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

- Introduction
 - VaR/TVaR definition
 - VaR/TVaR uses
 - Properties of risk measures
 - VaR/TVaR model specifications
- **Stranded London Whale**
- Standard VaR/TVaR estimators
- Robust Techniques for Estimating VaR/TVaR
 - Measuring VaR/TVaR model risk
 - Asymptotic efficiency testing
 - Real data testing

VaR/TVaR definition

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.

$$q_X(\alpha) = \inf\{x : \mathbb{P}[X \le x] \ge \alpha\}$$

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

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

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 $VaR(\alpha)$.

$$VaR(\alpha) = \inf\{I : \mathbb{P}[L > I] \le 1 - \alpha\}$$

Left and right quantiles symmetry

An alternative definition of VaR based on the P&L process rather than the loss process.

$$VaR(\alpha) = -\sup\{u : \mathbb{P}[P\&L \le u] \le 1 - \alpha\}$$

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

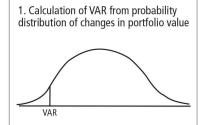
$$q_X^*(\alpha) = \sup\{x : \mathbb{P}[X \le x] \le \alpha\}$$

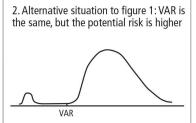
Both VaR definitions agree due to the left and right quantile symmetry.

$$q_X(\alpha) = -q_{-X}^*(1-\alpha)$$

VaR/TVaR definition

One VaR - different riskiness





VaR/TVaR definition

Going further into the tail

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.

VaR/TVaR uses

Various VaR applications

• Trading desk limits - α = .95, h = 1 day, and underlying process is L

$$VaR_{1d}(.95)$$

 Economic capital - α = .999, h = 1 year, and underlying process is L

$$ECapital = VaR_{1Y}(.999) - Expected Loss$$

• Maximum (peak) potential future exposure - α = .95, h = t is a variable, and underlying process is V_t^+

$$\mathsf{MPFE} = \max_{0 \le t \le T} \mathit{VaR}_t(.95)$$

VaR/TVaR uses

Regulatory 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.

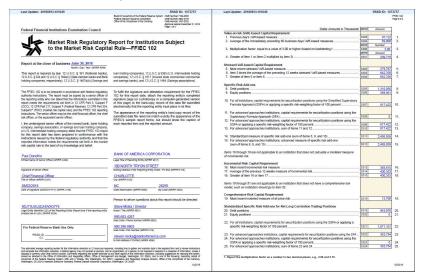
mRWA = max
$$\left\{ VaR_{10d}^{t_0}(.99), \frac{\beta}{60} \sum_{i=1}^{60} VaR_{10d}^{t_0-i\delta t}(.99) \right\}$$

- 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."

VaR/TVaR uses

FFIEC 102 report

https://www.ffiec.gov/nicpubweb/nicweb/NicHome.aspx



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 ρ outperforms VaR in this regard and obeys the following laws.

- **1** (Positive homogeneity) $\rho(aL) = a\rho(L), \forall a \ge 0$
- ② (Sub-additivity) $\rho(L_1 + L_2) \leq \rho(L_1) + \rho(L_2)$
- **3** (Normalized) $\rho(L + a1) = \rho(L) + a$, $\forall a \in \Re$
- **(Monotinicity)** $\rho(L_1) \leq \rho(L_2)$ if $L_1 \leq L_2$ in all scenarios

VaR/TVaR Extensions

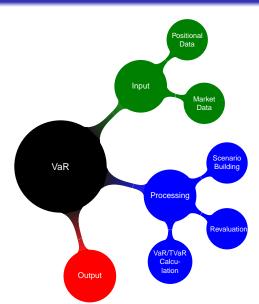
TVaR measure can be generalized by allowing more complicated increasing weighting functions w(u) defined on the interval [0, 1]. This leads us to the definition of the spectral risk measure (SRM).

$$SRM(L) = \int_0^1 w(u) VaR_u(L) du$$

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).

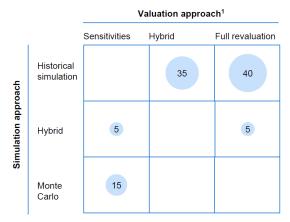
$$DM(L) = \int_0^1 VaR_u(L)dD(u)$$

VaR/TVaR production process



Common VaR models

Market-risk practices at 18 financial institutions, 2011, %



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



VaR/TVaR model specifications

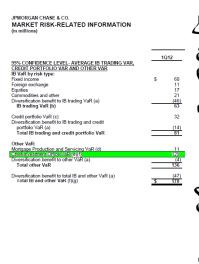
Specification attribute	Options
Revaluation methodology	Sensitivity based
	Full revaluation
VaR estimation methodology	Parametric
	Mixed
	Non-parametric
Simulation object	Relative change
	Absolute change
	Log change
Simulation calibration	Historical
	Market implied
Scenario Weighting	Equal
	Exponential
Confidence level scaling	Yes/No
Horizon scaling	Yes/No

What happened on 05/10/2012?

Testimony of Jamie Dimon Chairman & CEO, JPMorgan Chase & Co. Before the House Financial Services Committee Washington, D.C. June 19, 2012

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.

JPM VaR vs P&L



JPMORGAN CHASE & CO.
CORPORATE/PRIVATE EQUITY
FINANCIAL HIGHLIGHTS
(in millions, except headcount data)

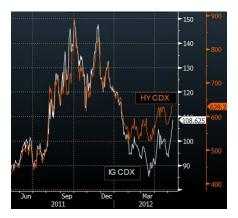
	2Q12			
\$	(3,576)			
	1,013			
	159			
	(2,404)			
_	(205)			
	(2,609)			
	(11)			
	652			
	1,317			
	1,969			
_	(1,410)			
_	559			
	(3.157)			
	(1,380)			
\$	(1,777)			
s	410			
	(3,434)			
	415			
3	(2,609)			
	107			
\$	197			
	(2(078) 104			
\$	(2,078)			
	(2(078) 104			
	\$			

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.

$$\sqrt{n}(V_1 - F^{-1}(p)) \stackrel{dist}{\to} N(0, p(1-p)/f^2(F^{-1}(p))), \qquad n \to \infty$$

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.

$$\sqrt{n}(T_1 - TVaR(p)) \stackrel{dist}{\rightarrow} N(0, \tau_1^2)$$

$$au_1^2 := \frac{1}{(1-p)^2} \int_{F^{-1}(p)}^{\infty} (u - VaR(p))^2 dF(u) - (TVaR(p) - VaR(p))^2$$

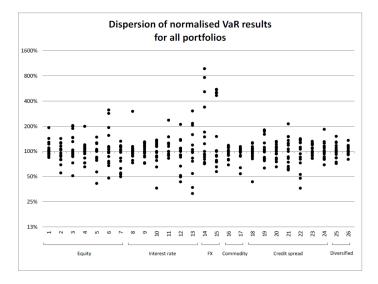
Theoretical VaR/TVaR under normal and Student's t distributions

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.

$$VaR_N(lpha) = \mu + \sigma\Phi^{-1}(lpha)$$
 $VaR_t(lpha) = \mu + \sigma t_{
u}^{-1}(lpha)$
 $TVaR_N(lpha) = \mu + rac{\sigma}{1-lpha} arphi(\Phi^{-1}(lpha))$
 $TVaR_t(lpha) = \mu + rac{\sigma}{1-lpha} g_{
u}(t_{
u}^{-1}(lpha)) \left(rac{
u + (t_{
u}^{-1}(lpha))^2}{
u - 1}
ight)$

Measuring VaR/TVaR model risk

VaR model variance



Measuring VaR/TVaR model risk

Statistical properties of VaR/TVaR estimators, elicitability

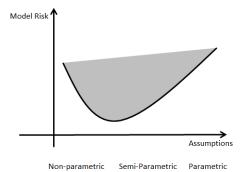
Financial industry tends to assess the quality of the VaR models through back testing. However, back testing is challenging when dealing with TVaR. It has been shown that VaR is elicitable whereas TVaR is not.

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



Testing framework

	VaR(Model,p)	TVaR(Model,p)
Theo		
Р	$V_{1,2,3}$	$V_{1,2,3}$
SP	$R_{1,2}$	$R_{1,2}$
NP	MME/MLE/UMVU	MME/MLE/UMVU

Enhanced non-parametric VaR/TVaR estimators

Theorem. (Joint asymptotic distribution of L-estimators) Let $X_1, X_2, \cdots, X_n \stackrel{iid}{\sim} F$ 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 $\widehat{\Theta}_{1n} = \frac{1}{n} \sum_{i=1}^n L_1(\frac{i}{n+1}) X_{(i)}, \ \widehat{\Theta}_{2n} = \frac{1}{n} \sum_{i=1}^n L_2(\frac{i}{n+1}) X_{(i)}$. If F is absolutely continuous with respect to Lebesgue measure with $E(X_1^2) < \infty$ then we can get.

$$\sqrt{n} \begin{pmatrix} (X_{(np)} - F^{-1}(p)) \\ \widehat{\Theta}_{1n} - J_1(L_1, F) \\ \widehat{\Theta}_{2n} - J_2(L_2, F) \end{pmatrix} \stackrel{dist}{\to} N \begin{pmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \Sigma \end{pmatrix}$$

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.

$$\sigma_2^2 = \sigma^2 \left\{ \frac{1}{4(g(0))^2} + \frac{(1-p)(2p-1)}{2(g(G^{-1}(p)))^2} \right\}$$

Robust non-parametric VaR estimators - 3rd variation

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

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

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

$$\sigma_3^2 = \sigma^2 \left\{ \frac{2\beta^2 \gamma}{(g(G^{-1}(\gamma)))^2} + \frac{(1-2\beta)^2}{4(g(0))^2} + \frac{2\beta(1-2\beta)\gamma}{g(0)g(G^{-1}(\gamma))} + \frac{(1-p)(2p-1)}{2(g(G^{-1}(p)))^2} \right\}$$

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.

Stranded London Whale

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

$$\tau_3^2 = \sigma^2 \left\{ \frac{2\beta^2 \gamma}{(g(G^{-1}(\gamma)))^2} + \frac{(1 - 2\beta)^2}{4(g(0))^2} + \frac{2\beta(1 - 2\beta)\gamma}{g(0)g(G^{-1}(\gamma))} \right\}$$

$$+ \frac{1}{2} \left\{ \frac{1}{(1 - \rho)^2} \int_{F^{-1}(\rho)}^{\infty} (u - VaR(\rho))^2 dF(u) \right\}$$

$$- (TVaR_\rho - VaR(\rho))^2$$

For $\beta = 0$ we get the asymptotic variance τ_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.

Parametric VaR estimators

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.

$$VaR_N^{UMVU}(\alpha) = \overline{X}_n + \frac{S\sqrt{n-1}}{D_n}\Phi^{-1}(\alpha)$$

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

$$VaR_t^{UMVU}(\alpha) = \overline{X}_n + \frac{S\sqrt{n-1}}{D_n(1-\alpha)}\varphi\left(\Phi^{-1}(\alpha)\right)$$

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 $U_1, U_2, \cdots, U_n \stackrel{iid}{\sim} G$, and $E(U_i) = 0$, with $Var(U_i) = \gamma_G^2 =: \gamma^2$. We will make the blanket assumption that $E(X_1^4) < \infty$. However, some of the results hold with only assuming finite variance. Since both VaR(p) and TVaR(p) are of the form, $\mu + \sigma c(p)$, with $E(U_1^2) > 0$, 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)

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

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

$$\begin{split} \sqrt{n}(R_1 - (\mu + c(p)\sigma)) &\stackrel{\textit{dist}}{\to} \textit{N}(0, \nu_1^2) \\ \nu_1^2 := \sigma^2 \left\{ \gamma^2 + 0.25(c(p))^2 (\kappa_G - 1) + \gamma c(p)\psi_G \right\}, \end{split}$$

where ψ_G and κ_G are the skewness and kurtosis of the distribution G.

Semi-parametric VaR/TVaR estimators - 2nd variation

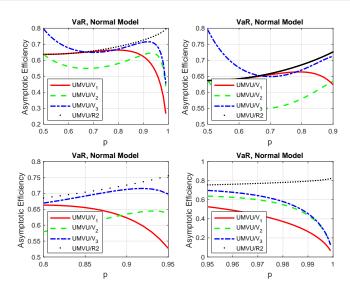
One may consider the following modification of the semi-parametric estimator for VaR(p) or TVaR(p).

$$R_2 := M_n + \frac{Y_n}{\delta} c(\rho),$$

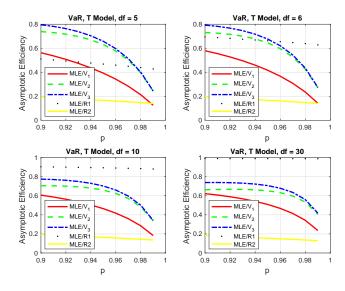
where $\delta = E|U_1 - M_G|$, M_G is the median and $Y_n = \frac{1}{n} \sum_{i=1}^n |X_i - M_n|$. When c(p) is available then one gets asymptotic normality.

$$\nu_2^2 := \sigma^2 \left\{ \frac{1}{4(g(M_G))^2} + \frac{Var(|U_1 - M_G|)(c(p))^2}{\delta^2} - \frac{c(p) M_G}{\delta g(M_G)} \right\}$$

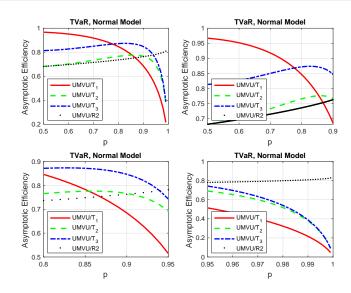
Comparative analysis - VaR for normal model



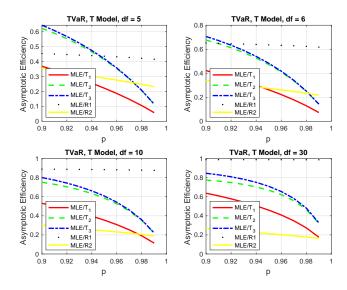
Comparative analysis - VaR for Student's t model



Comparative analysis - TVaR for normal model



Comparative analysis - TVaR for Student's t model



Real data testing

Sample variances of VaR estimators (2009-2018)

Tickers	V_1	V_2	V_3	R_2	UMVU
DB Commodity	58.54	43.98	45.48	38.65	42.64
Emerging Markets	205.68	171.62	175.81	112.79	211.33
High Yield	38.31	26.50	26.59	17.60	40.95
Developed (exc. North America)	97.72	80.72	81.07	63.39	98.32
Short Term Treasury	0.04	0.03	0.03	0.03	0.07
S&P 500	69.03	60.19	62.58	43.80	75.62

Real data testing

VaR breaches (2018)

Tickers	V_1	V_2	V_3	R_2	UMVU
DB Commodity	8.0	7.0	7.0	10.0	8.0
Emerging Markets	15.0	15.0	15.0	18.0	15.0
High Yield	6.0	6.0	6.0	8.0	5.0
Developed (exc. North America)	11.0	12.0	12.0	13.0	11.0
Short Term Treasury	7.0	7.0	7.0	7.0	7.0
S&P 500	16.0	16.0	16.0	16.0	16.0

Stranded London Whale

Conclusions

- Simple nonparametric estimators of VaR(p) and 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 VaR(p) and 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 VaR(p) and 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