

UNIVERSITY OF JYVÄSKYLÄ
School of Business and Economics

**WOULD STUDENT'S t -GARCH IMPROVE
VaR ESTIMATES?**

Master's Thesis

Spring 2005

Student: Abdelazim Reffat Mohamed

Supervisor: Professor\ Juhani Raatikainen

UNIVERSITY OF JYVÄSKLÄ – SCHOOL OF BUSINESS AND ECONOMICS

Author: Abdelazim Reffat Mohamed

Thesis title: Would student's t -GARCH improve VaR estimates?

Major subject: Economics

Month and year: Spring 2005

Master's thesis: 52 pages (including the cover pages)

ABSTRACT

In this study, GARCH volatility model was applied on Value at Risk methodology. The purpose is to compare between the performance of the plain GARCH (1,1) model that assumes normality of residuals, with the alternative student's distribution GARCH (1,1) model within the VaR framework at the 95% confidence level.

The two alternatives were empirically tested on a highly diversified portfolio consisting of two major European stock indices, The German stock index (DAX), and the French stock index (CAC). VaR estimates generated by each model were compared with actual P/L, and several backtesting techniques were applied to examine the validity of the two models.

Both models generated identical VaR estimates with respect to number and distribution of exceptions. The two models were equally accepted by the tests of Kupiec (1995) and Christoffersen (1998). Lopez (1999) Loss function preferred normal GARCH since it slightly reduced the cost of exceptions, but an empirical comparison has shown that normal GARCH generally over stated the level of risk. Keeping in mind the equality of results from both models, and the associated opportunity cost if normal GARCH was chosen, t -GARCH model should be preferred.

Keywords: VaR, EWMA, GARCH, student's t -distribution, backtesting.

Acknowledgments

This study has been carried out at the school of Business and Economics, University of Jyväskylä, Finland.

I am very grateful to Professor Juhani Raatikainen in having provided me with the opportunity to undertake this work and for his kind patience, valuable guidance, unlimited help and support.

I also wish to thank all individuals in the school staff and friends for their help and support.

My dearest thanks to my parents, and family for their endless help and support, and before all, my greatest gratitude to God for granting and providing me with all the gifts I enjoy in this life.

1. INTRODUCTION	2
2. <i>RISKMETRICS</i>TM VALUE AT RISK (VAR)	4
2.1. Value at Risk: Definition and Components	5
2.2. Value at Risk: Approaches of Estimation	7
2.3. <i>RiskMetrics</i> TM VaR: Parametric Assumptions and Calculation	8
2.4. <i>RiskMetrics</i> TM VaR Drawbacks	10
2.5. Improved <i>RiskMetrics</i> TM Volatility Measure	12
3. AUTOREGRESSIVE AND HETEROSCEDASTICITY	15
3.1. Autoregressive Conditional Heteroscedasticity (ARCH)	16
3.2. Generalized Autoregressive Conditional Heteroscedasticity (GARCH)	18
3.3. GARCH with a student's <i>t</i> - distribution	20
4. OBJECTIVE AND PROCEDURES OF EMPIRICAL RESEARCH	22
5. STATISTICAL INDICATORS	24
5.1. Normality Test	24
5.2. Testing for ARCH effects	26
6. PARAMETERS ESTIMATION AND VAR BACKTESTING	30
6.1. GARCH Parameters Estimation	30
6.2. Backtesting Value at Risk	33
7. DISCUSSION	42
8. CONCLUSION	46
9. LIST OF REFERENCES	47

1. INTRODUCTION

Before Markowitz (1952) presentation of risk as the dispersion or *standard deviation* around the average return or *mean* within the framework of Modern portfolio Theory, there was not any concrete measurement of financial risk. Instead, evaluating the risk of a financial asset involved deep investigation into financial and accounting figures, yielding at the end a more of qualitative than quantitative description of risk, and could not answer the vital questions of “how risky an asset is, compared to another asset?” Markowitz’s finding therefore laid the foundation of modern finance, and opened the way for better understanding of financial risk. Following Markowitz, William Sharpe’s Capital Asset Pricing Model (1961) presented the risk of a security as its covariance with respect to the general market index, known as *Beta*.

By early 1990s, Value at Risk (VaR) as a concept of Risk measurement has evolved when J.P. Morgan published its risk management methodology, widely known as *RiskMetricsTM VaR*. The concept gained a great popularity and became an integral risk management tool used by banks and other organizations, and a standard to monitor and control risk exposures.

Nevertheless, Value at Risk methodology came under tough criticism for the over simplified assumptions of normally distributed returns and constant variance that often leads to underestimation of risk. Empirical evidences have shown that the distributions of return rates of financial time series exhibit fat tails (Zangari 1996), and that volatility is time dependent known as *volatility clustering*. (Mandelbrot 1963)

In 1996, J.P.Morgan suggested the use of Exponential Weighted Moving Average (EWMA) of squared returns as volatility model to capture volatility clustering. The model has improved volatility measurement, but suffered the drawback of fixed parameters (Alexander 1998), and flat volatility forecast (Hubbert 2004).

ARCH model or Autoregressive Conditional Heteroscedasticity Model introduced in 1982 by Robert F. Engle, and for which he won the Nobel Prize in Economics in 2003, provided better volatility measurement with several desired properties. The model was developed further to the more parsimonious Generalized Autoregressive conditional Heteroscedasticity Model, or (GARCH) in short, by Bollerslev in 1986.

Despite the existence of numerous extensions for GARCH models in literature, the aim of this study will be to compare between the basic and original GARCH (1,1), assuming normally distributed residuals, and the alternative student's t-distribution. The two alternative models will be empirically applied on a highly diversified portfolio and compared within the Value at Risk context at the 95% confidence level. The validity of VaR estimates will be tested with Kupiec (1995) Proportion of Failure likelihood test, and Christoffersen (1998) likelihood test for conditional coverage. Then VaR estimates of both models will be compared with respect to Lopez (1999) Loss Function.

This paper will proceed as follow; Definition of Value at Risk, its historical and theoretical backgrounds, as well as, the *RiskMetricsTM EWMA* volatility model are presented in chapter 2. *ARCH*, *GARCH*, and *t-GARCH* volatility models are presented in Chapter 3. Chapters 4 will present the research objective and procedures. Statistical tests are presented in chapter 5. Parameters estimation, and GARCH-VaR backtesting procedures and results are presented in chapter 6, followed by the discussion in chapter 7, and conclusions in chapter 8.

Three software applications are used in the study. MICROSOFT EXCEL is used in calculating log returns series for the portfolio, as well as for comparing estimated VaR with actual P/Ls. PCGIVE is used to run the different statistical tests needed for the study as well as to estimate the parameters for the GARCH models. Finally, RISK DIMENSIONS is used to generate daily Value at Risk estimates over the two years backtesting period.

2. *RISKMETRICS*TM VALUE AT RISK (VaR)

Financial crisis during the 70s and 80s created an overwhelming motivation within large financial institution to develop well functioning Risk Management systems. There was also an urge for a Risk measure that can function as a benchmark among institutions and markets. The most successful and well-known system among them was the *RiskMetrics*TM system developed and used by J.P.Morgan.

The system was for internal use until J.P.Morgan took a bold step of making the system available for public in 1994. It is said that the system was invented when the chairman of J.P.Morgan, *Dennis Weatherstone*, asked his staff to provide him with a daily one-page report indicating risk and potential losses over the next 24 hours. (Dowd 1998:18)

The system originated from the portfolio theory introduced by Markowitz in 1952. The theory assumed that the expected return and Risk of an asset could be estimated, within the normal distribution framework, as the mean and the standard deviation respectively. An estimation of correlation coefficients between several assets is then used to build and optimize an efficient portfolio consisting of several assets that have low correlation with each other. The aim of J.P.Morgan however was to estimate the potential losses, and therefore the focus was on the left tail of the distribution and the estimated volatility of returns. (Dowd 1998:18)

The reasons behind the popularity of Value at Risk as a concept and methodology, in addition to its public release in 1994, lies in its simplicity of providing a single summary statistical figure of possible potential losses within a given horizon of time. Value at Risk (VaR) has provided a benchmarking tool to compare and aggregate risky positions. It served as a performance measurement used by management to evaluate the performance of business units and strategies. It also

served as determinant of capital adequacy; such a measure was strongly needed especially in the banking industry. Banks undertake risky financial positions on daily basis, and therefore in need for an adequate assisting tool to monitor the undertaken risk and determine the minimum capital requirement needed to provide a cushion against unexpected losses. (Wilson 1999:62-63)

2.1. Value at Risk: Definition and Components

“VaR summarizes the expected maximum loss over a target horizon with a given level of confidence interval ” (Jorion, 2001)

Value at Risk is a statistical figure, mostly presented in a monetary term, reflecting the potential change in the value of a single asset or a portfolio resulting from changes in market factors over a specified time interval.

For this statistical figure to come to light, it requires four components. a) The mark to market value of the asset or the portfolio; b) The expected volatility of the given asset/portfolio; c) The time horizon during which the expected losses are to be measured; and d) The degree of confidence for the reported expected losses.

a) Mark to Market:

The position has to be valued based on its current market value and not based on accrual accounting methods. Marking to Market has several benefits. It facilitates the process of profits and losses reporting and therefore making it harder for management to hide undesirable losses behind accrual valuation methods. It also provides management with an instant feedback regarding the investment strategy adopted, thus allowing management to swiftly change failing strategies before losses grow further. As far as VaR is concerned, marking to market reflects the day-to-day development and volatility of markets therefore gives a realistic and reliable value of the evaluated position. (Dowd 1998:8)

b) Expected Volatility:

The riskiness of a market or a position is measured by the level of volatility. It is the most vital component of Value at Risk measurement, and the central debate in most of Value at Risk literature. There are several methods of measuring volatility. The most basic method of estimating the volatility of an asset/portfolio is by estimating the standard deviation of its past returns based on Markowitz's portfolio theory. This method has been challenged on the ground that it assumes volatility is constant over the time, while empirical results show that volatility is time variant with periods of high volatility and periods of low volatility (see early work for Mandelbrot (1963)). Several alternatives to model varying volatility have been introduced. One of the most notable models is the ARCH model introduced by The Noble Prize Winner, Robert F. Engle (1982), and further developed to GARCH by Tim Bollerslev (1986).

c) Time Horizon:

The maximum potential loss depends on the time horizon in which the loss is expected to occur. Time horizon determines the span of time over which VaR is estimated. The choice of time horizon depends on the purpose of measuring VaR. For instance, The Basel Committee requires banks to use the shortest feasible forecast horizon, which is one market day. For illiquid instruments it may take some time to liquidate positions and a longer forecast horizon should be used. (Basel Committee, 1996).

d) Degree of Confidence:

Because VaR is just an estimation of maximum potential future loss, it does not rise up to the level of certainty, therefore it has to be associated with a given level of confidence. The confidence level is the probability that the loss is not greater than predicted by VaR. VaR measures are typically calculated for confidence levels ranging from 95% to 99.9%. The choice of confidence level depends on the purpose of use. For capital adequacy purposes relatively high confidence level is recommended. Basel Committee recommends the 99% confidence level (Basel

Committee, 1996). For internal reporting purposes lower confidence levels may be used. For instance, JP Morgan uses 95% confidence level, Citibank 95.4% confidence level and Bankers Trust reports VaR measures at 99% confidence level. (Dowd, 1998: 53)

2.2. Value at Risk: Approaches of Estimation

Broadly, there are two main approaches to estimate Value at Risk. The Parametric approach and, the Nonparametric approach.

The Parametric approach considers a basic assumption concerning the distribution of assets returns as a starting point, e.g., the normal distribution, the student t-distribution, or the Generalized Error Distribution (GED). These distributions have well-established statistical properties and have predetermined methods and procedures to estimate their parameters, e.g., average mean, standard deviation, skewness, and kurtosis. Furthermore, the use of Normal Distribution is suitable because it allows conversion to different confidence levels. Under the parametric approach, parameters of distribution are estimated, e.g., volatility is approximated by the standard deviation, and then VaR is obtained as a percentile of the estimated distribution corresponding to the confidence level (as formerly stated in point 1.1.). (Jorion, 2001:110-111, 121)

However, according to Dowd (1998), one problem with the parametric approach is its exposition to Model Risk, which refers to the risk that the statistical model may be incorrect, e.g., assets returns departing greatly from being approximated by the normal distribution.

On the other hand, nonparametric methods make no assumptions about the distributions of returns and the estimation of VaR is based solely on empirical distributions of returns, therefore, avoiding the problem of model risk. The most notably known nonparametric method is The Historical Simulation Method.

The Historical Simulation Method makes no assumption regarding the distribution of assets returns; instead, historical observations are let to define the shape of the

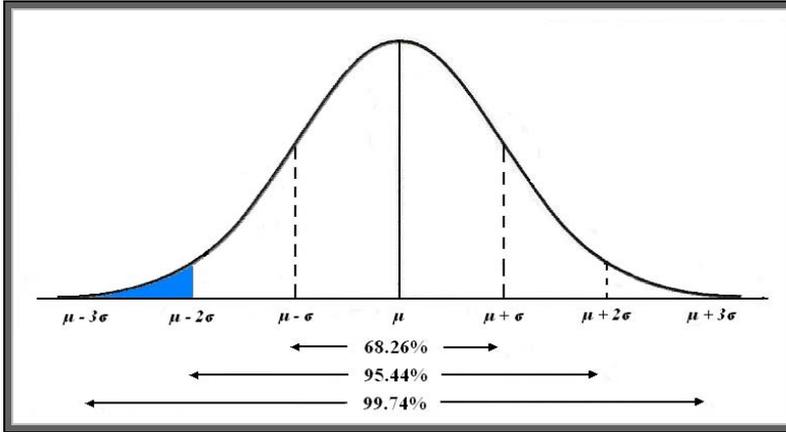
distribution. It assumes however, that the distribution of returns is constant over the sampling period and the future is sufficiently like the past. The first step of a Historical Simulation (HS) is to choose a window of past observations. Then, portfolio returns within this window are sorted in ascending order and the q -quantile of interest is given by the return that leaves $q\%$ of the observations on its left side and $(1-q)\%$ on its right side. If such a number falls between two consecutive returns, then some interpolation rule is applied. To compute VaR for the following day, the whole window is moved forward by one observation and the entire procedure is repeated.

The theoretical simplicity and straightforward implementation of Historical Simulation might be its most notable advantage since it does not need to estimate distribution parameters such as volatilities and correlations. Nevertheless, results obtained by this method depend to a great extent on the chosen window size of past observations. Short window size makes VaR sensitive to recent extreme market events.

Since Historical Simulation does not account for changes in market conditions, all observations included in the chosen window are given the same weight no matter how far into the past, causing slow response to changes in market conditions if an extreme observation falls inside the window. Consequently, choosing a long window size including such extreme but far events might lead to sustained periods of constant VaR until the event in question is dropped out from the sampling window.

2.3. *RiskMetrics*TM VaR: Parametric Assumptions and Calculation

*RiskMetrics*TM *Value at Risk* is a parametric method, assuming that (1) returns follow a conditional normal distribution with the property $X \sim N(\mu, \sigma^2)$, and that (2) the current position is frozen over the forecast horizon. With the normality assumption in hand, the computation process is simplified and reduced to the estimation of two parameters, namely *the mean* μ and *standard deviation* σ . A normal distribution has the property that approximately 68% of the area under the curve lies within one standard deviation ($\mu \pm 1 \sigma$), 95% within $\mu \pm 2 \sigma$, and 99% within $\mu \pm 3 \sigma$.



Graph 1. Normal distribution and Value at Risk

Since Value at Risk is concerned about the expected losses, it focuses on the left tail of the distribution as graph 1 shows. At a given a confidence level such as 95% ($1 - c$), VaR statement indicates that within the chosen time horizon, the probability for losses to reach (into the blue zone) are 5%. Formally, the probability that actual losses exceed Value at Risk estimates at any given level of confidence $1 - c$, and given time horizon h can be presented as:

$$\Pr[W_{t+h} - W_t < -VaR_W(h)] = c \quad (1)$$

Where W_t is the portfolio value at time t and $VaR_W(h)$ is the (negative) Value at Risk of the portfolio over the chosen horizon.

While Value at Risk can formally be presented as:

$$VaR_W(h) = W_t Z_\alpha \sigma_p \sqrt{\Delta t} \quad (2)$$

Where W_t is the portfolio/asset value at time t , Z_α is the constant that gives the appropriate one-tailed confidence interval ($1 - c$) for the standard normal distribution, e.g. $Z_\alpha = 1,96$ for a 95% level of confidence, σ_p is the annualized standard deviation of the portfolio/asset's returns, $\sqrt{\Delta t}$ is the holding period horizon (h) as a fraction of a year.

2.4. RiskMetricsTM VaR Drawbacks

Value at Risk came under strong criticism on two main fronts. First, for the assumption that returns are normally distributed, and second, for the estimated constant variance.

Following the work of Mandelbrot (1963) and Fama (1965), several empirical studies have shown and confirmed that (1) stock returns time series are Leptokurtic, which means that they show excess kurtosis and have fatter tails compared to that of a normal distribution; (2) The distribution of stock returns are skewed either to the right (positive skewness) or to the left (negative skewness); (3) The variance of stock returns are not constant over time, instead, volatility is time dependent and vary over time, which is known as *volatility clustering*.

Kurtosis is a measure of the peakedness of a distribution, while *skewness* is a measure of the degree of asymmetry of a distribution. The kurtosis of a normal distribution is 3. A distribution with excess kurtosis is more peaked and has fatter tails than the normal distribution. A positively/negatively-skewed distribution indicates large values at the direction of skewness. (Aczel, 2002: 38-39)

Within VaR framework, excess kurtosis indicates that tails die off in a slower rate than the normal distribution, implying a greater likelihood of large negative values as well as fatter tail implying that extreme outcomes may happen much more frequently, also a negative skewness indicates a long left tail and therefore large negative values than a normal distribution would suggest. (Jorion, 2001, 93)

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^T (r_i - \bar{r})^2}{(T - 1)} \quad (3)$$

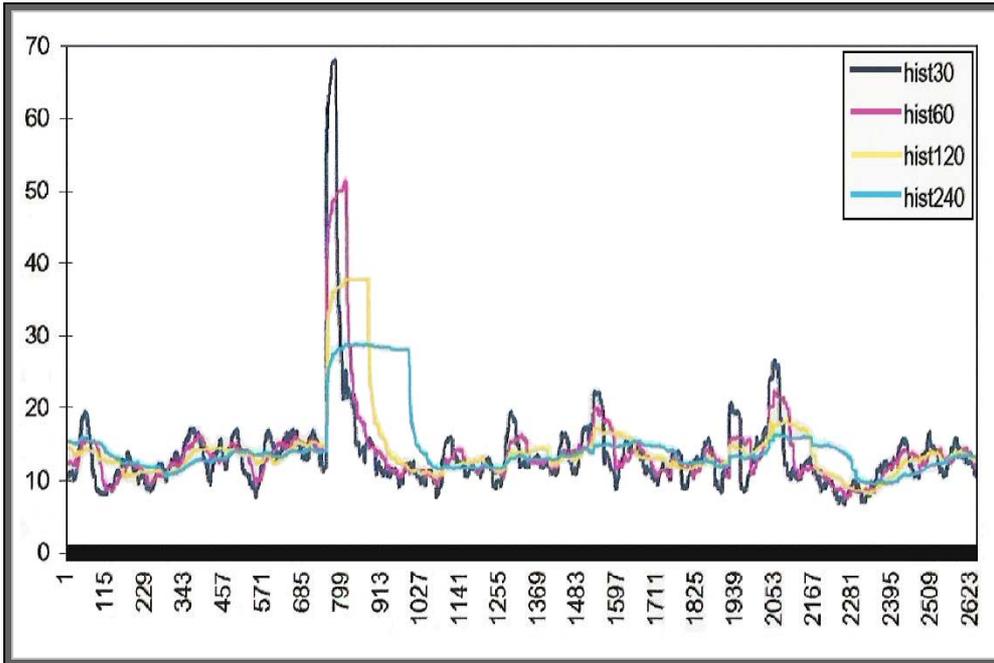
The simple unconditional constant variance as presented in equation (3) assumes that market factor distributions remain constant over time and independent of past realizations. However, Mandelbrot (1963) observed that volatility is time-dependent and varies over time. A phenomenon that became known as *volatility clustering*.

“At closer inspection, however, one notes that large price changes are not isolated between periods of slow change; ... In other words, large changes tend to be followed by large changes-of either sign-and small changes tend to be followed by small changes...” (Mandelbrot 1963:418)

Furthermore, the simple variance ignores the dynamic information in return series and gives equal weight to both early and recent observations. (Dowd 1998:94)

According to Alexander (1998), the problem with equally weighted averages is that extreme market events are treated equally with normal events no matter when the extreme event has taken place in the past. Consequently, one extreme event in any given day will keep the estimated volatility high for a long period even though actual market volatility has returned to normal long ago. Thus volatility estimate will be kept artificially high in periods of tranquility, creating a Ghost feature. (Alexander, 1998:127-128)

Graph (2) illustrates different shapes of Ghost Features created following Black Monday when Equally Weighted Volatility was estimated, using different length of historical observations for the London Stock Exchange (FTSE). The graph shows that the shorter a historical window of observations is used; a more realistic volatility estimate emerges since extreme events occurred in the past would eventually drop down. An interesting contrast can be seen between the volatility estimated using the most recent 30 past observations, and 240 past observations. Equally Weighted Volatility estimated with a 240 past observations created an unnecessary long lasting Ghost of high volatility following one extreme event (Black Monday), even though the market has recovered far earlier.



Graph 2. Historic volatilities of the FTSE from 1987 to 1995, showing “Ghost Features” of Black Monday and other extreme events (Alexander, 1998)

As a conclusion, maintaining the normality assumption in the presence of the excess kurtosis and negative skewness, as well as, adopting the constant variance estimate in the presence of volatility clustering will cause Value at Risk to greatly underestimate the risk level and therefore the expected losses. On the other hand, Value at Risk could overestimate the level of Risk due to the ghost feature created by the adoption of equally weighted averages.

2.5. Improved *RiskMetrics*TM Volatility Measure

In 1996, *RiskMetrics*TM introduced The Exponentially Weighted Moving Average (EWMA) as an improved methodology to forecast volatility. It takes into account the higher weight of recent observations and therefore reduces the Ghost Feature effect on volatility estimates. This approach has two important advantages. First, it allows volatility to react faster to shocks in the market since recent data carry more weight than data in the distant past. Second, following a shock, the volatility declines exponentially as the weight of the shock observation falls. In contrast, the use of a *simple moving average* leads to relatively instantaneous changes in the standard deviation once the

shock falls out of the measurement sample, which, in most cases, can be several months after it occurs. (Zangari, 1996: 87)

The exponential weighting is achieved by using a “smoothing constant” parameter λ often called the decay factor, constrained to be ($0 < \lambda < 1$) and, the expected volatility at time t using the Exponentially Weighted Moving Average (EWMA) is:

$$\hat{\sigma}_t^2 = (1 - \lambda) r_{t-1}^2 + \lambda \hat{\sigma}_{t-1}^2 \quad (4)$$

Where $\lambda \hat{\sigma}_{t-1}^2$ is the previous volatility forecast weighted by the decay factor, and $(1 - \lambda) r_{t-1}^2$ is the latest squared return weighted by $(1 - \lambda)$.

The closer the decay factor gets to unity the more weight is given to older observations and therefore the smoother the series becomes. The smaller the decay factor the greater weight is given to recent observations than older ones. If the decay factor is equal to unity the model is reduced to the equally weighted model. (Alexander, 1998:131)

Since the weight of older observations are declining exponentially, *conditional that the decay factor is less than unity*, it is therefore clear that the value of the decay factor plays a vital role in deciding the size of the moving window of past observations included and used to forecast the volatility, because no matter how far past observations are available, the older observations will eventually die off and drop down from the moving window. (Mina and Xiao, 2001: 15)

Mina and Xiao (2001) add that by using the idea that the magnitude of future returns corresponds to the level of volatility, one way to choose the appropriate value of the decay factor is obtained by comparing the volatility obtained with a certain decay factor to the magnitude of future returns.

According to Zangari (1996), *RiskMetricsTM* has chosen to use two different values for the decay factor. For daily data set, the decay factor is ($\lambda = 0,94$), and for monthly data set is ($\lambda = 0,97$). The reason for these different values is that *RiskMetricsTM* construct volatility and correlation forecasts on over 480 time series, which requires 480 variance forecasts, and 114,960 covariance forecasts. Since these parameters comprise a

covariance matrix, the optimal decay factors for each variance and covariance forecast are dependent on each other.

While the covariance matrix has to meet three conditions 1) the variances cannot be negative. $\sigma_1^2, \sigma_2^2, \sigma_3^2 \dots \sigma_n^2 \geq 0$, 2) the covariances σ_{xy}^2 and σ_{yx}^2 must be equal, which means that the matrix is symmetric, 3) the correlation coefficient between $r_{x,t}$ and $r_{y,t}$ is constrained to $-1 \leq \rho \leq +1$, the values of the optimal decay factors for each variance and covariance forecast must be consistent with the properties of the covariance matrix where they belong. However, as the number of covariance matrices grows large, as it is the case of *RiskMetricsTM*, the task of choosing optimal decay factors grows more complicated, therefore the optimal decay factors were restricted to the two previously mentioned values using the *Root Mean squared Error (RMSE) criterion* in which the variance is expressed explicitly as a function of the decay factor. Thus, the chosen decay factor is the one that produces the minimum variance forecast errors. (Zangari, 1996: 97-98)

However, Alexander (1998) criticizes *RiskMetricsTM* decay factor for daily data set as too high for most markets, while the decay factor for monthly data set leads to the emergence of “Ghost features” due to the 25-day equally weighted variance.

3. AUTOREGRESSIVE AND HETEROSCEDASTICITY

Autoregressive AR models gained great popularity in modeling financial time series. An autoregressive process attempts to regress a given time series on itself. In an autoregressive process, the observation y_t is explained and depends on its previous P observations. (Alexander, 1998:133)

An autoregressive process of order p , or $AR(p)$ can be expressed as:

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + u_t \quad (5)$$

Reduced to an $AR(1)$ can be expressed as:

$$y_t = c + \phi_1 y_{t-1} + u_t \quad (6)$$

In this model, the level of y_t is predicted by the conditional mean

$E(y_t | y_{t-1}) = c + \phi_1 y_{t-1}$ and therefore, conditional on past observation the conditional mean of y_t is allowed to change over time, while the error term u_t is assumed to be a Gaussian white noise with a zero mean $E(u_t) = 0$ and constant variance $Var(u_t) = \sigma^2$, unconditional on past observations. (Hamilton, 1994:657)

The assumption that error terms have a constant variance σ^2 is called homoscedasticity. This assumption has been strongly rejected; instead, empirical studies have found clear evidence showing that variance is changing over time or heteroscedastic $Var(u_t) = \sigma_t^2$. Several empirical studies have shown that financial time series exhibit varying volatilities over time in which periods of high (low) volatility tend to cluster together forming the volatility clustering phenomena, see Mandelbrot (1963) and Fama (1965).

3.1. Autoregressive Conditional Heteroscedasticity (ARCH)

Motivated by this important finding, Robert Engle (1982) introduced the Autoregressive Conditional Heteroscedasticity (ARCH) Model in which the variance of the error term is allowed to be conditional on past information and therefore allowed to vary over time.

Given an autoregressive process such as equation (6), according to Engle's ARCH model, the error term u_t can be decomposed into a systematic part and a random part such as $u_t = z_t \sqrt{h_t}$, where $z_t \sim (0,1)$ is a white noise with mean zero and variance unity, while h_t is a scaling factor that depends on the squared error term at time $t-1$ known as *ARCH* (1) :

$$h_t = \alpha_0 + \alpha_1 u_{t-1}^2 \quad (7)$$

Which can easily be expanded to an *ARCH* (p) :

$$h_t = \sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \dots + \alpha_p u_{t-p}^2 \quad (8)$$

Where h_t is the conditional variance, $\alpha_0, \alpha_1, \dots, \alpha_p$ are parameters subject to the constrain $\alpha_0 > 0$, $\alpha_1, \dots, \alpha_p \geq 0$ (for if equal zero, the variance is constant), to ensure nonnegative volatility. (Hamilton 1994: 659; Gujarati 1995:434)

Given such definition, the error term for the ARCH (1) model will have the following properties:

1- The mean of the error term $u_t = \text{zero}$:

$$u_t = z_t \sqrt{h_t} = z_t \sqrt{\alpha_0 + \alpha_1 u_{t-1}^2}$$

$$E_{t-1}[u_t] = E_{t-1}[z_t] \sqrt{\alpha_0 + \alpha_1 u_{t-1}^2} = 0$$

2- The error term u_t has conditional variance (heteroscedastic) given by:

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2$$

3- The unconditional variance for the error term is given by:

$$\sigma^2 = \frac{\alpha_0}{(1 - \alpha_1)}, \text{ Existing only if } \alpha_0 > 0, \text{ and } |\alpha_1| < 1$$

4- The error term has zero covariance:

$$E_{t-1}(u_t u_{t-1}) = u_{t-1} E_{t-1}(u_t) = 0$$

(Johnston and DiNardo, 1997: 195-202)

5- The ARCH process has excess kurtosis than a normal distribution, as shown by Bera and Higgins (1993):

$$\frac{E(u_t^4)}{\sigma_u^4} = 3 \left(\frac{1 - \alpha_1^2}{1 - 3\alpha_1^2} \right) > 3$$

This property implies that the ARCH process has heavier tails than the normal distribution, making the process attractive to model the volatility of financial time series, known for the characteristic of excess kurtosis and fat tails. (Bera and Higgins, 1993: 315)

Given the assumption that $z_t \sim i.i.d. N(0,1)$, the set of parameters, $\alpha_0, \alpha_1, \dots, \alpha_p$ denoted as θ , are then estimated with the Maximum Likelihood function of the normal distribution, which is more convenient to be expressed in its log term as:

$$l(\theta) = -\left(\frac{T}{2}\right)\log(2\pi) - \left(\frac{1}{2}\right)\sum_{t=1}^T \log(h_t) - \left(\frac{1}{2}\right)\sum_{t=1}^T \frac{u_t^2}{h_t} \quad (9)$$

(Engle 1982:994; Hamilton 1994: 661)

3.2. Generalized Autoregressive Conditional Heteroscedasticity (GARCH)

ARCH model is regarded as one of the most important development in modeling financial times series for its ability to capture the variability of volatility across time “volatility clustering”. However, the ARCH model suffered drawbacks in Empirical applications as it often required long lag length (P) in order to explain certain data. (Bollerslev 1992:9)

To come over this difficulty, Bollerslev (1986) introduced an alternative and a more flexible lag structure model in which the conditional variance depends, not only on the lagged squared error terms, but also on the lagged conditional variances. The new model is called Generalized Autoregressive Conditional Heteroscedasticity or *GARCH*(p, q) model.

$$h_t = \sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i u_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j} \quad (10)$$

Where $P > 0$ and $q \geq 0$, for if $q = 0$, the process is reduced to an *ARCH*(p), and if $q = p = 0$, the error term is a white noise random walk process. To ensure a well-defined process, it is also required that all the parameters to be nonnegative. That is $\omega > 0$, $\alpha_i \geq 0$ and $\beta_j \geq 0$. (Bollerslev 1986:309)

The simplest GARCH process *GARCH*(1,1) can be seen as an equivalent to an infinite *ARCH*(∞) model with geometrically declining parameters. (Bollerslev et al.1992:10)

To demonstrate this, consider the following GARCH (1,1) model:

$$h_t = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta_1 h_{t-1}$$

Replacing the term (h_{t-1}) with its recursive form, the model can be written as:

$$h_t = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta_1 (\alpha_0 + \alpha_1 u_{t-2}^2 + h_{t-2})$$

And by continuing the recursive substitution, the model will have the following general representation:

$$h_t = \sum_{i=1}^{\infty} \beta_1^i \alpha_0 + \alpha_1 \sum_{i=1}^{\infty} \beta_1^{i-1} u_{t-i}^2$$

In addition to this important feature, the GARCH (1,1) model has similar but slightly different properties than those of the ARCH model:

- 1- The zero mean of the error term: $E(u_t) = 0$.
- 2- The error term u_t has conditional variance (heteroscedastic) given by:

$$h_t = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta_1 h_{t-1}$$

- 3- To ensure that the model is covariance stationary, it is necessary that the sum of the parameters is less than one ($\alpha + \beta < 1$). Consequently, the unconditional variance for the error term is given by:

$$\sigma^2 = \frac{\alpha_0}{(1 - \alpha_1 - \beta_1)}$$

- 4- The error term has zero covariances:

$$E_{t-1}(u_t u_{t-1}) = u_{t-1} E_{t-1}(u_t) = 0$$

- 5- The process has excess kurtosis than a normal distribution, as shown by Bollerslev (1986):

$$\frac{E(u_t^4)}{\sigma_u^4} = 3 \left(\frac{1 - (\alpha_1 + \beta_1)^2}{1 - (\alpha_1 + \beta_1)^2 - 2\alpha_1^2} \right) > 3$$

Given a *GARCH* model such as equation (10), the used procedures to estimate the model's parameters are similar to those used for the ARCH model by maximizing the

log likelihood function as presented in equation (9). (Bollerslev 1986:315; Hamilton 1994: 661)

Several extensions to the GARCH model have been introduced in literature in recent years. One of these extensions can be of interest in relation with the *RiskMetrics*TM EWMA model is the Integrated GARCH model or I-GARCH, introduced by Engle and Bollerslev (1986), and motivated by the empirical finding that the sum of the parameters comes too close to unity with frequency financial data. In this model the sum of parameters is allowed to equal unity ($\alpha + \beta = 1$) therefore creating a non-stationary GARCH model. One interesting feature in this model is that when ($\alpha + \beta = 1$), denoting $\beta = \lambda$, the *I - GARCH* (1,1) will have the following presentation:

$$h_t = \sigma_t^2 = \omega + (1 - \lambda)u_{t-1}^2 + \lambda\sigma_{t-1}^2 \quad (11)$$

If the constant $\omega = 0$, the model is equivalent to an infinite EWMA such as the model used by *RiskMetrics*TM. (Alexander, 1998:131)

3.3. GARCH with a student's t - distribution

The Student's t -distribution is often seen as interesting and simple alternative to the normal distribution since it is characterized by heavy tails, which is a stylized fact in financial time series data and assets returns. According to Dowd (1998), student's t -distribution provides an easy way to capture uncertainty since it penalizes the lack of information regarding portfolio's standard deviation with a wider VaR confidence interval, and was proposed by Wilson (1993) as an alternative to the normal distribution.

Addressing this alternative assumption, Bollerslev (1987) introduced an extension to GARCH model in which the unconditional distribution of the error term u_t is assumed to be a non-Gaussian with heavier tails than the Normal distribution. Therefore, $u_t = v_t \sqrt{h_t}$, where $v_t \sim t(0, S_t, \nu)$ is a Student's t -distribution with mean zero, ν

degrees of freedom to be estimated as an additional parameter, and the scale parameter $S_t = h_t(\nu - 2)/\nu$. Therefore, the student's t-distribution PDF for the error term u_t is expressed as:

$$f(u_t) = \frac{\Gamma[(\nu+1)/2]}{\sqrt{\pi} \Gamma(\nu/2)} (\nu-2)^{-1/2} h_t^{-1/2} \left[1 + \frac{u_t^2}{h_t(\nu-2)} \right]^{-(\nu+1)/2} \quad (12)$$

Where $\Gamma(\cdot)$ is the gamma function.

The model's parameters are then estimated by maximizing the log likelihood function with respect to the set of parameters, including the degree of freedom, subject to the constraint $\nu > 2$.

The log likelihood function for the student's t-distribution can therefore be expressed as:

$$l(\theta) = T \log\{D\} - (1/2) \sum_{t=1}^T \log(h_t) - [(\nu+1)/2] \sum_{t=1}^T \log \left[1 + \frac{u_t^2}{h_t(\nu-2)} \right] \quad (13)$$

Where;
$$D = \frac{\Gamma[(\nu+1)/2]}{\sqrt{\pi} \Gamma(\nu/2)} (\nu-2)^{-1/2}$$

(Hamilton 1994: 662)

4. OBJECTIVE AND PROCEDURES OF EMPIRICAL RESEARCH

Given the theoretical presentation in chapters two and three, an empirical implementation for the two volatility methodologies (GARCH, and t -GARCH) is carried out in this part of the paper. Even though a plain GARCH model is capable of modeling leptokurtic residuals as shown in chapter three, several empirical studies suggested that a t -GARCH model would outperform a plain GARCH model on the grounds that residuals of financial time series exhibit higher kurtosis and fatter tails for a plain GARCH to model (see Bollerslev (1987) and Alexander (1998)). The main objective is to examine whether a t -GARCH would outperform a plain GARCH model when those two models are applied on Value at Risk at the 95% confidence level.

To achieve this objective, a hypothetical portfolio is constructed from two major European indices, namely, The French Stock Index (CAC) and the German Stock Index (DAX). For simplicity, the two indices have equal weight in the portfolio. Even though one index can be a sufficient proxy for a well-diversified portfolio, The reason however to construct the portfolio of these indices, in addition to avoid country specific risks, is that France and Germany can be seen as two major economic powers in Europe and therefore a portfolio consisting of those two indices can be a simple representation for the continent, as well as, a reflection of general global market condition and volatility.

The data used in the study is a daily time series covering the period from 16 February 1994 to 31 December 2002. The data set is then divided into two parts. The first part, from the beginning of the data set until the end of 2000 (approximately 7 years consisting of 1734 daily observations), is used to estimate the volatility parameters for each volatility model. The second part, starting from beginning of 2000 until the end

2002, is used for the backtesting of each estimated Value at Risk. The indices daily values are transformed into daily log-returns series with the following equation:

$$r_t = \log\left(\frac{P_t}{P_{t-1}}\right) = \log P_t - \log P_{t-1} \quad (14)$$

The new series of indices returns are then combined to form the hypothetical portfolio returns series consisting of the two indices return series with equal weights as follow:

$$R_{Portfolio} = (Weight_{DAX} \times R_{DAX}) + (Weight_{CAC} \times R_{CAC}) \quad (15)$$

Normality and ARCH tests are performed on the Portfolio returns and residuals respectively. Parameters for GARCH (1,1) and t-GARCH (1,1) models are estimated with the maximum likelihood method as presented in equation (9) for residuals assumed to be normality distributed, and equation (13) for residuals assumed to be distributed as a student t-distribution. Two years Value at Risk are then simulated at 95% confidence level, with one day forecast horizon for each volatility model. Estimated Value at Risk for each model is then compared with actual P/L, and backtested with Kupiec (1995) Likelihood Ratio Test for unconditional coverage, as well as Christoffersen (1998) test for conditional coverage, and Lopez “*Loss Function*” test .

5. STATISTICAL INDICATORS

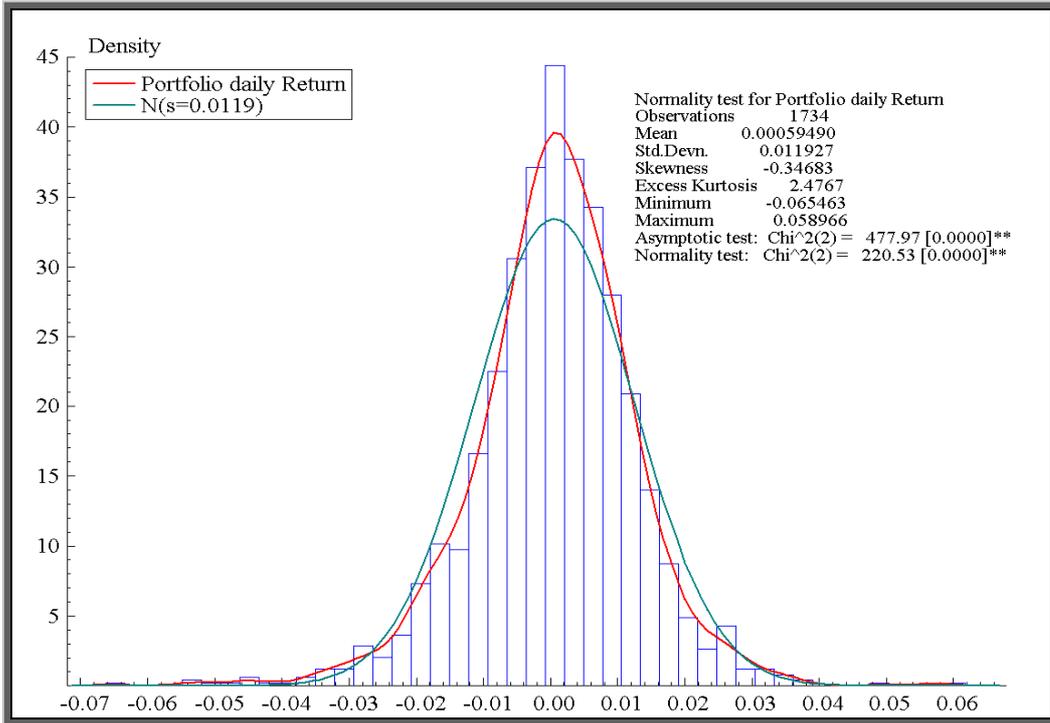
5.1. Normality Test

The Jarque-Bera (JB) test of normality is a test of excess kurtosis and skewness. A normal distribution has zero skewness, and kurtosis of three. The test, as presented in equation (16), is an asymptotic test and follows the chi-square distribution with 2 degrees of freedom.

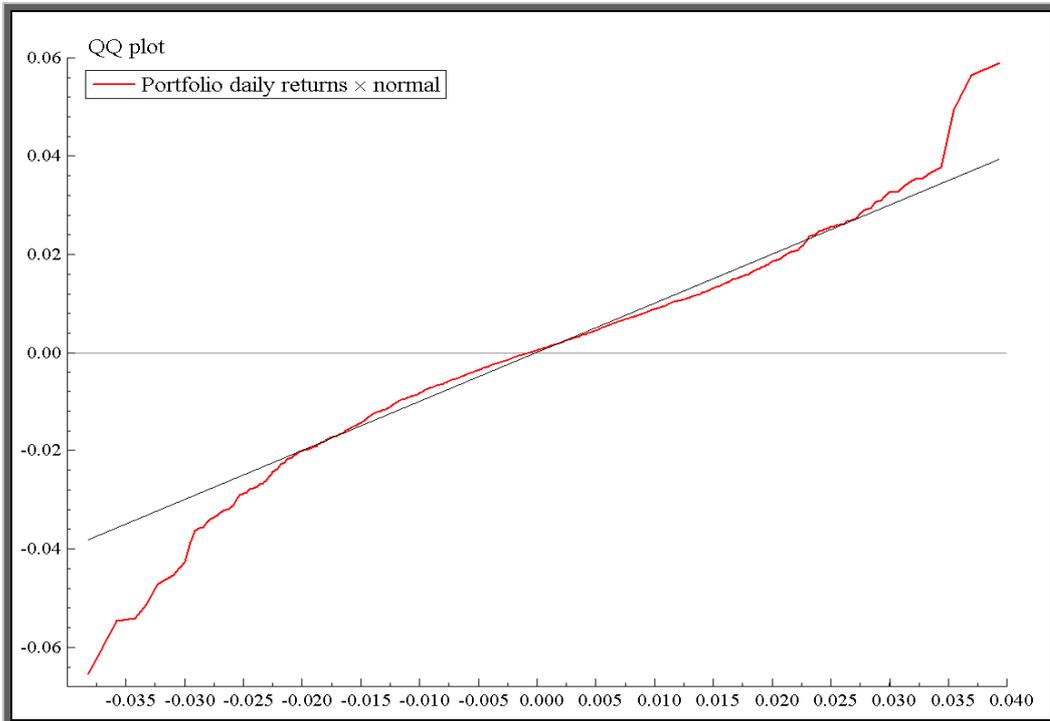
$$JB = n \left[\frac{S^2}{6} + \frac{K^2}{24} \right] \quad (16)$$

Where n is the number of observations, S represents the skewness of the distribution, and K represents the excess kurtosis.

Applying the (JB) normality test on the distribution of returns for the sampling period, as shown in graph (3), suggests that Portfolio's returns are not normally distributed. The value yielded by the normality test (220.53) has zero probability, therefore formally rejecting the null hypothesis of normality. The distribution shows significant excess kurtosis of (2,47) but relatively insignificant skewness of (-0,34). The distribution also has a heavy tails specially on the left side compared to a normal distribution. This can be seen more clearly with the QQ plot in graph (4) showing that the distribution greatly deviates from the normal distribution specially on the tails on both sides of the distribution but with even fatter tail on the left side as suggested in literature regarding the distribution of financial returns, implying a higher likelihood of negative returns than that suggested when Value at Risk is computed under the normality assumption.



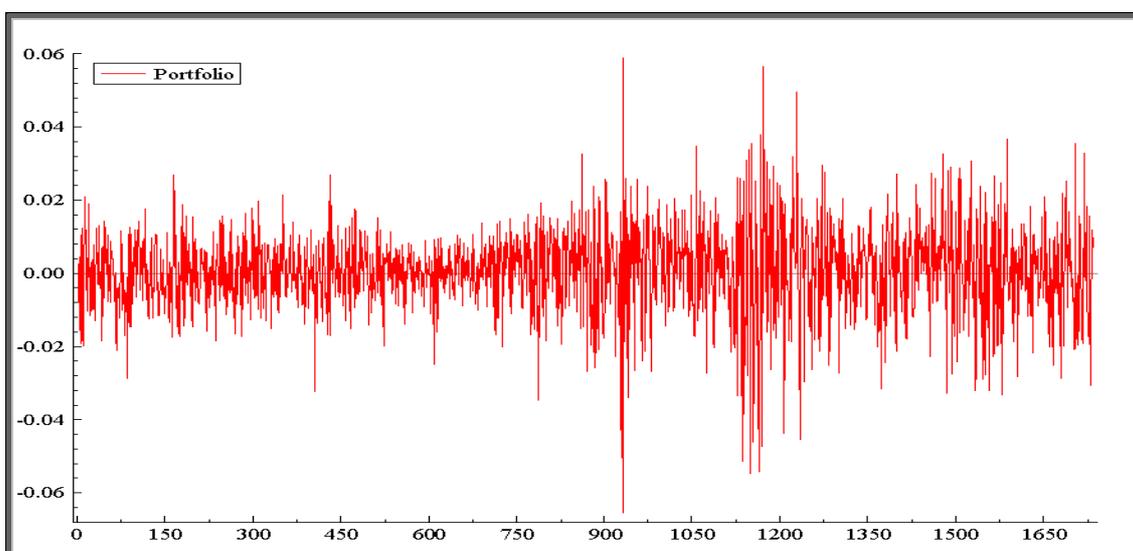
Graph 3: Portfolio distribution and Normality test



Graph 4: QQ plot for Portfolio against normal distribution

5.2. Testing for ARCH effects

ARCH models depend heavily on the error term of the mean equation. The basic assumption is that the error term is normally distributed but yet conditionally heteroscedastic. Unless it is proven that it has varying or heteroscedastic variance, the error term would otherwise be considered as a constant white noise and there would be no need for a varying volatility model to take place and a constant volatility model would be placed in use instead.

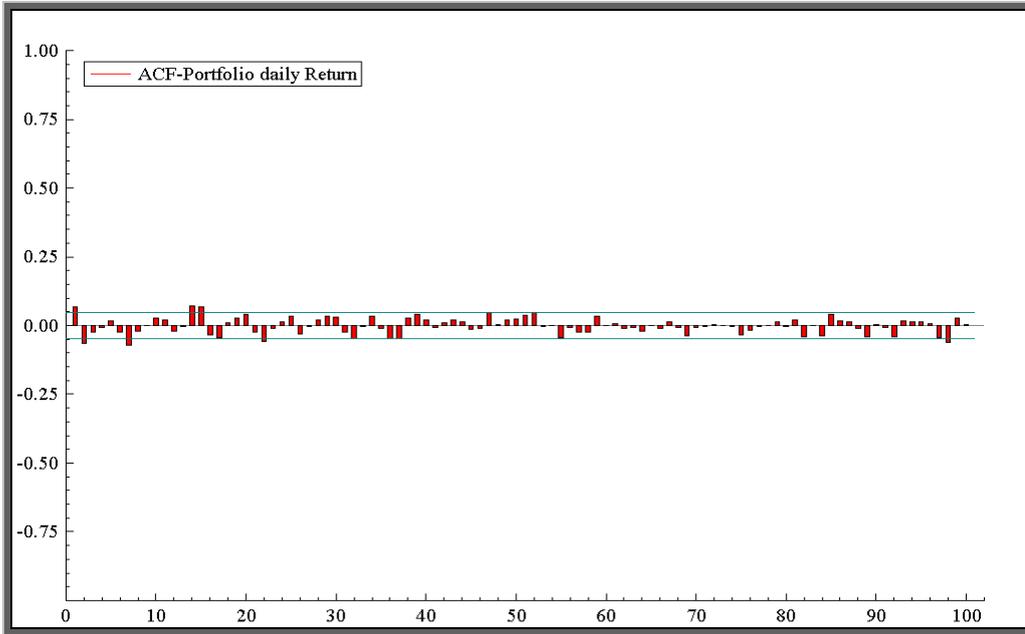


Graph 6: Portfolio's returns against time

A visual inspection of graph (6) show that portfolio's returns have varied greatly across time, with periods of extremely high volatility and periods of relatively lower volatility. For instance, the period from observation 900 till the end of the sampling period experienced very high volatility compared to a more tranquil period between observations 450 and 750 (nearly one year). This inspection suggests the existence volatility clustering, and therefore testing the existence of heteroscedasticity and ARCH effects would be of great interest.

In this study, the mean equation from which residuals are obtained and tested for the application of GARCH models was chosen to be a first order Autoregressive [AR (1)] as presented in equation (6). The main reason for choosing this representation for the mean, in addition to reduce the risk of possibly omitting a relevant variable before testing residuals for ARCH effects, is the stylized fact of significant existence of

autocorrelation in financial time series. This can be confirmed by inspecting graph (7) showing the significant existence of autocorrelation between the first 100 lags of daily returns.



Graph 7: ACF for daily returns up to 100 lags

To test for the existence of heteroscedasticity and ARCH effects in the sampling period, two tests are used; The Portmanteau statistic for autocorrelation, and The ARCH test.

The Portmanteau Q-statistic (also called the Ljung-Box statistic) tests the existence of residuals autocorrelation. The Null Hypothesis (H_0) is that there is no autocorrelation, while (H_1) is the existence of autocorrelation. The test follows the chi-square distribution under the null hypothesis and has the following representation.

$$Q_{LB} = n(n+2) \sum_{j=1}^p \left(\frac{r_j^2}{n-j} \right) \sim \chi_p^2 \quad (17)$$

Where n = sample size, p = lag length, and

$$r_j = \frac{\sum_{t=j+1}^n u_t u_{t-j}}{\sum_{t=1}^n u_t^2} \quad (18)$$

The test was applied on the sampling period several times with different lag lengths (10, 20, 30, 50, and 100 lags). Table (1) below shows the results of the Q_{LB} -statistic for the 6 different lag lengths. All the results suggest rejection of the null hypothesis in favor of the alternative hypothesis of existence of autocorrelation, implying a dependence relationship between the residuals and therefore departing from the white noise assumption. All values generated from the Q_{LB} -statistic are too large at both confidence levels (95% and 99%) with nearly zero probability except the first 10 lags in which the null hypothesis is rejected at only the 95% but not at the 99% confidence level.

Q Portmanteau Test statistic for residuals	
Portmanteau (10):	Chi ² (9) = 20.255 [0.0164]*
Portmanteau (20):	Chi ² (19) = 47.462 [0.0003]**
Portmanteau (30):	Chi ² (29) = 61.716 [0.0004]**
Portmanteau (50):	Chi ² (49) = 88.003 [0.0005]**
Portmanteau (100):	Chi ² (99) = 139.51 [0.0046]**

Table 1: ACF, PACF and Portmanteau statistic

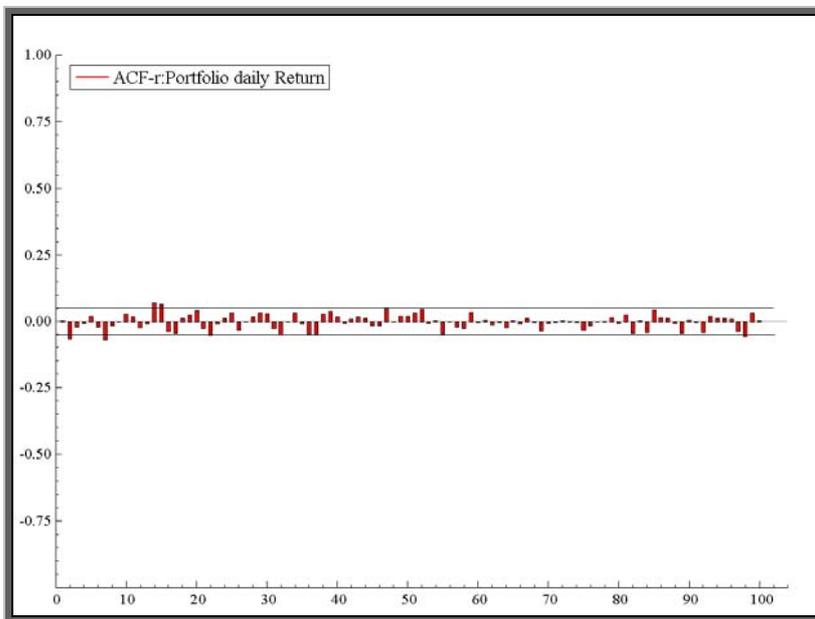
The next test applied on the data set is the ARCH test proposed by Robert Engle (1982), based on the Lagrange Multiplier (LM) principle. The null hypothesis is that the ARCH coefficients are equal zero, that is $H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_p = 0$. The test proceeds by regressing the squared residuals on a constant and P lagged residuals. From the residuals of this regression, a test statistic is calculated as TR^2 , where T is the number of observations and R^2 is the coefficient obtained from the regression. Under the null hypothesis, The LM statistic asymptotically follows a χ_p^2 distribution, which will be rejected if the test statistic exceeds the critical value.

Here the test was applied several times on the sampling period with different lags (from 2 to 10 lags) to make sure that ARCH effect indeed exists in the data.

Lags	F-Value	Probability
2	91.734	[0.0000]**
3	66.489	[0.0000]**
4	54.647	[0.0000]**
5	46.602	[0.0000]**
6	40.338	[0.0000]**
7	40.060	[0.0000]**
8	35.468	[0.0000]**
9	31.566	[0.0000]**
10	28.502	[0.0000]**

Table 2: ARCH Test results from 2 to 10 lags

As shown in table (2), The null hypothesis was strongly rejected at both confidence levels (95 and 99%) up to ten lags, and strongly suggesting the existence of ARCH effects.



Graph 7: Residuals ACF for the 100th lag.

Graph (7) provides an illustrative summary to the previously presented results, showing the strong existence of significant dependence between residuals up till the 100th lag respectively, and confirming the existence of ARCH effects in the data.

6. PARAMETERS ESTIMATION AND VaR BACKTESTING

6.1. GARCH Parameters Estimation

The parameters for the GARCH (1,1) and t-GARCH (1,1) models were estimated as presented in equations (9) and (13) respectively. To ensure models stationary, restriction was imposed on the sum of the parameters to be less than one ($\alpha + \beta < 1$). Values of estimated parameters for both models are presented in tables 3, and 4 below. The presented parameters are the constant, the ARCH, and the GARCH parameters in table (3), with an additional estimated parameter for the student's t-distribution degree of freedom in table (4). Each estimated parameter is followed by its robust standard error (a more precise standard error for heteroscedastic large sample size), and t-value. The sum of GARCH parameters, and the AIC (A goodness-of-fit criteria which weigh the better fit provided by an additional predictor variable against a penalty for its inclusion) complete the tables.

Estimated Parameters for a restricted GARCH (1,1) model			
Parameter	Value	Robust-SE	t-value
Constant	1.37547e-006	6.475e-007	2.12
ARCH (α_1)	0.0709637	0.01516	4.68
GARCH (β_1)	0.920121	0.01623	56.7
$\alpha_1 + \beta_1$	0.991084		
AIC	-6.23358737		

Table 3: Estimated Parameters for restricted GARCH (1,1) (modified PCGive output)

Estimated Parameters for a restricted Student t-GARCH (1,1) model			
Parameter	Value	Robust-SE	t-value
Constant	1.09176e-006	5.982e-007	1.83
ARCH (α_1)	0.0665699	0.01388	4.80
GARCH (β_1)	0.926782	0.01568	59.1
DF	≈ 14		
$\alpha_1 + \beta_1$	0.993352		
AIC	-6.24137135		

Table 4: Estimated Parameters for restricted student t -GARCH (1,1) (modified PCGive output)

As it can be seen from the results, the estimated parameters for the two models are very close. Also, all parameters are significantly different than zero at 95% confidence except for the constant of the t -GARCH model. As suggested in literature, the sum of $\alpha_1 + \beta_1$ parameters are on the brinks but less unity; a requirement for a covariance stationary, mean reverting process. Its closeness to unity however implies that the process is mean reverting but with slower rate. It can also be noticed that the t -GARCH model suggests slightly higher volatility persistence compared to plain GARCH model explained by the GARCH lag parameter β_1 . Expressed differently, the t -GARCH model expects that a market shock would take slightly longer time to die off compared to plain GARCH.

The difference between the parameters' values of the two models is possibly due to the fact that a GARCH model with student t -distribution is capable of capturing more extreme events on the left tail of the distribution compared to a GARCH model with normal innovations, thus, providing more precise estimations. This is supported by the Akaike criterion (AIC) that gives preference to the t -GARCH over plain GARCH with respect to data fit. The t -GARCH has approximately 14 degree of freedom suggesting that residuals distribution are not too far from normal (the higher the degree of freedom, the closer the distribution gets to normal).

Another interesting comparison would be between GARCH estimated parameters, and *RiskMetrics*TM EWMA decay factor. The predefined *RiskMetrics*TM decay factor for daily observations is (0.94), while GARCH's estimated volatility persistence parameters are approximately (0.92) for both GARCH models. Therefore, there is a reason to believe that *RiskMetrics*TM decay factor assumes higher volatility persistence than markets, as criticized by Alexander (1998). On the other hand, *RiskMetrics*TM EWMA returns parameter defined by (1- λ), reflecting volatility response to market shocks, appears to underestimate volatility responsiveness to market shocks compared to those estimated by GARCH models.

Finally, to verify that GARCH models have successfully modeled residuals heteroscedasticity, scaled residuals (u_t/σ_t) were tested for autocorrelation and ARCH effects.

Q Portmanteau Test statistic for scaled residuals			ARCH test for scaled residuals		
Lags	Value	Probability	Lags	Value	Probability
10	7.0884	0.6279	2	0.24856	0.7799
20	20.182	0.3227	4	0.72874	0.5723
30	29.746	0.4268	6	0.51812	0.7950
50	44.342	0.6622	8	0.69104	0.6998
100	111.06	0.1918	10	0.79362	0.6351

Table 5: Portmanteau and ARCH Test for scaled residuals

The results presented in table (5) for the Q and ARCH tests show that GARCH estimated conditional volatility has successfully modeled residuals' heteroscedasticity. The results of the Q test suggest accepting the null hypothesis of no autocorrelation in scaled residuals up to 100 lags. The same conclusion is obtained from the results of the ARCH test. The test results for 2, 4, 6, 8, and 10 lags all suggest accepting the null hypothesis of no presence of ARCH effect in scaled residuals.

6.2. Backtesting Value at Risk

6.2.1. Background and Literature Review

The Basel Committee on Banking Supervision imposed on banks and other financial institutions, such as investment firms, a minimum capital requirement to meet adverse market losses based on Value at Risk estimates. Recognizing their ability to develop more sophisticated methods to estimate VaR, in 1996, the Basel committee allowed financial institutions to use their own risk management models to determine their own VaR and capital requirement. For this purpose, the minimum capital requirement must be sufficient to cover possible losses over 10 days holding period with daily VaR estimated at 99% confidence level.

To verify the adequacy of the used VaR model of each institution, The Basel Committee required a backtest to be performed over a period of one year (250 trading days) in which realized day-to-day returns are compared to estimated VaR. Based on the number of violations or exceptions (realized losses exceeding estimated VaR) different scaling factors are imposed and multiplied by the estimated VaR to obtain the minimum capital requirement as shown in table (6).

Zone	Violations	Scaling Factor
Green Zone	0 to 4	3.00
Yellow Zone	5	3.40
	6	3.50
	7	3.65
	8	3.75
	9	3.85
Red Zone	10 or more	4.00

Table 6: violations over 1 year and their corresponding scaling factor (Jorion, 2001)

Apart from the Basel Committee's regulations, different holding periods and confidence levels may be used for internal purposes to evaluate the validity and compare between alternative VaR models, e.g. one day holding period and 95% confidence level.

The most basic used test to check the validity of VaR models is The Proportion of Failure Likelihood Ratio Test, introduced by Kupiec (1995). The test evaluates the model from two sides, e.g. a model under this test would be rejected if the number of

violations were too high, indicating the failure of the model to estimate VaR, and from the other side, it would be rejected also if it was too conservative with very few or no violations, since the former would underestimate the level of risk and drive the firm to undertake dangerously risky positions, while the later would lead to the loss of lucrative and feasible investments.

The test is based on the probability under the binomial distribution of observing x exceptions in the sample size T .

$$f(x) = \binom{T}{x} p^x (1-p)^{T-x} \quad (19)$$

An accurate VaR model should provide VaR estimates with unconditional coverage (\hat{p}) , given by the failure rate $\left(\frac{x}{T}\right)$, equal to the desired coverage (p) , given by the chosen confidence level (1% for 99%, and 5% for 95% confidence levels). Therefore, under the null hypothesis $H_0 : \hat{p} = p$, the appropriate likelihood ratio is given by:

$$LR_{uc} = -2\ln\left((1-p)^{T-x} p^x\right) + 2\ln\left((1-\hat{p})^{T-x} \hat{p}^x\right) \quad (20)$$

Which is asymptotically distributed Chi-square with one degree of freedom. Consequently, the null hypothesis would be rejected if $LR_{uc} > 3.84$.

VaR Confidence Level	Non-rejection region for number of violations x		
	$T= 255$ days	$T= 510$ days	$T= 1000$ days
99%	$x < 7$	$1 < x < 11$	$4 < x < 17$
97.5%	$2 < x < 12$	$6 < x < 21$	$15 < x < 36$
95%	$6 < x < 21$	$16 < x < 36$	$37 < x < 65$
92.5%	$11 < x < 28$	$27 < x < 51$	$59 < x < 92$
90%	$16 < x < 36$	$38 < x < 65$	$81 < x < 120$

Table 7: Non rejection regions for a given confidence levels and sample sizes (Jorion, 2000)

Even though the test is good in sorting out ill performing VaR models, it's unconditional coverage represents a major drawback since it does not take into account models ability to capture time varying volatility, and accepts models no matter how

exceptions are distributed over the sample. Exceptions clustering might be an alarming signal of a volatility model misspecification, and a failure to react to market shocks.

Addressing this issue, Christoffersen (1998) introduced the likelihood ratio test for conditional coverage, by extending the LR_{uc} to specify that exceptions must be independently distributed. The test first defines the indicator of exceptions as:

$$I_t \begin{cases} 1, & \text{if } r_t < VaR_{t|t-1} \\ 0, & \text{if } r_t \geq VaR_{t|t-1} \end{cases}$$

Then defines the number of days in which state i is followed by state j as T_{ij} , and the probability of observing an exception conditional on state i the previous day as π_i . A likelihood test for independence is then carried out to test the hypothesis that the failure rate is independently distributed. The test is calculated as:

$$LR_{ind} = -2 \ln \left(\frac{(1-\pi)^{(T_{00}+T_{10})} \pi^{(T_{01}+T_{11})}}{(1-\pi_0)^{T_{00}} \pi_0^{T_{01}} (1-\pi_1)^{T_{10}} \pi_1^{T_{11}}} \right) \sim \chi_1^2 \quad (21)$$

$$\text{Where; } \pi = \frac{T_{01} + T_{11}}{T}, \pi_0 = \frac{T_{01}}{T_{00} + T_{01}}, \text{ and } \pi_1 = \frac{T_{11}}{T_{10} + T_{11}}$$

Consequently, the likelihood test for conditional coverage is $LR_{cc} = LR_{uc} + LR_{ind}$, calculated as:

$$LR_{cc} = -2 \ln \left(\frac{(1-P)^{T_1} P^{T_0}}{(1-\pi_0)^{T_{00}} \pi_0^{T_{01}} (1-\pi_1)^{T_{10}} \pi_1^{T_{11}}} \right) \sim \chi_2^2 \quad (22)$$

Clearly Christoffersen's test provided a better approach to evaluate VaR models, but if one assumes that two or more models generated identical number, and distribution of exceptions, passing both tests (Kupiec and Christoffersen), which model would be preferred?

Lopez (1999) proposed an evaluation technique attempting to answer this question using a loss function that includes the magnitude of exceptions into the evaluation process. The proposed form of the loss function is:

$$\varphi_t \begin{cases} 1 + (r_t - VaR_{t|t-1})^2 & \text{if } r_t < VaR_{t|t-1} \\ 0 & \text{if } r_t \geq VaR_{t|t-1} \end{cases} \quad (23)$$

According to this loss function, the best model is the one that minimizes the cost of exceptions. It appears however that the test evaluates models from the regulator's perspective rather than the financial institution, as it prefers the model with the wider VaR band regardless of the opportunity cost banks and financial institutions would incur by setting aside excessive capital reserve during periods of "no exceptions".

Departing from failure rates and "exceptions based" evaluation techniques, Berkowitz (2000) proposed a different method of risk models evaluation that involves the transformation of all outcomes into an independently and identically distributed series based on Rosenblatt (1952) transformation.

$$X_t = \int_{-\infty}^{y_t} \hat{f}(u) du = \hat{F}(y_t) \quad (24)$$

Where y_t is the realized return and $\hat{f}(\bullet)$ is the forecasted density function. Then as shown by Rosenblatt, X_t is independent and identically distributed U(0,1). Berkowitz then applies an extension to Rosenblatt transformation in which an *iid* $N(0,1)$ series Z_t is created by transforming the observed portfolio returns with the inverse of the standard normal distribution function.

$$Z_t = \Phi^{-1} \left(\int_{-\infty}^{y_t} \hat{f}(u) du \right) = \Phi^{-1}(\hat{F}(y_t)) \quad (25)$$

The null hypothesis that observations are normally and independently distributed is then tested against an $AR(1)$, $Z_t - \mu = \rho(Z_{t-1} - \mu) + \varepsilon_t$, where the parameters μ , ρ are the conditional mean and $AR(1)$ coefficient respectively, while ε_t is a white noise random variable with mean zero and variance σ^2 .

Then to test the null hypothesis that $(\mu, \rho, \sigma^2) = (0, 0, 1)$, the appropriate likelihood ratio test is asymptotically distributed Chi-square with three degrees of freedom and has the following representation:

$$LR_{dist} = -2(L(0,0,1) - L(\hat{\mu}, \hat{\rho}, \hat{\sigma}^2)) \quad (26)$$

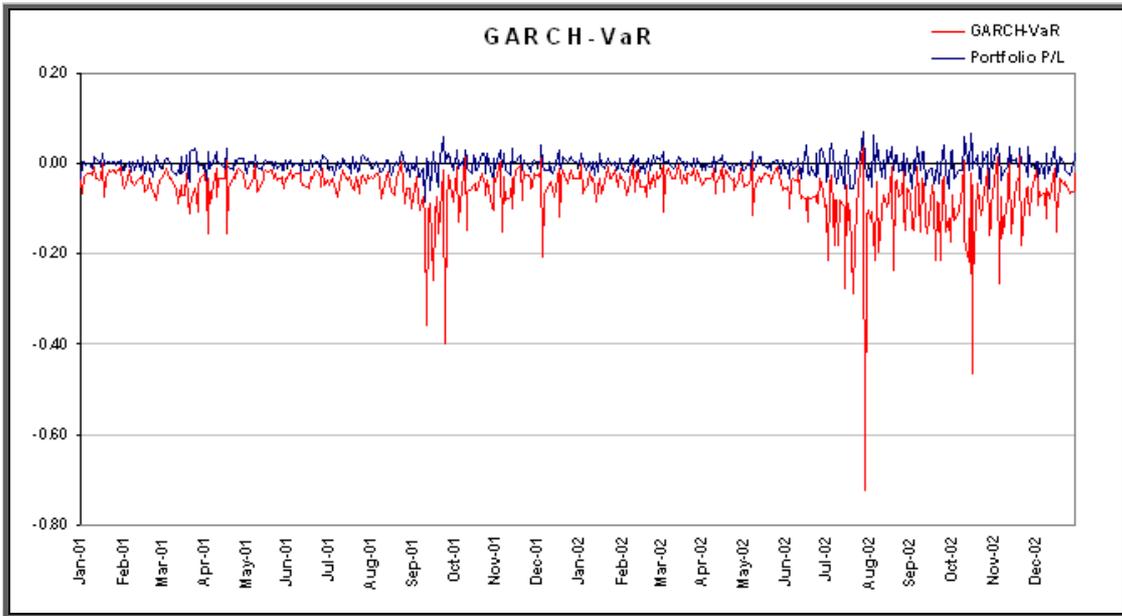
Even though this test has higher power compared to the other tests previously presented with respect to inaccurate risk models detection, according to Christoffersen (2002), and Campbell (2005), it requires more information concerning the shape of the left tail of return distribution, which is difficult to acquire.

In the following section, the first three presented tests will be applied on the data. First, Kupiec test will be applied to verify whether the models provide proper coverage according to the chosen confidence level. Christoffersen test will then be applied to examine independence, and Lopez loss function will finally be used to benchmark the two models with respect to better performance.

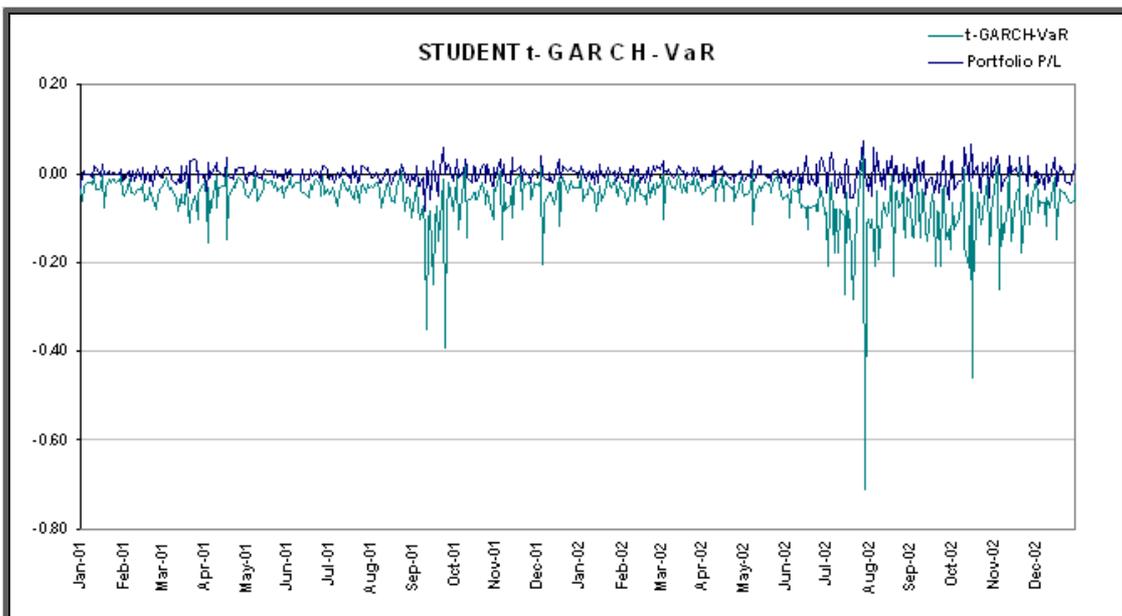
6.2.1. Results of Backtesting

Provided with the estimated parameters from the seven years sampling period for each volatility model, two years ahead of daily Value at Risk estimates were generated at 95% confidence level, covering the period from the beginning of 2001 till the end of 2002 (509 observations). Graphs (8 and 9) below plot daily actual profit and loss for the testing period against estimated Value at Risk with GARCH and *t*-GARCH models respectively (values presented in log returns).

Value at Risk estimates, generated by the two GARCH models, were nearly identical, and differences were totally trivial. Both models reported seven violations during the first year, and ten violations during the second year totaling seventeen violations over the testing period of 509 observations (2 years).



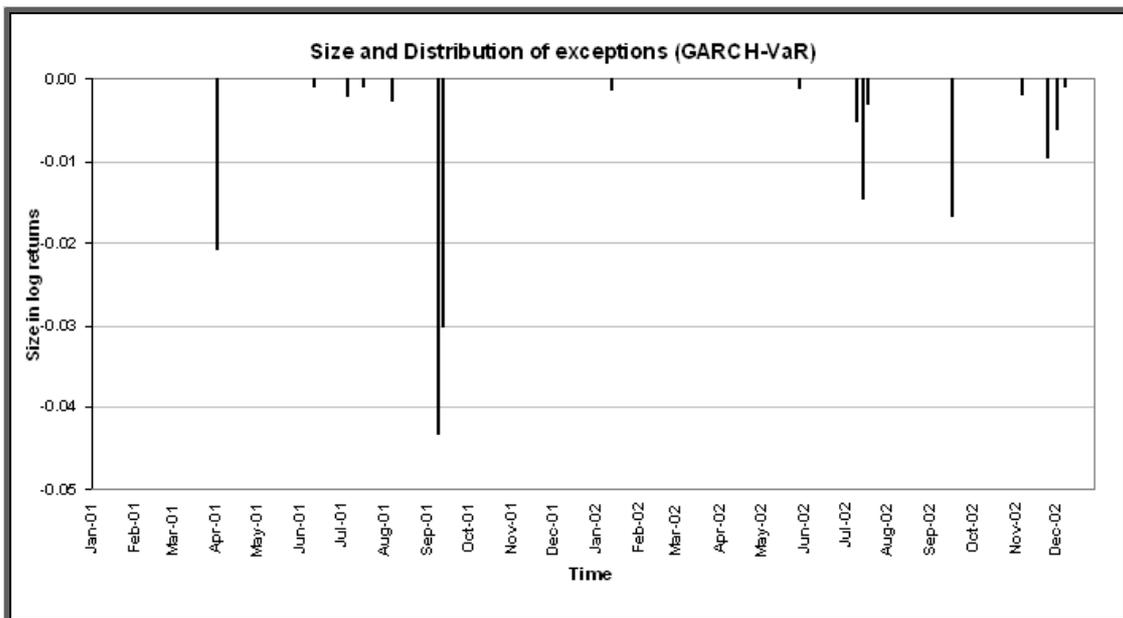
Graph 8: Actual daily P&L versus estimated daily GARCH Value at Risk



Graph 9: Actual daily P&L versus estimated daily t-GARCH Value at Risk

The results suggest clear indication toward the mutual acceptance for both models with respect to Kupiec’s test of unconditional coverage. Exceptions reported for each year separately (7 and 10 respectively) falls within the non-rejection region ($6 < x < 21$), and in aggregate for two years, 17 exceptions falls exactly above the minimum limit of 16 exceptions. This conclusion is confirmed by the value of the $LR_{uc} = 3.293$ for two years with (17) exceptions, which is less than 3.84.

As noted by Christoffersen (1998), a correct risk models must, not only report unconditional converge, but exceptions must also be independently dispersed, thus coverage must be conditionally independent. After applying Christoffersen’s likelihood for conditional coverage test on the data, the results obtained again suggest mutual acceptance for both models. The value of the likelihood ratio for independence is $LR_{ind} = 1.1750 < 3.84$, and the value for the conditional coverage test is $LR_{cc} = 4.47 < 5.99$, the critical value for the Chi-square with two degrees of freedom.

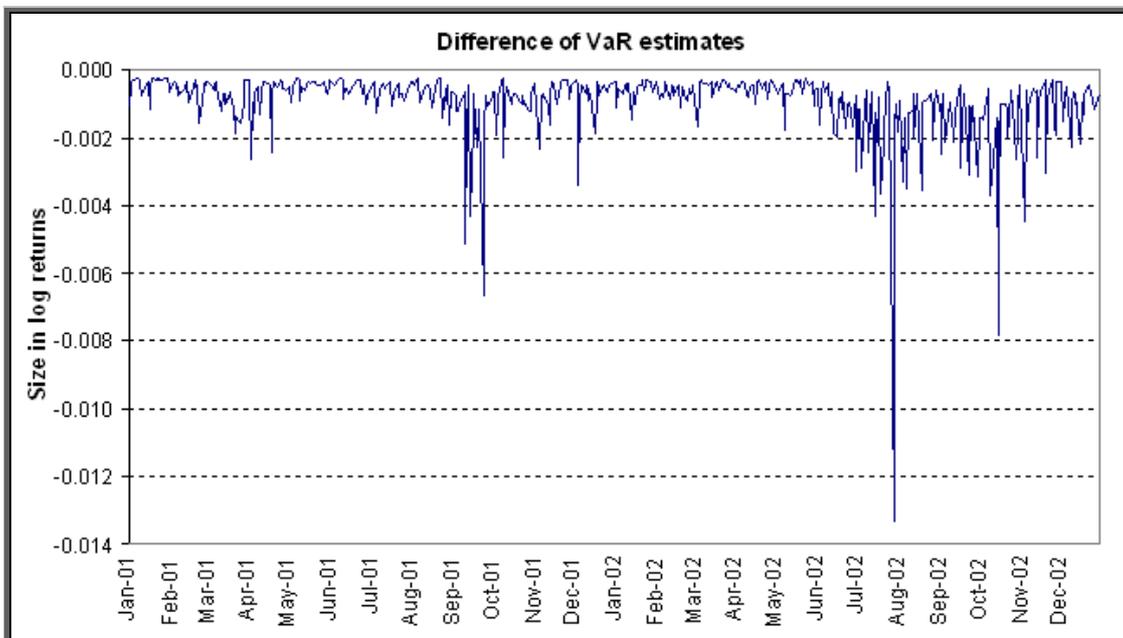


Graph 10: Size of exceptions beyond GARCH-VaR estimates

Graph (10) above illustrates the distribution of exceptions across the testing period, as well as their size below GARCH-VaR estimates. No consecutive exceptions occurred over the period, and the nearest exceptions were separated by two “no exception” observations. On the other hand, the graph shows that the magnitude of exceptions varied greatly across exceptions, and since there are small differences between the two VaR estimates, there is a motivation to examine whether one of the two models would be preferred according to Lopez’s Loss Function test.

The results obtained from the *Loss function test* suggests that normal GARCH is hardly preferred to student's *t*-GARCH since it minimized the cost of exceptions. The cost of normal GARCH exceptions is 17.0038 compared to 17.004 for student's *t*-GARCH. The results suggest that normal GARCH might has a slightly wider band than student's *t*-GARCH as well.

To examine this suggestion, a simple test was applied by finding the difference between the two VaR series (Normal GARCH *minus* student's *t*-GARCH). The results, as illustrated in graph (11) below, show that normal GARCH model has indeed a wider VaR band than student's *t*-GARCH, but with varying magnitude across the tested period.



Graph 11: Difference between GARCH and t-GARCH VaR estimates

As a conclusion, the results indicate that GARCH volatility models are generally powerful in modeling varying volatility, and their application to Value at Risk proved useful. Both models generated acceptable VaR estimates at the 95% confidence level, easily passing Kupiec's test for unconditional coverage. Furthermore, GARCH models responsiveness to market shocks prevented consecutive exceptions to occur, consequently passing Christoffersen's test for independence and conditional coverage.

Even though, the two tests equally validated the two models, they could not provide an answer to the preference question. This is due to the striking similarity of results with respect to number, and occurrence of exceptions from one side, and the tests' reliance on the number of exceptions regardless of their magnitude.

On the other hand, Lopez's *Loss Function* provided an answer for the preference question on behalf of normal GARCH model, since it produced less costly exceptions. However, as noted earlier, Lopez test appear to be more regulatory oriented, and does not pay attention to firms' opportunity cost of keeping excessive capital reserve during periods of "No Exceptions". Thus, given that both models are identical with respect to number and distribution of exceptions, the slim and weak preference of normal GARCH over student's *t*-GARCH, and the empirical evidence that normal GARCH generally over estimated volatility as presented in graph (11), for a profit maximizing financial institution, student's *t*-GARCH should be preferred over normal GARCH.

7. DISCUSSION

It is important when evaluating a parametric volatility model; that depends on a historical range of data, to have an in depth look into the events behind this data. This is simply because any parametric volatility model is not a magic stick, but depends on the nature of volatility experienced over the sampling period, and if it was too different than that of the testing period, the model would most likely fail to provide reliable volatility estimates.

In this study, the sampling data covered the period between 1994 till the end of 2000. A period that was rich of several regional markets shocks with global magnitude, starting with the Mexican crisis in 1994 –95 that caused varying degree of volatility in world's currency exchange and equity markets, known as the tequila effect. The crisis was triggered by a political unrest following the assassination of a presidential candidate, causing a sharp drop in investors confidence, and a large capital outflow, setting massive pressures on the Peso to be devaluated by the end of 1994.

As years 1995-96 passed relatively crisis free, years 1997-98 on the opposite witnessed a sequence of financial crisis across the globe. Starting with the shocking Asian crisis caused primarily by poor banking regulations and massive debt problem following the devaluation of several Asian currencies against the dollar in 1997. The Asian crisis was followed by the Russian public debt crisis and the near collapse of the giant LTCM in 1998, for which Value at Risk received a good proportion of blame.

The blame was for two reasons; the first one was the blind faith of LTCM management in its internal sophisticated mathematical risk management models in addition to its stunning 1 to 25-leverage ratio. The second reason was the automatic effect triggered by high VaR estimates causing banks to face two choices, either to increase capital reserve,

or get rid of risky positions. As banks preferred the latter at the same time, mounting pressures on prices to decrease caused LTCM to crash.

Even though no major financial crisis during 1999-2000, markets suffered relatively high volatility due to increasing oil prices, coinciding with the end of the IT bubble causing several world indices to decline.

These volatile market events have certainly contributed to the estimated GARCH models parameters, enabling the models to generate acceptable volatility and VaR estimates for a period that has seen extreme and unexpected events, particularly the September 11th attacks, followed by the US war in Afghanistan, and the general shift of US foreign policy spreading great uncertainty on world markets, causing high volatility experienced through all the remaining of the testing period.

Even though the Akaike criterion (AIC) gives relative preference to t -GARCH over plain GARCH, this preference was not *immediately* seen when the models were applied on Value at Risk at 95% confidence level. The two models generated identical results with respect to number and distribution of exceptions, leading both Kupiec test and, Christoffersen test, for unconditional and conditional coverage respectively, to equally accept the two models as correct.

One possible reason behind this obscurity is the high degree of freedom observed in residuals' distribution, implying that tails of the distribution, even though fatter, are still close to normal, and since a plain GARCH process has the property of excess kurtosis and heavier tail than normal, as shown by Bollerslev (1986), tail of the residuals' distribution was apparently within the reach of the plain GARCH model.

Obviously, the special characteristic of the tested portfolio has also contributed to this outcome. The observed degree of freedom for each index separately is approximately 10 and 16 for DAX and CAC respectively, which are exceptionally high compared to the general empirical convention, stating that returns and residuals distributions are significantly fat tailed with low degree of freedom. For instance, an empirical study by Goorbergh, and Vlaar (1999) on the Dutch stock market index (AEX) from 1983 to 1998 shows that the estimated degree of freedom of residuals equal to 5.7, and when

normal and t -GARCH models, along with other volatility models, were applied on Value at Risk, they passed Kupiec test at 95% confidence level, but surprisingly normal GARCH-VaR model recorded lesser exceptions than that of t -GARCH.

However, in this study it appears to be a special case and comparison between the two models was made difficult for traditional backtesting techniques. Thus, the implementation of Lopez Loss function was an important step toward answering this study's key question. The test has indeed detected differences between the two models, and preference was given to normal GARCH over t -GARCH since it minimized losses of exceptions. One must note however that the test preferred the safer model, focusing only on cost of exceptions, but ignoring the opportunity cost of accumulating unnecessary capital reserve during periods of "no exceptions". A comparison between the daily VaR estimates of the two models has shown that normal GARCH model has in general a wider band, which is an indication, or inference of risk over estimation.

Answering the preference question would therefore depend on one's own personal judgment and perspective. A conservative risk controller or supervisor might prefer the safer model, or normal GARCH in this case. On the other hand, given the similarity of results between the two models, a profit maximizing financial institution, would select t -GARCH since, 1) it provides acceptable and identical coverage as the safer model given the confidence level. 2) Its excess cost of exceptions is trivial compared with the safer model, and most importantly 3) it reduces the opportunity cost associated with capital reserve, and save extra capital for profitable investments.

However, suggested results cannot be generalized. This study's framework was strictly limited to compare the application of plain GARCH (1,1) and student's t -GARCH (1,1) on Value at Risk at the 95% confidence level, and there is a possibility of obtaining different results if other confidence levels were chosen.

There is no doubt that the exceptional characteristics of the tested portfolio returns and residuals distribution has partly contributed to this outcome, specially with respect to the high degree of freedom and closeness to normal.

Finally, The tests applied to backtest results (Kupiec and Christoffersen) focused only on the rate of exceptions, and accepted the models accordingly. Thus, the possibility that both models might be incorrect or misspecified cannot be ruled out if a more powerful test, such as that suggested by Berkowitz (2001) was applied instead.

8. CONCLUSION

GARCH models continue to prove superiority in modeling volatility clustering, and their application on Value at Risk provides more realistic and reliable estimates. Although several studies suggested that exchanging residuals' normality assumption with the fatter tailed student's t - distribution would improve the estimated volatility of financial returns, such improvement was not immediately present due to the exceptional characteristics of the tested portfolio.

Both models generated nearly identical VaR estimates, and were equally accepted by Kupiec and Christoffersen tests. Preference was only established when Lopez loss function was applied on the results, showing that normal GARCH slightly minimized losses of exceptions. However, a further investigation has shown that normal GARCH generated a wider band of VaR estimates which can be regarded as risk over estimation given the similarity between the two models with respect to number and distribution of exceptions.

A safer model might reduce the cost of exceptions when they occur, but would replace it with a sustained higher opportunity cost during periods of no exceptions. Taking this important aspect into account, the cost reduction of normal GARCH over t -GARCH was totally insignificant in this study. Thus, keeping in mind that both models produced identical results, t -GARCH should be preferred from the perspective of a profit maximization institution.

9. LIST OF REFERENCES

- Aczel, A. (2002). *Business Statistics*. 5th edition. McGraw-Hill
- Alexander, C. (1998). "Value at Risk". Risk Management and Analysis. Vol.1: Measuring and Modeling Financial Risk, pp. 125-171. John Wiley & Sons Ltd, UK.
- Bera, A., Hinggis, M. (1993). "ARCH Models: Properties, Estimation and Testing", Journal of Economic Surveys, Vol. 7, No. 4, pp. 307-366.
- Berkowitz, J. (2001). "Testing Density Forecasts With Applications to Risk Management", Journal of Business and Economic Statistics, 19, 465-474.
- Bollerslev, T. (1986), Generalized Autoregressive Conditional Heteroskedasticity, Journal of Econometrics, 69, pp. 542-547
- Bollerslev, T. (1987). A Conditional Heteroskedastic Time Series Model for Speculative Prices and rates of Returns, The Review of Economics and Statistics 31, pp. 307-327
- Bollerslev, T., Engle, R.F. and Nelson, D. (1994) "ARCH models" Handbook of Econometrics, Vol.4 (Eds. Robert F. Engle and Daniel McFadden). Amsterdam: Elsevier Science B.V., pp. 2959-3038
- Bollerslev, T., Chou, R. and Kroner, K. (1992), "ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence", Journal of Econometrics 52, pp. 5– 59.
- Campbell, S. (2005) "A Review of Backtesting and Backtesting Procedures" Finance and Economics Discussion Series, Divisions of Research & Statistics and Monetary Affairs, Federal Reserve Board, Washington, D.C.

Christoffersen P., (1998) “Evaluating Interval Forecasts” *International Economic Review*, 39, 841-862.

Christoffersen P., and Pelletier D., (2004), “Backtesting Value-at-Risk: A Duration-Based Approach”, *Journal of Empirical Finance*, 2, 84-108.

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

Engle, R., (1982), “Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation,” *Econometrica* 50, pp. 987– 1007.

Engle, R., (2001), “GARCH 101: The Use of ARCH/GARCH Models in Applied Econometrics” *Journal of Economic Perspectives* Vol. 15, No.4, pp. 157– 168.

Fama, E. (1965), The Behavior of Stock Market Prices, *Journal of Business* 38, 34-105.

Fan, J., (2003), *Nonlinear Time Series: Nonparametric and Parametric Methods*. Secaucus, NJ, USA: Springer-Verlag New York, Incorporated.

Franses, P. H. (2000) *Non-Linear Time Series Models in Empirical Finance*. Port Chester, NY, USA: Cambridge University Press.

Goorbergh, R., Vlaar, P. (1999). “Value-at-Risk Analysis of Stock Returns: Historical Simulation, Variance Techniques or Tail Index Estimation?” Netherlands Central Bank, Research Department.

Gourieroux, C. and, Jasiak, J. (2001) “Financial Econometrics”. Princeton University Press.

Gujarati, D.N. (1995) *Basic Econometrics*. Third edition. McGraw-Hill

Hamilton, J.D. (1994) *Time Series Analysis*, Princeton Univ. Press.

Hubbert, S. (2004). *Risk Management: Lecture three*. University of London, School of Economics, Mathematics and Statistics.

www.econ.bbk.ac.uk/faculty/hubbert/lecture3.pdf

Johnston, J., DiNardo, J. (1997). *Econometric Methods*, 4th edition, New York, McGraw-Hill

Jorion, P. (2000). Risk Management Lessons from Long-Term Capital Management. *The Journal of European Financial Management* 6 (September 2000): 277-300

Jorion, P. (2000) *The Value at Risk Field book: The Complete Guide to Implementing VaR*. McGraw-Hill.

Jorion, P. (2001). *Value at Risk: The Benchmark for Managing Market Risk*. McGraw-Hill

Lopez J.A. (1999a). "Regulatory Evaluation of Value-at-Risk Models", *Journal of Risk*, 1, 37-64.

Lopez J.A. (1999b). "Methods for Evaluating Value-at-Risk Models", *Federal Reserve Bank of San Francisco Economic Review*, 2, 3-17.

Mandelbrot, B. (1963), "The variation of certain speculative prices." *Journal of Business*, 36, pp. 394-419.

Markowitz, H. (1952). "Portfolio Selection". *The Journal of Finance*, Vol. VII, No.1, March 1952.

Mina, J. and, Xiao, J. (2001) "Return to RiskMetrics: The Evolution of a Standard". <http://www.riskmetrics.com/pdf/rrmfinal.pdf>

Sharpe, W. "Capital Asset Prices - A Theory of Market Equilibrium Under Conditions of Risk," *The Journal of Finance*, Vol. XIX, No. 3, September 1964, pp. 425-442.

Wilson, T. (1993). *Infinite Wisdom*, Risk Magazine, June.

Wilson, T. (1998). "Value at Risk. Risk Management and Analysis". Vol.1: *Measuring and Modeling Financial Risk*, pp. 61-124. John Wiley & Sons Ltd, UK.

Zangari, P. (1996). "RiskMetrics-Technical Document". 4th edition, pp. 77-25. <http://www.riskmetrics.com/pdf/td4e.pdf>