A Practical Guide to Market Risk Model Validations - Focusing on VaR and TVaR

Vilen Abramov¹
and M. Kazim Khan²

¹BB&T Corp, Charlotte, USA
²Kent State University, Kent, USA

Robust Techniques in Quantitative Finance
University of Oxford, 3-7 September 2018
Outline

1. Introduction
   - VaR/TVaR definition
   - VaR/TVaR uses
   - Properties of risk measures
   - VaR/TVaR model specifications

2. Stranded London Whale

3. Standard VaR/TVaR estimators

4. Robust Techniques for Estimating VaR/TVaR
   - Measuring VaR/TVaR model risk
   - Asymptotic efficiency testing
   - Real data testing
It’s just a quantile

In mathematical terms, VaR is just a quantile. The quantile function \( q_X(\alpha) \) of a random variable \( X \) can be defined as follows.

\[
q_X(\alpha) = \inf\{ x : P[X \leq x] \geq \alpha \}
\]

Using basic equation \( P[X \leq x] = 1 - P[X > x] \), we can rewrite \( q_X(\alpha) \) in the following manner.

\[
q_X(\alpha) = \inf\{ x : P[X > x] \leq 1 - \alpha \}
\]

In finance, one wants to make sure that the losses \( (L) \) do not exceed a certain level with a high probability. This leads to the following definition of \( VaR(\alpha) \).

\[
VaR(\alpha) = \inf\{ l : P[L > l] \leq 1 - \alpha \}
\]
An alternative definition of VaR based on the P&L process rather than the loss process.

\[ \text{VaR}(\alpha) = - \sup \{ u : \Pr[P&L \leq u] \leq 1 - \alpha \} \]

The second definition of VaR is based on the right quantile function.

\[ q^*_X(\alpha) = \sup \{ x : \Pr[X \leq x] \leq \alpha \} \]

Both VaR definitions agree due to the left and right quantile symmetry.

\[ q_X(\alpha) = -q^*_{-X}(1 - \alpha) \]
One VaR - different riskiness

1. Calculation of VAR from probability distribution of changes in portfolio value

2. Alternative situation to figure 1: VAR is the same, but the potential risk is higher
VaR gives a lower bound for the extreme unlikely losses. The actual losses can exceed VaR. What’s the expected loss in extreme cases? It can be measured as follows.

\[ TVaR(\alpha) = \frac{1}{1 - \alpha} \int_{\alpha}^{1} VaR(u) \, du \]

This can be interpreted as integration over extreme losses with equal weights of \( 1/(1 - \alpha) \). Tail VaR is also known as Expected Shortfall, Conditional VaR, Average VaR, and Expected Tail Loss.
Various VaR applications

- Trading desk limits - $\alpha = .95$, $h = 1$ day, and underlying process is $L$
  \[ VaR_{1d}(.95) \]

- Economic capital - $\alpha = .999$, $h = 1$ year, and underlying process is $L$
  \[ \text{ECapital} = V aR_{1Y}(.999) - \text{Expected Loss} \]

- Maximum (peak) potential future exposure - $\alpha = .95$, $h = t$
  is a variable, and underlying process is $V_t^+$
  \[ \text{MPFE} = \max_{0 \leq t \leq T} V aR_t(.95) \]
Regression usage of VaR/TVaR

- 1996-1997 - Basel II (market risk amendment) introduced VaR-based capital concept. It has been also adopted by the FRB in the form of market risk capital rule. VaR-based capital is calculated using $\alpha = .99$, $h = 10$ days, and underlying process $L$.

\[
mRWA = \max \left\{ \text{VaR}_{10d}^{t_0}(.99), \frac{\beta}{60} \sum_{i=1}^{60} \text{VaR}_{10d}^{t_0-i\delta t}(.99) \right\}
\]

- 2009-2012 - Basel 2.5 introduced stressed VaR. FRB revised market risk capital rule to include stressed VaR as well.

- 2011-2019 - FRTB introduced TVaR (expected shortfall) to ensure a more prudent capture of "tail risk."
Properties of risk measures

Coherent risk measures

From the risk measurement perspective, VaR has two major drawbacks - lack of sensitivity to the size of the extreme losses and lack of sub-additivity. A coherent risk measure $\rho$ outperforms VaR in this regard and obeys the following laws.

1. (Positive homogeneity) $\rho(aL) = a\rho(L)$, $\forall a \geq 0$
2. (Sub-additivity) $\rho(L_1 + L_2) \leq \rho(L_1) + \rho(L_2)$
3. (Normalized) $\rho(L + a1) = \rho(L) + a$, $\forall a \in \mathbb{R}$
4. (Monotonicity) $\rho(L_1) \leq \rho(L_2)$ if $L_1 \leq L_2$ in all scenarios
TVaR measure can be generalized by allowing more complicated increasing weighting functions $w(u)$ defined on the interval $[0, 1]$. This leads us to the definition of the spectral risk measure (SRM).

$$SRM(L) = \int_0^1 w(u) \text{VaR}_u(L) du$$

The SRM can be re-written using anti-derivative of weighting function $w(u)$ (called distortion function), $D(u) = \int_0^u w(s) ds$. This leads us to the following definition of distortion risk measure (DM).

$$DM(L) = \int_0^1 \text{VaR}_u(L) dD(u)$$
VaR/TVaR model specifications

**VaR/TVaR production process**

- **Input**
  - Positional Data
  - Market Data

- **Processing**
  - Scenario Building
  - Revaluation
  - VaR/TVaR Calculation

- **Output**
Common VaR models

<table>
<thead>
<tr>
<th>Simulation approach</th>
<th>Valuation approach(^1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sensitivities</td>
</tr>
<tr>
<td>Historical simulation</td>
<td></td>
</tr>
<tr>
<td>Hybrid</td>
<td></td>
</tr>
<tr>
<td>Monte Carlo</td>
<td></td>
</tr>
</tbody>
</table>

1 Banks are deemed to use the sensitivities approach if they use it exclusively, hybrid if they use it at least 30 percent of the time, and full revaluation if less than 30 percent.

Source: McKinsey Market Risk Survey and Benchmarking 2011
VaR models categorization

- Direct P&L Modeling
  - Known Distribution $F(x)$
    - Estimated Volatility
    - GARCH Family
    - Stochastic Volatility
    - Realized Volatility
- Risk Factors Modeling
  - Historical
    - Volatility Scaling
    - Time Weighting
    - Equal Weighting
  - Monte Carlo
- EVT
  - $\delta - \gamma$
    - Normal
- Block Maxima
  - Peak over Threshold
- Quantile Regression
### VaR/TVaR model specifications

<table>
<thead>
<tr>
<th>Specification attribute</th>
<th>Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>Revaluation methodology</td>
<td>Sensitivity based</td>
</tr>
<tr>
<td></td>
<td>Full revaluation</td>
</tr>
<tr>
<td>VaR estimation methodology</td>
<td>Parametric</td>
</tr>
<tr>
<td></td>
<td>Mixed</td>
</tr>
<tr>
<td></td>
<td>Non-parametric</td>
</tr>
<tr>
<td>Simulation object</td>
<td>Relative change</td>
</tr>
<tr>
<td></td>
<td>Absolute change</td>
</tr>
<tr>
<td></td>
<td>Log change</td>
</tr>
<tr>
<td>Simulation calibration</td>
<td>Historical</td>
</tr>
<tr>
<td></td>
<td>Market implied</td>
</tr>
<tr>
<td>Scenario Weighting</td>
<td>Equal</td>
</tr>
<tr>
<td></td>
<td>Exponential</td>
</tr>
<tr>
<td>Confidence level scaling</td>
<td>Yes/No</td>
</tr>
<tr>
<td>Horizon scaling</td>
<td>Yes/No</td>
</tr>
</tbody>
</table>
In December 2011, as part of a firmwide effort in anticipation of new Basel capital requirements, we instructed CIO to reduce risk-weighted assets and associated risk. To achieve this in the synthetic credit portfolio, the CIO could have simply reduced its existing positions; instead, starting in mid-January, it embarked on a complex strategy that entailed adding positions that it believed would offset the existing ones.
CIO strategy

- JPMorgan’s CIO sold substantial amounts of CDX IG index exposure in the first quarter of the year. It also bought CDX HY protection, with total notional trade sizes running to tens of billions of dollars.

- In the economic downturn scenario, HY companies’ spreads will widen more than those of IG companies.

- A substantial divergence occurred between IG and HY indices. The two indices moved in tandem in the past. The historical relationship between the indices has been broken.
CDX IG and CDX HY correlation breakdown

In periods of heightened market volatility, risk factors’ correlations can differ substantially from those seen in normal periods. This is so-called "correlation breakdown" effect.
Sample VaR - 200-year-old problem

A number of banks utilize historical approach. A common choice for the VaR estimator is the sample quantile.

\[ V_1 := F_n^{-1}(p) = X_{(np)} \]

In the last part of the Second Supplement (1818) to the monumental Théorie Analytique des Probabilités, Laplace derived the asymptotic distribution of a single order statistic. Laplace then compared the sample mean and median estimators on the basis of the variances of their asymptotic distributions. It is now well known that \( X_{(np)} \) is asymptotically normally distributed.

\[ \sqrt{n}(V_1 - F^{-1}(p)) \xrightarrow{dist} N(0, p(1 - p)/f^2(F^{-1}(p))), \quad n \to \infty \]
Sample TVaR

Now turning towards the nonparametric estimator of TVaR, the sample tail $T_1$.

$$T_1 := \frac{1}{n(1-p)} \sum_{i=(np)+1}^{n} X_{(i)}$$

It can be shown that.

$$\sqrt{n}(T_1 - TVaR(p)) \xrightarrow{dist} N\left(0, \tau_1^2\right)$$

$$\tau_1^2 := \frac{1}{(1-p)^2} \int_{F^{-1}(p)}^{\infty} (u - VaR(p))^2 dF(u) - (TVaR(p) - VaR(p))^2$$
Theoretical VaR/TVaR under normal and Student’s t distributions

If the loss $L$ has normal $N(\mu, \sigma^2)$ or Student’s t $(t(\nu, \mu, \sigma^2))$ distributions, the VaR and TVaR can be expressed in the following manner.

$$VaR_N(\alpha) = \mu + \sigma \Phi^{-1}(\alpha)$$
$$VaR_t(\alpha) = \mu + \sigma t^{-1}_\nu(\alpha)$$
$$TVaR_N(\alpha) = \mu + \frac{\sigma}{1 - \alpha} \varphi(\Phi^{-1}(\alpha))$$
$$TVaR_t(\alpha) = \mu + \frac{\sigma}{1 - \alpha} g_\nu(t^{-1}_\nu(\alpha)) \left( \frac{\nu + (t^{-1}_\nu(\alpha))^2}{\nu - 1} \right)$$
Introduction
Stranded London Whale
Standard VaR/TVaR estimators
Robust Techniques for Estimating VaR/TVaR

Measuring VaR/TVaR model risk

VaR model variance

Dispersion of normalised VaR results for all portfolios

Equity
Interest rate
FX
Commodity
Credit spread
Diversified
Financial industry tends to assess the quality of the VaR models through back testing. However, back testing is challenging when dealing with TVaR. It has been shown that VaR is elicitable whereas TVaR is not. Model validation teams need to address the theoretical soundness of the VaR/TVaR estimation models. They need to answer the following questions. How good are the standard non-parametric estimators, as well as their semi-parametric and parametric counterparts? Are they unbiased, sufficient, consistent, asymptotically efficient? What are the "best" estimators of VaR/TVaR?

In our analysis we will focus on the asymptotic efficiency of the VaR and TVaR estimators. This will allow us to avoid the pitfalls of the back testing approach. We will then put proposed VaR estimators to the real test.
Measuring VaR/TVaR model risk

**VaR model risk smile**

![Diagram showing the relationship between model risk and assumptions, with different risk levels for non-parametric, semi-parametric, and parametric models.]
### Testing framework

<table>
<thead>
<tr>
<th>Theo</th>
<th>$\text{VaR(Model,p)}$</th>
<th>$\text{TVaR(Model,p)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>$V_{1,2,3}$</td>
<td>$V_{1,2,3}$</td>
</tr>
<tr>
<td>SP</td>
<td>$R_{1,2}$</td>
<td>$R_{1,2}$</td>
</tr>
<tr>
<td>NP</td>
<td>MME/MLE/UMVU</td>
<td>MME/MLE/UMVU</td>
</tr>
</tbody>
</table>
Theorem. (Joint asymptotic distribution of $L$-estimators)
Let $X_1, X_2, \cdots, X_n \overset{iid}{\sim} F$ with the corresponding order statistics $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$. For piecewise smooth functions $L_1(x), L_2(x), x \in [0, 1]$ consider the order statistics $\hat{\Theta}_{1n} = \frac{1}{n} \sum_{i=1}^{n} L_1\left(\frac{i}{n+1}\right)X_{(i)}, \hat{\Theta}_{2n} = \frac{1}{n} \sum_{i=1}^{n} L_2\left(\frac{i}{n+1}\right)X_{(i)}$. If $F$ is absolutely continuous with respect to Lebesgue measure with $E(X_1^2) < \infty$ then we can get.

$$\sqrt{n} \begin{pmatrix}
(X_{(np)} - F^{-1}(p)) \\
\hat{\Theta}_{1n} - J_1(L_1, F) \\
\hat{\Theta}_{2n} - J_2(L_2, F)
\end{pmatrix} \overset{dist}{\to} \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \Sigma\right)$$
Robust non-parametric VaR estimators - 2nd variation

An improvement in efficiency may be possible under the location-scale model, $F(x) = G((x - \mu)/\sigma)$, when $G$ has a symmetric density. Optimizing $M_n + \alpha X_{(np)} - (1 - \alpha) X_{(n(1-p))}$ we get:

$$V_2 := M_n + \frac{X_{(np)} - X_{(n(1-p))}}{2}$$

where $M_n$ is the sample median. It is also a consistent and asymptotically normal estimator with the following asymptotic variance.

$$\sigma^2_2 = \sigma^2 \left\{ \frac{1}{4(g(0))^2} + \frac{(1 - p)(2p - 1)}{2(g(G^{-1}(p)))^2} \right\}$$
Asymptotic efficiency testing

**Robust non-parametric VaR estimators - 3rd variation**

By expanding on the media portion, we obtain another consistent and asymptotically normal estimator.

\[ V_3 := \beta X_{(\gamma n)} + (1 - 2\beta) X_{(0.5 n)} + \beta X_{(n(1-\gamma))} + \frac{X_{(np)} - X_{(n(1-p))}}{2}, \]

where \( \beta, \gamma \leq \frac{1}{2} \).

\[ \sigma^2_3 = \sigma^2 \left\{ \frac{2\beta^2 \gamma}{(g(G^{-1}(\gamma)))^2} + \frac{(1-2\beta)^2}{4(g(0))^2} + \frac{2\beta(1-2\beta)\gamma}{g(0)g(G^{-1}(\gamma))} + \frac{(1-p)(2p-1)}{2(g(G^{-1}(p)))^2} \right\} \]

Further refinements could be achieved by more sophisticated estimators of order statistics, however, the resulting improvements may be minor and parametric model dependent, if any.
Robust non-parametrics TVaR estimators - 2nd and 3rd variations

Just as in the case of $VaR(p)$, one may modify sample TVaR estimator a bit further when the distribution $G$ has a symmetric density, $g$.

\[
T_2 := M_n + \frac{1}{2n(1-p)} \left( \sum_{i=(np)+1}^{n} X(i) - \sum_{i=1}^{(n(1-p))} X(i) \right)
\]

\[
T_3 := \beta X_{(\gamma n)} + (1 - 2\beta) X_{(0.5n)} + \beta X_{(n(1-\gamma))} + \frac{1}{2n(1-p)} \left( \sum_{i=(np)+1}^{n} X(i) - \sum_{i=1}^{(n(1-p))} X(i) \right),
\]

for $\beta, \gamma \leq \frac{1}{2}$. These estimators are consistent and asymptotically normal with the following asymptotic variances.
Robust non-parametrics TVaR estimators - 2nd and 3rd variations

\[
\tau_3^2 = \sigma^2 \left\{ \frac{2\beta^2\gamma}{(g(G^{-1}(\gamma)))^2} + \frac{(1 - 2\beta)^2}{4(g(0))^2} + \frac{2\beta(1 - 2\beta)\gamma}{g(0) g(G^{-1}(\gamma))} \right\} \\
+ \frac{1}{2} \left\{ \frac{1}{(1 - p)^2} \int_{F^{-1}(p)}^{\infty} (u - \text{VaR}(p))^2 dF(u) \right\} \\
- (TVaR_p - \text{VaR}(p))^2
\]

For \( \beta = 0 \) we get the asymptotic variance \( \tau_2^2 \) for the estimator \( T_2 \).

Following this idea of separately estimating the location and scale terms, one may build more sophisticated estimators, however, the gains may or may not be worth the effort, and the optimal choices may be model dependent.
The estimators presented earlier are non-parametric ones. They do not require any assumptions about the underlying distributions. One can also use parametric approach to derive VaR and TVaR. Let’s consider standard examples of normal distribution and $t$-distribution with unknown parameters. The quantiles can be expressed as follows.

\[ \text{VaR}_N^{UMVU}(\alpha) = \bar{X}_n + \frac{S\sqrt{n-1}}{D_n} \Phi^{-1}(\alpha) \]

where \( D_n = \mathbb{E} \left[ \sqrt{\chi^2_{(n-1)}} \right] \).

\[ \text{VaR}_t^{UMVU}(\alpha) = \bar{X}_n + \frac{S\sqrt{n-1}}{D_n(1-\alpha)} \varphi \left( \Phi^{-1}(\alpha) \right) \]
Semi-parametric VaR/TVaR estimators

Now consider a location-scale representation of $F$ in terms of random variables, $X_i = \mu + \sigma U_i$, where $U_1, U_2, \cdots, U_n \overset{iid}{\sim} G$, and $E(U_i) = 0$, with $\text{Var}(U_i) = \gamma^2_G =: \gamma^2$. We will make the blanket assumption that $E(X_i^4) < \infty$. However, some of the results hold with only assuming finite variance. Since both $\text{VaR}(p)$ and $\text{TVaR}(p)$ are of the form, $\mu + \sigma c(p)$, with $E(U_i^2) > 0$, for appropriate expressions for $c(p)$ that depend on $G$, we consider estimators of $\mu + \sigma c(p)$. 
Semi-parametric VaR/TVaR estimators - 1st variation

Applying the method of moments estimation technique to the empirical distribution function $F_n$ we may consider a very simple method of moments semi-parametric estimator (MME)

$$R_1 = \bar{X}_n + \frac{c(p)}{\gamma} S$$

When $c(p)$ is available then one gets asymptotic normality and not just consistency.

$$\sqrt{n}(R_1 - (\mu + c(p)\sigma)) \xrightarrow{\text{dist}} N(0, \nu_1^2)$$

$$\nu_1^2 := \sigma^2 \left\{ \gamma^2 + 0.25(c(p))^2(\kappa_G - 1) + \gamma c(p) \psi_G \right\},$$

where $\psi_G$ and $\kappa_G$ are the skewness and kurtosis of the distribution $G$. 
Asymptotic efficiency testing

**Semi-parametric VaR/TVaR estimators - 2nd variation**

One may consider the following modification of the semi-parametric estimator for VaR\((p)\) or TVaR\((p)\).

\[
R_2 := M_n + \frac{Y_n}{\delta} c(p),
\]

where \(\delta = E|U_1 - M_G|\), \(M_G\) is the median and \(Y_n = \frac{1}{n} \sum_{i=1}^{n} |X_i - M_n|\). When \(c(p)\) is available then one gets asymptotic normality.

\[
\nu_2^2 := \sigma^2 \left\{ \frac{1}{4(g(M_G))^2} + \frac{\text{Var}(|U_1 - M_G|)(c(p))^2}{\delta^2} - \frac{c(p) M_G}{\delta g(M_G)} \right\}
\]
Asymptotic efficiency testing

Comparative analysis - VaR for normal model

![Graphs showing comparative analysis of VaR for normal model.](image_url)
Asymptotic efficiency testing

Comparative analysis - VaR for Student’s t model

![Graphs showing asymptotic efficiency for different degrees of freedom (df) with various estimators: MLE/V1, MLE/V2, MLE/V3, MLE/R1, MLE/R2 for Student’s t model with df = 5, 6, 10, and 30.](image)
Asymptotic efficiency testing

Comparative analysis - TVaR for normal model
Asymptotic efficiency testing

Comparative analysis - TVaR for Student’s t model

TVaR, T Model, df = 5

TVaR, T Model, df = 6

TVaR, T Model, df = 10

TVaR, T Model, df = 30
## Sample variances of VaR estimators (2009-2018)

<table>
<thead>
<tr>
<th>Tickers</th>
<th>$V_1$</th>
<th>$V_2$</th>
<th>$V_3$</th>
<th>$R_2$</th>
<th>UMVU</th>
</tr>
</thead>
<tbody>
<tr>
<td>DB Commodity</td>
<td>58.54</td>
<td>43.98</td>
<td>45.48</td>
<td>38.65</td>
<td>42.64</td>
</tr>
<tr>
<td>Emerging Markets</td>
<td>205.68</td>
<td>171.62</td>
<td>175.81</td>
<td>112.79</td>
<td>211.33</td>
</tr>
<tr>
<td>High Yield</td>
<td>38.31</td>
<td>26.50</td>
<td>26.59</td>
<td>17.60</td>
<td>40.95</td>
</tr>
<tr>
<td>Developed (exc. North America)</td>
<td>97.72</td>
<td>80.72</td>
<td>81.07</td>
<td>63.39</td>
<td>98.32</td>
</tr>
<tr>
<td>Short Term Treasury</td>
<td>0.04</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.07</td>
</tr>
<tr>
<td>S&amp;P 500</td>
<td>69.03</td>
<td>60.19</td>
<td>62.58</td>
<td>43.80</td>
<td>75.62</td>
</tr>
</tbody>
</table>
### VaR breaches (2018)

<table>
<thead>
<tr>
<th>Tickers</th>
<th>$V_1$</th>
<th>$V_2$</th>
<th>$V_3$</th>
<th>$R_2$</th>
<th>UMVU</th>
</tr>
</thead>
<tbody>
<tr>
<td>DB Commodity</td>
<td>8.0</td>
<td>7.0</td>
<td>7.0</td>
<td>10.0</td>
<td>8.0</td>
</tr>
<tr>
<td>Emerging Markets</td>
<td>15.0</td>
<td>15.0</td>
<td>15.0</td>
<td>18.0</td>
<td>15.0</td>
</tr>
<tr>
<td>High Yield</td>
<td>6.0</td>
<td>6.0</td>
<td>6.0</td>
<td>8.0</td>
<td>5.0</td>
</tr>
<tr>
<td>Developed (exc. North America)</td>
<td>11.0</td>
<td>12.0</td>
<td>12.0</td>
<td>13.0</td>
<td>11.0</td>
</tr>
<tr>
<td>Short Term Treasury</td>
<td>7.0</td>
<td>7.0</td>
<td>7.0</td>
<td>7.0</td>
<td>7.0</td>
</tr>
<tr>
<td>S&amp;P 500</td>
<td>16.0</td>
<td>16.0</td>
<td>16.0</td>
<td>16.0</td>
<td>16.0</td>
</tr>
</tbody>
</table>
Conclusions

- Simple nonparametric estimators of $\text{VaR}(p)$ and $\text{TVaR}(p)$, $V_1, T_1$ can have extremely low efficiency.
- One can improve stability of the simple nonparametric estimators by using more order statistics.
- Semi parametric estimators of $\text{VaR}(p)$ and $\text{TVaR}(p)$ can give some protection against the model risk and still have reasonably high efficiency.
- While taking a considerable model risk, uniformly minimum variance unbiased or maximum likelihood estimators of $\text{VaR}(p)$ and $\text{TVaR}(p)$ can be constructed under some typical parametric models.
- By using alternative estimators, one can both improve the back testing results and minimize the variance. This would lead to a more accurate and stable capital calculations.