Economic Notes by Banca Monte dei Paschi di Siena SpA, vol. 30, no. 2-2001, pp. 167–181

Non-parametric VaR Techniques. Myths and Realities

GIOVANNI BARONE-ADESI - KOSTAS GIANNOPOULOS*

VaR (value-at-risk) estimates are currently based on two main techniques: the variance-covariance approach or simulation. Statistical and computational problems affect the reliability of these techniques. We illustrate a new technique – filtered historical simulation (FHS) – designed to remedy some of the shortcomings of the simulation approach. We compare the estimates it produces with traditional bootstrapping estimates.

(J.E.L.: G19).

1. Introduction

Early VaR estimates were linear multipliers of variance-covariance estimates of the risk factors. This class of market risk techniques soon became very popular, mainly because of their link to modern portfolio theory. However, during worldwide market crises, users noticed that early models failed to provide good VaR estimates. In addition, variance-covariance VaR techniques require a large number of data inputs; all possible pairwise covariances of the risk factors must be included in a portfolio. To process all the necessary information demands much computer power and time. Factorization methods provide only partially satisfactory answers.

The early VaR models are also referred as *parametric* because of the strong theoretical assumptions they impose on the underlying properties of the data.[1] One such assumption is that the density function of risk factors influencing asset returns must conform to the multivariate normal distribution.[2]

* Giovanni Barone-Adesi, Università della Svizzera Italiana and City Business School, email: Giovanni.Barone-Adesi@lu.unisi.ch; Kostas Giannopoulos, Westminster Business School, email: giannok@wmin.ac.uk.

[1] Parametric VaR models are based on strong theoretical assumptions and rules. They impose that the distribution of the data (daily price changes) conforms to a known theoretical distribution. The most popular of these models is the exponential smoothing (ES), see RiskMetrics®.

[2] The normality assumption is frequently used because the normal distribution is well described; it can be defined using only the first two moments and it can be understood easily. Other distributions can be used, but at a higher computational cost.

The empirical evidence however, indicates that speculative asset price changes, especially the daily ones, are rather non-normal. Excess kurtosis will cause losses greater than VaR to occur more frequently and be more extreme than those predicted by the Gaussian distribution. Many risk managers remember the large losses they faced during the Mexico (1996), Asian (1997) and Russian (1998) market crises. During these periods, negative returns several standard deviations beyond the 2.33 threshold predicted by the normal distribution[3] were recorded within only a few days. A large number of markets were crashing together; the correlation forecasts used to calcualted VaR failed to predict such a synchronous crash. That resulted in further VaR failure.

The problems of earlier models spurred the search for better estimates of VaR. A number of recent VaR techniques are based on non-parametric or a mixture of parametric and non-parametric statistical methods. The family of historical simulation (HS) models belongs to the former group. The filtered historical simulation (FHS) as developed by Barone-Adesi *et al.* (1998) and Barone-Adesi *et al*. (1999, 2000) belongs to the second group. This paper analyses the assumptions on which these models are based. In addition we compare the VaR estimates produced by the above models in linear and nonlinear portfolios.

2. Literature Review

Regulators require that financial institutions backtest their internal VaR models (Basle Committee on Banking Supervision, 1995). Although the popularity and the use of HS has increased during the last few years, reports of backtests conducted by users are not publicly available. Some researchers, however, used smaller portfolios to backtest HS. Van den Goorbergh and Vlaar (1999) used rolling windows of different lengths (250, 500, 1,000 and 3,038 days) over a 15-year period to backtest HS daily data on the AEX (Dutch equity index). They found that the failure rate, i.e. the probability that actual losses exceed VaR, is often exceeding the corresponding left-tail probabilities. Van den Goorbergh and Vlaar found that results are sensitive to the selection of the window length.[4]

In another study, Vlaar (2000) investigates the accuracy of various VaR models on Dutch interest rate based portfolios. He concluded that HS produced satisfactory results only when a long history is included in the data sample.

Brooks and Persand (2000) investigated the sensitivity of VaR models due to changes in the sample size and weighting methods. They used a set of equally weighted portfolios each containing two asset classes selected from a

[3] At 99% probability.

[4] They strongly criticized the use of HS in predicting extreme events (far left on the tail) when the window is not of a substantial length.

set of national equity indices, bond futures and FX rates.[5] They found strong evidence that VaR models could produce very inaccurate estimates when the 'right' historical data sample length is not selected.

Perhaps the most comprehensive model comparison study published to date has been carried out by Hendricks (1996). Hendricks used 4,255 daily observations of eight FX rates against the US dollar and several performance criteria to study the performance of twelve VaR models. The twelve VaR models were grouped in three categories, equally weighted moving average, exponentially weighted moving average and HS approaches. He found that none of the twelve approaches is superior to others in every criterion. Furthermore, Hendricks reported that risk measures from the various VaR approaches for the same portfolio on the same data could differ substantially. Differences in the accuracy across models were also sensitive to the choice of the level of probability used in the VaR calculation. When a 95% probability was used in the VaR calculation, Hendricks found that the three approaches produced accurate risk measures. However, when a 99% probability was used there was a large discrepancy in the risk calculation between the three approaches. In general, the three approaches predict only between 98.2 per cent and 98.5 per cent of the outcomes. Hendricks failed to single out any VaR approach and he predicted that a more accurate VaR model may be created by combining the best features of each single approach.

Pritsker (2000) reviews the assumptions and limitations of HS and weighted HS (Boudoukh *et al*., 1998). He points that both methods associate risk with only the lower tail of the distribution. In an example, he showed that after the crash of 1987 the estimated VaR of a short equity portfolio, as computed by HS or weighted HS, did not increase. The reason is that the portfolio recorded a huge profit during the day of the crash. Pritsker goes further by formulating some interesting properties of the HS and weighted HS. He showed that if the portfolio's return follow a GARCH(1,1) process, then at a 1-day VaR horizon and 99% confidence level, the HS and weighted HS methods fail to detect increases in VaR about 31 per cent of the time. In a simulated example, he showed that the VaR on a short equity portfolio did not increase during the days after the crash of 1987.

Barone-Adesi *et al*. (1999) carried out an extensive backtest analysis for the FHS model. They used economic and statistical criteria to analyse the breaks on 100,000 daily portfolios held by financial institutions. The portfolios consisted of interest rate futures and options on futures trading at LIFFE during 1996 and 1997, as well as purely plain vanilla swaps and mixed portfolios invested on futures, options and swaps. Overall, their findings sustain the validity of FHS as a risk measurement model.

[5] The indices used were S&P500 and FTSE100. The bond portfolios were the 30-year US Treasury bonds, UK gilt and long-term German Bund. The FX rates used were dollar/Swiss franc, dollas/Dmark and dollar/yen. The US Treasury bond rates were obtained from near-month futures prices on the 30-year interest rate. No information was given for the other two bonds.

3. Historical Simulation (HS)

HS (bootstrapping) is being increasingly used in the risk management industry. It consists of generating scenarios by sampling historical returns associated with each risk factor included in the portfolio. The aggregate value of all linear and derivative positions produces a simulated portfolio value. The procedure is repeated many times using all past returns.[6]

HS does not require any statistical assumption beyond stationarity of the distribution of returns or, in particular, their volatility.[7] Since the estimated VaR is based on the empirical distribution of historical returns of each individual risk factor, it reflects a more realistic picture of a portfolio's past risk. At least one year of recent daily returns must be used in HS. A longer period is more appropriate when available, but availability of historical data is often problematic for the universe of (linear) contracts or other risk factors. In its simplest form, HS can be shown by the following example. Given a data set of historical returns Θ, we draw an element e^*

$$e^* = \{e_1^*, e_2^*, \ldots, e_T^*\} \qquad e_i^* \in \Theta$$

where $i = 1, \ldots, T$ refers to past days to form a simulated price for asset Y:

$$Y_{T+1}^* = Y_T + Y_T^* e^* \tag{1}$$

The process in (1) is repeated and the simulated price series Y^* is recursively updated to the last day of the VaR horizon. This sequence of simulated prices for day $T+1, T+2, \ldots, T+N$ forms a simulated pathway or scenario for the risk factor Y. Here is a simple example:

Assuming that Y_T is 100 and the vector e^* has values

$$\{-0.01053, -0.00759, -0.00408, 0.00474, 0.00093, 0.00921, 0.01712, -0.00443, 0.01342, -0.00304\}$$

The sequence for simulated prices, $Y_{T+1}^*, \ldots, Y_{T+10}^*$ for the first simulation run will be

Day:	1	2	3	4	5
Price:	98.94707	98.19566	97.79456	98.25811	98.34967

Day:	6	7	8	9	10
Price:	99.25578	100.9552	100.5078	101.8569	101.547

[6] Past returns are drawn with or without replacement.

[7] As we are going to see later this becomes a limitation which under certain market conditions results in underestimation of VaR.

Derivative securities are priced from the knowledge of the simulated price of the underlying asset by their full re-evaluation under a pricing model at each node. Derivative values can be aggregated to form a simulated value of all positions on each contract.

HS relies on a uniform distributions to sample innovations from the past.[8] These innovations are applied to current asset prices to simulate their future evolution. Once a sufficient number of different paths have been explored, it is possible to compute a portfolio's risk without making arbitrary distributional assumptions. That makes HS especially useful in the presence of abnormally large historical losses. The final outcome of HS is the simulated distribution of portfolio values at the desired horizon. VaR is a percentile of that distribution.

HS has a number of advantages. It is easy to understand and to implement. It uses the empirical distribution of past returns to generate realistic future scenarios. Nonlinear positions can be re-priced under each scenario.[9] Furthermore, HS does not require the computation of a covariance matrix. In fact, statistical dependencies across simultaneous asset returns can be kept by taking returns from the same day in the historical record for all the assets at each node in the simulation.

HS's ability to predict future losses may be undermined, however, if the distribution of any risk factors is not i.i.d.[10] Using a constant volatility model to calculate VaR when the distribution of returns is not stationary, as is the case with most daily financial time series, could be very misleading.[11] In fact, the probability of having a large loss is then not equal across different days. During days with higher volatility we would expect larger than usual losses. That contrasts with HS, where the volatility of N days is proportional to the square root of time, i.e. constant volatility is assumed over any period.

[8] HS is also known as bootstrapping simulation. For a detailed discussion about this simulation technique see Efron and Tibshirani (1993).

[9] This is a more efficient treatment of derivative security risk when compared to models that use linear approximations. But as we can see later, it still presents the weakness of assuming that implicit volatility is constant.

[10] Independently and identically distributed. When returns are i.i.d. and the moments of the distribution are known, any inferences made about potential portfolio losses will be accurate and unchanging over time. Stationarity implies that the probability of occurrence of a specified loss is the same for each day. Independence implies that the size of price movement in one period will not influence the movement of any successive prices. These properties combined with the normality assumption simplify the VaR calculation for any holding period and probabilty. Under these hypotheses, VaR for longer periods can be found by multiplying the VaR estimate at a shorter, i.e. daily, horizon by the square root of the number of days t in the period of interest.

[11] HS assumes that past and present moments of the density function of returns of a specific risk factor are constant and equal. For a detailed discussion of the properties of HS, see Pritsker (2000).

Today there is a large body of evidence suggesting that the distribution of speculative price changes is fat-tailed with changing conditional moments. Empirical studies have found that there is a tendency, known as volatility clustering, for large price changes to be followed by more large changes. An increase in portfolio volatility during the risk measurement horizon affects the portfolio VaR.[12] Unfortunately, HS does not take into account such a change.

4. Filtered Historical Simulation (FHS)

The above limitations of HS, a limited set of outcomes and unresponsiveness to changes in market volatility, can be overcome with the use of filtered historical simulation (FHS) (Barone-Adesi *et al.*, 1999). FHS is a generalized HS. It has all the positive properties and overcomes most of the HS weaknesses. In FHS the stationarity assumption is relaxed; historical returns[13] are first standardized by volatility estimated on that particular day (hence the name of filtered),

$$\eta_t = \frac{\varepsilon_t}{\sqrt{\hat{h}_t}}$$

This filtered process yields approximately i.i.d. returns (residuals) suited for HS.

Before filtered returns are used as innovations, they are scaled (multiplied) by the current conditional forecast of volatility; thus they reflect current market conditions:

$$Y^*_{T+n} = Y^*_{T+n-1} + Y^*_{T+n-1}\eta\sqrt{h_{T+n}} \tag{2}$$

where

$$\eta^* = \{\eta^*_1, \eta^*_2, \ldots, \eta^*_T\} \qquad \eta^*_i \in \Theta$$

where $i = 1, \ldots, T$ refers to past days and $\sqrt{h_{T+n}}$ is the simulated conditional volatility for VaR day n and is estimated recursively by a time-series model, such as

$$h_{T+n} = \omega + \alpha(\zeta^*_{T+n-1})^2 + \beta h_{T+n-1} \tag{3}$$

[12] For some types of investments, two or three consecutive adverse price changes may be sufficient to ruin the investor, i.e. contingent claims and investments with leverage. In a leveraged portfolio, investors borrow at a fixed rate and invest in a risky asset, usually equity. Leverage increases positive expected return because leverage magnifies volatility. The larger the leverage factor, the higher the gains will be in the case of positive returns. But large adverse returns increase the downside risk and chances of a disaster. Thus the VaR is far more important in the management of leveraged investments.

[13] Since returns are usually serial correlated in practice we replace Y_t with ε_t, where

$$\varepsilon_t = Y_t - \mathrm{E}[Y_t/\Phi_{t-1}]$$

where $\zeta^*_{T+n} = \eta^* \sqrt{h_{T+n}}$ is computed as in (2).

The 'unconditional' or 'unfiltered' HS described earlier is a special case of FHS, which holds when returns are i.i.d.[14] A major advantage of FHS over HS is that the filtering process increases the range of outcomes beyond the historical record through a change of scale. In other words, FHS provides a systematic approach to generate extreme events not present in the historical record, completing the tails of the distribution. FHS requires, therefore, a shorter historical record than HS to simulate the tails of the distribution of returns. Because this process is essentially an extrapolation, its validity must be carefully tested. Barone-Adesi *et al.* (2000) present evidence of the adequacy of this procedure for risk management.

The FHS approach can be adapted to stress testing because it simulates the whole distribution of security returns. It is not limited to the observed returns as is regular bootstrapping. Therefore, it is possible to sample from more extreme points in the tails of the multivariate distribution by increasing the number of simulation runs. As an example, applying FHS over a 10-day horizon using a database of 500 daily returns, the number of different possible pathways is 500^{10} for any given set of initial conditions. Therefore, it is possible to generate an almost arbitrarily large number of points of the simulated distribution. This is especially important for portfolios of derivatives that may experience more stress at points not on the tails of the distribution.

Finding the most stressful conditions for a given portfolio may be slow, because the probability of each pathway is the inverse of the number of pathways. Of course, the problem of finding the most stressful conditions is easier for linear portfolios, for which the most stressful conditions are given by the largest negative or positive returns that can be pre-selected in the simulation procedure.

5. HS or FHS? An Empirical Investigation

In this example, we investigate the difference in VaR estimates between HS and FHS. Three hypothetical portfolios are selected. The first portfolio consists of a long position on the S&P100 while each of the other two consists of a short European call option on the same basis. The historical data set used in the simulation covers the period from 1 January 1997 until 26 November 1999. Our scope is to compare VaR estimates for the three portfolios using the two alternative methodologies. Figure 1 shows the daily returns and conditional volatilities from 2 January 1997 until 18 December 1999 from the

[14] The algorithm can easily be expanded to generate parallel pathways on a multi-asset environment. This does not require the use of a correlation matrix. For a full description see Barone-Adesi *et al.* (1999).

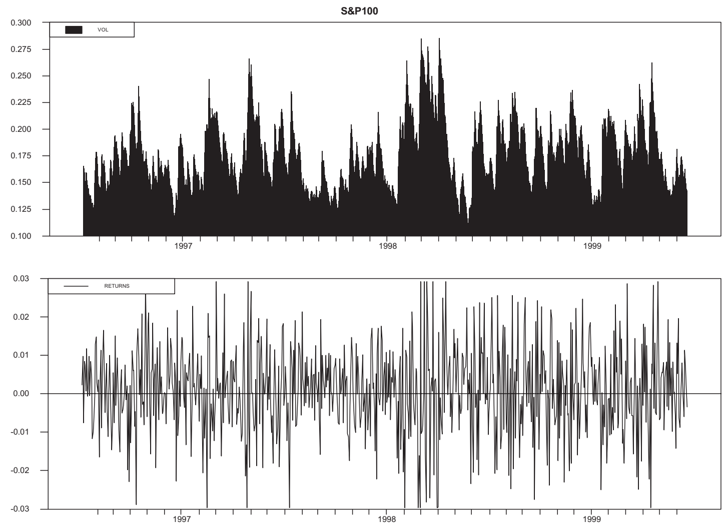

Figure 1: Conditional Volatility and Returns for the S&P100

S&P100. The conditional volatility of the returns, shown on the top graph, is rather non-constant, a fact that contradicts the assumption on which HS is based.

The value of the linear portfolio, the S&P100 index, at the close of business on the 26 November is \$753.56. The first step in our simulation consists in generating the simulated scenarios for the S&P100 index using the two alternative methods. To form the innovations, e, in (1), a random return[15] is selected from the historical data set.[16] Updating (1) generates multiperiod scenarios. A total of 5,000 simulated multiperiod scenarios, for 1, 2, . . ., 20 business days ahead, are generated.

Dividing each of the daily returns with that day's volatility forms the set of innovations, η^*, utilized by the FHS simulation, as shown in Figure 2.

Using the above data set of standardized residuals and the FHS algorithm, we generate 5000 multiperiod scenarios for days 1, 2, . . ., 20.

To investigate any possible discrepancy on VaR between the two, HS and FHS, methods, we carry out the FHS using two alternative risk scenarios on the last trading day. Under the first set of simulations, we set the volatility on 26 November 1999 to a low, but not unrealistic, level of 7 per cent per annum. On the second sets of simulations, we assume that there was a market crisis

[15] Rather, these are zero mean returns.

[16] The data are drawn with replacement. There are 770 observations in the data set. At each simulation run, on average, each observation will be drawn approximately 7 times.

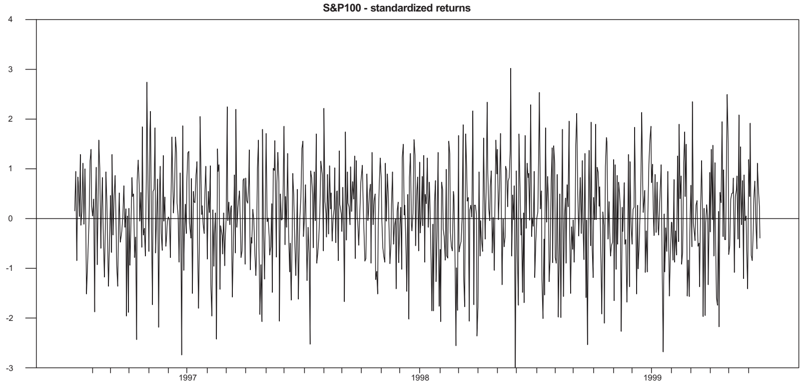

Figure 2: Standardized Returns

and the market volatility on the last trading date jumped to 30 per cent per annum.

Figure 3 shows the histogram from the simulated prices for day 1 as produced in each of the three simulations. The top histogram displays the distribution of returns from the HS, while the other two histograms report returns from the FH under each of the two, low and high, volatility regimes prevailing on the last trading day.

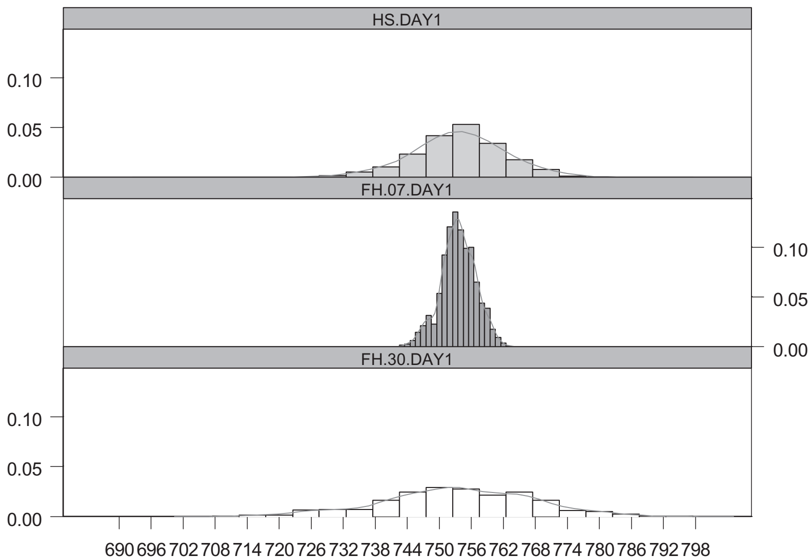

Figure 3: S&P Simulated Prices for Day 1

Figure 4 shows the histograms for each set of simulated prices corresponding to a 10-day horizon. As one would expect, in a volatile period – when the volatility on the last trading date is 30 per cent per annum – the forecast interval is wide, and in quiet periods it is narrow. This is reflected in the forecast produced by the FHS. On the contrary, HS makes interval forecasts that are static, taking no notice of the last trading date's risk level.

The VaR at 99% is the lower percentile in the set of simulated prices. Table 1 reports the various VaR estimates for days 1, 5, 10 and 20. The second column from the left shows the VaR at 99% as estimated by the HS. The other two columns report the FHS under two different risk states prevailing at the close of business on 26 November 1999. The historical volatility in the historical data sample was 19.5 per cent per annum.[17]

As expected, the VaR estimates produced under the HS occur between the other two estimates as computed by the FHS algorithm, e.g. using two different volatility values on the close of business on 26 November 1999. The 1-day VaR is proportional to the initial volatility used in this simulation process. Due however, to the mean-reverting volatility process that characterizes the parameterization of conditional volatility on all GARCH models, we observe that the VaRs estimated with the FHS algorithm have a tendency, in the long run, to converge towards the HS VaR estimates. The ratio of 1-day VaR estimates

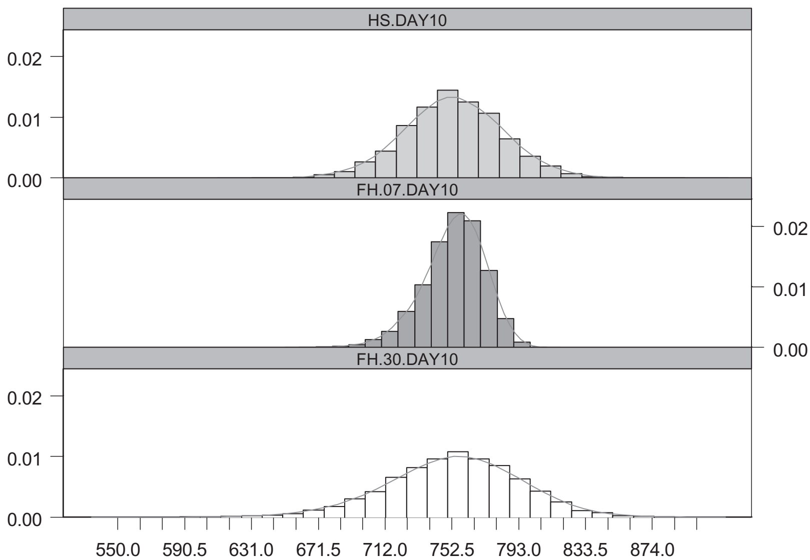

Figure 4: S&P100 Simulated prices of Day 10

[17] By calibration the average GARCH volatility is equal to the historical one.

Table 1: VaR Estimates (at 99%) on a Linear Portfolio

Horizon (days)	HS	FHS @ 7%	FHS @ 30%
1	23.87	8.30	35.67
5	57.96	30.08	72.91
10	73.66	50.47	102.04
20	97.52	87.63	129.23

computed by FHS with initial volatility of 7 per cent per annum and that computed by the HS is 0.35. This is very close to the ratio between the initial volatilities used by the two methods. On a 10-day VaR, the same FHS estimates closes the gap, with the ratio rising to 0.68 and, on a 20-day VaR, the ratio increases to 0.90. We draw similar conclusions by analysing the ratio between the HS and that of the FHS that has an initial volatility of 30 per cent per annum. This observation indicates that the bias present in the HS estimates becomes smaller at longer horizons.

The discrepancies in VaR between the two methods, HS and FHS are obvious in the above example. The magnitude in the discrepancies, however, increases when the portfolio contains predominantly nonlinear instruments. The next example uses two alternative portfolios, each containing a single, short, European call option. The option in the first portfolio has a strike of \$678 and is 'in-the-money'. The second call is chosen to be 'out-of-the-money' with a strike of \$828. Both options expire in twenty business days. We used the Black and Scholes model (1973) to find their market value[18] on 26 November 1999. The market value for the 'in-the-money' call is estimated at \$77.54 and the value for the 'out-of-the-money' call is estimated at \$0.85. These numbers are the values of the two portfolios at the end of the business on 26 November 1999.

Using the same simulated scenarios for the linear portfolios, we computed sets of 5,000 pay-offs and profits for the nonlinear portfolio. The profit at VaR horizon t for a single simulation run is

$$\text{Profit}_t = -(\text{Max}[P_t - X, 0], -C) \tag{4}$$

where X is the strike, C is the market value of the call on 26 November 1999 and P_t is the simulated value of the underlying asset at VaR horizon t for that simulation run. The simulated profits for the in-the-money call and for 10-day horizon are shown in Figure 5. The top histogram displays the HS profits and the other two the profits for the FHS under low and high volatility regimes prevailing on the last trading date.

[18] Assuming that the implied volatility equals the historical volatility on 26 November 1999, i.e. 19.5 per cent per annum.

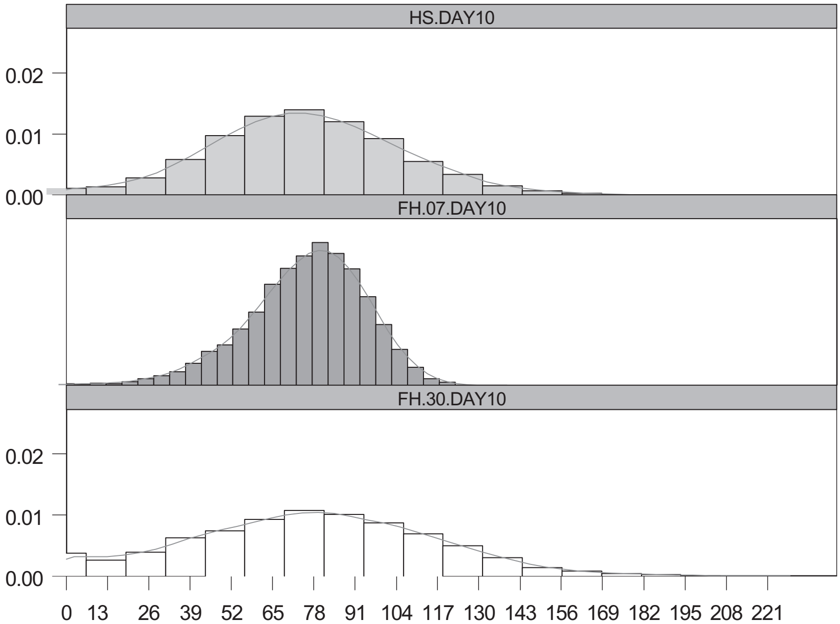

Figure 5: Simulated Pay-off for the 'in-the-money' Call at 100-day VaR

The upper percentile on each of the simulated set of prices less the call premium is our VaR estimate at 99% probability. These values, together with the ones for day 1, 5 and 20, are reported in Table 2. Note that the effect of low volatility is relatively stronger at short horizons, but it is still important at longer horizons. In summary, neglecting to adjust HS to current market conditions through FHS may cause substantial errors in the computation of risk.

Table 3 reports the equivalent VaR values for the portfolio invested in the short out-of-the-money call option, with a strike of $828.

Table 2: VaR Estimates (at 99%) on a Short in-the-money European Call

Day	HS	FHS-07	FHS-30
1	19.2815	5.60252	26.43452
5	48.60474	19.43428	59.78121
10	70.23241	33.21167	80.42265
20	97.42933	59.60564	106.8141

Table 3: VaR Estimates (at 99%) on a Short out-of-the-money European Call

Day	HS	FHS-07	FHS-30
1	0.0	0.0	0.0
5	0.0	0.0	0.0
10	0.0	0.0	5.28878
20	22.55335	0.0	29.50686

The effect of different volatility on VaR estimates is even more important at the longer horizon.

6. Conclusions

In principle, filtered historical simulation (FHS) can produce risk measures that are consistent with the current state of markets at any arbitrarily large confidence level. These impressive gains allow FHS to dominate historical simulation (HS) easily. In fact, HS fails to condition forecasts on the current state of the markets. The difference between the two simulation methods is magnified by the presence of options in the portfolio.

Although the superiority of FHS on HS is obvious, further development may mitigate some of the current limitations of FHS. These limitations of FHS rest mostly on two of the assumptions on which it is based: that correlations across asset returns are not related to the scale of returns; and that the scaling process accurately describes the tails of the distribution of asset returns. Neither of these assumptions is literally true. Therefore, our current FHS technique should be considered as a first-order approximation to a simulation engine accurately describing portfolio returns.

REFERENCES

G. BARONE-ADESI - F. BOURGOIN - K. GIANNOPOULOS (1998), "Don't Look Back", *Risk*, 11, August, pp. 100–04.

G. BARONE-ADESI - K. GIANNOPOULOS - L. VOSPER (1999), "VaR Without Correlations for Non-Linear Portfolios", *Journal of Futures Markets*, 19, August, pp. 583–602.

G. BARONE-ADESI - K. GIANNOPOULOS - L. VOSPER (2000), "Backtesting the Filtered Historical Simulation", unpublished manuscript.

BASLE COMMITTEE ON BANKING SUPERVISION (1995), "An Internal Models-based Approach to Market Risk Capital Requirements", Basle, (available at http://www.bis.org).

F. BLACK - M. SCHOLES (1973), "The Pricing of Options and Corporate Liabilities", *Journal of Political Economics*, May, pp. 637–59.

J. BOUDOUKH - M. RICHARDSON - R. WHITELAW (1998), "The Best of Both Worlds", *Risk*, 11, May, pp. 64–7.

C. BROOKS - G. PERSAND (2000), "The Pitfalls of VaR Estimates", *Risk*, May, pp. 63–6.

B. EFRON - R. J. TIBSHIRANI (1993), *An Introduction to Bootstrapping*, Chapman and Hall, New York.

D. HENDRICKS (1996), "Evaluation of Value-at-Risk Models Using Historical Data", *Federal Reserve Bank of New York Economic Policy Review*, April, pp. 39–69.

M. PRITSKER (2000), "The Hidden Risks of Historical Simulation", Federal Reserve Board, Washington.

R. VAN DEN GOORBERGH - P. VLAAR (1999), "Value-at-Risk Analysis of Stock Returns. Historical Simulation, Variance Techniques or Tail Index Estimation?" Mimeo.

P. VLAAR (2000), "Value at Risk Models for Dutch Bond Portfolios", *Journal of Banking and Finance*, 24, pp. 1131–54.

Non-technical Summary

Historical simulation (bootstrapping) is being used increasingly in the risk management industry. It consists of generating scenarios by sampling historical returns associated with each risk factor included in the portfolio. The aggregate value of all linear and derivative positions produces a simulated portfolio value. The procedure is repeated many times using all past returns.

Historical simulation does not require any statistical assumption beyond stationarity of the distribution of returns or, in particular, their volatility. Since the estimated value-at-risk (VaR) is based on the empirical distribution of historical returns of each individual risk factor, it reflects a more realistic picture of a portfolio's past risk.

Historical simulation's ability to predict future losses, however, may be undermined if the distribution of any risk factors is not i.i.d. (independently and indentically distributed). Using a constant volatility model to calculate VaR when the distribution of returns is not stationary, as is the case with most daily financial time series, could be very misleading. During days with higher volatility, we would expect larger than usual losses. That contrasts with historical simulation, where the volatility of N days is proportional to the time, i.e. constant volatility is assumed over any period.

There is a large body of evidence suggesting that the distribution of speculative price changes is fat-tailed with changing conditional moments. Empirical studies have found that there is a tendency for large price changes to be followed by more large changes. An increase in portfolio volatility within the risk measurement horizon affects the portfolio VaR. Unfortunately historical simulation does not take into account such a change.

The above limitations of historical simulation – limited set of outcomes and unresponsiveness to changes in market volatility – can be overcome with the use of filtered historical simulation (FHS). Filtered historical simulation has all the positive properties and overcomes most of the historical simulation weaknesses. In filtered historical simulation, the stationarity assumption is relaxed; historical returns are first standardized by volatility estimated on that particular day.

The filtering process yields approximate i.i.d. returns (residuals) suited for historical simulation. Before filtered returns are used as innovations, they are scaled (multiplied) by the current conditional forecast of volatility; thus, they reflect current market conditions.

The 'unconditional' or 'unfiltered' historical simulation is a special case of FHS, which holds when returns are i.i.d. A major advantage of FHS over historical simulation is that the filtering process increases the range of outcomes beyond the historical record through a change of scale. In other words, FHS provides a systematic approach to generate extreme events not present in the historical record, completing the tails of the distribution. FHS requires therefore a shorted historical record than historical simulation to simulate the tails of the distribution of returns.

The FHS approach can be adapted to stress testing because it simulates the whole distribution of security returns. It is not limited to the observed returns as is regular bootstrapping. Therefore, it is possible to sample from more extreme points in the tails of the multivariate distribution by increasing the number of simulation runs.