{\Psi_*} \otimes \mathbf{I}_M] \\ &\cdot \prod_{i=0}^{p} N_0^{\infty}[\alpha_i|\alpha_i^*, \mathbf{\Sigma}_{\alpha\otimes\mathbf{I}_{\tilde{M}}}] \cdot \prod_{i=1}^{q} N_0^{\infty}[\mathbf{\Phi}_i|\mathbf{\Phi}_i^*, \mathbf{\Sigma}_{\Phi\otimes\mathbf{I}_{\tilde{M}}}]. \end{aligned} \tag{9}$$

As prior distribution for the GARCH coefficients we use the positive truncated normal distribution (N_0^{∞}) since the variance components of the GARCH equation showed be positive. For the VAR regression coefficients we use the "tightness prior" of Litterman (1986) since the GARCH coefficients have to be positive for the prior means we assume $\alpha_* = 0.01\mathbf{I}_{1+p+q}$ and for the VAR coefficients $B_* = \mathbf{0}$, $\Psi_* = \mathbf{0}$ and for the prior precision matrices we assume the following diagonal tightness structure $\Sigma_{\mathbf{B}*}^{-1} = diag(\varepsilon, 1, \ldots, k)$, $\Sigma_{\mathbf{\Psi}*}^{-1} = diag(1, \ldots, r)$, $\mathbf{\Sigma}_{\alpha i}^{-1} = i\mathbf{I}_{\tilde{M}}$, $\mathbf{\Sigma}_{\phi i}^{-1} = i\mathbf{I}_{\tilde{M}}$, and for the inverse variance of the

intercepts we choose ε to be a small number like 10^{-6}.
Using the Marginal Likelihood (marg. lik.) as criteria, we estimate a multivariate VAR(2)-GARCH(1,1)-M(2) for the monthly returns vector of the MSCI North America, MSCI Europe and MSCI Pacific indices from February 1995 until September 1999.
Thus the forecasts of the variance matrix is obtained from

$$vech\mathbf{H}_{t+1} = \alpha_0 + \sum_{i=1}^{q} \alpha_i vech(\mathbf{u}_{t-i+1}\mathbf{u'}_{t-i+1}) + \sum_{j=1}^{p} \mathbf{\Phi}_j vech\mathbf{H}_{t-j+1} \tag{10}$$

where E_t is the conditional expectations operator

$$\mathbf{y}_{t+1} = \beta_0 + \sum \beta_i E_t \mathbf{y}_{t-i+1} + \sum \Psi vech\mathbf{H}_{t-1+1} \tag{11}$$

	marg. lik.
VAR(1)-GARCH(1,1)-M(1)	-314.99
VAR(1)-GARCH(2,1)-M(1)	-314.98
VAR(2)–GARCH(1,1)-M(1)	-312.22
VAR(2)–GARCH(1,1)-M(2)	-312.20^{*}
VAR(1)–GARCH(2,1)-M(2)	-312.21
VAR(2)-GARCH(2,2)-M(2)	-312.21

Table 2: Marginal likelihoods for different orders of the model. The star (*) indicates the maximum value.