Faculdade de Economia da Universidade Nova de Lisboa

What Is The Value Of Value-at-Risk After All?
A Conditional Approach Using Quantile Regressions

Carla Sofia Nobre Peixoto

Master in Finance, nr. 87

A Work Project under the supervision of:

Professor Pedro Santa-Clara

12 June 2009

What Is The Value Of Value-at-Risk After All?
A Conditional Approach Using Quantile Regressions[*]

Carla Peixoto

Abstract

In this Work Project, I propose a new approach to VaR estimation based on quantile regressions which does not require any distributional assumptions. I assume that there exist some state variables that capture persistent changes in risk. This methodology intends to solve the problem of lack of conditionality in VaR models and to capture volatility clustering where existing VaR models currently fail. I compare the out-of-sample performance of existing methods in predicting daily VaR for the S&P 500. I conclude that none of the methodologies developed so far produce satisfactory results in timing unexpected increases in market volatility. Moreover, alternative out-of-sample evaluation techniques yield to opposite results regarding the best VaR model. Nonetheless, in general, the GARCH model outperforms all the remaining models.

Keywords: Value-at-Risk, Conditional approach, Quantile Regression, Out-of-sample forecasting.

[*] I am grateful to Professor Pedro Santa-Clara for all his support and assistance in the development of this Work Project. I am also grateful to Professors Miguel Ferreira, José Faias, Filipe Lacerda and Tiago Botelho for helpful discussions and suggestions. I would further like to thank Professor José Ferreira Machado for valuable advice regarding quantile regression inference.

"Models work in normal markets, and normal markets are defined as those within which models work" – **Christopher Finger** (2009), RiskMetrics Group

2007 was bad, 2008 was worse, and 2009 is catastrophic: The World economy is experiencing what may be considered the deepest downturn since the Great Depression of the 1930s as a result of the crisis in the credit market and banking system (Campbell (2009)). The current financial crisis is partly due to financial institutions being highly leveraged and exposed to risk. Thus, the need to develop accurate market risk management tools has never been greater.

Market risk is the possibility of incurring in losses resulting from fluctuations in the market prices of financial securities. The time-varying nature of financial return volatility is well documented in the literature (starting with the findings of Schwertz (1989) and analyzed for example by Figlewski (1997) and Jondeau and Rockinger (2003)). During the current 2007-2009 financial crisis, we have been witnessing an unprecedented increase in market volatility. In particular, the Chicago Board Options Exchange's (CBOE) VIX index, which measures market volatility implied by 30-day S&P option prices, increased from 22.5% at the beginning of 2008 to 80.9% on November 20, 2008. As a result, commercial banks aiming to maintain a constant level of risk in their trading portfolios had to lower considerable their equity exposure. For instance, UBS justifies its large net loss during 2008 as a result of increased risk exposure that *"remained greater than UBS risk capacity"*.

The increased volatility in financial markets and the growing complexity and innovation in financial instruments have encouraged researchers, practitioners and regulators to develop risk management tools. In this context, and despite some criticisms,[1] Value-at-Risk

[1]One of the major shortcomings of VaR as a risk management tool is the lack of information about the maximum possible loss when VaR is exceeded. As a result, Expected Shortfall (or Conditional Value-at-risk) has been proposed in the literature as an alternative measure of risk See Acerbi and Tasche (2002), Artzner et. al (1999), Basak and Shapiro (2000), Berkowitz (2001), Chernozhukoy and Umantsev (2001), Granger (2002), Pflug (2000) and and Taylor (2008a, 2008b) for further discussion about criticisms on VaR and for Expected Shortfall.

(hereafter VaR) has emerged as the standard measure for assessing market risk both for internal risk management and for regulation purposes (Dowd (1998) and Jorion (2006)). In particular, the Basel Committee on Banking Supervision (1996) requires U.S. commercial banks to allocate regulatory capital as direct function of VaR estimates to cover potential financial losses.[2] Therefore, proving accurate VaR estimates is crucial to avoid excessively low capital levels, which can lead to bankruptcy in extreme situations. Examples of failures in risk management that resulted in the collapse of institutions include Long-Term Capital Management (LTCM) hedge fund, Barings bank, Metallgesellschaft, Orange County, Daiwa bank and Allied Irish Banks (AIB) (see Jorion (2006) for details on these risk case-studies). More recently, the collapse of Lehman Brothers, Northern Rock and Bear Stearns as well as the severe problems faced by giants like AIG, Lloyds, Merrill Lynch and HBOS warm up for the consequences of underestimation of market risks in the Economy.[3]

Intuitively, VaR summarizes the worst expected loss over a given holding period that will not be exceeded with a certain confidence level (Jorion (2006), page 17). More specifically, conditional on the information available up to time t, the h-period-ahead VaR forecast corresponds to the negative τ-quantile of the conditional return distribution[4]:

$$VaR_{t+h}^{\tau} = -Q_{\tau}(r_{t+h}|\Omega_t) = -\inf_x\{x \in \mathbb{R} : \mathrm{P}(r_{t+h} < x|\Omega_t) \geq \tau\} \quad , \; 0 < \tau < 1 \qquad (1)$$

where $Q_{\tau}(.)$ denotes the quantile function, r_{t+h} is the return on an asset or portfolio in the holding period $t+h$, and Ω_t represents the information available up to time t.

[2] The Basel Committee on Banking Supervision demands U.S. commercial banks to set minimum capital requisites as a function of market risk, credit risk and operational risk through the 1996 Market Risk Amendment to the 1988 Basle Capital Accord. Moreover, it penalizes for model inaccuracy by increasing capital requirements through a multiplicative factor. For more details see Basle Committee on Banking Supervision (1996) and Federal Register (1996).

[3] In brief, international banks have been hedging their exposures to credit via Credit Default Swaps (CDS) contracted with insurance companies like AIG. With this mechanism, banks became excessively leveraged while insurance companies were able to record huge profits. Since CDS are unregulated contracts, there was no capital requirement to cover the risk of these securities as long as they maintain the AAA-rating (easily assured by commercial banks, despite the existence of "subprime toxic assets"). As AIG did not properly collaterized the risk of CDS, it run out of cash as soon as investors realized that CDS were not AAA-rated but also included "subprime toxic assets".

[4] VaR is usually expressed as a positive quantity, although it represents a loss.

The extensive usage of VaR by financial institutions is explained by its conceptual simplicity as it reduces market risk to a single number. Nevertheless, one important question arises: how well do VaR measures perform empirically? As emphasized by Alan Greenspan (1996), "*Disclosure of quantitative measures of market risk, such as value-at-risk is enlightening only when accompanied by a thorough discussion of how the risk measures were calculated and how they related to actual performance*". In fact, although several statistical techniques have been proposed since the Market Risk Amendment to the Basle Accord (including RiskMetrics and Historical Simulation), none of the methodologies developed so far presents satisfactory results. Several empirical studies conclude that existing bank VaR models perform poorly and are easily outperformed by simple econometric models (see Berkowitz and O'Brien (2002), Pérignon and Smith (2007), and Pérignon, Deng and Wang (2007) for consistent evidence). More recently, during the 2007-2009 financial crisis, the World's experienced an extraordinary number of exceptions,[5] providing further evidence on the failure of risk management models. For example, during 2008, BNP Paribas and Credit Suisse reported 7 and 24 exceptions, respectively (both at the 99% confidence level), while Goldman Sachs and Morgan Stanley experienced 13 and 18 exceptions, respectively (at the 95% confidence level). More seriously, UBS experienced the most remarkable failure in risk management by recording 50 exceptions on its VaR estimates at the 99% confidence level! Put differently, UBS observed once per week what it was thought to be a one-in-a-hundred event! (see Campbell (2009) for more details on VaR exceptions during 2008). This is in line with the findings of Berkowitz and O'Brien (2002) that provide evidence that exceptions of VaR in commercial banks tend to cluster over time, suggesting that bank VaR models are not able to capture the time-varying nature of banks' trading portfolios. As a result, an exception today tends to be followed by an exception tomorrow. Such VAR

[5] An exception is any actual trading loss higher that the VaR estimate. Theoretically, a bank that calculates VaR at a 99% confidence level should expect to experience two or three exceptions within a year. At 95% confidence level, exceptions should occur roughly once a month.

violation clustering is evidence of lack of conditionality in bank VAR systems (See also Jackson, Maude and Perraudin (1997)). The challenge of forecasting VaR consists then on capturing time-varying conditional quantiles of return distribution. As emphasized by Peter Davies, "*To have a future in risk management, one needs to induce the future in risk measurement",* which suggests that effective risk management models must rely on an analysis of future market conditions.

In this Work Project, I propose an alternative and innovative approach to calculate VaR, RQVaR, based on quantile regressions introduced by Koenker and Bassett (1978). I assume that there exist some state variables that are able to forecast the risk of financial securities by capturing unexpected and persistent changes in market volatility.[6] This methodology intends to solve the problem of lack of conditionality in VaR models and to capture volatility clustering where existing VaR models currently fail. Moreover, and unlike fully parametric models, I model directly the quantile corresponding to VaR instead of the whole distribution. Since there is evidence that log-returns are not normally distributed, this procedure avoids making unrealistic assumptions.

This empirical study compares the out-of-sample performance of alternative VaR models: Historical Simulation, Exponential Smoothing, RQVaR and GARCH. To assess the predictive performance of VaR models, I use daily data on the S&P 500, which is representative of a diversified portfolio of volatile assets. Overall, I conclude that none of the methodologies is able to produce accurate forecasts for the S&P 500, especially in periods of crisis. If models are ranked according to the number of exceptions,[7] GARCH (1,1) and Historical Simulation outperform Quantile Regressions and Exponential Smoothing. Nonetheless, if models are ranked according to the magnitude of exceptions,[8]

---

[6] In fact, several authors have shown that variables like the level of interest rates, term spread, implied volatility and the default spread have forecasting power for both the first and second moments of returns of stocks and bonds.
[7] The actual trading losses higher VaR estimates are named exceptions.
[8] The magnitude of exceptions concerns the amount of the losses that are beyond VaR estimates.

Historical Simulation presents the worst performance, followed by Exponential Smoothing. In this context, I show that the most widely used method by U.S. commercial banks - Historical Simulation - produce conservative VaR estimates that, even though they are exceeded with an acceptable probability, produce a high amount of unexpected losses. By contrast, Quantile Regression and GARCH (1,1) models are able to successfully time the increases in market volatility, presenting the lowest magnitude of exceptions.

The remainder of this Work Project is structured as follows. Section 1 briefly summarizes the major existing statistical methods to VaR estimation. Section 2 introduces the alternative Quantile Regression approach I am proposing. In Section 3, I present the statistical procedures to compare the out-of-sample performance of the different forecasting models. Section 4 describes the data and methodology used. In Section 5, I discuss the empirical results of the alternative VaR models. The final section provides concluding remarks and briefly explores some relevant issues for further research.

1. Alternative VaR estimation methods

In practice, existing approaches to estimate VaR seek to fit within the characteristics of financial data, which can be summarized by three well documented stylized facts: (i) volatility clustering, *"…in which large changes in volatility tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes"* (Mandelbrot (1963)); (ii) negative skewness; and (iii) excess kurtosis compared to a normal distribution, indicating heavier tails and a sharper peak (see for example Harvey and Siddique (1999), Jondeau and Rockinger (2002) and Mittnik and Paolella (2003) for evidence of excess kurtosis, asymmetry and volatility clustering in financial returns). Pérignon and Smith (2009) conclude that trading revenues of a sample of commercial banks share the stylized facts that characterize financial returns, namely high volatility, negative skewness, excess kurtosis, volatility clustering and modest autocorrelation.

Existing approaches to estimate VaR can be classified as follows: (i) Historical Simulation; (ii) Fully Parametric models; (iii) Semi-parametric models; and (iv) Quantile Regression. Good reviews of recent VAR literature include Andersen et al (2007), Engle and Manganelli (2004), Kuester et al (2006), Duffie and Pan (1997), Jorion (2006).

1.1 Historical Simulation

Historical Simulation (henceforth HS) estimates VaR simply by evaluating the empirical quantile of the distribution of returns based on past data. (see Christoffersen (2003) and Jorion (2006)). Formally, the one-day-ahead VaR estimate is given by the negative empirical τ-quantile, $\hat{Q}_\tau(.)$, of a moving window of κ observations up to time t:

$$\widehat{VaR}\,^{\tau}_{t+h} = -\hat{Q}_\tau(r_t, r_{t-1}, \dots, r_{t-\kappa-1}) \tag{2}$$

The popularity of HS at commercial banks has been noted by Pritsker (2006), Berkowitz and O'Brien (2002), Berkowitz, Christoffersen and Pelletier (2008) and Pérignon, Deng and Wang (2008). Importantly, HS does make any explicit assumption about the distributional model generating the returns. The main shortcoming of HS is the implicit assumption that future returns behave similarly to past returns. This approach ignores the fact that returns are not independent and identically distributed (i.i.d.) and, more seriously, it fails in capturing volatility clustering.[9] Moreover, VaR estimates based on HS may present predictable jumps due to discreteness of extreme jumps. For a detailed analysis on the advantages and disadvantages of HS see Dowd (2002) and Christoffersen (2003).

[9] The empirical quantile estimator is consistent only if the window size, κ, goes to infitnity. However, if the window size is too large the VaR estimates will not be able to capture the current volatility cluster. Andersen et al (2005) argue that including a crash in the sample may not change significantly the VaR estimate through historical simulation, if the new second smallest return is similar to the previous one.

More recently, Filtered Historical Simulation was introduced by Barone-Adesi, Bourgoin and Giannopoulos (1998).[10] In addition, Boudoukh, Richardson and Whitelaw (1998) propose a hybrid approach to HS, in which they combine volatility models and HS methodologies.

1.2 Fully parametric models

Unlike HS, fully parametric models characterize the whole distribution of returns, by specifying the return process as:

$$r_t = \mu_t + \varepsilon_t \qquad \text{and} \qquad \varepsilon_t = \sigma_t z_t \tag{3}$$

where μ_t denotes the expected value of returns at time t, ε_t the innovations of returns, σ_t the standard deviation and z_t is a i.i.d. process with zero mean and variance of one[11].

1.2.1 GARCH models

The GARCH (Generalized Autoregressive Conditional Heretoskedasticity) model introduced by Bollerslev (1986) assumes that variances are time-varying and follow a predictable process. The simplest symmetric GARCH (1,1) model states that the conditional variance depends on the latest innovation, ε_{t-1}, and on the previous conditional variance, σ_{t-1}^2:

$$\sigma_t^2 = \omega + \alpha\, \varepsilon_{t-1}^2 + \beta\, \sigma_{t-1}^2 \tag{4}$$

[10] For empirical evidence on the superior performance of this method relatively to the traditional HS see for example Barone-Adesi, Giannopoulos and Vosper (1999, 2002), Pritsker (2001) and Kuester et al (2005))

[11] Unconditional parametric models set $\mu_t \equiv \mu$ and $\sigma_t \equiv \sigma$, assuming that returns are i.i.d. and have a constant expected value and variance. This is an unrealistic assumption, given the time-varying nature of expected returns and volatility of financial returns (see for example, Schwertz (1989)). To overcome this problem, conditional homoskedasticity parametric models allow for a time-varying mean, possibly captured by an ARMA (p,q) process (as described in Kuester e tal (2005)). However, these models are of relatively marginal use for risk management. Thus, I will not consider them.

Although several extensions of this model have been proposed in the literature, I will only consider GARCH (1,1), given its superior out-of-sample performance compared with other (and more complex) volatility models (see Hansen and Lundle (2004) for a detailed analysis of out-of-sample predictive ability of GARCH-family models). For notation convenience, I will denote GARCH (1,1) simply by GARCH.

GARCH is a nonlinear model, whose parameters are estimated by maximum likelihood usually based on a Gaussian distribution for the return innovations, ε_t (see Duffie and Pan (1997)).[12]

The one-day-ahead VaR estimate using a GARCH model is given by:

$$\widehat{VaR}\,^{\tau}_{t+h} = -\,\sigma^{GR}_{t+h}\,.\,\Phi_t \tag{5}$$

where σ^{GR}_{t+h} is the volatility forecast by the GARCH in equation (4) and Φ_t is the τ-quantile of a standard normal distribution.

1.2.2 Exponential Smoothing approach

The Exponential Smoothing approach (aforementioned ES) introduced by J.P. Morgan (1995) models the variances of financial returns at time t as an exponentially weighted moving average (EWMA) of past squared innovations[13]:

$$\sigma_t^2 = \sum_{\mathrm{j}=0}^{n} (1-\lambda)\lambda^{\mathrm{j}}\ \varepsilon^2_{t-1-j} \quad , \quad \lambda < 1 \tag{6}$$

12 The assumption of Gaussian disturbances in equation (3) is consistent with the normality assumption of log-returns in the finance literature. However, since market returns exhibit excess kurtosis relative to a normal distribution, Baillie and Bollerslev (1989) suggest the use of a t-distribution for GARCH models in order to accommodate fat tails. Other authors suggest alternative distributions for the return disturbances, like Generalized Error Distribution (GED) or skewed t-distribution. In this analysis I will focus on the case of Normal innovations, since the use of alternative distributions imply higher estimation error that is not compensated by superior out-of-sample performance (Hansen and Lundle (2004)). Moreover, even if the true distribution is not normal, the parameters estimated by maximum likelihood are consistent (as demonstrated by Bollerslev and Wooldridge (1992)). This method is called quasi-maximum likelihood.

13 Since it is common to assume that the average daily return is zero innovations coincide with the returns. For more details on this methodology see J.P. Morgan RiskMetrics Technical Document (1995).

where λ denotes the decay factor, ε_{t-1-j}^{2} the past squared innovation of order j, and n the number of days used in the estimation.

Put differently, ES places geometrically declining weights on past observations, assigning greater importance to recent observations.[14] The estimation of VaR using ES is particularly simple because future volatility of returns relies only on one parameter, the decay factor, which in practice is not estimated. It is common to adopt a decay factor of 0.94 for daily data, which is the optimal value chosen by J.P. Morgan to best fit the data.[15] Since ES requires the use of a single parameter, it implies less estimation error than GARCH. As a result, ES presents superior robustness of this model relatively to GARCH.

Similarly to GARCH approach, the daily VaR using ES is estimated as:

$$\widehat{VaR}_{t+h}^{\tau} = -\,\sigma_{t+h}^{ES}\,.\,\Phi_t \tag{7}$$

where σ_{t+h}^{ES} denotes the volatility estimated by the ES model in equation (6) and Φ_t is again the τ-quantile of a standard normal distribution.

As emphasized by Engle and Manganelli (2004b), both the ES and GARCH-type models tend to underestimate VaR, as the normality assumption for the innovations is not consistent with the fat-tailed behavior of financial returns.[16]

1.3 Semi-parametric models

Recently, alternative methodologies have been introduced in the literature, including applications of Extreme Value Theory (EVT) to VaR (see for eg. Danielsson and de Vries (2000), and Gourieroux and Jasak (1998)). EVT parametrically models the tails of return

[14] This model is also called Integrated GARCH (IGARCH), being a special case of GARCH models (Jorion (2006)).
[15] Fleming, Kirby and Ostdiek (2001) and Foster and Nelson (1996) have also found an optimal decay factor equal to 0.94 for daily data, using nonparametric estimation techniques.
[16] They further identify three types of misspecification: the variance equation may not be well-specified; (ii) the distribution for the log-likelihood function may be wrong, and (iii) the standardized residuals may not be i.i.d.

distribution instead of the whole distribution of returns. Alternatively, McNeil and Frey (2000) suggest fitting a GARCH model to time series of returns and then applying the EVT to the standardized residuals.[17] Other procedures include Hull and White (1998) approach to HS and Monte Carlo simulation.

1.4 Quantile Regression

Quantile regression approach (hereafter QR) introduced by Koenker and Basset (1978) models directly the quantile of interest for VaR estimation instead of the whole distribution of returns.[18] As a result, and unlike parametric models, it does not require any distributional assumptions for the behavior of returns.

The intuition behind QR is that we are able to model the conditional τ-quantile of return distribution, as a function, f, of the information X_t available up to time t:

$$\widehat{VaR}\,^{\tau}_{t+h} = -f_{\tau}(X_t; \beta_{\tau}) \tag{8}$$

where $f(.)$ and the parameter vector β_{τ} directly depend on τ.

Applications of QR to risk management include Engle and Manganelli (1999) Conditional Autoregressive Value at Risk (CAViaR) model, Chernozhukov and Umantsev (2000) empirical study, and Taylor (1999) VaR estimation in the context of exchange rates.

In this Work Project, we propose an alternative approach using QR, RQVaR further developed in Section 2.

[17] For an overview of EVT models see for example Christoffersen (2003).
[18] This approach is consistent with the findings of Bao, Lee and Saltoglu (2004), who conclude that a model that provides superior density forecasts for the whole return distribution may not necessarily meet the needs of risk managers who care much more about the tails.

2. RQVaR – Regression Quantile Value-at-Risk

As an alternative to existing models, I propose a novel approach for VaR estimation using QR. I assume that there exist some state variables that forecast one-day-ahead VaR of risky assets, by timely capturing dramatic and persistent changes in market volatility. Put differently, I propose a conditional approach that intends to timely signal investors of risk increases using lagged market variables. Examples of such state variables include VIX, term spread, TED spread and default spread, whose relationship in predicting risk and return has been widely studied in the literature.

More specifically, I estimate the relationship between the state variables and VaR by directly modeling the quantile of interest for VaR, as described by equation (8).[19] The choice of the appropriate functional form for function f should yield a close approximation to the population quantiles (as discussed in Chernozhukov and Umantsev (2001)). I consider the linear functional form[20]:

$$\widehat{VaR}\,_{t+h}^{\tau} = -\left[\beta_{0,\tau} + \beta_{1,\tau}\, X_t\right] \tag{9}$$

where X_t denotes a given state variable.

Koenker and Bassett (1978) show that the τ-th regression quantile estimator can be calculated by minimizing the average of asymmetrically weighted absolute errors (with weight τ on positive errors and weight $(1-\tau)$ on negative errors).

[19] The use of quantile regression is justified by the empirical evidence on different estimated coefficient across quantiles, suggesting that regressors may have different impacts on the dependent variable at different locations of the conditional distribution. As discussed in Engle and Manganelli (2004), quantile regression include a special case of least absolute deviation (LAD), which is considered a more robust method than ordinary least squares (OLS) whenever errors have a fat tail distribution.

[20] I consider several functional forms in this Work Project: linear, quadratic and exponential relationship between the state variables and VaR. Nonetheless, the linear form has proven to yield the closest approximation to the population quantiles. Thus, I will not present the results for the other functional forms.

Formalizing, the τ-th regression quantile estimator is defined as:

$$\hat{\beta}_\tau = argmin\left\{\sum_{r_t \geq X_t'\beta_\tau} \tau |\, r_t - X_t'\beta_\tau| + \sum_{r_t < X_t'\beta_\tau} (1-\tau)|\, r_t - X_t'\beta_\tau|\right\} \tag{10}$$

where X_t denotes a matrix of ones and one single state variable.

To calculate the standard errors of the estimates and assess the econometric significance of estimated coefficients, I apply bootstrapping.[21]

In line with the methodology proposed by Connor (1997), I further apply shrinkage to the estimated parameters to increase the robustness of the estimated parameters and reduce sampling error. However, since shrinkage does not improve out-of-sample performance of VaR estimates, I do not present here the results.

3. Comparing and testing the out-of-sample performance of alternative VaR models

To assess the predictive performance of alternative VaR models, I follow Christoffersen (1998) framework, which is designed to evaluate the accuracy out-of-sample interval forecasts. According to Christoffersen (1998), if VaR exceptions occur with the correct conditional and unconditional probability the VaR forecasts are efficient and cannot be improved:

$$E\left[\left(r_{t+h} < -\widehat{VaR}\,_{t+h}^{m\,\tau}\right) \mid \Omega_t\right] = \tau \tag{11}$$

where $\widehat{VaR}\,_{t+h}^{m\,\tau}$ denotes the VaR estimate on day $t+h$, at τ confidence level, derived from model m for a one-day-ahead return.

[21] In QR framework, the asymptotic variance matrix of the estimator depends on the density of the error. In this context, the bootstrap distribution is shown to converge weakly to the limit distribution of the QR estimator in probability.

To identify specific model inadequacies, I compute a set of individual tests instead of a joint test on the accuracy of VaR estimates. Moreover, I develop a dynamic strategy test, according to which an investor is able to keep the level of risk of his portfolio fixed.

3.1 Unconditional Coverage Test

The unconditional covered test proposed by Kupiec (1995) evaluates the adequacy of the VaR model by testing whether the number of observed model exceptions is in line with the significance level used in the VaR calculation. Accurate VaR estimates should exhibit an unconditional exception rate, $\hat{\tau} = N/T$, equal to τ%, where N denotes the number of observed exceptions using a given model, T the total number of observations, and τ the confidence level used for the VaR estimate. The appropriate hypothesis for this test is:

$$H_0: \hat{\tau} = \tau \tag{12}$$

Under the null hypothesis, the likelihood ratio statistic to test equation (10) is:

$$LR_{uc} = 2[log(\hat{\tau}^N(1-\hat{\tau})^{T-N}) - log(\tau^N(1-\tau)^{T-N})] \tag{13}$$

which is asymptotically distributed $\chi^2(1)$.

The unconditional covered test only considers the number of exceptions over the entire period. But, if returns exhibit time-dependent heteroskedasticity, there might be an important cluster effect in the exceptions that is neglected by this test. Thus, it is important to evaluate the independence and conditional covered tests.

3.2 Independence Test

Several tests have been proposed in the literature for the independence of VaR exceptions, including runs tests and the Ljung-Box test (Ljung and Box (1978)). More recently, a test based on the time between exceptions was proposed by Danielsson and Morimoto (2000). Under the null hypothesis, model exceptions are not serially correlated, meaning that an exception today has no influence on the probability of a violation tomorrow. The test statistic is the likelihood ratio statistic for the null hypothesis of serial independence:

$$LR_{ind} = 2\left[log\left(\hat{\tau}_0{}^{T_{00}}(1-\hat{\tau}_0)^{T_{01}}\hat{\tau}_1{}^{T_{11}}(1-\hat{\tau}_1)^{T_{10}}\right) - log\left(\hat{\tau}^{T_{01}+T_{11}}(1-\hat{\tau})^{T_{00}+T_{10}}\right)\right] \quad (14)$$

where $\hat{\tau}$ denotes the unconditional exception rate, $\hat{\tau}_0$ the exception rate conditional on no exception in the previous period, $\hat{\tau}_1$ the exception rate conditional on exception in the previous period, and T_{ij} denotes the number of observations in state j after having been in state i in the previous period. This statistic is asymptotically distributed $\chi^2(1)$.

3.3 Conditional Coverage Test

Christoffersen (1998) proposes a joint test of unconditional coverage and serial independence to test adequacy of VaR models. The relevant test statistic is:

$$L_{cc} = L_{uc} + L_{ind} \quad (15)$$

where L_{uc} is the statistic from equation (10) and L_{ind} the statistic calculated using equation (11). The L_{cc} statistic has an asymptotic distribution $\chi^2(2)$.

3.4 Dynamic Strategy Test

In this section, I develop a dynamic strategy based on VaR estimates. Consider a portfolio made of the S&P 500 (for which we are able to forecast VaR with alternative models) and the risk-free asset. The return on this portfolio, r_p, is given by:

$$r_p = w_1 r_{S\&P} + (1-w_1)r_f \quad (16)$$

where $r_{S\&P}$ denotes the daily return on the S&P 500 asset, w_1 the weigth invested in the S&P 500, and r_f the daily risk-free rate.

If VaR models are accurate, an investor is able to adjust the positions on the S&P 500 and on the risk-free rate in order to keep the VaR of the portfolio fixed at a given level, α.[22] The intuition behind the dynamic strategy consists of reducing the exposure on the S&P 500 when the VaR model forecasts risk to be high, and increase the exposure when the VaR forecast is low. If VaR is accurately estimated, this would result in a dynamic strategy with a series of portfolio returns with constant VaR.

The appropriate hypothesis testing is:

$$H_0: \widehat{VaR}_{rp}^{\tau} = -Q_\tau(r_p) = \alpha \tag{17}$$

where $Q_\tau(r_p)$ denotes the empirical τ-quantile of the distribution of portfolio returns, and α the desired level of VaR of the portfolio.

4. Data and VaR out-of-sample forecast

I evaluate the out-of-sample forecasting performance of alternative models in determining the 95% and 99% VaR of the S&P 500 index, which is representative of a diversified portfolio of volatile assets. The data comprises daily closing prices[23] on the S&P 500, from January 2, 1990 to May 5, 2009, yielding to 4876 log returns.[24]

In this empirical analysis, I select a subset existing VaR methodologies based on the frequency of usage by commercial banks. Pérignon and Smith (2007) show that HS (or related techniques) is used by 73% of the institutions that disclose VaR methodology. The

[22] To perform this test, we considered arbitrarily that α= 3%, meaning that I adjusted the weights on the S&P 500 and on the risk-free rate in order to keep the VaR of the portfolio constant at 3%. For simplicity, this is valid both for 95% and 99% VaR. A detailed demonstration of this test is presented in Annex 1.

[23] The closing prices are already adjusted for dividends, stock splits and other corporate actions.

[24] Log-return, r_t, are calculated as $r_t = \ln(P_t) - \ln(P_{t-1})$, where $P_t$denotes the closing price of the S&P at time *t*.

second most popular method is Monte Carlo simulation and other related parametric models, like ES. Accordingly, I compare the forecasting performance of HS, ES, RQVaR that I am proposing, and GARCH. I consider this latest model given the extensive literature on the superior performance of this model relatively to other VaR methods[25].

For HS, I calculate the empirical τ-quantile of a moving window of 252 observations (approximately one year), which allows to balance the tradeoff between consistency of VaR estimates and effectiveness in capturing volatility clustering, as described in section 1.2.

For the GARCH model, I consider an initial window of 2520 observations to estimate the first coefficients ω, α and β on equation (4). The coefficients are re-estimated every day using an expanding window.[26] VaR is subsequently calculated according to equation (5).

In what concerns ES, I set λ to 0.94 in equation (6) to calculate daily volatility of the S&P 500 and forecast VaR as described by equation (7).[27]

For RQVaR, I consider an initial window of 2520 observations to estimate the first coefficients β_τ, which are also re-estimated using an expanding window updated every day[28]. One-day-ahead VaR is estimated according to equation (8). I analyze the univariate relationship between the 5% and 1% quantile of the return distribution and several state variables, including VIX (implied volatility on S&P 500 index options), TED spread (difference between the three-month T-bills and three-month LIBOR), Term spread (difference between the long-term government bond and the T-bill), and default spread

---

[25] As argued in Poon and Granger (2003), this selection of models may lead to an unclear consensus on which model best forecast VaR. Since only a subset of alternative models are compared, with potential bias to the method being developed by the authors, and since there may not be uniform forecast evaluation techniques, it is difficult to reach a clear consensus on which model best forecasts VaR.

[26] The estimation of the GARCH coefficients was performed in MATLAB 7.1.using the Statistical Toolbox and loops to re-optimize the parameters every day.

[27] Although I consider an expanding scheme for ES, in practice, only the most recent observations are taken into account in the one-day-ahead volatility forecast.

[28] All the computations were done using MATLAB 7.1. I use the function *fminsearch* as optimization algorithm. To perform the daily re-estimation of the parameters, β_τ, I use loops to compute the recursive quantile functions.

(difference between BAA and AAA-rated corporate bond yields in US). A detailed description of these variables is presented in Annex 2.

All data used in this Work Project is from Bloomberg. The out-of-sample forecast period is from December 8, 1998 to May 5, 2009 leading to VaR 2365 estimates.

5. Empirical Analysis

I examine the VaR forecasting performance for the S&P 500 from December 8, 1998 to May 5, 2009. Table 1 reports some relevant summary statistics. Returns exhibit negative skewness and excess kurtosis, providing evidence on the violation of the normality assumption, consistent with the stylized facts documented in the literature.

5.1 RQVaR estimation

Figures 1 and 2 present the out-of-sample estimated coefficients for RQVaR using VIX as state variable, at the 95% and 99% confidence level, respectively.[29] Several conclusions can be drawn from the analysis of these figures. Firstly, the estimated model coefficients are not constant during the whole out-of-sample period, responding to changes in market conditions. Both the conditional expected quantile (expressed by the constant term, β_0) and the sensitivity of VaR estimates to the level of VIX verified in the market in the previous day (given by β_1 estimates) vary when there are abnormal market conditions. This is particularly true in periods of crisis and financial distress. For example, in Figures 1 and 2, we can see the rapid adjustments in the estimated parameters that occurred both during the speculative bubble of the "dot-com" in 2001 and in the current financial crisis of 2007-2009. Secondly, we can conclude that estimated coefficients are more stable at the 99%

[29] I do not present the results for the remaining state variables as they present relatively worse out-of-sample VaR forecasts than VIX.

VaR estimation than at the 95% VaR, suggesting that the sensitivity of VaR estimates to changes in the level of VIX change less when lower quantiles of the distribution are forecasted. All coefficients are statistically significant at 5% confidence level, as the estimates lie in the 95% standard error interval constructed with bootstrapping. As a result, out-of-sample forecast can be performed with the estimated coefficients.

5.2 Comparative out-of-sample performance of alternative VaR models

The out-of-sample forecast performance of alternative VaR models is presented in Figures 3-6 and Figures 7-10, corresponding to the 95% and 99% VaR estimates respectively. In general, all models tend to underestimate the frequency of extreme returns, although the performance varies substantially across different models. In fact, none of the methodologies present a number of exceptions consistent with the confidence level used in the VaR estimation. Nonetheless, if we rank models according to the number of exceptions, GARCH and HS outperform RQVaR and ES, both at the 95% and at the 99% confidence level forecasts.[30] At the 95% confidence level, GARCH model generates the lowest exception rate (5.3%) followed by HS (5.9%), while ES and RQVaR exhibit higher exception rates (6.0% and 7.9%, respectively). Similarly, at the 99% confidence level, GARCH and HS outperform ES and RQVaR (1.4% and 1.7% compared with 1.9% and 2.0%). Recalling that exceptions at the 95% confidence level should occur in 5% of the observations, all models present exception rates higher than what it is desirable for accurate models. This is evidence that all methodologies underestimate VaR. The results of the appropriate unconditional and conditional covered tests are examined in Section 5.3.

Importantly, the previous analysis did not take into account the amount of losses when VaR is exceeded, but only how frequently VaR is exceeded. Although the magnitude of

[30] Recall that the correct unconditional exception rate at the 95% confidence level is 5%, and at the 99% is 1%.

exceptions is commonly disregarded in the out-of-sample evaluation of VaR models, a careful analysis of the amount of losses beyond VaR should be of primary interest for risk management (see Ferreira and Lopez (2002), Berkowitz (2001), and Basak and Shapiro (2000) for further discussion). Seriously enough, losses higher that the VaR forecast can result in the collapse of financial institutions as there might not be enough capital to cover such extreme and unexpected losses. Table 2 presents the average size of exceptions of the four models considered in this Work Project. I measure the size of exceptions as the difference between the VaR estimate and the actual loss. Thus, alternative models can produce different magnitude of exceptions.

Evaluating model performance according to the average magnitude of exceptions yields to interesting opposite results: Although RQVaR presents the highest exception rate, it produces the lowest average magnitude of exceptions at the 95% confidence level (the average percentage loss beyond VaR given by RQVaR 0.65%). At the 99% confidence level the average percentage loss beyond VaR is 0.64% and this model is outperformed by ES (0.63%). Both GARCH and ES exhibit similar results to RQVaR. By contrast, HS provides the highest average magnitude of exceptions (0.88% and 0.89%). This implies that although VaR is exceeded 5.9% of the times at the 95% confidence level, when it is exceeded the average difference between VaR estimate and actual loss is 0.88%. Bearing in mind that HS is the most widely used method by U.S. commercial banks in assessing market risk, we can conclude that banks using this approach may be misallocating regulatory and economic capital as a consequence of inaccurate VaR estimates. Seriously enough, these large and unpredictable losses can result in the collapse of financial institutions. In its 2008 annual report, UBS justifies the remarkable failure of its risk management systems by notifying that HS *"does not respond quickly to periods of heightened volatility"* like the current financial crisis.

In this context, an important discussion about the correct techniques for VaR models evaluation arises. Different out-of-sample evaluation methods lead to opposite results.

5.3 Statistical tests on the out-of-sample performance of VaR models

Tables 3 and 4 report the results of the unconditional covered test for VaR estimates at the 95% and 99% confidence level, respectively[31]. At the 95% confidence level, GARCH is the only model that never rejects the null hypothesis correct unconditional probability of exceptions (see Table 3). It is interesting to find that the standard normal distribution generates volatility forecasts that perform well at the 5% quantile, despite the excess kurtosis usually found in returns. At the 99% confidence level, all models reject the null hypothesis (see Table 4). These results again confirm the superior performance of GARCH model in forecasting VaR for stocks

Table 5 reports the results of independence and conditional covered test for the alternative models at the 95% confidence level.[32] None of the approaches is able to produce i.i.d. VaR exceptions, suggesting that an exception today has impact on an exception tomorrow. As a consequence, none of the methodologies is able to successfully capture exception clustering.

Figures 11 and 12 report the results of the dynamic strategy test. None of the models is able to produce accurate forecasts for VaR of the S&P 500. As a result, an investor would not be able to follow a dynamic strategy with constant VaR based on model predictions for the risk of S&P 500. More specifically, and confirming previous results, ES and HS underestimate VaR (especially during the financial crisis period) and suggest investors to take a higher exposure, on average, than what would be optimal for the S&P 500.

---

[31] This test is computed for each 3 years of data (756 observations), i.e. for every for every period of 3 years of data I checked whether the exception rate observed within that period was statistically equal to the significance level (5% and 1%). In the last test we considered 853 observations since the division of the data is not exact.

[32] These tests at the 99% confidence level could not be computed.

5.4 Tradeoff between correct model and estimation error

Two striking differences between the alternative approaches are obvious. First, both RQVaR and GARCH produce VaR forecasts that are consistent with the behavior of financial returns. From Figures 3 and 4 (for the 95% VaR) and Figures 7 and 8 (for 99% VaR), we can conclude that both models are able to capture the systemic risk of the S&P 500 by timing the increases in market volatility. This clearly contrasts with the performance of HS. However, RQVaR and GARCH produce noisier VaR estimates than EWMA and HS, because the first two models involve parameter optimization. The noise in the estimates is due to the sampling error in the estimation of the model coefficients within a given and finite sample. In fact, the coefficients of RQVaR are optimized to minimize the asymmetrically weighted absolute errors in a specific sample. Similarly, GARCH parameters are optimized to maximize the log-likelihood function again in a certain sample. Therefore, the estimated coefficients express not only the fundamental relationship between the variables of interest but also the noise arising from sampling error. When applied for out-of-sample forecasting, the sampling error part is not robust, which leads to in sample overfitting versus poorer out-of-sample performance.

6. Conclusion

I examine the out-of-sample VaR estimates for the S&P 500, using different alternative statistical approaches. Given the current evidence on extraordinary failure of risk management models in capturing systemic risk during 2008, I propose an alternative approach to estimate VaR, RQVaR. This methodology aims both to solve the problem of lack of conditionality in VaR models and to avoid distributional assumptions that can be too unrealistic for certain financial securities held by commercial btanks.

Overall, I conclude that none of the methodologies is able to produce accurate forecasts for the S&P 500, especially in periods of crisis. Moreover, different evaluation techniques yield to different results, which can be problematic in determining capital requirements, from which the solvency of financial institutions depend on. The most striking results concern HS and RQVaR. On one hand, HS presents one of the lowest exception rates but it produces the worst fit of the returns' behavior (see Figures 6 and 10). On the other hand, RQVaR exhibits the highest exception rate during the out-of-sample period, but is able to successfully capture the movements in returns over time.

Moreover, this Work Project provides evidence on the superior performance of the GARCH model relatively to all models in terms of exception rates and the fitting of returns' behavior. This is consistent with the findings of Berkowitz and O'Brien (2002), Pérignon and Smith (2007), and Pérignon, Deng and Wang (2007). However, it is important to emphasize that this empirical study I use the S&P 500 to evaluate the accuracy of alternative VaR models. As noted by Pérignon and Smith (2007), concluding that a given VaR model outperforms the others is only useful if the data used in the horse race (in this case the S&P 500) is closely related with banks' trading portfolios. Actually, according to Berkowitz and O'Brien (2007) bank's trading positions are (i) complex and affected by non-standard risk factors, (ii) frequently rebalanced, and (iii) very different across banks. Thus, a GARCH model applied to trading positions of commercial banks might be less useful than shown in this Work Project. In fact, Figlewski (1997) demonstrates that GARCH models when applied to other securities (rather than stocks) for short horizons have less powerful forecasting power as the normality assumption for return innovation is to unrealistic to produce consistent estimates.

In this context, since RQVaR does not require any distributional assumptions to estimate conditional time-varying quantiles of future return distribution, it might be of some interest to investigate the out-of-sample performance of this approach in estimating the VaR of more complex and badly-behaved securities that constitute banks' trading portfolios. To improve the out-of-sample forecasts of RQVaR, future research may consider the introduction of an adaptive component that adjusts for model exceptions Suggestions of adaptive components can be found in Engle and Managanelli (2004), who propose the use of an autoregressive component or, alternatively, an indicator function that increases VaR forecast when VaR is exceeded and decrease it slightly otherwise.

REFERENCES

Acerbi, Carlo, and Dirk Tasche. 2002. "Expected Shortfall: a natural coherent alternative to Value at Risk". *Economic Notes*, 31: 379-388.

Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber, and David Heath. 1999. "Coherent measures of risk". *Mathematical Finance*, 9(3):203–228.

Andersen, Torben G., Tlim Bollerslev, Peter F. Christoffersen and Francis X. Diebold. 2007. "Practical Volatility and Correlation Modeling for Financial Market Risk Management" in M. Carey and R. Stulz (eds.), *The Risks of Financial Institutions,* University of Chicago for NBER 513-548.

Baillie, Richard, and Tim Bollerslev. 2000. "The forward premium anomaly is not as bad as you think". *Journal of International Money and Finance* 19, 471-488.

Barone-Adesi, Giovanni, Frederick Bourgoin, and Kostas Giannopoulos. 1998. "Don't look back". *Risk*, 11 (August): 100–104.

Barone-Adesi, Giovanni, Kostas Giannopoulos and Les Vosper. 1999. "VaR without correlations for nonlinear portfollios". *Journal of Futures Markets,* 19: 583–602.

Barone-Adesi, Giovanni, Kostas Giannopoulos and Les Vosper. 2002. "Backtesting Derivative Portfolios with Filtered Historical Simulation (FHS)". *European Financial Management*, 8: 31-58.

Bao, Yong, Tae-Hwy Lee and Burak Saltoglu. 2004. "A Test for Density Forecast Comparison with Applications to Risk Management". Technical report, University of California, Riverside.

Basak, Suleymand, and Alexander Shapiro. 2000. "Value-at-Risk based risk management: Optimal policies and asset prices". *Review of Financial Studies,* 14:371-405.

Basle Committee on Banking Supervision. 1996. "Overview of the Amendment to the Capital Accord to Incorporate Market Risks." Available at http://www.bis.org

Berkowitz, Jeremy. 2001. "Testing density forecasts with applications to risk management". *Journal of Business and Economic Statistics*, 19: 465–474.

Berkowitz, Jeremy, and James O'Brien. 2002. "How Accurate are Value-at-Risk Models at Commercial Banks?" *Journal of Finance*, 57: 1093–1111.

Berkowitz, Jeremy, and James O'Brien. 2007. "Estimating Bank Trading Risk: A Factor Model Approach" in M. Carey and R. Stulz (eds.), *The Risks of Financial Institutions,* University of Chicago for NBER 59-91.

Berkowitz, Jeremy, Peter Christoffersen and Denis Pelletier. 2007. "Evaluating Value-at-risk models with Desk-Level Data". Working Paper, McGill University.

Bollerslev, Tim. 1986. "Generalized autoregressive conditional heteroskedasticity". *Journal of Econometrics*, 31(3): 307-327.

Bollerslev, Tim, and Jeffrey Wooldridge. 1992. "Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time-Varying Covariances". *Econometric Reviews* 11: 143-172.

Boudoukh, Jacob, Matthew Richardson and Robert Whitelaw. 1998. "The best of both worlds". *Risk,* 11 (May): 64–67.

Campbell, Alexander. 2009. "The Risk of Value-at-Risk". *Risk*, 22(4):42-46.

Chernozhukoy, Victor, and Len Umantsev. 2001. "Conditional Value-at-Risk: Aspects of Modeling and Estimation." *Empirical Economics* 26:271–292.

Christoffersen, Peter F. 1998. "Evaluating Interval Forecasts." *International Economic Review,* 39: 841–862.

Christoffersen, Peter F. 2003. *Elements of Financial Risk Management.* Amsterdam: Academic Press.

Connor, Gregory. 1997. "Sensible return forecasting for portfolio management." *Financial Analyst Journal* 53, 44—51.

Danielsson, Jon, and Yuji Morimoto. 2000. "Forecasting Extreme Financial Risk: A Critical Analysis of Practical Methods for the Japanese Market." *Monetary and Economic Studies* 18(2):25–48.

Dowd, Kevin. 2002. *Measuring Market Risk*. Chichester: Jonh Wiley & Sons.

Duffie, Darrel, and Jun Pan, 1997. "An overview of Value-at-Risk" Journal of Derivatives 4:7-49

Engle, Robert F., and Simone Manganelli. 2004a. "CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles." *Journal of Business and Economic Statistics* 22:367–381.

Engle, Robert F., and Simone Manganelli. 2004b. "A Comparison of Value-at-Risk Models in Finance" in G. Szego (ed.), *Risk Measures for the 21st Century*, 123-144. Chichester: Jonh Wiley & Sons.

Federal Register. 1996. "Risk-Based Capital Standards: Market Risks." Vol. 61: 47357-78.

Ferreira, Miguel, and José Lopez. 2005. "Evaluating Interest Rate Covariance Models Within a Value-at-Risk Framework". Journal of Financial Econometrics, 3(1):126-168.

Finger, Christopher. 2009. "VAR Is From Mars, Capital Is From Mars". *Research Monthly*, April, RiskMetrics Group.

Figlewski, Stephen. 1997. "Forecasting Volatility." *Financial Markets, Institutions and Instruments* 6: 1–88.

Fleming, Jeff, Chris and Barbara Ostdiek. 2001. "The economic value of volatility timing". *Journal of Finance*, 56: 329-352.

Foster, Dean P., and Daniel B. Nelson. 1996. "Continuous record asymptotics for rolling sample variance estimators". *Econometrica*, 64: 139-174.

Granger, Clive W. J. 2002. "Some comments on risk". *Journal of Applied Econometrics*, 17: 447-456.

Harvey, Campbel, and Akhtar Siddique. 1999. "Autoregressive conditional skewness". *Journal of Financial and Quantitative Analysis*, 34: 465-488.

Hansen, Peter R., and Asger Lunde. 2005. "A forecast comparison of volatility models: does anything beat a GARCH (1,1)?". *Journal of Applied Econometrics,* 20(7):873-889.

Hull, John, and Alan White. 1998. "Incorporating volatility updating into the historical simulation method for Value-at-Risk". *Journal of Risk*, 1 (1): 5–19.

J.P. Morgan. 1995. Riskmetrics Technical Manual, 3rd ed.

Jackson, Patricia, David Maude and William Perraudin. 1997. "Bank Capital and Value at Risk". *Journal of Derivatives*, 4:73-89.

Jondeau, Eric, and Michael Rockinger. 2003. "Conditional volatility, skewness, and kurtosis: existence, persistence, and comovements". *Journal of Economic Dynamics and Control*, 27: 1699-1737.

Jorion, Philippe. 2006. *Value at Risk: The new benchmark for managing financial risk*, 3rd ed. New York: McGraw Hill.

Koenker, Roger, and Gilbert Bassett. 1978. "Regression quantiles". *Econometrica*, 46: 33-50.

Kuester, Keith, Stefan Mittnik and Marc S. Paolella. 2006. "Value-at-Risk Prediction: A comparison of alternative strategies". *Journal of Financial Econometrics*, 4 (1): 53-89.

Kupiec, Paul. 1995. "Techniques for Verifying the Accuracy of Risk Measurement Models." *Journal of Derivatives,* 3: 73–84.

Mandelbrot, Benoit. 1963 "The Variation of Certain Speculative Prices", *Journal of Business* 36, 393-413.

McNeil, Alexander J., and Rüdiger Frey. 2000. "Estimation of tail-related risk measures for heteroskedastic financial time series: an extreme value approach". *Journal of Empirical Finance*, 7: 271-300.

Mittnik, Stefan, and Marc Paolella. 2003. "Prediction of financial downside risk with heavy tailed conditional distributions". In S. T. Rachev (ed.), *Handbook of Heavy Tailed Distributions in Finance*. Amsterdam: North-Holland.

Pérignon, Christophe, and Daniel R. Smith. 2007. "Which Value-at-Risk Method Works Best for Ban Trading Revenues?". Working Paper, Simon Fraser University.

Pérignon, Christophe, and Daniel R. Smith. 2009. "The Level and Quality of Value-at-Risk Disclosure by Commercial Banks". Working Paper, Simon Fraser University.

Pérignon, Christophe, Zi Yin Deng and Zhi Jun Wang. 2008. "Do Banks Overstate their Value-at-Risk?" *Journal of Banking and Finance*, 32(5): 783-794.

Pflug ,Georg. 2000. "Some remarks on the value-at-risk and the conditional value-at-risk" in S. Uryasev (ed.), *Probabilistic Constrained Optimization: Methodology and Applications*. Dordrecht: Kluwer Academic Publishers.

Pritsker, Matthew. 2006. "The hidden dangers of historical simulation". *Journal of Banking and Finance*, 30: 561-582.

Poon, Ser-Huang, and Clive W. J. Granger. 2003. "Forecasting volatility in financial markets: a review". *Journal of Economic Literature*, 41: 478-539.

Schwertz, G. William. 1989. "Why does stock market volatility changes over time?". *Journal of Finance,* 44, 1115-1153.

Taylor, James. 1999. "A quantile approach to estimating the distribution of multi-period returns". Journal of Derivatives 6: 64-78

Taylor, James. 2008a. "Estimating Value at Risk and Expected Shortfall Using Expectiles". *Journal of Financial Econometrics* 6: 231-252.

Taylor, James. 2008b. "Using Exponentially Weighted Quantile Regression to Estimate Value at Risk and Expected Shortfall". *Journal of Financial Econometrics* 6: 382-406.

APPENDICES

Annex 1 – Demonstration of the dynamic strategy test

Consider a portfolio constituted by a risky asset and the risk-free rate. The return on this portfolio, r_p, is given by equation (16):

$$r_p = w_1 r_{risky} + (1 - w_1) r_f \tag{18}$$

where r_{risky} denotes the return on the risky asset, w_1 the weigth invested in the risky asset and r_f the risk-free rate.

The VaR of this portfolio is given by:

$$VaR_p = w_1 VaR_{risky} + (1 - w_1) VaR_{r_f} \tag{19}$$

Since we have a position of $(1 - w_1)$ on the risk-free asset, there is no market risk exposure in this position, as:

$$VaR_{r_f} = 0 \tag{20}$$

Thus, the VaR of the portfolio can be reduced to the market risk on the risky asset:

$$VaR_p = w_1 VaR_{risky} \tag{21}$$

Consider that an investor aims to keep the VaR of the portfolio fixed at a given level, α, over time:

$$VaR_p = \alpha \tag{22}$$

As a consequence, the investor that is able to accurately forecast the VaR of the risky asset using a given model m, he is able to adjust the positions on the risky asset and on the risk-free rate, as described below:

$$\alpha = w_1 \widehat{VaR}_{risky}^{m} \quad \Leftrightarrow w_1 = \frac{\alpha}{\widehat{VaR}_{risky}^{m}} \tag{23}$$

Annex 2 – Detailed description of the data

All data used in this Work Project is downloadable from Bloomberg. Tickers are in brackets:

S&P 500: It comprises the *(SPX index)*

VIX: 30-day ahead implied volatility of S&P 500 index options. (*VIX Index*)

TED spread: Difference between the 3-month T-bills and three-month LIBOR (London Interbank Offered Rate). (*.TEDSPRD Index*)

Term spread: Difference between the 10-year government bond and the 3-month T-bill. *(USGG10YR Index and USGG3M Index)*

Default spread: Difference between BAA and AAA-rated 10-year corporate bond yields in US *(MOODCBAA Index and MOODCAAA Index)*. Both indices are constructed by Moody's.

Table 1 - Summary statistics for the S&P 500 returns for the out-of-sample period

Sample size	Mean	Std. Dev.	Skewness	Kurtosis	Min.	Max.
4875	0,00019	0,01172	-0,19533	9,48882	-0,09470	0,10957

Figure 1 – Estimated coefficients for 95% VaR using RQVaR

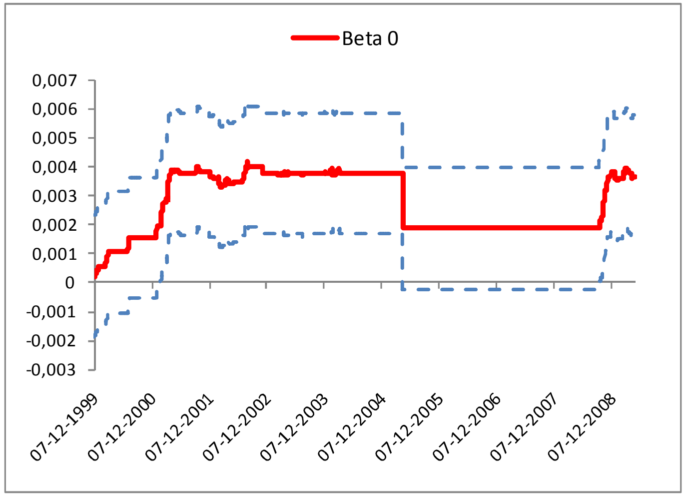


Figure 1A- Constant term, β_0

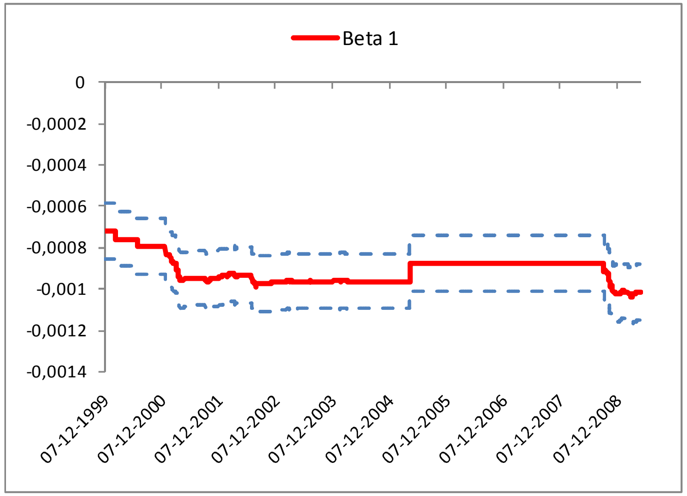


Figure 1B- Sensitivity of VaR estimate to VIX, β_1

The upper and lower bands on both figures represent a 95% confidence interval on the coefficients calculated with bootstrapping.

Figure 2 – Estimated coefficients for 99% VaR using RQVaR

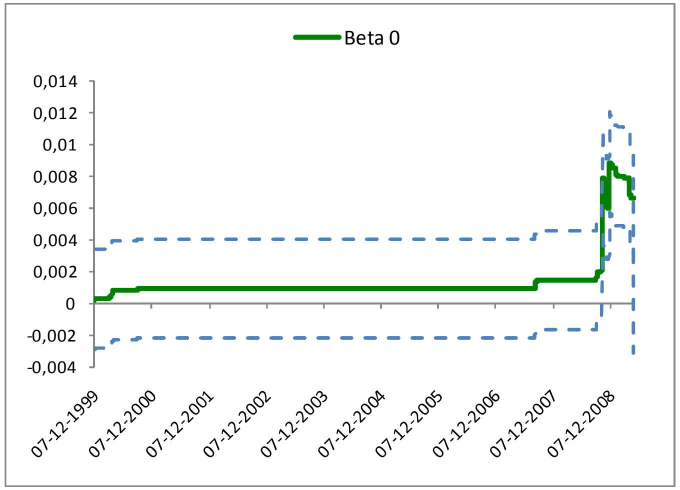


Figure 2A- Constant term, β_0

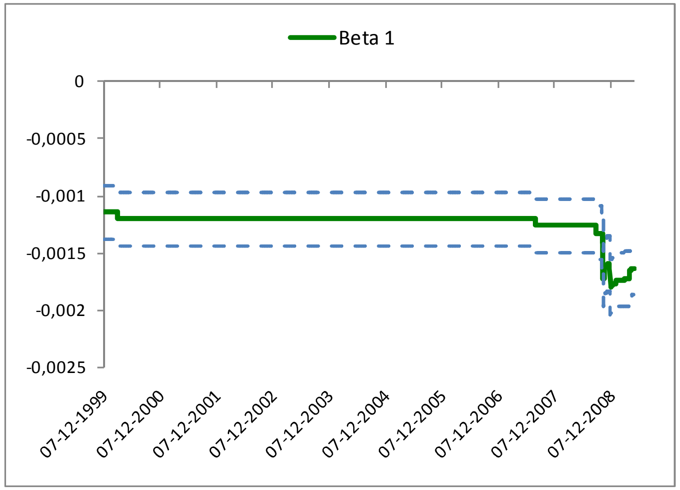
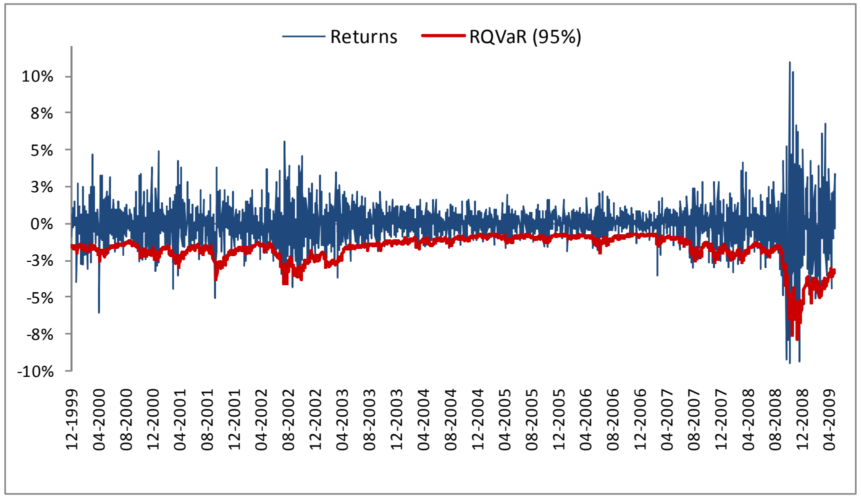


Figure 2B- Sensitivity of VaR estimate to VIX, β_1

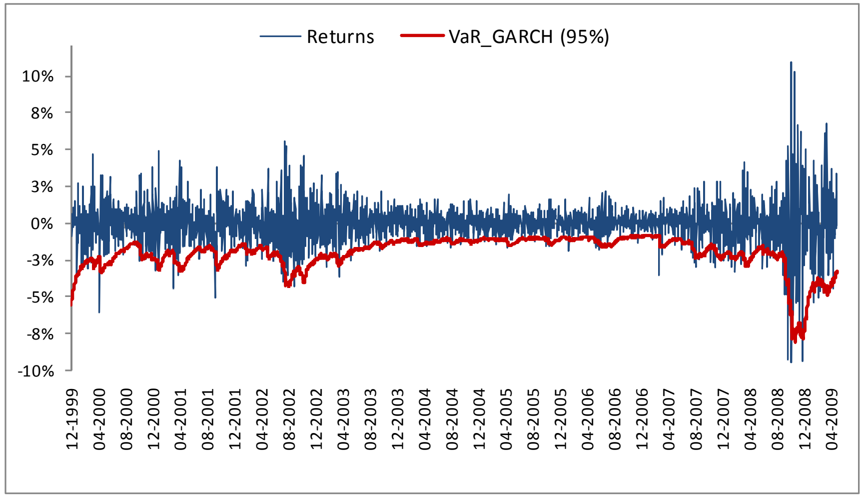
The upper and lower bands on both figures represent a 95% confidence interval on the coefficients calculated with bootstrapping.

Out-of-sample 95% VaR forecasts and realized returns from December 8th, 1999 to May 5th, 2009



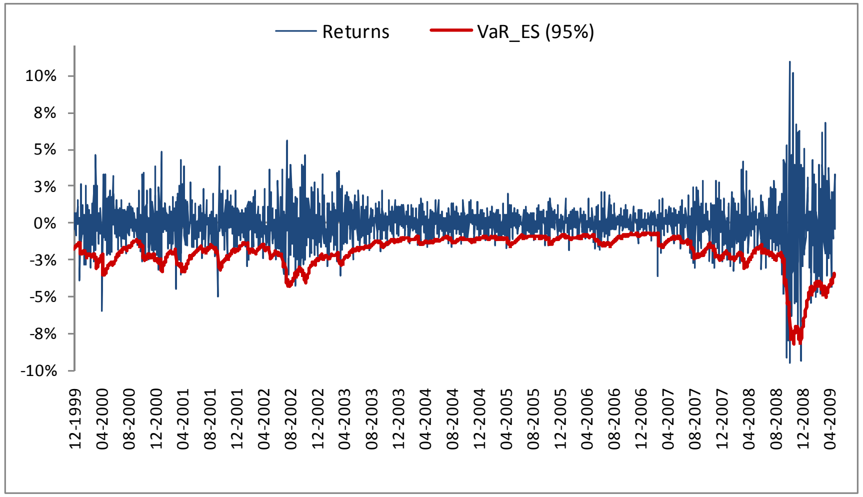
2365 observations; 187 exceptions; Exception rate of 7.9%

Figure 3 – Daily S&P 500 and 95% VaR forecasted with RQVaR, using VIX as state variable



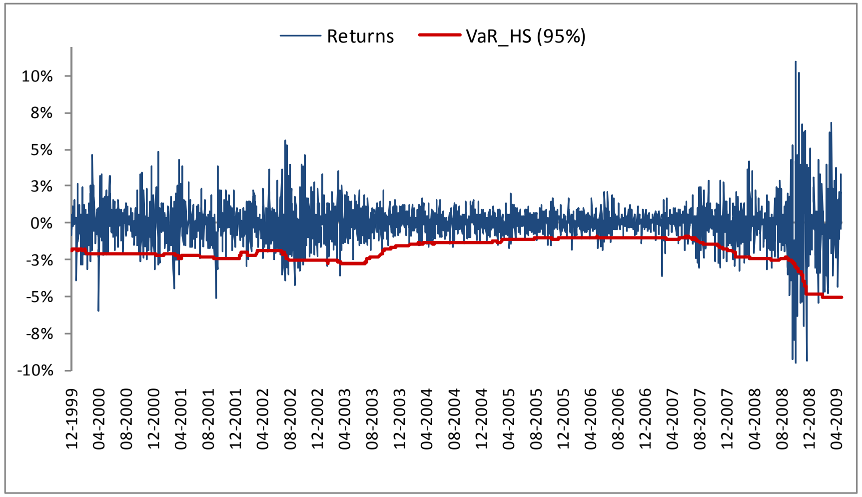
2365 observations; 126 exceptions; Exception rate of 5.3%

Figure 4 – Daily S&P 500 and 95% VaR forecasted with GARCH



2365 observations; 142 exceptions; Exception rate of 6.0%

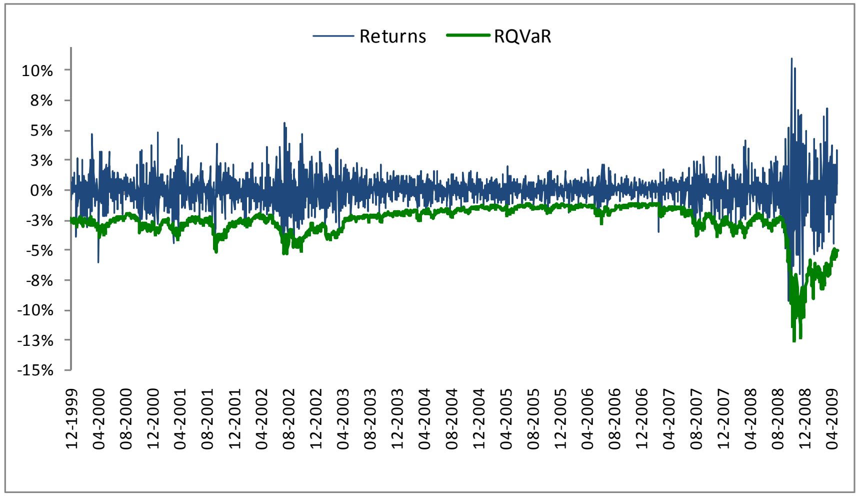
Figure 5 – Daily S&P 500 and 95% VaR forecasted with ES



2365 observations; 140 exceptions; Exception rate of 5.9%

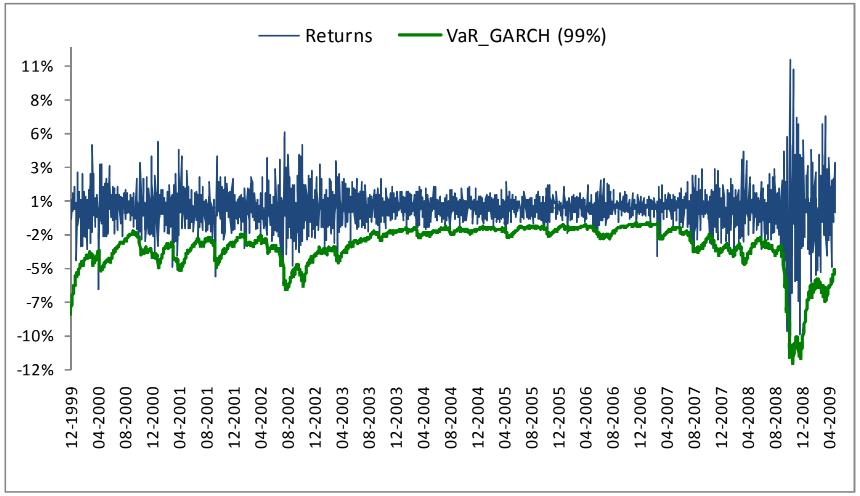
Figure 6 – Daily S&P 500 and 95% VaR forecasted with HS

Out-of-sample 99% VaR forecasts and realized returns from December 8th, 1999 to May 5th, 2009



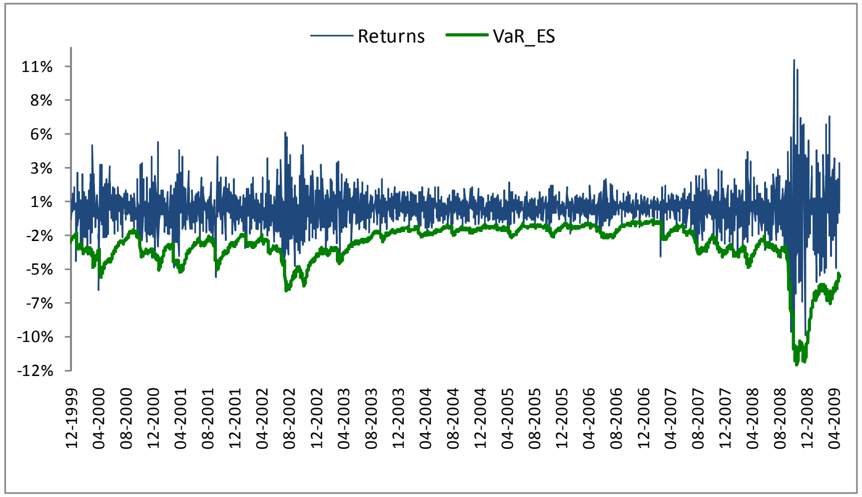
2365 observations; 48 exceptions; Exception rate of 2.0%

Figure 7 – Daily S&P 500 and 99% VaR forecasted with RQVaR, using VIX as state variable



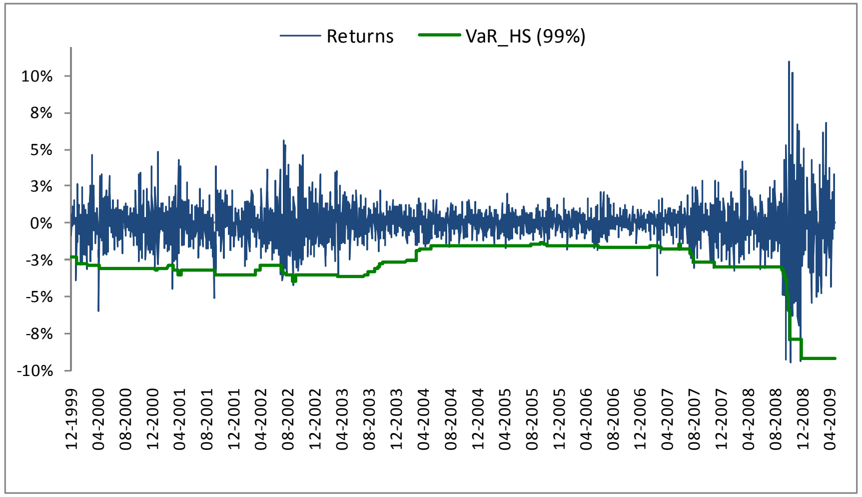
2365 observations; 33 exceptions; Exception rate of 1.4%

Figure 8 – Daily S&P 500 and 99% VaR forecasted with GARCH



2365 observations; 44 exceptions; Exception rate of 1.9%

Figure 9 – Daily S&P 500 and 99% VaR forecasted with ES



2365 observations; 41 exceptions; Exception rate of 1.7%

Figure 10 – Daily S&P 500 and 99% VaR forecasted with HS

Table 2 - Average size of exceptions of alternative VaR models

	HS	GARCH	ES	RQVaR
95% VaR	-0,0088	-0,0067	-0,0064	-0,0065
99% VaR	-0,0089	-0,0068	-0,0063	-0,0064

Table 3 – Unconditional covered test (95% VaR estimates)

	RQVaR	GARCH	ES	HS
Nr. Exceptions (3y)	66 48 73	44 29 53	46 36 60	49 21 70
Exception Rate (3y)	8,73% 6,35% 8,56%	5,82% 3,84% 7,01%	6,08% 4,76% 7,03%	6,48% 2,78% 8,21%
Unc. Covered Test Statistic	18,29 2,68 18,92	1,02 2,34 5,75	1,76 0,09 6,63	3,21 9,30 15,60
Critical value (5%)	3,84	3,84	3,84	3,84
Result	Reject Ho Don't Reject Ho Reject Ho	Don't Reject Ho Don't Reject Ho Reject Ho	Don't Reject Ho Don't Reject Ho Reject Ho	Don't Reject Ho Reject Ho Reject Ho

Table 4 – Unconditional covered test (99% VaR estimates)

	QR	GARCH	ES	HS
Nr. Exceptions (3y)	15 5 28	7 4 22	10 9 25	12 5 24
Exception Rate (3y)	1,98% 0,66% 3,70%	0,93% 0,53% 2,91%	1,32% 1,19% 2,93%	1,59% 0,66% 2,81%
Unc. Covered Test Statistic	18,59 46,85 2,93	39,29 51,20 8,13	30,07 32,91 8,97	24,98 46,85 10,13
Critical value (1%)	3,84	3,84	3,84	3,84
Result	Reject Ho Reject Ho Don't Reject Ho	Reject Ho Reject Ho Reject Ho	Reject Ho Reject Ho Reject Ho	Reject Ho Reject Ho Reject Ho

The tests were performed for block of 3 years of data (756 observations). The last period includes more observations because there were observations left;
Ho: Exception rate (3y) = 5%

Table 5 - Independence and conditional covered test (95% VaR estimates)

	RQVaR	GARCH	ES	HS
Independence test	107,29	73,54	76,34	80,29
	79,17	51,61	62,25	38,81
	119,99	96,81	108,31	109,66
Conditional Covered test	125,58	74,56	78,10	83,49
	81,85	53,95	62,35	48,11
	138,91	102,56	114,94	125,25
Critical value (5%)	5,99	5,99	5,99	5,99
Result	Reject Ho	Reject Ho	Reject Ho	Reject Ho
	Reject Ho	Reject Ho	Reject Ho	Reject Ho
	Reject Ho	Reject Ho	Reject Ho	Reject Ho

The tests were performed for block of 3 years of data (756 observations). The last period includes more observations because there were observations left; Ho: Serial independence of exceptions, and Exception rate (3y) = 5%

Figure 11 - Dynamic strategy test (95% VaR)

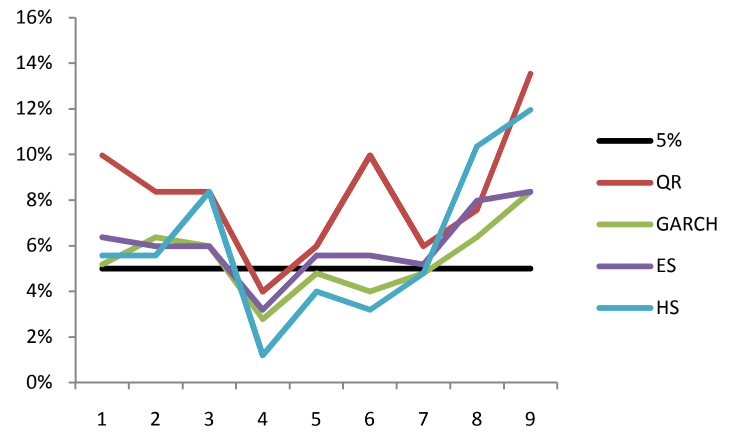


Figure 12 - Dynamic strategy test (99% VaR)