

An Approach for Comparison of Methodologies for Estimation of the Financial Risk of a Bond, Using the Bootstrapping Method

ChongHak Park*, Mark Everson, and Cody Stumpo

Business Modeling Research Group
e-Technology Department
Ford Research Laboratory
Ford Motor Company

*Computational Finance Program
Department of Statistics
Purdue University

Abstract

The Ford Motor Company's current cash holdings of approximately \$25B are primarily invested in bonds with maturities ranging from a few months to five years. Due to the large size of this cash holding, losses in the bond market could result in a substantial negative impact to quarterly profits. To gauge the size of this risk, the Portfolio Management Department in Treasury, uses a program called RiskManager[®], which utilizes a concept called value at risk (VaR). VaR summarizes the expected maximum loss to a financial portfolio under various conditions. The value at risk methodology rests upon a number of assumptions about the behavior and statistical properties of financial market variables. One important component of a VaR model is how it measures the fluctuations in value (volatility) of market instruments like bonds. The accuracy of these volatility estimates directly affects the end-user's confidence in the risk level determined by the VaR calculation. We have applied the statistical methodology called bootstrapping to compare the accuracy of the volatility estimates usable in the VaR model.

In this report, we look at the meaning of VaR, including aspects of its calculation methodology. Then we introduce the bootstrapping method. Bootstrapping allows one to infer the standard error of statistical parameters based upon a limited sample size. Using the bootstrapping method we can approximate the standard error of the volatility measure used in VaR, and so make an estimate of the potential error in the VaR value itself. In other words, we gauge how reliable the risk level indicated using VaR is. We determine, for a sample historical period, the standard error of five different possible volatility estimates for VaR using the bootstrapping method. This allows us to examine if there is a better (lower standard error) approach that could be used in volatility estimation. Finally, we conduct backtesting to check the bootstrapping result by using one year of historical data. Based upon the bootstrapping method using limited historical data for a five-year bond, we have found the best volatility estimate to be one based on the simple variance of past results. This is in contrast to the method used in RiskManager[®], which is an exponentially weighted moving average technique. However, the variance technique, in our limited backtesting analysis, was found to be somewhat conservative in predicting risk, while the other was overly aggressive. We conclude that in practice, the difference between the two methodologies is relatively small, and would not drive substantially different risk management behavior.

Introduction

Ford's cash reserves total approximately \$25 Billion, mostly in treasury bills and AA-rated corporate bonds, with a range of maturities ranging from a few months to over five years. So, even a marginal improvement in return or risk for cash management is of significant interest. First of all, keeping the risk of this portfolio within acceptable limits is important, since stock analysts and investors will use Ford's profit volatility as

one measure of the quality of our stock. Although risks in the cash portfolio are typically smaller than other risks the Company faces, in the event of a major sell-off in the bond market it could be an important factor in quarterly profits. Secondly, having a good measure of the risk is important in examining risk-return tradeoffs. If, for example, it were possible to achieve one additional basis point (0.01%) in return, and hold the risk level constant, that would equate to \$2.5 Million of extra income. An excellent risk measure is required to judge risk/return decisions at this level.

As one of several risk management tools, Ford Motor Company utilizes a program called RiskManager[®], which directly applies the concept of Value at Risk (VaR). VaR (1-3) was created as an easy to understand single number to measure the risk of a portfolio. For an extended general explanation of VaR see Reference 3. The VaR number is an estimate of the maximum loss to a financial portfolio over a given time period, within some specified confidence interval. To achieve this, the VaR methodology typically uses some amount of recent historical data to project a future distribution of returns for the market instruments in question (in this case, bonds).

The need to determine the risk of the Company's cash portfolio is driven not only by business, but also by legal requirements. In 1997, the Securities and Exchange Commission introduced an annual reporting requirement whereby certain corporations (Ford among them) are required to report several types of market risk including risk in the cash portfolio. Corporations were allowed a choice between several approaches that would provide investors a way to understand each company's risks in this area. VaR is one of the methodologies that Ford may use in the future to satisfy this market risk estimation requirement.

Most value at risk implementations assume that future returns on financial assets are normally distributed around some mean value, essentially like a one-dimensional random walk. Once the volatility (standard deviation) of an instrument has been calculated for a specific granularity of time (typically one day), the Gaussian distribution can yield a confidence level for any specific loss in the next time period (day). Or alternatively, each confidence interval is associated with a Value at Risk. The Risk Manager chooses a confidence level, such as 99%, and the distribution will provide that in 99% of the cases, a random walk would not produce a loss in excess of a particular value, which is called the Value at Risk. For example, this calculation might yield the result that on a daily basis the largest expected loss with a 99% confidence was \$10 million. VaRs for time periods longer than a day can be gotten directly from the daily VaR by multiplying by the square root of time. This is because the variance of a random walk grows linearly with time, and thus the volatility (standard deviation) grows as the square root of time. If there are 25 trading days in a month, the monthly VaR will simply be 5 times the daily VaR.

VaR also takes into account the correlations between instruments in a portfolio. Again, this calculation relies on recent historical data and assumes the distribution of returns of any two instruments in the portfolio is jointly normal. Less than perfect correlation between instruments will lower the VaR of the portfolio significantly from the naive sum of the VaRs of each instrument. In the most extreme case, a correlation of -1, one instrument acts to perfectly counter-balance the other to produce a riskless position (if the two are of equal magnitude). For the US bond market though, the diversification benefit achievable by buying bonds of various maturities is generally not very large.

In the remainder of this report, we will examine in some depth the meaning of VaR including specific methods for calculating it. Then we will use the bootstrapping method, described next, to assess the standard error of the VaR volatility estimates.

The bootstrapping method (4-7) was initially proposed by Efron (1979) as a nonparametric randomization technique that draws from the observed distribution of the data to model the distribution of a statistic of interest. Essentially, bootstrapping treats the observed result as a complete population, and investigates how statistical parameters like the mean and variance could change if this population were used as a source to create many different samples. With sufficiently large sample size, bootstrapping allows the discovery of a better estimate of the statistics of the population than the sample could by itself. The idea is a simple one, requiring only computing power to accomplish. The bootstrap method has wide application. However, it has its limitations. Most important among these is that the original sample must be sufficiently large to be reasonably representative of the underlying distribution from which the sample was collected. We use the asymptotic properties of the bootstrap to calculate the volatility to be used in VaR and compare the standard error for five different statistics to check which one is a better volatility estimate.

We have carried out this analysis to examine the application of bootstrapping techniques to the calculation of value at risk. The calculations were performed for a single representative bond over a limited time period. Due to the limited sample size, this work should be considered as a proof of concept,

rather than something that can be immediately applied to risk management practice. The specifications of the bond we used are shown below.

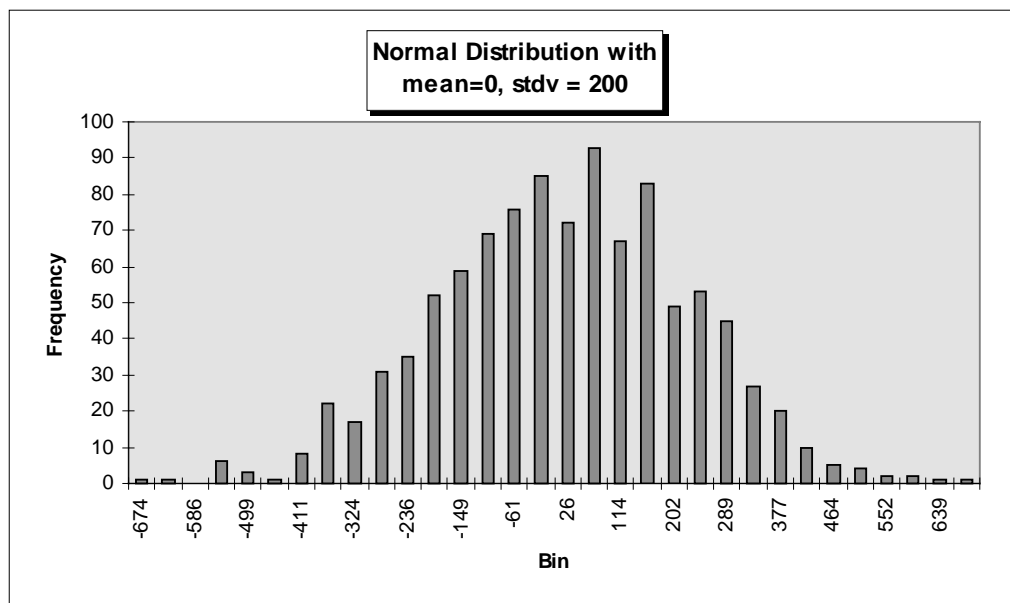
Principal Settlement	Maturity	Coupon	Frequency	Yield
\$100,000 6/1/1999	6/1/2004	5%	Annual	US Corp. AA

As we see in the next section, this bond is equivalent to the zero coupon bonds in terms of risk. A zero-coupon bond is a bond that pays its full face value at maturity, with no other payments.

1. Value at Risk (VaR)

1.1 What is VaR?

Value-at-Risk (1-3) is a number that represents the potential change in a portfolio's future value, as mentioned earlier. How this change is defined depends on the horizon over which the portfolio's change in value is measured and the confidence interval chosen by the RiskManager[®]. We focus on one-day VaR with 95% confidence interval. Let's take a look at a simple example. Suppose we want to find the 5th percentile of daily change in the price (p_t), of a bond, under the assumption that p_t is normally distributed with mean=0 and standard deviation =200 (arbitrary units). We generated random numbers with the given normal distribution, resulting in the histogram:



We know, by elementary statistics, that

probability($p_t < -1.65\sigma + \mu_t$) = 5%. Where σ is the standard deviation of the distribution, and μ_t is the mean of the distribution.

Notice that when $\mu_t=0$, we are left with the standard result that is the basis for short-term horizon VaR calculation, i.e.,

probability($p_t < -1.65\sigma$) = 5%.

We can easily find the 5th percentile in the histogram above and this is the one day, 95% VaR. For the distribution shown it is of order -330, since about 95% of samples from the distribution will be greater than

or equal to -330.

1.2 How to calculate VaR?

1.2.1 Cash Flows

Cash flows are the building blocks for describing any financial position. Once the cash flows are determined, they are marked-to-market, which means determining the present value of the cash flows given current market rates and prices. Therefore, it is necessary to get the current market rates, including the current yield curve for the appropriate kind of bond, and also a zero-coupon yield curve for them. The zero-coupon rate is the relevant rate for discounting cash flows received in a particular future period.

We might use a cash flow map to express the cash flows of interest rate positions. In this map, fixed income securities can be easily represented as cash flows given their standard future stream of payments. In practice, this is equivalent to decomposing a bond into a stream of zero-coupon instruments.

Let's take a look at one simple example. As we mentioned previously, let's consider a bond with a par value of \$100,000, a maturity of 5 years and an annual coupon rate of 5%. Assume that the bond is purchased at time 0 and that coupon payments are paid on an annual basis at the end of each year. We can draw the cash flow table as follows:

Year	1	2	3	4	5
Cash Flow	\$5,000	\$5,000	\$5,000	\$5,000	\$105,000

We can represent the cash flows of the simple bond in our example as cash flows from five zero-coupon bonds with maturities of 1, 2, 3, 4, and 5 years. This implies that on a risk basis, there is no difference between holding the simple bond or the corresponding five zero-coupon bonds.

1.2.2 Computation of VaR

There are two analytical approaches to measuring VaR: simple VaR for linear instruments and delta-gamma VaR for nonlinear instruments. Here, we are going to focus on the simple VaR method, which is appropriate for our bond portfolio. This derivation closely follows that in the RiskMetrics Technical Document (8).

In the simple VaR approach, we assume that returns on securities follow a conditionally multivariate normal distribution and that the relative change in a position's value is a linear function of the underlying return. Defining VaR as the 5th percentile of the distribution of a portfolio's relative changes, we compute VaR as 1.65 times the portfolio's standard deviation of \ln (daily return), where the multiple 1.65 is derived from the cumulative normal distribution to match the 5% one-side tail. This standard deviation depends on the volatilities and correlations of the underlying returns and on the present value of cash flows.

Based on the previous idea, we now formally derive the relationship between the relative change in the value of a position and an underlying return for linear instruments. We denote the relative change in value (the return) of the i th position, at time t , as $r_{i,t}^*$.

Note that in the case of fixed income instruments, the underlying value is defined in terms of prices on equivalent zero-coupon bonds. Alternatively, underlying returns could have been defined in terms of yields. For example, in the case of bonds, there is no longer a one-to-one correspondence between a change in the underlying yield and the change in the price of instruments. In fact, the relationship between the change in price of the bond and yield is nonlinear. Since we only deal with zero-coupon bonds we focus on these. Furthermore, we will work with continuous compounding.

Assuming continuous compounding, the price, P_t , at time t , of a zero-coupon bond with current price P_0 is

given by the expression below. The bond is assumed to mature in N periods, and have a yield y_t .

$$P_t = P_0 e^{-y_t N}$$

A second order approximation to the relative change in P_t yields

$$r_t^* = -y_t N (\Delta y_t / y_t) + \frac{1}{2} (y_t N)^2 (\Delta y_t / y_t)^2$$

Now, if we define the return r_t in terms of relative yield changes, i.e., $r_t = (\Delta y_t / y_t)$, then we have

$$r_t^* = -y_t N (r_t) + \frac{1}{2} (y_t N)^2 (r_t)^2$$

The equation above reveals two properties:

If we ignore the second term on the right-hand side we find that the relative price change is linearly related, but not equal, to the return on yield. However when we do include the second term, there is a nonlinear relationship between return, r_t , and relative price change.

Now we look at the general formula to compute VaR for linear instruments, such as our simple bond portfolio. The example provided below deals exclusively with the VaR calculation at the 95% confidence interval using the data provided by Risk-Metrics.

Let's consider our single bond that consists of 5 cash flows, for which we have volatility and correlation forecasts.

Denote the relative change in value of the n^{th} position by $r_{n,t}^*$. We can write the change of the portfolio $r_{p,t}^*$ as

$$r_{p,t}^* = \sum_{n=1}^5 \omega_n r_{n,t}^* = \sum_{n=1}^5 \omega_n \delta_n r_{n,t}$$

where ω_n is the total (nominal) amount invested in the n^{th} position. For example, suppose that the total current market value of a portfolio is \$100 and that \$10 is allocated to the first position. Then $\omega_1 = \$10$.

Now, suppose that the VaR forecast horizon is one day. In RiskMetrics, the methodology used in RiskManager[®], the VaR on a portfolio of simple linear instruments can be computed by 1.65 times the standard deviation of $r_{p,t}^*$ - the portfolio return, one day ahead. The expression of VaR is given as follows.

$$VaR_t = \sqrt{\vec{\sigma}_{t|t-1} R_{t|t-1} \vec{\sigma}_{t|t-1}^T}$$

where

$$\vec{\sigma}_{t|t-1} = [1.65\sigma_{1,t|t-1}\omega_1\delta_1 \quad 1.65\sigma_{2,t|t-1}\omega_2\delta_2 \quad \dots \quad 1.65\sigma_{5,t|t-1}\omega_5\delta_5]$$

is the individual VaR vector (1x5) and

$$R_{t|t-1} = \begin{bmatrix} 1 & \rho_{12,t|t-1} & \dots & \rho_{15,t|t-1} \\ \rho_{21,t|t-1} & 1 & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \rho_{51,t|t-1} & \dots & \dots & 1 \end{bmatrix}$$

is the 5x5 correlation matrix of the returns on the underlying cash flows. The correlation matrix indicates the closeness with which one asset type's price change follows another. In this case the different asset classes are the five zero coupon bonds with different maturity.

Note that, in this report, we are going to review the accuracy of volatility estimates (i.e.

$\sigma_{1,t|t-1}, \sigma_{2,t|t-1}, \dots, \sigma_{5,t|t-1}$), which directly specifies volatilities for market instruments in the VaR methodology.

As a final point, note that in line with RiskMetrics, we compute price volatility rather than yield volatility. Now that we know how to calculate VaR, we will describe how we can use the bootstrapping method to check the accuracy of VaR volatility estimates.

2. The Bootstrapping Method

2.1 Introduction to the Bootstrap

The bootstrap (4-7) is a computer-based method for assigning measures of accuracy to statistical estimates.

Suppose we obtain observations x_1, x_2, \dots, x_n independently from some distribution F and we wish to estimate the mean of F . Then naturally we would estimate the mean of the distribution by the mean of the sample \bar{x} . The accuracy of the estimate \bar{x} is measured by the estimated standard error:

$$SE = \frac{s}{\sqrt{n}} \approx \left\{ \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n(n-1)} \right\}^{1/2} \quad (1)$$

Standard error is an excellent first step forward thinking critically about statistical estimates. Unfortunately standard errors have a major disadvantage: for most statistical objects other than mean, there is no formula like the above to provide estimated standard errors. In other words, it is hard to assess the accuracy of an estimate, other than one for the mean.

For example, suppose we wish to compare two independent experiments by the medians of the two groups. The question is how accurate are the sample medians as the estimate of the medians of the distributions? Answering such questions is where the bootstrap, and other computer-based techniques, comes in.

Suppose we observe a random sample $x = (x_1, x_2, \dots, x_n)$ from some large sample of an unknown probability distribution function F , and we wish to estimate a parameter of interest θ on the basis of x . For instance, θ might be the median of the distribution function F . We can clearly obtain an estimate of θ , $\hat{\theta}$, by determining the median of sample x . We wish to assess the accuracy of the estimate $\hat{\theta}$. The bootstrap estimate of standard error, invented by Efron in 1979, allows us to do this.

A bootstrap sample $x^* = (x_1^*, x_2^*, \dots, x_n^*)$ is obtained by randomly sampling n times from the original data points x_1, x_2, \dots, x_n . For example, if the bootstrap sample consisted of three points $x = (1, 2, 3)$ one possible bootstrap sample would be $x^* = (2, 3, 2)$ where the value 2 happened to be selected randomly from the original sample twice, and the value 1 was never selected.

The bootstrap algorithm begins by generating a large number of independent bootstrap samples $x^{*1}, x^{*2}, \dots, x^{*B}$, each of size n . Typical values for B , the number of bootstrap samples, range from 50 to 200 for standard error estimation.

Corresponding to each bootstrap sample is a bootstrap replication of s , namely $s(x^{*b})$, the value of the statistic x evaluated for x^{*b} . If $s(x)$ is the sample median, for instance, then $s(x^{*b})$ is the median of the bootstrap sample.

Many data analysis problems, including the present one, involve data structures that are time series. For instance, the bootstrap algorithm can be adapted to general data structures including time series (see for example, Ref 6, pp385-412).

Most models for time series assume that the data are stationary, in which case the joint distribution of any subset of them depends only on their times of occurrence relative to each other and not on their absolute position in the series. In our data, the log differences of daily bond price--which is $\log(P(t+1)/P(t))$ --is assumed to be stationary. Under the assumption, we could generate a bootstrap sample from the log differences. Therefore, we can use the bootstrap technique under the hypothesis that the log differences of bond prices are close to Gaussian white noise.

2.2 Bootstrap Estimate of Standard Error

The bootstrap estimate of standard error is the standard deviation of the bootstrap replications,

$$SE_{boot} = \left\{ \sum_{b=1}^B [s(x^{*b}) - s(.)]^2 / (B-1) \right\}^{1/2}$$

where B is the number of bootstrap replications and $s(.) = (\sum s(x^{*b})) / B$. Suppose $s(x)$ is the mean \bar{x} . In this case, the weak law of large numbers tells us that as B gets very large, the formula above approaches

$$\left\{ \sum_{i=1}^n (x_i - \bar{x})^2 / n \right\}^{1/2}$$

Which approaches the value in formula (1) as n becomes large.

It is easy to write a bootstrap program that works for any computable statistic. With these programs in place, a data analyst is free to use any statistic, no matter how complicated, with the assurance that the statistic's error can be estimated.

Standard errors are the simplest measure of statistical accuracy, but the bootstrap can also assess more complicated accuracy measures--such as biases, prediction errors, and confidence intervals. However in this report, we focus on the standard errors. The price of using the bootstrap method for estimating the accuracy of a statistic is simply an increase in computational cost.

Bootstrap methods depend on the creation of a bootstrap sample. Let \hat{F} be the sample, putting probability $1/n$ on each of the observed values x_i , $i=1, \dots, n$. A bootstrap sample is defined to be a random sample of size n drawn from \hat{F} , say $x^* = (x_1^*, x_2^*, \dots, x_n^*)$,

$$\hat{F} \rightarrow (x_1^*, x_2^*, \dots, x_n^*).$$

The star notation indicates that x^* is not the actual data set x , but rather a randomized, or resampled, version of x . Another way to see the process above is that the bootstrap data points $x_1^*, x_2^*, \dots, x_n^*$ are a random sample of size n drawn from the population of n objects (x_1, x_2, \dots, x_n) . Some of the objects from the original sample may exist two or more times in a particular resampled version of x , and others will not

be present at all.

Corresponding to each bootstrap data set x^* is a bootstrap replication of $\hat{\theta}$, the statistic of interest.

$$\hat{\theta}^* = s(x^*)$$

The quantity $s(x^*)$ is the result of applying the same function $s(\cdot)$ to x^* as was applied to x .

The bootstrap estimate of $SE_F(\hat{\theta}^*)$, the standard error of a statistic $\hat{\theta}$, is a plug-in estimate that uses the sample \hat{F} in place of the unknown distribution F . The bootstrap estimate of $SE_F(\hat{\theta})$ is defined by

$$SE_{\hat{F}}(\hat{\theta}^*)$$

In other words, the bootstrap estimate of $SE_F(\hat{\theta})$ is the standard error of $\hat{\theta}$ for the data set of size n . The expression above is called the ideal bootstrap estimate of standard error of $\hat{\theta}$.

The bootstrap algorithm is a computational way of obtaining a good approximation to the value of $SE_{\hat{F}}(\hat{\theta}^*)$.

2.3 Bootstrap implementation using S-plus® (or R)

It is easy to implement bootstrap sampling on the computer using a statistical analysis package like S-Plus® or R (a similar open-source program). A random number generator selects integers i_1, i_2, \dots, i_n , each of which equals any value between 1 and n with probability $1/n$.

The bootstrap algorithm works by drawing many independent bootstrap samples, evaluating the corresponding bootstrap replications, and estimating the standard error of $\hat{\theta}$ by the empirical standard deviation of the replications.

The result is called the bootstrap estimate of standard error, denoted by \hat{SE}_B , where B is the number of bootstrap samples used. We summarize the algorithm as follows.

- Select B independent bootstrap samples $x^{*1}, x^{*2}, \dots, x^{*B}$, each consisting of n data values drawn with replacement from x . [For estimating standard error, the number B will ordinarily be in the range 25-200]
- Evaluate the bootstrap replication corresponding to each bootstrap sample, $\hat{\theta}^*(b) = s(x^{*b}), b = 1, 2, \dots, B$.
- Estimate the standard error $SE_F(\hat{\theta})$ by the sample standard deviation of the B replications

2.4 Comparison of the five different VaR volatility estimates

We compare the performance of five different estimators of the population variance through simulation. These five methods are sample variance, mean absolute deviation, median absolute deviation, interquartile range estimator, and exponential weighted moving average.

Suppose that Y is a daily return on a bond, measured by its price change, and n is the number of days

(samples). Now we briefly introduce those five methods as follows:

1. Sample Variance

$$\hat{\sigma} = \sqrt{\frac{\sum (Y_i - \bar{Y})^2}{n-1}}$$

2. Mean Absolute Deviation

$$\hat{\sigma} = \sqrt{\frac{\pi}{2}} d, \text{ where } d = \frac{1}{n} \sum_i |Y_i - \bar{Y}|$$

$$d \rightarrow \sqrt{\frac{2}{\pi}} \sigma$$

The constant is chosen for normalization since

3. Median Absolute Deviation (MAD)

$$\hat{\sigma} = \frac{MAD^2}{0.6745^2}, \text{ where } MAD = \text{median}_i \{|Y_i - \text{median}_j(Y_j)|\}.$$

The constant $0.6745 = \Phi^{-1}(0.75)$ since $MAD \rightarrow \text{median}\{|Y - \mu|\} = \sigma \Phi^{-1}(0.75)$

Where $\Phi^{-1}(x)$ is the probability that a variable with a standardized normal distribution is $< x$.

4. Interquartile Range Estimator (IQR)

$$\hat{\sigma} = \frac{IQR^2}{1.3490^2}, \text{ where } IQR = Y_{[3n/4]} - Y_{[n/4]}$$

The constant $1.3490 = \Phi^{-1}(0.75) - \Phi^{-1}(0.25)$ is chosen since for the normal $IQR \rightarrow \sigma[\Phi^{-1}(0.75) - \Phi^{-1}(0.25)]$

This is a good measure of dispersion for non-normal distributions, measuring the difference between the 25th %ile and 75th %ile results. Note that it takes into account only the central half of the distribution, and ignores the outlying data points.

5. Exponentially Weighted Moving Average

This is the volatility estimation method used in the VaR model.

$$\sigma = \sqrt{(1 - \lambda) \sum_{t=1}^T \lambda^{t-1} (Y_t - \bar{Y})^2}$$

The parameter λ ranges from 0 to 1 and is often referred to as the decay factor. This parameter determines the relative weights that are applied to the observations (returns) and the effective amount of data used in estimating volatility.

Now we compare the accuracy of each estimation method by bootstrapping. We calculated the accuracy of daily return-volatility estimates on cash flows for year 1 to year 5 of the hypothetical bond that we assumed at the beginning of this report.

2.5 Application of the bootstrap method to the five volatility estimation approaches

We summarize the steps in applying the bootstrap method to the different methods for estimating volatility here.

1) We have the bond price data and express it as follows:

$$P=(p(1), p(2), \dots, p(n))$$

Now since the data P are serially correlated and therefore it is inappropriate to use the bootstrapping method. (See the last paragraph in section 2.1) But we can avoid this difficulty if we calculate and use the daily return for our analysis. The daily return is assumed to be independent and identically distributed (iid) normal.

2) To get returns, we take the log difference of bond price between two consecutive days.

For example, to get the first daily return, we get $r(1) = \log(p(2)/p(1))$. Now we can express the daily returns as follows:

$$R=(r(1), r(2), \dots, r(n-1))$$

Now we can use bootstrapping to estimate the standard error for the five different volatility estimation methods such as IQR, MAD, Median Absolute Deviation, Normal Variance, and EWMA.

2.6 Results for the five different VaR volatility estimates

The bootstrapping test gave us the following results. Table 1 shows the standard error for each method in the first rows. The second rows are again standard errors, but normalized to the variance method's standard error for that instrument.

Table 1.

Daily return on 1-year zero coupon bond

var	mad	mean.ad	iqr	EWMA
4.04e-05	6.52e-05	4.14e-05	5.50e-05	6.59e-05
1.00	1.61	1.02	1.36	1.63

Daily return on 2-year zero coupon bond

var	mad	mean.ad	iqr	EWMA
8.42e-05	1.47e-04	8.67e-05	1.13e-04	1.75e-04
1.00	1.74	1.03	1.34	2.080457

Daily return on 3-year zero coupon bond

var	mad	mean.ad	iqr	EWMA
0.000131	0.000256	0.000146	0.000241	0.000313
1.00	1.95	1.11	1.83	2.38

Daily return on 4-year zero coupon bond

var	mad	mean.ad	iqr	EWMA
0.000179	0.000364	0.000192	0.000354	0.000523
1.00	2.03	1.07	1.97	2.92

Daily return on 5-year zero coupon bond

var	mad	mean.ad	iqr	EWMA
0.000185	0.000456	0.000207	0.000498	0.000559
1.00	2.46	1.12	2.68	3.01

As can be seen in the results above, the standard error for variance is consistently less than that of any other method, including the exponential weighted moving average, for each of the five different years. Therefore,

based on the bootstrapping analysis, the best parameter estimator is simple variance. Although the bootstrapping method can be used to check the accuracy of any parameter estimation method by definition, we should pay attention to the usage of the method to estimate the accuracy of the exponentially weighted moving average -- since there are weights λ involved in the EWMA calculation. EWMA assumes the order of the time series is relevant, whereas the bootstrapping method assumes that the original sample has been drawn as simple random sample. Clearly, to directly compare the EWMA bootstrapping result to the others, we need to prove mathematically whether the bootstrapping method can be directly applied to the exponential weighted moving average method. However, we have not yet been able to prove this. Given the lack of direct proof that the bootstrapping methodology is appropriate for volatilities calculated using an exponentially weighted moving average approach, the only practical way to test the results is through backtesting using an historical sample.

3. Backtesting

The bootstrapping analysis indicated the (simple) variance was a better volatility estimation method than exponential weighted moving average. At this point, we need to check if this result really makes sense for historical data. Therefore, we have performed some backtesting. We conducted backtesting only on the variance and exponential weighted moving average methods since these are the most interesting methods we are looking at.

In order to do the backtesting, we used data for bond prices from June 1, 1998 to June 1, 1999 with 6 months of rolling historical data to calculate the volatility estimates. For example, we use data between June 1, 1998 to October 31, 1998 to forecast volatility on November 1, 1998. We can perform the same procedure until June 1, 1999, which enables us to do have 130 backtested samples.

We calculated the daily actual change of the bond price by using the market yield data in RiskManager[®]. In other words, we discounted the cash flow for our hypothetical 5-year bond with the spot interest rate at each time and subtracted the price of the day from that of the next day, which would give us the daily price change.

We will provide a simple example in order to better explain this procedure. Let's say that at $t=0$ we calculated the one-day VaR using two different methods and obtained \$400 for the variance method and \$380 for the exponential weighted moving average method. Then we find the result for the actual price change between $t=0$ and $t=1$ is \$90 of loss. We can take a large number of such samples ($t=0,1,\dots$), and see if the expected fraction of cases exceeds the given VaR confidence interval in both the case of the simple variance method and the exponential weighted moving average method. For a good volatility estimation method, there should be close to the expected fraction of exceedances of the confidence interval. For instance a one-sided 95% confidence interval should be exceeded on average 5% of the time. Based on this idea, we review our results.

As can be seen in figure 1 below, we compared the two different methods in terms of the number of data points outlying the threshold of the 95% confidence level. The actual returns for each of 130 days are shown as large squares. The actual returns vary from gains of nearly \$700 to losses of approximately \$800. On the plot are shown the 95% confidence VaR estimates for each day based upon the simple variance (STDV for standard deviation) and exponentially weighted moving average (EWMA). In the case of the variance method, there are 5 actual data points (circled) outside the threshold, while there are 8 data points (circled and squared taken together) exceeding the line for the exponential weighted moving average method.

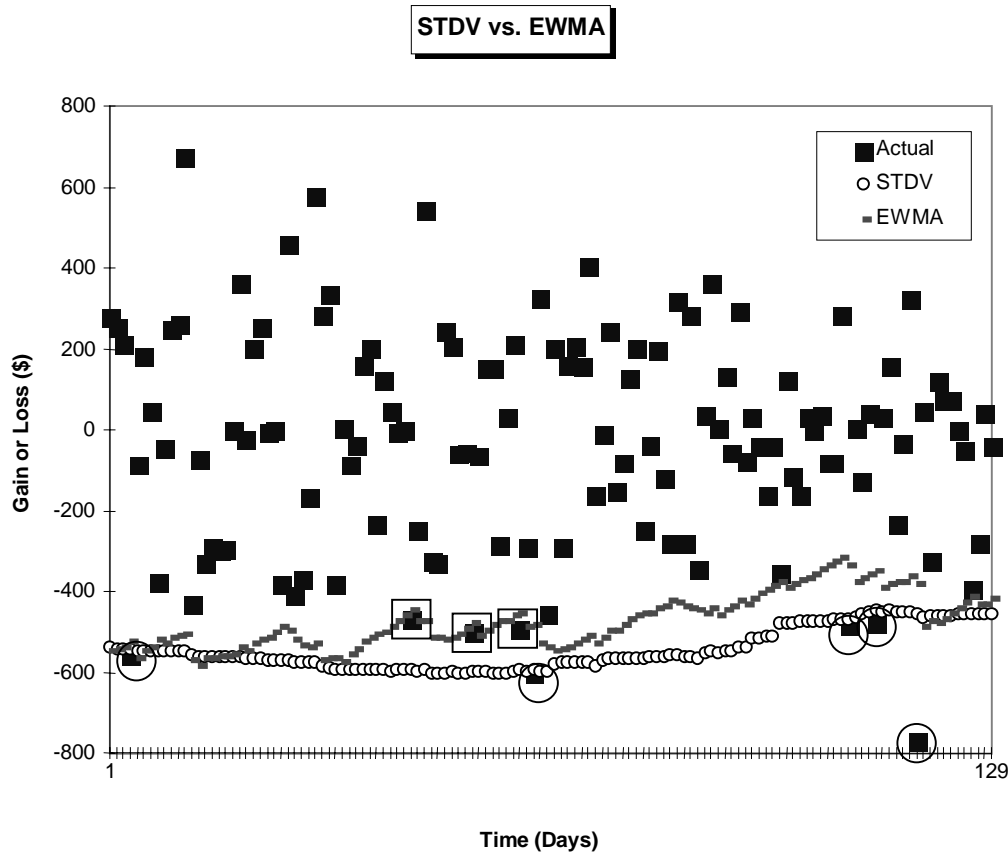


Figure 1. Gain or loss on a bond (solid squares) versus time, and predicted 95% risk levels using VaR methodology with simple variance (STDV, small circles) and EWMA (dashes) versus time. Both VaR numbers attempt to characterize the amount of expected daily loss for a 95% worst-case event assuming a \$100,000 five-year bond. See text for full discussion.

We can now compare the backtesting results for the case of simple variance vs. EWMA for our set of 130 predictions. Given that there are 130 predictions, the mean number of exceedances is expected to be 6.5 ($5\% \times 130$) for a perfect volatility forecast. Clearly, the results of five exceedances for simple variance, and eight exceedances for EWMA bracket the expected mean. Based on these results we can ask what the likelihood is of observing that number of exceedances if each of our forecasting methods were in turn assumed to be correct. We can use the binomial distribution with $p = 0.05$ to compare the observed number of exceedances of the 95% confidence interval for each case. For the case of simple variance, there were 5 exceedances during the sample period. This should happen 14.7% of the time for a binomial distribution. For the case of EWMA there were eight exceedances, which should happen approximately 12.1% of the time. So simple analysis of this limited data set cannot establish one method or the other as necessarily more accurate for forecasting volatility.

The plot does illustrate one contrast between value at risk numbers derived from simple variance and EWMA that we have seen in several different studies. Specifically, the estimated standard deviation (and so the predicted risk) is virtually always less using the EWMA method than the simple variance. This can be intuitively explained. The EWMA method gives less weight to all but the most recent events. Since large changes in prices tend to happen only infrequently, the EWMA method will estimate a standard deviation greater than that of the simple variance method only when one of these rare events has happened quite recently. For this reason we think of the simple variance VaR methodology as somewhat more conservative for risk management than VaR using an EWMA approach.

In addition to the accuracy of the volatility estimate, we need to consider the opportunity cost of keeping our risk within a pre-specified level for a portfolio of bonds. In general a bond portfolio will have a return that scales with its risk. So if we were to change the nature of our bond holding based on our risk measure, for instance based on a VaR threshold, we would in general be also changing the return of the portfolio. As we know from the analysis above, most of the time, the opportunity cost for holding a portfolio with a given VaR value suggested by the variance method is larger than that suggested by the exponential weighted moving average method. In other words, if we use the variance approach as a risk management tool, the circled lines in the graph above will be the threshold, and the hedging decision based on the "VaR with variance" method will be almost strictly larger than that obtained using the "VaR with EWMA" method. Therefore, the variance approach results in a presumably safer position than the EWMA approach, for any actual VaR, but this would be at the cost of reduced return.

Current practice in Treasury includes the understanding that VaR values based on EWMA can range widely due to fluctuating market conditions. The measured VaR is therefore used primarily as an indicator of the general risk level of the portfolio. Therefore, the difference between EWMA and simple variance in our context, usually significantly less than 20% of the VaR, is not really sufficient to result in a change in risk management practice given Treasury's approach to using VaR for risk management.

4. Summary and conclusion

The bootstrapping method shows that the parameter estimate using the normal standard deviation method is superior to the exponential weighted moving average method in terms of accuracy for our test case using a single bond. Backtesting indicated the quality of volatility forecast to be about equal between EWMA and the simple variance method. Based on our limited data, using simple variance resulted in VaR values that were somewhat more conservative than ideal (5 exceedances vs. 6.5 expected). For EWMA VaRs were less conservative than ideal (8 exceedances vs. 6.5 expected). Differences on this magnitude would not be expected to change behavior in Treasury's use of VaR for risk management. In any case, the conclusions here would require further work to substantiate, given the limited amount of historical data, and the use of only a single bond in our study.

References

1. J.P Morgan/Reuters, "RiskMetrics – Technical Documentation", 4th ed, J.P Morgan, 1996
2. Philippe Jorion, "Value at Risk", 1st ed, Irwin, 1997
3. Mark P. Everson, Christophe G. E. Mangin, Suzhou Huang, Cody Stumpo and JoAnn M. Schwartz, Evaluating Strategies for Foreign Exchange Risk Reduction, Research Highlights, to be published.
4. Efron, B. and Tibshirani, R. J. (1993). An Introduction to the Bootstrap. San Francisco: Chapman & Hall.
5. Shao, J. and Tu, D. (1995). The Jackknife and Bootstrap. New York: Springer-Verlag.
6. Davison, A.C. and Hinkley, D.V. (1997). Bootstrap Methods and Their Application. Cambridge University Press.
7. Tony Cai, lecture notes on "Advanced Statistical Methodology with Application in Finance", Statistics Department, Purdue Univ. 1999
8. RiskMetrics, Web page (<http://www.riskmetrics.com>)
9. Venables & Ripley, "Modern Applied Statistics with S-Plus", Springer-Verlag.