Practical Value at Risk and Expected Shortfall Estimation for Securities Market

Yam-wing SIU

Department of Economics and Finance, the Hang Seng University of Hong Kong.

Abstract

Formally, Value at risk (VaR) measures the worst expected loss over a given horizon under normal market conditions at a given confidence level (Jorion, 2007). However, there are several weaknesses of VaR. Therefore, under Basel's latest revised market risk framework, it can be seen that Basel has shifted banks' capital regulatory requirements from value-at-risk to an expected shortfall approach. The choice of the holding period and confidence level are relatively subjective and they may depend on regulatory requirement. Very often, daily data are used to compute VaR and ES and scale up to required time horizon with square root of time adjustment. This gives rise to two important questions when we perform VaR and ES estimations: 1) whether non-overlapping and overlapping of data windows should be used for determining VaR and ES and 2) whether the values of VaR and ES (ie. "loss") should be interpreted as within i days or exactly on i^{th} day.

Preliminary numerical simulation results, after matching the descriptive statistics of Standard and Poor's 500 Index, show that, in determining the proportionality of the values of VaR and ES versus the holding period, using overlapping windows is just as fine as (if not better than) using non-overlapping windows. It is because the way we determine values of VaR in this study is simply locating □-percentile in the cumulative loss distribution for exactly on i^{th} day or maximum cumulative loss distribution for within i days respectively. There is no regression estimate of volatility.

Keywords: Basel; expected shortfall; market risk, numerical simulation; value at risk

1. Introduction

Value at risk (VaR) traces its roots to the many financial disasters of the early 1990s. Formally, it measures the worst expected loss over a given horizon under normal market conditions at a given confidence level (Jorion, 2007). To put it in layman's terms, it could be described as: Under normal market conditions, the most of a portfolio of $100 million can lose over a month is about $2.5 million (or 2.5%) at 95 percent confidence level. However, there are well-known weaknesses in the value-at-risk measure: 1) it does not give any information about potential losses when the value-at-risk has been exceeded; 2) it also fails to exhibit subadditivity if the tails of the distribution are heavy enough. An alternative or complement to value-at-risk is expected shortfall (ES). It measures the expected size of the loss given that the loss exceeds VaR. Under Basel's latest revised market risk framework, it can be seen that Basel has shifted banks' capital regulatory requirements from value-at-risk to an expected shortfall approach. The choice of the holding period, 1 month or 1 day, and confidence level are

1

relatively subjective and they depend on regulatory requirement. Very often, daily data are used to compute VaR and ES and scale up to required time horizon with square root of time adjustment (i.e. the proportionality of VaR and ES versus time is 0.5). This gives rise to two important questions when we perform VaR and ES estimations. The first question is whether non-overlapping and overlapping of data windows should be used for determining VaR and ES. The second question is on the time horizon. That is whether the values of VaR and ES (ie. "loss") should be interpreted as exactly on i^{th} day or within i days.

2. Data and Methodology

Financial data are available in time series form. Modeling and simulating financial time series would enable us to take better financial decisions in the areas of investment, risk management and others. Simulation is a very important topic and it requires generating random numbers. It is often useful to create a model using simulation. Usually this takes the form of generating a series of random observations (often based on a specific statistical distribution). One very important point when generating random observations is that it can match the stylized facts of the underlying statistical distribution.

In this study, daily closes of S&P500 Index (SPX) from 3 Jan 2000 to 31 Dec 2015 that converted to 4,173 daily log returns are used to measure VaR and ES. This set of results will be denoted as Actual. In addition, numerical simulation with 1,000,000 independent normally distributed random variables, □, that matches the standard deviation of actual SPX data, $\varepsilon \sim N(\mu = 0, \sigma = \sigma_{Actual})$, (denoted NS hereafter) is used to generate daily log returns to analyse impact of 1) non-overlapping and overlapping windows; and 2) exactly on i^{th} day and within i days on the estimations of VaR and ES.

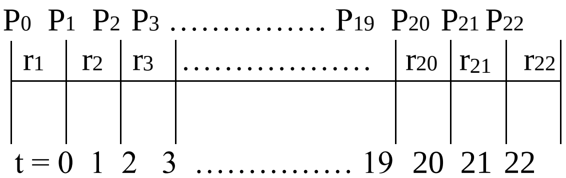
2.1 Non-overlapping versus Overlapping Windows

With 1,000,000 data generated, window size ranging from 1 day to 22 days[1] can be constructed. The corresponding numbers of windows are shown in Table 1.

Table 1. Window size and numbers of windows used in this study

Window size in day(s)	**1**	**2**	□□	**5**	□□	**10**	□□	**22**
# of non-overlapping windows	1,000,000	500,000	□□	200,000	□□	100,000	□□	45,454
# of overlapping windows	1,000,000	999,999	□□	999,996	□□	999,991	□□	999,979

Figure 1: Timeline for daily price and log return of simulated data



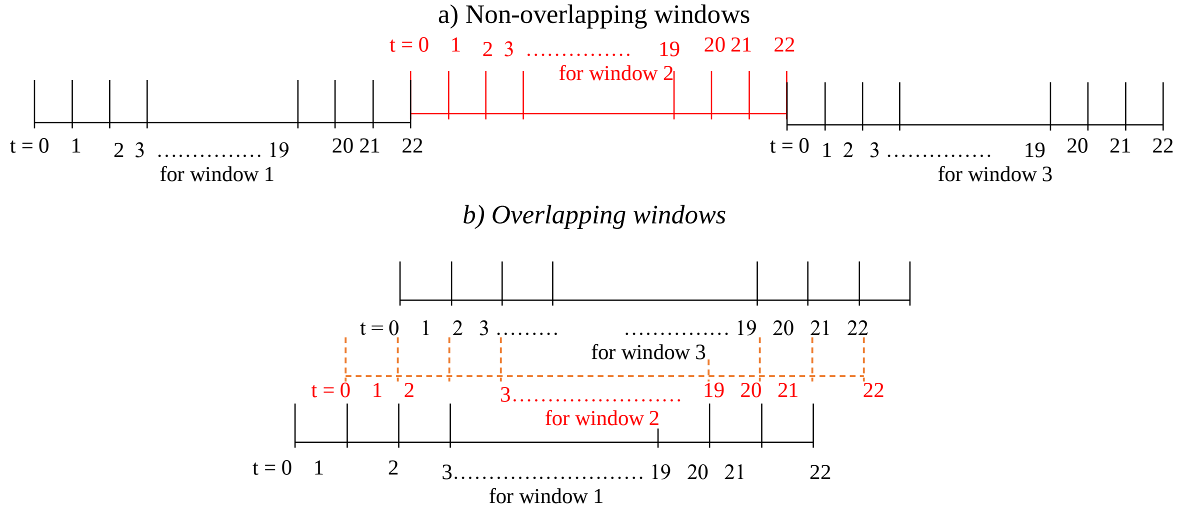
Suppose today is t = 0 and the price is P_0, tomorrow is t = 1 and the price is P_1, and r_1 is the daily log return of tomorrow's price with respect to today's one. With 1,000,000 daily log

[1] 22 days is selected to match one month's time horizon.

returns simulated, we can construct data window of various sizes. Let us use a window size of 22 as an example. Graphically, each window can be represented in figure 1.

With the window size fixed, next we need to decide whether the windows are non-overlapping or overlapping in this study. Non-overlapping and overlapping windows can be illustrated in figure 2 whereas the numbers of non-overlapping windows created by 1,000,000 daily log returns are much smaller than those with overlapping windows especially as the window size increases. In time series analysis, many researchers such as Christensen and Prabhala (1998) argue that using overlapping samples will causal serial correlation and lead to underestimating the standard errors for the coefficients on regression analysis. Whether using overlapping windows in this study will have implication on the proportionality of VaR and ES versus time and the values of VaR and ES will be discussed in §3.

Figure 2: Illustration of non-overlapping and overlapping windows



2.2 Value at Risk Estimation exactly on ith day

For value at risk estimation exactly on, say, 22nd day, we are inspecting the price, P_{22}, at exactly on i = 22nd day and see whether it exceeds a pre-determined value. Mathematically, the return (+ve for gain and -ve for loss) on 22nd day with respect to today is:

$$\frac{P_{22} - P_0}{P_0} \approx \ln\left(\frac{P_{22}}{P_0}\right) = r_1 + r_2 + \cdots + r_{22}$$

where $r_i = ln\ ({r_i}^{r}{}_{-i_1})$ = daily log return. The loss can also be described as Cumulative Loss (or cumulative return[1]) exactly on i = 22nd day.

Repeat the process for different window and we can compute 45,454 such cumulative losses in the case of non-overlapping windows of 22 days. The value at risk with (1- □) confidence level is the □-percentile where □ = 5%[3] as shown in figure 3.

3

Figure 3: Value at risk determined from distribution of cumulative returns

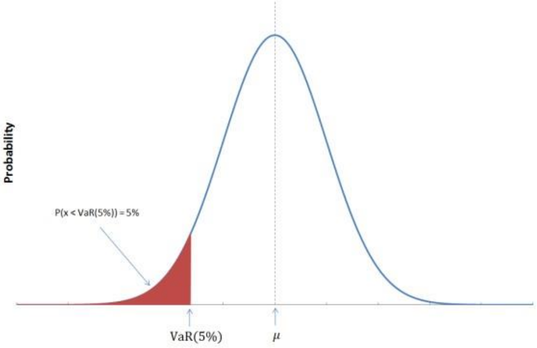
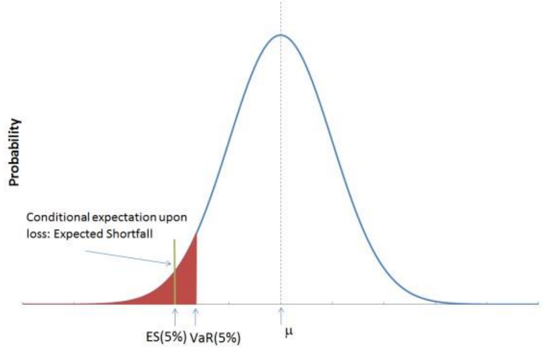


Figure 4: Expected shortfall conditional on exceeding value at risk



2.3 Value at Risk Estimation within i days

For value at risk estimation within, say, 22 days, we are inspecting the prices P_1 to P_{22} over and within all 22 days and see whether at any one day it exceeds a pre-determined value. Mathematically, the maximum cumulative loss within 22 days with respect to today is:

$$\text{Min}\left\{\frac{P_1 - P_0}{}, \frac{P_2 - P_0}{}, \cdots, \frac{P_{22} - P_0}{}\right\}_0 \quad 0$$

Note that Minimum function is used here when the sign of the returns is considered.

[1] Cumulative return is a more general term as it could be a gain on 22nd day. [3] Confidence level of (1-□) is equivalent to significance level of □.

P P P_0

$$\approx \mathrm{Min}\ \{ln\left(\frac{P_1}{P_0}\right), ln\left(\frac{P_2}{P_0}\right), \cdots, ln\left(\frac{P_{22}}{P_0}\right)\}$$

$$= \mathrm{Min}\{r_1, (r_1 + r_2), \cdots, (r_1 + r_2 + \cdots + r_{22})\}$$

Again repeat the process for different window and we can compute 45,454 such maximum cumulative losses in the case of non-overlapping windows of 22 days. And the value at risk can be computed as described before.

2.4 Expected Shortfall Estimation exactly on i^{th} day and within i days

An alternative or complement to VaR is expected shortfall (ES). It measures the expected size of the loss given that the loss exceeds VaR. Expected shortfall has several advantages over VaR that it is a coherent risk measure and it is a better risk measure in terms of tail risk, Yamai and Yoshiba (2002, 2005). Graphically, the expected shortfall can be represented in figure 4 where the expected shortfall is the average of losses conditional on exceeding the value at risk. When the losses are cumulative losses exactly on i^{th} day, the average of these losses exceeding a prescribed value at risk level is expected shortfall exactly on i^{th} day. When the losses are maximum cumulative losses within i days, the average of these losses exceeding a prescribed value at risk level is expected shortfall within i days.

3. Results and Discussions

As described in §2, numerical simulation together with the actual data of SPX are used in this study. Let us inspect the results for the simulation of 1,000,000 random variables with normal distribution that matches the standard deviations of SPX data (NS), $\varepsilon \sim N(\mu = 0, \sigma =$

σ_{Actual}). The main difference between the two datasets are that SPX is negatively skewed with skewness of around -0.188 and fat-tailed with kurtosis of 8.424 as shown in figure 5 where the histogram of the simulation (— black line with dots) together with that of actual SPX data (— red line and gray-shaded) are plotted. Other statistics of SPX and the simulation are shown in table 2.

Figure 5: Histogram of simulation and actual data of SPX

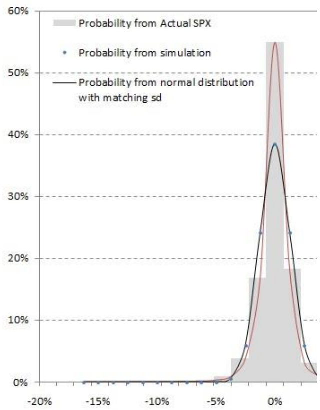


Table 2: Statistics of SPX data and simulation

Statistics	SPX (Actual)	Simulation (NS)
# of data	4,173	1,000,000
Skew	-0.188	-0.0027
Kurt	8.424	0.00177
Max	10.96%	5.87%
Min	-9.47%	-6.65%
Daily Mean	0.0081%	0.0000%
Daily SD	1.244%	1.244%
Annualized Mean	2.1166%	0.0000%
Annualized S.D.	20.054%	20.054%

Following the procedures described in §2, the value at risk and expected shortfall for time horizon exactly on i^{th} day and within i days for significance levels of 1%, 2.5% and 5% can be computed. Significance levels of 1%, 2.5% and 5% are chosen because banks are required to report VaR at 1% and ES at 2.5% significance levels for holding period of i = 10 according to Basel Accord. Also significance level of 5% is commonly used by risk managers of financial institutions. With value at risk and expected shortfall for i = 1 to 22 days determined, we can plot VaR and ES versus i in a log-log graph. The purpose of plotting the graph in log-log scale is to check whether VaR and ES grows proportionally with square root of time[2] or other exponential power.

3.1 Non-overlapping windows versus overlapping windows

A typical set of results look like those presented in figure 6 where results for NS simulation for value at risk and expected shortfall exactly on i^{th} day with non-overlapping windows are listed and plotted in log-log scale.

As mentioned, VaR at 1% and ES at 2.5% significance levels for holding period of i = 10 are of importance. From the best fit lines of figure 6, we can see that the slopes are around 0.50 for VaR and ES at all three significance levels. It means VaR and ES grow with proportionally with the square of time when they are determined exactly on i^{th} day with non-overlapping windows.

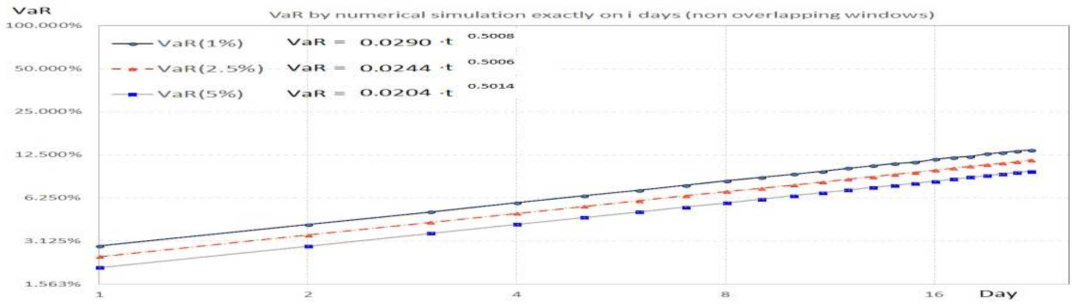
As described in §2, for simulation (NS), VaR and ES can be determined in four ways: non-overlapping versus overlapping, and exactly on i^{th} day and within i days. Table 3 summarize

[2] The slope of the log-log graph will be 0.5 if VaR or ES grows proportionally with square root of time.

the values of the slopes of VaR and ES at various significance levels and VaR at 1% and ES at 2.5% significance levels for holding period of i = 10.

Figure 6:a) VaR and b) ES for simulation NS for exactly on i^{th} day (non-overlapping windows)

a) VaR for simulation NS exactly on i^{th} day for significance levels of 1%, 2.5% and 5%



b) ES for simulation NS exactly on i^{th} day for significance levels of 1%, 2.5% and 5%

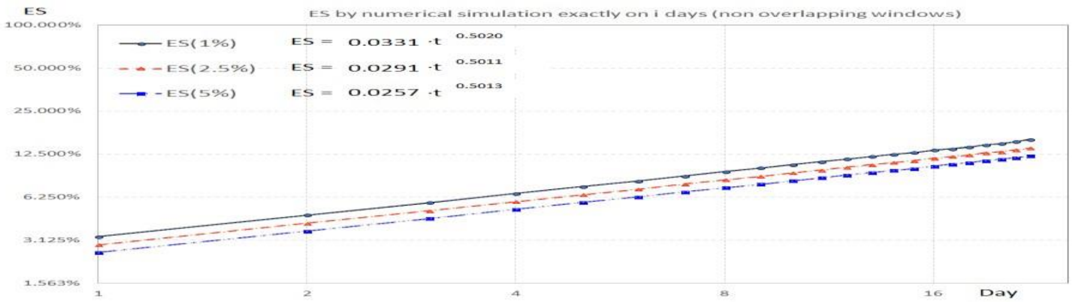


Table 3 - Slopes of VaR and ES versus i for different □, VaR(1%) and ES(2.5%) for i = 10 for NS simulation

		Exactly on ith day			Within i days		
Non Overlapping	Slopes of VaR (1%;2.5%;5%)	0.50	0.50	0.50	0.52	0.53	0.54
	Slopes of ES (1%;2.5%;5%)	0.50	0.50	0.50	0.52	0.52	0.53
	VaR(1%) / ES(2.5%) with i = 10	9.1693%	/	9.2383%	9.5834%	/	9.6377%
Overlapping	Slopes of VaR (1%;2.5%;5%)	0.50	0.50	0.50	0.52	0.53	0.54
	Slopes of ES (1%;2.5%;5%)	0.50	0.50	0.50	0.52	0.52	0.53
	VaR(1%) / ES(2.5%) with i = 10	9.1750%	/	9.2246%	9.5769%	/	9.6234%

It can be seen from table 3 that the results are practically the same between non-overlapping windows and overlapping windows. For VaR and ES exactly on i^{th} day, the slopes are all at 0.50 no matter they are determined with non-overlapping and overlapping windows. For VaR and ES within i days, the slopes are around 0.52-0.54, again, no matter they are determined with non-overlapping and overlapping windows. It is because there is no regression for coefficients in this study. Instead, the way we determine values of VaR or ES in this study is simply locating □-percentile in the Cumulative Loss for exactly on i^{th} day or Maximum Cumulative Loss for within i days.

Table 4 - Slopes of VaR and ES versus i for different □, VaR(1%) and ES(2.5%) for i = 10 for actual data

		Exactly on ith day			Within i days		
Non Overlapping	Slopes of VaR (1%;2.5%;5%)	0.53	0.49	0.49	0.60	0.58	0.59
	Slopes of ES (1%;2.5%;5%)	0.51	0.49	0.48	0.55	0.56	0.56
	VaR(1%) / ES(2.5%) with i = 10	9.6161%	/	10.3875%	16.0865%	/	14.4819%
Overlapping	Slopes of VaR (1%;2.5%;5%)	0.49	0.48	0.47	0.58	0.57	0.58
	Slopes of ES (1%;2.5%;5%)	0.49	0.48	0.48	0.55	0.57	0.57
	VaR(1%) / ES(2.5%) with i = 10	10.3767%	/	11.0791%	13.2269%	/	13.7550%

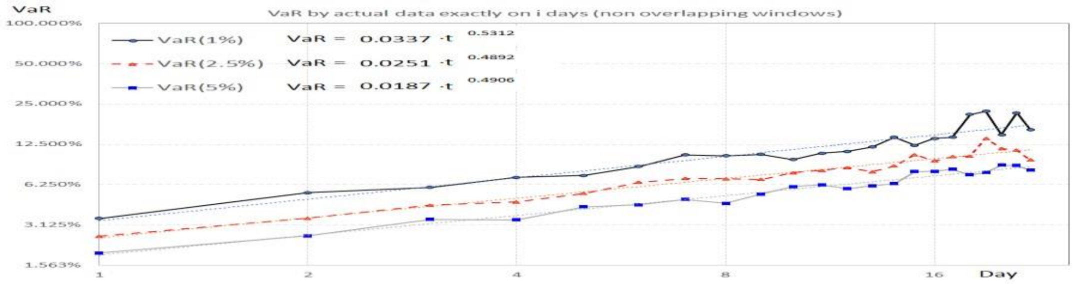
Having seen the results for NS simulation, let us inspect the results for actual data. It can be seen from table 4 that, similar to NS simulation, the slopes of VaR and ES are generally higher for within i days than those for exactly on i^{th} day. However, unlike the results for simulation NS, the slopes are not the same between non-overlapping windows and overlapping windows. But it must be emphasized that the ranges of results for overlapping windows are much closer than those for non-overlapping windows. For example, for results for exactly on i^{th} day, the slopes of VaR for overlapping windows for □ = 1%, 2.5% and 5% are 0.49, 0.48 and 0.47 respectively while they are much wider at 0.53, 0.49 and 0.49 respectively for non-overlapping windows. The reason that the ranges of results for non-overlapping windows are wider and that the results are less "tight" than those of NS simulation could be due to the small sample sizes as shown in table 5. For example, even for overlapping windows, there are only 4,173 windows for window size of 1 day and 4,152 windows for window size of 22 days. These numbers will drop to 4,173 and only 189 respectively for non-overlapping windows. Hence, the lines for VaR and ES diagram in figure 7 is more rugged towards to larger window size. Thus, given a limited number of daily closes of actual data and the results obtained from NS simulation, it is recommended to use overlapping windows to determine the proportionality of VaR and ES versus time.

Table 5. Window size and numbers of windows used in this study for actual data

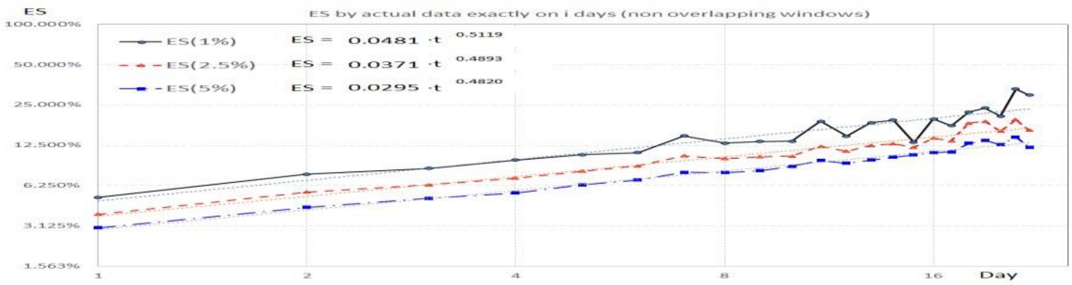
Window size in day(s)	**1**	**2**	□□	**5**	□□	**10**	□□	**22**
# of non-overlapping windows	4,173	2,086	□□	834	□□	417	□□	189
# of overlapping windows	4,173	4,172	□□	4,169	□□	4,164	□□	4,152

Figure 7:a) VaR and b) ES for actual data for exactly on i^{th} day (non-overlapping windows)

a) VaR for actual data exactly on i^{th} day for significance levels of 1%, 2.5% and 5%



b) ES for actual data exactly on ith day for significance levels of 1%, 2.5% and 5%



3.2 Exactly on ith day versus Within i days

A closer inspection of the results in tables 3 and 4 indicates that both the magnitudes and the proportionality of VaR and ES versus time is higher for results of within i days than those for exactly on ith day. Therefore, from the risk management point of view, it is recommended to adopt within i days for the calculation of capital regulatory requirement under Basel as it is more prudent. Notably, even in NS simulation where the data are independent, the proportionality for VaR and ES versus time for within i days is not 0.5. Thus, risk managers are suggested not to scale up VaR and ES to the required time horizon with square root of time adjustment. Instead, it should be determined/estimated separately or use a higher proportionality constant.

4. Conclusions

In this study we have obtained VaR and ES under a simulated distribution that matches the standard deviation of actual daily closes of S&P500 Index (SPX) data from 3 Jan 2000 to 31 Dec 2015. We analysed the impact of 1) non-overlapping and overlapping windows; and 2) within i days and exactly on ith day on the estimations of VaR and ES. Preliminary results show that, in determining the proportionality of the values of VaR and ES versus the holding period, using overlapping windows is just as fine as (if not better than) using non-overlapping windows. It does not suffer from the serial correlation of Christensen and Prabhala (1998) because the way we determine values of VaR or ES in this study is simply locating □-percentile in the Cumulative Loss for exactly on ith day or Maximum Cumulative Loss for within i days.

It is because the way we determine values of VaR in this study is simply locating □-percentile in the cumulative loss distribution for exactly on ith day or maximum cumulative loss distribution for within i days respectively. There is no regression estimate of volatility. Also, from the risk management point of view, it is recommended to adopt within i days for the calculation of capital regulatory requirement under Basel as it is more prudent. For actual data of SPX, VaR at □=1% over 10 holding days = 13.23% while ES at □ =1% over 10 holding days = 13.76%. It translates to a slight increase of 4% in regulatory capital requirement when Basel shifted to an expected shortfall (ES) measure from a value-at-risk (VaR) measure.

References

Christensen, B.J. and Prabhala, N.R. (1998). The relation between implied and realized volatility. Journal of Financial Economics, 50, pp.125-150.

Jorion, P. (2007). Value at Risk – The New Benchmark for Managing Financial Risk, 3rd ed., McGraw-Hill Education – Europe.

Yamai, Y. and Yoshiba, T. (2002). On the validity of value-at-risk: Comparative analysis with expected shortfall. Monetary and Economic Studies, 20 (1), pp.57-86.

Yamai, Y. and Yoshiba, T. (2005). Value-at-risk versus expected shortfall: A practical perspective. Journal of Banking and Finance, 29 (2), pp.997-1015.