The Historical Simulation Method for Value-at-Risk: A Research Based Evaluation of the Industry Favorite

Article *in* SSRN Electronic Journal · April 2012

DOI: 10.2139/ssrn.2042594

Meera Sharma

The Historical Simulation Method for Value-at-Risk:

A Research Based Evaluation of the Industry Favourite

Abstract

This paper surveys the literature relating to the historical simulation method of calculating VaR. The historical simulation method is the most popular method for VaR calculation in the banking industry. In India also banks are in the process of implementing VaR methods for market risk calculation.

Thirty Eight papers are surveyed to understand the performance measures for VaR methods and the comparative performance of Historical Simulation VaR methods. The performance measures are broadly divided into unconditional coverage and conditional coverage measures. While regulatory requirements are limited to unconditional coverage measures, conditional coverage measures have been developed to spot the phenomenon of exception clustering. The performance of historical simulation - the basic and modified methods - in comparison with other methods is surveyed through available studies.

The historical simulation approach is found to provide superior unconditional coverage among a wide variety of alternate methods ranging from the simple variance covariance approach to the sophisticated GARCH, explaining its popularity in the industry. This superiority translates into lesser likelihood of regulatory penalties since the regulatory back testing framework is based on unconditional coverage.

The advantage of superior performance by historical simulation is lost when it is measured on conditional coverage measures (joint tests of unconditional coverage and independence). However, the sophisticated conditional volatility models are not much better than the historical simulation in conditional coverage. A modification to the historical simulation method, the filtered historical simulation method emerges as the best performer using conditional coverage criteria.

This study has an important contribution to make to available research. First it compiles the performance measures of VaR methods. Second it surveys the modifications to the historical simulation approach. Third and most important it presents a comparative picture of the most popular approach vis a vis other methods on a variety of performance parameters.

Key Words: Historical Simulation, Conditional coverage, Market risk, Unconditional coverage, Value at Risk.

The Historical Simulation Method for Value-at-Risk:

A Research Based Evaluation of the Industry Favourite

This paper surveys the research available on the historical simulation method to calculate VaR. The second and third Basel Accord (International Convergence of Capital Measurement and Capital Standards, 2006 and Revisions to the Basel II Market Risk Framework, 2009) have laid down market risk capital requirements for trading books of banks. The market risk capital calculations can be done using either the standardized measurement method or the Internal Models approach. The internal models approach allows banks to calculate a market risk charge based on the output of their internal Value-at-risk models.

The migration of banks to the sophisticated internal models approach from the standardized approach is a desired outcome from the point of view of better risk management practices. A survey on Indian banks (Transition to the Internal Models Approach for Market Risk – a Survey Report, 2010) reported that while none of the Indian banks have adopted the internal market approach for market risk capital calculation, all of those surveyed either have or are in the process of building a VaR system for internal purposes. Around 25% of the banks surveyed were planning to migrate to the internal model approach before December 2010, 38% had set end of 2011 as the deadline. The banks surveyed reported that methodology selection was the key challenge in implementation of the internal markets approach.

Miura and Oue (2001) present a useful classification scheme for VaR methodologies along dual dimensions of distributional assumptions and dependence assumptions. The result is a 3 by 2 matrix with normal, non normal and non parametric along one dimension and i.i.d. and time dependence along the other dimension. The historical simulation approach (in its basic form) is a non parametric, i.i.d. approach that uses the empirical distribution of past returns to generate a VaR. The i.i.d., normal model is also called the variance-covariance (VC) approach or equally weighted moving average method (EQMA) when rolling windows are used. Methods that incorporate time dependence are the exponentially weighted moving average model (EXMA) and the GARCH method. Each of these methods can be used with a normal or non normal distributional assumption. Kalvyas et al (2004) outline the methodologies for the popular approaches. Manganelli and Engle (2001) review the assumptions behind the various methods and discuss the theoretical flaws of each.

The historical simulation approach has emerged as the most popular method for Value-at-risk calculation in the industry. Perignon and Smith (2006) survey the VaR disclosures of a cross section of 60 US, Canadian and large international banks over 1996-2005 and report that 73 percent of banks that disclosed their VaR methodology used historical simulation. The second most popular was Monte Carlo (MC) simulation (14%). A survey report on Indian banks (Transition to the Internal Models Approach for Market Risk – a Survey Report, 2010) states that of the 30 banks that participated in the survey 67% of the banks used historical ‘methods, 5% used Monte Carlo (MC) simulation and 3% used variance-covariance models.

The widespread popularity of historical simulation warrants attention. More so because researchers have ignored this humble method to focus on more sophisticated ones, for example the ARCH/GARCH methods of calculating conditional volatility for VaR.

This paper surveys the research available on the historical simulation method for VaR calculation. Thirty Eight papers are surveyed. The papers are selected on the basis of their citations in available research on VaR and on the rigor of their analysis - all papers use more than ten year lengths of daily data except Boudoukh et al (1998) which is cited for its contribution to methodology. Since an important focus area for this paper is the comparative performance of VaR methods, long historical periods are needed to conduct thorough back tests.

The research paper is organized to throw light on three questions. The first question examined is the characteristics of the basic HS and modifications of the basic method. The second question addressed is the unconditional coverage performance of historical simulation method or its modifications in comparison with other methods. The third question deals with the conditional performance of the method or modified methods in comparison with other methods. The remaining part of this paper is organized around these three discussions.

I. Alterations to the Historical Simulation VaR Method

One advantage of historical simulation (HS) is that since it does not use distributional assumptions to model the VaR it performs better particularly in comparison with methods that use the normal distribution assumption since financial returns data are thick tailed. It is also easier to compute and can incorporate correlations among assets empirically. A problem with HS is that owing to discreteness of extreme returns and very few observations in tails the VaR measures are expected to be highly volatile and erratic. Danielsson and De Vries (1997) observe that the under / over prediction of VaR by HS is more severe in case of an individual stock than an index. The underlying assumption in a HS being that returns are iid is another problem with the approach, making VaR estimates unresponsive to recent innovations in volatility. Modifications to the HS have aimed at bettering the problem of discreteness of extreme returns and the low responsiveness to recent volatility.

Boudoukh et al (1998) modify the historical simulation approach by assigning exponentially declining weights (as in EXMA) to the most recent observations. This method has been referred to as BRW in this paper. Hull and White (1998) improve the historical simulation method by altering it to incorporate volatility updating. They adjust the returns in the historical sample with the ratio of the current daily volatility to the historical volatility, both estimated using a conditional volatility model like GARCH or EXMA (Hull and White (1998) use EXMA with λ=0.94). This altered method (modifying HS with GARCH) called the filtered historical simulation (FHS) by Barone-Adesi et al (1999), effectively makes the HS more responsive to current data. The proponents of this approach claim it requires lesser data compared to the HS.

Barone-Adesi et al (1999) and Barone-Adesi et al (2000) demonstrate the application of Filtered Historical Simulation to portfolios with changing weights and derivatives. FHS deals with the problems associated with modelling asset return correlations particularly when using multivariate econometric models of conditional volatility like GARCH. The correlation structure is modelled empirically by preserving the dependencies of multiple asset returns through a sample of multiple returns on a given day.

Butler and Schachter (1998) use a Gaussian kernel to attach a normal pdf to the points of a return data set. The motivation is to modify the historical simulation method so that it can be implemented with short data history which results in a more responsive VaR measure. The procedure is applied to actual trading data of a financial institution using a data set of 100 trading days. The kernel based estimate is not significantly different from the usual HS The authors also compare the modified HS with a parametric (normal) VaR and again find no significant differences in the VaR estimates of the methods. The advantage of this method is that standard errors of the VaR estimates are also generated, enabling the precision of the VaR to be calculated.

II. Unconditional Performance Measures for VaR

Performance measures of VaR estimates can be categorized broadly into measures of unconditional coverage and conditional coverage. Unconditional coverage measures the number of exceptions to VaR in comparison with the confidence level of the VaR method. For example a 99% VaR method should result in 2.5 (1%) exceptions over a back testing period of 250 trading days. Higher number of exceptions indicate inadequate coverage while a lower number indicate excess coverage (a conservative VaR figure). This is the most commonly used performance measure for VaR methods. Conditional coverage measures the independence of the exceptions, exception clustering being undesirable. This section surveys the unconditional performance measures of VaR. Conditional coverage measures are surveyed in section IV of this paper.

a) Performance measures based on hypothesis tests

The importance of unconditional coverage derives from the fact that the back testing approach of the Basel committee relies on it to judge the performance of a VaR model (Basel Committee for banking Supervision, 1996). Often called the 'traffic light' approach it divides the number of exceptions that a model reports into green, yellow and red zones based on a judgmental trade-off between type I and type II error assuming that the probability of observing an exception is binomially distributed. A model in the green zone is acceptable to supervisors while the other two zones result in penalties imposed on the bank.

Based on the same distributional assumptions as the Basel back test is the likelihood ratio (LR_{uc}) statistic detailed by Christoffersen (1998) described along with other unconditional coverage measures by Campbell (2007). The LR_{uc} statistic follows a χ^2 distribution asymptotically. Campbell (2007) suggests the Pearson's Q test be used to measure the unconditional coverage performance of a VaR method over multiple confidence levels.

b) Performance measures based on forecast accuracy

A number of measures have been developed using the average number of exceptions over a period of time. The bias of the unconditional coverage, calculated using the average number of exceptions over a period of time, is the basic measure. Campbell (2007) defines a 'z' statistic based on the average number of exceptions which follows an approximately standard normal distribution in finite samples.

Bias is defined differently by Hendricks (1996) who calls the average distance of VaR estimates of different models from the average VaR of all models on a given day the mean relative bias. Hendricks (1996) has also used root mean squared relative bias, annualized percentage volatility of VaR, average multiple of tail event to VaR, and maximum multiple of tail event to VaR.

Pritsker (1996) show how non-parametric confidence intervals can be calculated around VaR measures. Pascual, Romo and Ruiz (2001) also construct confidence intervals for VaR estimates and find that confidence intervals constructed around the historical simulation method are ex ante, on average, very wide and have low effective coverage in comparison to the promised confidence level. This is attributed to the HS being unable to capture the persistence in volatility.

c) Performance measures based on magnitude of exceptions - economic loss functions

Jackson et al (1997) use a measure called capital shortfall (capital required by VaR estimate less actual loss sustained) to measure the ability of VaR methods to predict large, `spike' losses in portfolio values. This measure is used to compare the performance of different VaR methods in high volatility episodes. Related to the concept of capital shortfall is the quadratic loss function introduced by Lopez (1999). Unlike the capital shortfall it is modelled on squared distance between the observed returns and the forecasted VaR values if an exception occurs. Lopez (1999) finds that the loss function has higher power than the LR_{uc} statistic in classifying a variety of VaR methods, including HS.

Sarma et al (2003) supplement Lopez's (1999) loss by adding a cost of capital to the excess capital maintained in case no exception occurs.

III. Unconditional Coverage Performance of the Historical Simulation Method for VaR

The literature on comparison of the historical simulation method with others can be categorized on the basis of the measure of performance and the methods compared. Early literature focused on the traditional methods of EQMA, EXMA, HS and MC and used the unconditional coverage rates (defined as fraction of exceptions) as performance measures. These studies span the currency, equity and bond markets. All the studies find that the HS provides better coverage than the other methods. This could be on account of better modeling of fat tails in the data. In particular, the advantage that the EXMA has - of time varying volatility - does not translate into better coverage owing to the use of normality assumption to calculate VaR. These studies are summarized in sub section a) below. A comparison of the performance of HS with more advanced methods like the GARCH and EVT (extreme value theory) based methods is summarized in sub section b). The GARCH with a normal distributional assumption (GARCH-N) applied to the residuals does not perform as well as the HS in comparisons of unconditional coverage rates. The performance of the GARCH with a student's t distributional assumption (GARCH-t) performs better than the GARCH-N. The performance of the EVT method reported by researchers is as good or better than the HS.

Studies discussed in sub sections a) and b) only compare the fraction of exceptions with the VaR percentile, without conducting a rigorous hypothesis testing of the adequacy of unconditional coverage. Studies that do test for the null of correct unconditional coverage are summarized in sub section c). The record of the traditional methods in hypothesis tests is that VC, MC, EXMA and GARCH-N appear to fail these tests. The HS has a mixed performance with some studies reporting that the null of correct unconditional coverage cannot be rejected.

a) Comparison of HS with basic methods (EQMA, EXMA, MC)

Hendricks (1996) compared twelve different VaR methods, namely equally weighted moving average (EQMA) (with 5 lengths of historical observation periods), exponentially weighted moving average (EXMA) (with three values of the decay factor), and historical simulation (HS) (with 4 lengths of historical observation period) for 1,000 randomly chosen foreign exchange portfolios over the period 1983-94. For the 99% VaR it was observed that the historical simulation approach (except with a 125 day short observation period) provided better coverage than the other two VaR methods. The historical simulation VaR with 1250 was the best performer of all methods. However, a caveat to the study by Vlaar (2000) by noting that the period under study was a period of declining volatility, implying that VaR measures using longer periods were higher, resulting in better coverage. In contrast to Hendricks (1996), Hoppe (1998) reports that shorter observation periods (as short as 30 days) provide VaR measures with better unconditional coverage.

Jackson et al (1997) calculate VaRs using different methods for 79 different return series for an actual bank's trading portfolio (with fixed income, foreign exchange and equity security positions). They find that the historical simulation method with long data windows (12 to 24 months) consistently outperforms the parametric methods that use the normality distribution (EQMA with window lengths of 3 to 24 months, EXMA with λ = 0.94 and 0.97) for 99% VaRs in providing better unconditional coverage. This result is reinforced by their comparison of 'capital shortfall' generated by different VaR methods on days of large 'spike' losses. The HS reports a capital surplus on more number of days (double the number of parametric models) and the magnitude of the surplus is higher in majority of instances.

Vlaar (2000) uses 17 years of data for portfolios of Dutch government bonds to compare 99% VaR over a 10 day horizon for HS, VC and MC methods. Like Hendricks (1996) he also finds that the best performing method is the historical simulation with a 1250 day historical observation period. He attributes this partly to the declining volatility of interest rates over the period which results in higher VaR and lower exceptions.

Bias is defined by Hendricks (1996) as the average distance of VaR estimates of different models from the average VaR of all models on a given day the mean relative bias. He reports that the mean relative bias of historical models is more than that of the other approaches for 99% VaR measures. This difference is not discernible for 95% VaR.

b) Comparison of HS with advanced methods (GARCH, EVT)

The HS continues to show better unconditional coverage when compared with sophisticated methods like GARCH. This appears to be due to the assumption of normality used in calculation of VaR from GARCH volatility estimates. Most studies report than EVT matches HS in unconditional coverage.

The superior unconditional coverage performance of HS compared to GARCH and EVT methods has been reported by Ouyang (2009). He uses daily returns of the Shanghai Synthesized Index and the Shenzhen Component Index to examine the performance of five different 99% VaR methods (EXMA, EQMA, GARCH(1,1), HS and an extreme value model (EVT) based on Pareto distribution) using two lengths of observation periods - one and three years. HS provides the best coverage performance for both the observation periods and for both the indices, matched only by the extreme value model. Ouyang (2009) is one of the few studies to compare the relative performance of models on the criteria of Basel back testing approach. His results imply that HS method could appeal to banks since it is least likely to invite penalty as per supervisory backtesting rules. Interestingly, while the coverage performance of HS (and other models) deteriorates when historical observation period is 3 years, the HS still remains the best performer, matched only by the EVT method. It is important to note that the coverage performance of HS is better than the models that use conditional volatility like GARCH (1,1) and EXMA. This could be because Ouyang (2009) converts the conditional volatility generated by these models to VaR estimates using a normal distribution assumption. Also the coverage measures are unconditional. Being an empirical method, the HS performs the best because it models the fat tails in the data the best, matched only by the EVT method.

Danielsson (2000) uses data spanning 15 years for equities, bonds, commodities and foreign exchange to generate daily 99% VaR forecasts from four methods, namely GARCH with normal and student t assumptions, historical simulation and EVT. The GARCH with normal distribution assumption has the worst performance (poorest unconditional coverage) among the models across all asset classes. The GARCH with Students t is much better than GARCH with normal in all asset classes. In the case of S&P 500 the GARCH with Student's t has performance comparable to that of HS and EVT models. However, in case of commodities (oil), Exchange rate (GBP/USD) and individual stock (Microsoft) the performance of GARCH with Student's t assumption, though better than GARCH with normal assumption, is not as good as HS and EVT. As in the case of other studies, the GARCH with normal assumption performs well at 95% confidence levels. There is no significant difference in the performance of the EVT in comparison to HS.

McNeal and Frei (1998) combine the GARCH and EVT (with generalized pareto distribution) methods calling it the 'conditional EVT method (CEVT)' and back test the method using data ranging from 16 to30 years in length for two indices, one exchange rate and one commodity. For the 99% VaR the GARCH with a normal distribution assumption for the innovations systematically provides inadequate coverage for all the series; the GARCH with a t distribution assumption is better than the GARCH with normal distribution assumption; the EVT and GARCH with t distribution are equally good; and the CEVT is better than the EVT and both the GARCH specifications. While they do not include the HS in their comparisons the comparison between the GARCH with normal and t distributions underlines the fact that GARCH with normal distribution assumption for the innovations performs poorly at 99% owing to fat tails in data. Danielsson and de Vries (2000) also report that the GARCH with normal assumption performs very poorly for tail estimates compared to the GARCH with t distribution assumption.

c) Hypothesis testing of unconditional coverage measures

Huang and Tseng (2009) compare the performance of seven methods (HS, MC, two GARCH, two kernel methods (KE) applied to left tails of distributions) using data between 1980 and 2007 on 37 stock indices, 12 from developed countries and 25 from emerging markets. The historical simulation method fails the test of unconditional coverage in 9 out of 12 developed country indices. The MC and GARCH fail in all the indices. Both the KE models pass the test of unconditional coverage in all the developed market indices. Of the emerging market indices the hypothesis of correct unconditional coverage is accepted in 20 out of 25 markets for the historical simulation method (the authors donot try to interpret the better performance of HS for emerging market indices). The KE methods are still better with more number of acceptances but none of the other (MC, GARCH) methods report correct unconditional coverage for any of the emerging market indices either. The important points that emerge from this study are - the traditional methods are not able to stand upto a rigorous scrutiny of their unconditional coverage, with the exception of the HS method in emerging markets; and the KE methods outperform all traditional methods in hypothesis tests.

Sheedy (2008) used daily return data ranging from 16 to 30 years for five equity indices (S&P500, FTSE100, HSI, Nikkei and ASX200) to generate VaR estimates using six methods, three unconditional (VC using normal and Students t and HS) and three conditional (GARCH-N, GARCH-t and FHS). The results of hypothesis testing of unconditional coverage (using Christoffersen (1998) LR_{uc} statistic) are insightful. For 99% VaR of long and short equity portfolios the HS has better unconditional coverage than the GARCH-N and

GARCH-t. It is also better than VC with normal and Student's t. While the HS outperforms even in 95% VaR, this advantage is not as significant as in case of 99% VaR.

Not much literature is available on performance of HS using confidence interval construction. Pascual, Romo and Ruiz (2001) find that confidence intervals constructed around the historical simulation method are ex ante, on average, very wide and have low effective coverage in comparison to the promised confidence level. This is attributed to the HS being unable to capture the persistence in volatility.

d) Unconditional coverage performance of modified HS methods

The alteration of HS to incorporate declining weights improves the unconditional coverage performance of the model. The altered method outperforms conditional volatility methods like GARCH (even with t distribution), EXMA and is as good as tail modelling approaches like EVT.

Boudoukh et al (1998) modify the historical simulation approach by assigning exponentially declining weights (as in EXMA) to the most recent observations. They compare the historical simulation and the altered approach with EXMA and Variance-covariance approaches using daily returns data on one exchange rate, one stock index, one bond index and spot oil prices from 1991 to 1997. The HS is still the best performer as far as unconditional coverage (bias defined as average proportion of exceptions) is considered. However, the altered method significantly enhances the performance of the EXMA by reducing its bias to levels comparable to the HS.

Angelidis et al (2007) look at large and small cap portfolios from DJ Euro Stox for time period 1987 to 2005. They split the time period into two in order to investigate the robustness of VaR models over time. The methods covered by them are, VC, GARCH (with variety of specifications including Student's t distributions), HS, EVT and FHS). They find that only the FHS method is consistent in acceptance of unconditional coverage hypothesis across both time periods for short positions of long and short portfolios at 99% VaR.

Sheedy (2008) used daily return data ranging from 16 to 30 years for five equity indices (S&P500, FTSE100, HSI, Nikkei and ASX200) to generate VaR estimates using six methods, three unconditional (VC using normal and Students t and HS) and three conditional (GARCH with normal and student's t and FHS). The results of hypothesis testing of unconditional coverage (using Christoffersen (1998) LR_{uc} statistic) for 99% VaR of long and short equity portfolios the FHS has best unconditional coverage compared to other methods.

IV. Conditional coverage performance measures of VaR

Conditional coverage performance measures are important in spotting VaR methods that result in clustered exceptions. Clustered exceptions imply that the VaR method does not capture the innovations in volatility.

Hendricks (1996) has used correlation between VaR and absolute value of outcome as measures of conditional VaR performance. The autocorrelation of the VaR exceptions are used as conditional coverage performance measures by Boudoukh (1998) and Hull and White (1997). The Ljung-Box statistic is used to measure the statistical significance of the autocorrelation coefficient.

Christoffersen (1998) models the exceptions as a Markov chain. A likelihood ratio (LR_{cc}) statistic is used to test the null hypothesis of conditional coverage (joint hypothesis of independence and unconditional coverage). The test allows only for dependence of order 1.

A linear regression test proposed by Christoffersen (1998), Clements and Taylor (2002) and applied by Engel and Manganelli (2004) allows lags of order greater than one to be modelled in an autoregressive specification of exceptions. The specification can be expanded to include variables other than the autoregressive terms as well. A likelihood ratio statistic or a Wald statistic is used to test for the joint significance of the coefficients. Christoffersen (1998) also recommends the J test from the Hansen's (1982) GMM framework though applications of the same to VaR forecasts were not found in literature.

Clements and Taylor (2002) propose a modification of Christoffersen's (1998) test to incorporate multi period dependence. For this modification the periodicity of the dependence must be known.

V. Conditional coverage performance of Historical Simulation VaR and its modifications

In the case of papers that test the null hypothesis of correct conditional coverage it is important to note that the results reported are joint tests for unconditional coverage and independence of exceptions. The advantage of better modelling of tails of the returns distribution by HS is lost when the conditional coverage performance measures are used. However, the modification to HS, namely FHS outperforms conditional volatility models like EXMA and GARCH in conditional coverage performance.

Hendricks (1996) uses correlation between VaR and absolute value of outcome to report that the historical simulation method performs poorly in comparison to the exponentially weighted moving average measure, indicating that it is slower to reflect recent changes in risk levels. Moreover, longer observation periods show lower correlations, so much so that the historical simulation method VaR with 1250 days is uncorrelated with the absolute value of outcome. Pritsker (2006) compares the HS, FHS and BRW (all modifications of the HS) for daily returns of exchange rates over a 13 year period to find that the FHS responds the best to changes in volatility. Neither Hendricks (2006) nor Pritsker (2006) carry out tests of conditional coverage. Literature using hypothesis testing for conditional coverage are discussed in the paragraph below.

Alexander and Sheedy (2006) use daily returns data of 20 year length for three currency pairs (USD/JPY, GBP/USD and AUD/USD) to generate VaR estimates (99%, 99.5% and 99.9%) using VC-N (variance covariance method with normal distribution), VC-t (variance covariance method with Student's t), HS (smoothed with a kernel), GARCH-N, GARCH-t and FHS. They find that the HS performs better than VC (normal) and GARCH-N in tests of unconditional coverage. While VC-N and GARCH-N fail tests of conditional coverage, FHS passes the tests for all series under consideration. VC-t and GARCH-t perform better than their normal counterparts in conditional coverage, but not as consistently as the FHS.

Sheedy (2008) used daily return data ranging from 16 to 30 years for five equity indices (S&P500, FTSE100, HSI, Nikkei and ASX200) to generate VaR estimates using six methods, three unconditional (VC using normal and Student's t and HS) and three conditional (GARCH with normal and student's t and FHS). The results of hypothesis testing of conditional coverage (using Christoffersen (1998) LR_{cc} statistic) are insightful. For 99% VaR of long and short equity portfolios the FHS has best conditional coverage (joint test of unconditional coverage and independence). This advantage doesnot exist in case of 95% VaR with the conditional normal performing equally well.

Angelidis and Degiannakis (2005) look at six different data series - two equity, 2 commodities and 2 exchange rates - from 1989 to 2003. They find that the filtered historical simulation method is an improvement over parametric (VC, GARCH and EXMA) and non parametric (HS) method for 99% VaRs using S&P 500 data in case of long positions using both conditional and unconditional coverage tests.

Conclusions

This paper surveys the literature to understand the performance of the historical simulation method of VaR calculation - the most popular method of VaR calculation. The research covers unconditional coverage and conditional coverage performance measures. Historical simulation provides superior unconditional coverage in comparison to the simple (variance covariance, Monte Carlo, exponentially weighted average approaches) as well as the sophisticated GARCH approach (with normal distribution) for VaR at 99% confidence levels. This superiority is attributed to the better modelling of fat tails in the data by the empirical distribution. The unconditional coverage performance of the HS method is paralleled or bettered by the EVT and FHS approaches. This advantage carries to rigorous tests of the hypothesis of correct unconditional coverage. Most Methods excepting HS, EVT and FHS exhibit rejection of the null of unconditional coverage.

However, the apparent superiority of historical simulation method vanishes when conditional coverage performance measures are used. The HS mostly fails these tests, but so do most of the other methods including the GARCH-N. Interestingly, a modified historical simulation method, the FHS is the best performer in most conditional coverage tests.

In summary, as long as regulatory back tests favour unconditional coverage performance measures of VaR estimates, banks will have no incentive to change their historical simulation based VaR models since the method outperforms or equals most other methods in unconditional coverage.

There are a number of caveats to this study. Not too much literature is available on comparative performance of VaR methods using conditional coverage performance measures. The survey in this paper doesnot cover papers on VaR of portfolios with derivatives.

References

Alexander, C. and Sheedy, E. (2008) ‘Developing a stress testing framework based on market risk models, Journal of Banking and Finance, 32, 2220-2236.

Angelidis, T. and S. Degiannakis (2005) ‘Modeling Risk for Long and Short trading Positions’, The Journal of Risk Finance, 6, 3, pp.226-238.

Angelidis, T., A Benos, and S. Degiannakis (2007) ‘A robust VaR model under different time periods and weighting schemes’, Rev Quant Finan Acc, 28, pp.187-201.

Barone-Adesi, G. And Giannopoulos, K. (2001), ‘Non-parametric VaR techniques: myths and realities’, Economic Notes, Vol.30, banca Monte dei Paschi di Siena, Siena, pp.167-81.

Barone-Adesi, G., Giannopoulos, K. And Vosper, L. (1999). 'VaR without Correlations for Portfolios for Portfolios of Derivative Securities', Journal of Futures Markets, 19, 583-602.

Bollerslev, T., Chou, R.Y. and Kroner, K.F. (1992) 'ARCH Modeling in Finance: A Selective Review of the Theory and Empirical Evidence', *Journal of Econometrics*, No. 52, pp. 5-59.

Boudoukh, J, M. Richardson, and R. Whitelaw. (1998) 'The Best of Both Worlds,' *Risk,* May, pp 64-67.

Butler, J. S. and B. Schachter, 'Estimating Value at Risk with a Precision Measure by Combining Kernel Estimation with Historical Simulation', *Review of Derivatives Research*, 1998, 1, pp. 371-390.

Campbell, S.D. (2007) 'A Review of Backtesting and Backtesting Procedures', *The Journal of Risk,* 9, 2, pp.1-17.

Christoffersen, P. (1998) 'Evaluating interval forecasts', *International Economic Review,* 39, pp.841-62.

Christoffersen, P. and Gonsalves, S. (2005) 'Estimation risk in financial risk management', *The Journal of Risk,* 7, 3, pp.1-28.

Clements M.P., and Taylor N. (2002). 'Evaluating interval forecasts of high-frequency financial data.' Manuscript, University of Warwick.

Danielsson, J. (2002) 'The Emperor has no clothes: limits to risk modeling', *Journal of Banking and Finance*, 26, pp. 1273-96.

Danielsson, J. and C De Vries (1997) 'Value at Risk and extreme returns', FMG Discussion Paper No. 273, Financial Markets Group, London School of Economics.

Engle, R.F. (2001) 'GARCH 101: The use of ARCH/GARCH models in Applied Econometrics', *Journal of Economic Perspectives,* vol. 15, No.4, pp. 157-168.

Engle, R.F., and Manganelli, S. (2004) 'CAViaR: conditional autoregressive value-at-risk by regression quantiles', *Journal of Business and Economic Statistics,* 22, pp.367-81.

Hansen, L.P. (1982) “Large Sample Properties of Generalized Method of Moments Estimators,” *Econometrica,* 50, 1029-54.

Hendricks, D. (1996) ‘Evaluation of Value-at-Risk Models Using Historical Data’, *Economic Policy Review*, pp. 39-69.

Herlin, C. and S. Tokpavi (2007) ‘Backtesting Value-at-Risk Accuracy: a simple new test’, *The Journal of Risk*, 9, 2, pp. 19-37.

Hoppe, R. (1998) ‘VaR and the unreal world, *Risk*, Vol. 11, pp.45-50.

Huang, A.Y. and T. Tseng (2009) ‘Forecast of value at risk for equity indices: an analysis from developed and emerging markets’ *The Journal of Risk Finance*, Vol. 10 No. 4, 2009, pp. 393-409.

Hull, J., & White, A. (1998). ‘Incorporating volatility updating into the historical simulation method for value-at-risk’, *Journal of Risk*, 1(Fall), 5-19.

Jackson, P., Maude, D. and Perraudin, W. (1997) ‘Bank Capital and Value at Risk’, *Journal of Derivatives,* 4 (Spring), pp. 73–89.

Kalvyas, L. Dritsakis, N., Siriopoulos C. and Grose, C. (2004) ‘Selecting Value at Risk Methods According to their Hidden Characteristics’, *Operational Research*, Vol. 4, No.2, pp.167-189.

Lopez, J.A. (1999), ‘Methods for evaluating value-at-risk estimates’, *Economic Review*, vol. 2, pp.3-17.

Manganelli, S. And Engle, R.F. (2001). Value at risk models in finance. Working paper series 75, European Central Bank.

McNeal, A. And Frei, R. (1998). Estimation of tail related risk measures for heteroskedastic financial time series: an extreme value approach, Zurich:ETH.

Miura, R. and S. Oue (2001) ‘Statistical Methodologies for Market Risk Measurement’, *Asia Pacific Financial Markets*, pp.305-319.

Ouyang, Z. (2009). 'Model Choice and Value at Risk Estimation', *Qual Quant,* 43, pp.983-991.

Pritsker, M. (1995) 'Evaluating Value at Risk Methodologies: Accuracy versus Computational Time,' Board of Governors of the Federal Reserve System. Mimeographed.

Pritsker, M. (2006) 'The hidden dangers of historical simulation', *Journal of Banking and Finance*, 30, pp.561-582.

Sarma, M., Thomas, S. and A. Shah (2003) 'Selection of Value-at-Risk Models', *Journal of Forecasting*, 22, pp.337-358.

Sheedy, E. (2008). 'Why VaR Models Fail and what can be done', Macquarie University Applied Finance Center Research Paper no. 34.

Sollis, R. (2009) 'Value at Risk: A Critical Overview', *Journal of Financial Regulation and Compliance, Vol. 17, No.4,* pp. 398-414.

Supervisory Framework for the Use of Backtesting in Conjunction with the Internal Models Approach for Market Risk Capital Requirements (1996). Basel committee for Banking Supervision, http://www.bis.org/publ/bcbsc223.pdf.

Transition to the Internal Models Approach for Market Risk – a Survey Report. Aptivaa. www.aptivaa.com/pdfs/SurveyReport.pdf (2000). Accessed on 23 August 2011.

Vlaar P (2000) 'Value at risk models for Dutch bond portfolios' *Journal of Banking and Finance,* 24: pp. 131–154.

Wagner, N. And Marsh, T.A. (2005) 'Measuring tail thickness under GARCH and an application to extreme exchange rate changes', *Journal of Empirical Finance*, vol. 12, pp.165-85.