

Threshold models for Value at Risk Estimation

Modelli a Soglia per la Stima del Valore a Rischio

La Rocca M., Storti G., Vitale C.

Dipartimento di Scienze Economiche e Statistiche

University of Salerno, Italy, vitale@unina.it

Riassunto: Si affronta il problema della stima e la costruzione di un intervallo di confidenza k passi in avanti del Valore a Rischio (VaR) di un attivo o di un portafoglio finanziario. La stima e la previsione dei livelli e della volatilità è effettuata utilizzando un particolare modello a soglia con due regimi. S'illustra come questo sia diverso e più flessibile da quello usato da Tong (1978) per modellizzare la media condizionata. Si analizzano le proprietà probabilistiche e statistiche dei due approcci e si mostra come il modello proposto presenti una serie di proprietà desiderabili, quali facilità d'uso e generalità. Sono, infine, presentati i risultati di un'applicazione a dati reali su un portafoglio di 15 attivi.

Keywords: Value at Risk, volatility, portfolio, forecasts, threshold models.

1. Introduction

Given the returns process $\{r_t\}$, let $F_t(r_t|\mathcal{I}_{t-1})$ be the cumulative distribution function (cdf) of r_t conditioned by the information set available at time $t-1$, $\mathcal{I}_{t-1}$. If $F_t(\cdot)$ and the associated probability density function (pdf) $f_t(r_t|\mathcal{I}_{t-1})$ are continuous, for a given $\alpha \in (0, 1)$, there is only one $r_t(\alpha)$ such that

$$F_t^{-1}(r_t(\alpha)|\mathcal{I}_{t-1}) = \alpha \quad (1)$$

and the Value at Risk (VaR) is given by $\text{VaR}_{t-1}(1 - \alpha) = r_t(\alpha)$. Actually we don't know the process generating returns, but we can only observe a time series of finite length N that can be used for estimating VaR. The popularity of this risk measure among academic researchers and practitioners is mainly due to two reasons. First, VaR is easily interpretable: it gives the maximum loss expected at time t , which will be exceeded with probability α . The second reason is related to the Basle Agreement on Banking Supervision (Basle Committee on Banking Supervision, 2003), which has established rigorous guidelines for risk estimation and management.

On a theoretical ground, it has been shown that in general VaR does not satisfy the sub-additivity property and, hence, it is not a *coherent* risk measure (Artzner et al., 1999). In the literature, several different coherent risk measures have been proposed as an alternative to VaR (Dowd and Blake, 2006). However, from a statistical perspective, most of these can still be expressed as functions of percentiles implying the same difficulties encountered for VaR estimation.

Furthermore, if $F_t(r_t|\mathcal{S}_{t-1})$ belongs to the family of elliptical distributions, like the Normal and Student's T, VaR is a coherent risk measure. Percentiles, as, in general, any order statistic, are not smooth and their estimation implies some approximation problems even in the case of iid observations. It is known that if $(x_1, x_2, \dots, x_N)$ is an iid sample from $X \sim F(x)$, with F and the associated pdf $f(x)$ being continuous, if x_α is the α percentile of $F(x)$ and $\hat{x}_\alpha$ is its sample counterpart computed from its empirical cdf, it follows that

$$\sqrt{N}(\hat{x}_\alpha - x_\alpha) \xrightarrow{L} N(0, \alpha(1-\alpha)/f^2(x_\alpha)) ; \quad \alpha \in (0, 1) \quad (2)$$

and the error in the estimation of x_α by $\hat{x}_\alpha$ is $O(N^{-1/2})$. For finite N , the quality of this approximation deteriorates as α moves towards the extremes of the interval $(0, 1)$. Also it is characterized by a positive bias increasing as α decreasing and as the tails of $f(x)$ become heavier (Inui et al., 2005). These drawbacks affect also the estimation of VaR for $\alpha \leq 0.05$ since the (conditional) distributions of financial returns are typically characterized by heavy tails. In order to improve the properties of $\hat{x}_\alpha$ several procedures have been proposed based on the use of approximation methods, such as the Cornish–Fisher expansion, or resampling methods. However, we observe that the Cornish–Fisher approximation reduces the error if $f(x)$ is known while the performance of resampling methods is not as satisfactory as in the case of smooth statistics (Ho and Lee, 2005). On the other hand financial returns are far from being iid:

$$r_t = E(r_t|\mathcal{S}_{t-1}) + [\text{var}(r_t|\mathcal{S}_{t-1})]^{1/2} \varepsilon_t \equiv \mu_t + \sigma_t \varepsilon_t \quad (3)$$

with $\varepsilon_t \stackrel{\text{iid}}{\sim} D(0, 1)$. The associated VaR is then given by

$$\text{VaR}_{t-1}(1; \alpha) = E(r_t|\mathcal{S}_{t-1}) + \varepsilon_\alpha [\text{var}(r_t|\mathcal{S}_{t-1})]^{1/2} \quad (4)$$

where $\varepsilon_\alpha = F_\varepsilon^{-1}(\alpha)$ and $F(\varepsilon_t)$ is the cdf of ε_t . Equation (4) refers to a single asset but it can be typically generalized to the case of a portfolio with $m \geq 1$ assets. Let r_t^j be the return of the asset j at time t and v_{t-1}^j its portfolio weight at time t , computed at $t-1$,

$$\mathbf{v}_{t-1}' = (v_{t-1}^1, \dots, v_{t-1}^m), \quad v_{t-1}^j \geq 0; \quad \mathbf{r}_t = (r_t^1, \dots, r_t^m); \quad \mathbf{1}' = (1, 1, \dots, 1); \quad \mathbf{v}_{t-1}' \mathbf{1} = 1,$$

it follows $r_t = \mathbf{v}_{t-1}' \mathbf{r}_t$ and

$$\text{VaR}_{t-1}(1; \alpha) = \mathbf{v}_{t-1}' E(\mathbf{r}_t|\mathcal{S}_{t-1}) + \varepsilon_\alpha [\mathbf{v}_{t-1}' \text{var}(\mathbf{r}_t|\mathcal{S}_{t-1}) \mathbf{v}_{t-1}]^{1/2} \quad (5)$$

If the modeller is interested in risk management and not only in its numerical estimation, it is necessary to analyze $\text{var}(\mathbf{r}_t|\mathcal{S}_{t-1})$ in a multivariate setting which

requires the estimation of the inter-asset correlations.

Formulas (4) and (5) can be estimated in a parametric as well as in a non-parametric setting. In both cases, efficient estimates of the time-varying portfolio conditioned means and variances are required in order to reconstruct the unobservable standardized innovations sequence ε_t and estimate VaR for the relevant forecast horizon ($k \geq 1$). In section 2 we will deal with problems related to univariate VaR estimation using ARMA-GARCH models combined with appropriate distributional assumptions on ε_t . In order to simultaneously capture asymmetries in the levels and volatility, in section 3 we suggest a particular threshold model (Amendola et al., 2006), which is obtained through a modification of the structure originally proposed by Tong (1978) for conditioned mean modelling. The probabilistic properties of the model are also analyzed. Section 4 presents the results of an application to a portfolio of 15 assets. Section 5 briefly discusses the main problems required by the generalization of the proposed model to a multivariate setting and provides conclusions.

2. VaR estimation in ARMA-GARCH models

Let us assume that the level of returns can be modelled as an ARMA and that volatility follows a GARCH model of some order (Engle, 1982; Bollerslev, 1986):

$$r_t = \frac{\theta_o}{\phi(1)} + \frac{\theta(B)}{\phi(B)} e_t; \quad e_t = \sigma_t \varepsilon_t; \quad \sigma_t^2 = \frac{\omega}{\beta(1)} + \frac{\alpha(B)}{\beta(B)} e_t^2;$$

assuming that the usual conditions for stationarity, invertibility and positivity of the conditional variance hold. Following standard practice, by direct equation (4), we approximate k-steps ahead VaR through the following scaling rule

$$\text{VaR}_N(k; \alpha) = E(r_{N+k} | \mathcal{I}_N) + \varepsilon_\alpha [\text{var}(r_{N+k} | \mathcal{I}_N)]^{1/2} \quad (6)$$

where $(r_{N+k} | \mathcal{I}_N) = \sum_{h=1}^k (r_{N+h} | \mathcal{I}_N)$; letting $r_N(h)$ be the predictor of r_{N+h} and $\sigma_N^2(h)$ that of σ_{N+h}^2 , conditional on information at time $N+h-1$, it follows that

$$\begin{cases} E(r_{N+k} | \mathcal{I}_N) = \sum_{h=1}^k r_N(h) \\ \text{var}(r_{N+k} | \mathcal{I}_N) = \text{var}\left(\sum_{h=1}^k e_N(h)\right) = \sigma_N^2(k) + (1 + \psi_1)^2 \sigma_N^2(k-1) + \dots + \left(\sum_{j=0}^{k-1} \psi_j\right)^2 \sigma_N^2(1) \end{cases} \quad (7)$$

$$e_N(h) = r_{N+h} - r_N(h) = \sum_{j=0}^{h-1} \psi_j e_{N+h-j}; \quad 1 + \psi_1 B + \psi_2 B^2 + \psi_3 B^3 + \dots = \frac{\theta(B)}{\phi(B)}; \quad \psi_o = 1.$$

In usual applications, the autocorrelation structure of the levels is not very persistent and

a low order ARMA model is usually required. On the other hand, for volatility prediction, it has been observed how in many cases a GARCH (1,1) model is able to produce sufficiently accurate results. These considerations lead us to focus on the following examples.

1. For $r_t = \theta_o + e_t$; $e_t = \sigma_t \varepsilon_t$; $\sigma_t^2 = \omega + \alpha e_{t-1}^2 + \beta \sigma_{t-1}^2$, we have $r_N(h) = \theta_o$;

$$E(r_{N+k} | \mathfrak{I}_N) = k\theta_o; \quad \text{var}(r_{N+k} | \mathfrak{I}_N) = \sum_{j=1}^k \sigma_N^2(j);$$

$$\sigma_N^2(j) = \begin{cases} \omega + \alpha e_N^2 + \beta \sigma_N^2; & \text{if } j = 1 \\ \omega + (\alpha + \beta) \sigma_N^2(j-1); & \text{if } j > 1. \end{cases}$$

2. For $r_t = \theta_o + e_t - \theta e_{t-1}$; $e_t = \sigma_t \varepsilon_t$; $\sigma_t^2 = \omega + \alpha e_{t-1}^2 + \beta \sigma_{t-1}^2$, it is

$$E(r_{N+k} | \mathfrak{I}_N) = \begin{cases} \theta_o + \theta e_N; & \text{if } k = 1 \\ k\theta_o; & \text{if } k > 1 \end{cases}; \quad \text{var}(r_{N+k} | \mathfrak{I}_N) = \theta^2 \sigma_N^2(k-1) + \sum_{j=1}^k \sigma_N^2(j).$$

If we take (2) into account, it is possible to derive an approximate¹ asymptotic confidence interval for $\text{VaR}_N(k; \alpha)$ at the 95%:

$$\hat{E}(r_{N+k} | \mathfrak{I}_N) + \left[\hat{\varepsilon}_\alpha \mp 2 \left(\frac{\alpha(1-\alpha)}{N f^2(\hat{\varepsilon}_\alpha)} \right)^{1/2} \right] [\hat{\text{var}}(r_{N+k} | \mathfrak{I}_N)]^{1/2} \quad (8)$$

where $f(\cdot)$ is the pdf of the standardized innovation ε_t , while $\hat{\varepsilon}_\alpha$ is the order α percentile of the innovation distributions. The class of GARCH models has been further generalized to capture asymmetric (*leverage*) effects in the conditional variance dynamics. In particular, the EGARCH(1, 1) (Nelson, 1991) accounts for asymmetric dynamics by means of the following recursion:

$$\log(\sigma_t^2) = \omega + \alpha \frac{|e_{t-1}|}{\sigma_{t-1}} + \beta \log(\sigma_{t-1}^2) + \gamma \frac{e_{t-1}}{\sigma_{t-1}}.$$

3. VaR estimation in Threshold Models

Several papers have documented the presence of asymmetry in financial time series (So et al., 2002, among the others). The class of threshold models proposed by Tong (1978) offer a useful tool for modelling asymmetric dynamics. These models are very flexible

¹ The approximation is due to the variability in the estimation of the conditional mean and variance of aggregated returns. For forecast horizons $k > 1$, the scaling rule in (6) is evidently an approximation of the unknown k -steps ahead VaR.

and can allow several features including asymmetry but also leptokurtosis, limit cycles and other features. We focus on the class of two regimes SETARMA models:

$$r_t = \frac{\theta_o^{(2)} + (\theta_o^{(1)} - \theta_o^{(2)})I_{t-d}}{\phi^{(2)}(1) + (\phi^{(1)}(1) - \phi^{(2)}(1))I_{t-d}} + \frac{\theta^{(2)}(B) + (\theta^{(1)}(B) - \theta^{(2)}(B))I_{t-d}}{\phi^{(2)}(B) + (\phi^{(1)}(B) - \phi^{(2)}(B))I_{t-d}} e_t \quad (9)$$

where $I_{t-d} = 0$ if $r_{t-d} < c$ and $I_{t-d} = 1$ elsewhere, $d \geq 1$. The generalization to a higher number of regimes is straightforward. The first results on the probabilistic properties of (9), with $q_1=q_2=0$, can be found in Chan and Tong (1985) who show that a sufficient condition for the weak stationarity of r_t is given by

$$\max \left(\sum_{i=1}^p \phi_i^{(1)}, \sum_{i=1}^p \phi_i^{(2)} \right) < 1; \quad p = \max(p_1, p_2) \quad (10)$$

Lim (1999) shows that this condition still holds even for $q_i > 0$; furthermore, if the order m moment of e_t exists, the same holds for r_t . Differently, if

$$\sum_{i=1}^p \max(|\phi_i^{(1)}|, |\phi_i^{(2)}|) < 1 \quad (11)$$

r_t is strictly stationary and ergodic while the invertibility of these models has been investigated by Ling and Tong (2005).

The sufficient conditions in (10) and (11) are very restrictive and they are often not verified. Similarly, in order to reproduce asymmetries and limit cycles in the conditional variance we use

$$\sigma_t^2 = \frac{\omega^{(2)} + (\omega^{(1)} - \omega^{(2)})I_{t-d}}{\beta^{(2)}(1) + (\beta^{(1)}(1) - \beta^{(2)}(1))I_{t-d}} + \frac{\alpha^{(2)}(B) + (\alpha^{(1)}(B) - \alpha^{(2)}(B))I_{t-d}}{\beta^{(2)}(B) + (\beta^{(1)}(B) - \beta^{(2)}(B))I_{t-d}} e_t^2. \quad (12)$$

Ling (1999) proves, under standard regularity conditions, that if (11) is satisfied and

$$\sum_{i=1}^s \max(\alpha_i^{(1)}, \alpha_i^{(2)}) + \sum_{i=1}^s \max(\beta_i^{(1)}, \beta_i^{(2)}) < 1 \quad (13)$$

with $s = \max(s_1, s_2)$ and $u = \max(u_1, u_2)$, then (r_t, e_t) is stationary and ergodic. The condition (13) is a very restrictive one. Furthermore, the simultaneous estimation of the model parameters implies some difficulties (Chen et al., 2005).

In this paper we consider a different threshold model proposed in Amendola et al (2006) which allows to analyze separately the two regimes simplifying (9) and (12). It must be recognized that the complex structure of (9) and (12) is due to the presence of the binary process I_t in the denominator. If we use alternative formulations separating the dynamics of the two regimes, the model becomes much more tractable, as far as its statistical properties are concerned with the possibility of applying most of the results holding for the traditional ARMA-GARCH.

Let us consider:

$$r_t = \frac{\theta_0^{(2)}}{\phi^{(2)}(1)} + \left(\frac{\theta_0^{(1)}}{\phi^{(1)}(1)} - \frac{\theta_0^{(2)}}{\phi^{(2)}(1)} \right) I_{t-d} + \left[\frac{\theta^{(2)}(B)}{\phi^{(2)}(B)} + \left(\frac{\theta^{(1)}(B)}{\phi^{(1)}(1)} - \frac{\theta^{(2)}(B)}{\phi^{(2)}(1)} \right) I_{t-d} \right] e_t$$

with $e_t = e_{2t_2} + (e_{1t_1} - e_{2t_2})I_{t-d}$, that for $e_{1t_1} = e_{2t_2} = e_t$ and $\phi^{(1)}(B) = \phi^{(2)}(B) = 1$ becomes the SETMA model by Ling and Tong (2005). For the conditional variance we set:

$$e_t = [\sigma_{2t_2} + (\sigma_{1t_1} - \sigma_{2t_2})I_{t-d}] [\varepsilon_{t_e}^{(2)} + (\varepsilon_{t_e}^{(1)} - \varepsilon_{t_e}^{(2)})I_{t-d}] = \sigma_t \varepsilon_t$$

$$\sigma_t^2 = \frac{\omega^{(2)}}{\beta^{(2)}(1)} + \left(\frac{\omega^{(1)}}{\beta^{(1)}(1)} - \frac{\omega^{(2)}}{\beta^{(2)}(1)} \right) I_{t-d} + \left[\frac{\alpha^{(2)}(B)}{\beta^{(2)}(B)} + \left(\frac{\alpha^{(1)}(B)}{\beta^{(1)}(B)} - \frac{\alpha^{(2)}(B)}{\beta^{(2)}(B)} \right) I_{t-d} \right] e_t^2$$

with $\varepsilon_{t_i}^{(i)} \stackrel{iid}{\sim} D_i(0, 1)$, $i=1, 2$; $\varepsilon_{t_1}^{(1)}$ and $\varepsilon_{t_2}^{(2)}$ mutually independent; $t = t_1 + (t_2 - t_1)I_{t-d}$. It

can be easily verified that $r_t = r_{1t} + (r_{2t} - r_{1t})I_{t-d}$; $\sigma_t^2 = \sigma_{1t}^2 + (\sigma_{2t}^2 - \sigma_{1t}^2)I_{t-d}$. Both the levels and volatilities evolve independently within each regime and their dynamics are not influenced by the regime structure of the overall model. This implies that, if the regimes are stationary and ergodic then the overall process is stationary and ergodic. This condition is less restrictive than (10) e (13). The price to pay for the simplicity and flexibility of this model is that the observations occurring in the two regimes are unequally spaced and the sampling frequency within each regime varies randomly as a function of r_{t-d} . In section 4 we will further investigate this issue and its possible impact on the applicability of the proposed model.

Under the assumption that the error in regime $i = 1, 2$ is distributed as a standardized Student's distribution with g_i degrees of freedom, letting $\mathbf{r}$ be the observed returns vector, L the overall log-likelihood, L_i the log-likelihood function associated to regime $i = 1, 2$, it can be immediately verified that

$$\begin{aligned} -2\log(L(\theta_1, \theta_2|\mathbf{r})) &= -2\log(L_1(\theta_1|\mathbf{r})) - 2\log(L_2(\theta_2|\mathbf{r})); \\ -2\log(L_i(\theta_i|\mathbf{r})) &= (N_i + 1 - s_i) \log \left[(g_i - 2) \pi \left(\frac{\Gamma(g_i/2)}{\Gamma((g_i + 1)/2)} \right)^2 \right] + \\ &+ \sum_{t=s_i}^{N_i} \left[\log \left(1 + \frac{(r_t^{(i)} - \mu_t^{(i)})^2}{(g_i - 2)\sigma_{it}^2} \right) - \log(\sigma_{it}^2) \right] \end{aligned} \quad (14)$$

where $\mu_t^{(i)}$ is the conditional mean of regime $i = 1, 2$; $\max(p_1, p_2, q_1, q_2, 1) = s_1 + s_2 + 1$; $N_1 + N_2 = N - d$. The minimization of (14), which can be achieved by performing two separate optimizations within each regime, provides an estimate of model parameters. In conclusion, if the conditions for the stationarity, invertibility and for the positivity of the

conditional variance ($\sigma_{it} > 0$) hold for both regimes, it is possible to generate h ($\leq d$) steps ahead optimal forecasts by the formulas:

$$\begin{aligned} E(r_{N+h}|\mathcal{S}_N) &= E(r_{1N+h}|\mathcal{S}_N) + [E(r_{2N+h}|\mathcal{S}_N) - E(r_{1N+h}|\mathcal{S}_N)] I_{N+h-1} \\ \text{var}(r_{N+h}|\mathcal{S}_N) &= \text{var}(r_{1N+h}|\mathcal{S}_N) + [\text{var}(r_{2N+h}|\mathcal{S}_N) - \text{var}(r_{1N+h}|\mathcal{S}_N)] I_{N+h-1} \end{aligned}$$

given that, in this case, the value of I_{N+h} is known since r_{N+h-d} is observed. Differently, for $h > d$, we replace I_{N+h} with its expectation p . In practice the parameters of the two models involved in the forecasts are unknown and it is necessary to replace them with their maximum likelihood estimates derived from the minimization of (14) and with $\hat{p} = N_1 / N$. The estimated errors $\hat{\epsilon}_t = \hat{\epsilon}_{1t} + (\hat{\epsilon}_{2t} - \hat{\epsilon}_{1t})I_{t-1}$ can then be used to generate k -steps ahead prediction of the conditional mean and variance and, hence, of the associated VaR

$$\hat{\text{VaR}}_N(k; \alpha) = \hat{E}(r_{N+k}|\mathcal{S}_N) + \hat{\epsilon}_\alpha [\hat{\text{var}}(r_{N+k}|\mathcal{S}_N)]^{1/2} \quad (15)$$

with $\hat{\epsilon}_\alpha = \hat{\epsilon}_{1\alpha} + (\hat{\epsilon}_{2\alpha} - \hat{\epsilon}_{1\alpha})I_{N+k-1}$ and $\hat{\epsilon}_{i\alpha}$ being the percentile of the standardized innovations distribution in regime $i = 1, 2$ which is assumed to be distributed according to a Student's T with $\hat{g}_i$ degrees of freedom. If the model used for estimating levels and volatility is well specified, an approximate asymptotic 95% confidence interval for VaR, by equation (2), is given by:

$$\hat{E}(r_{N+k}|\mathcal{S}_N) + \left[\hat{\epsilon}_\alpha \mp 2 \left(\frac{\alpha(1-\alpha)}{N^* f^2(\hat{\epsilon}_\alpha)} \right)^{1/2} \right] [\hat{\text{var}}(r_{N+k}|\mathcal{S}_N)]^{1/2} \quad (11)$$

where $N^* f^2(\hat{\epsilon}_\alpha) = N_1 f_1^2(\hat{\epsilon}_{1\alpha}) + (N_2 f_2^2(\hat{\epsilon}_{2\alpha}) - N_1 f_1^2(\hat{\epsilon}_{1\alpha})) I_{N+k-1}$ and $f_i(\hat{\epsilon}_{i\alpha})$ is the pdf of a Student's T in regime i with $\hat{g}_i$ degrees of freedom evaluated at $\hat{\epsilon}_{i\alpha}$. Alternatively, the confidence interval for VaR can be generated via simulation resorting to an adequately chosen Bootstrap technique. The estimated model can be validated by the usual diagnostic tests and in particular by checking that the autocorrelations of $(\epsilon_{it}|\mathcal{S}_{t-1})$ and $(\epsilon_{it}^2|\mathcal{S}_{t-1})$ are compatible with the iid assumption.

4. An application to real data

In this section, based on daily historical data, we build a virtual portfolio composed of 15 financial assets and predict the associated VaR up to a ten days ahead forecast horizon using the best GARCH type model. The results are compared with those obtained by means of a threshold model for conditional mean and variance with the residuals distributed according to a Student's T distribution with different degrees of freedom parameter within each regime.

The data we use for this application are the adjusted closing prices observed for each

asset on the Milan Exchange Stock Market from 07/01/2002 to 15/02/2007 for a total of $N=1328$ observations (fig. 1, left panel). The assets we consider are: Alitalia, Espresso Gruppo, Intesa San Paolo, Juventus, Luxottica, Marzotto, Mediaset spa, Generali ASS, Bastogi, Benetton Grup, Unipol Priv, Campari, Eni, Fiat, Gemina. The portfolio weights are assumed to be time varying and proportional to the daily traded volume (fig. 1, right panel). Returns are calculated as the first difference of the log-transformed prices (fig. 2, left panel). The histogram (fig. 2, right panel) reveals a remarkable kurtosis and a slight negative skewness. The sample autocorrelation functions (SACF) of returns and squared returns have been plotted in fig. 3. For the levels, the SACF supports the use of a low order MA model while the dependence structure of squared returns is much more persistent. We used the Schwartz criterion to identify an MA(1)-EGARCH(1,1) model whose parameters have been estimated by Gaussian Quasi Maximum Likelihood (values in brackets are asymptotic standard errors):

$$r_t = 0.0008 + e_t - 0.0730 e_{t-1}; e_t = \sigma_t \varepsilon_t; \varepsilon_t \sim N(0, 1); \varepsilon_{.05} = -1.6449; f(\varepsilon_{.05}) = 0.1031$$

(.0027) (.0270)

$$\log(\sigma_t^2) = -0.4303 + 0.1945 \frac{|e_{t-1}|}{\sigma_{t-1}} - 0.0479 \frac{e_{t-1}}{\sigma_{t-1}} + 0.9685 \log(\sigma_{t-1}^2)$$

(.078) (.023) (.01314) (.0075)

The estimated model has been used to predict 5% VaR up to ten days ahead. The forecasts with the associated 95% prediction interval have been reported in fig. 4 (left). The predicted VaR ranges from 1.78% (1 day) to 4.90% (10 days). We compare these estimates with those obtained from the proposed threshold model assuming an MA(1), for the conditional mean, and a GARCH(1,1) model, for the conditional variance, within each regime. The time plots of the returns falling within each of the two regimes, with $N_1=611$ and $N_2= 716$, have been reported in fig. 5. The differences between the two regimes are more evident if we consider the associated histograms (fig. 6) and SACF of squared returns within each regime (fig. 7, left panel). The SACF patterns show how the volatility tends to be more persistent in the first regime rather than in the second. The estimated model parameters for $d = 1$, $c = 0$ are:

$$r_{1t_1} = 0.0017 + e_{1t_1}; e_{1t_1} = \sigma_{1t_1} \varepsilon_{1t_1}; \varepsilon_{1t_1} \sim N(0, 1);$$

(.0004)

$$\sigma_{1t_1}^2 = 0.0000118 + 0.196 e_{1t_1-1}^2 + 0.741 \sigma_{1t_1-1}^2$$

(.0000034) (.044) (.052)

$$r_{2t_2} = e_{2t_2} - 0.063 e_{2t_2-1}; e_{2t_2} = \sigma_{2t_2} \varepsilon_{2t_2}; \varepsilon_{2t_2} \sim \left(\frac{g-2}{g} \right)^{1/2} t_{(g)}; g = 6.79$$

(.037) (1.75)

$$\sigma_{2t_2}^2 = 0.0000053 + 0.103 e_{2t_2-1}^2 + 0.861 \sigma_{2t_2-1}^2$$

(.0000026) (.029) (.038)

$$\varepsilon_{2,.05} = ((g-2)/g)^{.5} t_{.05} = -1.59871; f(\varepsilon_{2,.05}) = (g/(g-2))^{.5} f(t_{.05}) = 0.0865873$$

confirming our prior belief on the structural differences between the two regimes. This is evident from the histograms of the residuals within each regime, which are

characterized by substantially different tail behaviours. In fig. 7 (right panel) we have represented the frequency distribution of the run lengths within the negative and positive regimes, respectively. In the case of equally spaced observations, this frequency should be equal to 1 for a run length equal to 1 and 0 elsewhere. A practical way of measuring the distance from this benchmark is to use the following relative distance measure, which compares the observed cumulated frequencies with their theoretical counterparts in the case of equally spaced observations

$$0 \leq v_i = 1 - \sum_{j=1}^{H-1} F_j^{(i)} / (H-1) \leq 1, i=1, 2$$

where H is the maximum observed run length and $F_j^{(i)}$ are the cumulated frequencies of run lengths within regimes $i=1, 2$. In our case it is $H=9$, $v_1 = 0.138$, $v_2 = 0.1$ which are very close to 0 implying that the information loss deriving from using regimes with unequally spaced observations is negligible with respect to the case of equally spaced data points. Again, the estimated threshold model was used to predict VaR up to 10 days ahead (fig. 4, right panel) starting from $I_N = 0$ and using $\hat{p} = 0.4604$ to compute the predicted values of $\text{VaR}_N(k, \alpha)$ for $k \geq 2$. The range of variation of VaR is similar to what observed for the MA-EGARCH model, going from a minimum of 1.7% (1 day) to a maximum loss of 4.6% (10 days). At the same time, however, the threshold model allows for a deeper interpretation of the market dynamics allowing to condition the term structure of risk on the state of the market.

5. Concluding remarks and further developments

The above-discussed models can be used for the estimation of risk measures such as the VaR. In order to use these models for risk management purposes, however it is necessary to establish a link between the target risk measure and the set of portfolio weights $\mathbf{v}_{t-1}$. Such a link is missing in equation (4) while it is present in equation (5), which, however, requires full knowledge of the conditional variance matrix of portfolio returns Σ_t . This consideration motivates the use of multivariate conditional heteroskedastic models. In theory the threshold model discussed in section 4 could be easily extended to a multivariate setting specifying appropriate multivariate threshold models for both conditional mean and variances. On an operational ground, however, several problems arise. Multivariate conditional heteroskedastic models are affected by the curse of dimensionality since the number of parameters tends to increase rapidly, often in an exponential fashion, as m increases (Bollerslev, 1988). To limit the number of parameters to be estimated, it is then necessary to impose restrictive and, in many cases, untestable constraints on the model dynamics. The semi-parametric approach, known as GEWMA, is perhaps the most widely diffused solution among practitioners (see e.g. Longerstay, 1996). Other notable examples of parsimonious multivariate conditional heteroskedastic models include the CCC, proposed by Bollerslev (1990), the OGARCH, proposed by Alexander (2001), the DCC, proposed by Engle (2002) and the related ADCC model, proposed by Cappiello et al. (2002). Alternatively, several non-parametric approaches have been proposed. Nevertheless, even in this setting, some important computational difficulties and probabilistic problems, related to the properties

of the estimators, arise. The estimates obtained by non-parametric approaches are not as easily interpretable as their parametric counterparts. A critical review is provided by Fan (2005).

References

- Amendola A., Niglio M., Vitale C. (2006) The moments in threshold autoregressive moving average models, *Statistics & Probability Letters*, 76, 625-633.
- Alexander C. (2001) Orthogonal GARCH, in: *Mastering Risk*, Alexander C. (Ed), Financial Times – Prentice Hall, London, 2, 21-38.
- Artzner P., Delbaen F., Eber J-M, Heath D. (1998) Coherent measures of risk, *Mathematical Finance*, 9, 203-228.
- Basle Committee in Banking Supervision (2003) Overview of the new Basel capital accord. Consultative document, *Bank for International Settlements*, Basel
- Bollerslev T. (1986) Generalized autoregressive conditional heteroskedasticity, *J. of Econometrics*, 92, 307-327.
- Bollerslev T. (1990) Modelling the coherence in short-run nominal exchange rate: a multivariate generalized ARCH, *Review of Economics and Statistics*, 72, 488-505.
- Bollerslev T., Engle R., Wooldridge J. (1988) A capital asset pricing model with time varying covariances, *J. of Political Economy*, 96, 116-131.
- Cappiello L.R., Engle R., Sheppard K. (2002) Asymmetric dynamics in the conditional correlations of global equity and bond returns, *NYU, preprint*.
- Chan K.S., Tong H. (1985) On the use of the deterministic Lyapunov function for the ergodicity of stochastic difference equations, *Advances in Applied Probability*, 17, 666-678.
- Chen C.W.S., So M.K.P., Gerlach R.H. (2005) Assessing and testing for threshold nonlinearity in stock returns, *Australian N.Z. J. of Statistics*, 47, 473-488.
- Dowd K., Blake D. (2006) After VaR: the theory, estimation, and insurance applications of quantile – based risk measures, *The J. of Risk Insurance*, vol. 73, pp. 193-229.
- Engle R. (1982) Autoregressive conditional heteroskedasticity with estimates of variance of U.K. inflation, *Econometrica*, 50, 987-1008.
- Engle R. (2002) Dynamic conditional heteroskedasticity models, *J. of Business & Economic Statistics*, 20, 339-350.
- Fan J. (2005) A selective overview of nonparametric methods in financial econometrics (with discussion), *Statistical Science*, 20, 317-357.
- Ho Y.H.S., Lee S.M.S. (2005) Iterated smooth bootstrap confidence intervals for population quantiles, *The Annals of Statistics*, 33, 437-462.
- Inui k., Kijima M., Kitano A. (2005) VaR is subject to a significant positive bias, *Statistics & Probability Letters*, 72, 299-311.
- Li C.W., Li W.K. (1996) On a double threshold autoregressive conditional heteroskedastic time series model, *J. of Applied Econometrics*, 11, 253-274.
- Ling S. (1999) On the probabilistic properties of a double threshold ARMA conditional heteroskedastic model, *J. of Applied Probability*, 36, 688-705.
- Ling S., Li W.K. (1997) Fractional ARIMA-GARCH time series models, *J. of the American Statistical Association*, 92, 1184-1194.
- Ling S., Tong H. (2005) Testing for a linear MA model against threshold MA models, *The Annals of Statistics*, 33, 2529-2552.

- Longerstay J. (1996) *Introduction to RiskMetricsTM* – Technical Document – Fourth edition, Morgan Guaranteed Trust Company, New York.
- Nelson D.B. (1991) Conditional heteroskedasticity in asset returns: a new approach, *Econometrica*, 59, 347-370.
- So M.K.P., Li W.K., Lam K. (2002) A threshold stochastic volatility model, *J. of Forecasting*; 21, 473-500.
- Tong H. (1978) On a threshold model, in: *Pattern Recognition and Signal Processing*, Chen C.H. (Ed.), Amsterdam: Sijhoff and Noordhoff.

Appendix

Figure 1: Adjusted closing prices from 07/01/02 to 15/02/07 (left) and portfolio weights (right)

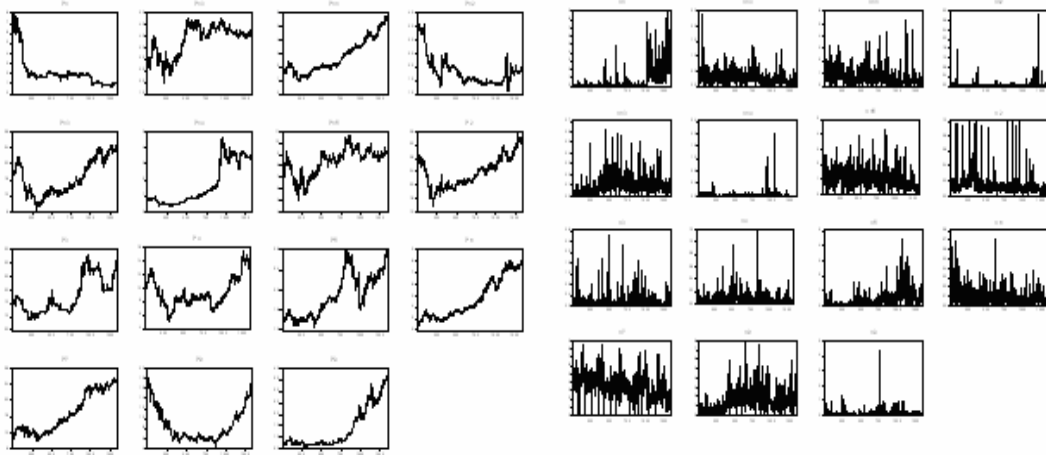


Figure 2: Daily portfolio returns: time plot and histogram

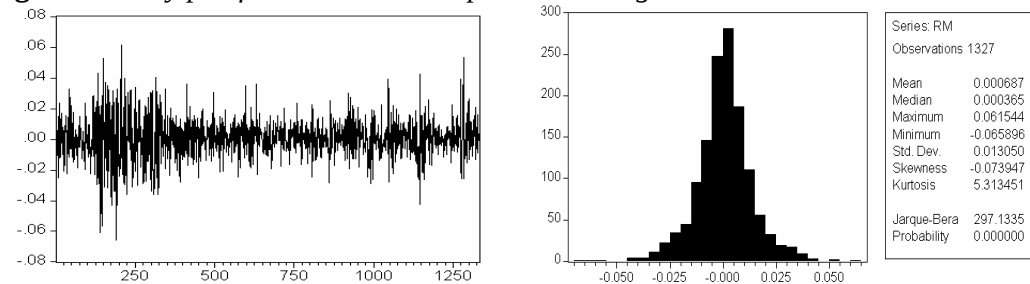


Figure 3 – SACF of raw (left) and squared (right) portfolio returns

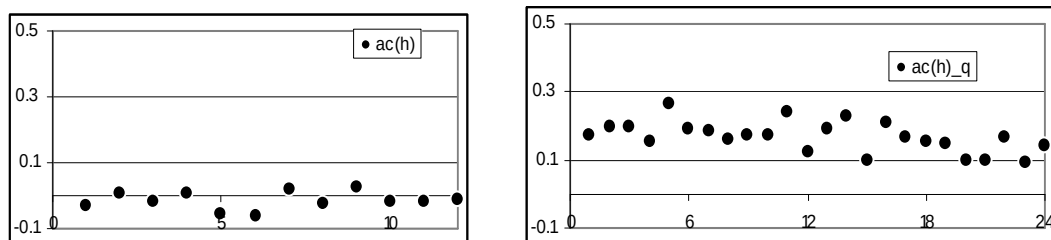


Figure 4: 5% VaR with asymptotic 95% confidence bands: EGARCH (left) and Threshold model (right)

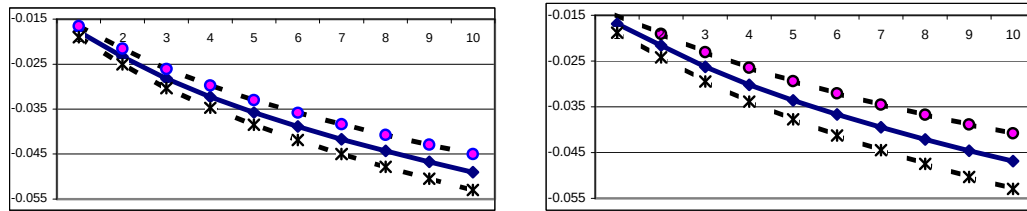


Figure 5: Returns within the first (left) and second regime (right).

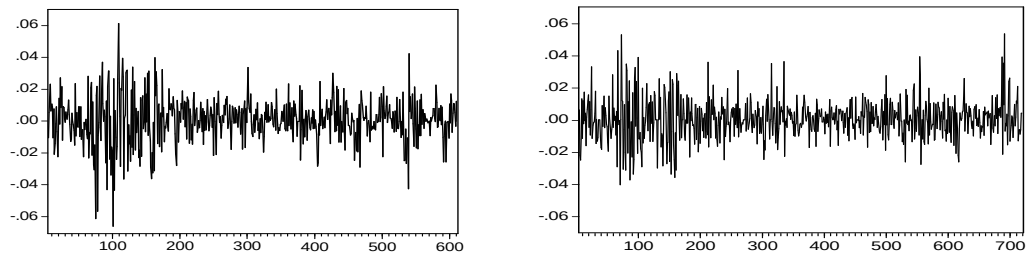


Figure 6: Histograms of first (left) and second regime returns (right).

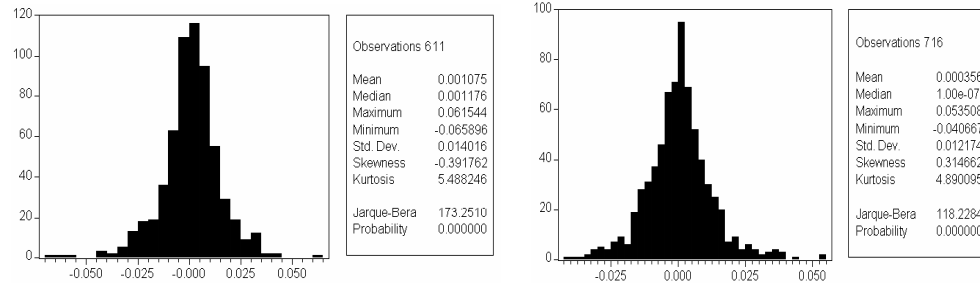


Figure 7: SACF of squared returns (left) and frequency distribution of run lengths (right) within each regime.

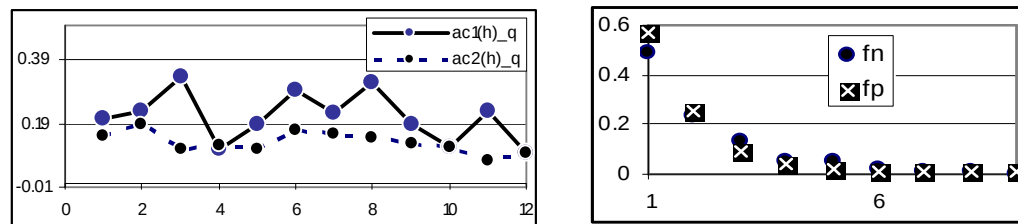


Figure 8: Histograms of first (left) and second regime (right) residuals.

