Available online at www.sciencedirect.com
ScienceDirect
Journal of Empirical Finance 15 (2008) 64 – 79
Journal of EMPIRICAL FINANCE
www.elsevier.com/locate/jempfin

Volatility of stock price as predicted by patent data: An MGARCH perspective☆

William W. Chow [a], Michael K. Fung [b,*]

[a] Hong Kong University of Science and Technology, China
[b] Hong Kong Polytechnic University, China

Accepted 25 October 2006
Available online 21 May 2007

Abstract

This paper proposes to model stock price volatility and variations in innovation effort using a Multivariate GARCH structure designed to extract information for risk prediction. The salient feature is that the model order, alongside other parameters, is endogenously determined by the estimation procedures. Using stock prices of U.S. computer firms, it is found that the model can pick up the correlation between the two variables and aid in producing accurate Value-at-Risk estimates.

JEL classification: C110; C320; G100; O320
Keywords: Multivariate GARCH; Reversible jump MCMC; Innovation, Patents; Value-at-Risk

1. Introduction

Publicly traded corporations are bundles of assets, both tangible and intangible, whose values are determined everyday in financial markets. Under the efficient market hypothesis, the stock market valuation of a firm can efficiently capitalize the expected future cash flow generated by currently held assets. Generally speaking, both tangible and intangible assets should increase market values. In past literature, innovation is commonly viewed as an investment that can increase the firm's intangible assets. Griliches (1981) uses a panel of 157 U.S. firms and finds a positive association between patenting activity and market value. A series of subsequent studies show similar findings, see, for example, the survey by Hall (1999).

Dukes (1976) finds that investors adjust reported earnings for the full expensing of research and development (R&D). Similarly, Ben Zion's (1978) results suggest that differences between market and book values are cross-

☆ The work described in this paper was substantially supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region (Project No. G-U129). The authors would also like to thank Francis T. Lui of HKUST and the referee for valuable comments that helped improve the article. We are indebted to Miles Spink of the Hong Kong Polytechnic University for reviewing and editing the article. All errors are ours.
* Corresponding author. School of Accounting and Finance, Hong Kong, Polytechnic University, Hung Hom, Kowloon, the Hong Kong SAR, China. Tel.: +852 2766 7102; fax: +852 2653 3947.
E-mail address: afmikef@polyu.edu.hk (M.K. Fung).

0927-5398/$ - see front matter
doi:10.1016/j.jempfin.2006.10.003

sectionally correlated with R&D expenditures. Hirschey and Weygandt (1985) and Shane and Klock (1997) observe that Tobin's Q is cross-sectionally correlated with R&D intensity. Hirschey and Spencer (1992), and Chauvin and Hirschey (1993, 1994) find that the coefficients for R&D expenses in cross-sectional valuation models are significantly positive. Using an event methodology, Woolridge (1988) and Chan et al. (1990) demonstrate a positive investor reaction to firms' R&D announcements. Sougiannis (1994) finds that lagged and current R&D expenses are both valued by the market. Green et al. (1996), and Stark and Thomas (1998) find that R&D expenses could improve the ability to explain market values of U.K. companies, and hence, there is a clear R&D component in market value.

This paper offers a methodological framework to extract information intertwined in stock prices and innovation activities of firms and suggests ways to make practical use of the modelling results. Specifically, we propose modelling the variations of stock prices and those of innovation activities using a Multivariate (bivariate, to be precise) GARCH model. Because of the dimensions of variables we try to model, a univariate GARCH specification[1] would not be applicable. We are not interested in the relationship between levels of stock prices and patents granted *per se*, which many of the works quoted above have looked into, but aim more at devising a model that can extract useful information with due predictive power. From a practical perspective, the validity of any reduced form model would be deemed refutable if the estimated results could not produce good prediction performance. In addition, just as other econometric models facilitate forecasting and policy analysis, an empirical financial model like the one proposed here should be accessible to practitioners in the financial industry.

A reduced form model in nature, the suggested structure is not without theoretical underpinnings. Indeed, the relationship between R&D and stock volatility is well studied in past literature. The prospects of many R&D-intensive companies, particularly those that have few tangible assets, are tied to the success of new, untested technologies and hence are highly unpredictable. Fixed expenditures on R&D are usually required at the outset to develop a new technology into a new product, but the outcome is far from certain. The return to R&D, if any, is likely to take a long time to materialize, but the life-cycle of the resulting product may be short in the face of fast-changing technology. A study by Schwert (2002) shows that technology can explain the unusual volatility of the Nasdaq portfolio (relative to the S&P portfolio) since mid-1999, but firm size and immaturity of firms cannot. Hollifield (2002) provides an explanation for this phenomenon: technologies have the flavor of a real option, and the volatility of an option can be much higher than the volatility of the underlying assets because the former represents a highly leveraged position.

Weeds (2002) state that R&D-related risks take two distinct forms: the technological success of a research project is probabilistic and the economic value of an innovation may evolve stochastically over time. Since the costs of inventing a new production process or developing a new product are fixed, these fixed commitments together with uncertain revenues increase the risk of the cash flow to equity. Therefore, the increase in operating leverage resulting from R&D activities may increase the volatility of a firm's performance, a claim that is supported by the empirical findings of Chan et al. (2001).

Asymmetric information on the potential costs and benefits of R&D may also produce a positive relationship between R&D and stock volatility. Under the overreaction hypothesis, [see for example Larson and Madura, 2003; Veronesi, 1999; Jegadeesh and Titman, 1995; Bloom and VanReenen, 2002], uninformed investors tend to overreact to the announcements of technological innovations or new products, and this overreaction will later be corrected by an opposite movement. Implied in this uncertain ultimate cash flow story is that investors will try to infer the distribution of these flows over time, and models that incorporate learning into asset pricing typically show that such a learning process tends to increase the volatility of asset returns [see Brennan and Xia, 2001; Timmerman, 1993; Veronesi, 1999, 2000]. Eden and Jovanovic's (1994) research is in a similar vein. Under standard hidden information settings, as in Akerlof's "lemons" problem, they show that when information is less precise, uninformed investors will value "good" companies higher than they would in the perfectly informed case and thus generate more volatility in the stock prices concerned. This offers a testable proposition in our context if we consider innovative efforts of firms can deliver information regarding their future valuations.

Apart from discussing estimation and testing issues, we also validate the informativeness of the model by conducting a Value-at-Risk (VaR) analysis using the estimates and see if innovations can improve efficiency in risk management. The empirical findings are then compared to those obtained by benchmark methods commonly used in the financial industry. Modelling and estimation are Bayesian in nature. A salient feature is that the model order (the

[1] The acronym GARCH will hereafter refer to a univariate GARCH process, and MGARCH represents the multivariate counterpart.

dimensions of the ARCH and GARCH parts) is endogenously determined using the Reversible Jump Markov Chain Monte Carlo (RJMCMC) algorithm. Bayesian analysis of GARCH type models has flourished for about the last decade, thanks to the propagation of the Markov Chain Monte Carlo (MCMC) method. Early works like Polasek and Jin (1997) and Bauwens and Lubrano (1998) focus on estimation of univariate GARCH models. In recent years, research coverage has extended to multivariate models; see for instance Vrontos, Dellaportas and Politis (2003a,b). The use of the RJMCMC algorithm in studying GARCH type models is not unprecedented, although application in multivariate models is rare. Vrontos et al. (2000) adopt it to evaluate the model selection problem of two distinct univariate models — GARCH and exponential GARCH.

To quantify innovation efforts of firms, we use patent data available in the literature so as to avoid making subjective and *ad hoc* definitions on the matter. Not only can patents reflect the intensity of innovation activities, they also represent acknowledgments by specialists in the fields under some standard review processes. While patent citations may be a natural choice, we realize that unless we can perfectly get around the truncation problem or the cohort effect the result may not be at all palatable. Instead, we proxy innovation effort by the gross number of patents granted by the USPTO. The choice is justified by the fact that these figures are actual and complete and indicate those research inputs recognized by the USPTO. The relevant data was obtained from the NBER Patent Citations Data File; see Hall et al. (2001) for details. Stock price data are from Datastream.

Section 2 describes the model setting and specifications of the prior distributions of parameters. This is followed by an exposition of the sampling algorithm. Estimation and model checking results are stated in Section 4. Section 5 compares the performance of the model in generating VaR estimates and those obtained using other benchmark methods. The summary of this article can be found in Section 6.

2. The model structure

2.1. Formulation

The variables of interest are modelled as a vector process containing a deterministic part and stochastic part. Specifically, the weekly vector time series y_t follows

$$y_t = \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \mu_t + \epsilon_t \tag{1}$$

where y_{1t}=rate of change in stock price at time between $t-1$ and t, and y_{2t}=rate of change in new patents granted between $t-1$ and t, for $t=1, \cdots, T$. μ_t is the deterministic vector specifying the unconditional means of y_t. The MGARCH element of the model concerns the stochastic part and can be illustrated by the following:

$$\epsilon_t = H_t^{1/2} z_t \tag{2}$$

where z_t has mean $E(z_t) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = 0_{2\times 1}$ and variance var(z_t)=I_2 being a 2×2 identity matrix. Thus, given past information available up to time $t-1$,

$$E(y_t|\mathcal{F}_{t-1}) = \mu_t,$$
$$\text{var}(y_t|\mathcal{F}_{t-1}) = \text{var}(\epsilon_t|\mathcal{F}_{t-1}) = H_t.$$

While most studies measuring stock prices will assume μ_t to be a zero vector, we leave it to the model to identify the values by writing $\mu_t = B\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ where B is a coefficient matrix. The model can be made appropriately parsimonious by restricting B to a 2×2 diagonal matrix

$$B = \begin{bmatrix} \beta_1 & 0 \\ 0 & \beta_2 \end{bmatrix}.$$

As for the conditional heteroskedastic elements, the BEKK representation[2] is used in this article whereby the conditional covariance matrix H_t can be defined as:

$$H_t = C^{*\prime}C^* + \sum_{k=1}^{K}\sum_{j=1}^{p} A_{jk}^{*\prime}\epsilon_{t-j}\epsilon_{t-j}^{\prime}A_{jk}^{*} + \sum_{k=1}^{K}\sum_{j=1}^{p} G_{jk}^{*\prime}H_{t-j}G_{jk}^{*} \tag{3}$$

with A_{jk}^* and G_{jk}^* being 2×2 parameter matrices and C^* being upper triangular. While the VEC representation[3] is more general, it is more difficult to foster a positive definite H_t in that case. The diagonal VEC looks attractive but despite the computational appeal, it rules out the possibility of explicitly modelling the dependence of the variance terms on one another and thus impedes the assessment of such relationships, which is a prime objective of this paper.

2.2. Dimension of the MGARCH

The dimension of a MGARCH model is characterized by the three parameters K, q and p. The parameter K determines the generality of the process while q and p are the lag orders of matrices A_{jk}^* and G_{jk}^*, respectively. Classical estimation will require the values K, q and p to be fixed *a priori*. We have the privilege of relaxing this as the Bayesian methodology we employ can endogenously determine these model parameters, although for simplicity we prefer to follow the common practice of setting $K=1$. Estimation of q and p necessitates the use of an algorithm that can search through and allow sampling from the parameter space of different dimensions. Towards this end, the RJMCMC emerges as a natural choice. There is more discussion on this in the next section.

2.3. Covariance stationarity

Engle and Kroner (1995) prove that the BEKK representation is covariance stationary if and only if all the eigenvalues of

$$\sum_{i=1}^{q}\sum_{k=1}^{K}(A_{ik}^*\otimes A_{ik}^*) + \sum_{i=1}^{p}\sum_{k=1}^{K}(G_{ik}^*\otimes G_{ik}^*) \tag{4}$$

are less than one in modulus. In other words, covariance stationarity can be ensured by imposing restrictions on the parameter matrices A_{ik}^* and G_{ik}^* alone. To facilitate the specification of prior distributions for the parameters, we state explicitly here the necessary conditions to guarantee stationarity. Consider the MGARCH(1, 1) model with $K=1$. Let

$$A_1^* = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, G_1^* = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}, \mathrm{vech}(H) = \begin{bmatrix} h_{11,t} \\ h_{12,t} \\ h_{22,t} \end{bmatrix}, \mathrm{vech}(C^{*\prime}C^*) = \begin{bmatrix} c_{11} \\ c_{12} \\ c_{22} \end{bmatrix}, \mathrm{vech}(\epsilon_t\epsilon_t{}') = \begin{bmatrix} \epsilon_{1,t}^2 \\ \epsilon_{1,t}\epsilon_{2,t} \\ \epsilon_{2,t}^2 \end{bmatrix}.$$

Then,

$$\begin{aligned} h_{11,t} &= c_{11} + a_{11}^2\epsilon_{1,t-1}^2 + 2a_{11}a_{21}\epsilon_{1,t-1}\epsilon_{2,t-1} + a_{21}^2\epsilon_{2,t-1}^2 + g_{11}^2h_{11,t-1} + 2g_{11}g_{21}h_{12,t-1} + g_{21}^2h_{22,t-1} \\ h_{12,t} &= c_{12} + a_{11}a_{12}\epsilon_{1,t-1}^2 + (a_{12}a_{21} + a_{11}a_{22})\epsilon_{1,t-1}\epsilon_{2,t-1} + a_{21}a_{22}\epsilon_{2,t-1}^2 + g_{11}g_{12}h_{11,t-1} \\ &\quad + (g_{12}g_{21} + g_{11}g_{22})h_{12,t-1} + g_{21}g_{22}h_{22,t-1} \\ h_{22,t} &= c_{22} + a_{12}^2\epsilon_{1,t-1}^2 + 2a_{12}a_{22}\epsilon_{1,t-1}\epsilon_{2,t-1} + a_{22}^2\epsilon_{2,t-1}^2 + g_{12}^2h_{11,t-1} + 2g_{12}g_{22}h_{12,t-1} + g_{22}^2h_{22,t-1}, \end{aligned}$$

[2] See Engle and Kroner (1995) or more recently Bauwens et al. (2006) for detailed description of the theoretical framework of the BEKK and VEC representations.

[3] VEC and BEKK are two general specifications of MGARCH models. The first is so named because the conditional variance matrix is represented in vectorized form, while the latter, as presented above, guarantees positive definiteness automatically and is named after the pioneer researchers in this area Baba, Engle, Kraft and Kroner.

Hence, the conditional variance h_{11} is influenced by lagged values of ϵ_2 via a_{21} and by lagged values of h_{22} via g_{21}. In the context of this paper, it is appropriate to impose the restriction $a_{12}=g_{12}=0$ since the variability in patents granted should not be affected by the variability in stock price *ex ante*. This constraint implies that the matrices A_1^* and G_1^* are lower triangular.

Given the lower triangular structure of the parameter matrices, $A_1^* \otimes A_1^* + G_1^* \otimes G_1^*$ will also be lower triangular with diagonal elements equal to the eigen-values λ_i, $i=1, \cdots, 4$. The stationarity condition of Eq. (4) thus implies the following

$$a_{11}^2 + g_{11}^2 < 1,$$

$$a_{22}^2 + g_{22}^2 < 1,$$

$$a_{11}a_{22} + g_{11}g_{22} < 1.$$

which will be automatically satisfied if

$$|a_{11}| + |g_{11}| < 1, \tag{5}$$

$$|a_{22}| + |g_{22}| < 1. \tag{6}$$

This condition can be generalized analogously to cases where q and p are bigger than 1.

2.4. Prior distributions

To ensure consistency in the terminology used in this article, parameters will henceforth refer to the various parameters of a certain MGARCH model. These include the model orders, the coefficient matrices and the various components of the conditional covariance matrix. The term model, on the other hand, will refer to a MGARCH model of a specific order.

We assume ϵ_t follow a Multivariate Normal distribution. Thus, $\epsilon_t|\mathcal{F}_{t-1} \sim N(0, H_t)$.

1. For β_i: A Normal prior of $\beta_i \sim N(0, \text{var}(\{y_{it}\}))$ for $i=1, 2$ and $t=1, \cdots, T$.
2. For C^*: We propose to block sample the elements of C^* using the reparameterization $C=C^{*\prime} C^*$ and the Wishart Distribution. Specifically, we assume $C^{-1} \sim W((\delta C_0)^{-1}, \delta)$ with $C_0=0.5 \times \text{diag}\,[\text{var}(\{y_{1t}\}), \text{var}(\{y_{2t}\})]$ and $\delta=5$, $t=1, \cdots, T$.
3. For $a_{ii,j}$ and $g_{ii,j}$: To impose the stationarity constraints on the diagonal elements of the MGARCH coefficient matrices, we generalize from Eqs. (5) and (6) to obtain the following:

$$\sum_{j=1}^{q} |a_{ii,j}| + \sum_{j=1}^{p} |g_{ii,j}| < 1 \tag{7}$$

for all i. The coefficients $a_{ii,j}$ and $g_{ii,j}$, conditional on q and p, are Uniformly distributed $\sim U\left(\frac{-1}{q+p}, \frac{1}{q+p}\right)$.
4. For $a_{21,j}$ and $g_{21,j}$: The off-diagonal elements $a_{21,j}$ and $g_{21,j}$ can be freely determined in principle. A Normal prior is assumed for each of these separately $\sim N(0, \omega\sigma^2)$ where $\sigma^2 = \left(\frac{2}{q+p}\right)^2/12$ which is the variance of the Uniform distribution specified above. $\omega=\text{var}(\{y_{1t}\})/\text{var}(\{y_{2t}\})$ is a scale factor that adjusts for the difference in the magnitudes of the two variables.
5. For q and p: The MGARCH orders $q \in (0, q_{\max})$ and $p \in (0, p_{\max})$ have Uniform priors $U(0, q_{\max})$ and $U(0, p_{\max})$, respectively with $q_{\max}=p_{\max}=12$.
6. For H_t: The conditional variance can be evaluated recursively using Eq. (3) for $t=(q \vee p)+1, \cdots, T$. The conditional variance for H_t, $t=1, \cdots, q \vee p$ will have to be determined *a priori*. We assume a Wishart prior for each of these initial conditional variance matrices H_t, $t=1, \cdots, q \vee p$: $H_t^{-1} \sim W((\nu\Sigma)^{-1}, \nu)$ where $\Sigma=\text{diag}\,[\text{var}(\{y_{1t}\}), \text{var}(\{y_{2t}\})]$ and degrees of freedom $\nu=10$.

3. The algorithm

3.1. Motivation

A comprehensive description of Bayesian inference and the MCMC methodology is beyond the scope of this article,[4] but some background information on these issues will facilitate understanding of material that is presented below. The essence of Bayesian modelling lies on the Bayes' Theorem, which states that the posterior distribution of parameters of interest are proportional to the product of prior distribution and the likelihood function of the data. In a way, the posterior distribution, or the probability distribution of parameters conditional on the observed data, strikes a balance between the prior belief one incorporates in the model and the statistical message the data deliver. Inference is then conducted using the posterior distribution, e.g. by evaluating its moments. While the posterior moments may be numerically tractable for simple and low dimensional problems, the situation may be very different in complex and high dimensional problems. Very often, in the latter case, the posterior will be of an unfamiliar functional form and inference will have to rely on computer simulation.

MCMC is a computational technique for sampling from complex probabilistic models. The idea is to set up a Markov chain with the desired posterior as a stationary distribution to which the chain eventually converges. Simulation is performed by running the chain repeatedly, and the Markov nature implies that ergodic averages of the samples drawn from each iteration can be used to approximate the posterior expectations. Standard MCMC methods like the Gibbs sampler and the Metroplis–Hastings chain can deal with cases of fixed parameter dimension very well. However, these may not be sufficient when the parameter spaces are subject to changes. Consider the example of variable selection for a regression model; the corresponding parameters involved will change as the number of covariates in the model varies. In cases like this, we need an algorithm that can search and sample from parameter spaces of different dimensionalities, so that inference on posterior moments can be conducted.

The RJMCMC method is a variant of the Metropolis–Hastings algorithm and was first proposed by Green (1995) to serve the need of handling problems of varying parameter dimensions. Indeed, early applications included working out the number of components needed in mixture models, variable selection for regression models and order selection for autoregressive models. An important area in which RJMCMC has excelled is model averaging or model selection; see for instance Avramov (2002). The model averaging approach is similar to other problems except that apart from parameter uncertainty, there is also model uncertainty at stake. For example, how likely will the true model be an autoregressive model of order 1 and how likely will it be of order N? By averaging over all possible models in the model space instead of relying on a single model, one can conduct inferences based on the average of posterior distributions under each model weighted by the corresponding posterior model probabilities. RJMCMC comes in handy in this respect. It allows jumps between models of different parameter/model dimensions and samples from models according to their posterior probabilities. Those with negligible probabilities will be rarely visited and only the most probable models are investigated frequently.

There is a striking similarity here in our application. Inasmuch as a model MGARCH(q, p) is distinct from one with different orders, say, MGARCH(q', p'), the RJMCMC is invoked to search through the parameter/model space in each iteration and draw samples from the posteriors of the most probable models. As these samples will eventually be integrated for inference purposes, we are in essence practicing a model averaging exercise.

3.2. Overview of RJMCMC

For ease of exposition, let us define $k=(q, p)$ and the corresponding subset of parameters as θ_k . Let us denote the remaining parameters as ϕ and evaluate the jump from dimension k to k' as follows:

1. Starting from the current order k, we propose a move $k \rightarrow k'$ by sampling k' from the transition probability distribution $J(k \rightarrow k')$.
2. Given k', we sample the parameter $\theta_{k'}$ from the proposal distribution $f(\theta_{k'}|k', Y, \theta_k, \phi)$.

[4] Gammerman (1997) is a good introductory text on these topics.

3. We then compute the acceptance ratio:

$$\mathcal{A} = ((k,\theta_k)\rightarrow(k',\theta_{k'})) = \min\left\{1, \frac{p(k',\theta_{k'}|Y,\phi)J(k'\rightarrow k)f(\theta_k|k,Y,\theta_{k'},\phi)}{p(k,\theta_k|Y,\phi)J(k\rightarrow k')f(\theta_{k'}|k',Y,\theta_k,\phi)}\right\} \tag{8}$$

4. The move $k\rightarrow k'$ is accepted with probability $\mathcal{A}((k,\theta_k)\rightarrow(k',\theta_{k'}))$. If the move is rejected, we stay at order k.
5. As in Troughton and Godsill (1998), we use a discretized Laplacian density for $J(k\rightarrow k')$ which gives $\Pr(k'|k)$

$$J(k\rightarrow k') \propto \exp\{-\lambda|k'-k|\},\ k'\in\{0,\cdots,k_{\max}\}$$

In RJMCMC, the jumps between state spaces of different dimensions are governed by the density function $J(k\rightarrow k')$ which gives the transition probability of jumping from model $\mathcal{M}_k$ to model $\mathcal{M}_{k'}$. In our context, these two models differ in terms of their MGARCH orders $k=(q,p)$. When structuring this density, a crucial issue is the detailed balance. This essentially means that each jump between state spaces must have a reverse jump defined so that the Markov chain can be made reversible and eventual convergence to the stationary distribution can be ensured. The discretized Laplacian density used in the paper obviously satisfies the detailed balance requirement since $J(k\rightarrow k')=J(k'\rightarrow k)$. Another feature is that the implied transition probability to a distinctly different model order will be significantly smaller than that to a neighboring model order.

3.3. Implementation issues

Define $\Theta=(q,p,C,\{A_1^*,\ldots,_q,G_1^*,\ldots,_p\},\{\beta_i\},H_{t=1,\ldots,q\vee p})$ and let $Y=[y_t]$, $t=1,\cdots,T$. The relevant likelihood is thus

$$p(Y|\Theta,y_0) \propto \prod_{t=1}^{T}|H_t|^{-1/2}\exp\left\{-\frac{1}{2}\epsilon_t{}'H_t^{-1}\epsilon_t\right\} \tag{9}$$

and the joint distribution is

$$\begin{aligned} p(\Theta|Y,y_0) \propto & \prod_{t=1}^{T}|H_t|^{-1/2}\exp\left\{-\frac{1}{2}\epsilon_t{}'H_t^{-1}\epsilon_t\right\}\times\frac{1}{q_{\max}p_{\max}} \\ & \times\left(\frac{q+p}{2}\right)^{2(q+p)}\times\exp\left\{-\frac{1}{2}\sum_{i=1}^{2}\frac{(\beta_i)^2}{\mathrm{var}(\{y_{it}\})}\right\} \\ & \times\exp\left\{-\frac{1}{2}\left[\sum_{r=1}^{q}\frac{(a_{21,r})^2}{\omega\sigma^2}+\sum_{s=1}^{p}\frac{(g_{21,s})^2}{\omega\sigma^2}\right]\right\} \\ & \times\prod_{t=1}^{q\vee p}|H_t|^{-(v-3)/2}\exp\left\{-\frac{1}{2}\mathrm{trace}(v\times\Sigma\times H_t^{-1})\right\} \\ & \times|C|^{-(\delta-3)/2}\exp\left\{-\frac{1}{2}\mathrm{trace}(\delta\times C_0\times C^{-1})\right\} \end{aligned} \tag{10}$$

To execute the RJMCMC sampling, note that we can write $k=(q,p)$, $\theta_k=(\{A_{1,\cdots,q}^*, G_{1,\cdots,p}^*\}, H_{t=1,\cdots,q\vee p})$, and $\phi=(C, \{\beta_i\})$. Thus,

$$\begin{aligned} p(k,\theta_k|Y,\phi) \propto & \prod_{t=1}^{T}|H_t|^{-1/2}\exp\left\{-\frac{1}{2}\epsilon_t{}'H_t^{-1}\epsilon_t\right\}\times\left(\frac{q+p}{2}\right)^{2(q+p)} \\ & \times\exp\left\{-\frac{1}{2}\left[\sum_{r=1}^{q}\frac{(a_{21,r})^2}{\omega\sigma^2}+\sum_{s=1}^{p}\frac{(g_{21,s})^2}{\omega\sigma^2}\right]\right\} \\ & \times\prod_{t=1}^{q\vee p}|H_t|^{-(v-3)/2}\exp\left\{-\frac{1}{2}\mathrm{trace}(v\times\Sigma\times H_t^{-1})\right\} \end{aligned} \tag{11}$$

and

$$J(k \to k') \propto \exp\{-\lambda_q|q'-q|\}\exp\{-\lambda_p|p'-p|\}, q' \in \{0,\cdots,q_{\max}\}, p' \in \{0,\cdots,p_{\max}\}, \lambda_q = \lambda_p = 0.5 \tag{12}$$

The proposal distribution $f(\theta_{k'}|k', Y, \theta_k, \phi)$ does not have to be the full conditional distribution of $\theta_{k'}$, although assuming this will simplify the evaluation of $\mathcal{A}((k,\theta_k) \to (k',\theta_{k'}))$ in which case the latter will be defined solely by k' because

$$\frac{p(k',\theta_{k'}|Y,\phi)}{p(\theta_k{}'|k',Y,\phi)} = p(k'|Y,\phi).$$

This would, however, require marginalizing elements – integrating out from the distribution the elements – of $\theta_{k'}$ from Eq. (11). Instead, defining

$$\theta_{k',1} \equiv \begin{bmatrix} a_{ii,1} \\ \vdots \\ a_{ii,q'} \\ g_{ii,1} \\ \vdots \\ g_{ii,p'} \end{bmatrix}, i = 1,2 \cdot \quad \theta_{k',2} \equiv \begin{bmatrix} a_{21,1} \\ \vdots \\ a_{21,q'} \\ g_{21,1} \\ \vdots \\ g_{21,p'} \end{bmatrix} \cdot \quad \theta_{k',3} \equiv H^{-1}_{t=1,\ldots,q\vee p},$$

the proposal distribution is assumed to take the form

$$f(\theta_{k'}|k',Y,\theta_k,\phi) = \prod_{j=1}^{3} f_j(\theta_{k',j}) \tag{13}$$

where

$$f_1(\theta_{k',1}) \sim \prod_{i=1}^{2}\left[\prod_{j=1}^{q'} N(a_{ii,j}|\mu^a_{ii},\sigma^2)\prod_{j=1}^{p'} N(g_{ii,j}|\mu^g_{ii},\sigma^2)\right] \times 1\left(\{a_{ii,j},g_{ii,j}\} \in \left(-\frac{1}{q'+p'},\frac{1}{q'+p'}\right)\right), \tag{14}$$

$$f_2(\theta_{k',2}) \sim \prod_{j=1}^{q'} N(a_{21,j}|\mu^a_{21},\omega\sigma^2)\prod_{j=1}^{p'} N(g_{21,j}|\mu^g_{21},\omega\sigma^2), \tag{15}$$

$$f_3(\theta_{k',3}) \sim \prod_{t=1}^{q\vee p} W\left(H_t^{-1}|(v\times\Sigma)^{-1},v\right). \tag{16}$$

where $1(\cdot)$ is an indicator variable and $\mu^a_{ii,j}$ is the mean of the nonzero elements of the iterated $a_{ii,j}$ sampled hitherto. $\mu_{ii,j}{}^g$, μ^a_{21} and μ^g_{21} are similarly defined. Thus, subsequent sampling from the proposal distribution will factor in the historic information retrieved by the chain. To save space, the detailed simulation procedures are given in Appendix A.

4. Empirical results

As an illustration of how the RJMCMC MGARCH(q, p) model can be applied to unveil possible associations between volatilities of stock price and patents granted, we ran the model using weekly data of Apple Computer, IBM and Unisys from Feb., 1984 to Dec., 1999. The data were obtained from Datastream and the NBER Patent Citations Data File, and were first converted to growth rates before implementation. A total of 10,000 iterations were performed for each stock and the initial 5000 were burn-ins. Posterior estimates were obtained by computing batch means using the remaining 5000 loops. These sampling outputs were first grouped into batches of size 10, one MCMC draw being selected from each batch to yield the estimates of interest. The major findings are summarized in Table 1.

Fig. 1 depicts the posterior distribution of the model order $k=(q,p)$ for the case of IBM. It can be seen that the model favors an order of $q=3$, $p=0$ suggesting thereby a Multivariate ARCH specification instead of a MGARCH one could best capture the inherent persistence in the data. Using the MCMC output, we can also obtain approximate model

Table 1
Posterior estimates of selected parameters of MGARCH model

Parameter	IBM		Apple		Unisys	
	Posterior mean (batch size=10)	95% HPD Interval	Posterior mean (batch size=10)	95% HPD interval	Posterior mean (batch size=10)	95% HPD interval
Model order	q=3, p=0		q=3, p=0		q=4, p=0	
β_1	0.0016	(−0.0008, 0.0041)	0.0023	(−0.0019, 0.0070)	6.3646e−004	(−0.0339, 0.0058)
β_2	−0.0019	(−0.0330, 0.0350)	0.0013	(−0.0039, 0.0362)	−0.0031	(−0.0560, 0.0550)
$a_{11,1}$	0.0055	(−0.3433, 0.3907)	0.0099	(−0.3855, 0.3596)	−0.0102	(−0.2818, 0.3399)
$a_{21,1}$	−0.0015	(−0.0217, 0.0170)	−3.7443e−004	(−0.0329, 0.0404)	6.2245e−004	(−0.0197, 0.0248)
$a_{22,1}$	−0.0444	(−0.2783, 0.3833)	−0.0400	(−0.3917, 0.3373)	0.0104	(−0.2508, 0.2672)
$a_{11,2}$	0.0023	(−0.1971, 0.2080)	−0.0283	(−0.2642, 0.1471)	0.0443	(−0.2338, 0.2747)
$a_{21,2}$	−0.0039	(−0.0247, 0.0078)	0.0047	(−0.0240, 0.0344)	0.0069	(−0.0114, 0.0320)
$a_{22,2}$	−0.0189	(−0.2511, 0.1646)	−0.0068	(−0.3017, 0.1767)	−4.7076e−004	(−0.2370, 0.2158)
$a_{11,3}$	−0.0187	(−0.2188, 0.1448)	−0.0123	(−0.1659, 0.1546)	−0.0201	(−0.2426, 0.1012)
$a_{21,3}$	−0.0001	(−0.0114, 0.0146)	−0.0067	(−0.0351, 0.0142)	−0.0074	(−0.0336, 0.0002)
$a_{22,3}$	−0.0136	(−0.1979, 0.1001)	−0.0022	(−0.2066, 0.1480)	−0.0037	(−0.2146, 0.1316)
c_{11}	0.0011	(−0.0007, 0.0015)	0.0038	(0.0024, 0.0055)	0.0041	(0.0025, 0.0059)
c_{21}	0.0001	(−0.0037, 0.0043)	3.0871e−004	(−0.0089, 0.0069)	−0.0035	(−0.0178, 0.0127)
c_{22}	0.2106	(0.1515, 0.2860)	0.2618	(0.1485, 0.3707)	0.6466	(0.4160, 0.8665)

posteriors $p(\mathcal{M}_i|Y)$ with which the Bayes factor can be constructed. The Bayes factor is defined as the ratio of posterior odds to the prior odds

$$B_{01} = \frac{p(\mathcal{M}_0|Y)/p(\mathcal{M}_1|Y)}{p(\mathcal{M}_0)/p(\mathcal{M}_1)}$$

and a proxy for $p(\mathcal{M}_0|Y)$ is

$$\widehat{p}(\mathcal{M}_0|Y) = \frac{1}{N}\sum_{n=1}^{N} 1\left(\mathcal{M}^{(n)} = \mathcal{M}_0\right)$$

where N is the total number of iterations. In our context, it will be interesting to see how the estimated order compares to MGARCH(0, 0) and MGARCH(1, 1). Assuming that various model orders are equal-probable *a priori*, $B_{(q=3,p=0)(q=p=0)}$= 1.27 and $B_{(q=3,p=0)(q=p=1)}$=25. Thus, the statistical evidence strongly favors the estimated order against MGARCH (1, 1), which is so often assumed in the literature, although the same cannot be asserted when compared to MGARCH(0, 0). The corresponding figures for Apple Computer and Unisys are {1.34, 49.79} and {1.45, 18.16} respectively.

In addition to the Bayes factors presented above, two other model adequacy exercises have been conducted. The first is the Bayesian Information Criterion (BIC), which is a traditional model selection criterion based on likelihood evaluation. Raftery (1996) defines it as

$$\log\widehat{P}_{\mathrm{Schwarz}}(Y) = \max_n \log p\left(Y|\Theta^{(n)}, y_0\right) - \frac{1}{2}d\log(rT)$$

where d is the number of parameters and r is the dimension of y_t. Again, comparison is done between the estimated model, and the two benchmark models MGARCH(0, 0) and MGARCH(1, 1). The BICs for these three specifications of IBM are {935.8, 997.08, 870.57}, while those for Apple Computer and Unisys are {368.98, 212.72, 449.35} and {−287.89, −84.475, −174.83} respectively. The results, not unexpectedly, discriminate against the estimated models in favor of more parsimonious counterparts, which the BIC is well known for. In fact, the maximum log-likelihoods, the first component in the above expression, give somewhat different indications. They are {1050.7, 1045.3, 1052.2} for IBM, {506.1, 416.56, 475.3} for Apple Computer, and {27.124, 8.1766, 6.7632} for Unisys. So the maximum log-likelihood favors the estimated models in virtually all three cases.

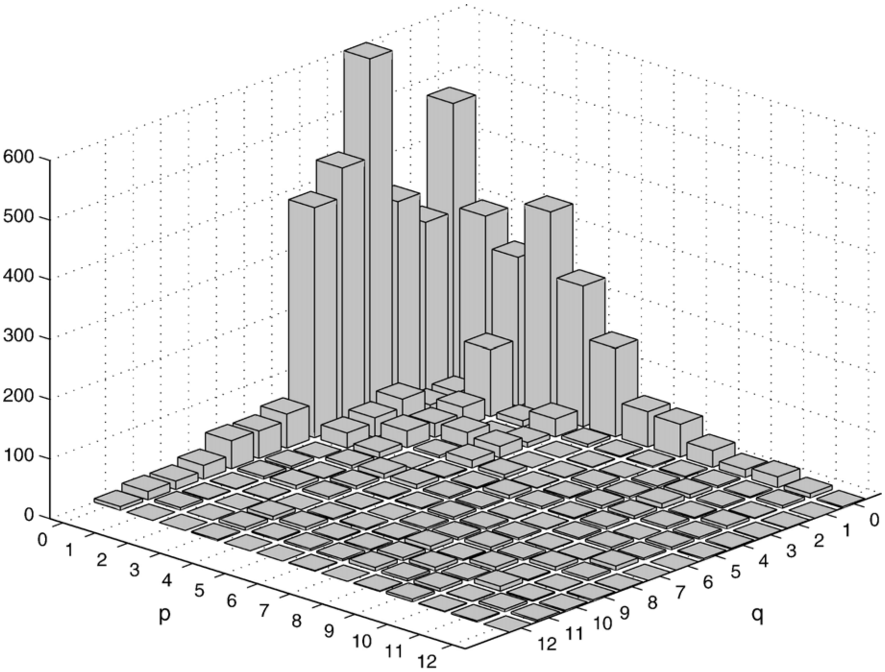

Fig. 1. Posterior distribution of model order (q, p).

The second exercise involves the conditional predictive ordinate (CPO), which is the conditional density of an observed data $y_{t,\text{obs}}$, given all but the t-th observed data $Y_{-t,\text{obs}}$. It facilitates the study of incremental contributions to the adequacy of the model by the t-th observation. A large CPO indicates support of a certain model by that observation and vice-versa. Gelfand (1996) provides both the definition of CPO and how that can be approximated using MCMC output. The CPO for model k and observation t is defined as

$$\text{CPO}_t^k = p_k(y_{t,\text{obs}}|Y_{-t,\text{obs}})$$

and can be approximated by a harmonic mean formula

$$\widehat{\text{CPO}}_t^k = \left\{\frac{1}{N}\sum_{n=1}^{N}\frac{1}{p(y_t|\Theta^{(n)})}\right\}^{-1}.$$

This approximation can then be used to construct CPO ratios for model k and k', $C_{t,(k)(k')} = \widehat{\text{CPO}}_t^k / \widehat{\text{CPO}}_t^{k'}$. If this ratio exceeds 1, the evidence supports model k, and vice-versa for a number smaller than 1. Of the $T=828$ observations we have for the three stocks, we conducted pairwise comparisons between the estimated model and the two benchmark models for each t. For IBM, the number of $C_{t,(q=3,p=0)(q=p=0)}$ that are greater than 1 equal 285, and those for $C_{t,(q=3,p=0)(q=p=1)}$ equal 490. The corresponding figures for Apple Computer are 611 and 584, whereas those for Unisys are 289 and 195 respectively. The results strongly support the estimated model for Apple Computer, give mixed indications for IBM and support the benchmark models for Unisys. This notwithstanding, the estimated models seem to perform better in terms of incorporating outliers when compared to the benchmark models. The sums of the CPO ratios over all t for IBM are {15689, 801.7}, {42733, 870.5} for Apple Computer, and {6.8026e+014, 801.5} for Unisys. So for all three cases, the estimated models explain the outliers in the sample much better than MGARCH(0, 0) although, again, the results are mixed when compared to MGARCH(1, 1).

A summary of posterior estimates of various parameters are highlighted in Table 1. As far as IBM is concerned, we see that $a_{21,j}$ for $j=1, \cdots, 3$ are all negatively signed, meaning that variations in patents granted will negatively impact on variations in stock prices. For Apple Computer, on the other hand, the relationship alters the sign depending on the

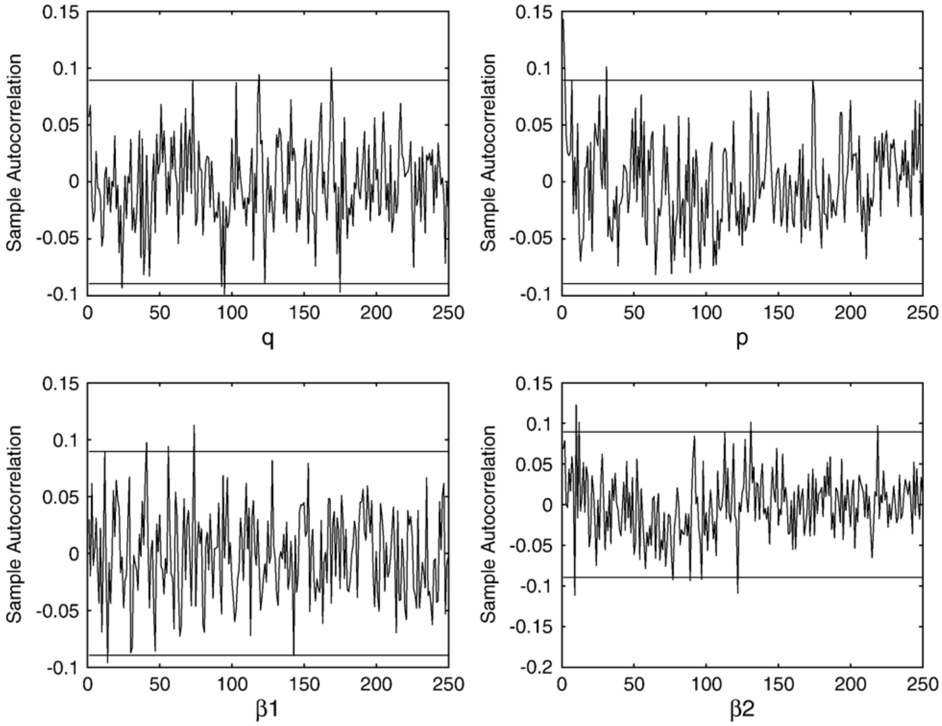

Fig. 2. Sample autocorrelation functions of model orders and betas with 95% confidence intervals.

lag. Apart from the posterior means, the 95% Highest Probability Density[5] (HPD) intervals for the concerned parameters are also listed for reference.

As for model diagnostics, we include in Figs. 2 and 3 the sample autocorrelation functions of selected parameters as a means to monitor convergence of the chain. Again, to save space, only the plots for IBM are included. Since the autocorrelation functions die out rapidly, sampling efficiency should not be a concern. In addition, the acceptance ratio for sampling model order k=(q, p) is 27.28% which falls within the standard ballpark of between 25% and 50% required for good mixing. The ratios for Apple and Unisys are 28.66% and 28.08% respectively. Thus, the RJMCMC should have explored and visited the subspace of k fairly thoroughly.

5. Application: VaR analysis

An integral part of our analysis is to assess the prediction power of the RJMCMC MGARCH model. We perform VaR analysis based on the empirical findings and compare the result with those obtained using other benchmark approaches in the financial industry. This metric is defined as the lower tail percentile for the distribution of the underlying asset. Apart from the industry's favorite, the RiskMetrics™, other approaches with whose results we compare include the Normal model, Historical Simulation and the GARCH(1, 1) model. A detailed introduction to these volatility measurement methods can be found in Jorion (2001).

We take out the last 100 weekly observations to construct the in-sample and estimate separately each of the benchmark models whereby a one-step-ahead 95%-VaR estimate is generated. The models are then re-estimated successively using a rolling-window of adding one observation each time. For the RJMCMC MGARCH model, instead of running the simulation exercise repeatedly, we employ the resampling approach suggested by Gelfand (1996) to obtain the out-of-sample forecasts.[6] The results of the VaR analysis are highlighted in Table 2. The first

[5] See Chen et al. (2000) for a description of the Chen-Shao algorithm used to compute the HPD intervals.

[6] An outline of the resampling procedure is provided in the Appendix.

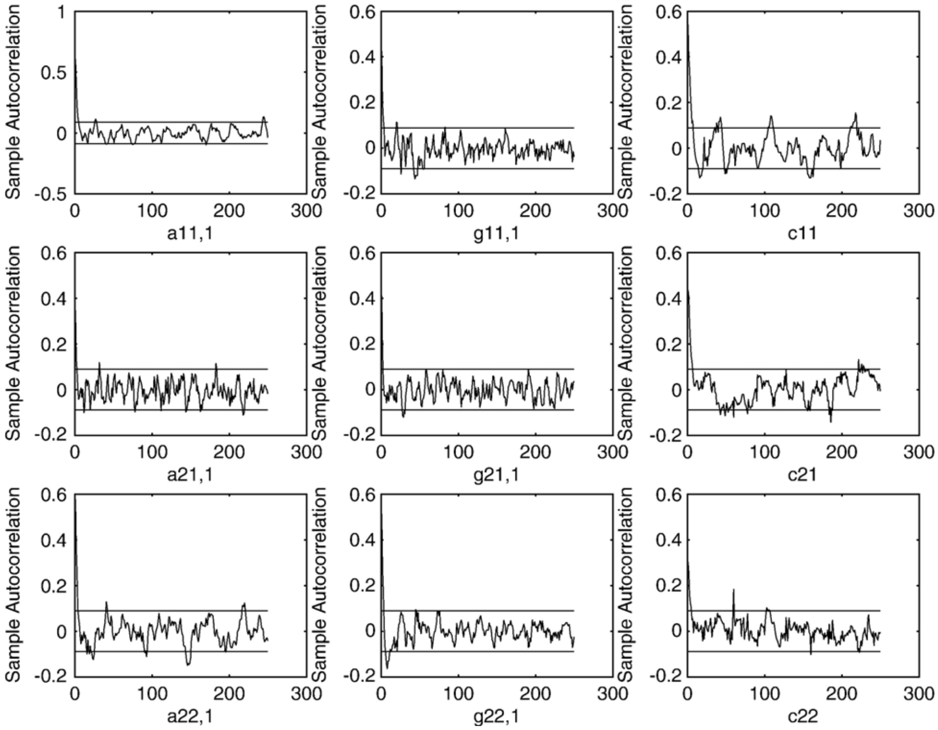

Fig. 3. Sample autocorrelation functions of A*1, G*1 and C with 95% confidence intervals.

reported statistic is an indictor that addresses the issue of accuracy. It is a binary loss function which measures the number of times the actual price change exceeds the VaR estimates:

$$L_{t+1} = \begin{cases} 1 & \text{if } \Delta P_{t+1} < \mathrm{VaR}_t \\ 0 & \text{if } \Delta P_{t+1} \geq \mathrm{VaR}_t \end{cases}$$

where ΔP_{t+1} = return of stock at period $t+1$. Thus the average of L_{t+1} should be 0.05 for 95%-VaR estimates. The average exceptions of the various models are, reported in triplet for IBM, Apple and Unisys respectively: {0.06, 0.02, 0.04} for the RJMCMC MGARCH model, {0.06, 0.05, 0.04} for the Normal model, {0.06, 0.06, 0.04} for the Historical Simulation model, {0.08, 0.04, 0.04} for the GARCH(1, 1) model and {0.07, 0.04, 0.04} for the

Table 2
Value-at-Risk analysis using MGARCH model

	IBM	Apple	Unisys
Binary loss (failure %)			
RJMCMC-MGARCH	6	2	4
Normal	6	5	4
Historical simulation	6	6	4
RiskMetrics™	8	4	4
GARCH(1, 1)	7	4	4
Mean relative bias			
RJMCMC-MGARCH	−0.0079	0.0838	0.0296
Normal	−0.0531	−0.0377	−0.0160
Historical simulation	−0.0331	−0.1564	−0.1196
RiskMetrics™	0.0409	0.0876	0.0377
GARCH(1, 1)	0.0532	0.0228	0.0682

RiskMetrics™ model. So the model considered in this article produces VaR estimates that are at least as accurate as other standard models if not significantly better.

The second statistic that can shed light on the usefulness of models is the Mean Relative Bias (MRB) introduced by Hendricks (1996). This captures the extent to which different models produce estimates of similar average size, and is defined as:

$$\mathrm{MRB}_i = \frac{1}{T}\sum_{t=1}^{T}\frac{\mathrm{VaR}_{it}-\overline{\mathrm{VaR}_t}}{\overline{\mathrm{VaR}_t}}, \quad \text{where} \qquad \overline{\mathrm{VaR}_t} = \frac{1}{N}\sum_{i=1}^{N}\mathrm{VaR}_{it},$$

with T being the forecast horizon and N the number of competing models. The MRB for the models can be found in Table 2. As an illustration, the figures for IBM are, respectively: −0.0079 for the RJMCMC MGARCH model, −.0531 for the Normal model, −0.0331 for the Historical Simulation approach, 0.0532 for the GARCH(1, 1) model and 0.0409 for the RiskMetrics™ model. This means that, for IBM, the RJMCMC MGARCH model produces estimates that are less than 1% smaller on average than those of other models, while those by the RiskMetrics™ model are about 4% larger. Inasmuch as a smaller VaR estimate implies lower capital cost, we can conclude that the RJMCMC MGARCH model is as accurate as the Normal and Historical Simulation models but would involve lower capital cost. As for the GARCH(1, 1) and the RiskMetrics™ model, there is a trade-off between lower costs and accuracy. Similar interpretations can be extended to the cases of Apple and Unisys.

To offer a different perspective of the VaR analysis, we plot out the VaR estimates (again for IBM only to save space) in Fig. 4. In each panel there, the prediction of the RJMCMC MGARCH model is superimposed with that of a competing model and the actual return series of IBM over the out-of-sample window. It can be seen that the RJMCMC MGARCH model tracks the volatility of the series much better than the Normal model and the Historical Simulation approach. When compared to the predictions of the RiskMetrics™ and the GARCH(1, 1) models, the RJMCMC MGARCH model seems to be able to differentiate between periods of positive and negative growth so that the VaR estimates are adjusted correspondingly.

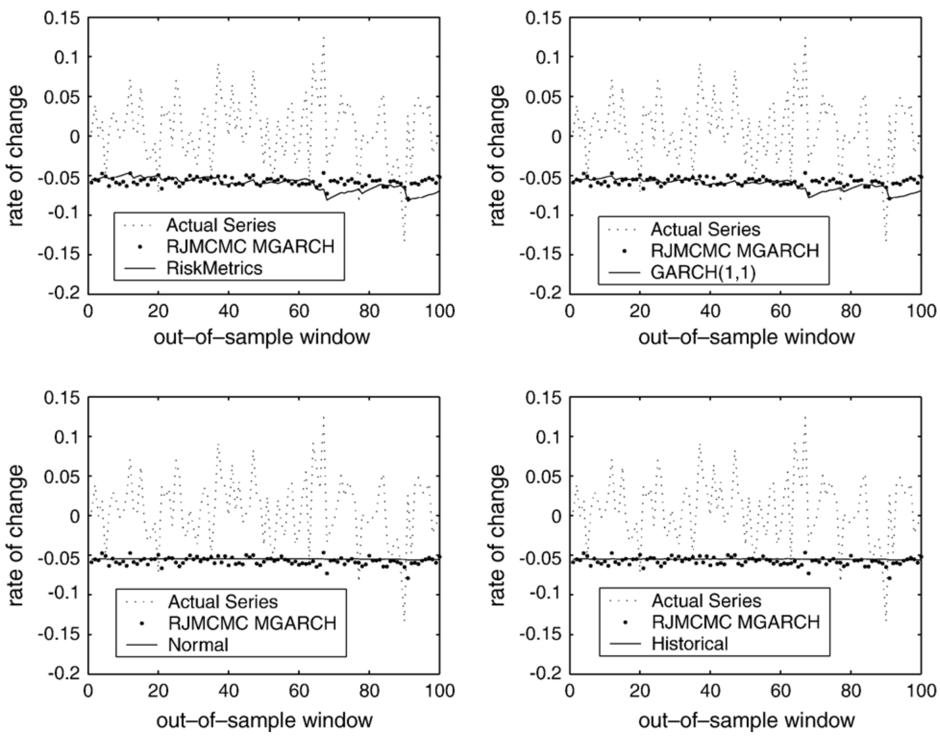

Fig. 4. Summary plots of Value-at-Risk analysis.

6. Conclusion

This paper offers a quantitative approach to analyzing possible associations of stock price changes and variations in innovative activities. The modelling requirement is relatively general with the only restrictions being the enforcement of a uni-directional causation from innovation to stock values and that of covariance stationarity. Instead of predetermining the model order as in other studies, we leave that an open issue and allow the model to search for the most probable dimension using the RJMCMC method. The statistical evidence suggests that in all cases we considered, a Multivariate ARCH model will suffice. This implies that while the assumption of MGARCH(1, 1) may be convenient and parsimonious, its validity should be further substantiated by some model comparison exercises; see for instance Osiewalski and Pipień (2004). In addition, variations in the patents granted turn out to be negatively related to the volatility of its stock prices for IBM, contrary to what Eden and Jovanovic (1994) would predict. The negative relationship is also there for Apple Computer and Unisys, although not at all lags. While the application of this model to VaR analysis may be limited only to equities, it compares favorably to other benchmark methods in rendering accurate VaR estimates. A useful extension of the methodology introduced here would be to model a portfolio of assets, not necessarily equities, whereby bi-directional impacts of one asset on the conditional variance of the other can be incorporated.

Appendix A

A.1. Simulation of parameters

Sample elements of $\Theta^{(m)}$ in each MCMC iteration $m=1, \cdots, M$ as follows:

1. In iteration m, sample q' and p' from Eq. (12) based on q and p from iteration $m-1$ as follows:
(a) Define $\gamma_{\max} \equiv \exp\left\{-\frac{1}{2}|q-q|\right\} = 1$.
(b) Define $\gamma_{\min} \equiv \exp\left\{-\frac{1}{2}|[(q_{\max}-q) \vee (q-0)] - q|\right\}$.
(c) Define $\alpha_{\max} \equiv (\gamma_{\max} - \gamma_{\min})(q_{\max} - 0)$.
(d) Generate $\alpha \sim U(0, \alpha_{\max})$ and $\gamma \sim U(\gamma_{\min}, \gamma_{\max})$.
(e) Return $q' = 0 + \mathrm{int}\left(\frac{\alpha}{\gamma_{\max}-\gamma_{\min}}\right)$ if $\exp\left\{-\frac{1}{2}|q'-q|\right\} \leq \gamma$. Otherwise repeat (a) to (e).
(f) Repeat process for sampling of p'.
2. Sample elements of $\theta_{k'}$ sequentially from proposal distributions Eqs. (14)–(16). Compute the acceptance ratio $\mathcal{A}((k,\theta_k)\rightarrow(k',\theta_{k'}))$ as in Eq. (8). Generate $u \sim U(0, 1)$. If $u \leq \mathcal{A}((k,\theta_k)\rightarrow(k',\theta_{k'}))$ then $\theta_k^{(m)} = \theta_{k'}$. Otherwise, $\theta_k^{(m)} = \theta_k^{(m-1)}$.
3. Sample C with a Independence Sampler and use the prior $pC^{-1} = W(C^{-1}|(\delta C_0)^{-1}, \delta)$ as the proposal. Evaluate

$$\alpha(C, C^{(m-1)}) = \min\left\{1, \frac{p(C|Y,k,\theta_k,\phi)}{p(C^{(m-1)}|Y,k,\theta_k,\phi)}\right\} \tag{17}$$

where

$$p(C|Y,k,\theta_k,\phi) \propto \prod_{t=1}^{T} |H_t|^{-1/2} \exp\left\{-\frac{1}{2}\epsilon_t' H_t^{-1}\epsilon_t\right\} \times \prod_{t=1}^{q\vee p} |H_t|^{-(v-3)/2} \exp\left\{-\frac{1}{2}\mathrm{trace}(v \times \Sigma \times H_t^{-1})\right\} \tag{18}$$

Generate $u \sim U(0, 1)$. If $u \leq \alpha(C, C^{(m-1)})$, then $C^{(m)} = C$. Otherwise, $C^{(m)} = C^{(m-1)}$.
4. Sample $\{\beta_i\}$, $i=1, 2$ with a Gibbs step. Define $Z \equiv [y_0, y_1, \cdots, y_{T-1}]$, $y \equiv \mathrm{vec}(Y)$, $\beta_+ \equiv \mathrm{vec}(B)$ and

$$\Omega^1 \equiv \begin{bmatrix} H_1^{-1} & 0 & \cdots & 0 \\ 0 & H_2^{-1} & & \vdots \\ \vdots & & O & 0 \\ 0 & \cdots & 0 & H_T^{-1} \end{bmatrix}.$$

Define also $\iota \equiv [1\ 0\ 0\ 1]'$, $\iota_{2T} \equiv$ stacking of $2T$ rows of ι', $\mu_+ \equiv 0_{4\times 1}$, and

$$\Sigma_+ \equiv \begin{bmatrix} \mathrm{var}(\{y_{1t}\}) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathrm{var}(\{y_{2t}\}) \end{bmatrix}, \Lambda \equiv \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Then the conditional for β_+ is

$$p(\beta_+|Y,k,\theta_k,C) \propto \exp\left\{ \begin{matrix} -\frac{1}{2}[y-((Z'\otimes I_2)\odot\iota_{2T})(\beta_+\odot\iota)]'\Omega^{-1} \\ \times[y-((Z'\otimes I_2)\odot\iota_{2T})(\beta_+\odot\iota)] \end{matrix} \right\} \times \exp\left\{ \begin{matrix} -\frac{1}{2}(\beta_t\odot\iota-\mu_+\odot\iota)'(\sum_+^{-1}\odot\Lambda) \\ \times(\beta_+\odot\iota-\mu_+\odot\iota) \end{matrix} \right\}$$

which is a Multivariate Normal distribution:

$$p(\beta_+|Y,k,\theta_k,C) \sim N(\beta_+\odot\iota|\,\widetilde{\mu}_+,\widetilde{\Sigma}_+) \text{ where}$$

$$\widetilde{\Sigma}_+ = \begin{bmatrix} \Sigma_+^{-1}\odot\Lambda + ((Z'\otimes I_2)\odot\iota_{2T})'\Omega^{-1} \\ \times((Z'\otimes I_2)\odot\iota_{2T}) \end{bmatrix}^{-1},$$

$$\widetilde{\mu}_+ = \widetilde{\Sigma}_+[(\Sigma_+^{-1}\odot\Lambda)(\mu_+\odot\iota) + ((Z'\otimes I_2)\odot\iota_{2T})'\Omega^{-1}y].$$

The conditional for the reduced coefficient vector $\beta \equiv$ nonzero elements of β_+ is then:

$$p(\beta|Y,k,\theta_k,C) \sim N(\beta|\,\widetilde{\mu}_\beta,\widetilde{\Sigma}_\beta) \text{ where} \tag{19}$$

$$\widetilde{\Sigma}_\beta^{-1} = \text{ non zero elements of } [\Sigma_+^{-1}\odot\Lambda + ((Z'\otimes I_2)\odot\iota_{2T})'\Omega^{-1}((Z'\otimes I_2)\odot\iota_{2T})], \tag{20}$$

$$\widetilde{\mu}_\beta = \widetilde{\Sigma}_\beta \times \text{non zero elements of } [(\Sigma_+^{-1}\odot\Lambda)(\mu_+\odot\iota) + ((Z'\otimes I_2)\odot\iota_{2T})'\Omega^{-1}y]. \tag{21}$$

Appendix B

Resampling procedures for estimating 95%-VaR for the RJMCMC MGARCH model:

- Define the cross-validation predictive density $p(y_r\,|Y_{(r)})$ where $Y_{(r)}$ denotes all elements of Y except y_r.
- Given MCMC output of parameters θ_j^* , $j=1, \cdots, M$, construct importance ratio

$$w_j = \frac{1}{p(y_r|Y_{(r)}x,\theta_j^*)}.$$

- Resample with replacement from the set of $\{\theta_j^*\}$ with probabilities proportional to $\{w_j\}$. This resulting sample (θ_j^{**}) is approximately from $p(\theta|Y_{(r)})$.
- Draw Y_j^{**} from $p(Y|\theta_j^{**})$ and the r-th element of Y_j^{**} is then a sample from $p(y_r|Y_{(r)})$.

References

Avramov, D., 2002. Stock return predictability and model uncertainty. Journal of Financial Economics 64, 423–458.

Bauwens, L., Lubrano, M., 1998. Bayesian inference on GARCH models using the Gibbs sampler. Econometrics Journal 1, C23–C46.

Bauwens, L., Laurent, S., Rombouts, J.V.K., 2006. Multivariate GARCH models: a survey. Journal of Applied Econometrics 21 (1), 79–109.

Ben Zion, U., 1978. The investment aspect of non-production expenditures: an empirical test. Journal of Economics and Business 30, 224–229.

Bloom, N., VanReenen, J., 2002. Patents, real options and firm performance. Economic Journal 112 (478), C97–C116.

Brennan, M., Xia, Y., 2001. Stock price volatility and equity premium. Journal of Monetary Economics 47 (2), 249–283.

Chan, S.H., Martin, J., Kensinger, J., 1990. Corporate research and development expenditures and share values. Journal of Financial Economics 26, 255–276.

Chan, L.K.C., Lakonishok, J., Sougiannis, T., 2001. The stock market valuation of research and development expenditures. Journal of Finance LVI (6), 2431–2456.
Chauvin, K.W., Hirschey, M., 1993. Advertising, R&D expenditures and the market value of the firm. Financial Management 22, 128–140.
Chauvin, K.W., Hirschey, M., 1994. Goodwill, profitability, and the market value of the firm. Journal of Accounting and Public Policy 13, 159–180.
Chen, M.H., Shao, Q.M., Ibrahim, J.G., 2000. Monte Carlo Methods in Bayesian Computation. Springer-Verlag, New York.
Dukes, R., 1976. An investigation of the effects of expensing research and development costs on security prices. In: Schiff, M., Sorter, G. (Eds.), Proceedings of the Conference on Topical Research in Accounting. New York University Press, New York.
Eden, B., Jovanovic, B., 1994. Asymmetric information and the excess volatility of stock prices. Economic Inquiry 32 (2), 228–235 (l).
Engle, R., Kroner, B., 1995. Multivariate simultaneous generalized ARCH. Econometric Theory 11, 122–150.
Gammerman, D., 1997. Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference. Chapman & Hall, London.
Gelfand, A.E., 1996. Model determination using sampling based methods. In: Gilks, W.R., Richardson, S., Spiegelhalter, D.J. (Eds.), Markov Chain Monte Carlo in Practice. Chapman & Hall, Boca Raton, pp. 145–161.
Green, J.P., Stark, A.W., Thomas, H.M., 1996. UK evidence on the market valuation of research and development expenditures. Journal of Business Finance and Accounting 23 (2), 191–216.
Green, P., 1995. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. Biometrika 82, 711–732.
Griliches, Z., 1981. Market value, R&D, and patents. Economics Letters 7, 183–187.
Hall, B.H., 1999. Innovation and market value. National Bureau of Economic Research Working Paper, vol. 6984.
Hall, B.H., Jaffe, A.B., Trajtenberg, M., 2001. The NBER patent citations data file: lessons, insights and methodological tools. National Bureau of Economic Research Working Paper, vol. 8498.
Hendricks, D., 1996. Evaluation of Value-at-Risk models using historical data. Federal Reserve Bank of New York Economic Policy Review April, pp. 39–69.
Hirschey, M., Spencer, S., 1992. Size effects in the market valuation of fundamental factors. Financial Analysts Journal 48 (2), 91–95.
Hirschey, M., Weygandt, J., 1985. Amortization policy for advertising and research and development expenditures. Journal of Accounting Research 23, 326–335.
Hollifield, B., 2002. Comment on stock volatility in the new millennium: how wacky is Nasdaq? Journal of Monetary Economics 49, 27–30.
Jegadeesh, N., Titman, S., 1995. Overreaction, delayed reaction, and contrarian profits. Review of Financial Studies 8 (4), 973–993.
Jorion, P., 2001. Value-at-Risk: The New Benchmark for Controlling Market Risk. McGraw-Hill, Chicago.
Larson, S., Madura, J., 2003. What drives stock price behavior following extreme one-day returns. Journal of Financial Research 26 (1), 113–127.
Osiewalski, J., Pipień, M., 2004. Bayesian comparison of bivariate ARCH-type models for the main exchange rates in Poland. Journal of Econometrics 123, 371–391.
Polasek, W., Jin, S., 1997. GARCH models with outliers. In: Klar, R., Opitz, O. (Eds.), Classification and Knowledge Organization: Proceedings of the 20th Annual Conference of the Gesellschaft fur Klassifikation e.V., University of Freiburg. Springer-Verlag, Berlin.
Raftery, A.E., 1996. Hypothesis testing and model selection. In: Gilks, W.R., Richardson, S., Spiegelhalter, D.J. (Eds.), Markov Chain Monte Carlo in Practice. Chapman & Hall, Boca Raton, pp. 163–187.
Shane, H., Klock, M., 1997. The relation between patent citations and Tobin's Q in the semiconductor industry. Review of Quantitative Finance and Accounting 9, 131–146.
Sougiannis, T., 1994. The accounting based valuation of corporate R&D. The Accounting Review 69 (1), 44–68.
Stark, A.W., Thomas, H.M., 1998. On the empirical relationship between market value and residual income in the U.K. Management Accounting Research 9, 445–460.
Schwert, G.W., 2002. Stock volatility in the new millennium: how wacky is Nasdaq? Journal of Monetary Economics 49, 3–26.
Timmerman, A.G., 1993. How learning in financial markets generates excess volatility and predictability in stock prices. Quarterly Journal of Economics 108, 1135–1145.
Troughton, P.T., Godsill, S.J., 1998. A reversible jump sampler for autoregressive time series, employing full conditionals to achieve efficient model space moves. Technical Report CUED/F-INFENG/TR.304. Cambridge University Engineering Department.
Veronesi, P., 1999. Stock market overreaction to bad news in good times: a rational expectations equilibrium model. Review of Financial Studies 12 (5), 975–1007.
Veronesi, P., 2000. How does information quality affect stock returns? Journal of Finance 55 (2), 807–837.
Vrontos, D., Dellaportas, P., Politis, D.N., 2000. Full Bayesian inference for GARCH and EGARCH models. Journal of Business and Economic Statistics 18 (2), 187–198.
Vrontos, D., Dellaportas, P., Politis, D.N., 2003a. Inference for some multivariate ARCH and GARCH models. Journal of Forecasting 22, 427–446.
Vrontos, D., Dellaportas, P., Politis, D.N., 2003b. A full-factor multi-variate GARCH model. Econometrics Journal 6, 312–334.
Weeds, H., 2002. Strategic delay in a real options model of R&D competition. Review of Economic Studies 69 (3), 729–747.
Woolridge, R., 1988. Competitive decline and corporate restructuring: is a myopic stock market to blame? Journal of Applied Corporate Finance 1, 26–36.