ISCTE Business School
Instituto Universitário de Lisboa

MODELING VOLATILITY: AN ASSESSMENT OF THE VALUE AT RISK APPROACH

Joana Bruno Vieira

Project submitted as partial requirement for the conferral of
Master in Finance

Supervisor:
Prof. Luís Oliveira, Prof. Auxiliar, ISCTE Business School, Departamento de Finanças

Co-supervisor:
Prof. José Dias Curto, Prof. Associado, ISCTE Business School, Departamento de Métodos Quantitativos

April 2012

Resumo

Value at Risk (VaR) tornou-se uma das mais populares técnicas de medição e controlo de risco, nomeadamente risco de mercado. Esta medida diz-nos qual a perda máxima esperada de um activo ou portfólio para um determinado período de tempo (h) dado um certo intervalo de confiança.

Nesta tese, pretende-se verificar a adequação de alguns modelos de heteroscedasticidade condicional para estimar e modelizar a volatilidade dos retornos. Para isso, consideraram-se os seguintes modelos: EWMA, GARCH, A-PARCH, E-GARCH e GJR-GARCH e diferentes índices e taxas de câmbio representativos de áreas geográficas distintas, também como dois activos com características particulares: o ouro e o petróleo. A performance dos modelos na estimação do VaR foi analisada com recurso às técnicas de backtesting nomeadamente ao teste de Kupiec (1995) e Christoffersen (1998).

Com este estudo é revelado que o método GARCH e GJR-GARCH conseguem prever o VaR de uma forma mais precisa do que os restantes modelos considerados para os dois níveis de confiança analisados (95% e 99%).

Palavras-Chave: Value at Risk, Volatility, GARCH, Backtesting

JEL Classification: C10, G17, G32

Abstract

The Value at Risk (VaR) became one of the most popular technics for risk measuring and control, especially for market risk. This type of measure tells us which is the maximum expected lost for an asset or portfolio, for a given period of time (h) and a certain confidence level.

In order to compute the VaR, the main purpose of this dissertation is to verify the suitability of some conditional heteroskedasticity models to estimate and model the volatility of returns. To do this, the following models were considered: EWMA, GARCH, A-PARCH, E-GARCH, GJR-GARCH and different indexes and exchange rates representative of different geographical areas as well as two assets with particular characteristics: gold and oil. The models' performance in the estimation of VaR was analyzed by using Kupiec (1995) and Christoffersen (1998) backtesting technics.

The study revealed that GARCH and GJR-GARCH models seem to be the most accurate way to predict the VaR when the two most commonly used confidence levels (95% and 99%) are used.

Key words: Value at Risk, Volatility, GARCH, Backtesting

JEL Classification: C10, G17, G32

Acknowledgements

I would like to thank my supervisors Prof. Luís Oliveira and Prof. José Dias Curto by the guidelines and support given over the completion of this thesis.

In addition, I would like to congrat my parents and some friends for their support and encouragement

Table of Contents

List of Tables

List of Figures

1. Introduction

As a consequence of the stock market crash in 1987, and as markets became more volatile, financial institutions were more interested in risk management. In this decade many institutions became more leveraged which led to an increase of bankruptcies. Markets turned unstable and measures to represent the market risk were then required (see Jorion, 1997).

These events led regulatory authorities to impose Risk-Based-Capital to control financial risks (see Dimson and Marsh, 1995 and Wagster, 1996).

The risk managers wanted to know: How can we quantify risk and how much capital do they need to cover the risks under your business?

There are several measures to quantify risk, namely: standard deviation, semi-standard deviation, Value-at-Risk (VaR), Expected Tailed Loss (ETL). In spite of the strengths and weaknesses of all the risk measures, VaR is the preferred one by the financial institutions, regulators, non-financial corporations and asset managers, as discussed by Jorion (1997).

In 1995 Basel Accords imposed VaR as the standard risk measure adopted by the financial industry for exposure of the market risk.

The advantage of this market risk metric is its easiness to be presented as well as to be understood by most of people. Additionally, it can be compared across different markets and different exposures. The VaR measures the expected maximum loss in value (or percentage) of a portfolio for a given period with some probability associated. For example, $VaR_{h\,day,\alpha\%} = V$. This means that we are $(1-\alpha)\%$ confident that the loss will not exceed V euros (or another currency) in the next h days. The significance level (α) is usually below 0.05 (5% percentile). RiskMetric[1] uses this percentile, but Basel Accords set 1% for the significance level. The $(1-\alpha)$ is the confidence level. The h is denominated as the risk horizon, meaning the number of trading days ahead for which we want to measure the risk of Profit and Losses (P&L) distribution. Basel Accords set $h = 10$ (two weeks). Lastly, the parameter V is the VaR, which can be presented as a percentage or an amount.

[1] Commercial implementation of a VaR measurement system developed by J.P. Morgan (1996).

This is a critical matter for financial institutions because it requires a good evaluation of VaR to predict possible losses and measure the risk accurately. The most important «input» for VaR estimation is the volatility. As volatility is not directly observable, there is lots of research attempting to deal with this problem. A good volatility forecast is very important for a good estimation of the portfolio risk and it may be interesting and profitable to know how a certain variable will evolve in the future.

The aim of this dissertation is to analyze and compare the different approaches to predict volatility in order to provide better estimation for computing Value-at-Risk. The focus will be answering the question: what is the best volatility model?

This thesis will use stock indexes all around the world, the most important exchange rates and two particular commodities: gold and oil. The data set includes daily returns from January 2, 2000 to December 31, 2011. It is important to remind that our sample contains one of the most important crisis that remains to these days: subprime crisis. So, it might be interesting to see if there are any differences in the performance of the models during the periods of high volatility. To do this, we will divide the sample into two sub periods (before crisis and during the crisis) to see if there are differences in the models' performance. The VaR will be estimated based on six different methodologies in which four of them only differ in the choice of the GARCH model.

This dissertation is organized as follow: Section 2 describes parts of the existing literature under the methodology; the theoretical background is described in section 3; on section 4 the data and methodology are presented. Finally, in section 5 the main findings are discussed.

2. Literature Review

According to Christoffersen, Hahn and Inoue (2001), two important developments to risk management have occurred in the recent past: Engle (1982) developed the first models to measure and forecast volatility and J.P.Morgan (1996) introduced RiskMetrics. The latter remained as a benchmark for measuring market risk and led to improvements made by Jorion (1997), Duffie and Pan (1997) and Dowd (1998).

Hopper (1996) defined Value-at-Risk as the expected largest loss that a portfolio is likely to suffer during all but truly exceptional periods. More precisely, VaR is the maximum loss that an institution can be confident to lose in a certain fraction of time over a particular period.

VaR can be defined by:

$$Prob[r_t < -VaR] = \alpha \tag{1}$$

where r_t is the log return and α the significant level.

VaR became the most popular measure to quantify the market risk and determine capital request because it is easier to communicate, it lets different financial markets to be compared (which is useful due to internationalization and diversification), and it is possible to evaluate a portfolio in terms of risk-return relation.

There are different approaches for computing VaR. Determining what it is the best methodology for VaR estimation becomes an empirical question. The choice may depend on the number of underlying assets, their type and the exact objective of the calculation. VaR estimation methods are usually divided into parametric and non-parametric categories. Parametric methods are based on statistical assumptions about the risk factor distribution while non-parametric methods are based on simulations (see Amman and Reich, 2001).

According to Carol (2008), there are three main methodologies to estimate the VaR:

1. Parametric Linear
2. Historical Simulation
3. Monte Carlo Simulation

Risk managers prefer to use non-parametric methods especially Historical Simulation (HS) because it is the easiest to implement (Wiener, 1999) while Monte Carlo is the most challenging method. In fact, almost three quarters of banks use Historical Simulation (see Perignon and Smith, 2006).

This method was introduced by Boudoukh et al. (1998) and Barone-Adesi et al. (1999). Unlike the parametric method, this approach makes few distribution assumptions and it is relatively simple to implement (see Danielsson and Vries, 1997; Dowd, 1998; Manganelli and Engle, 2001). The idea is simply to use historical market data to estimate VaR for the current portfolio. It does not make any assumption about the parametric form of the returns' distribution and does not require the estimation of volatility and correlations since the changes in the portfolio over time has all the information that we need to compute the VaR. This approach takes into account the fat tails stylized fact of the empirical returns' distribution, as the data "speaks" by itself. Thus, Historical Simulation virtually applies to any type of instrument and, unlike the parametric method, it is not limited to linear portfolios[2] (see Carol, 2008). Other advantage is the inclusion of risk factors dynamic behavior which is assumed to be very simple and sometimes unrealistic in the parametric form.

The major drawback associated with Historical Simulation is the assumption that the future P&L distribution will be identical to the past distribution as data is considered equally relevant. The method assumes that history repeats itself. Data can contain events that will not appear in the future, so they have no relevance for the sample (see Manganelli and Engle, 2001). In this case the sample size is a crucial choice because it must be large (it is common to use more than five years of historical data). Sometimes, there are not enough available data especially when new assets are introduced to the portfolio. In addition, when facing high confidence levels, it is important to have a long sample but looking further into the past will decrease the relevance in the VaR estimation because the empirical distribution will reflect less the market condition.

Another problem is that Historical Simulation approach assumes that returns are independent and identically distributed (i.i.d.), and so it does not allow for time-varying volatility (see Sarma, 2003).

[2] Daily changes in the portfolio value are linearly dependent on daily changes in market variable. Portfolios that contain options are non-linear.

This paper focuses in this method. The main purpose is to show how these drawbacks can be overcome using volatility updating.

According to Dowd (1998), a convenient solution to these issues is to use weighted Historical Simulation which gives lower importance to observations that lie further in the past.

The major problem with equally weighted data is that extreme events (positive or negative) can influence the VaR estimate. So, Boudoukl, et. al. (1998) introduced the hybrid approach which combines “The Best of Both Worlds”: Risk Metrics introduced by J.P.Morgan (1996) and Historical Simulation. In the hybrid approach, instead of giving the same weight to all data whether it happened in a nearly or a longer past, it gives more importance to recent observations, assuming that recent volatilities have more influence in the volatility forecast than older ones. The weight assigned to the observations decreases exponentially over time and depends from the smoothing constant λ that ranges between 0 and 1. This Exponential Weighted Moving Average (EWMA) is used in RiskMetrics to update the underlying variance-covariance matrix.

As was said before, the Historical Simulation requires a large sample of financial database. The problem related with this is that market conditions change while the time goes by. For example, if the volatility of a stock had been stable but in the last two years was very high, the volatility we expect so see in the present will be underestimated if you give to both periods the same importance.

To overcome this problem Duffie and Pan (1997) and Hull and White (1998) recommend a volatility adjustment method. This methodology was designed to weight the returns in a way to adjust their volatility to the current volatility. To do this we can use the EWMA or Generalized Autoregressive Conditional Heteroscedasticity (GARCH).

The GARCH model has several advantages in relation to EWMA as will be explained in the next paragraphs.

As mentioned before the assumption that returns are i.i.d. is very unrealistic, so the volatility and correlation forecasts that came out of these idea are also mistaken. The financial data shows volatility clustering[3] behavior (see Mandelbrot, 1963 and Fama, 1965). Engle (1982) introduced autoregressive conditional heteroscedasticity (ARCH) and Bollerslev (1986) and Taylor (1986) its generalization (GARCH) that capture the volatility clustering effect. The GARCH model takes into consideration the dynamic properties of returns, i.e., the conditional variance is a function that includes past innovations (see Angelidis, Benos and Degiannakis, 2004).

The simplest GARCH(1,1) model is given by:

$$\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\delta_{t-1}^2 \qquad (2)$$

where σ_t^2 (conditional variance) is a function given by the square of past errors (ε_{t-1}^2) and past conditional variance (σ_{t-1}^2). It is required that conditional variance will be always positive, so there are some restrictions on the parameters: $\omega > 0$ and $\alpha, \beta \geq 0$.

The sum of parameters α and β measures the persistence. If this sum is close to one, volatility shocks are quite persistent over time, if it is minor than one the variance is stationary and bigger than one the variance is explosive.

The GARCH model is applied in financial data series because can modelize the empirical distribution of financial returns such as volatility clustering, excess kurtosis (fatter tails than normal distribution) and heteroscedasticity conditional errors.

The risk metric technique mentioned before uses a particular autoregressive moving average process, exponentially weighted moving average (EWMA) which is a special case of the GARCH model. In this technique ω=0 and α+β=1:

$$\sigma_t^2 = \lambda\delta_{t-1}^2 + (1-\lambda)\varepsilon_{t-1}^2 \qquad (3)$$

J.P.Morgan (1996) considers λ equal to 0.94 for daily data and 0.97 for monthly data and assumes that standardized residuals are normally distributed.

The advantage of using GARCH instead of EWMA is that we do not make a subjective choice of the value for the exponential smoothing constant λ because the parameters of the GARCH model are estimated from the sample data.

[3] Large change in the price of an asset are followed by other large change and small changes are often followed by small change.

GARCH is a symmetric model, which means that positive and negative shocks have the same impact on volatility, but according to Bollerslev et al. (1992) this is not observed in financial returns. Moreover, the price variation is negatively correlated with volatility variation, i.e., when the asset price rises up the variance of the return gets down. These facts along with periods of persistent high volatility, which are followed by periods of persistent low volatility, were called Leverage Effect by Black (1976).

Due to these facts were developed extensions on the GARCH model to capture these asymmetric effects in the financial data.

The most common extensions in use are the GJR-GARCH developed by Glosten, Jagannathan and Runkle (1993), E-GARCH presented by Nelson (1991) and A-PARCH introduced by Ding, Granger and Engle (1993).

The purpose of this dissertation is to test which method is more appropriate to estimate the volatility. This can be done using various methods such as backtesting, stress testing or other techniques.

Backtesting is a method used to validate a model and it basically verifies if the actual losses exceed or not the expectations, i.e., we have to compare the VaR that is given by the method with the portfolio return. This reality checks (Jorion, 1997) are used by risk managers who need to see if their VaR forecast is good. Backtesting is also crucial to the Basel Committee which uses this technique to understand what VaR models they will let the financial institutions use to estimate the capital requirements. Is necessary to verify if the models are appropriate to predict the losses, ensuring that the required capital is sufficient to cover potential losses. The Basel Committee imposed a backtesting to all VaR models, where the actual losses cannot exceed four times the projected losses, considering an annual sample.

There are some different testing methods that have been proposed for backtesting. The statistical tests that will be used to do the model validation are the Kupiec test (1995) which examines the frequency of returns exceeded by VaR estimates and Christoffersen test (1998) that takes into account the independence of exceptions. This characteristic is important to notice, since a good model is indeed able to react quickly to any change in volatility. This subject is very important in the actual market environment where volatile prices are likely to increase and investors seek for a risky portfolio.

Despite its wide use in financial market and the acceptance of VaR as a risk measure, the method has been broadly criticized. Therefore, we should expect many assumptions and simplifications. In addition, VaR only provides the maximum losses and do not establishes any information about the exceeding that can occur. In fact, according to Jorion (1997), the VaR measure is computed to provide the expected loss only under normal conditions, which means that there are some deficiencies in estimating losses, when considering times of inconstant markets. For this reason, the backtesting it is very important.

3. Technical Background

As described in the introduction, the focus of this paper is related with the Historical Simulation since this is the method preferred by most banks as we do not have to make any assumption about the parametric form.

The aim of this dissertation is to evaluate six methodologies of measuring Value-at-Risk, each of which differs in the form they treat volatility. The next chapter will provide a summary of the technical background of the subject.

3.1 Definition VaR

VaR can be defined as the expected maximum loss of a portfolio that will not exceed with some confidence level $(1-\alpha)$ associated.

$$Prob[r_t < -VaR] = \alpha \tag{4}$$

where α is a significance level.

Let P_t be the price of a stock in the time t. The observed return is given by:

$$r_t = \ln P_t - \ln P_{t-1} \tag{5}$$

For a holding period h, the return can be written by:

$$r_t = \ln P_{t+h-1} - \ln P_{t-1} \tag{6}$$

In the equation (4) the VaR does not require to be negative, since we are dealing with losses which are a negative values by definition.

VaR can be expressed by a percentage or by an absolute amount. To estimate the VaR in absolute amount, we can simply multiply the VaR by the portfolio value:

$$VaR_{t(in\ amount)} = VaR_{t(in\ \%)} \times V_t \tag{7}$$

It is possible to define VaR in terms of a probability density function $F_{n,t}(.)$ being represented by its percentile, i.e. the left part of the VaR function. In this case, it is necessary to assume a return distribution, which is generally the normal distribution or student-t.

3.2 Estimating Historical VaR

The $100\ \alpha\%\ h\ day$ historical VaR, in value terms, is the α percentile of an empirical h day discounted P&L distribution or return in the case when VaR is expressed as a percentage.

The computation of historical VaR is very simple as we can see: firstly, we must choose a sample size n and compute the h-days returns for the assets. Following, we have to rank these returns from smallest to largest and obtain the cumulative distribution function (each observation has a probability of 1/n) or use the formula percentile in Excel. The historical VaR is the symmetric value that was computed before.

According to this method, the change in the risk portfolio is related with the historical past, as the past returns are used to predict future returns.

3.3 Equally Weighted Moving Average

In this method we give the same weight for all observations (1/n), old and recent. This is not the best approach, because it does not reflect the current market conditions as will be explained bellow. This methodology, also known as Historical Approach has been popular since the 1990s but has a number of drawbacks.

The historical VaR based on equally weighted returns depends on the choice of the sample size. In this method, this choice is the most important factor to estimate the VaR. For example, to estimate $VaR_{20\ days,1\%}$ we need at least 100 monthly observations (more or less 8 years of data), so as we look further into the past we find ghost features such as the golden years of finance and the most popular financial crisis. As the same weight was given to the returns any extreme market movement will have the same effect on VaR estimative, whether it had happened yesterday or years ago. So, equally weighted returns are not advisable for any VaR model.

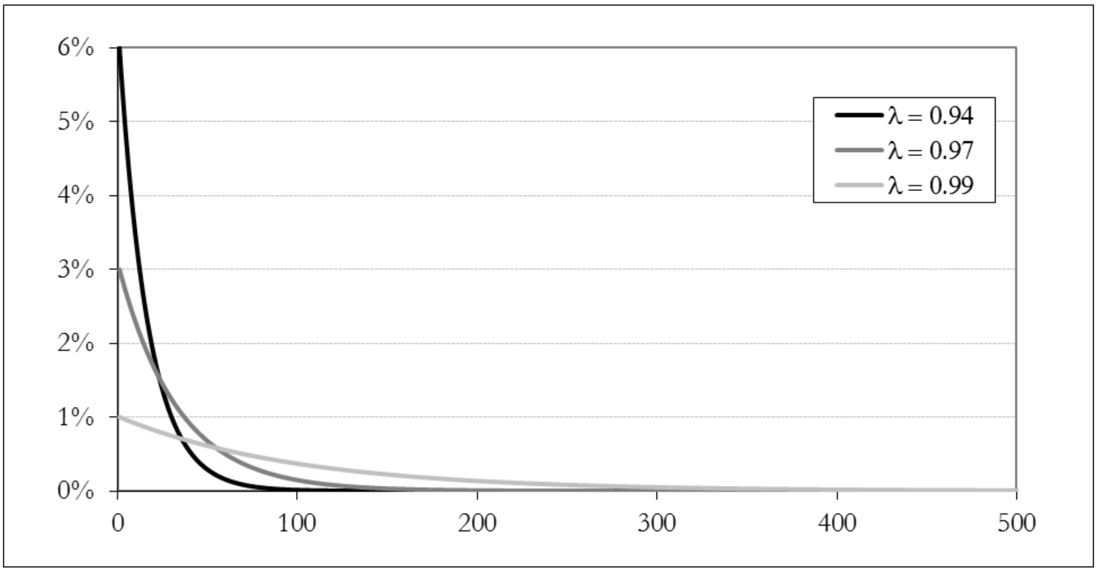
3.4 Exponentially Weighted Moving Averages

The EWMA is the base model of risk software RiskMetric developed by J.P.Morgan (1996) which helped to popularize this method in the mid of 1990s.

In an exponentially weighted moving average approach, instead of giving the same weight for all of the observations, we put more weight in the most recent observations through the fix smoothing constant denoted by λ. The choice of λ is subjective, it has to be between 0 and 1, the larger the value of λ less we overweight recent observations and underweight past returns.

Figure 1: Exponential Probability weight of return

In this figure we can see the lower the value of λ more we overweight recent observations and underweight returns from the past. For example, using a decay factor of 0.94 we give a weight of 6% to the last return while with a decay factor of 0.99 we give only 1%. As we can see observations over 2 or 3 months have little effect in VaR estimation.



The conditional variance is given by:

$$\hat{\sigma}^2{}_t = \lambda\sigma^2{}_{t-1} + (1-\lambda)r^2{}_{t-1} \tag{8}$$

Where:

$\sigma^2{}_t$: Variance calculated in day t

λ: Decay factor

Thus, an extreme return that happened further into the past becomes less important in the average. This approach does not suffer from the “ghost features”.

The weight given to the observations exponentially decrease over time, and the weight attributed to latest observations versus older observations is measured through the smoothing constant(λ). There is no “correct” decay factor. The problem with this approach is that a choice of λ can influence the VaR estimation: if all largest returns took place long time before the VaR estimate then higher values of λ will increase the value of VaR.

3.5. GARCH Models

The Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model captures the volatility clustering of returns and is commonly used by financial institutions.

There are some extensions of GARCH Models such as GJR-GARCH, E-GARCH, A-PARCH that are described in this section.

3.5.1 Symmetric normal GARCH model

The symmetric normal GARCH model assumes the dynamic behavior of the conditional variance, i.e., the value of the period t is conditioned by past values.

A general GARCH (q,p) model can be written as:

$$r_t = \mu_t + \varepsilon_t \tag{9}$$

$$\varepsilon_t | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots | \sim N(0, h_t) \tag{10}$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-1}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-1}^2 \tag{11}$$

where (9) is the conditional mean equation and (11) is the conditional variance equation.

The conditional mean equation (9) is written as a function of exogenous variables and an error term.

The conditional variance (11) is given as function of squared past errors (ARCH term) and the past conditional variances (GARCH term):

Where:

- ω – intercept (constant term)
- ε_{t-1}^2 – error from the previous period
- δ_{t-1}^2 – conditional variance from t-1

The equation above is relative GARCH (q,p) which refers to the presence of the order GARCH term and the order ARCH term. An ordinary ARCH model is a special case of this equation where the GARCH term is equal to zero. Generally, GARCH (1,1) is enough to describe the volatility of financial data.

The sum of ARCH and GARCH coefficients (α+β) measure the rate of convergence, if the sum is close to one, volatility shocks are quite persistent over time.

This model is applied in financial data because it can capture the volatility clustering, excess of kurtosis and heteroscedastic conditional errors.

As previously stated, this model is a symmetric one, so, the impact on the conditional volatility of a positive shock is the same than a negative shock. Due to this, along with the leverage effect (negative correlation between price variation and volatility variation) the GARCH model extension has been developed to capture this behavior in financial data.

3.5.2 GJR-GARCH

In this model was added an extra leverage parameter I_{t-1}. This binary variable aims at modeling the asymmetry between positive and negative market shocks. A negative market shocks has a greater volatility impact than a positive market shock with the same amplitude.

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha\varepsilon_{t-1}^2 + \sum_{j=1}^{p} \beta\sigma_{t-1}^2 + \gamma\varepsilon_{t-1}^2 I_{t-1} \tag{12}$$

Where $I_{t-1} = \begin{cases} 1 & if\ \varepsilon_{t-1} < 0 \\ 0 & otherwise \end{cases}$

The impact of positive shocks market are measured by the ω parameter and the negative shocks are measured by $\omega + \gamma$. If $\gamma > 0$, there is leverage effect and negative shocks cause an increase in conditional volatility.

3.5.3 Exponential GARCH

This model does not impose constraints on the coefficients (not negative) because it formulates the conditional variance equation is in terms of the log of the variance. The log can be negative but the variance is always positive.

$$\ln(\sigma_t^2) = \omega + \sum_{j=1}^{q} \beta_j \ln(\sigma_{t-1}^2) + \sum_{i=1}^{p} \alpha_i \left|\frac{\varepsilon_{t-i}}{\sigma_{t-i}}\right| + \sum_{k=1}^{r} \gamma_k \frac{\varepsilon_{t-k}}{\delta_{t-k}} \tag{13}$$

If $\gamma < 0$, the correlation between price and conditional volatility are negative (leverage effect) and vice-versa.

3.5.4 Assymetric Power GARCH[4]

In this model, the power parameter δ can be estimated rather than imposed and γ parameter is added to capture the asymmetry.

$$\sigma_t^\delta = \omega + \sum_{i=1}^{p} \alpha_i \left(|\varepsilon_{t-1}| - \gamma \varepsilon_{t-1}\right)^\delta + \sum_{j=1}^{q} \beta_j \, \sigma_{t-1}^\delta \tag{14}$$

A positive (negative) value of the γ parameter means that past negative (positive) shocks have a deeper impact on current conditional volatility than past positive (negative) shocks.

[4] All the other models are particular cases of APARCH, depending from the estimate for δ. See Laurent, S., 2003. "Analytical derivates of the APARCH model", Computational Economics.

4. Backtesting Methods

As seen from the previous section, there are many methods to model and forecast conditional volatility. Each method computes different VaR estimations, so it is crucial to estimate the accuracy of each one separately. In order to evaluate the quality of the estimates, the models should be backtested in the appropriate way.
Backtesting is the statistical procedure where actual returns are compared to corresponding VaR estimates of a given period. For instance, in a 100-daily sample, using a 95% confidence level, it is expected to have five exceptions $(5\% \times 100)$, this kind of test is named as tests of unconditional coverage.

Besides it is important, this property is not the only one required, these exceptions should be spread out over time, i.e., independent of each other. This clustering effect on exceptions indicates that the model does not accurately capture the changes in volatility and correlations of the market. Thus, we have to do conditional coverage tests on the data.

In case of the failure of these tests, the assumptions considered can be wrong, which lead the model not to be accurate in predicting the unforeseen events as well as it may be necessary review the entire methodology.

Afterwards, it will be presented the statistical test used by the backtesting Kupiec proportion of failures-test (1995) and Christoffersen interval forecast test (1998).

4.1 Kupiec Test

The Kupiec test method is based on the failure rate, i.e., it quantifies the number of times that the VaR estimate is exceed by the observed returns and simultaneously it compares the corresponding percentage with the level of significance (α). Denoting the number of exceptions as N and the total number of observations as T, the failure rate is computed by dividing N by T. In an ideal situation, this rate should converge towards the significance level (α) with an increased sample size.

However, this method does not consider the returns distribution or by how much the VaR is exceeded but just how often this occurs. Each observation produces a VaR violation exception or not defined as Binary function:

$$I_t = \begin{cases} 1 \; if \; r_t < VaR_t(\alpha) \\ 0 \; if \; r_t \geq VaR_t(\alpha) \end{cases}$$ [5]

The null hypothesis for this test is:

$$H_0: E[I_t] = p = \alpha$$

where p is the failure rate.

The aim of this test is to find out whether the observed failure rate, p, is significantly different from α.

Furthermore, Kupiec test is a likelihood-ratio (LR) test; therefore, this test is asymptotically chi-squared distributed with one degree of freedom, and is expressed in the following way:

$$LR_{uc} = -2\ln[(1-\alpha)^{T-N}\alpha^N] + 2\ln[(1-p)^{T-N}p^N] \sim \mathcal{X}^2_{(1)}$$

If the values of the above test exceed the critical value of the $\mathcal{X}^2_{(1)}$ distribution (3,841 for 95% confidence level and 6,635 for 99% confidence level), the null hypothesis is rejected and the model becomes inaccurate to forecast extreme returns (those exceeding the VaR).

Kupiec define as non-rejection regions for a 95%[6] confidence level, with different probability levels, p and number of observations, as we can see in Table 1.

Table 1: Nonrejection regions for Kupiec test under different confidence levels and sample sizes

This table shows the non-rejection regions for the Kupiec Test at 95% confidence level. As expected the number of failures decreases at higher confidence levels. In addition the range of acceptance is lowest when the sample size increases.

Probability Level p	VaR Confidence Level	Nonrejection Region for Number of Failures N		
		T = 255 days	T= 510 Days	T=1000 days
1%	99%	N < 7	1 < N < 11	4 < N < 17
2,5%	97,5%	2 < N < 12	6 < N < 21	26 < N < 36
5%	95%	6 < N < 21	16 < N < 36	37 < N < 65
7,5%	92,5%	11 < N < 28	27 < N < 51	56 < N < 92
10%	90%	16 < N < 36	38 < N < 65	81 < N < 120

[5] VaR is by definition a negative value

[6] Note that the confidence level of the backtest is not related with the confidence level used in the VaR calculation.

If the percentage of exceptions are systematically higher than α, we come out with the conclusion that the model underestimates the VaR; unlike, if it is systematically inferior to α, the model is overestimating the VaR.

However, there are some critical pointed out against this test. The main critique is that this test only considers the frequency of losses and ignores the number of the times they occur. Those critics go further and highlight the consequences: the test will not be accurate as it does not capture the volatility clustering. Thus, according to Campbell (2005), the backtesting should not rely only on tests of unconditional coverage.

4.2. Christoffersen Test

The previous described test does not test whether VaR model is capable of reacting to changes both in volatility and correlation in a way that exceptions occur independently of each other.

As a consequence, Christoffersen (1998) presented the conditional coverage test, which purpose is to deal with the problem mentioned above. This test has two properties: unconditional coverage and independence. The procedure is the same as in the Kupiec's test; moreover, it extends the test to include also a separate statistic for independence of exceptions. According to these changes, the VaR model is accurate if the probability of occurring a loss in a certain day is independent from the probability of occurring a loss in the previous day.

First and similar to what was done in the previous test, it is necessary to define the variableI_t is 1 if the VaR is exceeded; and 0 if VaR is not exceeded:

$$I_t = \begin{cases} 1 \; if \; VaR \; is \; exceeded \\ \quad 0 \; otherwise \end{cases}$$

Then, it is defined n_{ij} as the number of days when I_t= j (j[7]=0,1) occurred, assuming that I_{t-1}=i (i[6]=0,1) condition verifies on the previous day. Thus, n_{00} presents the number of consecutive no exceptions; n_{01} presents a non-exception followed by an exception and so on. We can see these outcomes in the following table:

[7] (0 = no exception, 1 = exception)

	$I_{t-1} = 0$	$I_{t-1} = 1$	
$I_t = 0$	n_{00}	n_{10}	$n_{00} + n_{10}$
$I_t = 1$	n_{01}	n_{11}	$n_{01} + n_{11}$
	$n_{00} + n_{01}$	$n_{10} + n_{11}$	N

In addition, let p represent the probability of observing an exception if an exception occurred, or not the day before:

$p_1 = \frac{n_{01}}{n_{00}+n_{01}}$ $\quad$ $p_0 = \frac{n_{11}}{n_{10}+n_{11}}$ $\quad$ and $\quad$ $p = \frac{n_{01}+n_{11}}{n_{00}+n_{01}+n_{10}+n_{11}}$

The p_1 represents a proportion of failures in which the last return was not exceeded and p_0 represents a proportion of failures in which the last return was exceeded.

The null hypothesis of this test is that the probability p_1 equals p_0, i.e., today's exceptions should not depend on the exception occurred in the previous day.

The statistic of this test is the following likelihood-ratio:

$$LR_{ind} = -2\ln[(1-p)^{(n_{00}+n_{10})}p^{(n_{01}+n_{11})}] + 2\ln[((1-p_0)^{(n_{00})}p^{(n_{01})}(1-p_1)^{n_{10}}p_1^{n_{11}}] \sim \mathcal{X}^2_{(1)} \quad (14)$$

The combination of the two statistical tests has both characteristics of a good VaR model: correct failure rate must be equal to α and independence of exceptions.

$$LR_{cc}{}^{8} = LR_{uc} + LR_{ind}$$

Each of component is chi-square distributed, with one degree of freedom; thus LR_{cc} also follows a chi-square distribution but with two degrees of freedom ($\mathcal{X}^2_{(2)}$).

In this case, if the test's value is lower than the critical value (5,991 for a 95% confidence level and 9,2103 for a 99% confidence level), then we do not reject the null hypothesis.

[8] Assuming that they are independent.

The evaluation of this test should be done in combination and individually because if null hypothesis was rejected we know the reason for this (inaccurate coverage, clustered exceptions or both). Campbell (2005) said that in some cases it is possible that the model fail in the separate test but pass in the combined test.

5. Data and Methodology

In this chapter will be presented the data and methodology used to compute the VaR through the methods previously described, with the aim of examining the accuracy of the models to estimate the VaR.

5.1. Data

The data set comprises daily returns of several representative economic series of stock indexes, exchange rates and two particular commodities during the period from January 1, 2000 to December 31, 2011. Note that the sample encompasses the most recent crisis in the financial market, the subprime crisis that began in the summer of 2007 and which consequences are still reflected nowadays. During this period many financial institutions suffered large losses all around the world making it reasonable to consider an estimation of the accuracy of the VaR when predicting these losses. So, we will consider two sub periods: before crisis period (from 01-01-2000 to 31-07-2007) and during crisis period[9] (from 1-08-2007 to 31-12-2011) in the analysis. The data embodies the following financial assets:

- Eurostoxx 50 – Representative European market index
- S&P 500 - Representative American market index
- Nikkei 225 - Representative Japanese market index
- FTSE 100 - Representative British market index
- EUR/USD – Exchange rate euro - American dollar
- EUR/JPY - Exchange rate euro - Japanese yen
- EUR/CHF - Exchange rate euro - Swiss franc
- EUR/GBP - Exchange rate euro - British pound
- Gold
- Brent

[9] The onset of the financial crisis is generally accepted to be late July 2007. On August 2007, the European Central Bank provided the first large emergency loan to banks in response to increasing pressures in the euro interbank market.

Consequently, we expect to find interesting tail behaviors all over the world, as seen in the following table and charts. The data were obtained through Bloomberg.

Table 2: Data Analysis

This table summarizes the main descriptive statistics and also shows the Jarque-bera statistic used for to test normality. The null hypothesis of normality is rejected at any confidence level (exception for EuroStoxx indexes in Panel B). For the most of the assets concerned the kurtosis is higher than three and the skewness is negative which is characteristic of financial data.

Panel A: Full Period

	EuroStoxx	S&P	Nikkei	FTSE	EUR/USD	EUR/GBP	EUR/JPY	EUR/CHF	Gold	Brent
Average	-0,024%	-0,005%	-0,027%	-0,007%	0,011%	0,010%	0,001%	-0,009%	0,055%	0,045%
Std. Deviation	1,609%	1,386%	1,602%	1,328%	0,672%	0,519%	0,801%	0,432%	1,175%	2,548%
Kurtosis	4,176	7,047	6,546	5,592	1,127	2,736	6,452	58,746	5,301	3,932
Skewness	0,026	-0,158	-0,389	-0,139	-0,030	0,054	-0,132	2,655	-0,093	-0,210
Minimum	-8,208%	-9,470%	-12,111%	-9,266%	-2,522%	-3,133%	-5,670%	-3,243%	-7,240%	-16,545%
Maximum	10,438%	10,957%	13,235%	9,384%	3,465%	3,148%	7,004%	8,391%	10,243%	16,410%
Observations	3070	3017	2947	3029	3111	3111	3111	3111	3093	2995
Jarque Bera	177,18	2072,78	1618,15	857,75	455,19	10,55	1554,08	406483	686,86	130,48
p-value	0,000	0,000	0,000	0,000	0,000	0,003	0,000	0,000	0,000	0,000

Panel B: Before Crisis Period

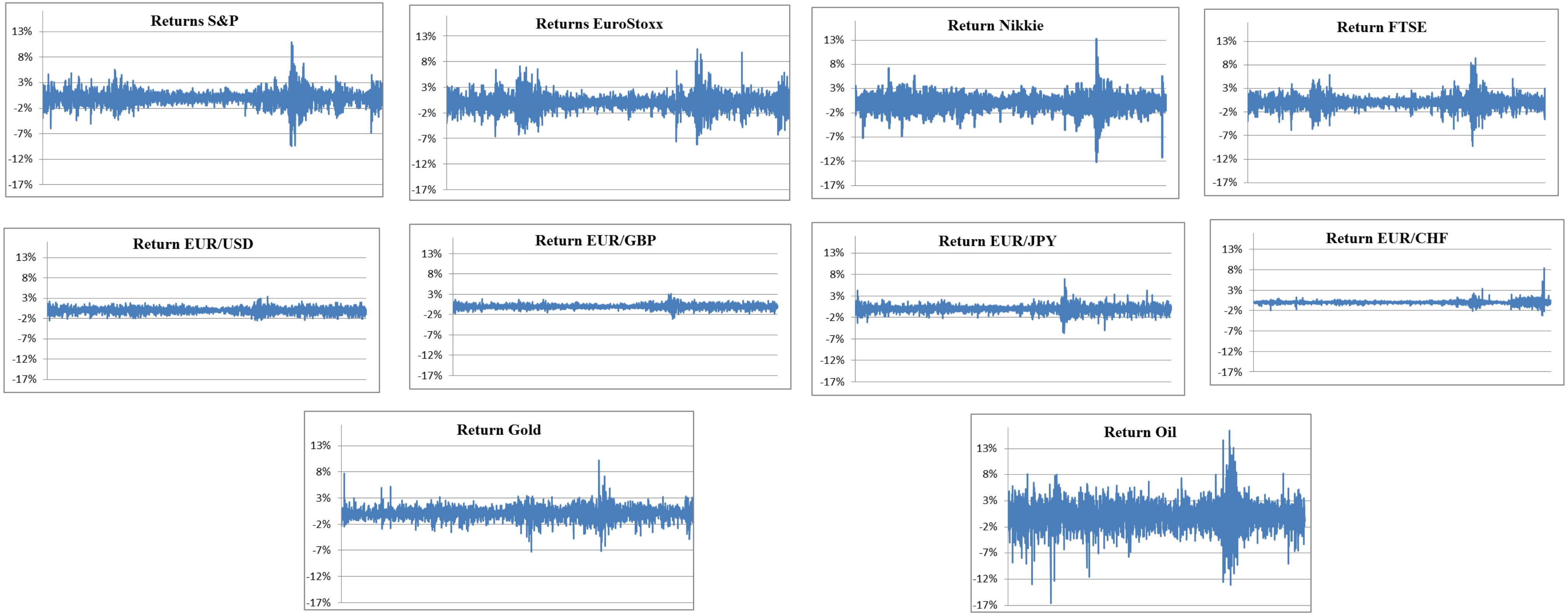
	EuroStoxx	S&P	Nikkei	FTSE	EUR/USD	EUR/GBP	EUR/JPY	EUR/CHF	Gold	Brent
Average	-0,006%	0,000%	-0,007%	-0,005%	0,027%	0,013%	0,027%	0,004%	0,043%	0,055%
Std. Deviation	1,450%	1,112%	1,381%	1,136%	0,588%	0,424%	0,628%	0,247%	1,021%	2,329%
Kurtosis	2,931	2,674	1,779	3,272	0,790	1,159	2,797	3,087	4,653	2,946
Skewness	-0,051	0,059	-0,178	-0,243	-0,075	0,078	-0,157	-0,130	-0,159	-0,539
Minimum	-6,620%	-6,005%	-7,234%	-5,885%	-2,522%	-1,815%	-3,232%	-1,663%	-7,240%	-16,545%
Maximum	7,078%	5,573%	7,222%	5,904%	2,290%	1,909%	4,292%	1,327%	7,649%	8,113%
Observations	1959	1925	1889	1935	1987	1982	1985	1983	1964	1905
Jarque Bera	1,23	9,64	127,34	25,02	406,05	282,03	11,55	6,20	231,85	92,35
p-value	0,270	0,004	0,000	0,000	0,000	0,000	0,002	0,023	0,000	0,000

Panel C: Crisis Period

	EuroStoxx	S&P	Nikkei	FTSE	EUR/USD	EUR/GBP	EUR/JPY	EUR/CHF	Gold	Brent
Average	-0,056%	-0,015%	-0,063%	-0,014%	-0,016%	0,006%	-0,044%	-0,026%	0,074%	0,022%
Std. Deviation	0,019	0,018	0,019	0,016	0,008	0,007	0,010	0,008	0,014	0,029
Kurtosis	4,322	5,927	7,281	5,504	0,783	2,019	4,965	26,673	4,466	4,306
Skewness	0,1087	-0,2252	-0,4736	-0,0643	0,0495	0,0498	-0,0262	2,4673	-0,0616	0,1157
Minimum	-8,208%	-9,470%	-12,111%	-9,266%	-2,434%	-3,133%	-5,670%	-3,243%	-7,195%	-13,065%
Maximum	10,438%	10,957%	13,235%	9,384%	3,465%	3,148%	7,004%	8,391%	10,243%	16,410%
Observations	1110	1091	1057	1093	1123	1128	1125	1127	1128	1089
Jarque Bera	83,01	398,74	846,71	286,22	230,49	45,72	181,21	27458,76	101,75	79,79
p-value	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000

Figure 2: Asset daily returns (Full Period)

The volatility clustering generally associated to financial data: strong (weak) variation are more probable to be followed by strong (weak) variation, it is noted in the assets under consideration. The graphs show the existence of periods of higher and lower volatility, demonstrating that the volatility is not constant and varies over time. The oil asset is the asset with higher volatility, according to the standard deviation.



As we can observe in the tables above, considering full period, gold was the asset that presented higher daily average returns (0,055%) whereas the index Nikkei presented the lowest return in average (-0,03%), being also the second index with higher volatility measure according to the standard deviation estimates (1,602%). When comparing all the assets, Brent has the higher standard deviation (2,55%) as perceived through the chart above. The minimum and maximum returns were also observed in this commodity.

Regarding the differences between the period before and during the crisis, in all assets except gold, were found a decrease in average returns. Considering volatility, the period during crisis presents a higher standard deviation for all indices comparing with the period before the crisis, revealing that we are in the presence of a period of great uncertainty for the markets.

Analyzing the measures of skewness and kurtosis, seven in ten assets have negative skewness, meaning that the distribution is asymmetric on the left side, having heavier tails. When considering the kurtosis, we observe that this measure exceeds three in all assets except the exchange rates EUR/USD and EUR/GBP meaning that distribution is leptokurtic which is to say that the distributions tends to have a heavy tails compared to the normal distribution. It is important to emphasize the value of kurtosis EUR/CHF exchange rate (56.116) meaning that distribution is peaked. Analyzing the two sub periods this increase was justified by the crisis period where the value has increased compared to the period before the crisis. Considering the two sub periods we conclude that this increase was justified by the crisis where the value has increased significantly compared to the previous period.

To confirm this conclusion, the Jarque-Bera statistic used to test the normality, was also computed. The null hypothesis of normality is rejected at any significance level reinforcing the idea of the non-normality. The only exception is the EuroStoxx index in the before crisis period where the null hypothesis is not rejected, which means that based on the sample there is no statistical evidence to reject the null of normality. Furthermore, this conclusion can be confirmed through the values of kurtosis near 3 and skewness near 0.

This characteristics are common among financial assets distributions i.e., kurtosis higher than three and negative skewness indicating that the assets returns does not follow a normal distribution as is assumed in some methods of VaR calculation.

5.2. Methodology

Based in the last known price of the data, the log returns used to estimate the VaR were also computed:

$$r_t = 100 \times log\left(\frac{P_t}{P_{t-1}}\right) \tag{15}$$

With regard to all the approaches discussed earlier, an estimation of the conditional variance was made (except for the equally weighted moving average in which no volatility update must be made) and posteriorly was computed the volatility adjusted returns through the following formula[10]:

$$r_{tj}^* = \delta_{Tj} \times \frac{r_{tj}}{\delta_{tj}} \tag{16}$$

Where:

r_{tj} : Historical return of variable j on day t of the period covered by the historical sample

δ_{tj}^2 : Historical GARCH/EWMA estimate of the daily variance of the percentage change in variable j made for day t at the end of day t-1

δ_{Tj}^2 : Most recent GARCH/EWMA estimate of the daily variance

An adjustment of the historical changes to the current volatility was used for computing the VaR instead of the actual historical percentage changes in market variables.

The backtesting procedure was computed by comparing daily profits and losses with daily VaR estimates using a time period of one year, i.e. 250 trading days.

[10] Hull, J. and A. White, 1998, Incorporating Volatility Updating into the Historical Simulation Method for Value at Risk, Journal of Risk.

According to Jorion (2001) a confidence level of 95% is sufficient for backtesting purposes but in this study will be considered a confidence level of 99% in resemblance with the research made by Hull and White (1998). For each market variable and approach a calculation of the $VaR_{1\,day}$ was performed, with levels of 1% and 5% of significance, and two indicator functions were defined, I_t and J_t for day t. I_t $(J_t) = 1$ if the observed percentage variation would be less than the 5(1) percentile on day t, otherwise the value would be zero, as discussed before in the backtesting technique.

For exponential weighted moving average method, we will set a decay factor at 0.94 according to J.P.Morgan (1996).

All the parameters used in GARCH models were computed through E-views software. On the GARCH models estimation is assumed that errors follow a Student's t distribution. This distribution is usually considered instead to the normal distribution due heavier tails characteristics of the financial assets.

6. Empirical Results

The main objective of this dissertation is to find the best approach to model the volatility to be use in VaR estimation and verifying if changes were observed during the before crisis period and crisis period.

6.1. Full Period

After the estimation of VaR considering six different approaches for two confidence levels, the values of the two statistical tests presented in the section 4 were computed in order to achieve the main conclusions. The values can be analyzed in table 3.

Note that in some models it is not possible to compute the independence test, because the VaR was not exceeded on two consecutive days, i.e., the n_{11} is equal to zero.

When analyzing the results computed from the tests of Kupiec and Christoffersen (LRcc), considering a confidence level of 99%, the HS and E-GARCH models are the only rejected, since their test value is lower than the critical value (9,21). Comparing the models not rejected, we can conclude that the GARCH and EWMA model performed better since it was not statically rejected by any asset, relying on a confidence level of 99%. On average, APARCH model present a better result however is rejected for forecast oil commodity which is the asset with higher volatility. The GJR-GARCH is also rejected for the same asset and EUR/GBP exchange rate.

For 5% tails, the same models were not rejected, but on average, GJR-GARCH and APARCH models had a better performance when compared with the GARCH, since the average of the joint test is lower. Regarding the exchange rate forecasts EUR/GBP, considered a confidence level of 95%, these were rejected in all the proposed models. The GJR-GARCH model is only rejected only for the asset mentioned earlier, the APARCH is rejected to this asset and the EUR/USD exchange rate while the GARCH is rejected also for EUR/JPY. On the EWMA model we reject three of the four exchange rates (EUR/GBP, EUR/JPY and EUR/GBP).

Table 3: Kupiec and Christoffersen Test resultds with a confidence level of 99% and 95%

This table synthesizes the unconditional coverage test (LR_uc), the independence test (LR_ind) and the join test results(LR_cc) for both confidence levels considered. In **bold** we have the values where critical values were not exceeded. Assuming a confidence level of 99%, A-PARCH has better performance on average but it is rejected in the oil commodity, so the best models are GARCH and EWMA. Based on a confidence level of 95%, GJR-GARCH performed better. The rejected models were the E-GARCH and HS.

α=1%	HS			EWMA			GARCH			APARCH			E-GARCH			GJR-GARCH		
	LR_uc	LR_ind	LR_cc	LR_uc	LR_ind	LR_cc	LR_uc	LR_ind	LR_cc	LR_uc	LR_ind	LR_cc	LR_uc	LR_ind	LR_cc	LR_uc	LR_ind	LR_cc
EuroStoxx	11,6011	**1,2246**	12,8257	**1,1307**	**0,5894**	**1,7201**	**2,0040**	**0,4604**	**2,4644**	**2,5259**	-	**2,5259**	11,6011	**1,2246**	12,8257	**2,5259**	-	**2,5259**
S&P	14,7884	**1,2825**	16,0709	**4,2075**	**2,7984**	**7,0059**	**3,5346**	**0,5042**	**4,0387**	**2,9150**	**0,5705**	**3,4856**	13,6048	-	13,6048	**3,5346**	-	**3,5346**
Nikkei	7,2321	8,1871	15,4192	**3,3760**	**2,5415**	**5,9175**	**2,2078**	**0,4728**	**2,6806**	**0,8942**	**0,6817**	**1,5759**	8,1533	7,8557	16,0090	**2,7637**	**0,4131**	**3,1768**
FTSE	16,9567	7,6608	24,6175	**4,7705**	**0,2436**	**5,0141**	**4,0588**	**0,2875**	**4,3463**	**4,0588**	-	**4,0588**	18,2470	5,2608	23,5077	**5,5330**	-	**5,5330**
EUR/USD	**5,5321**	-	**5,5321**	**4,0757**	-	**4,0757**	8,0766	-	**8,0766**	**2,8222**	**0,3639**	**3,1862**	**6,3334**	-	**6,3334**	**6,3334**	-	**6,3334**
EUR/JPY	**2,2748**	**0,0005**	**2,2753**	**3,4230**	-	**3,4230**	**4,0757**	-	**4,0757**	**3,4230**	-	**3,4230**	**2,8222**	**0,0039**	**2,8261**	**3,4230**	-	**3,4230**
EUR/CHF	19,3601	8,0953	27,4554	**2,2748**	**2,7227**	**4,9975**	**3,4230**	**0,3142**	**3,7372**	**2,8222**	-	**2,8222**	16,7834	**5,4411**	22,2245	**4,7792**	-	**4,7792**
EUR/GBP	**6,3334**	17,1863	23,5197	**4,0757**	**2,2489**	**6,3246**	**4,0757**	**0,2686**	**4,3443**	**6,3334**	**1,8346**	**8,1680**	**6,3334**	17,1863	23,5197	8,0766	**4,9070**	12,9836
Gold	13,4835	**3,0509**	16,5344	**0,7056**	-	**0,7056**	**2,9433**	-	**2,9433**	**2,3831**	-	**2,3831**	14,6483	**3,0475**	17,6958	**1,8777**	-	**1,8777**
Oil	10,5242	7,0420	17,5662	**1,4674**	**3,1302**	**4,5976**	6,7035	**1,8508**	**8,5544**	7,5889	**1,7214**	9,3103	9,5003	**4,0245**	13,5248	7,5889	**1,9898**	9,5787
Average	10,8087	**5,3730**	16,1817	**2,9507**	**1,4275**	**4,3782**	**4,1103**	**0,4159**	**4,5262**	**3,5767**	**0,5172**	**4,0939**	10,8027	**4,4044**	15,2072	**4,6436**	**0,7310**	**5,3746**
α=5%	**HS**			**EWMA**			**GARCH**			**APARCH**			**E-GARCH**			**GJR-GARCH**		
	LR_uc	LR_ind	LR_cc	LR_uc	LR_ind	LR_cc	LR_uc	LR_ind	LR_cc	LR_uc	LR_ind	LR_cc	LR_uc	LR_ind	LR_cc	LR_uc	LR_ind	LR_cc
EuroStoxx	11,0051	8,5397	19,5448	**1,8463**	**0,0832**	**1,9295**	**1,4195**	**1,9330**	**3,3525**	**0,7304**	**0,1860**	**0,9164**	11,0051	8,5397	19,5448	**0,4695**	4,3706	**4,8401**
S&P	**3,1974**	12,0394	15,2368	**0,5987**	**0,0280**	**0,6267**	**0,1848**	**0,1091**	**0,2939**	**0,3626**	**0,0618**	**0,4244**	**3,5027**	9,8578	13,3605	**0,1848**	**0,1091**	**0,2939**
Nikkei	**0,7858**	7,4999	8,2856	**0,3925**	**1,5839**	**1,9764**	**0,3925**	**0,0423**	**0,4349**	**0,0769**	**0,0020**	**0,0789**	**0,6400**	6,1320	6,7720	**0,0769**	**0,7702**	**0,8470**
FTSE	**1,4500**	7,6608	9,1108	**0,3706**	4,6944	**5,0651**	**0,9027**	**0,0031**	**0,9058**	**0,6081**	**0,1276**	**0,7357**	**1,0711**	8,1827	9,2538	**0,3706**	**0,4604**	**0,8311**
EUR/USD	**2,5387**	5,6238	8,1626	**2,8080**	4,5006	7,3087	4,0147	**1,8134**	**5,8282**	**0,7130**	8,2221	8,9352	**2,8080**	4,5006	7,3087	4,3484	**0,0007**	**4,3490**
EUR/JPY	**1,5930**	16,0639	17,6570	**1,0242**	9,4232	10,4474	**0,7130**	10,0002	10,7132	**0,5781**	3,9775	**4,5556**	**1,8096**	16,0614	17,8709	**0,5781**	5,3201	**5,8983**
EUR/CHF	11,3610	11,2735	22,6345	**0,1784**	4,7131	**4,8914**	**0,7130**	**0,9475**	**1,6606**	**0,1784**	**0,0122**	**0,1906**	10,8330	11,5861	22,4190	**0,5781**	**0,0023**	**0,5804**
EUR/GBP	**0,8618**	11,2828	12,1446	**1,0242**	6,2492	7,2734	**1,3899**	5,8086	7,1985	**1,0242**	9,5653	10,5895	**1,2003**	9,1422	10,3425	**1,3899**	7,3325	8,7224
Gold	5,0472	**0,6536**	**5,7008**	**1,3788**	**0,3598**	**1,7386**	**2,2713**	**0,1446**	**2,4160**	**2,0281**	**0,0003**	**2,0285**	6,2015	**0,6513**	6,8528	**1,7983**	**0,1656**	**1,9638**
Oil	**2,5873**	6,6904	9,2777	**0,7133**	5,7395	**6,4528**	**0,0390**	3,8609	**3,8999**	**0,1712**	3,9380	**4,1092**	**2,8647**	5,0382	7,9028	**0,0043**	**3,3502**	**3,3546**
Average	4,0427	8,7328	12,7755	**1,0335**	**3,7375**	**4,7710**	**1,2041**	**2,4663**	**3,6703**	**0,6471**	**2,6093**	**3,2564**	4,1936	7,9692	12,1628	**0,9799**	**2,1882**	**3,1681**

The historical simulation and E-GARCH model are the ones that performed worse when forecasting VAR, since the test was rejected in eight assets, for a confidence level of 99%. When taking into account a 95% confidence level, the results are even worse once the test for the EGARCH model was rejected in all assets as we can see in the following table.

Table 4: Number of non-rejected assets

This table presents the number of non-rejected assets in the unconditional coverage test (Kupiec Test), the independence (Christoffersen test) and the joint test (Combined test) considering all the proposed models. Considering the combined test, EWMA and GARCH models were not rejected for all of the assets at 99% confidence level. Regarding a confidence level of 95%, the model with less assets not rejected is the GJR-GARCH while the E-GARCH is rejected in all assets.

	HS	EWMA	GARCH	APARCH	E-GARCH	GJR-GARCH
Combined Test						
α=1%	2	10	10	9	2	8
α=5%	1	7	8	8	0	9
Kupiec Test						
α=1%	3	10	8	9	3	8
α=5%	7	10	9	10	7	9
Christoffersen Test						
α=1%	5	10	10	10	7	10
α=5%	1	5	7	6	1	7

In conclusion, at 99% confidence level, the GARCH and EWMA models performed better since they are not rejected in any of the considered assets. For 5% tails, GJR-GARCH model had a better performance of capturing VaR forecast.

As mentioned before, the test analysis must be done individually in order to determine the cause of rejection towards the test. Further information on this analysis can be found in table 3 and table 4 presented before.

The cause of the rejection of oil commodity and EUR/GBP exchange rate, considering a confidence level of 99%, is due to the inaccurate coverage, meaning that the VaR estimation is exceeded more times than the significance level.

The rejection of the A-PARCH and GRJ-GARCH models to predict the VaR on the exchange rate EUR/GBP (to a level of significance of 95%) is due to clustering exception since the test of independence is rejected in both models. For the rejections of GARCH and EWMA model the reason is the same.

Looking for the tests individually, in average terms we keep the previous conclusions. It is important to notice that for the exchange rate EUR/USD the GJR-GARCH and GARCH models were rejected in Kupiec test but were acceptable on combined regarding a significance level of 5%. Still, on average the APARCH model has a better performance on a significance level of 95%. Considering the significance level of 99% the EWMA performed better.

Regarding the independence test, on average none of the models is rejected for 99% level confidence level. However, when considering 95% level of confidence, the HS and E-GARCH are rejected. On average, GARCH model performs better in capturing volatility clustering considering 1% level of significance but when using a 5% level of significance, GJR-GARCH presents a lower average value meaning a better performance against the others. In this test it is importance to notice the rejection of EWMA model, this model was rejected in five of ten assets (Table 4) meaning that at 5% significance level this model do not capture the changes in the volatility.

Still, it is necessary to reinforce that for some models that last two exceptions do not apply, since the Christoffersen test is inept. From this, we can conclude that the time independence exception is captured by the model, which may not correspond to reality.

6.2. Sub Periods

This section will analyze the main differences between the two sub periods considered (before crisis period and crisis period) in order to check if there are significant differences in the performance of the models due to increased volatility. In this period which begins in August 2007, market was characterized by a higher level of volatility, equity markets all over the world performed poorly. The conclusions will be presented based in the joint tests (Table 6).

Firstly at 99% confidence level, considering the before crisis period, on average all approaches were not rejected. The method that performed better had been the GJR-GARCH model followed by APARCH since these models presents the lowest test value. In these models, along with EWMA and GARCH model, all assets were not rejected. The approach with the worst performance was HS which indicate an underestimation of VaR in this method. The assets rejected in the E-GARCH and HS are the same (EUR/GBP and oil).

Regarding of the period during the crisis, all the approaches were not rejected but the average test value of the HS method is very close to the critical value (9,21) what would be expected given that in this period there was an increase in volatility and in this method none volatility update is done. The methods that perform better were GARCH, GJR-GARCH and EWMA model. These models are not rejected for all assets while APARCH model was rejected for three assets (S&P, EUR/CHF and oil) as we can see in the following table. EurStoxx, S&P, FTSE indexes and oil commodity are the rejected assets by all the remaining models (E-GARCH and HS).

Table 5: Number of non-rejected assets

This table indicates the number of not rejected assets in the joint test (Combined test) considering all the proposed models and the two sub periods. The "scenario" is better in before crisis period where most of the assets were not rejected. The GJR-GARCH is the best model since none of the assets is rejected in all the scenarios considered.

Combined test	HS	EWMA	GARCH	APARCH	E-GARCH	GJR-GARCH
Before Crisis Period						
α=1%	8	10	10	10	8	10
α=5%	6	9	10	8	7	10
Crisis Period						
α=1%	5	10	10	7	5	10
α=5%	3	8	9	7	4	10

Table 6: Results of the joint test regarding the two confidence levels and the two sub periods

This table summarizes the join test results for both of confidence levels considered. In **bold** we have the values where critical values were not exceeded. This section is only presented the combined test since the main objective is to check what are the main differences between the models in two sub periods. On average, considering the before crisis period, the GJR-GARCH model presents better performance having an account a confidence level of 99%, although with a confidence level of 95% the GARCH shows a better performance. The situation is reverse in crisis period. The worse models were the same, the only difference is that in before crisis period, the values of the test were smaller and the models are not rejected.

Panel A: Before Crisis Period

α=1%	HS	EWMA	GARCH	APARCH	E-GARCH	GJR-GARCH	α=5%	HS	EWMA	GARCH	APARCH	E-GARCH	GJR-GARCH
EuroStoxx	**3,3382**	**1,7238**	**1,8302**	**1,3064**	**3,3382**	**0,8423**	EuroStoxx	10,6490	**1,2207**	**1,1016**	**2,0900**	10,4588	**1,7660**
S&P	**7,1133**	**2,5559**	**2,5559**	**7,1133**	**2,1100**	**0,5997**	S&P	9,5498	**1,9902**	**1,5424**	9,5498	9,0552	**1,3901**
Nikkei	**3,2816**	**3,8979**	**3,1221**	**1,7517**	**3,8979**	**3,1221**	Nikkei	**4,2078**	**1,6181**	**1,5961**	**1,5541**	**2,9972**	**3,0227**
FTSE	**8,7209**	**5,2221**	**3,4661**	**2,0353**	**9,8660**	**2,7084**	FTSE	8,6655	**1,3823**	**0,1431**	**1,0037**	9,4881	**0,9085**
EUR/USD	**1,6729**	**1,6729**	**2,9805**	**1,6729**	**2,2844**	**3,7577**	EUR/USD	**1,6489**	**2,3695**	**2,4990**	**2,3745**	**1,7328**	**4,7938**
EUR/JPY	**0,3897**	**2,2999**	**2,2999**	**3,7778**	**0,7268**	**3,7778**	EUR/JPY	**5,0883**	**3,4146**	**2,2691**	**1,2092**	**5,0883**	**2,1974**
EUR/CHF	**2,6652**	**0,1576**	**1,6993**	**1,1713**	**2,6652**	**1,6993**	EUR/CHF	**3,3582**	**0,9279**	**0,5214**	**1,3138**	**3,3582**	**0,8373**
EUR/GBP	15,7615	**5,6245**	**2,6710**	**3,1723**	15,7615	**2,6710**	EUR/GBP	7,3902	**2,2686**	**3,4183**	6,9127	**5,2545**	**3,5064**
Gold	**7,2525**	**0,4575**	**3,1895**	**1,2775**	**7,2525**	**2,4664**	Gold	**2,6124**	6,7461	**4,7978**	**4,4722**	**2,6124**	**4,7978**
Oil	10,7104	**5,8145**	**6,1534**	**5,9248**	10,7311	**6,1534**	Oil	**5,7217**	**5,4785**	**4,2130**	**5,8475**	**5,0728**	**3,8944**
Average	**6,0906**	**2,9427**	**2,9968**	**2,9203**	**5,8634**	**2,7798**		**5,8892**	**2,7417**	**2,2102**	**3,6328**	**5,5118**	**2,7115**

Panel B: Crisis Period

α=1%	HS	EWMA	GARCH	APARCH	E-GARCH	GJR-GARCH	α=5%	HS	EWMA	GARCH	APARCH	E-GARCH	GJR-GARCH
EuroStoxx	11,5490	**1,2091**	**1,9656**	**1,9656**	11,1121	**3,5911**	EuroStoxx	8,4709	**1,3167**	**0,6188**	**0,6188**	7,1519	**1,5377**
S&P	9,9259	**5,3787**	**1,3671**	9,9259	9,9259	**1,3671**	S&P	8,6668	**0,9773**	**1,1786**	8,6668	6,6918	**0,8328**
Nikkei	**7,9782**	**4,0900**	**2,8410**	**2,8410**	**8,2500**	**3,0410**	Nikkei	6,4429	**0,3342**	**0,7063**	**0,6432**	**5,9292**	**0,6432**
FTSE	12,8191	**3,2056**	**2,8734**	**1,3499**	15,8861	**1,3499**	FTSE	**4,0604**	**4,1285**	**1,2162**	**0,3256**	**4,0604**	**0,3256**
EUR/USD	**2,7164**	**4,9077**	**4,9077**	**1,8339**	**2,7164**	**2,7164**	EUR/USD	9,3324	7,7938	8,2739	7,8920	9,3324	**5,8735**
EUR/JPY	**0,0071**	**1,0927**	**0,0071**	**0,0669**	**0,0071**	**0,0669**	EUR/JPY	7,0172	8,5669	**4,9283**	**4,9283**	7,0172	**2,8275**
EUR/CHF	25,6023	**8,8663**	**3,7587**	10,1521	20,6697	**3,7587**	EUR/CHF	22,8608	**3,8806**	**2,3025**	**3,1286**	22,8608	**1,0055**
EUR/GBP	**2,7469**	**3,6718**	**1,0703**	**4,0786**	**2,7469**	**4,0786**	EUR/GBP	**3,1574**	**3,9387**	**4,0938**	**3,9498**	**3,1574**	**3,9862**
Gold	**5,8473**	**0,1639**	**0,1639**	**0,0055**	**5,8473**	**0,0721**	Gold	**1,0359**	**1,0219**	**0,2274**	**1,2452**	**1,0359**	**1,2452**
Oil	12,1013	**1,3845**	**1,3845**	12,1013	9,9773	**4,2630**	Oil	7,8557	**1,4386**	**1,6664**	7,8557	6,7369	**1,4386**
Average	**9,1294**	**3,3970**	**2,0339**	**4,4321**	**8,7139**	**2,4305**		7,8900	**3,3397**	**2,5212**	**3,9254**	7,3974	**1,9716**

Finally at 95% confidence level, regarding the before crisis period, on average all the methods were not rejected despite, the HS and E-GARCH average test being close to the critical value (5,99). The best approach to forecast volatility is GARCH and GJR-GARCH model. In these models, none of assets were rejected. The EWMA model was rejected the asset gold while in APARCH approach was rejected also the gold and the S&P index.

When considering the during crisis period, some changes are observed. The HS and E-GARCH model were rejected in 7 and 6 of the 10 assets respectively (Table 5). In this case, the GJR-GARCH was the model that best performing followed by GARCH. The exchange rate EUR/USD is only not rejected using the GJR-GARCH model. This is also the asset rejected by GARCH while the EWMA was rejected by the same exchange rate and the exchange rate EUR/JPY. Finally, A-PARCH was also rejected for S&P index and oil.

Comparing the two sub periods, we can conclude that the best models to predicting volatility is GARCH and GJR-GARCH models which is consistent with the conclusion made in the analysis of the full period. Considering before crisis period at a 99% confidence level the model that shows the best performance is GJR-GARCH, assuming a 95% confidence level the best model is GARCH, the situation is reversed in the crisis period.

7. Conclusions

The VaR has been considered the most popular market risk measure among financial institutions. All VaR estimation methods use historical data to forecast the future performance of financial assets. Furthermore, the methods under VaR estimation require assumptions that in general are not supported by data, and this is the main reason why this risk measure has been the subject of some criticism.

The most commonly used method to estimate VaR is historical simulation; this approach is the most used by financial institutions due to fewer assumptions needed. The main difference between VaR estimation methodologies is how the assets volatility is computed.

In the theoretical section we discussed some models to predict volatility in order to estimate the VaR from which different results were obtained. Six methodologies were tested and compared by using backtesting techniques.

Regarding the methodologies considered the models that present a better performance to forecast VaR were the GARCH assuming a confidence level of 99%, while considering a significance level of 95% the GJR-GARCH shows better performance.

The models with the worse performance are the Historical Simulation (HS), what would be expected given the disadvantages referred in the literature, and the E-GARCH as can be seen based on the results of the tests.

Among the assets under consideration, the exchange rates EUR/GBP, EUR/JPY and EUR/EUR/USD were the assets in which the models reveal more failures especially with a confidence level of 95%.

Contrary to the common empirical financial result, the asymmetric conditional volatility models did not show superiority in terms of VaR forecasting accuracy when compared to the symmetric GARCH model.

Regarding the differences between the two sub periods analyzed, it can be said that the conclusions are similar for both of them: the models with the best performance in the crisis period and the before crisis period are the same ones as in the full period. The only exception to the conclusions is that the worse performance models considered in both crisis period and full period are not rejected in the before crisis period.

For further developments it would be interesting to analyze other significance levels less than 1% and estimate the VaR by extreme value theory. This is the latest method applied to financial theory. In addition, the VaR prediction could have been tested through the use of other backtesting techniques as Basel traffic light approach or mixed Kupiec-test by Hass (2001).

References

Alexander, C. (2008), *Market Risk Analysis, Volume II: Practical Financial Econometrics*: Wiley.

Alexander, C. (2008), *Market Risk Analysis, Volume IV: Value at Risk Models*: Wiley.

Amman, M. and Reich, C. (2001), VaR for nonlinear financial instruments – linear approximation or full Monte Carlo?, *Financial Markets and Portfolio Management* Vol. 15.

Angelidis, T., Benos, A. and Degiannakis, S. (2004), The use of Garch Models in VaR Estimation, *Statistical Methodology* vol.1.

Barone-Adesi, G., Giannopoulos, K. and Vosper, L. (1999), VaR without correlations for nonlinear portfolios, *Journal of Futures Markets* 19.

Black, F. (1976), Studies of Stock Market Volatility Changes, *Proceedings of the American Statistical Association*, Business and Economic Statistic Section.

Bollerslev, T. (1986), Generalized Autorregressive Conditional Heteroskedasticity, *Journal of Econometrics* 31.

Bollersev, T., Chou, R.Y. and Kroner, K.F. (1992), ARCH modeling in finance: a selective review of the theory and empirical evidence", *Journal of Econometrics* 52.

Boudoukh, J, M. Richardson, and R. Whitelaw (1998), *The Best of Both Worlds*, RISK, May.

Campbell, S. (2005), A Review of Backtesting and Backtesting Procedure, *Finance and Economics Discussion Series*, Divisions of Research & Statistics and Monetary Affairs, Federal Reserve Board, Washington D.C.

Christofferssen, P. (1998), Evaluating Interval Forecasts, *International Economic Review* 39.

Christoffersen, P., Hahn, J. and Inoue, A. (2001), Testing and Comparing Value-at-Risk Measures, *Journal of Empirical Finance* vol. 8.

Danielson, J. and Vries, C. G. (1997), Value at Risk and Extreme Returns, *Manuscript*, London School of Economics.

Dimson, E. and Marsh, P. (1995), Capital Requirements for Securities Firm, *Journal of Finance* vol 50, 821-851.

Ding, Z., C. W. J. Granger, and R. F. Engle (1993), A Long Memory Property of Stock Market Returns and a New Model, *Journal of Empirical Finance* 1.

Dowd, K. (1998), *Beyond Value at Risk: The New Science of Risk Management*, Chichester: John Wiley&Sons.

Duffie, D. and Pan, J. (1997), An overview of Value at Risk. *The Journal of Derivatives*.

Engle, R. and Patton, A. (2001), What good is a volatility model?, *NYU Working Paper No. S-DRP-01-03*.

Engle, R. (1982), Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometria* vol. 50, 987.

Eviews, EViews 6 User's Guide II.

Fama, E. (1965), The behavior of stock-market prices, *The journal of Business*.

Glosten, L., R. Jagannathan, and D. Runkle (1993), On the Relation Between Expected Value and the Volatility of the Nominal Excess Return on Stocks, *Journal of Finance* 48.

Hopper, G.P. (1996), Value at Risk: A New Methodology for Measuring Portfolio Risk, *Federal Reserve Bank of Philadelphia* - Business Review 37 (4), 19-31, July/August.

Hull, J. and A. White (1998), Incorporating volatility updating into the historical simulation method for value at risk, *Journal of Risk*.

Jorion, P. (1997), *Value at Risk: The New Benchmark for Controlling Market Risk*, Chicago: Irwin.

J.P. Morgan (1996), *Risk Metrics TM-Technical Document*, Fourth edition, New York: Morgan Guaranty Trust Company.

Kupiec, P. (1995), Techniques for Verifying the Accuracy of Risk Management Models, *Journal of Derivatives* 3.

Laurent, S., 2003. "Analytical derivates of the APARCH model", Computational Economics.

Mandelbrot, B. (1963), The variation of a certain speculative prices, *Journal of Business* 36.

Manganelli, S. and Engle, R. (2001), Value at Risk Models in Finance, *Working Paper* nº 75, European Central Bank.

Nelson, D. B. (1991), Conditional heteroskedasticity in asset returns: a new approach, *Econometrica* 59.

Perignon, C. and Smith, D. (2006), The level and quality of value-at-risk disclosure by commercial banks, *Working paper*, Simon Fraser University, Vancouver

Sarma, M., Thomas, S. and Shah, A. (2003), Selection of Value-at-Risk Models, *Journal of Forecasting* vol. 22.

Taylor, S. (1986), *Modelling Financial Time Series*, Wiley, Chichester.

Wagster, J. (1996), Impact of the 1988 Basle Accord on International Banks, *Journal of Finance* vol. 51, 1321-1346.

Wiener, Z. (1999), *Introduction to VaR (Value-at-Risk)*, Risk Management and Regulation in Baking, Kluwer Academic Publishers, Boston.