

3-1-2009

Market risk VaR historical simulation model with autocorrelation effect: A note

Wantanee Surapaitoolkorn
SASIN Chulalunkorn University

Recommended Citation

Surapaitoolkorn, Wantanee (2009) "Market risk VaR historical simulation model with autocorrelation effect: A note," *International Journal of Banking and Finance*: Vol. 6: Iss. 2, Article 9.
Available at: <http://epublications.bond.edu.au/ijbf/vol6/iss2/9>

This Journal Article is brought to you by the Faculty of Business at ePublications@bond. It has been accepted for inclusion in International Journal of Banking and Finance by an authorized administrator of ePublications@bond. For more information, please contact [Bond University's Repository Coordinator](#).



MARKET RISK VaR HISTORICAL SIMULATION MODEL WITH AUTOCORRELATION EFFECT: A NOTE

Wantanee Surapaitookorn
SASIN Chulalunkorn University, Thailand

Abstract

The modern market risk model using Value at Risk (VaR) method in the banking area under the BASEL II Accord can take different forms of simulation. In this paper, historical simulation will be applied to the VaR model comparing the two different approaches of Geometric Brownian Motion (GBM) process and Bootstrapping methods. The analysis will use correlation plots and examine the effects of the autocorrelation function for stock returns.

Keywords: BASEL II Accord, Market Risk Model, VaR Model, Stochastic Process, Historical Simulation, Bootstrapping

JEL Classification: C15, C51, C88

1. Introduction

In all aspects of life, we often want to have the ability to predict the future with some certainty. The most reliable techniques used for forecasting the future are probability theory and statistical applications. In the banking and finance world, we often combine quantitative engineering techniques such as quantum physics, classical econometrics theory and operational research techniques to help with predictions of the movement of liquidity price in financial markets.

In the area of financial market risk, the main technique used to predict the movement of prices is based on the so-called Value at Risk (VaR) model. This model has been used widely in the banking industry after the completion of the BASEL I accord in 1996 and ever since it was publicly introduced by J. P. Morgan in 1997 (see Pearson (2002) for more details). The mathematical methods behind this model are based on the stochastic process using geometric Brownian motions which will be introduced in the next section. More detailed explanations of the BASEL accord and the market risk model can be found in Surapaitookorn (2007).

The main feature of the VaR model is the performance by calculating risk statistics using a bank's portfolio value with a correlation method like the parametric VaR or simulation methods using the historical or the Monte Carlo (MC) techniques. Two essential criteria for choosing the simulation method are the accuracy of the estimated output and the timing required to create such output (see Glasserman (2004) for more details). Simulation method requires fast computer machine and software technology as well as the best risk model specialist (often referred to as "financial engineer") who can handle the data applications, the stochastic process techniques, the simulation methods and who can understand the market risk drivers well. In this analysis, the historical simulation method will be used. It will be explained in Section 3.

One of the most important concepts in modelling market risk is the use of adequate future financial market price returns. Financial market data required for the estimation of VaR often come in terms of market prices and spread rates such as interest rates, foreign exchange (FX) rates, equity index rates, and implied volatilities. These prices and rates shall be referred to as 'market return rates'. Market price returns provided by the bank often come in daily high frequency date rates. In this quantitative analysis, the equity index rates will be used. They are introduced in Section 4.

In this paper, the estimation of future expected value of financial assets using financial market data will be analysed using the Autocorrelation (ACF) produced from the historical simulation of the VaR Model. The stochastic processes involved in the VaR model will be introduced in the next section. The aim is to see the effects of the simulation path generated from financial market data.

2. Stochastic Process for Market Risk VaR Model

There are a lot of VaR applications used for the calculation of risk factors in measuring financial market risk. Among them is McNeil et al. (2005). The purpose of using the VaR Model is to find the expected loss of a portfolio over some time horizon with a given level of probability at the maximum level using stochastic processes. In this section, we begin with the key theory behind the VaR model known as the stochastic process.

In the world of finance, the uncertainty of the future can be quantified using the so-called stochastic process. Many genius mathematicians and scientist like Gauss, Wiener, Levy, and Ito have contributed the most extraordinary mathematical techniques such as Gaussian distribution, Wiener process, Brownian motions, Levy processes, and Ito's Lemma respectively to the modern world of financial banking. See Mikosch (2006) for further details.

All of these techniques are based on the method of stochastic process which is proven to be the most powerful tool for forecasting the future of important financial applications in the banking sector. Stochastic process is used in simulation methods to obtain the maximum values of the required market risk statistics. The simplest stochastic process can be defined using mathematical notations as

$$Y_t = \{Y_t(\alpha), t \in T, \alpha \in A\} \quad (1)$$

where Y_t is the random variable which can be any market return rate; t refers to time interval frequency type of data like monthly, daily, hourly, second-by-second (or tick-by-tick); and A is defined as some space in the process.

A. Geometric Motion Processes

One of the most useful stochastic processes that are well used in financial engineering is the geometric motion. The word geometric often refers to exponential form of mathematics. There are three useful geometric motions used in the market risk banking system. We will define them in this section.

In the market risk system, Y_t can take different forms of stochastic process depending on the type of market data returns. Suppose that Y_t represents the stock price process at time t . In financial mathematics, the stock price process Y_t is often of the Geometric exponential form given by

$$Y_t = Y_0 e^{X_t}, \quad (2)$$

or

$$X_t = \log(Y_t), \quad (3)$$

where, X_t is defined as the rate of log-return of the stock price. The log-return process X_t can be modelled by a stochastic diffusion equation (SDE), for example an Ornstein-Uhlenbeck (OU) process, or a Levy process, for example Variance Gamma process.

The three geometric motions are defined as followed:

(i) Geometric Brownian Motion (GBM) Process

Geometric Brownian Motion (GBM) process is named after *Robert Brown*, the biologist whose research dates to the 1820s. It was Norbert Wiener (1923) who introduced Brownian motion to the mathematical world. If X_t is a Brownian with or without drift, then equation (3) is called Geometric Brownian Motion (GBM). The SDE for Y_t can be written as

$$dY_t = \mu Y_t dt + \sigma Y_t dZ_t, \quad (4)$$

where, Z_t is a standard Brownian Motion. In this case, Y_t has a Log-Normal distribution, since X_t has a normal distribution. Therefore, equation (4) is sometimes called *Log-Normal* model. In practice, we often use this geometric process for equity index rates or stock price returns.

(ii) Geometric Ornstein-Uhlenbeck (OU) Process

The Ornstein-Uhlenbeck (OU) Process is named after *Leonard Salomon Ornstein* and *George Eugene Uhlenbeck*, the Dutch physicist whose research dates to the 1920s. It is also known as the mean reverting stochastic process.

In any market risk model, the log-return X_t is an OU Process with dynamics SDE as follows

$$d \log(Y_t) = \alpha(\mu - \log(Y_t))dt + \sigma dZ_t, \quad (5)$$

or

$$dX_t = \alpha(\mu - X_t)dt + \sigma dZ_t, \quad (6)$$

where, $Z(t)$ is a standard Brownian Motion, α is the speed of reversion, μ is the (log) mean reversion level, and σ is the volatility. Then equation (3) may be called Geometric OU Process. In practice we often use this geometric process for interest rates.

(iii) Geometric Levy Process

The Levy Process is named after the French mathematician *Paul Levy*. This is the stochastic continuous time process based in probability theory. *Wiener Process* is one of the most well-known examples of the Levy Process which has to have stationary independent increments as well as starting the value at zero. Another good example of the Levy Process is the *Poisson Process*.

If X_t is a Levy Process, for example X_t is a Variance Gamma Process, then a Brownian Motion is a special Levy process, therefore (4) is also a Geometric Levy Process.

3. Historical Simulation and Bootstrapping Method

Historical simulation is one of the most well-known simulation methods used for market risk modelling. In historical simulation, rather than generating random numbers from a machine, the actual past history of market data are used so as to reflect realities like fat tails, persistence, and other common stylized facts of volatility. In historical simulation, we assume that past history is repeated and predict the future using the past data. The data used can be referred to as “non-overlapping” or “overlapping” data.

In this section, we firstly analyse these two types of data using the historical simulation. Bootstrapping method used to help with the correlation between paths as well as the limitation of data series will be introduced at the end of this section.

(i) Non-overlapping data

For *non-overlapping data*, consider the following:

- The 1st simulation run is to be constructed from $[1, \dots, n]$
- The 2nd simulation run is to be constructed from $[n+1, \dots, 2n]$
- and so on until.....
- The N^{th} / n (if N is a multiple of n) simulation run is to be constructed from $[N-n+1, \dots, N]$

In practice, more paths are generated in order to increase accuracy of the VaR. This is where the overlapping historical periods can be used.

(ii) Overlapping data

For *overlapping data*, consider a n day VaR calculation (e.g. 10 days). Obviously, the simulation must construct many paths composed of n observations. Assume that for each market risk factor there exist N historical observations. Currently the paths will be sampled as follows.

- Observations $[1, \dots, n]$ can be used to construct the first path.
- Observations $[2, \dots, (n+1)]$ can be used to construct the second path, and so on until
- The $N-n+1$ path is constructed from observations $[(N-n+1), \dots, N]$; so that $N-n+1$ ($> N/n$) simulation paths are generated from this overlapped data.

Let take a discrete example. Suppose that the stock prices for the last 3 days were 100, 105.13, 111.63 and today's stock price is 119.72, so that the daily stock returns for the past 3 days were 5%, 6% and 7%.

If this data is used to run the historical simulation and predict the stock price for tomorrow and 3 days from today, we have:

- At time node 1 for day 1 :
 $119.72 * \exp(0.05) = 125.86$, and
- At time node 2 for day 3 :
 $125.86 * \exp(0.06) * \exp(0.07) = 125.86 * \exp(0.13) = 143.33$

So the rate change data used for time node 1 is 5% and that for time node 2 is 13%.

For example, assume that there are 9 days of past data and we want to estimate the values for the next 3 days. From the data of days 1, 2 and 3, we can generate one path for future 3 days, from days 4, 5, 6 for the 2nd path, and from days 7, 8, 9 for the 3rd path. Now, we have 3 historical simulation paths, and at each future time node of day 1, 2 and 3, we have 3 samples each so that we can compute the statistics for the values.

In historical simulation, the process often allows for overlapping. This means that it is possible to use [days 1, 2, 3] for the 1st path, [days 2, 3, 4] for the 2nd path, [days 3, 4, 5] for the 3rd path, etc. This way, more simulation paths are

generated. The only disadvantage is that the paths are correlated. The purpose of this analysis is to see if any distortion results due to this correlation. To avoid the correlation problem, (for example, 3-day data can be randomly picked from 9 days of data) the bootstrapping methods may be introduced.

Bootstrapping Method

In historical simulation method, the main problem often found is the lack of historical data. To avoid this problem, we introduce a statistical technique known as *Bootstrapping*. Bootstrapping method assumes that daily returns are independent and can be used in our historical simulation. For example, we can randomly sample the rate of returns from the historical data and repeat many times to extract more information from the given data. The basic idea of the bootstrapping method is to take some real data and then create a large number of replicate data sets by sampling with replacement (with equal probability) from the original sample. Thus each new, replicate data set is a slightly perturbed version of the original one. Sampling with replacement means that every sample is returned to the data set after sampling. Therefore a particular data point from the original data set could appear multiple times in a given bootstrap sample.

4. Quantitative Analysis for Market Risk Model

In this section, we will outline the correlation and autocorrelation functions, proposed analysis and the analysis carried out for this paper. The US equity returns using the 10 assets of 1000 observations is used as financial market data. MATLAB software is chosen to help with constructing the market risk VaR model in order to generate the simulated paths using the three different methods of sampling from the historical data.

(i) Correlation and Autocorrelation Functions (ACF)

The linear correlation of two variables X , and Y can be defined as

$$\rho_{xy} = \frac{Cov(X,Y)}{[\text{var}[X]\text{var}[Y]]^{\frac{1}{2}}}, \text{ where } Cov(X,Y) = E[XY] - E[X]E[Y] \quad (7)$$

The correlation value lies between -1 to $+1$, where high positive or negative correlation is possible depending on the market movements. Let us consider the following three examples.

- In the *overlapping data*, if the market is keep moving up, so that the rate of return for a week is 0.1, 0.2, 0.4, 0.5, 0.7, then the one-day lagged data could be 0.2, 0.4, 0.5, 0.7, 0.8. Obviously they are highly positively correlated.
- If the market moves up and down, the correlation between the initial data and the lagged data could be negatively correlated. For example,

if the daily returns are 0.1, 0.0, 0.2, 0.1, 0.2, and the one-day lagged data are 0.0, 0.2, 0.1, 0.2, 0.0, then the correlation is highly negative.

- For *Non-overlapping real market data*, it is possible to obtain either positive or negative correlation values. If the rates were moving up last week and moving up again this week, then there is a positive correlation between last week and this week. If the rates were moving up last week and moving down this week, then there is a negative correlation between last week and this week.

The Auto-Correlation Function (ACF) or serial correlation occurs when we know that observations that are recorded sequentially over time are not independent of one another. For example, if a series $X = (X_t, X_{t-1})$, then X_t, X_{t-1} share information. This happened in overlapped historical periods. The ACF can be defined as

$$\rho_t = \frac{\text{Cov}(X_t, X_{t-1})}{[\text{var}(X_t) \text{var}(X_{t-1})]^{1/2}}, \quad (8)$$

(ii) Proposed Analysis

It is proposed that the autocorrelation effect can be quantified by constructing a simple stand-alone simulation of this process in MATLAB. In particular, historical market equity data will be used to construct the sample paths (in the manner outlined above). The terminal value of each sample path will be used to value a simple contract (for example a vanilla option) and the VaR of this contract calculated.

The simulation will run for calculating the VaR for a 10-day horizon. We will then analyse the resulting sample paths generated for ACF or serial autocorrelation. In addition to this, each of the simulated paths will be analysed for evidence of systematic bias in their correlations. For example, the correlations between paths and their autocorrelation functions can be determined.

This method will then be contrasted with the three further procedures for sampling from the historical data provided.

- (i) Selecting non-overlapping periods. That is, observations $(1..n)$ will be used to construct the first path, observations $(n+1)..,2n$ will be used to construct the second path, and so on until the N/n th path is constructed from observations. $N-n+1, \dots, N$.
- (ii) A sampling procedure based on the Bootstrap method will be utilised (where samples of n observations will be chosen repeatedly at random from the set of N observations).
- (iii) Each of these procedures can also be undertaken using simulated equity data generated from a GBM process as detailed in subsection 2.1.

(iii) Analysis

For procedures (i) and (ii) using the non-overlapping and the overlapping periods via the Bootstrapping method, there exists a new window N_w for each generated new path. The number of new windows can be considered as

$$N_w = \left(\frac{Tob - Horizon}{Tau} \right) + 1, \tag{9}$$

where *Tob* represents the number of total observations, *Horizon* represents the sample size within each window, and *Tau* represents number of increments for the next simulation path (i.e. 1 day, 2 days,...etc.). Note that when horizon equals *Tau*, we have a non-overlapping path.

For procedure (iii), by generating simulated data using a GBM process, we will perform the same analysis on this simulated data (which has the same moments as the actual market data) as we do for the historical data for comparison (so that the length of the simulated time series is identical to that of the historical data).

The GBM process can be considered as follows:

$$S_t = S_{t-1} \exp\left(\mu - \frac{\sigma^2}{2}\right)\Delta t + \sigma\sqrt{\Delta t}\varepsilon \tag{10}$$

where $S_t - S_{t-1} = \Delta S$ is the change in stock price, S is the small interval of time, Δt , and ε is a random drawing from a standardised normal distribution. A parameter μ represents the expected rate of return per unit time from the stock and a parameter σ represents the volatility of the stock price.

5. Empirical Results

For Bootstrapping methods, simulated values of variables mentioned in equation (9) for the historical paths give a mixture of negative and positive values. Using the two following examples, the results are shown below.

- (i) Taking the ‘Tau’, ‘Horizon’ to be 500, and the ‘Tob’ to be 2500, there are the non-overlapping historical periods. The number of windows is 5. The correlation values is displayed below:

1. Using the GBM process,

1	0.071894	0.090361	-0.020968	-0.0013005
0.071894	1	0.025145	0.026059	-0.056777
0.090361	0.025145	1	-0.011327	-0.02862
-0.020968	0.026059	-0.011327	1	-0.015625
-0.0013005	-0.056777	-0.02862	-0.015625	1

2. Using the Bootstrapping method,

1	0.041655	0.098772	0.033606	0.08557
0.041655	1	-0.015867	-0.016676	0.016893
0.098772	-0.015867	1	-0.026674	-0.008783
0.033606	-0.016676	-0.026674	1	0.010165
0.08557	0.016893	-0.008783	0.010165	1

By examining the correlation plots (not shown) for these outputs it was observed that the plots are not adequate. This shows that we should increase the number for variables *Tau* and *Horizon*. The second analysis is detailed below.

- (ii) Taking the *Tau* = 10, *Horizon* = 1500, and *Tob* = 2500 there are overlapping historical periods. The number of window is 101. The correlation values for the first 5 columns (i.e 5 windows) and the first 5 paths are displayed below.

1. Using GBM process,

1	0.015563	0.039555	0.038064	0.0037525
0.015563	1	0.015907	0.038379	0.034823
0.039555	0.015907	1	0.012851	0.038529
0.038064	0.038379	0.012851	1	0.0078944
0.0037525	0.034823	0.038529	0.0078944	1

2. Using the Bootstrapping method,

1	0.0081298	-0.013902	-0.032916	0.017007
0.0081298	1	0.013858	-0.011403	-0.034304
-0.013902	0.013858	1	0.021605	-0.010686
-0.032916	-0.011403	0.021605	1	0.022989
0.017007	-0.034304	-0.010686	0.022989	1

The correlation plots displayed in Figure 1 represent the simulation of Path 1. These plots suggest that there exists a serial correlation between paths around zero. The extreme point of value 1 comes from the diagonal matrix of 1 where each path is correlated to itself. This is just one example of one particular path; however, it is true for all historical simulation paths. At this stage, let us examine the autocorrelation (ACF) plots below to see whether our model is a suitable model for our data set.

By taking the lag at 50, the ACF plots for Path 1 show little correlation for the paths at all lag. Slightly more auto-correlation at lag 14 occurred in the second plot below. Again, the ACF plots for all paths are similar to the Path 1. This means that the historical simulation model may be of value in fitting with our financial stock data.

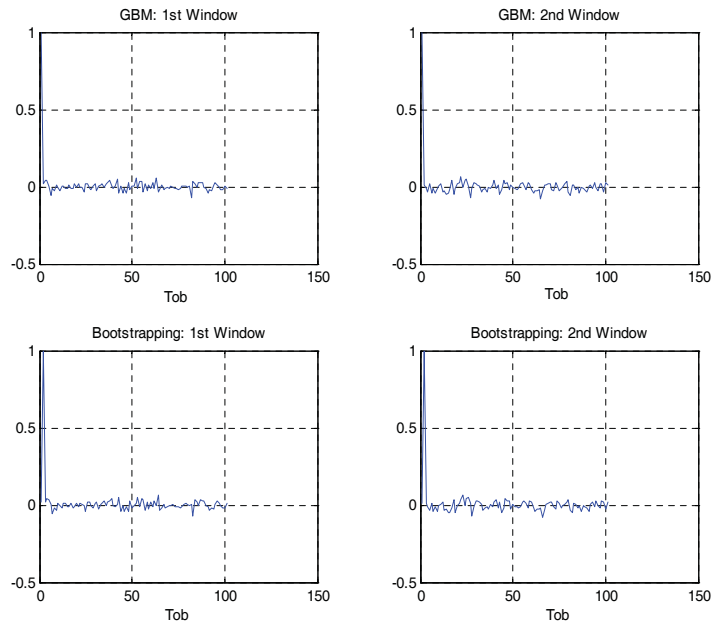


Figure 1: Correlation Plots using the GBM process and the Bootstrapping method

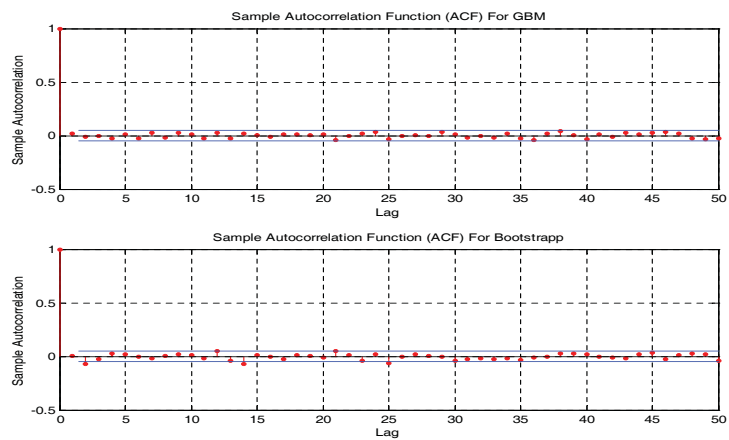


Figure 2: ACF Plots using the GBM process and the Bootstrapping method

6. Conclusion

These empirical results suggest that the GBM process and Bootstrapping methods provide similar output in terms of correlation values and the ACF plots. That is, the correlation values are not consistent. The non-overlapping paths should provide correlation values to be relatively larger than the overlap paths. Although this is not always the case, it seems that the values of correlation are very similar throughout (i.e. around zero) for all methods.

Author statement: Submitting author is Wantanee Surapaitookorn, Sasin Graduate Institute of Business Administration, Chulalongkorn University, Bangkok, Thailand, 2007. Tel: +66-2-2183893; E-mail: Wantanee.surapaitookorn@sasin.edu.

References

- Glasserman, P. (2004). *Monte Carlo Methods in Financial Engineering, Application of Mathematics: Stochastic Modelling and Applied Probability*, Springer.
- McNeil, A. J., Frey, R., and Embrechts P. (2005). *Quantitative risk management: Concepts, techniques, tools*. Princeton Series in Finance, Princeton University Press, Princeton and Oxford.
- Mikosch, T. (2006). *Elementary stochastic calculus with finance in view*. Advanced Series on Statistical Science & Applied Probability Vol. 6.
- Pearson, N. D. (2002). *Risk budgeting: Portfolio problem solving with value-at-Risk*. John Wiley & Sons.
- Saunders, A., Boudoukh, J., and Allen, L. (2003). *Understanding market, credit and operational risk: The value at risk approach*. Blackwell Publishing.
- Surapaitookorn, W. (2007, July-September). Quantitative review on the present and future of financial risk modelling and the role of BASEL II Accord. *SCMS Journal of Indian Management*, 4 (3).