The estimation of Market VaR using Garch models and a heavy tail distributions

Tagliafichi Ricardo A.[*]

Av. Cordoba 1646 – 6to 209
1055 – Buenos Aires
Argentina

Telephone: 054-11-4811-3185
Facsimile: 054-11-4811-3185
E-mail: tagliafi@mbox.servicenet.com.ar

Topic 4

Abstract *(about 30 pages)*
The new rules presented in the documents published by Basel II, obliges to the actuaries to think in this new challenge that is to provide answers to the Institutions that need comply with these new rules based in three risk principal components, Market Risk, Credit Risk, and Operational Risk
When we need to develop an estimation of Market VaR, we must predict the probability of a maximum loss. To comply with this objective we must predict the volatility for the next period and the probability associated with this value.
This paper contains a development of Garch theory and the application of different, symmetric and asymmetric models, to predict the volatility of financial series, accompanied with the theory of Extreme Value Theory, EVT, and others heavy tails distributions to estimate the probability that the maximum loss may be occurred.
In the first part I analyze the presence of different Garch models in the returns of stocks in several markets and compare the same with other models in use. In the second part it is presented the estimation of the probability associated with the volatility forecasted. The methods used are the Kupiec estimation of the probability the Extreme Value Theory, and other heavy tails distributions as Weibull, Pareto, Pearson, etc. In the third part there are an estimation of several methods, for different series of returns. In the fourth part there are presented the results and the different methods used. Finally in the last part there are the conclusions arrived

Keywords: *Arch, Garch, Egarch, Tarch, EVT (Extreme Value Theory) Kupiec, Pareto, Heteroscedasticity, VaR (Value at Risk), Market Risk, Kolmogorov Smirnov Test, Anderson Darling Test, Basel II*

---

[*] *Actuary. Master in Business Administration. Professor of Statistics for Actuaries, Statistics for Business Administration and Financial Mathematics in the University of Buenos Aires, Faculty of Economics. Professor of financial Mathematics and Software Applied to finance in University of Palermo, Faculty of Economics and Business Administration. Professor of Investments and Financial Econometric, MBA in finance Palermo University. Member of Global Association of Risk Professionals. Independent Consultant*

Introduction

The actuarial work in the insurance companies and in the pension funds, and in extension in the government controllers of these entities, is the developing of models to estimate adequate mathematical and technical reserves to attend the demands by the sinister produced. These models estimate the probability that an event will occur and in other words is the risk covered by an insurance policy.

These reserves may be invested in several portfolios but these investments have three principal risks, the market value risk, the credit risk, and the operational risk. The Basel Committee on Banking Supervision in several documents concluding with the QIS-3 (Quantitative Impact Study) Technical Guidance document and the QIS-3 Instructions, both on October 2002, treats these risks.

In a brief commentary I try to develop the idea under each type of risk.

1) **The Operational risk**

The operational risk includes the risk of fraud, trading errors, legal end regulatory risk and so on. This risk is one of the risks more injurious, given that affects indirectly to the financial results. There isn't a concrete definition on this risk. It is more useful use the approach given in the documents issued by the Committee of Banks of Basel.

2) **The Credit risk**

Credit Risk is viewed as one component of market risk, but nevertheless focus here in the market risk associated with changes in the prices or rates of underlying traded documents over short times horizons. The credit risk includes the risk of default of a counterpart on a long-term contract. When we think in Credit Risk we think in how much money we will loss next year

3) **The market risk**

The market risk, in the business mentioned in the first part, is the most important risk. Market risk meaning the risk of unexpected changes in prices or rates. When we think in Market Risk we think in how much money we will loss tomorrow or the next week.

When we compute the portfolio position, we have the value of the investments that covers the mathematical and technical reserves, but the question is: how is the value of the investments tomorrow, the next week or the next month?

After 1995, the Financial Risk or the Value at Risk has introduced several models to estimate these values. The components of these models are the assets returns, the volatility, the time horizon to predict the worst loss possible, the probability desired, and the liquidity risk[1].

The behavior of assets returns was developed in a previous work "The Garch model and their application to the VaR" where I was demonstrated that the returns following these conditions:

1) Don't have normal probability distribution, due the results of goodness of fit test, using the Kolmogorov Smirnov statistic.

2) Are correlated, because the tested for the Q Statistic presented by Ljung Box, determine the presence of black noise in the autocorrelation function and in the partial autocorrelation function.

3) The hypotheses that $H_o : \boldsymbol{s}_n = \boldsymbol{s}_1 \sqrt{n}$ against $H_1 : \boldsymbol{s}_n \neq \boldsymbol{s}_1 \sqrt{n}$, studied by Young in 1971, was applied a several assets and conclude that not accept the null hypothesis.

---

[1] The liquidity risk is known as a relationship between the market and the risk produced to liquidate a position. This is produced when the position in an assets is greater than the quantity that trades the market

In base on these concepts it is very difficult to model an interval of confidence, to calculate de worst loss possible, or in other form to estimate the roof of the "x percent" of the minimum returns.

The series returns present values of skew ness and kurtosis that permit to reject the presence of normal distribution, but introduce what type of distribution covers the presence of extreme values, specially the negative returns.

A plain vanilla of value at risk is the following case:

  1) The asset has a daily mean return of 0.5% and unconditional daily volatility is 2.5%

  2) The time horizon is 10 days and the probability of a maximum loss is 1%

  3) Supposed that the daily returns have normal probability distribution
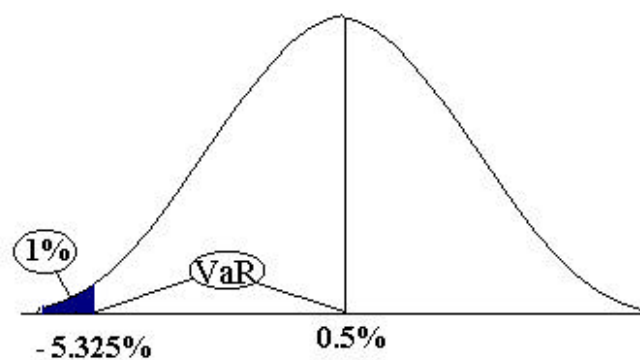
In the case exposed, it is obvious that the maximum loss possible is the extreme of the negative values as a result of 0.5 – (2.33 * 2.5). In consequence if we have 1.000.000 ∈, the maximum loss for one day is 53.250 ∈, for that reason the portfolio only covers reserves for 946.750 ∈

What happen if for liquidity market reasons or for comply with Basle regulations, we need estimate an horizon to predict of 10 days, in consequence if the volatility follows the law of $t^{0.5}$ then the maximum loss is $(0.5 * 10) – [2.33 * (2.5 * 10^{0.5})] = 13.4203\%$. Now the maximum loss possible in 10 days is 134.203 $\in$ and the portfolio value is 865.797 $\in$.

Really the company doesn't have 1.000.000 $\in$, probably only will have a 946.750 $\in$ tomorrow or 865.797 $\in$ in the next 10 days.

In view of these results, and taking in account that the reserve for VaR is necessary to compute in the balance, we have many problems to solve

1) If we apply the considerations done about the autocorrelation of daily volatility, we need to apply a model to predict the same.

2) Considering that the returns aren't *nid* and with strong presence of skew-ness and kurtosis, we need to think in a heavy tails distribution.

3) By the presence of liquidity risk it is important to determine how much time we need to liquidate the position. Basel Committee recommends a 10 days forecast, in consequence is appropriate to think how estimate the speed of increase the daily volatility.

I.   Modeling the volatility – The Arch family models

The first part to estimate VaR is the volatility forecast. As it has been explain in previous paper the volatility is a predictable process. If we regress the series returns on a constant the model is:

$$R_t = c + e_t$$

This constant is de mean of the series and the e_t is the error or the difference between the real value and the constant or mean. If we analyze these squared differences we are in presence of variance series or e_t^2. If the autocorrelation and the partial autocorrelation function of the variance series present a black noise is a signal that the variance is a predictable process.

5

The presence of Arch

There are two steps to determine the presence of arch in the model selected that may be a Regression Model or an Arima model. Once we select the right model we must go trough the following steps

First step

The analysis if the squared residuals are correlated with the intervention of the lags between time t and t-k, for that we use the ACF (autocorrelation function), and without the intervention of the lags between t and t-k, for that we use the PACF (partial autocorrelation function).

It is useful the Q-Statistic coefficient to make the following test of hypothesis for the autocorrelation function:

$H_0 : \rho_1 = \rho_2 = \rho_3 = \ldots\ldots = \rho_\kappa = 0$

$H_1 :$ some $\rho_\kappa \neq 0$

The statistic to be used is the Q-Statistic presented by Ljung Box represented by:

$$Q = n(n+2) \sum_{i=1}^{k} \frac{r_i^2}{n-i} \rightarrow c_k^2$$

Where k is the quantity of lags included in the analysis

When the Q-Statistic value is large, the area under the Chi Square distribution that exceeds this value is less than 0.05; in consequences we reject the null hypothesis.

The same test we use for the partial autocorrelation function to test if some of the parameters $\phi_{11}, \phi_{22}, \phi_{33}, \ldots, \phi_{\kappa\kappa}$ are distinct of zero.

Second step

If we pass the first step, we can advance with the second step to detect the presence of Arch. Engel in his seminal paper (1982) suggest the use of the Lagrange multiplier or the Arch LM Test. The methodology involves to fit e_t^2 by the regression of these squared residuals founded in the right model on a constant and on the k lagged values, that is:

$$\hat{e}_t^2 = a_0 + a_1 e_{t-1}^2 + a_2 e_{t-2}^2 + \cdots + a_k e_{t-k}^2$$

If there are no Arch effects, the estimated values of a_1 through a_k should be zero. Hence, this regression will have little explanatory power so that the coefficient of determination, R^2, will be

quite low. With a sample of T residuals, under the null hypothesis of no Arch errors, the test statistic TR^2 converges to a c_k^2

If TR^2 is sufficiently large, we reject the null hypothesis. It is to say that the coefficients a_l through a_k are jointly equal to zero and is equivalent to rejecting the null hypothesis of no Arch errors, or is to say there are Arch effects. In the same sense if TR^2 is sufficiently low, it is possible to conclude that there are not Arch effects.

A simple Arch model to forecast the conditional variance

After we detect the presence of Arch effects in the model proposed (regress the series on a constant) we could forecast the volatility for the next period using the conditional variance as follows:

$$s_t^2 = a_0 + \sum_{i=1}^{k} a_i s_{t-i}^2 + h_t$$

This is an Arch (k) model and forecasts the variance for to morrow conditioned to its past realizations.

Bollerslev presents the Garch model adding the concept that the volatility for tomorrow depends not only on the past realizations but it depends too on the errors of the volatility predicted. In consequence the model is the following:

$$s_t^2 = v + \sum_{i=1}^{p} a_i e_{t-1}^2 + \sum_{i=1}^{q} b_i s_{t-i}^2 + v_t$$

Where e_t^2 is the impact realization or the difference between the variance forecasted by the model and the real variance produced, s_t^2 is the variance forecasted by the model and v_t is a white noise process with mean zero and variance equals to 1

The market uses with successful in the middle of 90' the Garch (1,1) that is the traders do to take decisions on the behavior of different assets. The question: What will happen tomorrow? Is answered by the traders with this concept "*a little of the error of my prediction of today plus a little of the prediction for today*" and this is a Garch (1,1) model.

Everything seems to indicate until here that we have arrived to the most precise values for this estimate, and the obtained results help to predict the variance or the volatility for the period *t+1*. Analyzing the coefficients of the equation of the estimate of the variance a and b can we

calculate the non-conditional variance or the traditional variance considering that $s_t^2 = s_{t-1}^2 = s^2$.

Solving for s we obtain $s^2 = \dfrac{w}{1 - a - b}$

So that the pattern is stationary the sum of the parameters a and b should it be smaller to the unit. This sum of $a + b$ knows as persistence of the model and on this base this Garch (1,1) model can predict the variance for an horizon of t days. We don't forget that when we estimate the VaR for an asset, the presence of liquidity risk obliged to forecast the variance not only for one day, we need to forecast for t days.

If the volatilities of financial series of the capital market are analyzed we can observe that the

persistence of the model Garch (1,1) used it comes closer to the unit, but is what is noticed that

the values of a and b are different to each other.

Two alternatives are presented for the representation of the equation of the variance of the

model:

1) The equation of the conditioned variance estimated by a Garch (1,1) we can write it starting

   from considering that in the following way:

$$e_t^2 = v_t + s_t^2 \quad \therefore \quad s_t^2 = e_t^2 - v_t$$
$$e_t^2 - v_t = w + a e_{t-1}^2 + b\left(e_{t-1}^2 - v_t\right)$$
$$e_t^2 = w + (a + b) e_{t-1}^2 + v_t - b v_{t-1}$$

   Then the square error of an heteroscedasticity process seems an ARMA (1,1). The
   autoregressive root that governs the persistence of the shocks of volatility is the sum of
   $(a + b)$

2) It is possible in recursive form to substitute the variance of the previous periods in the right
   part of the equation and to express the variance like the sum of a pondered average from all
   the residuals to the last square.

8

$$s_t^2 = v + ae_{t-1}^2 + bs_{t-1}^2$$

$$s_t^2 = v + ae_{t-1}^2 + b(v + ae_{t-2}^2 + bs_{t-2}^2)$$

$$\vdots$$

$$s_t^2 = \frac{w(1-b^k)}{1-b} + a\sum_{j=1}^{k} b^{j-1}e_{t-j}^2 + b^k s_{t-1}^2$$

We can appreciate that in the Garch (1,1) model the variance decays in exponential form pondering the last residuals, giving him most importance to the residuals nearer and less importance to the residuals more distant. If $k \to \infty$ (k is the lagged value) then *(1 - b^k) = 1* because $b^k = 0$ also affecting to the last one adding of the right. In consequence the variance for the period t in recursive form is in the following way:

$$s_t^2 = \frac{w}{1-b} + a\sum_{j=1}^{k} b^{j-1}e_{t-j}^2$$

As the expressed, with a model Garch, it is assumed that the variance of the returns can be a predictable process. In a long term prediction horizon, using a Garch (1,1), the variance can be calculated for one period to t periods in the following way, taking as t the quantity of periods that there are between the horizon to be predict and $t+1$

$$s_{t,t}^2 = s_t^2 + s_{t+1}^2 + s_{t+2}^2 + \cdots + s_t^2$$

$$E_{t-1}(s_{t,t}^2) = E_{t-1}(s_t^2) + E_{t-1}(s_{t+1}^2) + E_{t-1}(s_{t+2}^2) + \cdots + E_{t-1}(s_t^2)$$

If it is calculated with a model Garch (1,1) the prediction for period number 2 using the conditional variance calculated for the first period, considering one has:

$$E_{t-1}(s_{t+1}^2) = E_{t-1}(v + ae_t^2 + bs_t^2) = v + (a+b)s_t^2$$

$$E_{t-1}(s_{t+2}^2) = E_{t-1}(v + ae_{+1t}^2 + bs_{+1t}^2) = v + (a+b)[v + (a+b)s_t^2] = v + v(a+b) + (a+b)^2 s_t^2$$

Substituting t period for the future one has the following formula that ensures the variance for the day t if the series have a behavior of a Garch (1,1)

$$E(s_t^2) = v\frac{1-(a+b)^t}{1-(a+b)} + (a+b)^t s_t^2$$

In the previous formula if w trends to zero the decay factor is $(a+b)^k$ for that reason an impact in the predicted variance decays following an exponential function. The shock disappears in a some days and trends to the non-conditional variance. The quantity of days is a function of the sum of $(a+b)$ or the persistence of the model.

In consequence the total sum of the variance for the t periods is the following

$$E_{t-1}(s_{t,t}) = \frac{v}{1-(a+b)}\left[(t-1) - \left\{(a+b)\frac{1-(a+b)^{t-1}}{1-(a+b)}\right\}\right] + \frac{1-(a+b)^t}{1-(a+b)}s_t^2$$

The extrapolation of the next day variance to a longer horizon is a complicated function of the variance process and the initial conditions. Thus the consideration of the volatility follows the law of $t^{0.5}$ fails due to the fact that the returns are not identically distributed.

Some results applying these equations are the following:

If we analyze a Garch (1,1) with the following coefficients:

$$s_t^2 = 0.001 + 0.03\,e_{t-1}^2 + 0.90\,s_{t-1}^2$$

If the values of $s_{t-1}^2 = 1.45\%$ and $e_{t-1}^2 = 2.4025\%$ the variance predicted for time t is 1.3781%. If we apply this result and estimate de total variance for a period of 10 days de value is 10,1963% very different of the result of the two ways proposed.

1) Considering the non-conditional variance $s^2 = \dfrac{w}{1-a-b}$ then $s^2 = 0.01428\%$, then de variance for 10 days is 0.01428 * 10 = 0.1428

2) Considering the variance estimated with a Garch (1,1), the last prediction for one day is 1.3781%, and applying the rule used in one we have that the variance for 10 days is 1.3781 * 10 = 13.781%

Both results are so far from the Garch (1,1) recursive form that estimates a variance for 10 days of 10.1963%.

At this time we don't arrive to the heaven but we are in the right way.

If we analyze the behavior of this model we can observe that the sign of the impact is not considered because to forecast the variance for tomorrow the model consider the impact of today as a squared value, then is not important the behavior of the market in the prediction model, for that reason the Garch model is call **symmetric model**.

In Engle and Ng (1993) the news impact curve analyzes how the asymmetry Garch models predicts the volatility for tomorrow based on today's impact. Nelson (1990) develop the idea of modify the Garch model due to the simple structure of Garch models. We have found evidence of negative correlations between the stocks returns and the changes in return volatility. In other words volatility tends to rise in response to "bad news" and tends to fall in response to good news". The analysts have fruitfully applied the Garch methodology in assets pricing models and in the volatility forecast. Risk Metrics [2] use a special Garch model when use the decay factor $l = 0.94$. The behavior of this model is similar to a Garch (1,1) with $a = (1-l), b = 1$ and $w = 0$. It is easy verify that the coefficients of a Garch (1,1) process are nonnegative due to the conditional variance is captured by a constant plus a weighted average (positive weights) of past squared residuals. These models elegantly capture the volatility clustering in assets returns, first noted by Mandelbroot (1963): " … large changes tend to be followed by large changes – of either sign - and small changes by small changes ". The Garch models has important limitations for example:

1) Researchers have found evidence that stocks returns that stock returns are negatively correlated with changes in returns volatility. As it was explained before in response to the type of excess of returns, positive or negative, the Garch models only are sensitive to the magnitude of the excess of returns and not to the sign of this excess of return.

2) An other limitation of Garch models results from the non negative constraints on $w, a,$ and b which are imposed to ensure that s^2_t remains positive, and if e_t increases the values of s^2_{t+m} for m ≥ 1

3) This third drawback in the Garch modeling is the concept of persistence of shocks to conditional variance. The problem is to estimate how long shocks persists on conditional variance. In Garch (1,1) models, shocks may persist in one norm and die out in another, so the conditional moments, may explode when the process itself is strictly stationary and ergodic.

To solve the restrictions of Garch models Nelson and Glosten Jaganathan and Runkle presents the asymmetric Garch models named Exponential arch (Egarch) and Threshold arch (Tarch) respectively.

---

[2] Risk Metrics is a trade mark of J. P. Morgan

The asymmetric models

The Exponential Arch (EGARCH)

Nelson (1990) was proposed an Exponential Garch model known as EGARCH, with the purpose to solve the restriction that presents the Garch (1,1) model, discussed previously.

The presentation for the conditional variance for an Egarch (1,1) is:

$$Ln\left(s_t^2\right) = w + b\, Ln\left(s_{t-1}^2\right) + a\frac{|e_{t-1}|}{|s_{t-1}|} + g\frac{e_{t-1}}{s_{t-1}}$$

In the left hand side is the *ln* of of the conditional variance. This implies that the leverage effect is exponential, rather than quadratic, and that forecasts of the conditional variance are guaranteed to be nonnegative. The presence of leverage effects is due to that $g < 0$. The impact is asymmetric if $g \ne 0$

If we generalize the model we obtain:

$$Ln\left(s_t^2\right) = w + \sum_{j=1}^{q} b_j\, Ln\left(s_{t-j1}^2\right) + \sum_{i=1}^{p}\left(a_i\frac{|e_{t-i}|}{|s_{t-i}|} + g_i\frac{e_{t-i}}{s_{t-i}}\right)$$

In Exponential Garch *ln* (s_t^2) is a linear process and its stationarity[3] and ergodicity[4] are easily checked. If the shocks to $\{ln(s_t^2)\}$ die out quickly, and remove the time varying component a then $\{ln(s_t^2)\}$ is strictly stationarity and ergodic

The Threshold Arch (TARCH)

Glosten, Jaganathan and Runkle (1993) presents the Tarch model and independently Zakoian presents the same model (1990) to forecast the variance. The conditional variance is specified as follows:

$$s_t^2 = v + ae_{t-1}^2 + ge_{t-1}^2 d_{t-1} + bs_{t-1}^2$$
$$where \quad d_{t-1} = 1\, if\; e_{t-1} < 0,\, and\; d_{t-1} = 0\; otherwise$$

---

[3] The process is stationarity when $E(y_t) = E(y_{t-s}) = m$
$$E[(y_t\text{-}m)] = E[(y_{t-s}\text{-}m)] = s_y^2$$
$$E[(y_t\text{-}m)(y_{t-s}\text{-}m)] = E[(y_{t-j}\text{-}m)(y_{t-j-s}\text{-}m)] = g\,(auto\ covariance\ at\ lag\ s)$$
Where μ, σ_y^2, and γ_s are constants
[4] The process is ergodic for the mean if g goes to zero quickly as j becomes large, and the process is ergodic for the second moments when:
$$[1/(T-j)]\sum_{i=j+1}^{T}(y_t - m)(y_{t-j} - m) \to g_j$$

In this model, good news are represented by $e_t > 0$, and bas news by $e_t < 0$, and have differential effects on the conditional variance equation. The good news only has an impact of a and bad news has impact in the sum of $a + g$. The leverage effect of bad news exists only when $g \neq 0$ in this case the news impact is asymmetric.

For a higher order of Tarch model the specification is the following:

$$s_t^2 = v + \sum_{i=1}^{q} a e_{t-i}^2 + g e_{t-1}^2 d_{t-1} + \sum_{i=1}^{p} b s_{t-i}^2$$

$$where \quad d_{t-1} = 1 \, if \, e_{t-1} < 0, \, and \, d_{t-1} = 0 \, otherwise$$

As synthesis of it was exposed we are in the presence of symmetric Garch (p,q) model and asymmetric models as Egarch (p,q) and Tarch (p,q) models. Engle and Ng (1993) presents a resume of several models that they try to explain the volatility structure.

Using the convenient software we can select the appropriate model to analyze the behavior of the volatility. The first step after accepts the presence of heteroscedasticity is test the significance of the type of model select between the asymmetric and symmetric Garch model.

The asymmetry test

To check if the model is correctly specified, we have two forms to detect the right model. One form is comparing the coefficients of the model. The log likelihood, the Akaike info criteria and the Schwarz Criteria aids to take a good model selection. The value of the log likelihood estimates when the model converges, and then is calculated by the following equation:

$$\ell = -\frac{T}{2} \left[1 + \log(2p) + \log\left(e'e / T\right) \right]$$

Where $e'e$ is the sum of the squared residuals of the model.

The AIC and SC ensures that the model complies with the condition established for an ARIMA model that balances the goodness of fit and parsimonious specification. The formulas to calculate these coefficients are the following:

$$\text{Akaike info criterion (AIC)} \quad -2\ell/n + 2k/n$$

$$\text{Schwarz criterion (SC)} \quad -2\ell/n + [k \log(n)]/n$$

Where k is the number of estimated parameters, n is the number of observations, and ℓ is the value of the lo g likelihood function using the k estimated parameters

The other form is making an asymmetry test, where we must do a cross correlation between the squared residuals of the Garch model and the standardized residuals of the same. The result of this cross correlation will be a white noise if the model is symmetric or in other words the Garch model is correctly specified, and a black noise is the model is asymmetric.

Now we have identified a model to forecast the volatility and we know a part of the problem, because the interval of confidence has two parts the volatility and the coefficient of the probability distribution that covers the limits of the interval of confidence due to a certain probability established previously.

II.   The probability distribution of the returns

In previous papers it was demonstrated that the returns not follows a normal distribution, and we reject the test of goodness of fit in the same sense. The behaviors of these time series have probability distribution with heavy tails. These distributions are treated in detail in the Annex 1 of this paper, and complete the forecast of VaR considering jointly the volatility and the probability distribution.

These heavy tails distributions have incidence in the estimation of VaR especially when we analyze the extreme limits of the distribution. Remember that VaR treats to estimate the possible maximum loss at certain level of confidence established previously. Normally this maximum probability is determined at 99% of confidence, and then we are analyzing the lower percentile of the distribution.

If we are in presence of a Garch model to predict de volatility, this is a symmetric model and in consequence the model does not detect the incidence of bad news. For example estimate the risk for the 5 years treasury bond.

Using the appropriate software, the models to forecast the volatility are shown in table I, in consequence of this results, we accept that the best model is an Egarch (1,1) model due to the values of the log likelihood and the Akaike Info Criterion and Schwarz Criterion.

If we apply the results and estimate the VaR for the first percentile, using the normal distribution and make a back testing using the last 755 observations, ended on 31st January 2003, we obtain the results exhibit on table II, for the different Arch models

Table I

Coefficient	Garch(1,1)	Tarch (1.1)	Egarch(1,1)
Mean	-0.100235	-0.145484	-0.181462
	(0.041346)	(0.043854)	(0.045529)
w	0.019106	0.009516	-0.002929
	(0.009783)	(0.007824)	(0.010497)
a	0.129244	0.030235	0.013904
	(0.017804)	(0.028565)	(0.011090)
b	0.879834	0.909609	0.998601
	(0.016776)	(0.015203)	(0.001713)
g	-	0.150783	-0.110551
		(0.031201)	(0.008830)
Log Likelihood	-1359.123	-1351.342	-1343.758
AIC	3.610922	3.592959	3.572869
SC	3.635434	3.623599	3.603509

(The numbers into the brackets are the standard deviation of the coefficients)

Table II

Number of returns that exceed the VaR Forecasted		
Garch (1,1)	**Tarch (1,1)**	**Egarch (1,1)**
6	**6**	**4**

Backtesting Garch(1,1) and Egarch(1,1) for 5 years Treasury Bond
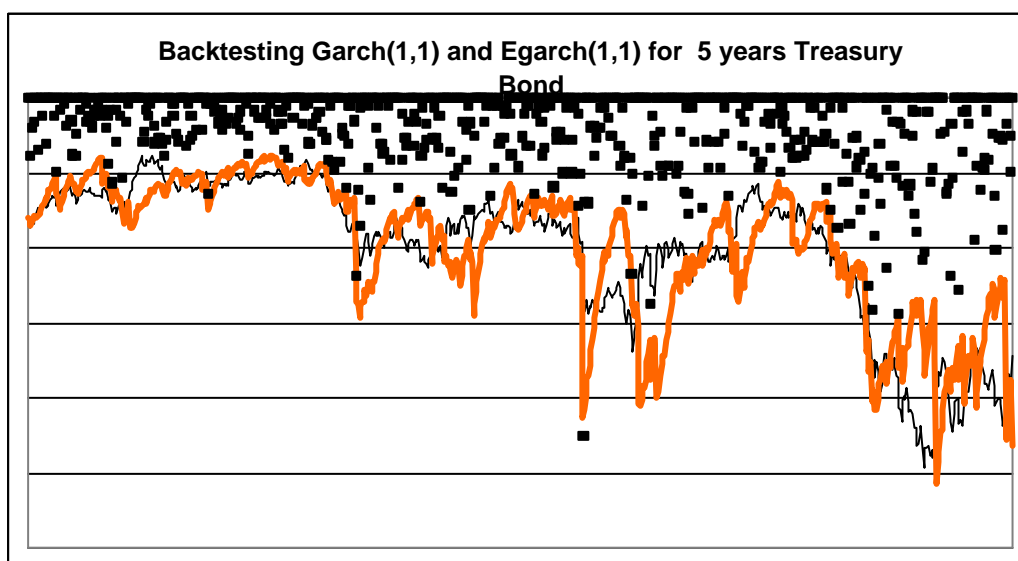
As a result of the back testing exhibit in the graph, we can observe the behavior of a Garch (1,1) with bold line and the behavior of the Egarch (1,1) with fine line. Both models comply with the objective, using the normal distribution, that there are 1% of observations at left of the maximum loss forecasted, 6 over 755 observations for Garch (1,1) and 4 over the same number of observations for Egarch (1,1).

The model selection accepts that the Egarch (1,1) model because it improves lightly the log likelihood and the Akaike and Schwarz Criterion.

To estimate the amount of VaR or the maximum loss possible, the normal distribution of probabilities jointly with the volatility forecasted by Garch family models produce good results because the number of observations over the limit in the low limit complies with de amount estimated.

The probability distributions of returns don't follow the normal distribution as I had demonstrated in a previous work. The distributions found for some time series are shown in table III

Table III

Time series	Period analyzed	Number of obs.	First Distribution Fitted	Second Distribution Fitted	Test goodness of fit			
					K S		A D	
					1st	2nd	1st	2nd
5 yr. T bond	02/01/00 30/01/03	755	Logistic	Extreme value Dist.	0.059	0.140	2.401	32.38
Bayer	06/06/02 18/03/02	201	Logistic	Extreme value Distr.	0.046	0.104	0.409	6.142
Dow Jones	04/01/99 18/03/03	1056	Logistic	Weibull	0.013	0.067	0.171	10.87
Bovespa[5]	04/01/99 18/03/03	1036	Logistic	Weibull	0.019	0.078	0.492	10.56
Merval	03/01/01 18/03/03	773	Logistic	Extreme value Distr.	0.049	0.127	3.079	30.55
IDP[6]	25/01/01 30/01/03	445	Logistic	Weibull	0.115	0.176	8.257	23.13

The test of goodness of fit used are the Kolmogorov Smirnov test (KS) and the Anderson Darling (AD). Both test of goodness of fit, have strong results specially when it is necessary to test

---

[5] See next page
[6] See next page

asymmetric and heavy tails distributions of probabilities. The Kolmogorov Smirnov test is a test that is independent of any gaussian distribution, and have the benefit that not need a great number of observations. The Anderson Darling test is a refinement of KS test, specially studied for heavy tails distributions. Once we fit the appropriate distribution based on the sample or the time series available we can simulate the values that will take the series using a sampling method Monte Carlo or Latin Hypercube. After 20.000 trials the value corresponding to the first percentile is presented in the Table IV and these results are compared with the values obtained by de interval of confidence using Arch models and normal probability distribution.

Table IV

Time Series	Value at 1% 1st distribution fitted	Value at 1% 2nd distribution fitted	Arch model	Values for 1st percentile using Arch an normal distribution	
				Maximum	Minimum
5 yr. T bond	- 1.02	-5.78	Egarch (1,1)	-1.534	-9.856
Bayer	-10.03	-21.78	Egarch (1,1)	- 3.915	-21.283
Dow Jones	- 3.35	- 3.90	Egarch (1,1)	- 1.449	- 6.512
Bovespa [7]	- 5.46	- 6.65	Tarch (1,1)	- 3.649	-15.103
Merval	- 6.50	-14.95	Egarch (1,1)	- 2.772	-14.462
IDP [8]	- 2.34	- 5.86	Egarch(1,1)	-3.697	- 7.865

The results of these arch models presented in table IV are detailed in the following part, but considering that the number of excess of values founded under the lower limit not exceed the 1% of the total observations, the combination of Arch model and normal distribution provides a good fit for the VaR forecast.

III.   The forecast of Market VaR

As a result of the previous considerations about the volatility forecast and the distribution fitted to compute de Market VaR, the conclusion arrived are the following:

1) The volatility it is presented in clusters and have an strong incidence of heteroscedasticity in the model used to forecast the volatility for the period *t+1*

---

[7] Bovespa is the Brazilian index of the stock exchange of San Pablo
[8] IDP is the index of the bonds that represent the public debt of Colombian Republic

2) The prices of stocks and the corresponding index market are agitated by bad news product of, terrorism, war, economic depression, adulterated balance sheet, etc, and in consequence the asymmetric models are the selected Arch models.

The use and asymmetric model, that weights the negative results, produce leverage in the volatility forecast, and directly the impact of negative returns open the parametric space and produces a great value of volatility.

In the following graphs, we can observe the volatility through the time pass.

Daily Volatility MERVAL Index

Daily Volatility of DOW JONES
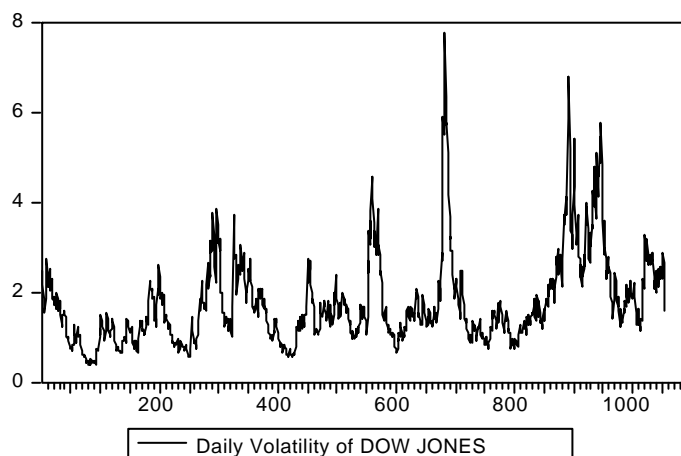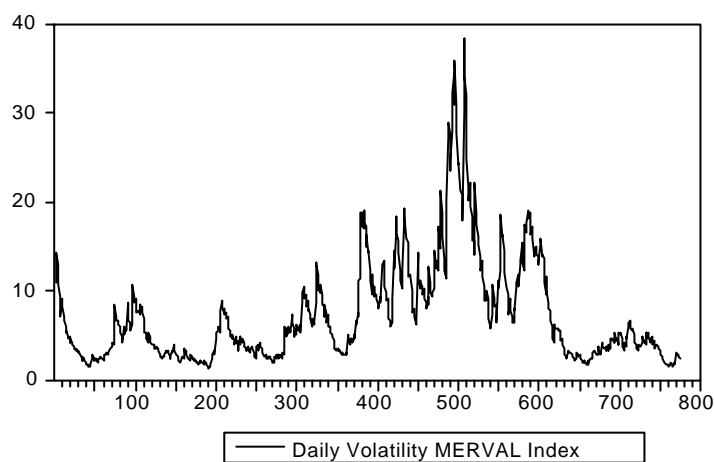
As an example we can observe in the Merval Index, the index of the Stock Exchange of Buenos Aires Argentina, the great volatility in the middle of the graph is a result of the change of government at the end of 2001 and principle of 2002. In the other graph, the Dow Jones index

18

has been agitated by September eleven, the crises of bankruptcy of Enron and MCI, the war notices in the final.

Then the asymmetric Garch model and the normal distribution of probabilities produce a best approach to forecast how much many we will lose tomorrow, and the distribution fitted for the returns only forecast the VaR for a long range of time. Remember that when we think in Market risk, we think in the result for tomorrow, not for the next year.

IV. Some results

The results, to estimate de volatility for the next period, of the time series analyzed are the following:

Table V

Series\Coeff.	5 years T bond	Bayer	Dow Jones	Bovespa	Merval	IDP
Model	Egarch (1,1)	Egarch (1,1)	Egarch (1,1)	Tarch (1,1)	Egarch (1,1)	Egarch (1,1)
Mean return		-0.233684 (0.194770)	-0.036214 (0.036492)	0.003047 (0.063416)	-0.051153 (0.081690)	0.096725 (0.053694)
w		0.174492 (0.194770)	-0.036214 (0.017712)	0.397125 (0.089869)	-0.121925 (0.016902)	0.359076 (0.031723)
a		-0.142854 (0.028628)	0.060276 (0.022571)	-0.001285 (0.014372)	0.198057 (0.025022)	0.188690 (0.028072)
b		0.974717 (0.006145)	0.974012 (0.006254)	0.833621 (0.034005)	0.983886 (0.006189)	0.578215 (0.074028)
g		-0.185586 (0.035579)	-0.129700 (0.014525)	0.163282 (0.021862)	-0.064811 (0.014907)	-0.181756 (0.181756)
Outliers [9]		1	10	10	13	3
Outliers / observations		0.5%	1 %	1%	1.1%	0.9%

The importances of these results are that the percentage of outliers over the number of observations not exceeds the 1% equivalent to the probability of the first percentile. Only in the case of the Merval index, it presents in the back test that the total of outliers are great the 1%. In this case is reasonably because at and of 2001 it was produced the end of the convertibility.

---

[9] The outliers are the negative returns that exceeds the VaR forecasted using de normal probability distribution for the first percentile and the conditioned volatility estimated by the Garch model

V.   Conclusions

The conclusions arrived are the following:

1) To estimate the volatility is necessary to develop a model that consider the movements of the volatility in the time series, for that reason, the traditional volatility don't follows the market variations, and the forecast of the market VaR have in certain moment excess of reserve and in other moments a lack of reserve, as we can observe in table IV and the graphs included in part IV when we do the estimation using this traditional volatility with a heavy tail distribution.

2) The asymmetric Garch models, like Tarch an Egarch model, not only fulfill with the movements of the volatility, as we can observe with the back testing presented, also it is not necessary to use the heavy tails distributions, because the negative impact or the negative returns are included by the model form.

3) The time series history, complies with the requirements of Basel Committee, to make the volatility forecast. It is easy to teach this models to the traders but not for the actuaries. The traders and the investors only the recent past import them.

4) The models as we can observe are dynamic, and is very important the revision of these models periodically

The final objective is present a exact reserve which covers the maximum loss possible, and this reserves may have a value that corresponds with the reality. When we introduce the VaR restrictions for the traders and for the managers, the forecast of the reserve may be a credible value at effect that traders and managers use this information for decision-making. Also the market VaR jointly with the credit risk and operational risk serves to calibrate the total risk of a performance of the company.

Bibliography

Bollerslev T., 1986, Generalized autoregressive conditional heteroscedasticity, *Journal of Econometrics* 31, 307-327

Basel Committee on Banking Supervision, 2001, The New Basel Capital Accord, *Bank for International Settlement*

Cruz Marcelo G., 2002, Modeling measuring and hedging operational risk, *John Wiley and Sons*

Duffie Darrell, Jun Pan, 1997, An Overview of Value at Risk, *The Journal of Derivatives,* Spring 1997

Embrechts Paul, Kluppelberg Claudia, Mikosch Thomas, 1997 Modeling Extremal Events for insurance and finance, *Springer*

Engle, Robert F., 1982 Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica* 50, 987-1007

Engle, R., y T. Bollerslev, 1986, Modeling persistence of conditional variances, *Econometric Review* 5, 1-50

Engle, R., y Victor K. Ng, , 1993, Measuring and testing the impact of News an Volatility, *The Journal of Finance* Vol. XLVIII, Nro. 5

Glosten, Lawrence, Ravi Jaganathan, and David Runkle, 1993, On the relationship between then expected value of the volatility of the nominal excess return on stocks, *The Journal of Finance* Vol. XLVIII, Nro. 5

Greene, William H., 1997, Econometric Analysis, *Prentice Hall*, New Jersey

Duan Jin-Chuan, 1995, The Garch option pricing model, *Mathematical Finance* Vol. V Nro.1

J.P. Morgan, 1994 Riskmetrics, Technical document Nro 4 *Reuters*

Jorion Phillippe, 2000, Value at Risk, *Mac-Graw-Hill*, New York

Jorion Phillippe, 2001, Handbook Financial Risk Manager 2001-2002, *John Wiley and Sons*

Nelson, D., 1990, Conditional heteroscedasticity in asset returns: A new approach, *Econometrica* 59, 347-370

Peters Edgar, 1994, Fractal Market Analysis, *John Wiley & Sons, Inc.* New York

Reiss R. D., Thomas M. 2001, Statistical Analysis of Extreme Values, *Birkhauser Verlag*

Tagliafichi Ricardo A., 2001, The Garch model and their application to the VaR, *XXXI International Astin Colloquium,* Washington 2001

Tagliafichi Ricardo A. 2002, Betas calculated with Garch models provides new parameters for a Portfolio selection with Efficient Frontier, *ICA Cancun 2002*

Tsay Ruey S. 2002, Analysis of Financial Time Series, *John Wiley and Sons*

Wiggins, J.B., 1987, Option Values under stochastic volatility: Theory and empirical tests. *Journal of Financial Economics* 19, 351-372

Zakoian Jean-Michel, 1992, Threshold Heteroskedastic models, *Journal of Economic Dynamics and Control* Nro 18

Annex 1

The heavy Tails Distributions

Exponential distribution

This distribution is used in the study of the reliability theory and the PDF y CDF as follows respectively:

$$f(x) = \frac{e^{-\frac{x}{l}}}{l}$$

$$F(x) = 1 - e^{-\frac{x}{l}}$$

The calculus of the parameter value, will be estimate by the following estimator:

$$\hat{l} = \frac{1}{\sum_{i=1}^{n} X_i / n}$$

Weibull Distribution

The Sweden Waloodi Weibull presents this distribution in 1939. The financial econometrics pay attention to this distributions in the last part of the last century. The PDF y CDF are the following in the same order:

$$f(x) = \frac{a}{b^a} x^{a-1} e^{-\left(\frac{x}{b}\right)^a} \quad , \quad F(x) = 1 - e^{-\left(\frac{x}{a}\right)^b}$$

The calculus of the parameters a and b is:

$$b = \frac{c \ln(a) - \log(b)}{c - 1} \qquad y \quad a = \frac{\ln(\ln(4))}{\ln(b) - \ln(b)}$$

$$c = \frac{\ln(\ln(4))}{\ln(\ln(4/3))} = -0.262167b$$

Where a and b are the 25 y 75 percentiles respectively

The Pareto distribution

The Pareto distribution is know as the Pareto's law who determine the income population distribution and the PDF y CDF are the following:

$$f(x) = \frac{aq}{(x+q)^{a+1}} \quad , \quad F(x) = 1 - \left(\frac{q}{x+q}\right)^{a}$$

To estimate the parameters we use the following expressions:

$$\hat{a} = 2\frac{\dfrac{\sum\limits_{i=1}^{n} x_i^2}{n} - \left(\dfrac{\sum\limits_{i=1}^{n} x_i}{n}\right)^2}{\dfrac{\sum\limits_{i=1}^{n} x_i^2}{n} - 2\left(\dfrac{\sum\limits_{i=1}^{n} x_i}{n}\right)^2} \quad , \quad \hat{q} = \frac{\dfrac{\sum\limits_{i=1}^{n} x_i}{n} - \left(\dfrac{\sum\limits_{i=1}^{n} x_i^2}{n}\right)^2}{\dfrac{\sum\limits_{i=1}^{n} x_i^2}{n} - 2\left(\dfrac{\sum\limits_{i=1}^{n} x_i}{n}\right)^2}$$

The logistic distribution

This distribution is a symmetric distribution with PDF and CDF estimate as follows:

$$f(x) = \frac{z}{a(1+z)^2} \quad para -\infty < x < \infty$$

$$a = scale\,factor = \sqrt{3s/p}$$

$$z = e^{-\left(\frac{x-m}{a}\right)}$$

$$F(x) = \frac{1}{1+e^{-z}}$$

The extreme value distribution EVD

The question is "*How much capital I will lose in a particular event*" or "*Which is the probability that it will produce an other September eleven*"

The answer is in the Extreme Value Distribution. The actuaries and financial specialists are using this distribution used by Embretchs in his famous work.

Let it be, X_1, X_2, ... X_n, random variables, observed in a certain period, like returns. This is a continuous random variable in the variable domain and discrete in the time domain. If the extremes are defined like the maximum and the minimum in the order of the n observations, we can define $X_{1,n}$ as the maximum value observed during the period and $X_{2,n}$ as then second value of the series. In consequence the maximum value of the random variable will be $Y = X_{1,n}$ and will

be standardized by the parameters of location (mode), scale and shape, selected to show an appropriate distribution of the standardized extremes.

Frechet, Gumbel and Weibull define three important extreme distributions. A really representation of those distributions is done by the Extreme Value Distribution o Generalized Extreme Value.

The three distribution parameters, $F_{m,e,y}$, arise as a maximum normal limit distribution id random variables

For the random variable $Y = X_{1,n}$,we have:

$$Z = \frac{(Y - m)}{y} \quad y \quad z = \frac{(z - m)}{y}$$

Where m is the mode and y is the scale the:

$$P(Y \le y) = F_{mey}(y) = F_{0ey}(z) = \exp\left\{-1\left(1 + e\,z^{-\frac{1}{z}}\right)\right\} \qquad 1 + e\,z \ge 0$$

Where e is shape parameter. Due to the values that take e the EVD tends this behavior

$$e = 0 \qquad \text{Gumbel Distribution}$$
$$e \,^3\, 0 \qquad \text{Frechet Distribution}$$
$$e \,£\, 0 \qquad \text{Weibull Distribution}$$

The purpose is to found a practice form to estimate the parameters values for use EVD. The most important parameter is e that describes the weight of the tails distribution. If the shape parameter e is significant then the EVD is significant too. If the shape parameter is not significant then a log normal distribution or a logistic distribution is a good distribution to descibe the behavior of the random variable.

One of the form to approximate the calculus of this three parameters, m, y, e, is the PWM or probability weighted moments, that follows this steps.

$$\hat{m}_r(m, y, e) = \frac{1}{n}\sum_{i=1}^{n} X_i U_i^r$$

Where U is a plotting position that follows a free distribution and k takes the probability as:

$$p_{k,n} = [(n-k)+0.5]/n.$$

The best approach to estimate the parameters m, y, and e, is presented by Hosking in 1985 following this formula:

$$m_r = \frac{1}{r+1}\left[m + \frac{y}{e}\left\{1 - \frac{\Gamma(1+e)}{(1+r)^e}\right\}\right] \quad e > -1, e \neq 0$$

The value of e arise from the iterative estimations suggested too by Hosking

$$\hat{m} = \frac{(2m_2 - m_1)e}{\Gamma(1+e)(1-2^{-e})}$$

$$\hat{y} = m_1 + \frac{\hat{m}}{\hat{e}}(1 - \Gamma(1-e))$$

$$\backslash e = 7.859\, c + 2.9554\, c^2$$

$$c = \frac{2m_2 - m_1}{3m_3 - m_1} - \frac{\log 2}{\log 3}$$

The values of the scale and location, y and m depends on the values that take se shape value e Once we obtain the parameter values it can estimate de maximum values or the minimum values of the series. To estimate the quantiles based on a certain level of confidence it is possible to estimate the value of the random variable as follows

$$\hat{x}_p = m - \frac{y}{e}(1 - (-\ln p)^{-e})$$

The Normal extended distribution of Kupiec

This distribution is an application of heavy tails due to a normal distribution. If the problem is to cover the probability of extreme events then is important to extend the parametric space that cover the extreme values of the random variable. Kupiec demonstrate that on base a normal distribution that it is possible to extend the tails of the distribution in form that contemplate the probability of a catastrophe.

The value that takes the abscissa named z of a standardized normal distribution extended by Kupiec is:

$$z_{Kupiec} = \frac{p(1-p)}{f(x)^2}$$

Where $f(x)$ is the ordered of a normal distribution

26

If we apply the Kupiec considerations for the first percentile, the value of z, is:

$$z_{Kupiec} = \frac{0.01(1-0.01)}{0.02665^2} = 3.7332$$

In consequence for the same probability the parameter space of the random variable is $3.7332\boldsymbol{s}$ against $2.33\boldsymbol{s}$ of a normal distribution.

Notas Varias