

ESTIMATION PORTFOLIO VAR WITH THREE DIFFERENT METHODS: FINANCIAL INSTITUTION RISK MANAGEMENT APPROACH

By

Apostolos Kiohos, Aris Dimopoulos***

*Doctoral Researcher Candidate in FI Risk Management University of Thessalia,
MSc Risk Management and Insurance City University Business School

**MSc Finance and Management University of Exeter

Abstract

Value at Risk (VaR) technique is very important for the measurement and control of market, credit and also operational risks. VaR has become an essential tool for financial institution risk management. VaR forecasts of financial instruments are also of great importance. The fact that volatility seems to cluster over time in a predictable way has very important implications for investors, risk managers and financial institutions. In order to generate daily VaR forecasts of equity portfolios for S&P 500, FTSE ALL SHARE and NIKKEI 500 and of over ten-years Government bond portfolios for US and the UK we use the exponentially weighted moving average (EWMA). EWMA model emphasizes in most recent observations by assigning heavier weights to them than to those from the distant past. In the latter EWMA estimates are used as inputs in three VaR estimation methods in order to produce forecasts based on each one of them. The three methods are: the Variance Covariance approach used by JP Morgan Risk Metrics, the Historical Simulation approach and the Hybrid method. Finally we use six backtesting techniques for the validation of the estimation and the evaluation of the forecasting performance. The results indicate that Historical Simulation out performs the other two competing approaches with close second the Risk Metrics approach. (JEL Classification: G10 G2 C52 C53).

Key Words: Risk Management, Value at Risk, Financial institution, Forecasting, EWMA, Backtesting.

1. Introduction

During the last few years, risk management has become one of the main concerns in financial industry. Financial institutions invest into developing reliable risk measurement and management techniques in order to measure and

control market, credit and operational risks. The main advanced technique includes the use of Value at Risk (VaR) models. These models calculate the worst expected loss of a portfolio of financial instruments at a pre-specified time and level of confidence. VaR exhibits the attractive property of summarizing market risks in one single number. This simplification is significant for risk managers because it makes this technique very informative and easily understood. There are four basic approaches of computing VaR: Variance Covariance approach, Historical Simulation approach, Monte Carlo Simulation and Extreme Value approach. However there are some drawbacks for these methods. One negative feature is that they use past data to provide estimates of the maximum future potential loss. An additional weakness is that the VaR models are based on distributional assumptions. Furthermore, risk managers have emphasized in the idea of complementing VaR estimates with stress testing. The definition that Kevin Dowd¹ gives for stress testing is: "Stress testing is a variety of different procedures that attempt to gauge the vulnerability of our portfolio to hypothetical events". These tests provide risk managers with information about the expected losses caused by specific events or scenarios.

Another aspect essential for risk management is the volatility forecasts of the portfolio returns. A firm needs a time dynamic forecast that will take into account the dynamic properties of variance such as volatility clustering. Good forecasts also provide better control of market financial risks and lead to good decisions. The forecasts' literature includes various methods. In this paper the Exponentially Weighted Moving Average (EWMA) model is employed in order to generate one-day VaR forecasts for five index portfolios. These forecasts are a weighted average of the previous time period's variance. The basic characteristic of EWMA is that it assigns heavier weight to the most recent observations than those from the distant past. EWMA forecasts are used as input in three VaR estimation methods in order to generate forecasts based on each of them. The three methods are the Historical Simulation approach, the Variance Covariance approach and the Hybrid approach, which was introduced by Christoffersen (1998). The first two approaches are standard estimation techniques and the third is an extension to Historical Simulation method (we use this method as an alternative to Hull-White approach). The aim of this study is to investigate which method performs best in forecasting VaR.

However, financial institutions in order to evaluate the performance and verify the accuracy of their estimates of the maximum losses use back testing techniques. Back testing involves a variety of methods since a definite test for

all the VaR estimates does not exist. Since there is no complete theory for the evaluation of the VaR forecasts we use various methods for the verification of their accuracy.

This paper consists of three parts. The first part introduces a general view of modern risk management and VaR. The second part discusses the basic methods for VaR estimation and gives a general view of our method of forecasting volatilities. The third part presents the data, the methodology used for the generation of VaR forecasts and the back testing techniques used to verify the accuracy of these forecasts. Finally the last section includes the empirical findings and the conclusions.

2. Value at Risk (VaR)

The measurement of the market risk of financial instruments' portfolio is very important for investors as well as financial institutions. The financial instability and the large losses that corporations faced over the past decade made the identification of risks and their measurement an urgent issue. VaR is a product of the effort of several financial institutions to measure and control risks of their activities during the late 1980's. The most famous of the VaR methodologies that were developed is the Risk Metrics approach, which was introduced by JP Morgan. In recent years, it has become not only a risk measurement method but also a process that helps decision-making and investment guiding. It is a methodology that has gained rapid acceptance as a valuable approach to risk management.

Value at Risk is defined as the maximum monetary loss of a portfolio due to market fluctuation at a given level of confidence over a given risk horizon. VaR uses statistical methods to derive a single number that summarizes market risks. It enables a firm to determine which investments offer the highest expected returns at the least expense of risk.

Nowadays, VaR methods are widely used not only by financial but also by non-financial institutions and fund managers. The popularity of these models is due to their appealing features. The first and most obvious advantage of VaR models is their ability to express the market risk of an entire portfolio in one number. This number represents the potential loss of the portfolio in monetary terms. Therefore, there is no need for special technical knowledge in order to understand it. One more interesting characteristic is that VaR allows for comparisons between different investment positions and risky activities. The fact that it is a single number makes easy the comparison of positions across

different countries, instruments (one can compare the risk for an equity with the risk of a bond), markets (currency, foreign exchange or interest rates) and type of risks. These comparisons enable firms and investors to adjust their investment strategy for risk. They can set position limits in order to control their risk exposure. Another advantage is that VaR takes into account the interaction between different risk factors. This means that if two risk factors offset (do not offset) each other then VaR methodology gives a low (high) risk estimate.

Although VaR methods are very popular, some negative aspects characterize them. The first disadvantage derives from the fact that these methods use past historical data to provide an estimate for the future. What happened in the past does not mean that will happen again in the future and hence it is easy for the estimation to be incorrect. Second, the VaR number can be calculated by using several methods. These methods try to capture volatility's behavior. However, there is an argument on which is the method that performs best. A third drawback for VaR is that the methods for computing it are based on different assumptions. These assumptions help us with the calculation of VaR but they are not always true (like distributional assumptions). Furthermore, there are many risk variables such as political risk, liquidity risk, personnel risk, regulatory risk, phantom liquidity risk and others that cannot be captured by the VaR methods.

VaR has several uses. First, it can be used by senior management to set risk targets and position limits. According to the desired level of risk a firm can set an overall VaR, which in turn represents the overall target risk. VaR is very helpful for the capital allocation. Since this number shows what is the maximum expected loss, the firm can determine its capital requirements at each different level or even at each individual investment. Thus, the riskier the activity the greater the VaR and consequently the greater the capital requirement and vice versa. One more attractive use is that VaR can be used to assess the risks of different investment opportunities before decisions are made, as well as to evaluate the performance of business units after the event. These uses make obvious that VaR methodology improves the process of risk management. By using such methods a risk manager can measure not only market risks but also other risks like credit, liquidity and cash flow risks. Finally, it enables firms to react quickly and appropriately to the capital adequacy regulation.

As the risk management systems evolve, regulators reexamine the capital standards on financial institutions. Banks and insurance companies are re-

quired to keep enough capital in order to cover their unexpected losses. Evidence of the great importance of VaR methods is their recognition by the regulators. The first important proposition came on 1988 from Basle Capital Accord, which was requiring tighter risk management systems. In 1994 Bank for International Settlements Fisher report advised financial intermediaries to disclose VaR measures publicly. The Derivatives Policy Group has also advised the use of VaR models. The most important proposal for the recognition of this methodology came from the Basle Accord on Banking Supervision in 1996, which allowed banks to adopt their own risk measurement models in order to calculate the minimum regulatory required capital to cover their market risks.

The estimation of VaR can be conducted by a variety of methods. This fact makes the verification of the veracity of the statistical assumptions underlying VaR calculations and the accuracy of VaR estimates necessary. This means that after VaR estimation one needs to follow a procedure called back testing in order to verify the accuracy of VaR models.

3. VaR and Forecasting methodology

There are various methods, which produce VaR measures. We can distinguish four different basic methods of calculating VaR: Variance Covariance approach, Historical Simulation approach, Monte Carlo Simulation and Extreme Value Theory. This chapter attempts to describe theoretically the methods, which we use in our research, and presents a brief summary of their advantages and disadvantages. Additionally, we present Exponentially Weighted Moving Average (EWMA), our methodology of forecasting variance.

According to the variance covariance approach all market risks are normal and the portfolio is a linear function of these normal risks. If normality holds the VaR is the multiple of portfolio's standard deviation and the portfolio standard deviation is a linear function of individual volatilities and covariances. In order to calculate VaR one needs to find the portfolio's standard deviation (using the variance covariance matrix and the weights of individual assets) and then multiply it by the value of the portfolio and the desired confidence level. The formula for VaR is:

$$VaR = -\alpha * P * \sigma_p \quad (1)$$

Where σ_p is the portfolio's standard deviation, P is the value of the portfolio and $\kappa(\alpha)$ is the desirable level of confidence (the $(1-\alpha)\%$ quantile of the standard normal distribution).

Before someone can use this approach to calculate VaR should consider some other issues, which concern the nature of returns. As in the case of bonds (and generally fixed income instruments) and derivatives, returns are not linear functions of risk factors. One solution to this problem is the delta-normal approach, which takes a first order approximation to the returns and then uses this approximation to calculate VaR. Delta-normal approach works by replacing the true positions with linear approximations and handling them in the same way as other linear positions. Hence, we are assuming that the non-linearity in our position is sufficiently limited so that we can ignore it and still produce accurate VaR estimates. This approach has some attractive aspects. The first is that maintains the linearity of the portfolio without adding any new risk factors. It also requires few additional data. By Delta-normal approach one can handle more complex positions without losing the benefits and convenience of the normal distribution.

However, one may need more precise VaR estimates without losing the convenience of working with the variance covariance method. This is the case of the second order or delta-gamma approach. Second order approximation assumes that returns are normally distributed but allows for a non-linear relationship between the portfolio's value and the underlying returns. According to this method VaR can be calculated in two steps. The first step is the calculation of the first four moments (mean, standard deviation, skewness and kurtosis) of the portfolio's return distribution. The second step is to find a distribution that has the same four moments as the ones from portfolio's return distribution and then to find the $a\%$ quantile. After these steps the calculation of VaR is brought back to equation (1).

Generally, for large portfolios where optionality is not the most important issue the delta normal approach is proper because it provides a fast and efficient calculation of VaR. On the other hand, for portfolios that are exposed to a small number of risks and with significant option components the delta gamma approach offers increased precision at a low computational cost.

Variance covariance approach has many attractive features. First, this method, because of normality, makes the calculation of VaR very simple. Although the calculations are easy to implement, the figures of VaR produced by

this approach are very informative both on confidence level and holding period. Besides these, variance covariance method is very informative about the expected tail losses. However, there are some drawbacks in this approach. Its main problem is the excess kurtosis, which means that the returns distribution may have fat tails. This derives from the assumption that normality holds and means that VaR will underestimate the expected losses. Moreover, the normal distribution assumes the symmetry of the returns distribution while in reality financial returns often exhibit asymmetric behavior. This negative aspect represents the problem of negative skewness. This is because the right tail of the distribution contains more data than the left one.

Historical Simulation Approach tries to find an empirical distribution of the rates of return assuming that past history carries out into the future. This method uses the historical distribution of returns of a portfolio to simulate the portfolio's VaR. Often historical simulation is called non-parametric approach, because parameters like variances and covariances do not have to be estimated, as they are implicit in the data. It is simple to implement if daily historical daily data have been collected. The choice of sample period influences the accuracy of VaR estimates. Longer periods provide better VaR estimates than short.

The first step on implementing this method involves identifying the instruments in a portfolio and collecting a sample of their historical returns. The second step is to calculate the simulated price of every instrument using the weights of the current portfolio (in order to simulate the returns in the next period). The third step assumes that the historical distribution that the returns follow is a good proxy for the returns in the next period. Historical Simulation uses the actual percentiles of the observation period as VaR measures. For instance for an observation period of 1000 days, the 95th percentile historical simulation VaR measure is the 51st largest logs observed in the sample of 1000 outcomes. That is because the five percent of the sample that should exceed the risk measure is equal to fifty losses.

One of the greatest advantages of this method is that it does not depend on assumptions about the distribution of returns. Therefore, the mistakes of assuming parametric distributions with thin tails where in reality the distributions of returns have fat tails are avoided. One more positive characteristic is that the data set reflects gamma, vega risks as well as the correlations and volatilities. Thus, there is no need for any parameter estimation. In relation to the previous characteristic Historical Simulation gives information about other

useful statistics such as skewness and kurtosis. This approach does not make any distinction between the type of position and market risk (that is there are not different models for equities, bonds and derivatives like in variance covariance approach).

Although Historical Simulation seems to have many attractive characteristics, there are some disadvantages. First, Historical Simulation results are dependent on the data set from the past, which may be too volatile or not, to predict the future. Hence, one cannot provide accurate estimates of VaR because what happened in past will not necessarily happen in the future. The same occurs when the period used for the estimation of VaR includes serious incidents (for example economic shocks), which are unlikely to happen in the future. This can also be reversed by assuming that important events can happen in future but because the data set (the observation period) does not include them the HS underestimates VaR. Second, Historical Simulation assumes that returns are independent and identically distributed. The data display time varying property of volatility. For example, by choosing a long time series there is a problem in VaR estimation because too much emphasis is placed on data from the distant past. With short time series the estimations will probably not be reliable (because of the small length of the observation period). Another drawback for this method is that it uses the same weights on all past observations. If an observation from the distant past is excluded the VaR estimates may change significantly.

Forecasts are of great importance and widely used in economics and finance. It is reasonable that good forecasts lead to good decisions. The fact that volatility seems to cluster over time in a predictable way has very important implications for investors, risk managers and financial institutions. VaR increases as the volatility increases and vice-versa. There are many approaches for forecasting VaR. This paper uses the exponentially weighted moving average to produce daily VaR forecasts.

The EWMA approach for forecasting volatilities emphasizes on recent observations by using exponentially weighted moving averages of squared deviations. This method attaches different weights to the past observations contained in the observation period. Since the weights decline exponentially, the most recent receive much more weight than the earlier ones. EWMA model calculates the volatility forecast for the next period as a weighted average of the previous period's volatility forecast and the current squared return. The formula for the EWMA model is:

$$\sigma_t^2 = \lambda \sigma_{t-1}^2 + (1 - \lambda) r_{t-1}^2$$

where σ_{t-1} represents the volatility of the returns at time $t-1$, r_{t-1} represents the return at time $t-1$ and λ is the decay factor. This method emphasizes that the volatility on a given day $t-1$ is actually used as a predictor for the volatility of the next day t . The EWMA model depends on the decay factor. The parameter λ ($0 < \lambda < 1$) determines the relative weights that are applied to the returns and the effective amount of data that are used in volatility estimation. In this paper the decay factor is set to be 0.94 just like Risk Metrics for daily data².

This method is equivalent to the Integrated GARCH or IGARCH (1,1) family of popular conditional models. It can be viewed as a special case of GARCH process, which is given by the equation:

$$\sigma_t^2 = a_0 + a_1 \sigma_{t-1}^2 + \alpha_2 r_{t-1}^2$$

where α_0 , α_1 and α_2 are parameters. By setting α_0 to 0 and α_1 and α_2 sum to unity the GARCH model becomes an exponentially weighted estimator.

EWMA approach explains volatility clustering, which means that abnormal volatility in one period is likely to lead to abnormal volatility in the next period. It also allows the volatility to vary from one period to another. A very important feature of the exponentially weighted estimator is that it can be written in recursive form in order to be used for making volatility forecasts. To derive this form one should assume that an infinite amount of data is available. This recursive form can be written as follows:

$$\sigma_t^2 = (1 - \lambda) \sum_{i=0}^{\infty} \lambda^{i-1} r_{t-i}^2$$

4. Data, Value at Risk Estimation Methods and Backtesting

4.1 Data

In this section we present the data and the VaR estimation and validation procedure. The standard EWMA estimator is used to generate daily VaR forecasts based on either Variance Covariance approach (or JP Morgan Risk Metrics approach) or Historical Simulation approach. The third method, which is Hybrid approach, produces daily forecasts by itself without using any volatility forecast models. The forecasts of the alternative methods are evaluated using seven techniques: The proportion of failures, the unconditional coverage, the

conditional coverage or independence, the correct coverage, the Lopez' magnitude of exceptions, the mean relative bias and the root mean squared relative bias.

The data used are bond and equity data for the period between 25 February 1987 and 1 February 2002 provided by Primark Datastream. The sample consists of daily prices for the following five indices: S&P500, FTSE ALL SHARE, NIKKEI 500, US Government Bonds over 10 years, UK Government Bonds over 10 years. For all five indices we have a total of 3900 observations. We need to have a powerful set of back testing techniques, hence, the daily observations ensure that. Furthermore, the equity indices represent three of the most significant markets worldwide. The bond indices are an average of all government bonds with time to maturity over 10 years⁴. The selection of these two indices is due to the fact that there are limited resources that can provide data with such long period and high density. For these time series we calculated geometric returns. The formula for the geometric returns is: $r_t = \ln P_t / \ln P_{t-1}$ where P_t is the price for day t . After that, we calculate EWMA. For this calculation we need an initial value for the volatility, which is set at zero. As mentioned before, EWMA estimates are used as inputs in the following models in order to generate forecasts based on each one of them.

4.2 Value at Risk Estimation Methods

Variance Covariance Approach

For the Variance Covariance approach, in this paper we employ one of the most widely used volatility estimation and forecasting methodology, the JP Morgan Risk Metrics⁵. This approach applies exponentially declining weights to the returns from distant past (and greater weights to more recent returns) in order to estimate conditional volatilities and correlations. Exponential smoothing allows for cyclical behavior of return volatility to be captured. Exponentially weighted moving average model can be considered as an improvement over the traditional volatility forecasting method, which is based on moving averages with fixed, equal weights.

Risk Metrics methodology assumes that returns on securities follow a conditionally (on the information set at time t) normal distribution and that the change in position's value is a linear function of the underlying return. Although this model is based on this unrealistic assumption and ignores the presence of fat tails in the probability distribution of returns, it is commonly found that it performs satisfactory. Nelson⁶ (1992) showed that even misspecified

models can estimate volatility rather accurately. He showed that if the return generating process is well approximated by a diffusion, a broad class of even misspecified ARCH models can provide consistent estimates of the conditional volatility. Another important aspect is that Risk Metrics technique not only reacts fast to shocks in the market as the most recent data carry a heavier weight but also after an abnormal return, the volatility declines exponentially as the weight of the shock observation falls. VaR is computed as the α quantile times the portfolio standard deviation. In order to provide daily VaR forecasts the α quantile is multiplied to the EWMA standard deviation, which is given by the following formula:

$$\sigma_t = \sqrt{\lambda \sigma_{t-1}^2 + (1 - \lambda) r_{t-1}^2}$$

where σ_t is the one-day volatility forecast, λ is the decay factor, r_t is the portfolio's returns and σ_{t-1} is the last day's volatility. Therefore the formula for the daily VaR forecasts is:

$$VaR_{t-1} = -\kappa(\alpha) * \sigma_{t-1} * V$$

where $\kappa(\alpha)$ is the α quantile of the standardized distribution and α is one minus the desired confidence level (1-a)% and V is the value of the portfolio.

Historical Simulation

In order to compute daily VaR forecasts we use an extension of the traditional Historical Simulation, which incorporates volatility clustering. This extension, which allows for volatility updating, was proposed by Hull and White in 1998⁷. The main feature of volatility clustering is that large returns tend to be followed by more large returns, small returns tend to be followed by more small returns. One consequence of volatility clustering is that it induces leptokurtosis in the unconditional distribution of portfolio returns. Therefore, it improves VaR estimates, since if we know that we are currently in a high volatility period, our estimates of VaR should be correspondingly higher. In order to incorporate volatility clustering into the calculation of VaR we need a model of time-varying volatility.

In their approach, Hull and White, use GARCH or exponentially moving average model in concurrence with the historical simulation in order to calculate VaR. They suggest that historical simulation can produce better estimators by taking into consideration the volatility changes experienced during the period covered by historical data. They consider a portfolio that depends on a

number of market variables. They did not use the actual historical percentage changes for calculating VaR but the historical changes, which have been adjusted and reflect the ratio of current daily volatility to the daily volatility at the time of the observation.

We compute the empirical quantile of the historical standardized distribution, that is the distribution of returns scaled by each days' estimated standard deviation from the model, and then we scale this up by the model's forecast of tomorrow's standard deviation. For the estimation of this model a rolling window of 1000 observations is used. Initially EWMA is estimated for all the observations but due to the fact that a rolling window is used, the first 1000 observations are discarded. This means that we calculate 2900 daily VaR forecasts for Historical Simulation method. Thus it is necessary to ignore the first 1000 observations for the other two VaR estimation methods in order to be able to compare them equally with Historical Simulation. Just like in Variance Covariance approach VaR is computed for 95% and 99% levels of confidence.

The daily VaR forecasts are provided by the following formula:

$$VaR_{t+1} = -\kappa(\alpha) * \sigma_{t+1} * V$$

where $\kappa(\alpha)$ is the α percentile of the standardized distribution and α is one minus the desired confidence level $(1-\alpha)\%$ and V is the value of the portfolio. The term σ_{t+1} represents the standard deviation of the returns at time t . The standard deviation estimate that is used to standardize returns is found from the EWMA model. These standardized returns give us the empirical distribution of portfolio's returns, which we use to produce VaR forecasts.

Hybrid Approach

The third method that is used to provide forecasts is the Hybrid approach. This method was introduced by Boudoukh, Richardson and Whitelaw (1998)⁸. In their paper they use elements from the methodologies of two very popular approaches such as the Risk Metrics and the Historical Simulation. It estimates the percentiles of the returns directly, using weights that decline on data from the distant past. This approach can be implemented in three steps:

1) Let $R(t)$ be the rate of return observed between the period $t-1$ and t . This method assigns a weight $[(1-\lambda)/(1-\lambda^K)]$, $[(1-\lambda)/(1-\lambda^K)] \lambda, \dots, [(1-\lambda)/(1-\lambda^K)] \lambda^{K-1}$ to each of the most recent K returns: $R(t), R(t-1), \dots, R(t-k+1)$, respectively. The constant term $[(1-\lambda)/(1-\lambda^K)]$ ensures that the weights sum to unity.

2) Put the returns in ascending order.

3) In order to get the $a\%$ VaR of the portfolio, start from the lowest return and keep accumulating the corresponding weights until the $a\%$ is reached. To get exactly the $a\%$ of the distribution linear interpolation is used between adjacent points.

We calculate VaR for 95% and 99% level of confidence for all five indices.

4.3 Backtesting

There are many measures have been proposed for the validation of a VaR model performance in forecasting. Hence, in order to evaluate the performance of competing models we present a variety of different measures, which provide an indication of model performance. These indicators attempt to capture the accuracy of the different models by evaluating the extent to which the proportion of losses that exceed the VaR estimates are consistent with the models' stated confidence level. Another valuation technique that we employ examines the variability of VaR estimates produced by the different VaR estimation methodologies.

Proportion of Failures

The failure or exception can be defined as the number of times for which the real portfolio returns are smaller than the estimated Value at Risk. A score of 1 is imposed when an exception occurs and 0 when it does not. Proportion of Failures test compares the total number of failures to the total accumulated sample size. For the sample of T observations the Proportion of Failures can be calculated from the following formula:

$$F\% = \frac{\sum_{t=1}^T N_t}{T}$$

where $t=1,2,\dots,T$ is the number of observations and N is the failure at time t .

Unconditional Coverage

Exceptions or Failures can be modeled as independent draws from a binomial distribution with probability of occurrence equal to N/T percent (exceptions/sample size). For a VaR forecast to be accurate its unconditional coverage ρ must equal the proportion of failures (N/T). In 1995 Kupiec⁹ came up

with a Likelihood Ratio (LR) to test the null hypothesis that the probability of exception is ρ against the alternative that the probability differs from ρ assuming that the exception process is independently distributed. Kupiec's LR is given in the following formula:

$$LR_{UC} = -2 \ln[(1-p)^{T-N} p^N] + 2 \ln[(1-N/T)^{T-N} (N/T)^N] \sim \chi_1^2$$

Conditional Coverage or Independence

The conditional coverage or independence test of VaR estimates counts the number of exceptions over the observed period without considering the information available at each point of time. If the returns show heteroskedasticity that depends on time then the conditional coverage becomes more important. In this case VaR models that do not take into account variance dynamics, produce VaR estimates that may have incorrect conditional coverage but correct unconditional coverage. In 1998 Christoffersen¹⁰ proposed a LR based on interval forecasts. This ratio allows the separation of the effects of the volatility's dynamic properties from the effects of the assumptions for the probability distribution of the returns. Christoffersen's likelihood ratio for the independence of the exceptions is given by the following formula:

$$LR_t = 2[lnL_A - lnL_0] \sim \chi_1^2$$

where

$$L_A = (1 - \pi_{01})^{T_{00}} \pi_{01}^{T_{01}} (1 - \pi_{11})^{T_{10}} \pi_{11}^{T_{11}}$$

$$L_0 = (1 - \pi)^{T_{00} + T_{10}} \pi^{T_{01} + T_{11}}$$

$$\pi_{ij} = \frac{T_{ij}}{T_{i0} + T_{i1}}$$

$$\pi = \frac{T_{00} + T_{11}}{T_{00} + T_{01} + T_{10} + T_{11}}$$

T_{ij} notation denotes the number of observations in state j after having been in state i the period before.

Correct Conditional Coverage

In the event that the conditional distribution of returns and the variance dynamic properties (for example time varying volatility) are captured by a VaR

model then the exceptions should be unpredictable. The exception series should show both correct unconditional coverage and serial independence. The likelihood ratio for correct coverage is a test for both of these properties and it is given by the following formula:

$$LR_{CC} = LR_{UC} + LR_i$$

The LR_{CC} is the sum of Kupiec's unconditional coverage and Christoffersen's independence tests.

Lopez Magnitude of exceptions

Lopez (1999)¹¹ introduced a new measure for the accuracy of VaR estimates based on the magnitude of exceptions. As an alternative to the hypothesis testing framework, he proposed an evaluation method that uses standard forecast evaluation techniques. Lopez indicates as an accuracy indicator of VaR estimates how well they minimize certain loss functions. He is introducing two loss functions that represent specific regulatory concerns. The first loss function is implied by the binomial method and is equal with:

$$LF_1 = \sum_{t=1}^T C_{m,t}$$

where $C_{m,t} = 1$ if $-VAR_t > X_t$, and $C_{m,t} = 0$ if $-VAR_t < X_t$. The appropriate benchmark price for the above function is simply the number of observations multiplied with $1-a$, where a is the desired level of confidence. The second loss function addresses the magnitude of the exceptions. That loss function is described by the following equations:

$$LF_2 = \sum_{t=1}^T C_{m,t}^*$$

where $C_{m,t}^* = 1 + (r_t - VAR_t)^2$ if $-VAR_t > X$ and $C_{m,t}^* = 0$ if $-VAR_t < X_t$.

A magnitude term is incorporated into the binomial loss function. The exception now takes a score $1+(r_{t+1}+VaR_{mt})$, where r_{t+1} is the return and VaR_{mt} is the Value at Risk forecast. The term $(r_{t+1}+VaR_{mt})$ is based on the magnitude of failures. Finally, Lopez' statistic calculates the sum of the number of exceptions and their squared distance from the corresponding VaR.

Mean Relative Bias (MRB)

The mean relative bias statistic was introduced by Daryll Hendricks in 1996. This statistic tries to capture the extent to which different VaR estimation methods generate estimates of similar average size. The mean relative bias for a portfolio is independent of the scale of the simulated portfolio because each of the daily relative bias calculations on which it is based is also scale dependent. The formula is:

$$MRB_i = \frac{1}{T} \sum_{t=1}^T \frac{VaR_{it} - \overline{VaR}_t}{\overline{VaR}_t}$$

$$\text{where } \overline{VaR} = \frac{1}{N} \sum_{i=1}^N VaR_{it}$$

Root Mean Squared Relative Bias (RMSRB)

This criterion was also introduced by Hendricks¹² in 1996 and is an extension of the Mean Relative Bias statistic. This statistic tries to capture the variability of the model's estimates. It examines the degree to which the risk measures tend to vary around the average risk measure. It is computed by taking the square root of the mean of the squares of the daily relative biases. The formula for the Root Mean Squared Relative Bias statistic is:

$$RMSRB_i = \sqrt{\frac{1}{T} \sum_{t=1}^T \left(\frac{VaR_{it} - \overline{VaR}_t}{\overline{VaR}_t} \right)^2}$$

$$\text{where } \overline{VaR} = \frac{1}{N} \sum_{i=1}^N VaR_{it}$$

5. Empirical Results

Tables 1 and 2 below report the proportion of failures of the VaR estimation methods for each of the five indices for the 95% and 99% levels of confidence, respectively. Examination of these two tables makes clear that in both levels of confidence Historical Simulation approach exhibits the lowest proportion of failures in the sample for all five indices and the Variance Covariance follows it closely. The percentage of exceptions in Hybrid approach is constantly higher than the other two. For the 95% level, Variance Covariance

figures are similar to those of Historical Simulation while for 99% level they are similar to the ones of hybrid approach. However the rate of exceptions in all three methods does not vary systematically.

TABLE 1.

Proportion of failures for 95% level of confidence

Proportion of failures	VCV	HS	Hybrid
S&P 500	0.051087	0.050742	0.061443
FTSE ALL SHARE	0.054884	0.052123	0.063169
NIKKEI 500	0.059027	0.049707	0.066966
UK GOV. BONDS	0.055230	0.050742	0.063169
USA GOV. BONDS	0.054194	0.047981	0.061443

TABLE 2.

Proportion of failures for 99% level of confidence

Proportion of failures	VCV	HS	Hybrid
S&P 500	0.021056	0.010356	0.023818
FTSE ALL SHARE	0.018295	0.009320	0.030376
NIKKEI 500	0.017259	0.008630	0.030031
UK GOV. BONDS	0.021056	0.010701	0.027960
USA GOV. BONDS	0.022782	0.011046	0.025544

Tables 3 and 4 below present the Unconditional Coverage (LR_{UC}) statistic for the 95% and 99% levels of confidence, respectively. Kupiec's likelihood ratio tests the null hypothesis that the proportion of exceptions N/T is equal to the desired significance level. Under this null hypothesis the LR_{uc} statistic has a Chi-square distribution with one degree of freedom. From the statistical tables of this distribution we obtain its critical values for the 1% and 5% significance levels, which are $\chi_{0.01}=6.6349$ and $\chi_{0.05}=3.84146$ respectively. If the likelihood ratio exceeds the critical value than the null hypothesis can be rejected for the desired significance level.

For the 5% level the null hypothesis can be rejected for hybrid approach for all the indices except for NIKKEI 500. For the same level the null hypothesis cannot be rejected for Historical Simulation (for all the indices). The hypothesis holds also for the Variance Covariance method except in NIKKEI 500

where is rejected. Generally for this level Historical Simulation seems to provide the best unconditional coverage. For the 1% level the null hypothesis can be strongly rejected for both Variance Covariance and Hybrid approaches. Again Historical Simulation generates the best unconditional coverage, as the hypothesis is not rejected for any of the indices.

TABLE 3.

Unconditional Coverage for 95% level of confidence

Unconditional Coverage	VCV	HS	Hybrid
S&P 500	0.071617	0.033436	7.467107
FTSE ALL SHARE	1.412218	0.271249	9.797491
NIKKEI 500	4.709232	0.005260	1.559491
UK GOV. BONDS	1.615580	0.033436	9.797491
USA GOV. BONDS	1.045502	0.251934	7.467107

TABLE 4.

Unconditional Coverage for 99% level of confidence

Unconditional Coverage	VCV	HS	Hybrid
S&P 500	27.141841	0.036564	40.264035
FTSE ALL SHARE	16.169199	0.138460	78.712759
NIKKEI 500	12.670801	0.576243	76.460776
UK GOV. BONDS	27.141841	0.140474	63.456167
USA GOV. BONDS	35.108175	0.309608	49.445764

Tables 5 and 6 below present the Conditional Coverage or Independence (LR_1) statistic for the 95% and 99% levels of confidence, respectively. Christoffersen's statistic for conditional coverage tests the null hypothesis that the sequence of VaR exceptions is independent (exceptions are serially uncorrelated). Under this hypothesis the independent likelihood ratio has also a Chi-squared distribution with one degree of freedom. Thus the critical values are the same as in the unconditional coverage. That is $\chi_{0.01} = 6.6349$ and $\chi_{0.05} = 3.84146$ for 1% and 5% significance level, respectively.

For the 5% level the null hypothesis holds for all three methods for the S&P 500, UK and US Government Bonds. However for FTSE ALL SHARE the hypothesis can be rejected for the Variance Covariance and Historical Simulation approaches. For NIKKEI 500 it can be marginally rejected for the Hybrid

approach, as the LR_1 is very close to the critical value $\chi_{0.05}$, but holds for the other two. The conditional coverage of the three methods varies across the indices. For the 1% level the null hypothesis holds for all the VaR approaches and Variance covariance seems to out perform the other two.

TABLE 5.

Independence for 95% level of confidence

Independence	VCV	HS	Hybrid
S&P 500	0.062797	0.337204	0.137770
FTSE ALL SHARE	6.740856	5.529450	0.462945
NIKKEI 500	2.917809	0.224865	3.985477
UK GOV. BONDS	2.536382	0.546555	0.055021
USA GOV. BONDS	0.148127	0.239264	1.096008

TABLE 6.

Independence for 99% level of confidence

Independence	VCV	HS	Hybrid
S&P 500	0.062919	0.896707	0.265732
FTSE ALL SHARE	0.016509	1.191926	5.213647
NIKKEI 500	1.127697	1.547100	0.607924
UK GOV. BONDS	0.043859	0.810118	2.590804
USA GOV. BONDS	0.164703	0.728966	0.040515

Tables 7 and 8 below present the Correct Conditional Coverage (LR_{CC}) statistic for the 95% and 99% levels of confidence, respectively. This likelihood ratio tests both the independence of exceptions and the correct coverage. It is a mixed test of these two hypotheses. The distribution of LR_{CC} test is asymptotically Chi-square with two degrees of freedom. The critical values at the 1% and 5% level are $\chi_{0.01}=9.21034$ and $\chi_{0.05}=5.9914$, respectively. Just like the two previous tests if LR_{CC} for the desired significance level is lower than the critical value, the hypothesis can be accepted.

For the 5% level Historical Simulation produces the best correct conditional coverage, as the null hypothesis can be accepted in all five return series. The hypothesis holds also for the Variance Covariance approach except in FTSE ALL SHARE and NIKKEI 500 where it can be rejected. Hybrid method only makes it acceptable for the NIKKEI 500 and S&P 500. For the 1% level

the null hypothesis can be strongly rejected for the Variance Covariance and Hybrid methods in all five indices while it holds everywhere for Historical Simulation.

TABLE 7.

Correct Conditional Coverage for 95% level of confidence

Correct Coverage		HS	Hybrid
S&P 500	0.134414	0.370640	7.604877
FTSE ALL SHARE	8.153074	5.800699	10.260436
NIKKEI 500	7.627041	0.230125	5.544968
UK GOV. BONDS	4.151962	0.579991	9.852512
USA GOV. BONDS	1.193628	0.491198	8.563115

TABLE 8.

Correct Conditional Coverage for 99% level of confidence

Correct Coverage	VCV	HS	Hybrid
S&P 500	27.204760	0.933271	40.529766
FTSE ALL SHARE	16.185708	1.330385	83.926406
NIKKEI 500	13.798499	2.123343	77.068699
UK GOV. BONDS	27.185700	0.950591	66.046971
USA GOV. BONDS	35.272878	1.038574	49.405250

Tables 9 and 10 below report Lopez' magnitude of exceptions statistic for 95% and 99% level of confidence, respectively. The score, which is imposed on the exceptions, increases together with the magnitude and thus can provide information for the lower tail of the distribution of returns. We compare the number of exceptions (that incorporate the magnitude term) in order to decide which of the three alternative VaR approaches is more accurate. That is, the approach with the lowest score is more accurate than the others. In both 1% and 5% levels Historical Simulation outperforms the other two methods by achieving the lowest score of exceptions. On the other hand Hybrid exhibits the higher score. However it is worth noticing that in 5% level the differences are not so large as in 1% level.

TABLE 9.

Magnitude of exceptions for 95% level of confidence

Magnitude of exceptions	VCV	HS	Hybrid
S&P 500	148.019984	147.019852	178.020315
FTSE ALL SHARE	159.007066	151.006549	183.007261
NIKKEI 500	171.022432	144.017575	194.023850
UK GOV. BONDS	160.003577	147.003233	183.003530
USA GOV. BONDS	157.009187	139.007989	178.011529

TABLE 10.

Magnitude of exceptions for 99% level of confidence

Magnitude of exceptions	VCV	HS	Hybrid
S&P 500	61.009774	30.005128	69.010792
FTSE ALL SHARE	53.002351	27.001624	88.004649
NIKKEI 500	50.008293	25.004837	87.012553
UK GOV. BONDS	61.001320	31.000656	81.001701
USA GOV. BONDS	66.003538	32.001433	74.004709

Tables 11 and 12 below report the Mean Relative Bias statistic for 95% and 99% confidence levels, respectively. This statistic is a measure of size for each VaR approach and it is expressed in percentage terms. For example the mean relative bias of a VaR estimation method is 0.20. This means that this method is 20% larger than the average of all three approaches. The MRB results for the 95% level of confidence indicate that the differences between the average sizes of the different VaR forecasts of the three approaches are not large. The estimates of all methods vary around the mean at a maximum of 3.7% (Hybrid for FTSE ALL SHARE) except in one occasion where Historical Simulation estimates are 5.68% larger than the average. For the 99% level of confidence the differences are slightly larger. The results for the size of the VaR measures are similar to those of 95% level. In both levels Variance Covariance seems to produce the smallest risk measures in most of the indices in contrast to Historical Simulation that generates the largest. Finally, Historical Simulation estimates are constantly larger than the average whereas Hybrid approach estimates are constantly smaller.

TABLE 11.

Mean Relative Bias for 95% level of confidence

MRB	VCV	HS	Hybrid
S&P 500	0.010779	0.011069	-0.021848
FTSE ALL SHARE	0.017961	0.018978	-0.036939
NIKKEI 500	-0.022338	0.056808	-0.034470
UK GOV. BONDS	-0.010677	0.025769	-0.015092
USA GOV. BONDS	-0.021280	0.029687	-0.008406

TABLE 12.

Mean Relative Bias for 99% level of confidence

MRB	VCV	HS	Hybrid
S&P 500	-0.056752	0.117750	-0.060998
FTSE ALL SHARE	-0.012056	0.101040	-0.088985
NIKKEI 500	-0.044791	0.133331	-0.088540
UK GOV. BONDS	-0.065378	0.113905	-0.048527
USA GOV. BONDS	-0.100333	0.141079	-0.040746

Tables 13 and 14 below report the Root Mean Squared Relative Bias statistic for 95% and 99% confidence levels, respectively. This test examines the degree to which the different risk measures tend to vary around the average risk measure. For 95% level of confidence the results vary. First, the risk measures produced by Hybrid method deviate most of the mean in all indices. For the S&P and FTSE ALL SHARE Variance Covariance estimates deviate more than those of Historical Simulation while for NIKKEI 500, UK Government Bonds and USA Government Bonds do not. For the 99% level Variance Covariance method produces the risk measure, which is the closest to the average of all three approaches for all indices, in contrast to Hybrid, which produces the furthest (except in USA Government Bonds where Historical Simulation deviates most).

TABLE 13.

Root Mean Squared Relative Bias for 95% level of confidence

RMSRB	VCV	HS	Hybrid
S&P 500	0.089278	0.088996	0.174621
FTSE ALL SHARE	0.089331	0.084486	0.167961
NIKKEI 500	0.078191	0.099702	0.158504
UK GOV. BONDS	0.093015	0.096805	0.182552
USA GOV. BONDS	0.088611	0.097226	0.173688

TABLE 14.

Root Mean Squared Relative Bias for 99% level of confidence

RMSRB	VCV	HS	Hybrid
S&P 500	0.108577	0.162742	0.197219
FTSE ALL SHARE	0.081859	0.135692	0.191474
NIKKEI 500	0.090815	0.164807	0.177356
UK GOV. BONDS	0.109101	0.154623	0.190427
USA GOV. BONDS	0.124966	0.169792	0.164649

6. Conclusion

VaR forecasts are of great importance for financial and risk management. The relative literature includes a variety of different methods. In this study we employ a broadly used forecasting methodology, the EWMA model. We use this method in combination with three different VaR estimation approaches and generate three different types of daily VaR forecasts based on each individual underlying approach for three equity and two bond portfolios. The aim of this study is to evaluate the forecasting performance of these methods and to determine which of them produces the best forecasts. In the evaluation procedure we use a number of back testing techniques that attempt to capture the accuracy and the variability of VaR estimates.

The tests that aim to evaluate the accuracy of VaR forecasts indicate that Historical Simulation generate the most accurate forecasts. Specifically, Historical Simulation exhibits the lowest proportion of failures and the lowest score of exceptions in Lopez' test. Additionally it produces the best unconditional and correct coverage. For the conditional coverage in 95% level of confidence the results vary, while for 99% Variance Covariance produces the best coverage. Between Variance Covariance and Hybrid the first appears to per-

form slightly better. The tests for the variability of VaR forecasts indicate that the differences between the average sizes of the different VaR estimates are not large for both 95% and 99% percent levels of confidence. However, Hybrid and Variance Covariance (in most cases) tend to produce estimates smaller than the average of all three methods while Historical Simulation produces constantly higher estimates. It is important to note that this is also the case even when confidence level varies, as they can have an effect on the performance of different VaR approaches.

Finally, this study considers three alternative approaches of estimating VaR and some of the tests that can verify their accuracy. They have been examined and evaluated separately and conclusions have been drawn as to the one which individual produces the best estimates. An interesting subject for further examination would be the investigation of *combined* approaches and how the accuracy of the estimates compares to those currently evaluated at an individual level in this study.

References

- Bodoukh J., Richardson M. and Whitelaw R., 1997, "Investigation of a Class of Volatility Estimators", *The Journal of Derivatives* 4, p.63-71.
- Bodoukh J., Richardson M. and Whitelaw R., 1998, "The Best of Both Worlds: A Hybrid Approach to Calculating Value at Risk", *Working Paper*.
- Bollerslev T., 1986, "Generalised Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics* 31, p.307-327.
- Basel Committee on Banking Supervision, 2001, The New Basel Capital Adequacy Accord, BIS, Switzerland.
- Basel Committee on Banking Supervision, 1999d, Performance Of Models-Based Capital Charges for Market Risk: July - December 1998, BIS, Switzerland.
- Christoffersen P., 1998, "Evaluating Interval Forecasts", *International Economic Review* 39, p.841-862.
- Dowd K., 1998, *Beyond Value at Risk: The new science of risk management*, Wiley.
- Haas M., 2001, "New Methods in Back testing", *Working Paper*.
- Hendricks D., 1996, "Evaluation of Value at Risk Models using Historical Data", *Federal Reserve Bank of New York Economic Review*, April. p.39-69.
- Hull J. and A. White, 1998, "Incorporating Volatility Updating into the Historical Simulation Method for Value at Risk", *Journal of Risk* 1, p.5-19.
- Jorion P., 1996, "*Risk²: Measuring the Risk in Value at Risk*", *Financial Analysts Journal* 52, November, 47-56.

- Jorion P., 2000, *Value at Risk: The New Benchmark for Controlling Market Risk*, Mc Graw Hill.
- JP Morgan, 1996, *Risk Metrics Technical Document*, 4th edition, New York.
- Kupiec P. H., 1995, "Techniques for Verifying the Accuracy of Risk Measurement Models", *The Journal of Derivatives* 3, p.73-84.
- Lopez J., 1999 "Methods for Evaluating Value at Risk Estimates", *Federal Reserve Bank of San Francisco Economic Review*, No. 2, p.3-17.
- Nelson D., 1992 "Filtering and Forecasting with misspecified ARCH models: Getting the right variance with the wrong model", *Journal of Econometrics* 52, p. 61-90.

Footnotes

1. Dowd K., 1998, "Beyond Value at Risk: The new Science of Risk Management", Willey, p. 121.
2. JP Morgan, 1996, *Risk Metrics Technical Document*, 4th edition, New York.
3. Bollerslev T., 1986, "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics* 31, p. 307-327.
4. For further information on how Primark Datastream indices are calculated see the database brochure.
5. JP Morgan, 1996, *Risk Metrics Technical Document*, 4th edition, New York.
6. Nelson D., 1992 "Filtering and forecasting with misspecified ARCH models: Getting the right variance with the wrong model", *Journal of Econometrics* 52, p. 61-90.
7. Hull J. and A. White, 1998, "Incorporating Volatility Updating into the Historical Simulation Method for Value at Risk", *Journal of Risk* 1, p. 5-19.
8. Bodoukh J., Richardson M. and Whitelaw R., 1998, "The Best of Both Worlds: A Hybrid Approach to Calculating Value at Risk", *Working Paper*.
9. Kupiec P. H., 1995, "Techniques for Verifying the Accuracy of Risk Measurement Models", *The Journal of Derivatives* 3, p.73-84
10. Christoffersen P., 1998, "Evaluating Interval Forecasts", *International Economic Review* 39, p.841-862.
11. Lopez J., 1999 "Methods for Evaluating Value at Risk Estimates", *Federal Reserve Bank of San Francisco Economic Review*, No. 2, p. 3-17.
12. Hendricks D., 1996, "Evaluation of Value at Risk Models using Historical Data", *Federal Reserve Bank of New York Economic Review*, April. p. 39-69.