Journal of Contemporary Management
Submitted on 22/03/2016
Article ID: 1929-0128-2016-03-62-15 Debasish Majumder

Proposing Model Based Risk Tolerance Level for Value-at-Risk – Is It a Better Alternative to BASEL's Rule Based Strategy?*

Dr. ***Debasish Majumder***
Assistant Adviser, Reserve Bank of India
B-37, RBI Officers' Quarters, Tardeo, Mumbai – 400 034, INDIA
Tel: +919833112190 E-mail: deb_neeta@yahoo.com

Abstract: Since the mid-1990s, BASEL committee of banking supervision sequentially published its guidelines for market risk management for commercial banks, recommending therein rule based strategy for risk tolerance level (or confidence level) for common tail measures. At present, committee's guidelines on confidence level are 99% and 99.9% for Value-at-Risk (VaR) and 97.5% for Expected Shortfall (ES). This rule is probably judgment-based keeping the possibility of becoming suboptimal in some circumstances. Moreover, the approach does not support variation of risk tolerance level across portfolios and across risk scenarios. This may lead to a more serious objection: the BASEL's rule-based strategy cannot differentiate a risky situation from a less risky situation. Conversely, a superior risk management practice might be one, where input parameters are determined by a quantitative process making the risk measure 'non-subjective to regulator's/ or risk modeler's preferences'. A generalized Pareto approximation of the tail part of any common distribution gives us the opportunity to estimate simultaneously the tail size and the starting point of the tail. In other words, the procedure allows us simultaneous estimation of VaR and the risk tolerance level. Minimizing the scope of human intervention in risk measurement, the approach may widen the applicability of tail-related risk models in institutional risk management.

Keywords: Value-at-Risk; Extreme value theory; Generalized Pareto distribution; Tail-related risk models

JEL Classifications: C51, G10, G32, G28

1. Introduction

With the introduction of the regulatory guidelines for market risk management for commercial banks in 1996, popularly known as amendment to BASEL-I accord, banks started to measure and apply capital charges in respect of their market risks in addition to their credit risks. Since then Value-at-Risk (VaR) was regarded as a primary tool to measure the regulatory capital. This risk measure is based on basics of fixing ex-ante a risk tolerance level (or a probability level) which is a small region in the extreme right tail of the underlying density curve. The common perception about this parameter is that since we are interested in protection against adverse market conditions, the risk tolerance level should be kept small. Accordingly, the BASEL committee of banking supervision (BCBS) in 1996 had prescribed the confidence level for VaR at 99%[1]. The resulting risk forecast was backtested by several researchers during the period 1997-2007, findings of whom

* The views expressed in this paper are of the author and not of the organization to which he belongs.

[1] risk tolerance level = 1- confidence level

were mostly supportive to the risk management guidelines by BCBS (Rossignolo et al., 2012, Sharma, 2012). But the scenario had changed after 2008. Market practitioners started questioning the reliability of this risk measure when nearly all financial institutions in USA during 2007-09 and Europe during 2010-12 recorded multiple consecutive exceptions, i.e. days in which ex post loss exceeded the ex-ante VaR. In those circumstances, back tests failed to confirm the correctness of VaR. This raised serious question not only on the empirical validity of the model for high volatile scenarios but also on internal consistency of the entire risk management process.

Before 2008, regulatory prescriptions by BCBS demanded VaR estimation to observe the following quantitative requirements (see BCBS, 1996, 2004, 2006; Rossignolo et al., 2012): (a) Daily-basis estimation; (b) confidence level set at 99%; (c) One-year minimum sample extension with quarterly or more frequent updates; (d) No specific models prescribed: banks are free to adopt their own schemes; (e) regular backtesting and stress testing programme for validation purposes. The severe stress in many developed and emerging markets since 2008 exposed weakness in the framework for capitalizing risks from trading activities. Consequently, BCBS revised their regulatory guidelines for market risk management through sequential prescriptions. The major revisions during 2009-2015 include: a) the boundary between the trading book and banking book was revised in the direction of imposing more stringent rule for internal transfer of assets between them; b) a more rigorous model approval process was to be followed in the case of internal model approach; c) standardized approach was revised to make it sufficiently risk-sensitive; d) a shift from Value-at-Risk (VaR) to an Expected Shortfall (ES) measure of risk under stress, confidence level for ES to be set at 97.5%; e) Adopting separate internal model to measure the default risk of trading book positions; f) default risk to be measured using a VaR model, the VaR calculation must be done weekly and would be based on a one-year time horizon at a one-tail with 99.9% confidence level; g) Incorporation of the risk of market illiquidity; h) higher model validation requirements: daily Profit and Loss attribution, model backtesting and regular model validation reporting (see BCBS, 2009a, 2009b, 2011, 2014, 2016). These guidelines reveal that BCBS followed rule based strategy for confidence level (or risk tolerance level) for Value-at-Risk (VaR) and Expected Shortfall (ES). Unfortunately, the committee did not disclose the logic behind those rules. An inevitable criticism might be: why does the percentages (confidence level) are chosen as 99% or 99.9% or 97.5% and not any other?

BCBS might have followed the common belief that the risk tolerance level should be kept small as possible to cover extreme downside risks. But a relevant question raised by Kerkhof & Melenberg (2004) and many others was: how small this parameter should be? The controversy may gain another dimension with a more serious question: are BASEL risk tolerance levels efficient to capture tail risk? It is relevant to quote an explanation by Linsmeier and Pearson (2000): “The conventional theory provides little guidance about the choice of the probability level. It is determined primarily by how the designer of the risk-management system wants to interpret the VaR number”. Effectively, the risk modeler reduces his losses in states that occur with (100-α)% probability, but ignores the α% of states that are not included in the computation of VaR (Berkelaar, Cumperayot and Kouwenberg (2002)). Theoretically, a loss belonging to the specified α% region signals insolvency or distress of a business organization (Jackson et al., 1997; Frenkel, Hommel & Rudolf, 2005). In practice, however, VaR violations are not always very harmful to investors and hence might not be extreme events in reality. This argument was supported by Dan ́elsson (2002) by stating that VaR violations frequently have little relevance to bankruptcy, financial crashes, or systemic failures. The reason might be that low probability events are not equally harmful in all circumstances. For example, we may consider two portfolios: one, a portfolio of Government treasury bills and, second, a market portfolio in a volatile stock market. For both these portfolios, notwithstanding the fact that the expected number of violations in 99% VaR would be one, this event in the former case is not equally undesirable as in the latter. Critics may pose a question:

"why do we keep the tail probability same for portfolios with diversified price risks?". Unfortunately, BASEL's rule-based risk tolerance level cannot differentiate a risky situation from a less risky situation and therefore does not support variation of that parameter across portfolios and across risk scenarios. Conversely, a superior risk management system might be one, where the risk tolerance level is not pre-assigned, but endogenously determined by the risk model. The rest of the paper is organized as follows: Section 2 provides disadvantages in rule based risk tolerance level. Section 3 proposed a model-based approach for the risk tolerance level. Section 4 describes the model. Section 5 provides empirical findings and Section 6 concludes.

2. Disadvantages in Rule-based Risk Tolerance Level

Following example illustrates how the choice of the risk tolerance level affects the final risk estimates. The daily VaR using historical simulation based on rolling windows of 500 trading days has been computed on daily losses based on 12 leading indices from different markets during a period of 30 years from 15th Feb 1985 to 16th Feb 2015. For this computation, three commonly used risk levels, e.g. 0.01, 0.05 and 0.10, are considered. Results indicate that a reduction in the risk level from 0.05 to 0.01 or from 0.10 to 0.01 translates into a heterogeneous increase in the daily VaR for select indices (Tables 1 and 2). Although this increase in the VaR remains at less than 2 percentage points for majority of trading days for all select indices, we can find a significant number of days for each index, for which the VaR difference (e.g. $VaR_{.01}$-$VaR_{.05}$ or $VaR_{.01}$-$VaR_{.1}$) lies in the higher class intervals, namely [2,3) or [3,4) or [4,5) or even more. Therefore, it is possible to discover a number of scenarios where we can find that a variation in choices of the risk level resulting in significant variations in the VaR estimates. When this variation is 3 percentage points or more, the change in risk level may attract a substantial additional risk capital for an equity investor. The number of such instances is not insignificant in the present example. For instance, in 991 trading days among the total of 7327 trading days, the VaR difference (e.g. $VaR_{.01}$-$VaR_{.05}$) for AEX index is equal to or more than 3 percentage points (Table 1). These findings are in the similar line of Degennaro (2008) who formed examples to illustrate that non-cooperative choices of the risk tolerance level by two investors would be resulting in a substantial variation in their VaR estimates. These incidents could be interpreted in a way of uncovering the economic cost of improper setting of the cutoff point for "likely" and "unlikely" events. Not surprisingly, the rule for setting the cutoff, as recommended by the BCBS, could be suboptimal, because the procedure is perhaps not backed by quantitative techniques. Often, the economic costs of this suboptimal policy is enormous, which may affect the financial health of an organization. Conversely, a model based selection of the cutoff is expected to produce better risk estimates. The additional advantage of this procedure is that it allows time variation of the risk tolerance level.

Table 1. Frequency distribution of the increase in VaR due to reduction in risk level from 0.05 to 0.01

Market Index	**Country**	**Number of trading days for which ($VaR_{.01}$-$VaR_{.05}$) lies in**						**Total Number of days**
		[5, ∞)	**[4,5)**	**[3,4)**	**[2,3)**	**[1,2)**	**[0,1)**	
S&P 500 Composite Index	USA	0	0	63	912	2,165	4,187	7,327
FTSE 100 Index	UK	0	0	0	1,396	1,664	4,267	7,327
DAX 30 Index	Germany	0	0	530	1,581	2,745	2,471	7,327
Dow Jones Industrial Index	USA	0	0	0	754	2,757	3,816	7,327

S&P/TSX Composite Index	Canada	0	0	96	853	1,789	4,589	7,327
NIKKEI 225 Index	Japan	0	0	374	316	5,393	1,244	7,327
HANG SENG Index	Hong kong	0	75	95	2,333	3,196	1,628	7,327
NYSE Composite Index	USA	0	0	464	537	1,981	4,345	7,327
OMXS Index	Sweden	0	11	490	984	3,809	2,033	7,327
AEX Index	Netherlands	0	12	979	675	3,595	2,066	7,327
NASDAQ Composite Index	USA	0	0	0	1,519	2,764	3,044	7,327
KOREA SE Composite Index	South Korea	0	0	417	1,116	4,280	1,514	7,327

Table 2. Frequency distribution of the increase in VaR due to reduction in risk level from 0.10 to 0.01

Market Index	**Country**	**Number of trading days for which ($VaR_{.01}$-$VaR_{.10}$) lies in**						**Total number of days**
		[5, ∞)	**[4,5)**	**[3,4)**	**[2,3)**	**[1,2)**	**[0,1)**	
S&P 500 Composite Index	USA	0	37	504	857	4,291	1,638	7,327
FTSE 100 Index	UK	0	0	528	1,011	3,938	1,850	7,327
DAX 30 Index	Germany	0	412	1,238	1,738	3,792	147	7,327
Dow Jones Industrial Index	USA	0	0	586	1,398	4,080	1,263	7,327
S&P/TSX Composite Index	Canada	0	152	353	945	3,892	1,985	7,327
NIKKEI 225 Index	Japan	0	499	37	2,537	4,170	84	7,327
HANG SENG Index	Hong kong	75	285	1,293	2,494	3,167	13	7,327
NYSE Composite Index	USA	6	464	61	1,065	4,073	1,658	7,327
OMXS Index	Sweden	0	497	497	1,956	3,950	427	7,327
AEX Index	Netherlands	0	991	470	1,480	3,949	437	7,327
NASDAQ Composite Index	USA	0	0	867	1,553	4,535	372	7,327
KOREA SE Composite Index (KOSPI)	South Korea	0	569	897	2,489	3,114	258	7,327

Note: These tables display the frequency distribution of the increase in VaR due to reduction in the risk levels from 0.05 to 0.01 (Table 1) and 0.10 to 0.01(Table 2), respectively. The daily VaR has been computed based on historical simulation using rolling windows of daily losses of previous 500 days for 12 stock market indices during a period of 30 years from 15th Feb 1985 to 16th Feb 2015.

3. A Model-based Risk Tolerance Level --- An Alternative to BASEL?

We have already stated that the theory of Value-at-Risk (VaR) is based on basics of selecting a region in the extreme right tail part of the loss distribution. The region corresponds to large abnormal losses which exceed VaR. These losses are rare in nature, however occurrence of them may trigger solvency of the business firm. It is implicit that losses belonging to the rest of the distribution generate risk, which is somewhat manageable by the business firm and hence do not trigger solvency. The critical question may arise: what would be the size of the specified region? In practice, in the case of commercial banks, the size is frequently pre-assigned by a banking regulator (Rossignolo et al., 2012; Kerkhof & Melenberg, 2004). In the case of non-banking institutions, the size is fixed either by a risk manager or a risk management committee. In both the cases, however, the size is exogenous to the risk modeler.

In an alternative approach, the present paper proposes that the size ought not to be pre-assigned, but may be endogenously determined by the risk model. In this framework, the size may vary with the shape of the loss distribution: the size would be larger for a heavy tailed distribution and would be smaller for a relatively thinner tailed distribution. One way to determine the size might be using the Pickands-Balkema-de Haan theorem which essentially says that, for a wide class of distributions, losses which exceed the high enough threshold follow the generalized Pareto distribution (GPD) (Balkema & de Haan, 1974; Pickands, 1975). Using this theorem, it is easy to establish that the extreme right tail part of a distribution asymptotically converges to the tail of a generalized Pareto distribution (GPD). This hypothesis reveals that we can always find a region in the extreme right tail of the loss distribution, for which the equivalent region from a suitable GPD is available.

The GPD is often regarded as a basic distribution to model the heavy tail properties of the data and, therefore, considered as an indispensable part of risk management in general and the VaR calculations in particular (Gençay & Selçuk, 2004; McNeil & Frey, 2000; Ghosh & Resnick, 2010; Goldberg et al., 2008). The GPD is specified by three parameters: location, scale and shape. For estimating these parameters, a variety of different methods were employed by several scholars (see Brazauskas & Kleefeld, 2009). In this distribution, the tail is characterized by a shape parameter, the reciprocal of which is known as the 'tail index'. The tail index signifies the rate of the tail decay, which is a measure of heaviness of the tail. Conversely, the location parameter specifies the starting point of the tail. However, for any other distribution, it is merely impossible to find a starting point of the tail unless the tail part could be approximated by an appropriate GPD. Pickands-Balkema-de Haan theorem provides the theoretical ground to model the tail part of any distribution through estimating the tail of a suitable GPD. According to this theorem, there exists a threshold, the data above which shows generalized Pareto behavior. The threshold would essentially be reasonably large to cover all events which are 'extreme' in nature. Conversely, we may argue that events belonging to the rest of the distribution are 'normal' or 'non-extreme' in nature. The approach requires a parametric estimation of the density function of two distributions and the non-parametric estimate of the cut-off point. The cut-off is a point which fits the tails of the two distributions (i.e the underlying distribution and the GPD) best. The point essentially separates ordinary realizations of the random variable from extreme realizations of the same variable. Therefore, the underlying distribution of the loss (X) can be separated into two parts: $X \geq \hat{y}$ is the risky region of the distribution, occurrence of such losses may trigger solvency of the business firm. All large unforeseen losses belong to this part. Conversely, $X < \hat{y}$ is the region of the distribution which does not cause extreme/ solvency risks. $\hat{y}$ is the threshold, above which the tail of a given distribution can be approximated by a GPD. In the line of our earlier analysis, $\hat{y}$ is the starting point of the tail for any common distribution. We can find the region $X \geq \hat{y}$ for a wide class of distributions which would be the 'tail region' and the threshold ($\hat{y}$) would be VaR of the given distribution. If we get a loss distribution for which the GPD approximation of the tail does not hold, we would not get a finite VaR. In this case, the loss distribution does not have tail/ solvency risk as such.

Our approach differs from conventional characterization of events of exceedance of VaR. Such events are conventionally those which belong to the pre-specified small region laid in the tail part of the density curve. However, based on theories of extreme values, we propose a more logical specification of events violating VaR: i) these events belong to the extreme tail part of the underlying loss distribution and ii) these events could be modeled as a tail of a suitable generalized Pareto distribution (GPD). (ii) poses an additional constraint to the conventional characterization. But, (i) and (ii) might be the basis of a more logical framework for identifying the size of the specified tail region (or risk tolerance level).

4. The Model

4.1 Behavior of losses exceeding a high threshold

Suppose X_1; X_2; ---; X_n are n independent realizations from a random variable (X) representing the loss with distribution function $F_X(x)$ having a finite or infinite right endpoint (x_0). We are interested in investigating the behavior of this distribution exceeding a high threshold (u). In the line of Hogg and Klugman (1984), the distribution function $\left(F_{Y_1^u}\right)$ of the truncated loss (Y_1^u) (truncated at the point u) can be defined as:

$$F_{Y_1^u}(x) = P\left[Y_1^u \leq x\right] = P\left[X \leq x/X > u\right] = \begin{cases} 0 & \text{if } x \leq u \\ \dfrac{F_X(x) - F_X(u)}{1 - F_X(u)} & \text{if } x > u \end{cases}$$

Based on $F_{Y_1^u}$, we can define the distribution function of the excesses over a high threshold u:

$$F_{Y^u}(x) = P\left[X - u \leq x/X > u\right] = \frac{F_X(x+u) - F_X(u)}{1 - F_X(u)} \tag{1}$$

$$\text{for } 0 \leq x < x_0 - u$$

Balkema and de Haan (1974) and Pickands (1975) showed that, for a large class of distributions, the generalized Pareto distribution is the limiting distribution for the distribution of the excesses, as the threshold (u) tends to the right endpoint. The distribution function of a three parameter generalized Pareto distribution with shape parameter (ξ), location parameter (u) and scale parameter (σ) has the following representation:

$$G_{\xi,u,\sigma}(x) = \begin{cases} 1 - \left(1 + \xi\left(x - u\right)/\sigma\right)^{-1/\xi} & \text{if } \xi \neq 0 \\ 1 - \exp(-\left(x - u\right)/\sigma & \text{if } \xi = 0 \end{cases}$$

where $\sigma > 0$, $(x - u) \geq 0$ when $\xi \geq 0$ and $0 \leq (x - u) \leq -\frac{\sigma}{\xi}$ when $\xi < 0$. Balkema and de Haan (1974) and Pickands (1975) asserted that we can find a positive measurable function $\sigma(u)$ such that

$$\lim_{u \to x_0} \sup_{0 \leq x < x_0 - u} \left| F_{Y^u}(x) - G_{\xi,\sigma(u)}(x) \right| = 0 \qquad (2)$$

If the equation (2) holds, we can get a sufficiently high threshold u, above which the distribution function of the excesses may be approximated by $G_{\xi,\sigma(u)}(x)$ for some values of ξ and $\sigma(u)$. This provides us a theoretical ground to claim that there exists a threshold, the data above which will show generalized Pareto behavior.

4.2 Identifying the tail region

The equation (1) and (2) suggest that for a sufficiently high threshold, it can be written:

$$F_X(x + u) \approx F_X(u) + G_{\xi,\sigma(u)}(x)(1 - F_X(u))$$

Setting y = x+u

$$F_X(y) \approx F_X(u) + G_{\xi,\sigma(u)}(y - u)(1 - F_X(u)) \qquad (3)$$

The right hand side of the equation (3) can be simplified in the form of a distribution function of a generalized Pareto distribution. Therefore, the equation (3) can be rewritten as:

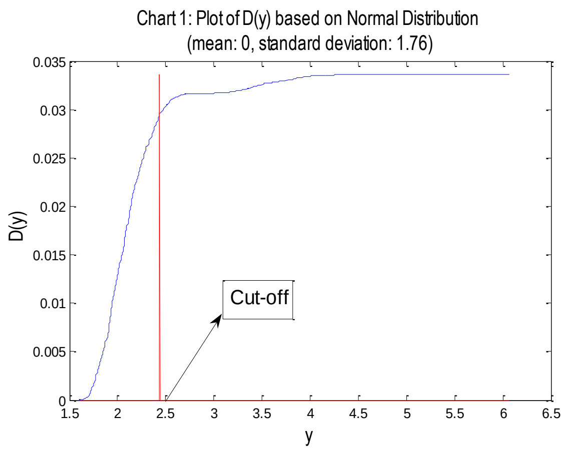
$$F_X(y) \approx G_{\xi,\tilde{\sigma}}(y - \tilde{\mu}) \qquad (4)$$

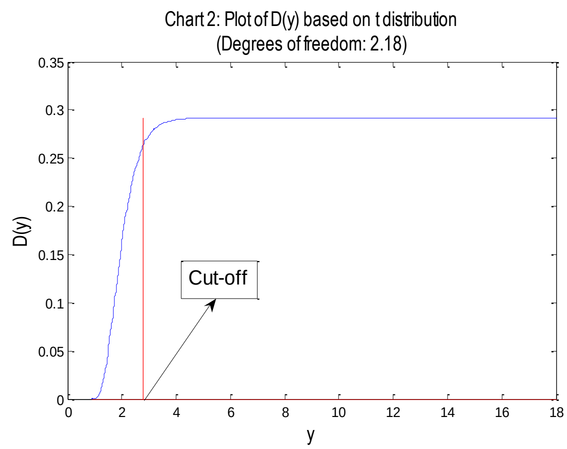
where $\tilde{\sigma} = \sigma(1 - F_X(u))^{\xi}$ and $\tilde{\mu} = u - \tilde{\sigma}\left((1 - F_X(u))^{-\xi} - 1\right)/\xi$.

Hence, if we can fit the GPD to the conditional distribution of the excesses above a high threshold, it can also be fitted to the tail of the original distribution above a certain threshold (also see Reiss & Thomas, 1997). However, the minimum threshold above which the GPD will hold may not be same for the conditional loss distribution of the excesses and the original loss distribution. If we can identify the minimum threshold (say $\hat{u}$) for the conditional loss distribution, it is possible get a cut-off (say $\hat{y}$) above which the given distribution can be approximated by a GPD. The procedure for estimating $\hat{y}$ is described below.

When u is fixed at $\hat{u}$, $\hat{y}$ would be the minimum value of y for which the equation (4) will hold. The deviation of $F_X(y)$ from $G_{\xi,\tilde{\sigma}}(y - \tilde{\mu})$ would, therefore, be non-zero for $y < \hat{y}$, which is expected to be zero for all $y \geq \hat{y}$. We may consider an indicator, viz. the cumulative square deviation for $y < y_0$, $D(y_0) = \sum_{y < y_0} \left[F_X(y) - G_{\xi,\tilde{\sigma}}(y - \tilde{\mu})\right]^2$, which might be useful for identifying $\hat{y}$. By its nature, $D(y_0)$ would be an increasing function of y_0 for $y_0 < \hat{y}$ and would be nearly flat for $y_0 \geq \hat{y}$. Therefore, the slope of the $D(y_0)$ would be positive for $y_0 < \hat{y}$, which would be almost zero for $y_0 \geq \hat{y}$. We can identify the cut-off point, $\hat{y}$, after which the slope of the $D(y_0)$ would be statistically insignificant (for detail, see Appendix B).

Results obtained for normal and t distribution are reported in Charts 1 & 2 respectively which would validate our postulate. Based on a set of random numbers of size 4000 from normal distribution with mean 0 and standard deviation 1.76, we have computed D(y) for various values of y. A plot of D(y) in respect of y reveals that D(y) is almost flat after a certain cut-off (Chart 1). Similar results can also be observed for t distribution with degree of freedom 2.18 (Chart 2).





Based on the above analysis, we can bifurcate the underlying distribution into two parts: $X \geq \hat{y}$ is the risky region of the distribution in the sense that this region could be approximated by the tail of an equivalent GPD. All large unforeseen losses would belong to this part. Conversely, $X < \hat{y}$ is the region of the distribution which does not cause severe tail risk. Losses belong to this region may generate risk which is somewhat manageable by the business firm and hence do not trigger solvency.

4.3 Measuring the Tail Risk

For a small quantile of order p, $P = 1 - F_X(\hat{y})$, we can write

$$P \approx (1 - F_X(\hat{u}))(1 - G_{\varsigma,\sigma(u_0)}(\hat{y} - \hat{u})) \qquad (5)$$

Our earlier analysis reveals that VaR_p is the p-th quantile point cut from the distribution, $F_X(.)$. Therefore,

$$VaR_p = \hat{y} \qquad (6)$$

Equations (6) and (7) lead to interesting inferences: when the distributional form of the underlying distribution F(x) is known, p and VaR_p can be estimated simultaneously. Value-at-Risk to be resulted by the above set of equations does not depend on the ex-ante setting of the tolerance probability and may be named as non-subjective VaR (VaR^{N-S}).

Proposition 1: Non-subjective VaR (VaR^{N-S}) exists for all common distributions.

Proof: The Pickands-Balkema-de Haan theorem holds if the underlying distribution function, $F_X(x)$, belongs to the maximum domain of attraction of a generalized extreme value distribution (GEV) (H). It is denoted by $X \in MDA(H)$. H must be one of the three possible forms: the Fréchet, the

Gumbel and the Weibull families. These three types of distributions can be nested into a single class of continuous probability distributions – the GEV distribution (Embrechts et al., 1997):

$$H_{\xi}(x) = \exp\left\{-(1+\xi x)^{-\frac{1}{\xi}}\right\} \text{if } \xi \neq 0$$
$$= \exp\left(-e^{-x}\right) \text{if } \xi = 0$$

where $(1+\xi x) > 0$. If $\xi > 0$, we have the Fréchet distribution; if $\xi < 0$ we have the Weibull distribution; $\xi = 0$ gives the Gumbel distribution. Gnedenko (1943) showed that if the tail of the $F_X(x)$ decays like a power function, then the distribution is in the domain of attraction of the Fréchet. This class of distributions is quite large and includes Pareto, Burr, loggamma, Cauchy and t-distributions as well as various mixture models. Distributions in the domain of attraction of the Gumbel, include normal, exponential, gamma and lognormal distributions. These distributions are medium-tailed distributions. Distributions in the domain of attraction of the Weibull are short tailed distributions such as the uniform and beta distributions. As stated in McNeil and Saladin (1997), all common continuous distributions belong to the above three categories. For all these distributions, we can find a threshold level, $\hat{u}$, above which the distribution function of the excesses may be approximated by a GPD. Simultaneously, it is possible get a cut-off (say $\hat{y}$) above which, the given distribution can be approximated by a GPD. Clearly, $VaR^{N\text{-}S}$ exists for all these distributions. Conversely, if we get a distribution, $X \notin MDA(H)$, we cannot find the threshold level, $\hat{u}$, above which the distribution function of the excesses may be approximated by a GPD. $VaR^{N\text{-}S}$ does not exist for this distribution.

4.4 Simulation study for threshold choice

Unfortunately, estimation of the model designed by us is feasible only when the threshold level, $\hat{u}$, above which the Pareto law is expected to hold, is known. When $\hat{u}$ is unknown, the model is unable to generate a simultaneous solution for Value-at-Risk and the risk tolerance level. Many scholars, however, developed empirical methods for estimating the threshold (Hill, 1975; Meerschaert & Scheffler, 1998). They proposed extreme value theory to model the tail part of the distribution, irrespective of the shape of the loss distribution.

When the form of the underlying loss distribution $F_X(.)$ is known, we can develop a procedure for estimating the threshold, $\hat{u}$, by a simulation study. We may recall our result in the preceding section that we can get a sufficiently high threshold u, above which the distribution function of the excesses $F_{Y^u}(x)$ can be approximated by the distribution function of a generalized Pareto distribution, $G_{\xi,\sigma(u)}(x)$. Initially, we fix u to some u' and generate 100 samples each of size 4000 from the underlying distribution F_X. If u' is the true threshold, then the deviation of $F_{Y^{u'}}(x)$ from $G_{\xi,\sigma(u')}(x)$ is expected to be zero for all $x \geq u'$ for the j th sample, j=1,2,----,100. We may consider an indicator, viz. the cumulative square deviation for $x \geq u'$, $D2(u') = \sum_{x \geq u'} \left[F_{Y^{u'}}(x) - G_{\xi,\sigma(u')}(x)\right]^2$, which might be useful for identifying the threshold. If u' is the true threshold, $D2(u')$ would be zero for each sample. Based on this indicator, we can form a Mean Squared Error (MSE):

$$MSE(u') = \frac{1}{100}\sum_{i=1}^{100}\frac{\{D2(u')\}}{n_i}$$

where n_i is the number of observation in the i th sample exceeding u'. $MSE(u)$ can be computed for various values of u starting from 0. In practice, values of u may be increased with a fixed increment, say 0.1. The best estimate of u (say $\hat{u}$**)** would be such that where $MSE(u)$ is minimum.

5. Empirical findings: A comparison between VaR and VaR$^{N\text{-}S}$

5.1 A comparison based on some standard distributions

VaR and VaR$^{N\text{-}S}$ based on innovations from normal and student's t distributions are reported in Table 3. In the case of normal distribution, an increase in standard deviation from 1.76 to 3.51 translates into a larger VaR and also VaR$^{N\text{-}S}$. The outcome is directionally similar in the case of t distribution when we reduce the degrees of freedom from 2.95 to 2.18[2]. Therefore, an increase in the volatility leads to an amplification of the computed VaR as well as VaR$^{N\text{-}S}$. Furthermore, it can be observed that for normal as well as t distributions, the equilibrium risk tolerance level in VaR$^{N\text{-}S}$ lies in-between 0.05 and 0.1 which leads VaR$^{N\text{-}S}$ to belong in-between VaR$_{.1}$ and VaR$_{0.05}$. For both the select distributions, the estimate of the equilibrium risk tolerance level and also VaR$^{N\text{-}S}$ are reasonably reliable to be accepted. The advantage of this measure is in eliminating subjectivity that is found in the conventional VaR. In the conventional VaR, the risk tolerance level (p) is choice-based and, therefore, this parameter does not necessarily depend on the tail risk of the distribution.

Table 3. A comparison between VaR and VaR$^{N\text{-}S}$ (based on Normal and t distributions)

Distribution	Conventional VaR			Non-subjective VaR		
	VaR$_{.01}$	**VaR$_{.05}$**	**VaR$_{0.1}$**	**Threshold ($\hat{u}$)**	**Risk tolerance level (p)**	**VaR$^{N\text{-}S}$**
Normal Distribution (Mean = 0, Standard deviation = 1.76)	4.09	2.89	2.25	1.6	0.065 (0.023)	2.72 (0.405)
Normal Distribution (Mean = 0, Standard deviation = 3.51)	8.16	5.77	4.49	2.2	0.063 (0.022)	5.468 (0.625)
Student's t Distribution (Degrees of freedom = 2.95)	4.60	2.36	1.64	0.9	0.074 (0.018)	2.03 (0.503)
Student's t Distribution (Degrees of freedom = 2.18)	6.23	2.76	1.81	0.9	0.071 (0.020)	2.40 (0.717)

Note: This table reports the conventional VaR with risk tolerance levels 0.01, 0.05 and 0.1 and non-subjective VaR (VaR$^{N\text{-}S}$) based on the normal and the student's t distributions. The threshold ($\hat{u}$), the risk tolerance level (p) and VaR$^{N\text{-}S}$ as shown in this table are average based on 50 estimates. The standard deviation of the estimate is shown in the parenthesis.

[2] Some authors allow non-integer degrees of freedom of Student's distribution in VaR computation (see Heikkinen, V. P., and Kanto A. (2002)).

5.2 A comparison based on market data

VaR and VaR^{N-S} based on daily losses on S&P 500 Composite Index for the period of 30 years, from 18th February, 1985 to 17th February, 2015, computed using four risk models separately for the full sample and simulated stress scenario are reported in Table 4. Stress scenarios are simulated in the line of Studer (1997,1999) and Breuer and Krenn (1999), who employed the Mahalanobis distance as a mathematical tool to choose stress scenarios (for detail see Appendix A). Additionally, the Conditional EVT framework proposed by McNeil and Frey (2000) was adopted to compute VaR^{N-S} for GARCH. Results indicate that in every occasion VaR^{N-S} lies in-between $VaR_{.1}$ and $VaR_{0.05}$ and hence is not a too arbitrary number to be accepted as a risk measure. The equilibrium risk tolerance level in VaRN-S lies in-between 0.05 and 0.1. Furthermore, for each risk model, VaRN-S in the stress scenario is greater than the estimate of the same for the full sample indicating that the new risk measure correctly captures riskiness of markets. VaRN-S, has natural advantages over VaR, as the latter relies on Judgmental selection of the risk tolerance level which may induce arbitrariness. The additional advantage of the new VaR measure might be that the methodology allows time variation of the risk tolerance level.

Table 4. A comparison between VaR and VaR^{N-S} based on S&P 500 Composite Index

Scenario	Model	Conventional VaR			Non-Subjective VaR		
		$VaR_{.01}$	$VaR_{.05}$	$VaR_{0.1}$	Thres-hold ($\hat{u}$)	Risk tolerance level (p)	VaR^{N-S}
Uncondi-tional	Normal	2.60 (0.066)	1.82 (0.038)	1.41 (0.030)	0.8	0.0609 (0.024)	1.75 (0.242)
	t location-scale	3.21 (0.215)	1.55 (0.056)	1.04 (0.034)	0.4	0.0724 (0.018)	1.33 (0.349)
	ARMA-GARCH with normal innovation	2.66 (0.066)	1.89 (0.038)	1.47 (0.029)	0.9	0.0638 (0.023)	1.80 (0.280)
	ARMA-GARCH with t innovation	3.28 (0.211)	1.61 (0.056)	1.10 (0.032)	0.8	0.0727 (0.017)	1.37 (0.251)
simulated stress scenario	Normal	4.41 (0.111)	3.12 (0.623)	2.43 (0.054)	1.5	0.0643 (0.022)	2.95 (0.420)
	t location-scale	4.53 (0.144)	3.01 (0.069)	2.28 (0.052)	1.8	0.0626 (0.021)	2.86 (0.480)
	ARMA-GARCH with normal innovation	4.48 (0.111)	3.19 (0.064)	2.50 (0.050)	1.5	0.0634 (0.022)	3.02 (0.409)
	ARMA-GARCH with t innovation	4.57 (0.138)	3.06 (0.067)	2.34 (0.050)	1.5	0.0670 (0.020)	2.82 (0.427)

Note: VaR and VaR^{N-S} are average based on 50 estimates. The standard deviation of the estimate is provided in the parenthesis.

6. Conclusion

A large class of common risk measures is based on basics of fixing ex-ante a risk tolerance level or a probability level which is a small region in the extreme right tail of the underlying density curve. In practice, this parameter is generally fixed by judgment/ or perception by a risk manager or a risk management committee or, in certain cases, an external regulatory body. Value-at-Risk (VaR) and Expected Shortfall are common examples of this category of risk models. In practice, the risk tolerance level (or confidence level) is commonly set in the line of recommendation by the BASEL committee of banking supervision. Alternatively, in some cases, financial practitioners/ researchers adopted commonly used percentages viz. 99%, 95% and 90% for this purpose. At present, the BASEL guideline is 99% and 99.9% confidence level for VaR and 97.5% confidence level for ES. Unfortunately, the committee did not disclose the logic behind this guideline. However, the choice of this parameter may have a significant impact on the risk measure, particularly in a scenario when a small change in the risk tolerance level translates into a large difference in the threshold. Nevertheless, such scenarios are not uncommon in financial markets.

We accumulated examples based on indices from different markets where a reduction in the risk tolerance level from 0.05 to 0.01 or from 0.10 to 0.01 led to a significant rise in the daily VaR. Often, this increase was more than 3 percentage points, which might attract a substantial additional risk capital for an equity investor. The number of such instances is not insignificant for each market and hence they are difficult to be ignored. They could be interpreted in a way of uncovering the economic cost of improper setting of the cutoff point for “likely” and “unlikely” events. The analysis might be relevant because we cannot ignore the possibility of the BASEL rule to be suboptimal in some circumstances, given that the methodology is probably not backed by quantitative techniques. Further, the approach does not support variation of this parameter across portfolios and across risk scenarios. Conversely, a superior risk management practice might be one, where input parameters are determined by a quantitative process which makes the risk measure ‘non-subjective to regulator’s/ or modeler’s preferences’. A generalized Pareto approximation of the tail part of any common distribution gives us the opportunity to estimate simultaneously the tail size and the starting point of the tail. In other words, the procedure allows us simultaneous estimation of VaR and the risk tolerance level. The model inherently detects the tail region which is a specified region from a given distribution converging asymptotically to the tail of an equivalent generalized Pareto distribution (GPD). The outcome of the model is a non-subjective VaR number which might be comparable across portfolios or across scenarios of an equity market. Our empirical study based on S&P 500 composite index reveals that the tail risk (or heaviness of the tail) of the loss distribution is well captured by the new risk measure in the normal scenario as well as in the stressed scenario. In addition, our approach might be a natural guide to the future research on computing model based risk tolerance level for Value-at-Risk measure and also for other tail related risk measures, which would perhaps widen the scope of this kind of risk models in institutional and regulatory policymaking.

Appendix A: Choice of stress scenarios

The usefulness of stress tests as a risk management tool crucially depends on the choice of stress scenarios which was traditionally picked by hand. Handpicked scenarios, however, might misrepresent risks, either because such scenarios are too implausible or because really dangerous scenarios might not have been considered (Breuer & Csiszár, 2013). For overcoming pitfalls of performing stress tests with handpicked scenarios, we adopted the approach by Studer (1997,1999) and Breuer and Krenn (1999), who used the Mahalanobis distance as a mathematical tool to choose stress scenarios. The Mahalanobis distance (Maha) of the realisation r from the expectation, E(r), is defined by:

$$\text{Maha}(r) = \sqrt{(r - E(r))^T \Sigma^{-1} (r - E(r))}$$

where Σ is the covariance matrix. The trust region is an ellipsoid for some given Mahalanobis radius k:

$$\text{ELL}(k) = \{ r : \text{Maha}(r) \le k \}$$

Given that the distribution of the risk factor is elliptical, Studer's systematic stress test method suggested the choice of suitable Mahalanobis ellipsoids as sets of plausible scenarios. Conversely, the less plausible scenarios are severe scenarios that lie over admissible domain of plausible scenarios (Breuer et al. (2009)). A high value of Maha implies a low plausibility of the scenario r. If we want to get more severe scenarios, we choose a higher k and get less plausible worst-case scenarios. If we want to get more plausible scenarios, we choose a lower k and get less severe worst-case scenarios.

Appendix B: The choice of the cut-off point $(\hat{y})$

The cut-off point $(\hat{y})$ is the minimum threshold above which the given distribution can be approximated by a GPD. For choosing $\hat{y}$, we may consider an indicator, the cumulative square deviation for $y < y_0$, $D(y_0) = \sum_{y<y_0} \left[F_X(y) - G_{\xi,\tilde{\sigma}}(y - \tilde{\mu}) \right]^2$, which would be an increasing function of y_0 for $y_0 < \hat{y}$ and expected to be flat for $y_0 \ge \hat{y}$. We can write:

$$D(y_0) = \begin{cases} f(y_0) & \text{for } y_0 < \hat{y} \\ \alpha + \beta . y_0 & \text{for } y_0 \ge \hat{y} \end{cases}$$

where $f' > 0$. α and β are constant. The statistical significance of β can be tested using a Wald test with statistic:

$$W = \frac{\hat{\beta}^2}{\text{Variance}(\hat{\beta})}$$

where $\hat{\beta}$ is the maximum likelihood estimate of β. W follows a χ^2 distribution with 1 degree of freedom. Initially, we choose a trial $\hat{y}$, an observation from the extreme right tail of the distribution of X. For an example, initial $\hat{y}$ might be the (N-15) th observation in a sample of size N where observations are arranged in an increasing order. Using 15 observations belonging to the extreme right tail part, we can evaluate the initial $\hat{\beta}$ and hence the initial W. Statistical significance may be checked for the initial estimate of W and also for subsequent estimates of W to be obtained in each step. We include one more observation in each step and then re-estimate W for the purpose of checking its significance. In practice, we may set trial values of $\hat{y}$ as (N-15) th observation, (N-16) th observation, (N-17) th observation and so on. Finally, we select the cut-off point, i.e. minimum $\hat{y}$ for which the W is insignificant at 5% level. Naturally, W is significantly different from zero if we select the value of $\hat{y}$ less than the cut-off**.**

References

[1] Balkema, A. A., and Haan L. de. (1974). “Residual lifetime at great age”, *Annals of Probability*, 2(5): 792-804.

[2] Basel Committee on Banking Supervision (1996). *Amendment to the Capital Accord to Incorporate Market Risks*. Basel: Bank for International Settlements.

[3] Basel Committee on Banking Supervision (2004). *International Convergence of Capital Measurement and Capital Standards*. Basel: Bank for International Settlements.

[4] Basel Committee on Banking Supervision (2006). *International Convergence of Capital Measurement and Capital Standards. A Revised Framework. Comprehensive Version.* Basel: Bank for International Settlements.

[5] Basel Committee on Banking Supervision (2009a). *Revisions to the Basel II Market Risk Framework*. Basel: Bank for International Settlements..

[6] Basel Committee on Banking Supervision (2009b). *Strengthening the Resilience of the Banking Sector, Consultative Document Issued for Comment by 16 April 2010*. Basel: Bank for International Settlements.

[7] Basel Committee on Banking Supervision (2011). *Revisions to the Basel II market risk framework*. Basel: Bank for International Settlements.

[8] Basel Committee on Banking Supervision (2014). Fundamental review of the trading book: A revised market risk framework, Consultative Document Issued for Comment by 31st Jan 2014, Basel: Bank for International Settlements.

[9] Basel Committee on Banking Supervision (2016). *Minimum capital requirements for market risk. Bank for International Settlements*, Basel: Bank for International Settlements.

[10] Berkelaar, A., Cumperayot, P., and Kouwenberg R. (2002). “The Effect of VaR Based Risk Management on Asset Prices and the Volatility Smile”, *European Financial Management*, 8(2): 139-164.

[11] Brazauskas, V., and Kleefeld A. (2009). “Robust and efficient fitting of the generalized Pareto distribution with actuarial applications in view”, *Insurance: Mathematics and Economics,* 45(3): 424-435.

[12] Breuer, T., and Csiszár I. (2013). “Systematic stress tests with entropic plausibility constraints”, *Journal of Banking & Finance*, 37(5): 1552-1559.

[13] Breuer, T., and Krenn G. (1999). “Stress Testing. Guidelines on Market Risk”, 5, Vienna: Oesterreichische Nationalbank, [Online] Available at: http://www.oenb.at/en/img/band5ev40_tcm16-20475.pdf.

[14] Breuer, T., Jandačka, M., Rheinberger, K., and Summer M. (2009). “How to Find Plausible, Severe, and Useful Stress Scenarios”, *International Journal of Central Banking*, 5(3): 205-224.

[15] Danélsson J. (2002). “The emperor has no clothes: Limits to risk modelling”, *Journal of Banking & Finance,* 26(7): 1273-1296.

[16] Degennaro, R. P. (2008). “Value-at-Risk: How Much Can I Lose by This Time Next Year?”, *The Journal of Wealth Management,* 11(3): 92-96.

[17] Embrechts, P., Klüppelberg, C., and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*. Berlin: Springer.

[18] Frenkel, M., Hommel, U., and Rudolf, M. (2005). *Risk Management: Challenge and Opportunity*. 2nd ed. Berlin: Springer.

[19] Gençay, R., and Selçuk F. (2004). "Extreme value theory and Value-at-Risk: Relative performance in emerging markets", *International Journal of Forecasting,* 20(2): 287-303.

[20] Gnedenko, B. V. (1943). "Sur la distribution limite du terme d'une serie aleatoire", *Annals of Mathematics*, 44(3): 423-453.

[21] Ghosh, S., and Resnick S. (2010). "A discussion on mean excess plots", *Stochastic Processes and their Applications,* 120(8): 1492-1517.

[22] Goldberg, L. R., Miller, G., and Weinstein J. (2008). "Beyond Value-at-Risk: Forecasting Portfolio Loss at Multiple Horizons", *Journal of Investment Management,* 6(2): 73-98.

[23] Heikkinen, V. P., and Kanto A. (2002). "Value-at-risk estimation using non-integer degrees of freedom of Student's distribution", *Journal of Risk,* 4(4): 77-84.

[24] Hill, B. M. (1975). "A Simple General Approach to Inference about the Tail of a Distribution", *Annals of Statistics,* 3(5): 1163-1173.

[25] Hogg, R. V., and Klugman S. A. (1984). Loss Distributions. New York: Wiley.

[26] Jackson, P., Maude, D., and Perraudin, W. (1997). "Bank capital and value-at-risk", *Journal of Derivatives,* 4(Spring): 177-184.

[27] Kerkhof, J., and Melenberg B. (2004). "Backtesting for risk-based regulatory capital", *Journal of Banking & Finance,* 28(8): 1845-1865.

[28] Linsmeier, T. J., and Pearson N. D. (2000). "Value-at-Risk", *Financial Analysts Journal,* 56(2): 47-67.

[29] McNeil, A. J., and Frey R. (2000). "Estimation of Tail-Related Risk Measures for Heteroscedastic Financial Time Series: An Extreme Value Approach", *Journal of Empirical Finance,* 7(3-4): 271-300.

[30] McNeil, A. J., and Saladin T. (1997). *The peaks over thresholds method for estimating high quantiles of loss distributions*. Proceedings of XXVII th International ASTIN Colloquium, Cairns, Australia, 23-43.

[31] Meerschaert, M. M., and Scheffler H-P. (1998). "A simple robust estimation method for the thickness of heavy tails", *Journal of Statistical Planning and Inference,* 71(1-2): 19-34.

[32] Pickands, J. (1975). "Statistical inference using extreme order statistics", *Annals of Statistics,* 3(1): 119-131.

[33] Reiss, R-D., and Thomas M. (1997). *Statistical analysis of extreme values*. Basel: Birkhäuser.

[34] Rossignolo, A. F., Fethi, M. D., and Shaban M. (2012). "Value-at-Risk models and Basel capital charges Evidence from Emerging and Frontier stock markets", *Journal of Financial Stability,* 8(4): 303-319.

[35] Sharma, M. (2012). "Evaluation of Basel III revision of quantitative standards for implementation of internal models for market risk", *IIMB Management Review,* 24(4): 234-244.

[36] Studer, G. (1997). "Maximum Loss for Measurement of Market Risk", Dissertation. Zürich: ETH Zürich, [Online] Available at: http://www.gloriamundi.org/picsresources/gsmlm.pdf.

[37] Studer, G. (1999). "Market risk computation for nonlinear portfolios", *Journal of Risk,* 1(4): 33-53.