

A multivariate GARCH-M model for exchange rates in
the US, Germany and Japan

Wolfgang Polasek and Lei Ren

Institute of Statistics and Econometrics
University of Basel
Holbeinstrasse 12, 4051 Basel, Switzerland
Email: Wolfgang@iso.iso.unibas.ch
October 15, 1999

Abstract

After the so-called Asia crisis in the summer of 1997 the financial markets were shaken by increased volatility transmission around the world. Therefore, in this paper we will analyse the daily exchange rates in New York, Germany, and Japan for the period of 2 years (June 21, 1996 to June 22, 1998). We estimate a VAR-GARCH in mean model and estimate the multivariate volatility effects between the time series. We are also interested in the question of whether or not the volatility of the 3 exc

1 Introduction

The analysis of the international financial system and the connection of markets have become a major topic in financial econometrics in recent years. The availability of daily data and the connectedness of financial markets have inspired analysis of the transmission mechanism in terms of mean and variances between time series (see ergo Engle et al. 1990). In the present paper we want to examine the exchange rates of Japan, Germany and the US. We estimate a multivariate VAR-GARCH-M (short: VARCH in mean) model to explore the relationships in the returns and the conditional variances (also called volatilities). The ARCH in mean structure is used to find out if the volatilities feed back into the mean equations. Since conditional variances can be interpreted as temporary increases or decreases in uncertainty, it would be not surprising to find that daily returns react to the changes in uncertainties in different international markets.

For the estimation approach we have chosen a Bayesian MCMC (Markov Chain Monte Carlo) method (see Gelfand and Smith 1990) since reliable methods for the likelihood estimation of the VARCH-M model seem to be difficult to obtain in closed form (see Liu and Polasek 1999). Furthermore, the MCMC approach allows us to introduce new concepts and to find exact (small sample) results for characteristics of the dynamic process, like the impulse response function or the predictive distributions.

In section 2 we introduce the basic VAR-GARCH-M model and in section 3 we present the estimation results. We show how the Gibbs sampler and the Metropolis step for the ARCH parameters are implemented in the simulation using the full conditional distributions. The program is available in BASEL package (see Polasek 1998). The lag orders of the model are estimated by the posterior marginal likelihood or the pseudo marginal likelihood criterion, since the method by Chib (1995) is not efficiently applicable.

The estimated model shows that there is a rich interaction pattern between the coefficients of the mean equation and the volatility equations. The ARCH-M coefficients exhibit a substantial reaction to volatilities and all the impulse response functions have a quick decay. The predictive distributions are compared to the usual VAR approach and they show considerable improvements. In a final section we conclude our approach.

2 Estimation and model selection

The modeling of financial time series has been enriched by the class of ARCH-M processes which were introduced by Engle, Lilien and Robins (1987). But with more processes to choose from, like stochastic volatility models or heavy tail distributions, model choice has become a more complicated problem. In this paper we suggest using the posterior and pseudo marginal likelihoods as model choice criteria for multivariate time series models.

The following section describes the VAR-GARCH-M processes from a Bayesian point of view (see also Pelloni and Polasek 1998).

2.1 The VAR-GARCH-M model

To describe the interactions of returns and conditional variances in a VAR model we extend the univariate ARCH-M model of Engle et al. (1987) to the multivariate case. Thus, we define a VAR(k) model of dimension M , i.e., the VAR(k)-GARCH(p,q)-M(r) model, in the following way:

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

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

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

where the parameters for each l are satisfying the stationarity condition

$$\sum_{m=1}^M \left(\sum_{i=1}^p \alpha_i^{lm} + \sum_{i=1}^q \phi_i^{lm} \right) < 1, \quad (3)$$

with all coefficients being positive: $\alpha_0^{lm} > 0$, $\alpha_i^{lm} \geq 0$, $\phi_i^{lm} \geq 0$ and $m, l = 1, \dots, M$.

The equation (1) can be written as

$$\mathbf{y}_t = \boldsymbol{\beta}_0 + \sum_{i=1}^k \boldsymbol{\beta}_i \mathbf{y}_{t-i} + \sum_{i=1}^r \boldsymbol{\Psi}_i \text{vec} h_{t-i} + \mathbf{u}_t = \boldsymbol{\mu}_t + \mathbf{u}_t, \quad (4)$$

where $\mathbf{y}_t = (y_{t1}, \dots, y_{tM})'$ is an $M \times 1$ vector of observed time series at time t , β_i ($i = 1, \dots, k$) and Ψ_i ($i = 1, \dots, r$) are fixed $M \times M$ coefficient matrices, $\beta_0 = (\beta_{10}, \dots, \beta_{M0})'$ is a fixed $M \times 1$ vector of intercept terms, $\boldsymbol{\mu}_t = (\mu_t^1, \dots, \mu_t^M)'$ is the $M \times 1$ vector of conditional means and $\mathbf{u}_t = (u_{t1}, \dots, u_{tM})^T$ is an $M \times 1$ vector of error terms.

The above model is rewritten as a multivariate regression system

$$\mathbf{Y} = \mathbf{B}\mathbf{X} + \Psi\tilde{\mathbf{H}} + \mathbf{U}, \quad (5)$$

where the coefficient matrices are defined as

$$\mathbf{B} = [\beta_0, \beta_1, \dots, \beta_k]_{(M \times (\tilde{M}k+1))}, \quad \Psi = [\Psi_1, \dots, \Psi_r]_{(M \times \tilde{M}r)}.$$

The regressor matrices are

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

with

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

and $\tilde{M} = M(M+1)/2$. 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

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

$$vech\mathbf{H}_t = \boldsymbol{\alpha}_0 + \sum_{i=1}^q \boldsymbol{\alpha}_i vech(\mathbf{u}_{t-i}\mathbf{u}'_{t-i}) + \sum_{j=1}^p \Phi_j vech\mathbf{H}_{t-j},$$

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

$$\begin{aligned} \mathbf{B} &\sim N_{M \times (1+\tilde{M}k)}[\mathbf{B}_*, \Sigma_{B_*} \otimes \mathbf{I}_M], \\ \Psi &\sim N_{M \times \tilde{M}r}[\Psi_*, \Sigma_{\Psi_*} \otimes \mathbf{I}_M], \end{aligned} \quad (7)$$

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 $\mathbf{A} = (\boldsymbol{\alpha}_0, \boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_q)$ and $\mathbf{\Phi} = (\phi_0, \phi_1, \dots, \phi_p)$

$$\begin{aligned}
p(\theta, \mathbf{Y}) &= N[\mathbf{Y} | \mathbf{B}\mathbf{X} + \mathbf{\Psi}\tilde{\mathbf{H}}, \text{diag}(\mathbf{H}_1, \dots, \mathbf{H}_T)] \\
&\cdot N[\mathbf{B} | \mathbf{B}_*, \boldsymbol{\Sigma}_{B_*} \otimes \mathbf{I}_M] \cdot N[\mathbf{\Psi} | \mathbf{\Psi}_*, \boldsymbol{\Sigma}_{\Psi_*} \otimes \mathbf{I}_M] \\
&\cdot \prod_{i=0}^p N[\boldsymbol{\alpha}_i | \boldsymbol{\alpha}_i^*, \boldsymbol{\Sigma}_{\boldsymbol{\alpha}_i}] \cdot \prod_{i=1}^q N[\boldsymbol{\Phi}_i | \boldsymbol{\Phi}_i^*, \boldsymbol{\Sigma}_{\boldsymbol{\Phi}_i}]. \tag{8}
\end{aligned}$$

Thus, we can derive the following full conditional distributions (f.c.d.) for the MCMC simulation process.

2.2 The full conditional distributions (f.c.d.)

a) The f.c.d. for the regression coefficients \mathbf{B}

The full conditional density can be written as

$$p(\mathbf{B} | \mathbf{Y}, \theta^c) = N_{M \times (1 + \tilde{M}k)}[\mathbf{B}_{**}, \mathbf{D}_{B_{**}}], \tag{9}$$

$$\mathbf{D}_{B_{**}}^{-1} = \mathbf{I}_M \otimes \boldsymbol{\Sigma}_{B_*}^{-1} + \langle \mathbf{x}'_t \mathbf{H}_t^{-1} \mathbf{x}_t \rangle,$$

$$\mathbf{B}_{**} = \mathbf{D}_{B_{**}} [\text{vec}(\boldsymbol{\Sigma}_{B_*} \mathbf{B}_* + \langle \mathbf{x}'_t \mathbf{H}_t^{-1} \tilde{\mathbf{y}}_t \rangle)],$$

where $\tilde{\mathbf{Y}} = \mathbf{Y} - \mathbf{\Psi}\tilde{\mathbf{H}}$ and $\theta^c = (\mathbf{\Psi}, \mathbf{A}, \mathbf{\Phi})$ denotes a vector of all parameters save the arguments of the full conditional distribution.

b) The f.c.d. for the regression coefficients $\mathbf{\Psi}$

The f.c.d. is given by

$$p(\mathbf{\Psi} | \mathbf{Y}, \theta^c) = N_{M \times \tilde{M}r}[\mathbf{\Psi}_{**}, \mathbf{D}_{\Psi_{**}}] \tag{10}$$

with

$$\mathbf{D}_{\Psi_{**}}^{-1} = \mathbf{I}_M \otimes \boldsymbol{\Sigma}_{\Psi_*}^{-1} + \langle \mathbf{x}'_t \mathbf{H}_t^{-1} \mathbf{x}_t \rangle,$$

$$\mathbf{\Psi}_{**} = \mathbf{D}_{\Psi_{**}} [\text{vec}(\boldsymbol{\Sigma}_{\Psi_*} \mathbf{\Psi}_* + \langle \mathbf{x}'_t \mathbf{H}_t^{-1} \tilde{\mathbf{y}}_t \rangle)]$$

and $\tilde{\mathbf{Y}} = \mathbf{Y} - \mathbf{B}\mathbf{X}$.

c) The f.c.d. for the GARCH coefficients

For the f.c.d. of α_i and ϕ_i we use the Metropolis-within-Gibbs step with a normal distribution which is obtained by an iteration proposal given by

$$vec\alpha_i \sim N[vec\hat{\alpha}_i, \hat{\Sigma}\alpha_i],$$

$$vec\Phi_i \sim N[vec\hat{\Phi}_i, \hat{\Sigma}\Phi_i],$$

and the f.c.d. is given by

$$p(\alpha, \Phi | \mathbf{Y}, \theta^c) = \prod_{t=1}^T N[\mathbf{y}_t | \boldsymbol{\mu}_t, \mathbf{H}_t] \quad (11)$$

with $\boldsymbol{\mu}_t$ given in (4) and the normal distribution being proportional to

$$N[\mathbf{y}_t | \boldsymbol{\mu}_t, \mathbf{H}_t] \propto |\mathbf{H}_t|^{-1/2} exp\left\{-\frac{1}{2}(\mathbf{y}_t - \boldsymbol{\mu}_t)' \mathbf{H}_t^{-1} (\mathbf{y}_t - \boldsymbol{\mu}_t)\right\}.$$

Note: If $\mathbf{H} = diag(\mathbf{H}_1, \dots, \mathbf{H}_T)$ is a $TM \times TM$, \mathbf{W} a $r \times T$, and \mathbf{V} a $T \times k$ matrix, then we define the special matrix

$$\begin{aligned} \langle \mathbf{w}_t \mathbf{H}_t \mathbf{v}_t \rangle_{rM \times kM} &= (\mathbf{W} \otimes \mathbf{I}_M) diag(\mathbf{H}_1, \dots, \mathbf{H}_T) (\mathbf{V} \otimes \mathbf{I}_M) \\ &= \begin{pmatrix} \sum_t \mathbf{w}_{1t} \mathbf{H}_t \mathbf{v}_{t1}, \dots, \sum_t \mathbf{w}_{1t} \mathbf{H}_t \mathbf{v}_{tk} \\ \dots \\ \sum_t \mathbf{w}_{rt} \mathbf{H}_t \mathbf{v}_{t1}, \dots, \sum_t \mathbf{w}_{rt} \mathbf{H}_t \mathbf{v}_{tk} \end{pmatrix}. \end{aligned}$$

2.3 Model selection

For the order selection of the VAR-GARCH-M model we suggest using the posterior marginal likelihood of Aitkin (1991) evaluated numerically with the MCMC output and the pseudo marginal likelihood criterion for computational reasons as in table 1.

Following Gelfand and Dey (1994) we define the pseudo marginal likelihood function via the conditional predictive ordinates (CPO's)

$$PsML = \prod_{t=1}^T CPO_t = \prod_{t=1}^T f(z_t | Z_{[t]})$$

with $Z_{[t]} = (y_{1-p}, \dots, y_t, \varepsilon_{1-q}, \dots, \varepsilon_t)$ the conditioning set and where the CPO's can be calculated from the MCMC output $\{\Theta_j, j = 1, \dots, L\}$ by

$$f(z_t|Z_{[t-1]}) \simeq \frac{\sum_{j=1}^L [\prod_{k=t+1}^T f(z_k|Z_{[k-1]}, \Theta_j)]^{-1}}{\sum_{j=1}^L [\prod_{k=t}^T f(z_k|Z_{[k-1]}, \Theta_j)]^{-1}} \quad (12)$$

with

$$f(z_t|Z_{[t-1]}, \Theta_j) = N[z_t|0, H_t^{(j)}], \quad t = 1, \dots, T.$$

3 Estimation results

In this section we present model estimates for our three dimensional VAR system. We have analyzed the exchange rates from 3 countries: The US\$/DM, US\$/Yen and DM/Yen from June 21, 1996 until June 22, 1998 (see figure 1). The exchange rates are transformed to returns (first differences of the logarithms) and are perfectly collinear. Nevertheless, if contemporaneous regressors are excluded, a 3-dimensional VAR-GARCH-M can be estimated, since the conditional covariance matrix is also estimated as a parameterization of past observations and is not freely estimated. Also, we use informative prior information of a rather simple form: $\mathbf{B}_* = \mathbf{0}$, $\mathbf{\Psi}_* = 0.01\mathbf{E}_{M \times \tilde{M}_r}$, $\mathbf{A}_* = 0.01\mathbf{E}_{M \times (1 + \tilde{M}_p)}$, $\mathbf{\Phi}_* = 0.01\mathbf{E}_{M \times \tilde{M}_q}$, $\mathbf{\Sigma}_{\mathbf{B}_*} = 0.01\mathbf{I}$, $\mathbf{\Sigma}_{\mathbf{\Psi}_*} = 0.01\mathbf{I}$, $\mathbf{\Sigma}_{\mathbf{A}_*} = 0.01\mathbf{I}$, $\mathbf{\Sigma}_{\mathbf{\Phi}_*} = 0.01\mathbf{I}$, where \mathbf{I} is the identity matrix and \mathbf{E} is a matrix of ones.

The model has been estimated using the Gibbs-Metropolis algorithm for a Bayesian vector ARCH model. The model selection by two marginal likelihood criteria in table 1 gives the same result. The stationarity of the returns was checked by classical unit root tests and the Bayesian method as in Polasek and Ren (1997).

1) US\$/DM exchange rate, first differences of logs (posterior standard deviations are given in parentheses)

$$\begin{array}{rcccccc} US\$/DM_t & = & 0.079 + & 0.270 & U/D_{-1} - & 1.4067 & U/Y_{-1} + & 0.047 & D/Y_{-1} - \\ & & (.215) & (.184) & & (1.070) & & (.217) & \\ & & & -1.714 & h^{U/D} - & 0.0608 & h^{U/Y} - & 1.040 & h^{D/Y} \\ & & & (1.225) & & (0.168) & & (1.126) & \end{array}$$

2) US\$/Yen exchange rate, first differences of logs

$$\begin{aligned}
 US\$/Yen_t = & -0.015 + & 0.811 & U/D_{-1} - & 0.007 & U/Y_{-1} - & 0.012 & D/Y_{-1} - \\
 & (.165) & (.607) & & (.190) & & (.117) & \\
 & & -0.269 & h^{U/D} + & 0.532 & h^{U/Y} + & 0.532 & h^{D/Y} \\
 & & (.132) & & (.259) & & (.259) &
 \end{aligned}$$

3) DM/Yen exchange rate, first differences of logs

$$\begin{aligned}
 DM/Yen_t = & 0.072 - & 0.404 & U/D_{-1} - & 0.030 & U/Y_{-1} + & 0.440 & D/Y_{-1} - \\
 & (.192) & (.300) & & (.194) & & (.385) & \\
 & & -0.067 & h^{U/D} + & 0.039 & h^{U/Y} + & 0.750 & h^{D/Y} \\
 & & (.210) & & (.187) & & (.625) &
 \end{aligned}$$

The first equation shows that the returns of the US\$/DM exchange rate are negatively influenced by their own lags and positively influenced by the DM/Yen returns of the last period.

The US\$/Yen returns are negatively influenced by their own returns from the previous period. For the DM/Yen returns we find that the coefficients of all other past returns have posterior t values which are smaller than 1, but there is a negative GARCH-M effect of the US\$/DM volatility of the previous period. The t ratio is close to 2 and the negative coefficient implies that a higher conditional variance (i.e. uncertainty) in the US\$/DM returns has a dampening effect on DM/Yen returns.

This interpretation of the lag 1 coefficient structure in the 3-dimensional exchange rate returns model shows a simple pattern of connections. This rather simple pattern is also present in simple decays of the impulse response functions (see Hamilton 1994). While the US\$/Yen returns are quite insensitive to shocks in their own or other past returns, the impulse response functions of the US\$/DM returns and the DM/Yen returns react with opposite sign effects: a positive shock in the US\$/DM returns leads to a negative response on DM/Yen returns and vice versa. Only the decay rate for shocks is different. Shocks in the US\$/DM and DM/Yen returns culminate on the second day and die out more slowly.

4 Conclusions

We have estimated a 3-dimensional model for exchange rates in the US, Germany and Japan. We found that a multivariate ARCH-M model is better than traditional VAR and VAR-GARCH models. We suggest estimating the model by MCMC models and comparing them by the posterior or pseudo marginal likelihood criterion. Both criteria point to the same model as the best: a lag 1 order for the autoregressive component, a lag 1 order for the ARCH-M component and a VARCH(2,2) model for the conditional variances. The one step ahead prediction of the VAR-GARCH-M model shows a smaller variance than the VAR predictions (see table 2) and is also better in terms of the MSE (see Polasek 1999). In summary we conclude that simple extensions of the traditional VAR model, which incorporates volatilities as explanatory variables, seem to play an important role in financial econometric models. One hypothesis for this phenomenon is that volatilities play an important role in the specification and formulation of uncertainty for exchange rates on a daily basis. In a further paper we report an extension of these results on stock market returns (see Polasek and Ren 1999).

5 References:

- Aitkin, M. (1991): Posterior Bayes Factors. *J.R.Statistical Soc. B*, 1, 111-142.
- Chib, S. (1995): Marginal likelihood from the Gibbs output. *JASA* 90, 1313-1321.
- Engle, R. F. (1995): ARCH, Selected Readings. *Oxford University Press, Oxford*.
- Engle, R. F., Lilien, D. M. and Robins, R. P. (1987): Estimating Time-Varying Risk Premia in the Term Structure, *Econometrica* 55/2 391-407, reprinted in *Engle (1995)* 24-41.
- Engle, R. F., Ito, T. and Lin, W. L. (1990): Meteor Showers or Heat Waves? Heteroskedastic Intra-Daily Volatility in the Foreign Exchange Market. *Econometrica* 58/3, 525-542.
- Gelfand, A.E., and Smith, A.F.M. (1990): Sampling based approaches to calculating marginal densities. *Journal of American Statistical Association*, 85, 398-409.
- Gelfand, A.E. and Dey, D.K. (1994): Bayesian model choice asymptotics and exact calculations. *JRSSB*, 56, 501-514
- Hamilton, J. (1994): Time Series Analysis, Princeton Univ. Press, New York.

k	r	p	q	posterior marginal likelihood	pseudo marginal likelihood
1	1	1	1	-4011.5323	-13.3142
2	1	1	1	-3827.6512	-13.6476
1	2	1	1	-3686.5379	-13.5332
1	1	2	1	-4275.6836	-13.9061
1	1	1	2	-3490.5725	-13.1242
1	1	2	2	-3153.7863*	-13.1011**
2	1	2	2	-3647.5625	-13.4572
2	1	1	1	-3809.1132	-13.7971
2	1	2	1	-4011.5239	-14.0116
2	2	2	2	-4152.6732	-14.5109

Table 1: The marginal likelihood for the VAR(k)-GARCH(p,q)-M(r) model (for $y_t^1 = \text{US\$/DM}$, $y_t^2 = \text{US\$/Yen}$, $y_t^3 = \text{DM/Yen}$)

Liu, S. and Polasek, W. (1999): Maximum likelihood estimation based on the method of scoring for a vector autoregressive ARCH-M model. *ISO-WWZ, University of Basel, mimeo.*

Pelloni, G. and Polasek, W. (1998): Intersectoral Labour Reallocation and Employment Volatility: A Bayesian analysis using a VAR-GARCH-M model. *ISO-WWZ, University of Basel, mimeo.*

Polasek, W. et al. (1998): The BASEL package, *ISO-WWZ, University of Basel. (www.unibas.ch/iso)*

Polasek, W. (1999): Forecast evaluations for multiple time series: A generalized Theil decomposition. *ISO-WWZ, University of Basel, mimeo. (www.unibas.ch/iso)*

Polasek, W., and Ren, L. (1997): Structural breaks and model selection with marginal likelihoods. *Proceedings of the Workshop on Model Selection, in Racugno W. (ed.), Cagliari 223-274.*

Polasek, W., and Ren, L. (1999): Multivariate GARCH-M model for stock returns in the US, Germany and Japan. *ISO-WWZ, University of Basel, mimeo.*

