Chapter 2

Literature Review and Theoretical Background

2.1 Value at Risk Model (VaR) and Extreme Value Theory Approach (EVT)

This study investigates Value at Risk of gold price return using Extreme Value Theory. Value at Risk (VaR) is a method of assessing risk that uses standard statistical technique routinely used in other technical fields. Formally, Value at Risk measures the worst expected loss over a given horizon under normal market conditions at a given confidence level. Based on the firm scientific foundations, Value at Risk provides users with a summary measure of market risk. For instance, a bank might say that the daily Value at Risk of its trading portfolio is $35 million at the 99 percent confidence level. In the other words, there is only 1 chance in a 100, under normal market conditions, for a loss greater than $35 million to occur. This single number summarizes the bank's exposure to market risk as well as the probability of an adverse move. (Jorion, 1997: 22)

2.1.1 Measuring Value at Risk (VaR)

Butler (1999) stated that there are three common methods of computing Value at Risk: Variance covariance, Historical, Stochastic or Monte Carlo simulation

1. Variance covariance

The easiest method is the variance covariance method. Since this approach involves using "published" information on volatility and correlation and then constructing an internal weighting matrix. The process is probably the most popular because it is simple to construct. A bank wishing to calculate its Value at Risk must simple construct a weight matrix, and then obtain the volatility and correlation data from JP Morgan, which publishes and regularly updates the data on the Internet (JP Morgan Risk Metrics). There are a few limitations, however. JP Morgan's approach is not suitable for options. Also, there is the assumption that relation between assets i.e. correlation coefficients are stable. This, of course, may not be true, particularly when there is a major upheaval like a stock market crash. Additionally, JP

Morgan's approach places an over reliance on the normal distribution curve. Very occasionally, asset returns are not normal and so JP Morgan's approach may give a biased result. That said, the RiskMetrics approach is intuitively appealing and widely used by risk managers. Rather than going for a state-of-the-art complex system, it sometimes makes sense to go for a model that is intuitively understandable and easy to implement. A slight compromise on precision can often lead to a substantial reduction in operational risk.

2. The historical method

Like the JP Morgan approach, the historical method is intuitively simple to understand. Risk managers simply keep a historical record of daily profit and loss within the portfolio and then calculate the fifth percentile for 95 percent or 1 percent for 99 percent VaR, As well as being simple, the historical approach is realistic. The same cannot be said for RiskMetrics because the volatilities and correlations are not actual figures, but estimates based on average over a specific time. In extreme situations, these averages may not be hold, so the RiskMetrics approach may not give a realistic result. The historical method is based on actual results and if, during the historical period, major market events happened, these would be picked up accurately by the historical system. A second advantage of the historical method is that it does not require "mapping". When constructing a weighting matrix for RiskMetrics, the instruments may not neatly fit into the model devised by JP Morgan. Therefore, awkward instruments must be broken down and "mapped" onto standard vertices published by RiskMetrics. This process, as well will see later, can be computationally cumbersome, and very often certain assumptions are necessary. With the historical approach, no mapping is necessary and there is no need to make assumption.

The main weakness of the historical approach is that it is unsuitable if the weights of the portfolio change, that is to say if the portfolio composition changes overtime. To overcome this, the historical approach can be augmented by the historical simulation approach. Here we use the current portfolio composition, but use historical market data. If a current portfolio consists of 70 percent Asset One and Asset Two for, say, the past 1000 days, and for each day calculate the value of the portfolio, keeping the current weights of 30 percent and 70 percent constant.

Obviously, this is more time consuming and, in the case of large portfolios, very demanding of computer resources.

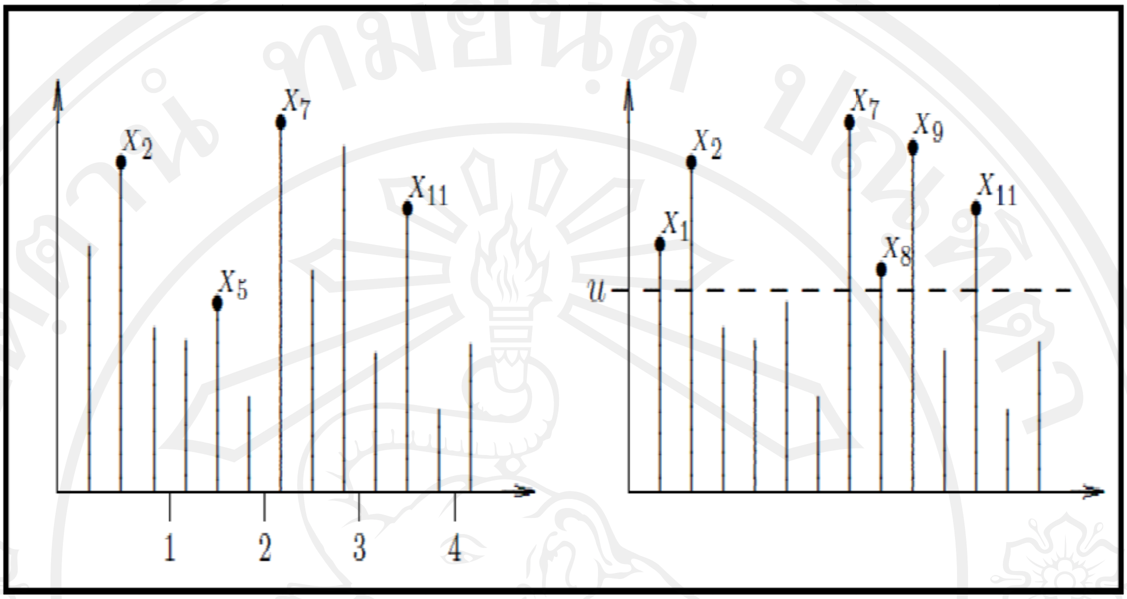
3. The stochastic or Monte Carlo simulation method

This procedure involves asking a computer to generate a series of share prices using a "random walk" approach. The procedure can be quite complex and, although in terms of precision it is the most effective, it suffers from the fact that it is time consuming and, like historical simulation, demanding of computer resources. Where the portfolio is enormous, we can end up with hardware constraints. Stochastic simulation is probably very appropriate when a portfolio is complex, particularly where it contains many options. Banks writing exotic options would have considerable difficulty in calculating VaR using RiskMetrics and, although the historical market prices were unique. In such cases, the historical path would not be representative of all possibilities and, therefore, prices generated from a stochastic system would give a more realistic result.

2.1.2 Extreme Value Theory Approach (EVT)

Extreme Value Theory models only the tail of the return distribution rather than the entire distribution with the extreme events. So, this approach can potentially perform better than other approaches in terms of forecasting unexpected extreme changes.

Modeling Extreme Value Theory, there are two ways if identified extremes in data. This paper is considered a random variable which may represent daily losses or returns. The first approach considers the maximum (or minimum) the variable takes in periods, for example months or years. In the left panel of figure 2, the observations: X_2, X_5, X_7 and X_{11} represent the block maxima for four periods. These selected observations are called "Block" (or per-period) maxima. The second approach focuses on the largest value variable over some high threshold. In the right panel of figure 2 shows the observations: X_1, X_2, X_7, X_8, X_9 and X_{11} are all exceed the threshold u, which lead to extreme events.



Source: Author's review

Figure 2.1 Block-maxima (left panel) and excess over a threshold u (right panel)

2.1.2.1 Block Maxima or Generalized Extreme Value Distribution (GEV)

This approach is the one the studies the limiting distributions of the sample extreme, which is presented under a single parameterization. . In this case, extreme movements in the left tail of the distribution can be characterized by the negative numbers (Jiahn-Bang Jang, 2007)

Given that X_i be the negative of the i th daily return of the gold prices between day i and day i-1. Define

$$X_i = -(\ln P_i - \ln P_{i-1})$$

where P_i and P_{i-1} are the gold prices of day i and day i-1 . Then, suppose that X_1, X_2, …, X_n be iid random variables with an unknown cumulative distribution function (CDF) $F(x) = \Pr(X_i \leq x)$. Extreme values are defined as maxima (or minimum) of the n independently and identically distributed random variables X_1, X_2, …, X_n.

Let X_n be the maximum negative side movements in the daily gold prices returns, that is

$$X_n = \max(X_1, X_2, \ldots, X_n)$$

Since the extreme loss movements are the focus of this study, the exact distribution of X_n can be written as

$$\Pr(X_n \le a) = \Pr(X_1 \le a, X_2 \le a, \ldots, X_n \le a)$$
$$= \prod_{i=1}^{n} F(a)$$
$$= F^n(a)$$

In practice the parent distribution F is usually unknown or not precisely known. The empirical estimation of the distribution F^n (a) is poor in this case. Fisher and Tippet (1928) derived the asymptotic distribution of F^n (a). Suppose μ_n and σ_n are sequences of real number location and scale measures of the maximum statistic X_n. Then the standardized maximum statistic,

$$Z_n^* = \left(\frac{X_n - \mu_n}{\sigma_n}\right) \quad (1)$$

Converges to z = (x - μ) / σ which has one of three forms of non-degenerate distribution families such as

$$H(z) = \exp\{-\exp[-z]\}, -\infty < z < \infty$$
$$H(z) = \exp\{-Z^{-1/\xi}\}, z > 0$$
$$= 0, \text{else}$$
$$H(z) = \exp\{-[Z]^{-1/\xi}\}, z > 0$$
$$= 1, \text{else} \quad (2)$$

These forms go under the names of Gumbel, Frechet, and Weibull respectively. While μ and σ are the mean return and volatility of the extreme values x, ξ is the shape parameter or called 1/ ξ the tail index of the extreme statistic distribution.

Embrechts et al. (1997) describe GEV distribution in detail, which fundamental types of extreme value distributions are defined by ξ:

1. If $\xi = 0$, the distribution is called the Gumbel distribution. In this case, the distribution spreads out along the entire real axis.

2. If $\xi > 0$, the distribution is called the Fréchet distribution. In this case, the distribution has a lower bound

3. If $\xi < 0$, the distribution is called the Weibull distribution. In this case, the distribution has an upper bound.

The Fisher and Tippett (1928) theorem suggests that the asymptotic distribution of the maxima belongs to one of the three distributions above, regardless of the original distribution of the observed data. Random variables fall into one of three tails shapes, fat, normal, and thin, depending on the various properties of the distribution. Thus, the tails of distributions are:

i. thin, that is, the tails are truncated,

ii. normal. In this case, the tails have an exponential shape,

iii. fat. The tails follow a power law.

Embrechts et al. (1997) proposed a generalized extreme value (GEV) distribution which included those three types and can be used for the case stationary GARCH processes. GEV distribution has the following form

$$H_{\xi}(X;\mu,\sigma) = \exp\{-\frac{\exp[-x-\mu)}{\sigma}]\}, -\infty \le x \le \infty;\ \xi = 0$$

$$= \exp\{-[1+\xi(x-\mu)/\sigma]^{-\frac{1}{\xi}}, 1+\frac{\xi(x-\mu)}{\sigma} > 0;\ \xi \neq 0 \qquad (3)$$

Then, suppose that block maxima $B_1, B_2, \ldots, B_k$ are independent variables from a GEV distribution, the log-likelihood function for the GEV, under the case of $\xi \neq 0$, can be given as

$$\text{In } L = -k \text{ In } \sigma - (1+\frac{1}{\xi})\sum_{i=1}^{k} \text{In}\{1+\ \xi\frac{(B_i-\mu)}{\sigma}\} - \sum_{i=1}^{k}\{1+\xi\frac{(B_i-\mu)}{\sigma}\}^{-1/\xi} \qquad (4)$$

For the Gumbel type GEV form, the log-likelihood function can be written as

$$\text{In } L = -k \text{ In } \sigma - \sum_{i=1}^{k}\frac{(B_i-\mu)}{\sigma} - \sum_{i=1}^{k}\exp\{\frac{(B_i-\mu)}{\sigma}\} \qquad (5)$$

As Smith (1985) suggested that, for $\xi > 0.5$, the maximum likelihood estimators, for ξ, μ, and σ, satisfy the regular conditions and therefore having asymptotic and consistent properties. The number of blocks, k and the block size form a crucial tradeoff between variance and bias of parameters estimation.

2.1.2.2 Peak over threshold or Generalized Pareto Distribution (GPD)

Fitting models with more data is better than less estimation, so Peaks over thresholds (POT) method utilizes data over a specified threshold. Jiahn-Bang Jang (2007) defined the excess distribution as

$$F_h(x) = \Pr(X-h < x \mid X > h) = \frac{[F(x+h)-F(h)]}{1-F(h)} \qquad (6)$$

where h is the threshold and F is an unknown distribution such that the CDF of the maxima will converge to a GEV type distribution. For large value of threshold h , there exists a function $\tau(h) > 0$ such that the excess distribution of equation (6) will approximated by the generalized Pareto distribution (GPD) with the following form

$$H_{\xi,\tau(h)}(x) = 1 - \exp\left(-\frac{x}{\tau(h)}\right), \xi = 0$$

$$= 1 - (1 + \frac{\xi x}{\tau(h)})^{-1/\xi}, \xi \neq 0 \qquad (7)$$

where $x > 0$ for the case of 0 , and $\xi \geq 0 \geq x$, and $0 \leq x \leq \tau(h) / \xi$ for the case of $\xi < 0$. Define $X_1, X_2, \ldots, X_k$ as the extreme values which are positive values after subtracting threshold h .

For large value of h, $X_1, X_2, \ldots, X_k$ is a random sample from a GPD, so the unknown parameters ξ and $\tau(h)$ can be estimated with maximum likelihood estimation on GPD log-likelihood function.

Based on equation (6) and GPD distribution, the unknown distribution F can be derived as

$$F(y) = \big(1 - F(h)\big)H_{\xi,\tau(h)}(x) + F(h) \qquad (8)$$

where $y = h + x$. F(h) can be estimated with non-parametric empirical estimator

$$\hat{F}(h) = \frac{n-k}{n}$$

where k is the number of extreme values exceed the threshold h . Therefore the estimator of (8) is

$$\hat{F}(h) = \left(1 - \hat{F}(h)\right)\hat{H}\left(x; \hat{\xi}, \hat{\tau}(h)\right) + \hat{F}(h) \qquad (9)$$

where $\hat{\xi}$ and $\hat{\tau}(h)$ are mle of GPD log-likelihood. High quantile VaR and expected shortfall can be computed using (9). First, define $F(VaR_q) = q$ as the probability of distribution function up to q^{th} quantile VaR_q, therefore,

$$\widehat{VaR}_q = \hat{F}^{-1}(q) = h + \hat{\tau}(h)\left\{\left[\left(\frac{n}{k}\right)(1-q)\right]^{-\hat{\xi}} - 1\right\}/\hat{\xi} \qquad (10)$$

Next, given that VaR_q is exceeded, define the expected loss size, expected shortfall (ES), as

$$ES_q = E\big(X \mid X > VaR_q\big) = VaR_q + E(X - VaR_q \mid X > VaR_q \qquad (11)$$

From (10), $\widehat{ES}_q$ can be computed using $\widehat{VaR}_q$ and the estimated mean excess function of GPD distribution. Therefore,

$$\widehat{ES}_q = \frac{\widehat{VaR}_q}{1-\hat{\xi}} + (\hat{\tau}(h) - \hat{\xi}h)/(1-\hat{\xi})$$

2.2 Literature Review

2.2.1 Value at Risk (VaR)

In terms of evaluation in Value at Risk, Baran and Witzany (2010) applied EVT in estimating low quantiles of P/L distribution and the results were compared to common VaR methodologies. The result confirms that EVT-GARCH is superior to other methods. Gençay and Selçuk (2004) investigated the Extreme Value Theory to generate Value at Risk to estimate and study the tail forecasts of daily returns for stress testing. Then, Bali (2003) studied how to estimate volatility and Value at Risk by an extreme value approach and determines the type of asymptotic distribution for the extreme changes in U.S. Treasury yields. In this paper, the thin-tailed Gumbel and exponential distribution is not as good as the fat-tailed Fréchet and Pareto distributions.

In the analysis of Stelios Bekiros and Dimitris Georgoutsos (2003) conducted a comparative evaluation of the predictive performance of various Value at Risk (VaR) models. Both estimation techniques are based on limit results for the excess distribution over high thresholds and block maxima respectively. The results we report reinforce previous ones according to which some "traditional" methods might yield similar results at conventional confidence levels but at very high ones the EVT methodology produces the most accurate forecasts of extreme losses. Moreover, Yamai and Yoshiba (2002) investigated the comparison between Value-at-Risk (VaR) and expected shortfall under market stress. The paper found that First, VaR and expected shortfall may underestimate the risk of securities with fat-tailed properties and a high potential for large losses. Second, VaR and expected shortfall may both disregard the tail dependence of asset returns. Third, expected shortfall has less of a problem in disregarding the fat tails and the tail dependence than VaR does.

2.2.2 Extreme Value Theory Approach (EVT)

Extreme Value Theory is most used in evaluation of Value at Risk in Financial markets. Martin Odening and Jan Hinrichs[1] (2010), who focused on Using Extreme Value Theory to estimate Value-at-Risk, examined problems that may occur when conventional Value at Risk (VaR) estimators are used to quantify market risks in an agricultural context. For example, standard Value at Risk methods, such as variance-covariance method or historical simulation, can fail when the return distribution is fat tailed. This problem is aggravated when long-term Value at Risk forecasts is desired. Extreme Value Theory is therefore proposed to overcome these problems. For a stock market study, Vladimir Djakovic, Goran Andjelic, and Jelena Borocki (2010**)** investigated the performance of Extreme Value Theory with the daily stock index returns of four different emerging markets. Research results according to estimated Generalized Pareto Distribution (GPD) parameters indicate the necessity of applying market risk estimation methods and it is clear that emerging markets such as those selected by the study have unique characteristics.

In the analysis of gold price return, the study of Jiahn-Bang Jang (2007) has been examined to illustrate the main idea of Extreme Value Theory and discuss the tail behavior. The results show that GPD model with threshold is a better choice. Also, Blake LeBaron and Ritirupa Samanta (2004) investigated the application of Extreme Value Theory (EVT) to construct statistical tests. The result shows that EVT elegantly frames the problem of extreme events in the context of the limiting distributions of sample maxima and minima. In financial market study, Neftci (2000) found that application of extreme distribution theory is well suitable to studying the extreme events in financial markets. Moreover, Alexander J. McNeil (1999) investigated Extreme Value Theory for risk managers. In this paper, the tail of a loss distribution is of interest, regarding risk in general; market, credit, operational or insurance risks, the POT method provides a simple tool for estimating measures of tail risk.

[1] Professor of Farm Management and PhD candidate, respectively. Department of Agricultural Economics, Humboldt University Berlin, Germany.