Estimating Operational Risk Capital with Greater Accuracy, Precision, and Robustness

J.D. Opdyke,
Head of Operational Risk Modeling,
Quantitative Methods Group, GE Capital
J.D.Opdyke@ge.com

Disclaimer

This work was completed by the sole author, J.D. Opdyke, when he was Senior Managing Director of DataMineit, LLC (aside from more recent reviews of the published journal paper). The views presented herein are the views of the sole author and do not necessarily reflect the views of DataMineit, LLC, GE Capital, or any other institution.

Contents

1. **Presentation Goals and Operational Risk Setting**

2. **AMA–LDA OpRisk Capital Defined**

3. **When is Capital Apparently a Convex Function of Severity Parameters?**

   a. Convexity of Severity VaR $\Rightarrow$ Convexity of ALD VaR $\Rightarrow$ Inflated Capital

   b. Multiple Checks for Convexity of Severity VaR

4. **When is this Capital Bias (Inflation) Material?**

5. **RCE (Reduced-bias Capital Estimator) Virtually Eliminates Capital Bias**

6. **Simulation Study: RCE vs. MLE**
   Severities: LogNormal, GPD, LogGamma, & Truncated Versions of Each

   a. <u>RCE More Accurate</u>: MLE Capital Bias can be ENORMOUS ($Billion+ for one unit of measure!)

   b. <u>RCE More Precise</u>:   RCE RMSE < MLE RMSE, RCE StdDev < MLE StdDev, RCE IQR < MLE IQR

   c. <u>RCE More Robust</u>:   RCE Robustness to Violations of iid > MLE

7. **Summary and Conclusions**

8. **References and Appendices**

1. Presentation Goals

I. **Demonstrate that Jensen's Inequality is the apparent source of systematically inflated operational risk capital estimates** under the most common implementations of Basel II/III's AMA-LDA, and that this bias often is very large: hundreds of $millions, and sometimes $billions at the unit-of-measure level.

II. **Develop a Reduced-bias Capital Estimator (RCE) that** i) dramatically mitigates this capital overstatement, ii) notably increases the precision of the capital estimate, and iii) consistently increases its robustness to violations of the (unsupported) i.i.d. presumption. With capital accuracy, precision, and robustness greater than any other current LDA implementation, RCE arguably **would unambiguously improve the most widespread OpRisk Capital Estimation Framework, and would be the most consistent with regulatory intent vis-à-vis an unbiased and more stable implementation under Basel II/III's AMA**.

Operational Risk
└─ **Basel II/III**
    └─ **Advanced Measurement Approaches (AMA)**
        └─ **Risk Measurement & Capital Estimation**
            └─ **Loss Distribution Approach (LDA)**
                ├ **Frequency Distribution**
                └ **Severity Distribution***
                    └─ **Aggregate Loss Distribution**

* For purposes of this presentation, and as is widespread practice, potential dependence between the frequency and serverity distributions is ignored. See Chernobai, Rachev, and Fabozzi (2007) and Ergashev (2008).

2. AMA–LDA OpRisk Capital Defined

- A la Basel II/III, Operational Risk Capital for large banks/SIFIs must be estimated with an Advanced Measurement Approaches (**AMA**) framework.

- In writing, AMA provides great flexibility, but in practice, there has been industry convergence to the Loss Distribution Approach (**LDA**).

- Under LDA, **severity and frequency distributions** representing the magnitude and number of OpRisk loss events, respectively, are estimated based on samples of OpRisk loss event data.

- The severity and frequency distributions are **convoluted** (rarely in closed form) **to obtain the Aggregate Loss Distribution**.

- **Estimated Capital is a VaR of the Aggregate Loss Distribution**: specifically, the quantile associated with its **99.9%tile**, or the 1-in-1000 year loss, on average. Capital is estimated for every cell of data (or "Unit-of-Measure" (UoM), typically defined by Line of Business and Event Type) and then aggregated to the enterprise level via dependence modeling. The focus in this presentation is UoM-level capital.

- In practice, frequency parameters have very little effect on **estimated capital**, which **is driven almost entirely by the severity parameter values** (see Degen's (2010) analytical result below).
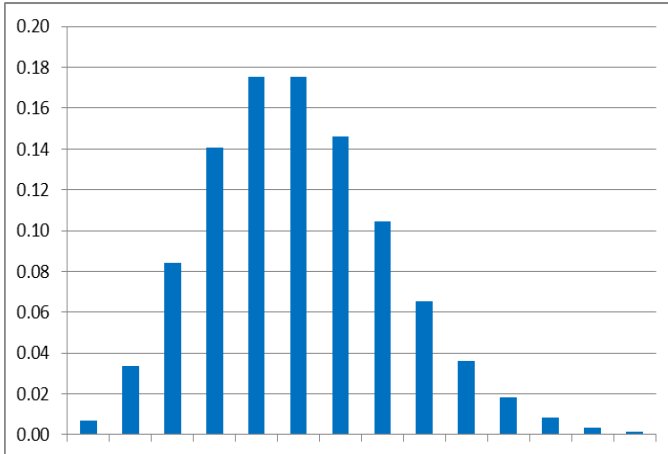
Loss Distribution Approach – For a given UoM:

Estimated Severity PDF – Truncated LogNormal (μ=10, σ=2.8, H=10k)

DCT = 10k
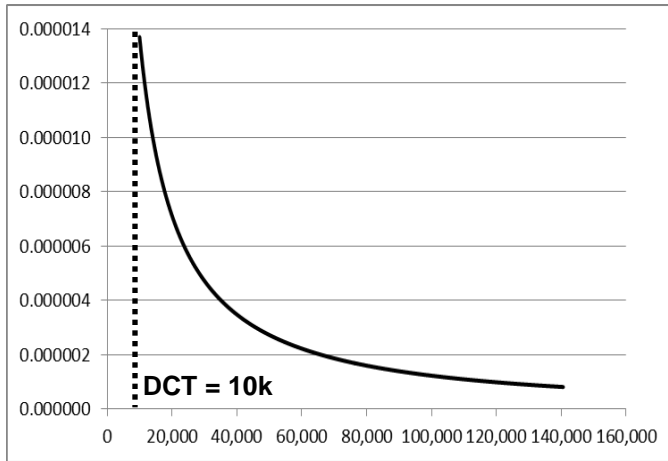
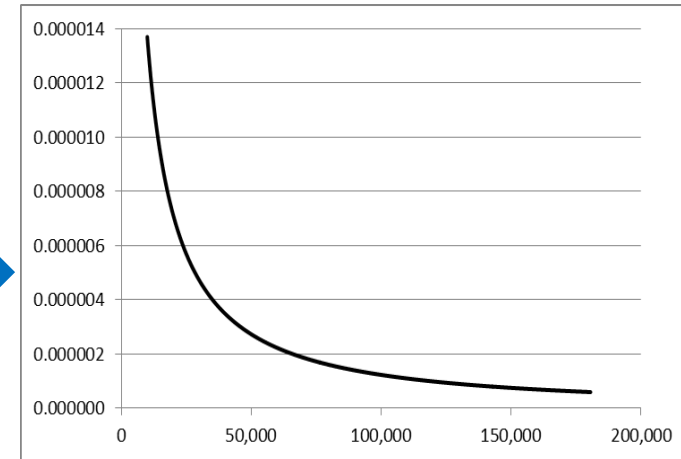**Estimated Frequency PMF – Poisson (annual λ=25)**

Convolution via simulation (in practice, rarely a closed form solution … but for the VaR there are good and widely accepted analytical approximations much faster than Monte Carlo simulation)

Aggregate Loss Distribution

Regulatory Capital = VaR at 99.9%tile

JSM 2014 August 2-7 BOSTON, MASSACHUSETTS

GE Capital

3.a. Apparent Convexity of Severity VaR $\Rightarrow$ Inflated Capital

- **Estimated Capital is Essentially a High Quantile of the Severity Distribution** as per Degen's (2010) Single Loss Approximation (SLA):

$$C_\alpha \approx F^{-1}\left(1 - \frac{1-\alpha}{\lambda}; \hat{\beta}\right) + \lambda\mu \qquad \text{where } \lambda = \text{frequency parameter and } \mu = \text{E}\left[X\right]$$

In other words, first term >> second term (see Appendix A for an improved Interpolated SLA (ISLA) from Opdyke, 2014).

- <u>PROPOSED:</u> **For this setting (heavy-tailed severities, certain parameter value ranges, and very high p = percentiles):**
IF Aggregate Loss Distribution (ALD) VaR (i.e. Capital) is a very slightly concave function of λ, the frequency parameter(s) (as shown empirically in Opdyke, 2014),
AND Severity VaR is a highly convex function of Severity Parameter Vector $\hat{\beta}$,
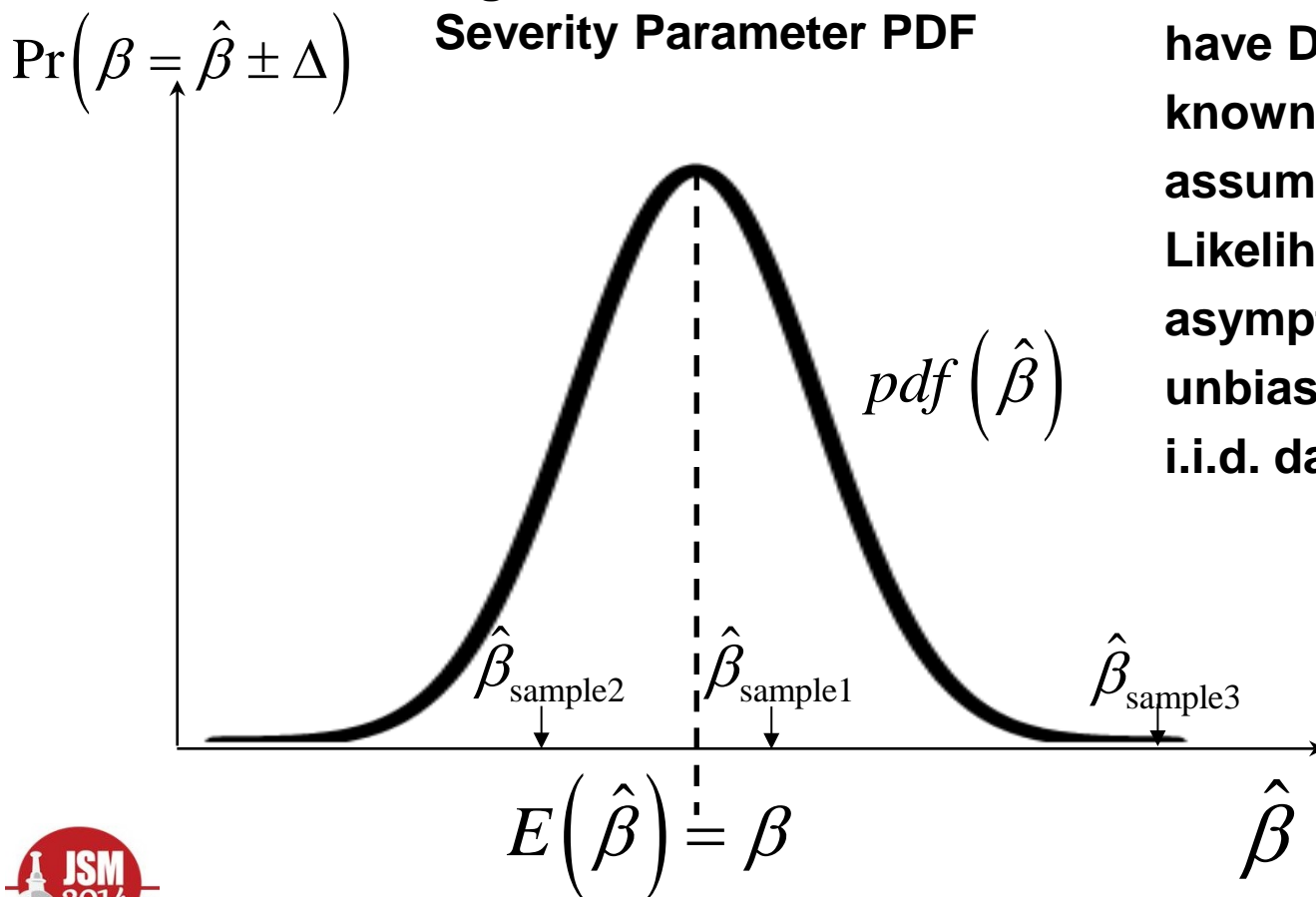**THEN ALD VaR (Capital) is a highly convex function of $\hat{\beta}$.**

- <u>NOTE:</u> **Severity VaR is much more extreme than ALD VaR, because for, say, $\lambda = 30$, and $\alpha = 0.999$ and $\alpha = 0.9997$, $p = \left[1 - \left(1-\alpha\right)/\lambda\right] = $ 0.999967 and 0.99999, respectively.**

- **Operational Risk Loss Event Data = a Sample, NOT a Population**

- **Therefore, true severity parameters, $\boldsymbol{\beta}$, will never be known.**

Figure 1:
Severity Parameter PDF

$$\Pr\left(\beta = \hat{\beta} \pm \Delta\right)$$

$$pdf\left(\hat{\beta}\right)$$

$$\hat{\beta}_{sample2} \quad \hat{\beta}_{sample1} \quad \hat{\beta}_{sample3}$$
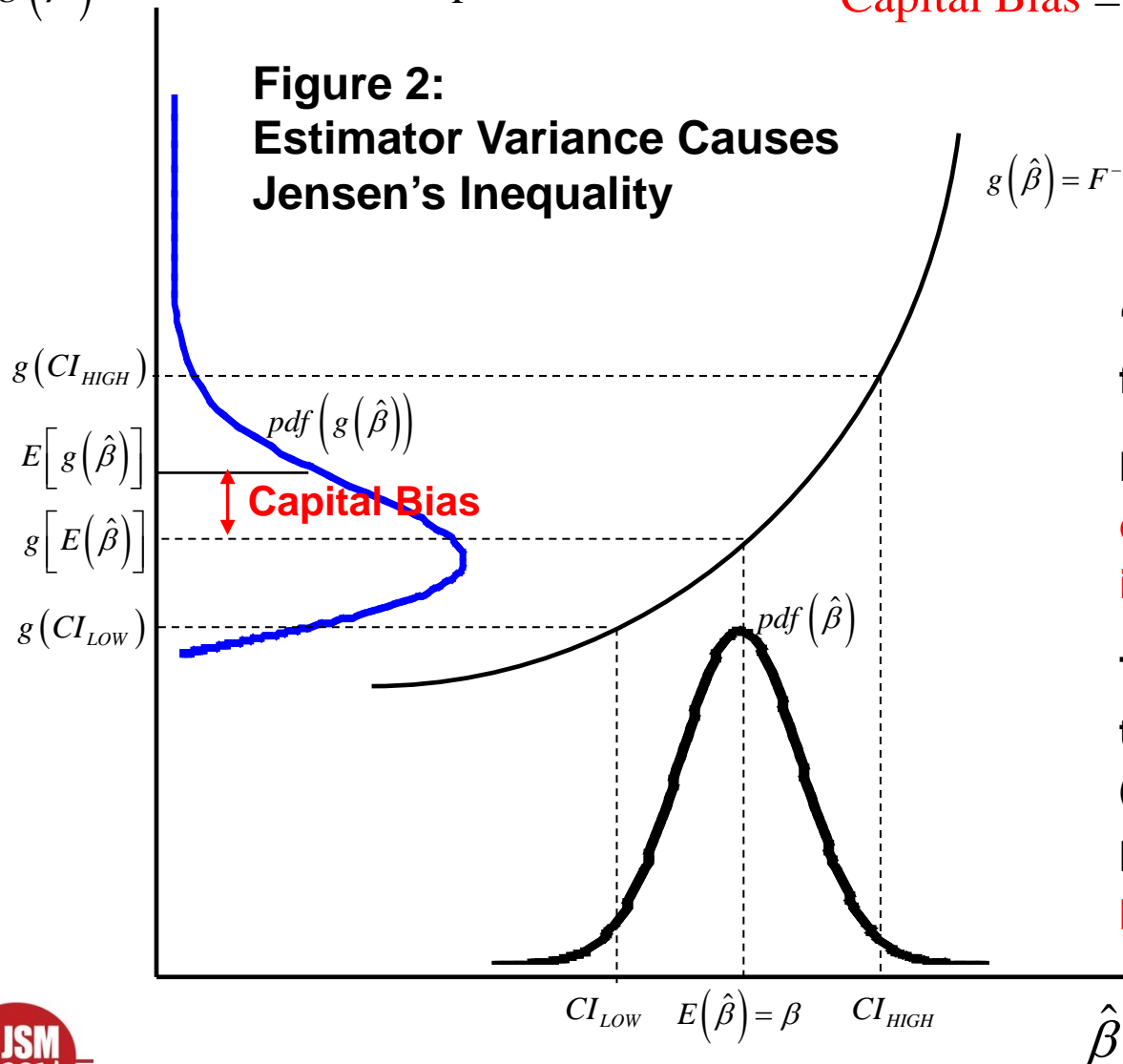
$$E\left(\hat{\beta}\right) = \beta \qquad \hat{\beta}$$

BUT Estimated Parameters, $\hat{\beta}$, have DISTRIBUTIONS that are known under specified assumptions (e.g. Maximum Likelihood Estimators (MLE) are asymptotically normal, unbiased, and efficient under i.i.d. data)

3.a. Apparent Convexity of Severity VaR $\Rightarrow$ Inflated Capital

$g(\hat{\beta}) = \hat{C}$ = Estimated Capital

Capital Bias $= \left(E\left[g(\hat{\beta}) \right] - g\left[E(\hat{\beta}) \right] \right) > 0$

Figure 2:
Estimator Variance Causes
Jensen's Inequality

$g(\hat{\beta}) = F^{-1}(p; \hat{\beta})$ = presumed shape of severity quantile/VaR

$g(CI_{HIGH})$

$pdf\left(g(\hat{\beta}) \right)$

$E\left[g(\hat{\beta}) \right]$

Capital Bias

$g\left[E(\hat{\beta}) \right]$

$g(CI_{LOW})$

$pdf(\hat{\beta})$

CI_{LOW} $E(\hat{\beta}) = \beta$ CI_{HIGH}

$\hat{\beta}$

"Jensen's inequality" first proved in 1906.

If VaR, g(), is convex in $\hat{\beta}$, **capital always will be inflated.**

This appears to be true for i) the relevant severities (heavy-tailed) with ii) relevant parameter ranges, WHEN iii) p is large (i.e. p>0.999).

* Graph based on Kennedy (1992), p.37.

3.b. Multiple Checks for Convexity of Severity VaR

- **Severity VaR is NOT a convex function of the severity parameter vector $\hat{\beta}$ globally, for all percentiles (p) and all severities. This is widely known and easily proved.**

- **However, Severity VaR appears always to be a convex function of $\hat{\beta}$ under, concurrently, BOTH i) sufficiently high percentiles ($p>0.999$) AND ii) sufficiently heavy-tailed severities (amongst those used in OpRisk modeling). Both conditions hold in AMA–LDA OpRisk Capital Estimation (see Appendix B), and the very strong empirical evidence is exactly consistent with the effects of convexity in that we observe Jensen's Inequality.**

- **Still, we would like to PROVE a) convexity in Severity VaR under these conditions, and b) convexity in VaR for all relevant severities.**

- **Still, we would like to PROVE a) convexity in Severity VaR under these conditions, and b) convexity in VaR for all relevant severities.**

- **Re: a), we can examine three things:**
  The shape of VaR as a function of the severity parameters…

  i. **individually (i.e. check for marginal convexity)**

  ii. **jointly (i.e. mathematically determine the shape of the multidimensional VaR surface)**

  iii. **jointly, based on extensive Monte Carlo simulation (i.e. examine the behavior of VaR as a function of joint parameter perturbation)**

3.b. Multiple Checks for Convexity of Severity VaR

- **Re: a), we can examine three things:
  The shape of VaR as a function of the severity parameters…**

  i. **individually (i.e. check for marginal convexity)**

     Analytically this is straightforward for those severities with closed-form VaR functions. For the LogNormal, for example,

$$VaR = ICDF = \exp\left(\mu + \sigma \Phi^{-1}(p)\right), \text{ so}$$

$$\partial^2 VaR \big/ \partial \mu^2 = VaR > 0$$

$$\partial^2 VaR \big/ \partial \sigma^2 = VaR \cdot \left[\Phi^{-1}(p)\right]^2 > 0$$
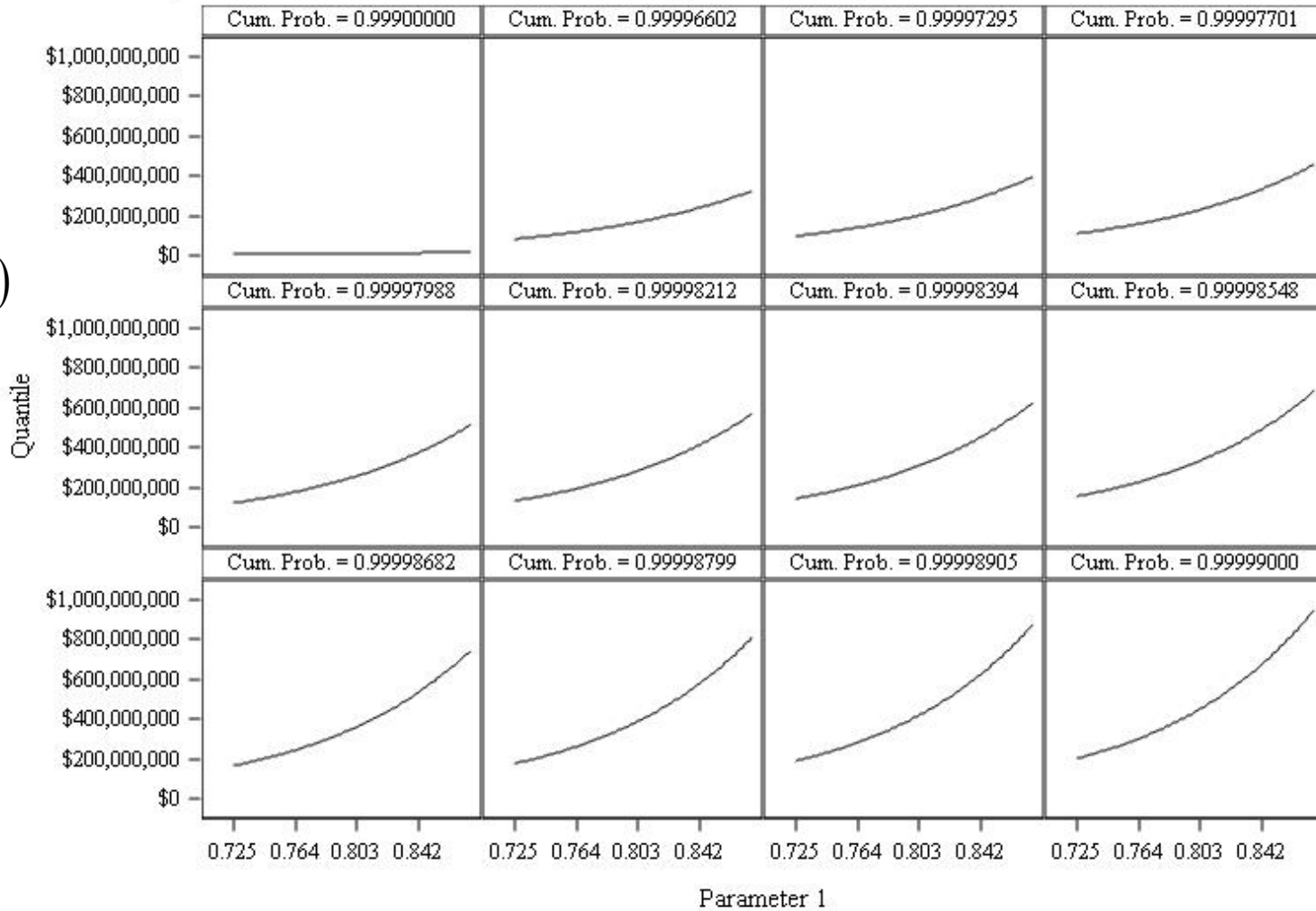
     However, this is not typically the case, especially for truncated distributions. But these marginal checks are easy to do graphically (NOTE that GPD also is straightforward analytically).

Figure 3a:

Severity = GPD, Threshold = 0k, Parm1 by Quantile by CumProb, Parm2 = 35000

$F^{-1}(p;\xi)$

Figure 3b:

Severity = GPD, Threshold = 0k, Parm2 by Quantile by CumProb, Parm1 = 0.8

$$F^{-1}(p;\theta)$$
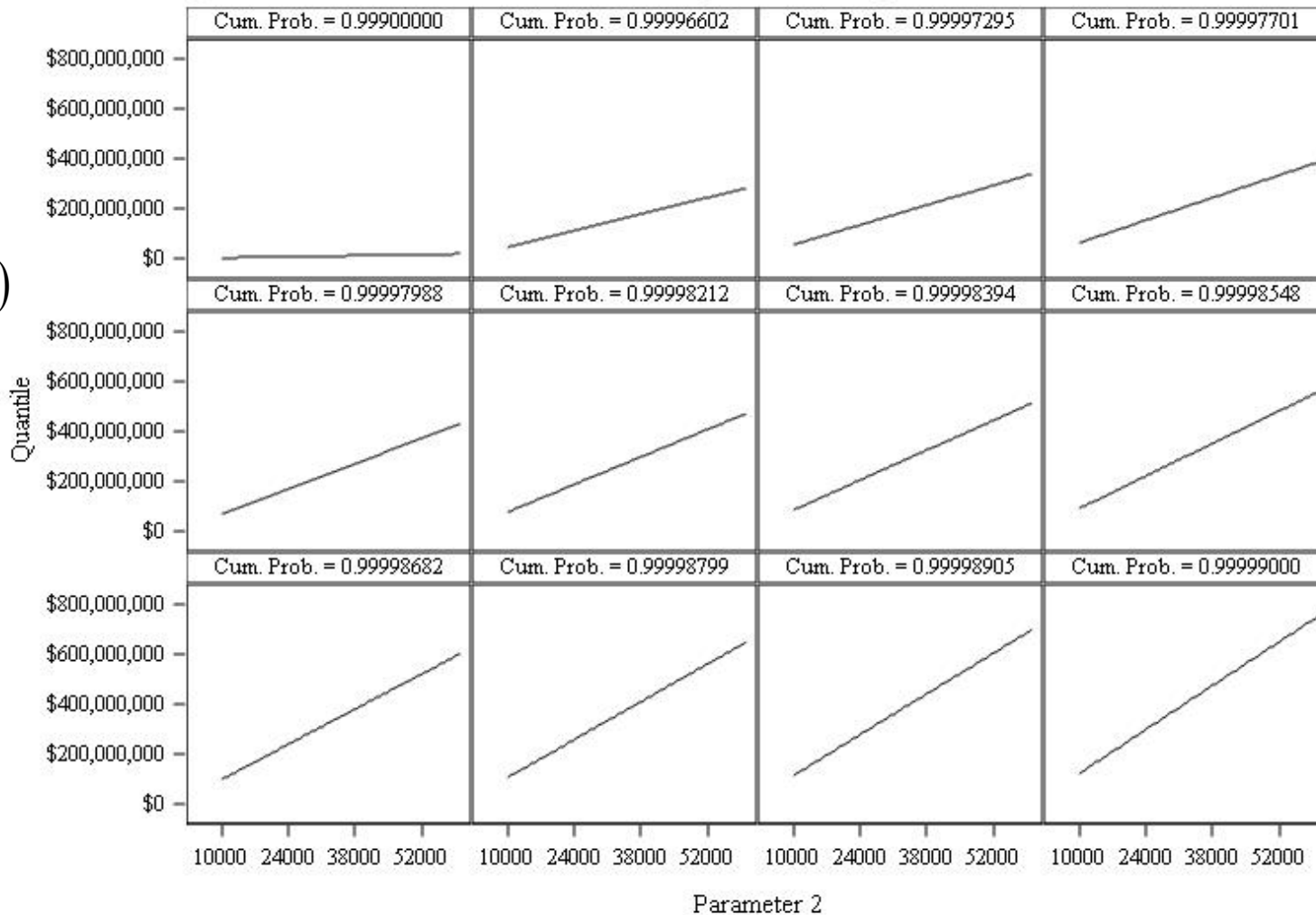
3.b. Multiple Checks for Convexity of Severity VaR

For GPD, for large p(>0.999): VaR is convex in ξ and linear in θ, so VaR APPEARS to be convex in parameter vector $\hat{\beta}$, implying systematic and consistent capital inflation. Note this convexity in ξ increases in p. Additional widely used severities are shown below.

TABLE 1: Marginal VaR Convexity/Linearity OVER RELEVANT DOMAIN ($p > 0.999$) by Parameter by Severity

Severity Distribution	VaR is Convex/Linear as Function of...			Relationship between Parameters
	Parameter 1	Parameter 2	Parameter 3	
1) **LogNormal (μ, σ)**	Convex	Convex		Independent
2) **LogLogistic (α, β)**	Linear	Convex		Independent
3) **LogGamma (a, b)**	Convex	Convex		Dependent
4) **GPD (ξ, θ)**	Convex	Linear		Dependent
5) **Burr (type XII) (ϒ, α, β)**	Convex	Convex	Linear	Dependent
6) **Truncated 1)**	Convex	Convex		Dependent
7) **Truncated 2)**	Linear	Convex		Dependent
8) **Truncated 3)**	Convex	Convex		Dependent
9) **Truncated 4)**	Convex	Linear		Dependent
10) **Truncated 5)**	Convex	Convex	Linear	Dependent

3.b. Multiple Checks for Convexity of Severity VaR

- **Re: a), we can examine three things:**
  The shape of VaR as a function of the severity parameters…

  i. **individually (i.e. check for marginal convexity)**

     For all commonly used severities in this space, VaR always appears to be a convex function of at least one parameter, and a linear function of the rest. This would be consistent with convex, or "convex-dominant" (see below) behavior when VaR is examined as a function of the severity parameters jointly.

3.b. Multiple Checks for Convexity of Severity VaR

- **Re: a), we can examine three things:**
  The shape of VaR as a function of the severity parameters…

  ii. **jointly** **(i.e. mathematically determine the shape of the multidimensional VaR surface)**

     This can be done via examination of the signs and magnitudes of the eigenvalues of the shape operator (which define its principal curvatures).

     This turns out to be analytically nontrivial, if not intractable under truncation, and even numeric calculations for many of the relevant severities are nontrivial given the sizes of the severity percentiles that must be used in this setting (because most of the gradients are exceedingly large for such high percentiles).

ii.   **jointly** **(i.e. mathematically determine the shape of the multidimensional VaR surface)**

So this research currently remains underway, and without this strict mathematical verification, attributions of capital inflation to Jensen's inequality are deemed "apparent" and/or "preliminary," as are those related to VaR's (apparent) convexity.

This scientifically conservative approach, however, belies the strong and consistent empirical evidence of capital inflation, and its behavior as being exactly consistent with the effects of Jensen's inequality (in addition to findings of marginal convexity). In other words, just because the specific multidimensional shapes of high-percentile VaR under these severities are nontrivial to define mathematically, we should not turn a blind eye toward strong empirical evidence that convexity dominates VaR's shapes as a joint function of severity parameters.

3.b. Multiple Checks for Convexity of Severity VaR

iii. **jointly, based on extensive Monte Carlo simulation** (i.e. examine the behavior of VaR as a function of joint parameter perturbation)

This is unarguably the most directly relevant of the three "checks" for convexity -- **EXAMPLE**:

a. simulate 10 years of i.i.d. losses generated under a Poisson frequency distribution, with $\lambda = 25$, and a LogNormal severity distribution with $\mu = 9.27$, $\sigma = 2.77$, estimating λ, μ, and σ using, say, maximum likelihood.

b. Use Degen (2010) to calculate RCap with $\alpha = 0.999$ and ECap with $\alpha = 0.9997$ based on the estimated λ, μ, and σ.

c. Repeat a. and b. 1,000 or more times.

d. The mean of the 1,000+ RCap/ECap estimates $E\left[g\left(\hat{\beta}\right)\right]$ will be about \$83m/\$203m **larger than "true" capital** $g\left[E\left(\hat{\beta}\right)\right]$ (\$603m, \$1,293m; see complete results in Table 4a below).

ANOTHER EXAMPLE:

a. simulate 10 years of i.i.d. losses generated under a Poisson frequency distribution, with $\lambda = 25$, and a GPD severity distribution with $\xi = 0.875$, $\theta = 47{,}500$, estimating λ, ξ, and θ using, say, maximum likelihood.

b. Use Degen (2010) to calculate RCap with $\alpha = 0.999$ and ECap with $\alpha = 0.9997$ based on the estimated λ, ξ, and θ.

c. Repeat a. and b. 1,000 or more times.

d. The mean of the 1,000+ RCap/ECap estimates $E\left[g\left(\hat{\beta}\right)\right]$ will be about \$249m/\$1,016m **larger than "true" capital** $g\left[E\left(\hat{\beta}\right)\right]$ (\$391m/\$1,106m; see complete results in Table 4e below).

iii. **jointly, based on extensive Monte Carlo simulation** (i.e. examine the behavior of VaR as a function of joint parameter perturbation)

As long as the percentiles examined are large enough (e.g. *p* > 0.999) and the severity parameter values large enough, the **estimates of severity VaR and Rcap/ECap consistently, across all severities used in AMA-based operational risk capital estimation, are notably inflated**. This inflation can be dramatic, not uncommonly into the hundreds of millions, and even billions of dollars, for each UoM (unit-of-measure) as shown below.

So let us **presume VaR convexity** and design a capital estimator accordingly **to mitigate the actual capital bias/inflation** of which it is the presumed source…

3.b. Multiple Checks for Convexity of Severity VaR

- **Still, we would like to PROVE a) convexity in Severity VaR under these conditions, and b) convexity in VaR for all relevant severities.**

- **So we are presuming a) based on very strong empirical evidence and incomplete mathematical evidence.**

- **For b), tackling ALL potentially relevant severities is nontrivial (if possible), but arguably unnecessary as the number of severities used in this setting are quite finite, and we can satisfy a) for each individually.**

  Note again that because capital (VaR of ALD) was shown empirically in Opdyke (2014) to be only a slightly concave function of the frequency parameter(s), the only source of capital inflation would appear to be strong convexity in severity VaR.

Convexity in Severity VaR $\Rightarrow$ Capital Bias is upwards … always!
Magnitude of Capital Inflation is Determined by:

a) **Variance of Severity Parameter Estimator:**
   Larger Variance (smaller $n<1,000$) $\Rightarrow$ Larger Capital Bias

b) **Heaviness of Severity Distribution Tail:**
   Heavier $\Rightarrow$ More Capital Bias
   (so truncated distributions $\Rightarrow$ more bias, ceteris paribus)

c) **Size of VaR Being Estimated:**
   Higher VaR $\Rightarrow$ More Capital Bias
   (so Economic Capital Bias > Regulatory Capital Bias)

This demonstrable empirical behavior is exactly consistent with Jensen's Inequality, and since most UoMs are heavy-tailed severities and typically $n < 250$, <u>AMA–LDA OpRisk capital estimation is squarely in the bias zone!</u>

4. When is Capital Bias (Inflation) Material?

NOTE: **LDA Capital Bias holds for most, if not all widely used severity parameter estimators** (e.g. Maximum Likelihood Estimation (MLE), Robust Estimators (OBRE, CvM, QD, etc.), Penalized Likelihood Estimation (PLE), Method of Moments, all M-Class Estimators, Generalized Method of Moments, Probability Weighted Moments, etc.).

NOTE: Because CVaR is a (provably) convex function of severity parameter estimates (see Brown, 2007, Bardou et al., 2010, & Ben-Tal, 2005), **switching from VaR to CVaR**, even if allowed, **does not avoid this problem** (and in fact, appears to make it worse).

NOTE: **Severities with E(x)=∞ also can exhibit such bias** (see GPD with ξ = 1.1, θ = 40,000 in Opdyke, 2014), even though counterexamples exist.

5. RCE – Reduced-bias Capital Estimator

I.   **Demonstrate that Jensen's Inequality is the apparent source of systematically inflated operational risk capital estimates …**

II.   **Develop a Solution…**

SOLUTION CHALLENGES / CONSTRAINTS:

1.   **It must remain consistent with the LDA Framework (even with new guidance (6/30/14) encouraging new methods, arguably the smaller the divergence from widespread industry practice, the greater the chances of regulatory approval).**

2.   **The same general method must work across very different severities.**

3.   **It must work when severity distributions are truncated to account for data collection thresholds.**

4.   **It must work even if $E(x)=\infty$ (or close, which is relevant for any simulation-based method).**

5.   **It cannot be excessively complex (or it won't be used).**

6.   **It cannot be extremely computationally intensive (e.g. a desktop computer, or it won't be used).**

7.   **Its range of application must encompass all commonly used estimators of severity (and frequency)**

8.   **It must work regardless of the method used to approximate VaR of the aggregate loss distribution.**

9.   **It must be easily understood and implemented using any widely available statistical software.**

10.   **It must provide unambiguous improvements over the most widely used implementations of LDA (e.g. MLE, and most other estimators) on all three key criteria – capital accuracy, capital precision, and capital robustness.**

5. RCE – Reduced-bias Capital Estimator

RCE (Reduced-bias Capital Estimator) is the only estimator designed to effectively mitigate LDA Capital Bias.

RCE simply is a scaler of capital as a function of the degree of empirical VaR convexity.

<u>RCE Conceptually Defined</u>:

<u>Step 1</u>: Estimate LDA-based capital using any estimator (e.g. MLE).

<u>Step 2</u>: Using 1), simulate _K_ iid data samples and estimate parameters of each

<u>Step 3</u>: Using 2), simulate _M_ data samples for each of the _K_ parameters, estimate capital for each, and calculate median for each, yielding _K_ medians of capital

<u>Step 4</u>:
RCE = median(_K_ medians) * [median(K medians) / weighted mean(_K_ medians)]^_c_

RCE Motivation:

RCE = median(*K* medians) * [median(K medians) / weighted mean(*K* medians)]^*c*

<u>First term</u>: The median of *K* medians is empirically close to "capital." The *K* medians simply trace out the VaR function (in 1-dimension, $g\left(\hat{\beta}\right)$ in Figure 2) just as do *K* capital estimates, but capital is more volatile than using another layer of sampling to obtain the *K* medians in <u>Step 3</u>.

<u>Second term</u>: The ratio of the median to the mean is an empirical measure of the convexity of VaR, $g\left(\hat{\beta}\right)$. This is used to scale down the first term (which is essentially capital) to eliminate inflation exactly consistent with the effects of Jensen's Inequality. The mean is weighted* based on the sampling (perturbation) method described below. The *c* exponent is a function of the severity chosen and the sample size, both of which are known ex ante under LDA.

* Due to the sampling method described below, the median in the numerator turns out to be empirically identical to a weighted median, and so for efficiency, the simple median is used.

5. RCE − Reduced-bias Capital Estimator

RCE Implemented:

Step 1: Estimate LDA-based capital using any estimator (e.g. MLE).

Step 2: Using 1), generate K parameter vectors based on the Var-Cov matrix using iso-density sampling (see Figure 4 below): use iso-density ellipses to select parameter values associated with a given probability, and change parameter values to reach these ellipses via the decrease-decrease, decrease-increase, increase-decrease, and increase-increase of both parameters by the same number of standard deviations (thus generating two orthogonal lines emanating from original parameter estimate in the normalized coordinate system). Opdyke (2014) uses ellipse percentiles = 1, 10, 25, 50, 75, 90, and 99, so K = 4*7=28, and two frequency percentiles for λ, 25 and 75, so total K = 28*2 = 56. Weights = $(1-p_{sev})*2*(1-p_{frq})$.

Step 3: Using the K parameter vectors from 2) (including the frequency parameters), generate another triplet of M parameter vectors for each (let $M=K$), and calculate capital for each, and take the median to get K medians of capital.

Step 4:
RCE = median(K medians) * [median(K medians) / weighted mean(K medians)]^c

FIGURE 4: Iso-density Perturbation of the Joint Severity Parameter Distribution
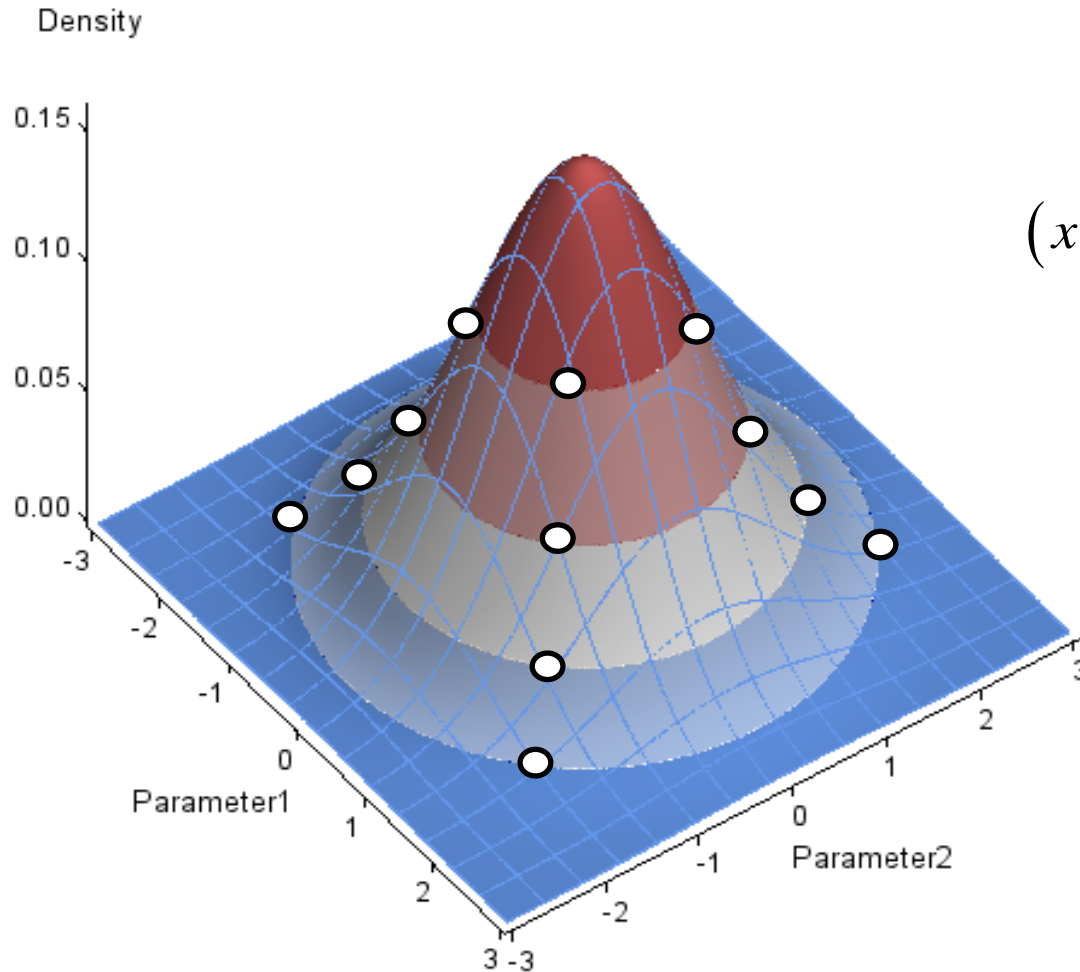
For multivariate normal (e.g. all M-class estimators), ellipses are given by:

$$\left(x - \mu \right)^T \Sigma^{-1} \left(x - \mu \right) \leq \chi_k^2 \left(p \right)$$

where **x** is a **k- (2-) dimensional vector**, **μ** is the known **k - dimensional mean vector (the parameter estimates)**, Σ is the **known covariance matrix (the inverse of the Fisher information of the given severity)**, and $\chi_k^2 \left(p \right)$ is the **quantile function for probability p of the Chi-square distribution with k degrees of freedom.**

5. RCE – Reduced-bias Capital Estimator

Finding *x* as the solution to $(x-\mu)^T \Sigma^{-1}(x-\mu) \le \chi_k^2(p)$ **can be obtained quickly via a convergence algorithm (e.g. bisection) or simply the analytic solution to the equation rather than the inequality (see Mayorov 2014). Simply change both parameters by *q* units of their respective standard deviations to obtain four pairs of parameter values on the ellipse defined by *p*: increase both parameters by *q* standard deviations** $(z_1 = z_2 = 1)$**, decrease both parameters by *q* standard deviations** $(z_1 = z_2 = -1)$**, increase one while decreasing the other** $(z_1 = 1, z_2 = -1)$ **, and decrease one while increasing the other** $(z_1 = -1, z_2 = 1)$**.**

$$q\#SD = \sqrt{\frac{\chi_k^2(p) \cdot \left(1 + z_1 z_2 \rho_{1,2}\right)}{2}}$$

where $\sigma_1(\sigma_2) = $ stdev of parameter 1 (2), and $\rho_{1,2}$ is Pearson's correlation of the parameter estimates.

Alternately, the eigenvalues and eigenvectors of Σ^{-1} **can be used to define the most extreme parameter values (smallest and largest) on the ellipses (corresponding to the largest/smallest eigenvalues) (see Johnson and Wichern, 2007), but this would likely change the values of *c* calculated below, and the above is arguably more straightforward.**

5. RCE − Reduced-bias Capital Estimator

Iso-density sampling (perturbation) makes RCE runtime feasible (1 to 3 seconds on a standard desktop PC):

Table 2: Runtime of RCE by Severity (seconds)

Severity*	Real Time	CPU Time
LogN	0.14	0.14
TLogN	1.10	1.10
Logg	1.13	1.12
TLogg	2.96	2.94
GPD	0.21	0.18
TGPD	1.35	1.35

The complexity of the Fisher information is the only thing that drives runtime (sample size is irrelevant).

* See Appendix C.

JSM 2014
August 2-7
BOSTON, MASSACHUSETTS

GE Capital

5. RCE – Reduced-bias Capital Estimator

Implementation NOTE:

It is important to avoid bias when using iso-density sampling in cases of incalculably high capital. For example, say the initial MLE parameters happen to be large, and then the 99%tile of the joint parameter distribution, based on the initial estimates, is obtained in <u>Step 2</u> of RCE's implementation; and then the 99%tile of THIS Fisher information is obtained in <u>Step 3</u>, based on the joint parameter distribution of the <u>Step 2</u> values. Capital calculated in <u>Step 3</u> sometimes simply will be too large to calculate in such cases. If ignored, this could systematically bias RCE. A simple solution is to eliminate the entire ellipse of values – along with all "larger" ellipses – when any one value on an ellipse is too large to calculate.

5. RCE – Reduced-bias Capital Estimator

How is *c(n, severity)* determined?:

Method 1: Conduct a simulation study to empirically determine the value of *c* for the relevant sample sizes and severities (both known ex ante within the LDA framework) using three sets of parameter values: the original estimates, and those corresponding to the 2.5%tile and the 97.5%tile of the joint parameter distribution, which yields a 95% confidence interval (a wider confidence interval can be used if desired). The value of *c(n, severity)* is chosen to yield true capital (or slightly above) for all three sets of parameter values.

Method 2: Use the simulation study conducted in Opdyke (2014) to select values of *c* for specific values of *n* and *severity* (see Table 3 and Figure 5 below).
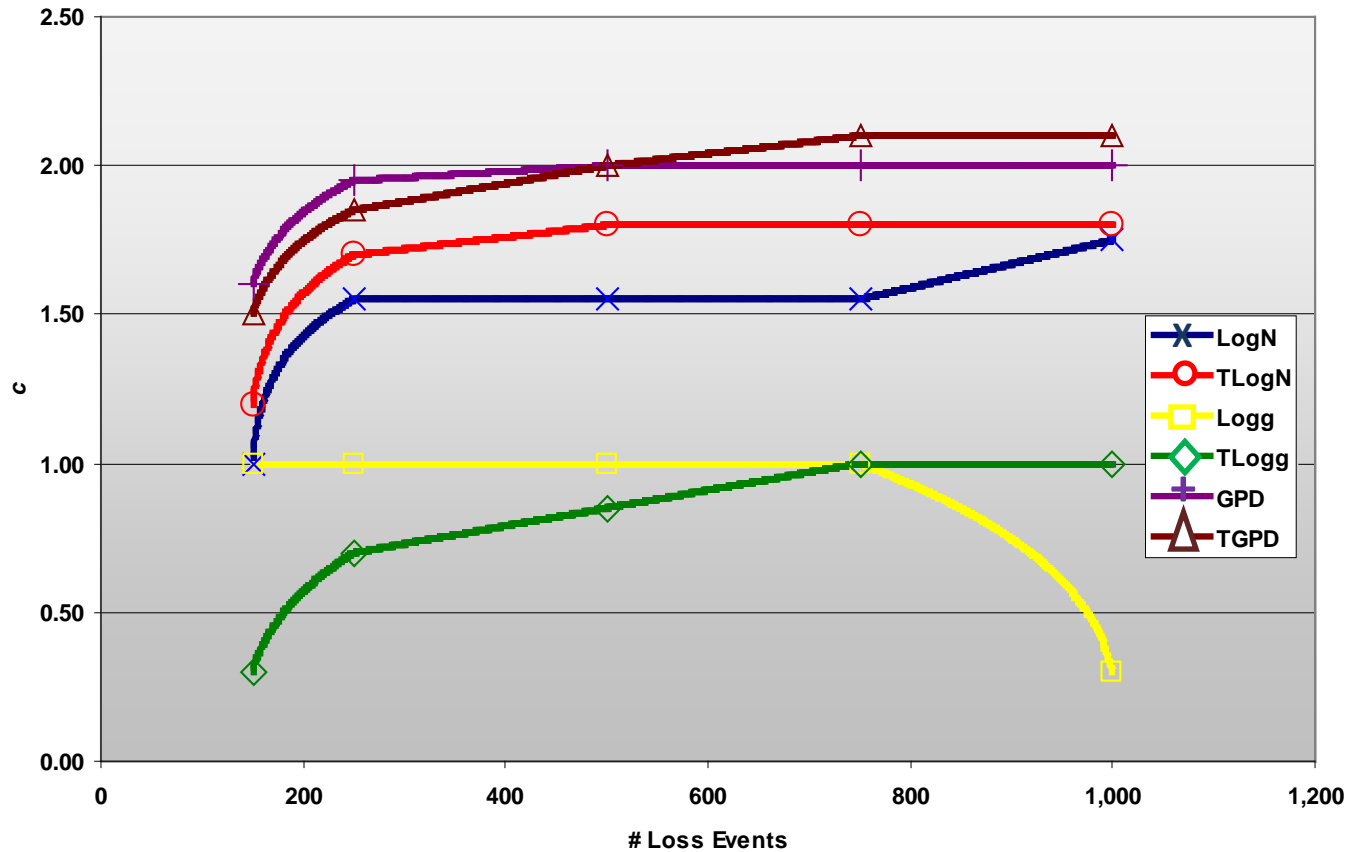
5. RCE – Reduced-bias Capital Estimator

Table 3:
Values of *c(n, severity)* by Severity by # of Loss Events
(Linear, and Non-Linear Interpolation with Roots Specified for Shaded Ranges)

N →	150	250	500	750	1000	Root
Severity						
LogN	1.00	1.55	1.55	1.55	1.75	8
TLogN	1.20	1.70	1.80	1.80	1.80	8
Logg	1.00	1.00	1.00	1.00	0.30	3
TLogg	0.30	0.70	0.85	1.00	1.00	3
GPD	1.60	1.95	2.00	2.00	2.00	10
TGPD	1.50	1.85	2.00	2.10	2.10	10

Figure 5:
Values of *c*(*n*, *severity*) by Severity by # of Loss Events

5. RCE – Reduced-bias Capital Estimator

NOTE: Unterlined **Unfortunately, other Bias-reduction/elimination strategies in the literature, even for VaR (e.g. see Kim and Hardy, 2007), do not appear to work for this problem.** * Most involve shifting the distribution of the estimator, often using some type of bootstrap distribution, which in this setting often results in negative capital estimates and greater capital instability. RCE-based capital is never negative, and is more stable than capital based on most, if not all other commonly used severity parameter estimators (e.g. MLE).

Also, given the very high percentiles being examined in this setting (e.g., Severity VaR = 0.99999 and higher), approaches that rely on the derivative(s) of VaR(s), perhaps via (taylor) series expansions, appear to run into numeric precision issues for some severities. So even when such solutions exist in tractable form, practical challenges may derail their application here.

* The only other work in the literature that appears to be similar in approach to RCE is the fragility heuristic (H) of Taleb et al. (2012) and Taleb and Douady (2013). Both RCE and H are measures of convexity based on perturbations of parameters: H measures the distance between the average of model results over a range of shocks and the model result of the average shock, while RCE is a scaling factor based on the ratio of the median to the mean of similar parameter perturbations. Both exploit Jensen's inequality to measure convexity: in the case of the fragility heuristic, to raise an alarm about it, and in the case of RCE, to eliminate it (or rather, to effectively mitigate its biasing effects on capital estimation).

SIMULATION STUDY: 1,000 (i.i.d.) Simulations of

- $\lambda = 25$ (Poisson-distributed average annual losses … so $n = 250$, on average, over 10 years)

- $\alpha = 0.999$ and 0.9997 for Regulatory and Economic Capital, respectively (so $[\, 1 - (1-\alpha)\,/\,\lambda\,] = 0.99996$ and 0.999988, respectively).

Selected Results of RCE capital vs. MLE capital:
  - LogNormal
  - LogGamma
  - GPD
  - Truncated LogNormal
  - Truncated LogGamma
  - Truncated GPD

6. Simulation Study: RCE vs. MLE

Table 4a:
RCE vs. LDA-MLE for LogNormal Severity (μ = 9.27, σ = 2.77, H=\$0k)*

(millions)	Regulatory Capital**			Economic Capital**	
	RCE	LDA-MLE		RCE	LDA-MLE
Mean*	$614	$686		$1,333	$1,498
True Capital	$603	$603		$1,293	$1,293
Bias (Mean - True)	$12	$83		$40	$205
Bias %	2.0%	13.8%		3.1%	15.8%
RMSE*	$328	$382		$764	$898
STDDev*	$328	$373		$763	$874

* 1,000 Simulations, n ≈ 250

** λ = 25; α = 0.999 RC; α = 0.9997 EC

6. Simulation Study: RCE vs. MLE

Table 4b:
RCE vs. LDA-MLE for Truncated LogNormal Severity (μ = 10.7, σ = 2.385, H=$10k)*

(millions)	Regulatory Capital**			Economic Capital**	
	RCE	**LDA-MLE**		**RCE**	**LDA-MLE**
Mean*	$700	$847		$1,338	$1,678
True Capital	$670	$670		$1,267	$1,267
Bias (Mean - True)	$30	$177		$71	$411
Bias %	4.5%	26.4%		5.6%	32.4%
RMSE*	$469	$665		$1,003	$1,521
STDDev*	$468	$641		$1,000	$1,464

* 1,000 Simulations, n ≈ 250

** λ = 25; α = 0.999 RC; α = 0.9997 EC

JSM 2014 BOSTON, MASSACHUSETTS August 2-7

GE Capital

6. Simulation Study: RCE vs. MLE

Table 4c:
RCE vs. LDA-MLE for LogGamma Severity (*a* = 25, *b* = 2.5, H=$0k)*

(millions)	Regulatory Capital**			Economic Capital**	
	RCE	**LDA-MLE**		**RCE**	**LDA-MLE**
Mean*	$466	$513		$1,105	$1,272
True Capital	$444	$444		$1,064	$1,064
Bias (Mean - True)	$11	$70		$42	$208
Bias %	2.5%	15.7%		3.9%	19.5%
RMSE*	$301	$355		$814	$984
STDDev*	$301	$348		$813	$962

* 1,000 Simulations, n ≈ 250

** λ = 25; α = 0.999 RC; α = 0.9997 EC

6. Simulation Study: RCE vs. MLE

Table 4d:

RCE vs. LDA-MLE for Truncated LogGamma Severity (a = 34.5, b = 3.15, H=$10k)*

(millions)	Regulatory Capital**			Economic Capital**	
	RCE	LDA-MLE		RCE	LDA-MLE
Mean*	$539	$635		$1,158	$1,437
True Capital	$510	$510		$1,086	$1,086
Bias (Mean - True)	$29	$125		$72	$350
Bias %	5.8%	24.5%		6.6%	32.2%
RMSE*	$397	$544		$941	$1,453
STDDev*	$396	$529		$938	$1,410

* 1,000 Simulations, n ≈ 250

** λ = 25; α = 0.999 RC; α = 0.9997 EC

JSM 2014
August 2-7
BOSTON, MASSACHUSETTS

GE Capital

6. Simulation Study: RCE vs. MLE

Table 4e:
RCE vs. LDA-MLE for GPD Severity ($\xi = 0.875$, $\theta = 47,500$, H=$0k)*

(millions)	Regulatory Capital**			Economic Capital**	
	RCE	LDA-MLE		RCE	LDA-MLE
Mean*	$396	$640		$1,016	$2,123
True Capital	$391	$391		$1,106	$1,106
Bias (Mean - True)	$5	$249		$24	$1,016
Bias %	1.2%	63.7%		2.2%	91.9%
RMSE*	$466	$870		$1,594	$3,514
STDDev*	$466	$834		$1,594	$3,363

* 1,000 Simulations, n ≈ 250

** λ = 25; α = 0.999 RC; α = 0.9997 EC

6. Simulation Study: RCE vs. MLE

Table 4f:

RCE vs. LDA-MLE for Truncated GPD Severity ($\xi = 0.8675$, $\theta = 50,000$, H=\$10k)*

(millions)	Regulatory Capital**			Economic Capital**	
	RCE	**LDA-MLE**		**RCE**	**LDA-MLE**
Mean*	$466	$737		$1,327	$2,432
True Capital	$452	$452		$1,267	$1,267
Bias (Mean - True)	$13	**$285**		$61	**$1,166**
Bias %	3.0%	**63.0%**		4.8%	**92.0%**
RMSE*	$576	$1,062		$1,988	$4,337
STDDev*	$576	$1,023		$1,988	$4,177

* 1,000 Simulations, n ≈ 250

** λ = 25; α = 0.999 RC; α = 0.9997 EC

JSM 2014 August 2-7 BOSTON, MASSACHUSETTS

GE Capital

Table 5:
Summary of Capital Accuracy by Sample Size:
MLE vs. RCE ($millions) (across 6 severities, Opdyke, 2014)

	+-------------------- ECap --------------------+			+-------------------- RCap --------------------+		
	Mean Absolute Bias		Median Absolute Bias		Mean Absolute Bias		Median Absolute Bias	
λ =	RCE	MLE	RCE	MLE	RCE	MLE	RCE	MLE
15	**7.8%**	**92.6%**	2.6%	82.3%	**5.9%**	**61.6%**	1.6%	58.1%
25	**3.4%**	**53.1%**	3.3%	40.6%	**2.4%**	**38.1%**	2.0%	30.6%
50	**2.8%**	**25.7%**	2.7%	17.7%	**2.0%**	**19.4%**	1.9%	14.3%
75	**1.2%**	**15.5%**	0.8%	10.7%	**0.8%**	**11.9%**	0.5%	8.7%
100	**0.9%**	**11.3%**	0.5%	7.9%	**0.5%**	**8.7%**	0.4%	6.1%
15	**$61**	**$825**	$18	$502	**$21**	**$228**	$5	$154
25	**$45**	**$727**	$29	$410	**$14**	**$209**	$8	$133
50	**$69**	**$617**	$52	$320	**$20**	**$182**	$15	$109
75	**$40**	**$526**	$14	$250	**$11**	**$157**	$3	$80
100	**$32**	**$485**	$15	$223	**$7**	**$142**	$5	$73

NOTE: Even when relative absolute bias of MLE decreases, actual bias $ still are notable.

SIMULATION STUDY: Conclusions
RCE vs. MLE-LDA

a) **RCE is Dramatically More Accurate**: LDA-MLE Bias can be ENORMOUS: $Billion+ just for one uom!

b) **RCE is Notably More Precise**: Sometimes <50% RCE RMSE < MLE RMSE, RCE StdDev < MLE StdDev

c) **RCE is Consistently More Robust**: RCE Robustness to Violations of iid > MLE (see non-iid simulation study in Opdyke, 2014)

7. Summary and Conclusions

- **Under an LDA framework, operational risk capital estimates** based on the most commonly used estimators of severity parameters (e.g. MLE) and the relevant severity distributions **are consistently systematically biased upwards, presumably due to Jensen's inequality (Jensen, 1906).**

- **This bias is often** material, sometimes inflating required capital by **hundreds of millions, and even billions of dollars.**

- **RCE is the estimator MOST consistent with regulatory intent** regarding a prudent, responsible implementation of an AMA–LDA framework in that it alone is not systematically and materially biased, let alone imprecise and non-robust.

- **RCE is the only capital estimator that mitigates and nearly eliminates capital inflation under AMA-LDA. RCE also is notably more precise than LDA-based capital under most, if not all severity estimators, and consistently more robust to violations of i.i.d. data (which are endemic to operational risk loss data). Therefore, with greater capital accuracy, precision, and robustness, RCE unambiguously and notably improves LDA-based OpRisk Capital Estimation by all relevant criteria.**

8. References and Appendices

- Bardou, O., Frikha, N., and Pages, G. (2010), "CVaR Hedging Using Quantization Based Stochastic Approximation," working paper.

- Ben-Tal, A. (2007) "An Old-New Concept of Convex Risk Measures: The Optimized Certainty Equivalent," *Mathematical Finance*, 17(3), 449-476.

- Brown, D., (2007) "Large Deviations Bounds for Estimating Conditional Value-at-Risk," *Operations Research Letters*, 35, 722-730.

- Chernobai, A., Rachev, S., and Fabozzi, F. (2007), Operational Risk: A Guide to Basel II Capital Requirements, Models, and Analysis, John Wiley & Sons, Inc., Hoboken, New Jersey.

- Daníelsson, J., Jorgensen, B. N., Sarma, M., Samorodnitsky, G., and de Vries, C. G. (2005), "Subadditivity Re-examined: The Case for Value-at-Risk," *FMG Discussion Papers*, www.RiskResearch.org.

- Daníelsson, J., Jorgensen, B. N., Sarma, M., Samorodnitsky, G., and de Vries, C. G. (2013), "Fat Tails, VaR and Subadditivity," *Journal of Econometrics*, 172(2), 283-291.

- Degen, M. (2010), "The Calculation of Minimum Regulatory Capital Using Single-Loss Approximations," *The Journal of Operational Risk*, Vol. 5, No. 4, 3-17.

- Degen, M., Embrechts, P., Lambrigger, D.D. (2007), "The Quantitative Modeling of Operational Risk: Between g-and-h and EVT," *Astin Bulletin*, 37(2), 265-291.

- Embrechs, P., and Frei, M., (2009), "Panjer Recursion versus FFT for Compound Distributions," *Mathematical Methods of Operations Research*, 69(3), 497-508.

- Ergashev B., (2008), "Should Risk Managers Rely on the Maximum Likelihood Estimation Method While Quantifying Operational Risk," *The Journal of Operational Risk*, Vol. 3, No. 2, 63-86.

- Jensen, J. L. W. V. (1906), "Sur les fonctions convexes et les inégalités entre les valeurs moyennes," *Acta Mathematica*, 30 (1), 175–193.

- Johnson, R., and Wichern, D., (2007), Applied Multivariate Statistical Analysis, 6th Ed., Pearson, Upper Saddle River, NJ.

- Kennedy, P. (1992), A Guide to Econometrics, The MIT Press, Cambridge, MA.

- Kim, J., (2010), "Conditional Tail Moments of the Exponential Family and its Related Distributions," *North American Actuarial Journal*, Vol. 14(2), 198-216.

- Kim, J., and Hardy, M., (2007), "Quantifying and Correcting the Bias in Estimated Risk Measures," *Astin Bulletin*, Vol. 37(2), 365-386.

- Mayorov, K., (2014), Email correspondence with the author, September, 2014.

8. References and Appendices

- Opdyke, J.D., and Cavallo, A. (2012a), "Estimating Operational Risk Capital: The Challenges of Truncation, The Hazards of Maximum Likelihood Estimation, and the Promise of Robust Statistics," *The Journal of Operational Risk*, Vol. 7, No. 3, 3-90.

- Opdyke, J.D. and Cavallo, A. (2012b), "Operational Risk Capital Estimation and Planning: Exact Sensitivity Analysis and Business Decision Making Using the Influence Function," in <u>Operational Risk: New Frontiers Explored</u>, ed. E. Davis, Risk Books, Incisive Media, London.

- Opdyke, J.D. (2014), "Estimating Operational Risk Capital with Greater Accuracy, Precision, and Robustness," *forthcoming, The Journal of Operational Risk*, Vol. 9, No. 4, Decmeber.

- Panjer, H. (2006), <u>Operational Risk Modeling Analytics</u>, Wiley Series in Probability and Statistics, John Wiley & Sons, Inc., Hoboken, NJ.

- Roehr, A., (2002), "Modelling Operational Losses", *Algo Research Quarterly*, Vol.5, No.2, Summer 2002.

- Smith, J. (1987), "Estimating the Upper Tail of Flood Frequency Distributions," *Water Resources Research*, Vol.23, No.8, 1657-1666.

- Taleb, N., Canetti, E., Kinda, T., Loukoianova, E., and Schmieder, C., (2012), "A New Heuristic Measure of Fragility and Tail Risks: Application to Stress Testing," *IMF Working Paper*, Monetary and Capital Markets Department, WP/12/216.

- Taleb, N., and Douady, R., (2013), "Mathematical Definition, Mapping, and Detection of (Anti)Fragility," *Quantitative Finance*, forthcoming.

- Under the Basel II/III AMA, estimated capital requirements are the Value-at-Risk (VaR) quantile corresponding to the 99.9%Tile of the aggregate loss distribution, which is the convolution of the frequency and severity distributions. This convolution typically has no closed form, but its VaR may be obtained in a number of ways, including extensive monte carlo simulations, fast Fourier transform, Panjer recursion (see Panjer (2006) and Embrechts and Frei (2009)), and Degen's (2010) Single Loss Approximation. All are approximations, with the first as the gold standard providing arbitrary precision, and SLA as the fastest and most computationally efficient. SLA is implemented as below under three conditions (only a) is relevant for severities that cannot have infinite mean):

a) if $\xi < 1$, $\qquad C_\alpha \approx F^{-1}\left(1 - \frac{1-\alpha}{\lambda}\right) + \lambda\mu$ where μ is the mean of F

b) if $\xi = 1$, $\qquad C_\alpha \approx F^{-1}\left(1 - \frac{1-\alpha}{\lambda}\right) + c_\xi \lambda \mu_F\left[F^{-1}\left(1 - \frac{1-\alpha}{\lambda}\right)\right]$ $\quad$ where $\quad c_\xi = 1, \ \mu_F(x) = \int_0^x \left[1 - F(s)\right] ds$

c) if $1 < \xi < 2$, $\ C_\alpha \approx F^{-1}\left(1 - \frac{1-\alpha}{\lambda}\right) - (1-\alpha) F^{-1}\left(1 - \frac{1-\alpha}{\lambda}\right) \cdot \left(\frac{c_\xi}{1 - 1/\xi}\right)$ $\qquad$ where $c_\xi = (1 - \xi)\dfrac{\Gamma^2\left(1 - 1/\xi\right)}{2\Gamma\left(1 - 2/\xi\right)}$

$\qquad$ ($\xi \geq 2$ is so extreme as to not be relevant in this setting)

(the above assumes a Poisson-distributed frequency distribution and can be modified if this assumption does not hold)

- When implementing the above it is important to note that the capital estimate diverges as $\xi \to 1$; specifically, for a) $C_\alpha \to +\infty$ as $\xi \to 1^-$ and for c) $C_\alpha \to -\infty$ as $\xi \to 1^+$. **Note that this divergence does not only occur for small deviations** from $\xi = 1$. For example, for GPD, divergence can be noticeable in the range of $0.8 < \xi < 1.2$. Therefore, one must utilize a nonlinear interpolation or an alternative derivation of Degen's formulae to avoid this obstacle. All results relying on SLA herein utilize the former **solution** – i.e. "**ISLA**" (see Opdyke, 2014) and were all tested to be within 1% of extensive monte carlo results (e.g. five million years' worth of monte carlo loss simulations).

Figures A1-A4: ISLA Correction for SLA Divergence at Root of ξ=1 for GPD Severity (θ = 55,000)

REGULATORY

ECONOMIC
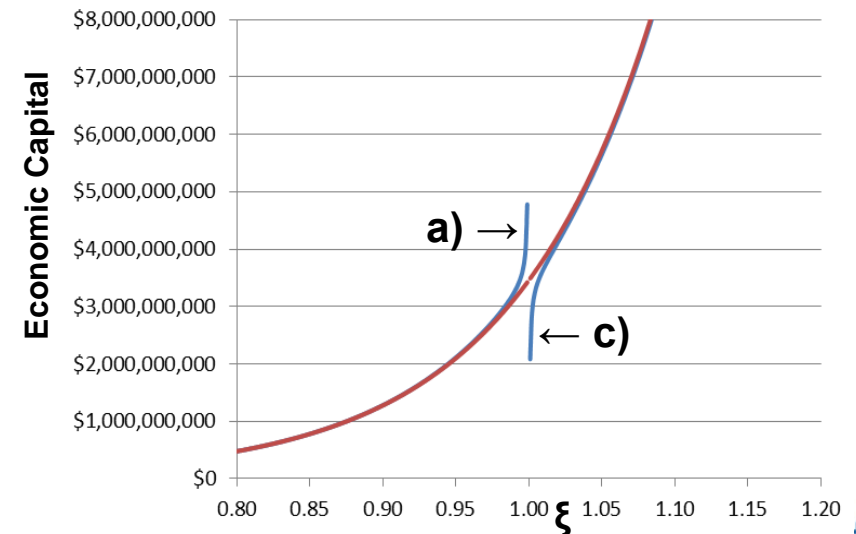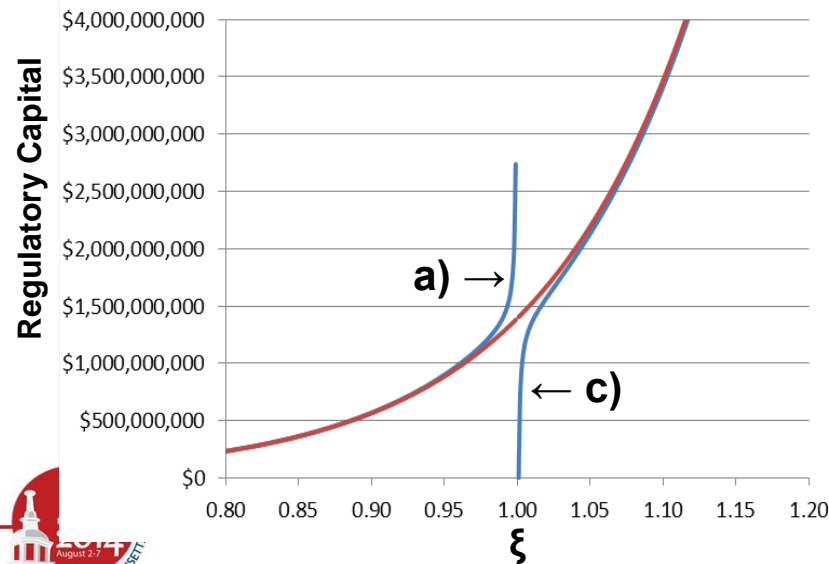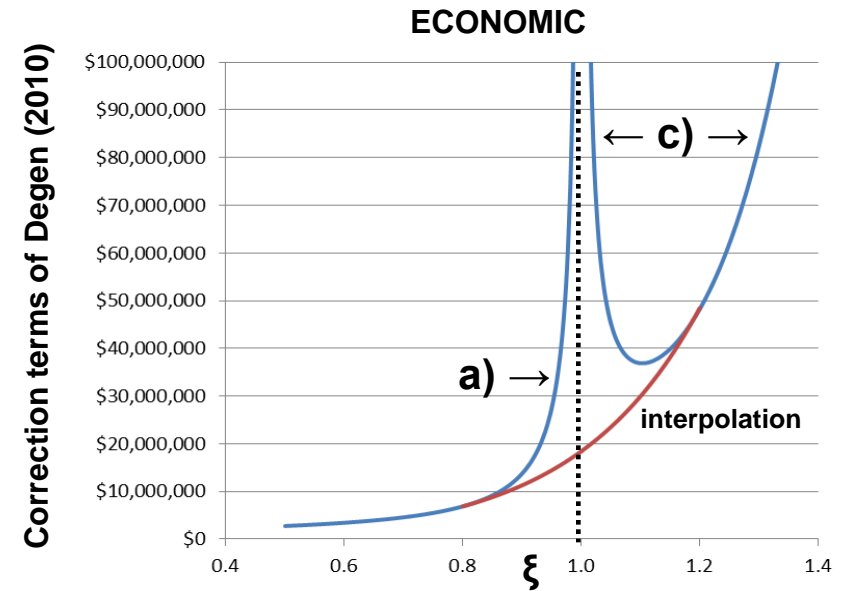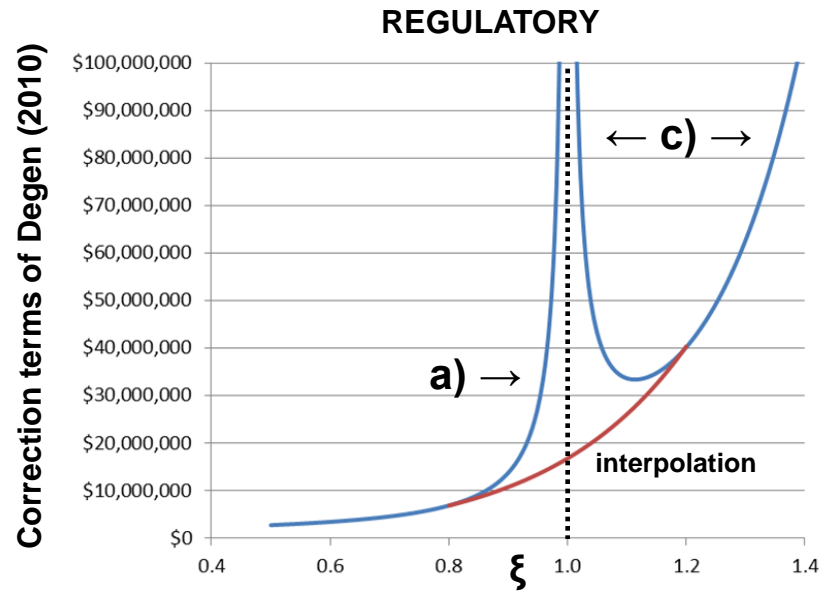
- As currently implemented per Basel II/III's AMA-LDA, operational risk capital is a value-at-risk (VaR) estimate (i.e. the quantile corresponding to p = 0.999, the 99.9%tile) of the aggregate loss distribution. As shown by Degen (2010), this is essentially a high quantile of the severity distribution. For those severities relevant to operational risk capital estimation, VaR always appears to be a convex function of the severity distribution parameter estimates **as long as the quantile being estimated is large enough** (e.g. corresponding to p>0.999; see Degen, Embrechts, & Lambrigger, 2007; Daníelsson et al., 2005; and Daníelsson et al., 2013). For the heavy-tailed severities examined above, in addition to two others sometimes used in this space (Burr type XII and LogLogistic), we see:

TABLE A1: Marginal VaR Behavior OVER RELEVANT DOMAIN (p > 0.999) by Severity

Severity Distribution	VaR is Convex/Linear as Function of...			Relationship between Parameters
	Parameter 1	Parameter 2	Parameter 3	
1) **LogNormal (μ, σ)**	Convex	Convex		Independent
2) **LogLogistic (α, β)**	Linear	Convex		Independent
3) **LogGamma (a, b)**	Convex	Convex		Dependent
4) **GPD (ξ, θ)**	Convex	Linear		Dependent
5) **Burr (type XII) (ϒ, α, β)**	Convex	Convex	Linear	Dependent
6) **Truncated 1)**	Convex	Convex		Dependent
7) **Truncated 2)**	Linear	Convex		Dependent
8) **Truncated 3)**	Convex	Convex		Dependent
9) **Truncated 4)**	Convex	Linear		Dependent
10) **Truncated 5)**	Convex	Convex	Linear	Dependent

As mentioned above (p.16), VaR empirical convexity increases in *p*: larger quantiles are associated with greater convexity.

- **PDF and CDF of LogNormal:**

$$f(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma x} e^{-\frac{1}{2}\left(\frac{\ln(x)-\mu}{\sigma}\right)^2} \qquad F(x;\mu,\sigma) = \frac{1}{2}\left[1+erf\left(\frac{\ln(x)-\mu}{\sqrt{2\sigma^2}}\right)\right] \qquad 0 < x < \infty,\ 0 < \sigma < \infty$$

- **Mean of LogNormal:** $E(X) = e^{\left(\mu+\sigma^2/2\right)}$

- **Inverse Fisher information of LogNormal:**

$$A(\theta)^{-1} = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2/2 \end{bmatrix}$$

- ## PDF and CDF of Truncated LogNormal:

$$g(x; \mu, \sigma) = \frac{f(x; \mu, \sigma)}{1 - F(H; \mu, \sigma)} \qquad G(x; \mu, \sigma) = 1 - \frac{1 - F(x; \mu, \sigma)}{1 - F(H; \mu, \sigma)}$$

$$H < x < \infty, \ 0 < \sigma < \infty$$
$$f(\ ) \text{ is LogNormal PDF and } F(\ ) \text{ is LogNormal CDF}$$

- ## Mean of Truncated LogNormal:

$$E(X) = e^{\mu + \sigma^2/2} \cdot \Phi\left(\frac{\mu + \sigma^2 - \ln(H)}{\sigma}\right) \cdot \frac{1}{\left[1 - F(H)\right]} \quad \text{where } \Phi(\ ) \text{ is the standard normal CDF.}$$

- ## Inverse Fisher information of Truncated LogNormal:

$$\text{Let } u = \frac{\ln(H) - \mu}{\sigma}, \quad j = \frac{-u^2/2}{\sqrt{2\pi}}, \quad J = \frac{j}{1 - \Phi(u)}, \text{ where } \Phi = \text{CDF of Standard Normal, and } INV = \frac{\sigma^2}{\left[2 + J \cdot (J - u) \cdot (u \cdot (J - u) - 3)\right]}$$

$$\text{Then } A(\theta)^{-1} = INV \cdot \begin{bmatrix} 2 + J \cdot u \cdot (1 - u \cdot (J - u)) & J \cdot (u \cdot (J - u) - 1) \\ J \cdot (u \cdot (J - u) - 1) & 1 - (J \cdot (J - u)) \end{bmatrix}$$

From Roehr (2002). Note that the first cell of this matrix as presented in Roehr, 2002, contains a typo: this is corrected in the presentation above.

8. Appendix C: Severity PDFs, CDFs, & Means for Capital Approximation

- ## PDF and CDF of Generalized Pareto Distribution (GPD):

$$f(x; \xi, \theta) = \frac{1}{\theta}\left[1 + \xi\frac{x}{\theta}\right]^{\left[-\frac{1}{\xi}-1\right]} \qquad F(x; \xi, \theta) = 1 - \left[1 + \xi\frac{x}{\theta}\right]^{\left[-\frac{1}{\xi}\right]} \qquad \text{assuming } \xi \geq 0, \text{ for } 0 \leq x < \infty; \ 0 < \theta < \infty$$

- ## Mean of GPD: $E(X) = \dfrac{\theta}{1-\xi}$ for $\xi < 1$ $\left(= \infty \text{ for } \xi \geq 1\right)$

- ## Inverse Fisher information of GPD:

$$A(\theta)^{-1} = (1+\xi)\begin{bmatrix} 1+\xi & -\theta \\ -\theta & 2\theta^2 \end{bmatrix}$$

From Smith (1987)

8. Appendix C: Severity PDFs, CDFs, & Means for Capital Approximation

- **PDF and CDF of Truncated GPD:**

$$g(x;\xi,\theta) = \frac{f(x;\xi,\theta)}{1 - F(H;\xi,\theta)} \qquad G(x;\xi,\theta) = 1 - \frac{1 - F(x;\xi,\theta)}{1 - F(H;\xi,\theta)}$$

assuming $\xi \geq 0$, for $H \leq x < \infty$; $0 < \theta < \infty$

$f(\ )$ is GPD PDF and $F(\ )$ is GPD CDF

- **Mean of Truncated GPD:** $E(X) = \dfrac{\theta}{\xi} \cdot \left(\dfrac{\left[1 - F(H)\right]^{-\xi}}{1-\xi} - 1 \right)$ for $\xi < 1$ $\left(= \infty \text{ for } \xi \geq 1\right)$

- **Inverse Fisher information of Truncated GPD:**

$$A(\theta)^{-1} = (1+\xi) \cdot \begin{bmatrix} (1+\xi) & -\theta\left(1 + (1+2\xi)\left(\dfrac{H}{\theta}\right)\right) \\ -\theta\left(1 + (1+2\xi)\left(\dfrac{H}{\theta}\right)\right) & \theta^2\left(2 + 2(1+2\xi)\left(\dfrac{H}{\theta}\right) + (1+\xi)(1+2\xi)\left(\dfrac{H}{\theta}\right)^2\right) \end{bmatrix}$$

From Roehr (2002)

8. Appendix C: Severity PDFs, CDFs, & Means for Capital Approximation

- ## PDF and CDF of LogGamma*:

$$f(x;a,b) = \frac{b^a \left(\log(x)\right)^{(a-1)}}{\Gamma(a) x^{b+1}} \qquad F(x;a,b) = \int_1^x \frac{b^a \left(\log(y)\right)^{(a-1)}}{\Gamma(a) y^{b+1}} dy \qquad 1 \le x < \infty; \; 0 < a; \; 0 < b$$

where $\Gamma(a)$ is the complete gamma function

- ## Mean of LogGamma: $E(X) = \left(\dfrac{b}{b-1}\right)^a$ for $b > 1$; otherwise $E(X) = \infty$

- ## Inverse Fisher information of LogGamma:

$$A(\theta)^{-1} = \frac{1}{\left(a/b^2\right) \cdot trigamma(a) - 1/b^2} \begin{bmatrix} a/b^2 & 1/b \\ 1/b & trigamma(a) \end{bmatrix}$$

From Opdyke and Cavallo (2012a)

*NOTE that a location parameter can be added to change the lower end of the domain to zero, but this is unnecessary in this setting. Also note that this is the "rate" or "inverse scale" parameterization of the LogGamma, which can also be defined with a "scale" parameterization wherein b = 1/b.

8. Appendix C: Severity PDFs, CDFs, & Means for Capital Approximation

- ## PDF and CDF of Truncated LogGamma*:

$$g(x;a,b) = \frac{f(x;a,b)}{1-F(H;a,b)} \qquad G(x;a,b) = 1 - \frac{1-F(x;a,b)}{1-F(H;a,b)}$$

$H \le x < \infty; \ 0 < a; \ 0 < b$

$f(\ )$ is GPD PDF and $F(\ )$ is GPD CDF

- ## Mean of Truncated LogGamma:

$$E(X) = \left(\frac{b}{b-1}\right)^a \cdot \frac{1-J\big(\log(H)(b-1);a,1\big)}{\big[1-F(H)\big]} \quad \text{for } b>1, \text{ otherwise } E(X) = \infty$$

where $J(\ )$ is the CDF of the Gamma distribution.

From Opdyke (2014)

- ## Inverse Fisher information of Truncated LogGamma:

- **Inverse Fisher info. of Truncated LogGamma*:** $A(\theta)^{-1} = \begin{bmatrix} A & B \\ B & D \end{bmatrix}^{-1}$ where

$$A = trigamma(a) - \frac{\left[\int_{1^+}^{H} \ln(b) + \ln(\ln(x)) - digamma(a) f(x) dx\right]^2}{\left[1 - F(H;a,b)\right]^2} - \frac{\left[1 - F(H;a,b)\right] \cdot \int_{1^+}^{H} \left[\ln(b) + \ln(\ln(x)) - digamma(a)\right]^2 - trigamma(a) f(x) dx}{\left[1 - F(H;a,b)\right]^2}$$

$$B = -\frac{1}{b} - \frac{\left[1 - F(H;a,b)\right] \cdot \frac{1}{b} \cdot F(H;a,b)}{\left[1 - F(H;a,b)\right]^2} - \frac{\left[1 - F(H;a,b)\right] \cdot \int_{1^+}^{H} \left[\ln(b) + \ln(\ln(x)) - digamma(a)\right] \cdot \left[\frac{a}{b} - \ln(x)\right] f(x) dx}{\left[1 - F(H;a,b)\right]^2}$$

$$-\frac{\int_{1^+}^{H} \ln(b) + \ln(\ln(x)) - digamma(a) f(x) dx \cdot \int_{1^+}^{H} \left(\frac{a}{b} - \ln(x)\right) f(x) dx}{\left[1 - F(H;a,b)\right]^2}$$

$$D = \frac{a}{b^2} - \frac{\left[\int_{1^+}^{H} \left(\frac{a}{b} - \ln(y)\right) f(x) dx\right]^2 + \left[1 - F(H;a,b)\right] \cdot \int_{1^+}^{H} \frac{a(a-1)}{b^2} - \frac{2a\ln(y)}{b} + \left[\ln(y)\right]^2 f(x) dx}{\left[1 - F(H;a,b)\right]^2}$$

From Opdyke and Cavallo (2012b)

*The digamma and trigamma functions are the first and second order logarithmic derivatives of the complete gamma function:
$\text{digamma(z)} = \partial/\partial z \ln\left[\Gamma(z)\right]$ and $\text{trigamma(z)} = \partial^2/\partial z^2 \ln\left[\Gamma(z)\right]$.

• Inverse Fisher information of Truncated LogGamma:

To avoid computationally expensive numeric integration, Opdyke (2014) derives the analytic approximation below:

$$A(\theta)^{-1} = \begin{bmatrix} A & B \\ B & D \end{bmatrix}^{-1} \quad \text{where}$$

$$A = \frac{1}{a^4 UIG^2} \times \left\{ \left[-(GHG2)^2 \right] \cdot (-z)^{2a} + 2a(-z)^a \cdot \left[-UIG \cdot GHG3 + a\Gamma(a) \cdot GHG2 \cdot \left(Log(-z) - digamma(a) \right) \right] \right.$$

$$\left. + a^4 \Gamma(a) \left[-\left(\Gamma(a) - UIG \right) \cdot \left(Log(-z) - digamma(a) \right)^2 + UIG \cdot trigamma(a) \right] \right\}$$

$$B = \frac{1}{a^2 b UIG^2} \times \left\{ t^{-b} \cdot GHG2 \cdot (-z)^{2a} - a^2 \left(t^b UIG^2 + \Gamma(a)(-z)^a \left(Log(-z) - digamma(a) \right) \right) \right\}$$

$$D = \frac{a}{b^2} + \frac{t^{-b}(-z)^a (1 - a - z)}{b^2 UIG} - \frac{t^{-2b}(-z)^{2a}}{b^2 UIG^2}$$

where...

- ## Inverse Fisher information of Truncated LogGamma:

where...

$t =$ data collection (truncation) threshold

$\eta = 0.001$

$adown = a - \eta$

$aup = a + \eta$

$z = -bLog[t]$

divide $a = diva = \dfrac{\Gamma(a+1)}{(-z)^a}$

divide $adown = divad = \dfrac{\Gamma(adown+1)}{(-z)^{adown}}$

divide $aup = divau = \dfrac{\Gamma(aup+1)}{(-z)^{aup}}$

$$GHG2 = divad \cdot J(-z; adown, 1) \frac{aup}{aup - adown} + divau \cdot J(-z; aup, 1) \frac{adown}{adown - aup}$$

$$GHG3 = divad \cdot J(-z; adown, 1) \left(\frac{aup}{aup - adown}\right)\left(\frac{a}{a - adown}\right) + diva \cdot J(-z; a, 1)\left(\frac{adown}{adown - a}\right)\left(\frac{aup}{aup - a}\right)$$

$$+ divau \cdot J(-z; aup, 1)\left(\frac{adown}{adown - aup}\right)\left(\frac{a}{a - adown}\right) \quad \text{where } J(\;) \text{ is the CDF of the Gamma distribution.}$$

$$UIG = \text{upper incomplete gamma function} = \Gamma(a, -z) = \Gamma(a)\left(1 - J(-z; a, b = 1)\right)$$

Estimating Operational Risk Capital with Greater Accuracy, Precision, and Robustness

The largest US banks and Systemically Important Financial Institutions are required by regulatory mandate to estimate the operational risk capital they must hold using an Advanced Measurement Approach (AMA) as defined by the Basel II/III Accords. Most of these institutions use the Loss Distribution Approach (LDA) which defines the aggregate loss distribution as the convolution of a frequency distribution and a severity distribution representing the number and magnitude of losses, respectively. Capital is a Value-at-Risk estimate of this annual loss distribution (i.e. the quantile corresponding to the 99.9%tile, representing a one-in-a-thousand-year loss, on average). In practice, the severity distribution drives the capital estimate, which is essentially a very large quantile of the estimated severity distribution. Unfortunately, when using LDA with any of the widely used severity distributions (i.e. heavy-tailed, skewed distributions), all unbiased estimators of severity distribution parameters generate biased capital estimates apparently due to Jensen's Inequality: VaR always appears to be a convex function of these severities' parameter estimates because the (severity) quantile being estimated is so large and the severities are heavy-tailed. The resulting bias means that capital requirements always will be overstated, and this inflation is sometimes enormous (sometimes even billions of dollars at the unit-of-measure level). Herein I present an estimator of capital that essentially eliminates this upward bias when used with any commonly used severity parameter estimator. The Reduced-bias Capital Estimator (RCE), consequently, is more consistent with regulatory intent regarding the responsible implementation of the LDA framework than other implementations that fail to mitigate, if not eliminate this bias. RCE also notably increases the precision of the capital estimate and consistently increases its robustness to violations of the i.i.d. data presumption (which are endemic to operational risk loss event data). So with greater capital accuracy, precision, and robustness, RCE lowers capital requirements at both the unit-of-measure and enterprise levels, increases capital stability from quarter to quarter, ceteris paribus, and does both while more accurately and precisely reflecting regulatory intent. RCE is straightforward to explain, understand, and implement using any major statistical software package.

CONTACT

J.D. Opdyke
Head of Operational Risk Modeling,
Quantitative Methods Group, GE Capital
J.D.Opdyke@ge.com
617-943-6463

GE Capital