INVESTMENT research

23 September 2015

by Zhili Cao
Research & Investment Strategy

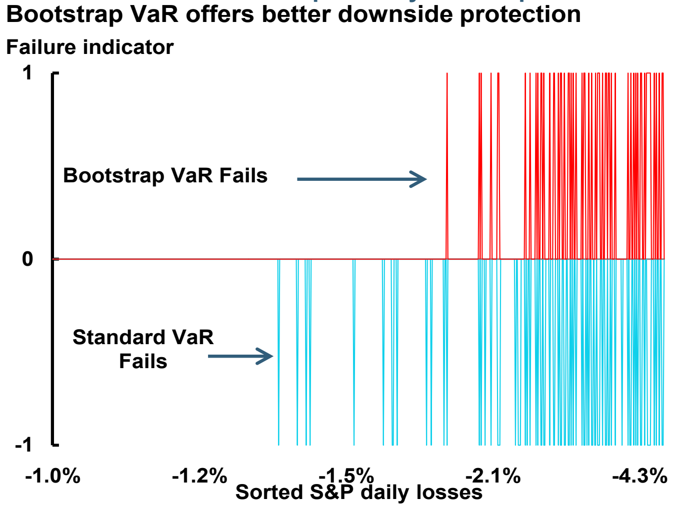
"Value at risk" might underestimate risk when risk bites. Just bootstrap it!

Key points

- Value at Risk (VaR) is one of the most widely used statistical tools to estimate a potential economic risk and is the cornerstone and common language of risk management in virtually all major financial institutions and regulators around the globe.
- Yet some caution is advised when using the conventional VaR. This measure systematically underestimates the risks of an investment. The rationale is simple: VaR models typically assume that the distribution of returns is normal (Gaussian). This is not true, the actual distribution has fat tails.
- A nonparametric Bootstrap VaR provides a more accurate picture of hypothetical losses and is thus a better gauge of market risk.
- Bootstrap VaR and market-implied volatility are strongly correlated, demonstrating the credibility and usefulness of Bootstrap VaR as a barometer of market sentiment / risk aversion.

Exhibit 1

Fat tail events are better captured by Boostrap VaR



Source: Datastream and AXA IM Research. Note: Standard VaR and Bootstrap VaR fail when their estimate of Value at Risk is lower (absolute value) than actual loss.

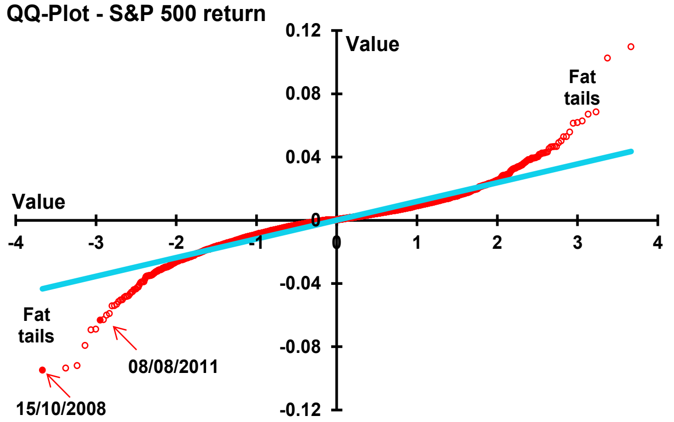
In this paper, we evaluate the conventional value at risk (VaR) methods, explain why the reality of financial markets limits its usefulness, and put forward an alternative VaR calculation method.[1]

Entering a turbulent financial world

VaR is one of the most widely used statistical measures of potential economic loss. It has been adopted as the cornerstone and common language of risk management by virtually all major financial institutions and regulators worldwide. Like many risk management tools, and calculations on which they are built, VaR has been designed for "normal times" e.g. based on assumptions that financial markets behave smoothly and follow certain distributions. However, in recent years – particularly since the global financial crisis – fat tails do exist, they usually occur without a clear warning bell. **We are living in ever more turbulent times - large, sudden risks (fat tail events) occur without warning and shake the foundations of financial markets worldwide.**

Exhibit 2 provides a quantile-quantile plot (QQ-plot)[2] of the S&P 500 Index daily returns from 1 January 2000 to 14 August 2015. A standard normal distribution would have a steady slope (blue line) while a strong nonlinear pattern similar to the curves at the extremes of the graph depicting S&P 500 returns (red line) indicates non-normal distribution: the presence of fat tails.

Exhibit 2
S&P 500 index returns suffer from fat tails



Source: Datastream and AXA IM Research

In light of a reality that is by no means normal in the statistical sense, risk management methods must be adapted to these more turbulent times. Rather than reject VaR outright, **we take the more reasonable view that VaR is a potentially useful tool to be used with caution.**

Value at risk : two alternative approaches

The VaR of a portfolio measures the amount an investor would lose with some small probability q, usually between 1% and 10%, over a given period. In this report, we set q equal to 1%. Thus, the value at risk represents a hypothetical loss. Now, we focus on **two alternative ways to calculate VaR**: 1) Generalized AutoRegressive Conditional Heteroskedasticity (GARCH) with normal distribution assumption method, and 2) Nonparametric bootstrap method. Here, the emphasis is placed on a portfolio's riskiness.

1. **Standard VaR**, with Normal assumption:

   Under the normality assumption of return with zero mean, value at risk equals to the product of quantile of probability q and the standard deviation of return (volatility). In order to allow the volatility to depicts a dynamic pattern, we implement a TGARCH process to calculate the dynamic conditional volatility, taking into account that volatility increases more after bad news than after good news.

2. **Nonparametric bootstrap VaR**:

   Instead of using a distribution-based approach to calculate VaR, **we loosen the distribution assumption and use the nonparametric bootstrap methodology to determine portfolio value at risk.** One can extend the existing measure to calculate the expected shortfall to gauge the potential loss.

   The idea of this bootstrap method is to mimic the return of S&P index returns at each time *t*[3] and we pick the lowest *q*-return that corresponds to the confidence level *q* to represent the VaR at time *t*. The advantage of this method is to allow us to estimate empirically the sampling distribution of a statistic without making assumptions about the form of the population and without deriving the sampling distribution explicitly.[4]

Evaluation of VaR

Given the importance of VaR estimates to the financial industry, evaluating the accuracy of the underlying model is a necessary exercise. The confidence level $1-q$ plays a crucial role in testing VaR's accuracy. That is to say, in a case of having 100 observations of VaR and *q* equals 1%, if the VaR has been violated only one time, we can conclude that this is a reliable VaR, otherwise, we should reject it.

To do so, suppose that we observe a time series of past *ex ante* VaR forecasts and past *ex post* returns, we can define the "HIT sequence" of VaR violation as:

$$H_t = \begin{cases} 1, & if\ r_t < VaR_t^q \\ 0, & if\ r_t \geq VaR_t^q. \end{cases}$$

[1] There are numerous other useful risk measures in the literature. However, these are not in the scope the research in this paper.
[2] Horizontal axis represents the normal theoretical quantile, vertical axis represents the data quantile.

[3] In this article, we implement bootstrap for 1000 times.
[4] See *Appendix 1 or see section 4.4 in "*[Mult-iCoVaR and Shapley value: A systemic risk measure](#)*"* for details.

If the model is correctly specified, H_t should be a Bernoulli distribution with a probability of 1% and independent and identically distributed (i.i.d.). As there are two constraints in this test, the likelihood ratio statistic follows a Chi-square distribution with degree of freedom of 2. The full description of statistical test is presented in *Appendix 2*.

Exhibits 1 and *3* provide the results for the two different methods of VaR calculation. The mean of standard VaR is -2.50%, which is higher than the mean of bootstrap VaR, -2.72%, but the difference between the two is a mere 22bps. Moreover, for each quartile, the VaR of bootstrap method is larger in absolute terms than the conventional VaR, as well as for minimum value and maximum value. Therefore, Boostrap VaR gives a larger cushion in case of an extreme event. This result is in line with expectations, since the bootstrap methodology does not assume any distribution assumption on S&P 500 index return, and it does take into account the fat tail events in VaR calculation. But, so far, we have not given any statistical test to judge which VaR is a more reliable gauge of expected loss that would allow investors to build an appropriate reserve.

Exhibit 3
Evaluation of VaR test result

	Standard VaR	bootstrap VaR	$\chi^2(2)$
Mean	-2.50%	-2.72%	
Median	-2.11%	-2.34%	
Min	-14.10%	-14.37%	
1Q	-2.85%	-3.05%	
3Q	-1.66%	-1.90%	
Max	-1.22%	-1.26%	
Likelihood ratio test	35.89	8.04	
Critical level – 1%			9.21

Source: AXA IM Research calculation

Statistical test

The last two lines in *Exhibit 3* show the results of evaluation of VaR based on a statistical test to determine the accuracy of both approaches.[5] For standard VaR, the test is greater than the critical value 9.21, which means that the hit function generated by conventional *VaR* does not follow an independent and identically distributed Bernoulli (1%). This is because, by assuming the normal distribution of S&P 500 index return, one would underestimate the tail risk during a crisis period, therefore the hit function will necessarily be violated more frequently. However, the bootstrap VaR did much better (the test is lower than the critical value): it fully captured the return movement in both calm and turbulent periods. As *Exhibits 3* clearly shows, **conventional VaR systematically underestimates the risk relative to bootstrap VaR and should be rejected**.

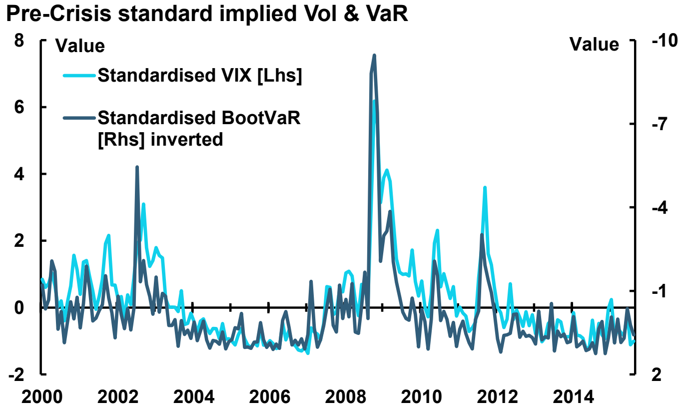
[5] A comprehensive construction of statistical test framework is provided in appendix.

Given the result of our evaluation of VaR, the bootstrap VaR appears to be the more reliable approach, but this method suffers a bigger computational burden; usually it takes more time than the conventional method. The clear advantage of a more reliable risk management tool is that it allows investors to build sufficient reserves in case of market turbulence or fat tail events.

Consistency with market implied volatility

Exhibit 4 shows the bootstrap VaR compared to the implied volatility of equity option prices as a proxy to risk aversion. After the dotcom bubble, both measures[6] were stable and moved within a tight band. However, with the onset of the crisis, both measure spiked. During the last two years, the two measures dropped to a pre-crisis level. Indeed, we can observe that bootstrap VaR and market-implied volatility depict the same pattern of evolution (the correlation between the two is 83% in absolute value), demonstrating the credibility and usefulness of bootstrap VaR as a barometer of market sentiment / risk aversion.

Exhibit 4
Bootstrap VaR is consistent with market implied vol



Source: Datastream and AXA IM Research

Conclusion

As we are in the middle of a turbulent financial market, it is no longer appropriate to make any specific distribution assumption when calculating relative market risks. By doing so, we may potentially underestimate the risk in calm period and maintain insufficient reserves to fight against a sudden market drop, which in turn would generate further market turmoil and amplify the pro-cyclicality effect in the market. Nonparametric method may overcome this shortcoming in calculating relative risk measures. Moreover, the bootstrap VaR can be a useful measure to gauge the risk appetite in the market.

[6] Here, both VIX and bootstrap VaR are standardized series relative to the pre-crisis mean and standard deviation.

Appendix 1: Bootstrap VaR

The evolution of the conditional variance dynamics (TGARCH) is given by:

$$r_t = \epsilon_t \sigma_t,$$
$$\sigma_t^2 = \omega + \alpha r_{t-1}^2 + \gamma r_{t-1}^2 I_{t-1}^- + \beta \sigma_{t-1}^2,$$

With $I_{t-1}^- = r_{t-1} \leq 0$. The model is estimated by Quasi-MLE which guarantees the consistency of estimator.

The key concept of bootstrap is that the population is to the sample as the sample is to the bootstrap sample. Then we proceed the bootstrap technique in the following way.

For a given series of returns $\{r_1, \dots, r_T\}$, consider a TGARCH model as in the previous case, whose parameters have been estimated by QMLE. Then we can obtain the standardized residuals, $\hat{\epsilon}_t = \frac{r_t}{\hat{\sigma}_t}$, where $\hat{\sigma}_t^2 = \widehat{\omega} + \hat{\alpha} r_{t-1}^2 + \hat{\gamma} r_{t-1}^2 I_{t-1}^- + \hat{\beta} \sigma_{t-1}^2$, and $\hat{\sigma}_1^2$ is long-run variance of the sample.

To implement the bootstrtap methodology, it is necessary to obtain bootstrap replicates $R_T^* = \{r_1^*, \dots, r_T^*\}$ that mimic the structure of original series of size T. R_T^* are obtained from the following recursion (Pascual et al. (2006))

$$\sigma_t^{*2} = \widehat{\omega} + \hat{\alpha} r_{t-1}^{*2} + \hat{\gamma} r_{t-1}^{*2} I_{t-1}^{*-} + \hat{\beta} \sigma_{t-1}^{*2},$$
$$r_t^* = \epsilon_t^* \hat{\sigma}_t^* \; for \; t = 1, \dots, T,$$

Where $\hat{\sigma}_1^{*2} = \hat{\sigma}_1^2$ and ϵ_t^* are random draws with replacement from the empirical distribution of standardized residuals $\hat{\epsilon}_t$. (It is necessary to sample with replacement, because on would otherwise simply reproduce the original sample). This bootstrap method incorporate uncertainty in the dynamics of conditional variance in order to make useful to estimate VaR. Given the bootstrap series R_T^*, we can obtain estimated bootstrap parameters, $\{\widehat{\omega}^b, \hat{\alpha}^b, \hat{\gamma}^b, \hat{\beta}^b,\}$. The bootstrap of historical values are obtained from following recursions,

$$\sigma_t^{b*2} = \widehat{\omega}^b + \hat{\alpha}^b r_{t-1}^{*2} + \hat{\gamma}^b r_{t-1}^{*2} I_{t-1}^{*-} + \hat{\beta}^b \sigma_{t-1}^{b*2},$$
$$r_t^{b*} = \epsilon_t^* \hat{\sigma}_t^{b*} \; for \; t = 1, \dots, T,$$

Where σ_1^{b*2} is the long-run variance of the bootstrap sample R_T^{b*}, note that he historical values is based on the original series of return and on the bootstrap parameters. We repeat the above procedure B times, and estimated $\widehat{VaR^*}_t(q)$ is k^{th}-order of series $\hat{r}_t^{b*}$, for b=1,…,B, where $k = B * q$.

Appendix 2: Evaluation of VaR

The objective is to test H_t are i.i.d. and Bernoulli(q). We need to define what type of serial correlation we want to test against. A simple alternative is a homogenous Markov chain. A simple first order binary valued Markov chain produces Bernoulli random variables which are not necessarily independent. It is characterized by a transition matrix which contains the probability that the state stays the same. The transition matrix is given by:

$$\Omega = \begin{bmatrix} \pi_{00} & \pi_{01} \\ \pi_{10} & \pi_{11} \end{bmatrix} = 1 - \begin{bmatrix} \pi_{00} & 1 - \pi_{00} \\ 1 - \pi_{11} & \pi_{11} \end{bmatrix},$$

Where $P[H_t = 1 | H_{t-1} = 1] = \pi_{11}$, $P[H_t = 1 | H_{t-1} = 0] = \pi_{01}$. In a correct specified model, the probability of a HIT in the current period should not be depend on whether the previous period was a HIT or not. In other words, the HIT sequence, $\{H_t\}$ is iid, and so that $\pi_{00} = 1 - q$ and $\pi_{11} = q$ when model is conditionally correct. The likelihood of Markov chain is

$$\mathrm{L}(\Omega) = (1 - \pi_{01})^{T_{01}} \pi_{01}^{T_{01}} (1 - \pi_{11})^{T_{10}} \pi_{11}^{T_{11}},$$

Where T_{ij} is the number of observations with a j following and i. The MLE estimator of π_{01} and π_{11} are

$$\hat{\pi}_{01} = \frac{T_{01}}{T_{00} + T_{01}}, \hat{\pi}_{11} = \frac{T_{11}}{T_{11} + T_{10}}.$$

Under independence, on has

$$\pi_{01} = \pi_{11} = \pi,$$

And

$$\Omega_0 = \begin{bmatrix} 1 - \pi & \pi \\ 1 - \pi & \pi \end{bmatrix},$$

While the MLE of π is $\hat{\pi} = T_1/T$. Hence, the likelihood ratio test of the independence assumption is given by

$$\mathrm{LR}_{ind} = 2[\log\left(L(\widehat{\Omega}) - \log\left(L(\widehat{\Omega}_0)\right)\right] \sim \chi^2(1),$$

Under the null, where

$$\widehat{\Omega} = \begin{bmatrix} 1-\hat{\pi}_{01} & \hat{\pi}_{01} \\ 1-\hat{\pi}_{11} & \hat{\pi}_{11} \end{bmatrix} \; and \; \widehat{\Omega}_0 = \begin{bmatrix} 1-\hat{\pi} & \hat{\pi} \\ 1-\hat{\pi} & \hat{\pi} \end{bmatrix}.$$

However, we want to test independence and Bernoulli(q) (conditional convergence test), i.e. $\pi_{01} = \pi_{11} = q$. The test is

$$\mathrm{LR}_{cc} = 2[\log(L(\widehat{\Omega}) - \log(L(q))] \sim \chi^2(2).$$

DISCLAIMER