Operational Risk Modeling: Three Material Improvements to AMA/LDA OpRisk Capital Estimation Frameworks

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

Disclaimer

Presented at the 18th Annual Operational Risk – North America conference, New York, New York, March 15-16, 2016.

The views presented herein are the views of the sole author and do not reflect the views of GE Capital or any other institution.

Contents

Contents

I. Beyond "Red-Light, Green-Light" KRIs…

Key Risk Indicator (KRI) Data
(a.k.a. Business Environment and Internal Control Factor (BEICF) Data) –

Econometric Methods for Establishing Direct, Material, Statistically Causal Relationships between KRIs and i. Capital, ii. Loss Frequency, and iii. Loss Severity COMPLETELY CONSISTENT AND COMPATIBLE WITH AMA/LDA:

KRI Data can (and should!) be used directly in capital modeling. **Establishing material, statistically causal relationships between KRIs and capital is the only way that operational risk management and mitigation efforts can have direct and desired effects on capital requirements.**

For example, this gives the operational risk capital analyst the means by which to make statements to, say, the head of the trading shop such as, "If you can decrease your system downtime by a standard deviation, or X%, I can take $40m in capital off the table for you, all else equal."

This is accomplished using multivariate econometric (regression) techniques to estimate frequency and severity parameters based directly on the KRI Data. This is directly analogous to knowing the drivers of, say, a PD model when estimating capital for credit risk.

I. OpRisk Management/Mitigation via KRI Regression

Econometric Methods for Establishing Direct, Material, Statistically Causal Relationships between KRIs and i. Capital, ii. Loss Frequency, and iii. Loss Severity COMPLETELY CONSISTENT AND COMPATIBLE WITH AMA/LDA:

WHY Multivariate Regression?

Multivariate regression is needed to control for covariance betwixt the KRI covariates. **Multivariate regression is the only way to estimate the effect of an independent variable (a particular KRI) on a dependent variable (capital) holding all else constant**, that is, without capturing the effects of other KRIs that to some degree move in tandem with the one in question.

Without a regression to "hold all else constant" and eliminate the confounding effect of, say trading volume, when estimating the effect of system downtime on operational risk capital, the estimate of the effect of system downtime will be biased, and inference based on it will be misinformed, and the mitigation efforts based on it will be misguided and likely ineffective.

OpRisk NorthAmerica

I. Beyond "Red-Light, Green-Light" KRIs…

Econometric Methods for Establishing Direct, Material, Statistically Causal Relationships between KRIs and i. Capital, ii. Loss Frequency, and iii. Loss Severity COMPLETELY CONSISTENT AND COMPATIBLE WITH AMA/LDA:

WHY Multivariate Regression?

This, of course, presumes that relationships (covariance) exist betwixt relevant KRIs, as it does in the real world (if it did not, there would be no need for multivariate regression here).

Multivariate regression also increases the precision with which we are able to estimate the frequency and severity parameters. We are using additional data in the estimation, which **will increase statistical power** (even though we are not increasing sample size in the form of additional loss events).

I. OpRisk Management/Mitigation via KRI Regression

Econometric Methods for Establishing Direct, Material, Statistically Causal Relationships between KRIs and i. Capital, ii. Loss Frequency, and iii. Loss Severity COMPLETELY CONSISTENT AND COMPATIBLE WITH AMA/LDA:

WHY Multivariate Regression? ONLY this approach provides

1. **Statistically Causal Relationships between KRIs and Capital, AND KRIs and LOSS FREQUENCY AND SEVERITY (… NOT JUST CAPITAL!!!)**
2. **Magnitude of Effect of Each KRI** on i. Capital AND ii. LOSS FREQUENCY AND SEVERITY Independent of other KRIs
3. **RELATIVE IMPORTANCE of Each KRI's Effect** on i. Capital AND ii. LOSS FREQUENCY AND SEVERITY Independent of other KRIs (key for $allocation for mitigation efforts)
4. **Direction of Effect of Each KRI** on i. Capital AND ii. LOSS FREQUENCY AND SEVERITY Independent of other KRIs
5. **Whether Effect of Each KRI is Statistically Significant** vis-à-vis i. Capital AND ii. LOSS FREQUENCY AND SEVERITY Independent of other KRIs
6. **Whether Effect of Each KRI is Material** vis-à-vis i. Capital AND ii. LOSS FREQUENCY AND SEVERITY Independent of other KRIs
7. **Increase in the Precision of the Estimate of Capital** (AND Frequency and Severity), all else equal

OpRisk NorthAmerica

I. Beyond "Red-Light, Green-Light" KRIs…

Econometric Methods for Establishing Direct, Material, Statistically Causal Relationships between KRIs and i. Capital, ii. Loss Frequency, and iii. Loss Severity COMPLETELY CONSISTENT AND COMPATIBLE WITH AMA/LDA:

REQUIREMENTS

KRI Data for modeling purposes must be disaggregated at the level of the loss event. In other words, it must be "**granular**," with data points for each KRI collected associated with each individual loss (or timing that concurs with the loss).

This is distinct from what many (non-modelers) refer to as "KRIs," which are typically highly aggregated, descriptive statistics that are tracked over time and used to guide operational risk management and mitigation efforts directly, rather than via an estimation process that links them to capital (or some other outcome measure). **Aggregated KRIs typically are used non-inferentially, to identify "Red Lights," "Amber Lights," and "Green Lights."**

I. OpRisk Management/Mitigation via KRI Regression

Econometric Methods for Establishing Direct, Material, Statistically Causal Relationships between KRIs and i. Capital, ii. Loss Frequency, and iii. Loss Severity COMPLETELY CONSISTENT AND COMPATIBLE WITH AMA/LDA:

METHODS

Frequency: Poisson and Negative Binomial Regression
- Time tested, decades old methods applied in many fields.
- However, doesn't move the capital needle nearly as much as severity.

Severity: Scale regression
- More recent, main difference is just the link function.
- DOES move the capital needle, sometimes dramatically.
- This is a Scale Regression, and so the Severity requires a scale parameter.

GAMLSS (Generalized Additive Models of Location, Scale, and Shape) Regression:
- Most general, covariates apply to location, scale, and shape parameters.
- In literature and applied use at least as long as Operational Risk has been a discipline (see Rigby and Stasinopoulos, 2001).

Econometric Methods for Establishing Direct, Material, Statistically Causal Relationships between KRIs and i. Capital, ii. Loss Frequency, and iii. Loss Severity COMPLETELY CONSISTENT AND COMPATIBLE WITH AMA/LDA:

METHODS

Frequency: Poisson and Negative Binomial Regression

$\ln\left(E\left[Y \mid x\right]\right) = \boldsymbol{\beta}'\mathbf{x}$ where $\mathbf{x}$ is a vector of regressor variables.

$$E\left[Y \mid x\right] = \exp\left(\boldsymbol{\beta}'\mathbf{x}\right) = \lambda \qquad f_X\left(x \mid \lambda\right) = \frac{\lambda^x \exp\left(-\lambda\right)}{x!} \qquad f_Y\left(y \mid \mathbf{x};\boldsymbol{\beta}\right) = \frac{\left(\boldsymbol{\beta}'\mathbf{x}\right)^y \exp\left(-\boldsymbol{\beta}'\mathbf{x}\right)}{y!}$$

$$L\left(\boldsymbol{\beta} \mid X,Y\right) = \prod_{i=1}^{n} \frac{\left(\beta' x_i\right)^{y_i} \exp\left(-\beta' x_i\right)}{y_i!}$$

$$l\left(\boldsymbol{\beta} \mid X,Y\right) = \ln\left[L\left(\boldsymbol{\beta} \mid X,Y\right)\right] = \sum_{i=1}^{n}\left(y_i \ln\left(\beta' x_i\right) - \beta' x_i - \ln\left(y_i!\right)\right)$$

$$l\left(\boldsymbol{\beta} \mid X,Y\right) = \sum_{i=1}^{n}\left(y_i \ln\left(\beta' x_i\right) - \beta' x_i\right) \qquad \frac{\partial l\left(\boldsymbol{\beta} \mid X,Y\right)}{\partial \boldsymbol{\beta}} = 0, \text{ no closed form solution,}$$

OpRisk NorthAmerica

Econometric Methods for Establishing Direct, Material, Statistically Causal Relationships between KRIs and i. Capital, ii. Loss Frequency, and iii. Loss Severity COMPLETELY CONSISTENT AND COMPATIBLE WITH AMA/LDA:

METHODS

Severity: Scale Regression

$Y \sim \Im(\theta, \Omega)$ such that θ is affected by the regressors as

$$\theta = \theta_0 \cdot \exp\left(\sum_{i=1}^{k} \beta_i x_i \right)$$

where θ_0 is the base value of the scale parameter,

$\Im$ is the distribution of Y with nonscale parameters Ω and scale parameter θ

and x_i are k regressors and β_i are the corresponding parameters.

θ is a scale parameter iff $f\left(x;\theta,\beta\right)=\frac{1}{\theta} f\left(\frac{x}{\theta};1,\beta\right)$ and $F\left(x;\theta,\beta\right)=F\left(\frac{x}{\theta};1,\beta\right)$.

I. Beyond "Red-Light, Green-Light" KRIs…

Econometric Methods for Establishing Direct, Material, Statistically Causal Relationships between KRIs and i. Capital, ii. Loss Frequency, and iii. Loss Severity COMPLETELY CONSISTENT AND COMPATIBLE WITH AMA/LDA:

METHODS

GAMLSS Regression

if $Y_i \sim f\left(y_i;\mu_i,\sigma_i,\tau_i\right)$; $i=1,\ldots,N$; and X_{ikj_k} are j_k covariates; $k=1,\ldots,p$ parameters;

$g_k\left(\theta_k\right)=\eta_k=h_k\left(X_k,\beta_k\right)$ and μ_i,σ_i, and τ_i are location, scale, and shape parameters, θ_k

$$g_1\left(\mu\right)=\eta_1=h_1\left(X_1,\beta_1\right)=\beta_{11}+\beta_{12}X_{i12}+\cdots+\beta_{1j_1}X_{i1j_1}$$

$$g_2\left(\sigma\right)=\eta_2=h_2\left(X_2,\beta_2\right)=\beta_{21}+\beta_{22}X_{i22}+\cdots+\beta_{2j_2}X_{i2j_2}$$

$$g_3\left(\tau\right)=\eta_3=h_3\left(X_3,\beta_3\right)=\beta_{31}+\beta_{32}X_{i32}+\cdots+\beta_{3j_3}X_{i3j_3}$$

$$\hat{\theta}_k=\underset{\theta_k\in\Theta}{\arg\max}\left(\sum_{i=1}^{N}\log\left[f\left(y_i\mid\mu_i,\sigma_i,\tau_i\right)\right]\right)$$

for the parametric version, and a penalized log likelihood for the semi-parametric version.

GAMLSS can include both linear & non-linear effects.

OpRisk NorthAmerica

I. OpRisk Management/Mitigation via KRI Regression

Econometric Methods for Establishing Direct, Material, Statistically Causal Relationships between KRIs and i. Capital, ii. Loss Frequency, and iii. Loss Severity COMPLETELY CONSISTENT AND COMPATIBLE WITH AMA/LDA:

IDENTIFYING THE RIGHT KRIs

Everybody knows what factors drive a PD model in credit risk. Why do we not automatically know the KRIs that drive operational risk capital?

- OpRisk is much Newer

- A much broader question (and Risk Type) across a much broader range of very different businesses, products, and specific risk types.

- This question needs to be answered for each UoM, and while there may be some overlap, there will be considerable differences in the things that drive trading-related operational risk losses in EDPM event type vs., say, class action litigation losses in the CPBP event type.

I. Beyond "Red-Light, Green-Light" KRIs…

Econometric Methods for Establishing Direct, Material, Statistically Causal Relationships between KRIs and i. Capital, ii. Loss Frequency, and iii. Loss Severity COMPLETELY CONSISTENT AND COMPATIBLE WITH AMA/LDA:

IDENTIFYING THE RIGHT KRIs

But smart operational risk researchers have developed useful guidelines in the search for and development of material and statistically significant KRIs (**although the only way to "know" this, ultimately, is to use them in the regression**).

Econometric Methods for Establishing Direct, Material, Statistically Causal Relationships between KRIs and i. Capital, ii. Loss Frequency, and iii. Loss Severity COMPLETELY CONSISTENT AND COMPATIBLE WITH AMA/LDA:

IDENTIFYING THE RIGHT KRIs: Cruz (2012) – useful KRIs must be

a. **Objective**: Quantitative and objectively measurable.
b. **Simple**: The simpler they are, the more likely they are to be used. Also, highly complex KRIs confound issues of causality with issues of functional form and possible overfitting.
c. **Identifiable**: Explicitly identified with a cost center (or no one will maintain and champion it)
d. **Representative**: …of a critical path process (or no one will maintain/champion it)

(“maintaining” KRIs in current feedback loops is discussed below…)

I. Beyond "Red-Light, Green-Light" KRIs…

Econometric Methods for Establishing Direct, Material, Statistically Causal Relationships between KRIs and i. Capital, ii. Loss Frequency, and iii. Loss Severity COMPLETELY CONSISTENT AND COMPATIBLE WITH AMA/LDA:

IDENTIFYING THE RIGHT KRIs: Chapelle (2013) – useful KRIs must be

a. **KEY**: "Essential/Relevant/Few" … less is more. Many for the sake of comprehensiveness is an utter waste.

b. **RISK**: Must address RISK, i.e. the possibility of a negative outcome, not losses that have already taken place (e.g. customer satisfaction – actually a performance indicator with little predictive power re: operational risk … must address the root cause of a negative outcome, not the symptom).

c. **INDICATOR**: "Predictor" is better word… must be a predictor of the future, not a descriptor of the past or present. Must indicate an increase in the likelihood of an operational risk loss. Control failures are good candidates.

I. OpRisk Management/Mitigation via KRI Regression

<u>Econometric Methods for Establishing Direct, Material, Statistically Causal Relationships between KRIs and i. Capital, ii. Loss Frequency, and iii. Loss Severity COMPLETELY CONSISTENT AND COMPATIBLE WITH AMA/LDA</u>:

Chapelle (2012) – good candidates for predictive, forward-looking KRIs:

i. **Failed KPIs**: Poor performance is often a source of risk.

ii. **Failed Controls**: Most obvious source of preventative indicators: the role of a control is to reduce risk, so failed KCI is also a KRI. Also, failed performance of a control function – e.g. a back office – is a KPI, a KRI, and a KCI (**<u>note that to a multivariate regression, data is data: it does not matter whether it is called a KCI or a KRI!</u>**).

iii. **Cause of the Cause**: Root causes. Why the "human error?" Lack of skill? Lack of training? Lack of attention (longer than mandated hours? History of poor individual performance?).

iv. **Environmental KRIs**: Relate to context in which business operates, e.g. macro economic environment, financial markets, etc. Active awareness of broader environment, not merely tracking firm-centric, descriptive "KRIs."

I. Beyond "Red-Light, Green-Light" KRIs…

Econometric Methods for Establishing Direct, Material, Statistically Causal Relationships between KRIs and i. Capital, ii. Loss Frequency, and iii. Loss Severity COMPLETELY CONSISTENT AND COMPATIBLE WITH AMA/LDA:

Chapelle (2012) – be SMART when designing/defining predictive, forward-looking KRIs:

i. **Specific**: KRIs must reflect specific business activity.

ii. **Measurable**: Especially for modeling, must be quantifiably measurable.

iii. **Actionable**: Without a clear path to mitigation, it serves no purpose.

iv. **Realistic**: Actions indicated by changes in KRIs must be reasonable and credible. For example, "zero tolerance" of a risk is not realistic. Risk mitigation appropriately reflects well defined risk appetite: no risk, no return.

v. **Timely**: Must be as close to real time as possible to maximize utility to the business.

I. OpRisk Management/Mitigation via KRI Regression

Econometric Methods for Establishing Direct, Material, Statistically Causal Relationships between KRIs and i. Capital, ii. Loss Frequency, and iii. Loss Severity COMPLETELY CONSISTENT AND COMPATIBLE WITH AMA/LDA:

Proactively Selecting and Maintaining the Right KRIs:

KRIs must be current. They must be maintained – continually assessed and reevaluated, simultaneously, together – in a dynamic feedback loop based on current data.

I. Beyond "Red-Light, Green-Light" KRIs…

Econometric Methods for Establishing Direct, Material, Statistically Causal Relationships between KRIs and i. Capital, ii. Loss Frequency, and iii. Loss Severity COMPLETELY CONSISTENT AND COMPATIBLE WITH AMA/LDA:

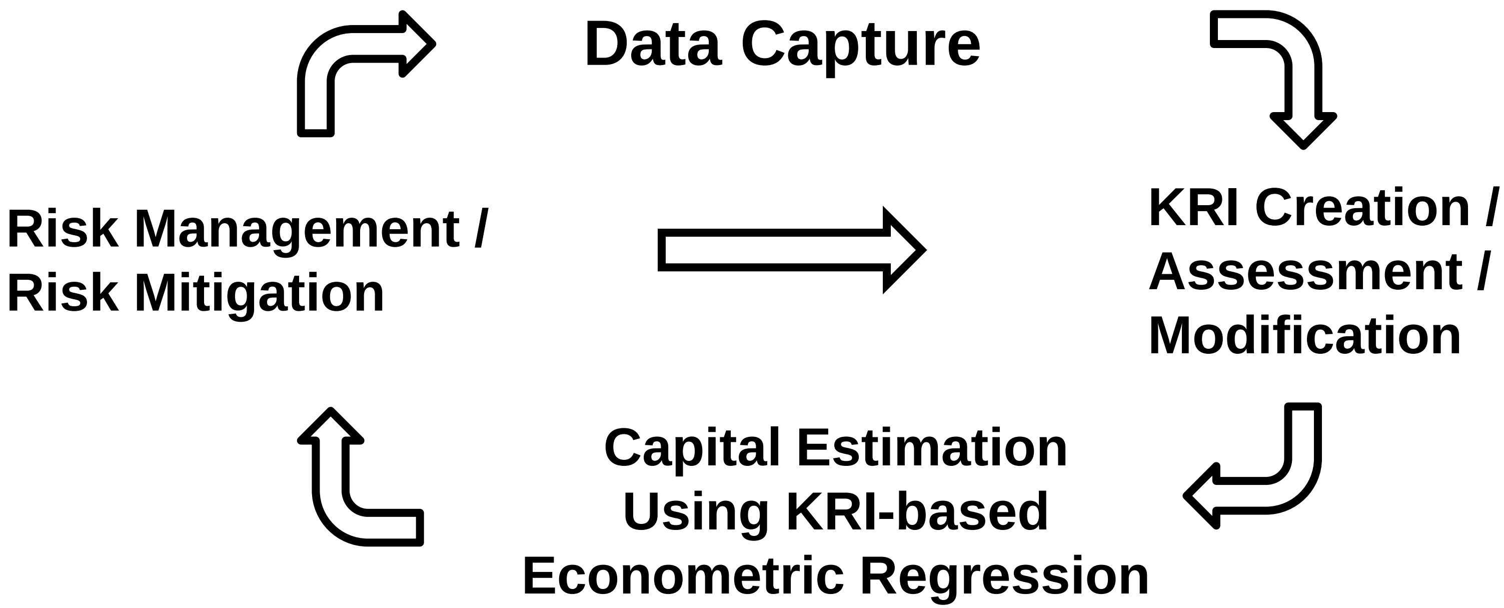
Proactively Selecting and Maintaining the Right KRIs:
Dynamic Feedback Loop

- **Data Capture**: Informed by subject matter expertise and technical modeling expertise (and previous results).
- **KRI Creation / Assessment / Modification**: Informed by SME and technical modeling expertise (and previous results).
- **Capital Estimation Using KRI-based Regression**: Informed by technical modeling expertise.
- **Risk Management / Mitigation**: If KRIs are well designed, results from capital modeling seemlessly dictate mitigation efforts.
- **New Data Capture Efforts**: The loop continues, modified and informed by results from the prior stages / loops.

I. OpRisk Management/Mitigation via KRI Regression

Econometric Methods for Establishing Direct, Material, Statistically Causal Relationships between KRIs and i. Capital, ii. Loss Frequency, and iii. Loss Severity COMPLETELY CONSISTENT AND COMPATIBLE WITH AMA/LDA:

Proactively Selecting and Maintaining the Right KRIs:
Dynamic Feedback Loop



I. Beyond "Red-Light, Green-Light" KRIs…

Econometric Methods for Establishing Direct, Material, Statistically Causal Relationships between KRIs and i. Capital, ii. Loss Frequency, and iii. Loss Severity COMPLETELY CONSISTENT AND COMPATIBLE WITH AMA/LDA:

Proactively Selecting and Maintaining the Right KRIs:
Dynamic Feedback Loop

This is **DYNAMIC**. Macro environments change, the firm's risk profile changes, and capital SHOULD change based on risk mitigation efforts guided by KRI-based capital modeling results. The KRIs that exhibit predictive power vis-à-vis capital will change over time (at least in part because the mitigation based on them was effective!), old will be discarded or modified, and new ones will be used, as long as the loop is actively maintained, owned, and fed current data and modeling expertise.

This is a labor intensive process, but the returns dwarf the required investment.

This also addresses a major criticism of operational risk capital models as generally backward-looking: KRI-based regression uses admittedly historical loss event data for a forward-looking purpose – to PREDICT and mitigate losses!

I. OpRisk Management/Mitigation via KRI Regression

Econometric Methods for Establishing Direct, Material, Statistically Causal Relationships between KRIs and i. Capital, ii. Loss Frequency, and iii. Loss Severity COMPLETELY CONSISTENT AND COMPATIBLE WITH AMA/LDA:

WHY Multivariate Regression? SUMMARY: ONLY this approach provides

1. **Statistically Causal Relationships between KRIs and Capital, AND KRIs and LOSS FREQUENCY AND SEVERITY (… NOT JUST CAPITAL!!!)**
2. **Magnitude of Effect of Each KRI** on i. Capital AND ii. LOSS FREQUENCY AND SEVERITY Independent of other KRIs
3. **RELATIVE IMPORTANCE of Each KRI's Effect** on i. Capital AND ii. LOSS FREQUENCY AND SEVERITY Independent of other KRIs (key for $allocation for mitigation efforts)
4. **Direction of Effect of Each KRI** on i. Capital AND ii. LOSS FREQUENCY AND SEVERITY Independent of other KRIs
5. **Whether Effect of Each KRI is Statistically Significant** vis-à-vis i. Capital AND ii. LOSS FREQUENCY AND SEVERITY Independent of other KRIs
6. **Whether Effect of Each KRI is Material** vis-à-vis i. Capital AND ii. LOSS FREQUENCY AND SEVERITY Independent of other KRIs
7. **Increase in the Precision of the Estimate of Capital** (AND Frequency and Severity), all else equal

I. Beyond "Red-Light, Green-Light" KRIs…

Econometric Methods for Establishing Direct, Material, Statistically Causal Relationships between KRIs and i. Capital, ii. Loss Frequency, and iii. Loss Severity COMPLETELY CONSISTENT AND COMPATIBLE WITH AMA/LDA:

1. **GAMLSS and related Regression Models are the only way to Estimate OpRisk Capital SCIENTIFICALLY, OBJECTIVELY, and INFERENTIALLY to directly and quantitatively and CAUSALLY LINK it to DRIVERS OF OPRISK, thus allowing responsible, targeted, and measurable mitigation and management. THIS is OpRisk's endgame!**

2. **This is NOT mathematical voodoo!! – related regressions are referenced in current, published guidance from the Federal Reserve Board (see BGFRS, 2013, p.28) and have been used in applied settings longer than OpRisk has been a risk type/discipline (sometimes much longer).**

3. **And this is NOT merely "aspirational"!!! No more effort/resources are required to define and obtain covariates AT THE LEVEL OF THE OPRISK LOSS EVENTS than are dedicated to creating "rolled up" aggregated KRIs USED IN RED-LIGHT, GREEN-LIGHT "analysis"**

I. OpRisk Management/Mitigation via KRI Regression

Econometric Methods for Establishing Direct, Material, Statistically Causal Relationships between KRIs and i. Capital, ii. Loss Frequency, and iii. Loss Severity COMPLETELY CONSISTENT AND COMPATIBLE WITH AMA/LDA:

4. **Yes, of course there are many challenges associated with effectively implementing GAMLSS models in this setting. But it HAS been done (see Cruz, 2012, and Shevchenko, 2014), and ANY non-trivial empirical analysis worth implementing will have challenges: that has not stopped us from tackling it for other risk types, so why should it stop us for OpRisk? It shouldn't, and doesn't, as long as we invest in OpRisk where we need to: i. our data (the right type, of the highest quality, and maximum quantity (i.e. maximum appropriate capture)); and ii. our quants (great data's worthless if we don't have the technical sophistication to maximize its value).**

5. **Finally, the ROI here dwarfs that of any other endeavors related to Operational Risk Measurement or Management because it most effectively accomplishes and integrates BOTH! How can we effectively MANAGE and MITIGATE if we cannot scientifically, objectively, and accurately MEASURE the direction, magnitude, and causality of OpRisk Drivers? This can only be done with regressions.**

I. Beyond "Red-Light, Green-Light" KRIs…

Econometric Methods for Establishing Direct, Material, Statistically Causal Relationships between KRIs and i. Capital, ii. Loss Frequency, and iii. Loss Severity COMPLETELY CONSISTENT AND COMPATIBLE WITH AMA/LDA:

Oft-heard "responses":

1. This regression stuff is just "aspirational"
2. That may be fine for credit and market risk, but that's way too hard to do for OpRisk
3. CRO's are "comfortable" with Red-Light, Green-Light
4. CRO's wouldn't understand this complicated regression stuff
5. Its too hard to collect that data
6. Our rolled-up KRI's are not correlated at all… we know, we checked them with the =CORREL function in EXCEL! So we don't need regression here…
7. No way, we'd have to adjust our data collection software … the vendor says adding a few (KRI) columns of data for each OpRisk loss event will cost $25m
8. We've been aggregating KRIs for years like everybody else… that can't be wrong!
9. Well, maybe you have a point, but its too late to do it the right way now.
10. [MY FAVORITE…] "Our job is a higher calling then just estimating capital."

Contents

II. Eliminating Upward Approximation Bias in Capital

- **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 a closed-form formula) as the fastest and most computationally efficient. SLA is shown below under three tail index conditions (not that only a. is relevant for severities that cannot have infinite mean):**

OpRisk NorthAmerica

II. Eliminating Upward Approximation Bias in Capital

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

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

where $c_\xi = (1-\xi)\frac{\Gamma^2(1-1/\xi)}{2\Gamma(1-2/\xi)}$

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

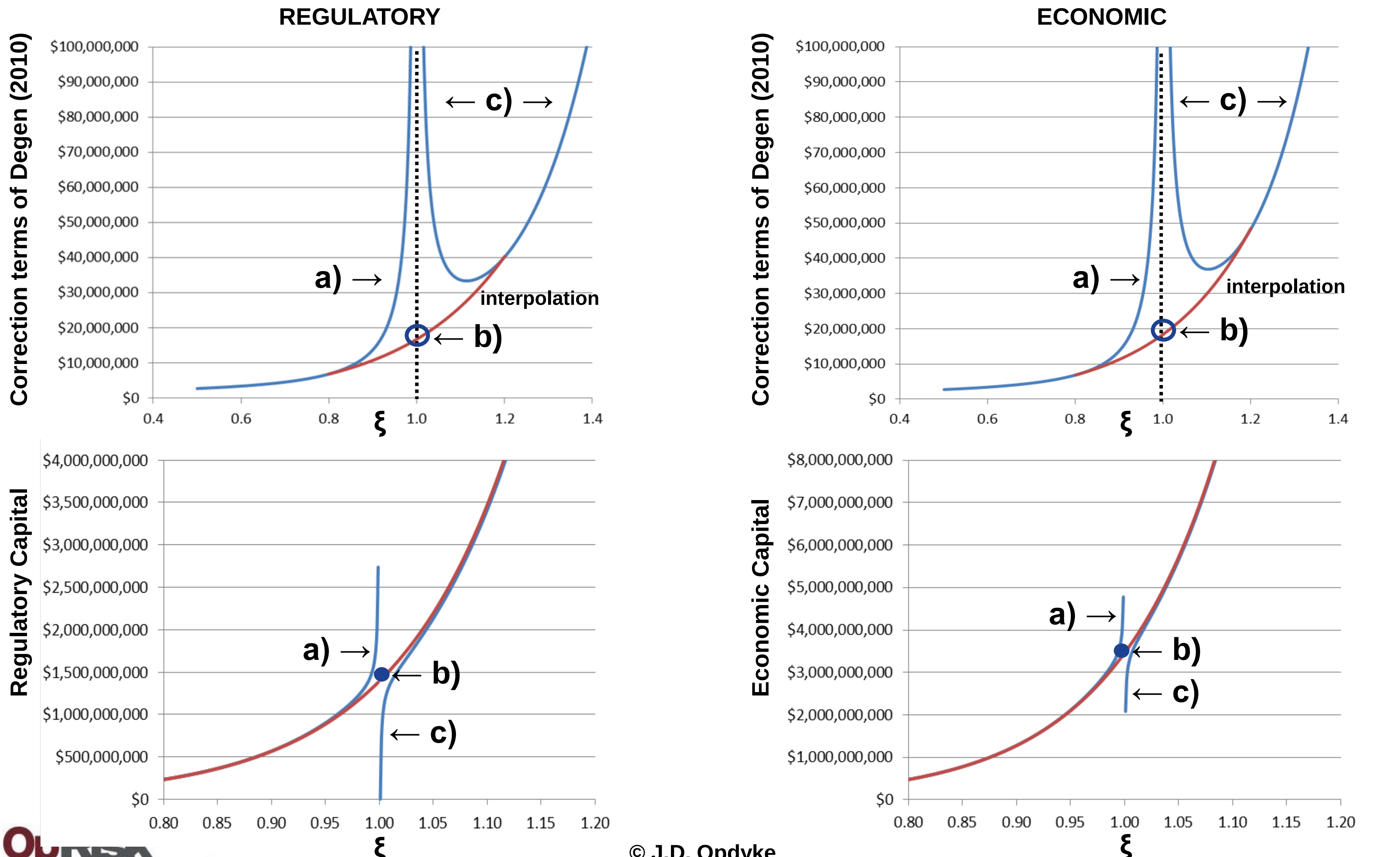
($\xi \geq 2$ is so extreme as to not be relevant in this setting)

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

OpRisk NorthAmerica

II. Eliminating Upward Approximation Bias in Capital

Figures A1-A4: **ISLA2** **Correction for SLA Divergence at Root of ξ=1 for GPD Severity (θ = 55,000)**



II. Eliminating Upward Approximation Bias in Capital

- **When implementing the above it is critical to note that the capital estimate diverges as $\xi = 1$; 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 $\xi \to 1$ AND DOES NOT DISAPPEAR ASYMPTOTICALLY. For example, for GPD, divergence can be noticeable in the range of $0.8 < \xi < 1.2$. Therefore, one must utilize a an alternative derivation of Degen's formulae to avoid this obstacle.**

- **Opdyke and Mayorov (forthcoming, 2016) show that for $\xi = 0.99$, upward capital bias is on the order of magnitude of tens of millions of USD, and that even when the tail index does not approach $\xi \to 1$, for example, $\xi = 0.85$, which is not at all uncommon, SLA's upward bias in as few samples as 1,000 when running simulations is often orders of magnitude larger than true capital. In other words, it does not take many samples to hit one that is yields parameter estimates with a tail index arbitrarily close to a value of one, making SLA's use in any kind of random sampling or simulations extremely unreliable.**

OpRisk NorthAmerica

II. Eliminating Upward Approximation Bias in Capital

- To solve this problem, because SLA is so widely used, Opdyke (2014) developed a straightforward, fast, and accurate non-linear interpolation ("ISLA," Interpolated SLA), which was generalized and made even more accurate by Opdyke and Mayorov (forthcoming, 2016) in "ISLA2." The ISLA2 completely eliminates SLA's systematically upward capital bias and provides an extremely fast and accurate approximation.
- The perturbative approach of Hernandez et al. (2012) appears to be the only other method that is simultaneously comparably fast and accurate, although for banks/sifi's already using SLA, ISLA2 arguably would be preferred.

OpRisk NorthAmerica

Contents

III. Mitigating Upward Estimation Bias in Capital

1. **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 $\Longrightarrow$ Convexity of ALD VaR $\Longrightarrow$ 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. RCE More Accurate: MLE Capital Bias can be ENORMOUS ($Billion+ for one unit of measure!)
   b. RCE More Precise: RCE RMSE < MLE RMSE, RCE StdDev < MLE StdDev, RCE IQR < MLE IQR
   c. RCE More Robust: RCE Robustness to Violations of iid > MLE
7. **Alternate Estimators**
8. **Summary and Conclusions**

OpRisk NorthAmerica

1. Goals and OpRisk Setting

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.**

OpRisk NorthAmerica

1. Goals and OpRisk Setting

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 severity distributions is ignored. See Chernobai, Rachev, and Fabozzi (2007) and Ergashev (2008).

2. AMA–LDA OpRisk Capital Defined

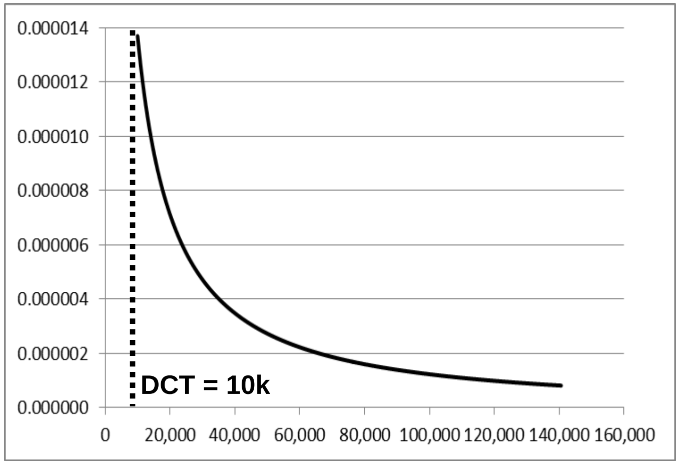
- A la Basel II/III, Operational Risk Capital for large banks/SIFIs must be estimated with an Advanced Measurement Approaches (AMA) framework.
- In writing, AMA provides great flexibility, but in practice, there has been industry convergence to the Loss Distribution Approach (LDA).
- Under LDA, severity 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).

OpRisk NorthAmerica

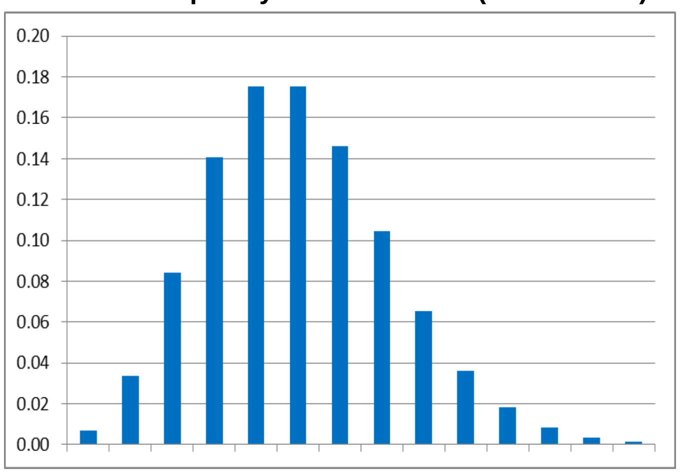
2. AMA–LDA OpRisk Capital Defined

Loss Distribution Approach – For a given UoM:

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

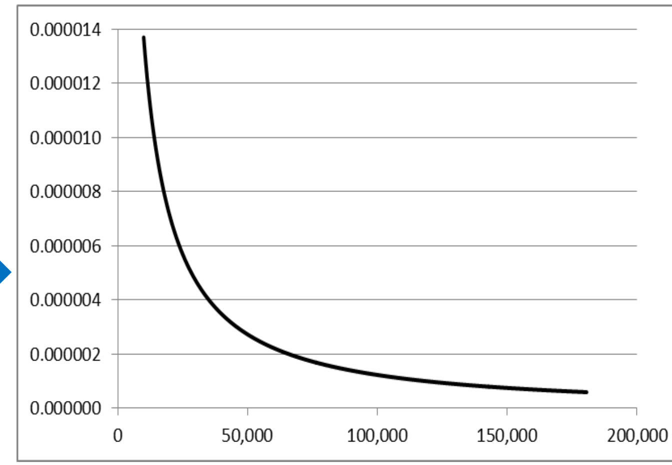


Estimated Frequency PMF – Poisson (annual λ=25)



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

OpRisk NorthAmerica

3.a. Apparent Convexity of Severity VaR ⇒ Inflated Capital

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

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

In other words, first term >> second term (see Section 2. for improved approximations of Opdyke, 2014 (ISLA), Hernandez et al., 2012, and Opdyke and Mayorov, 2016 (ISLA2)).

- **PROPOSED: 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 sufficiently convex function of severity parameter vector $\hat{\beta}$ for Jensen's inequality to hold
  THEN ALD VaR (Capital) is a sufficiently convex function of $\hat{\beta}$ for Jensen's inequality to hold.

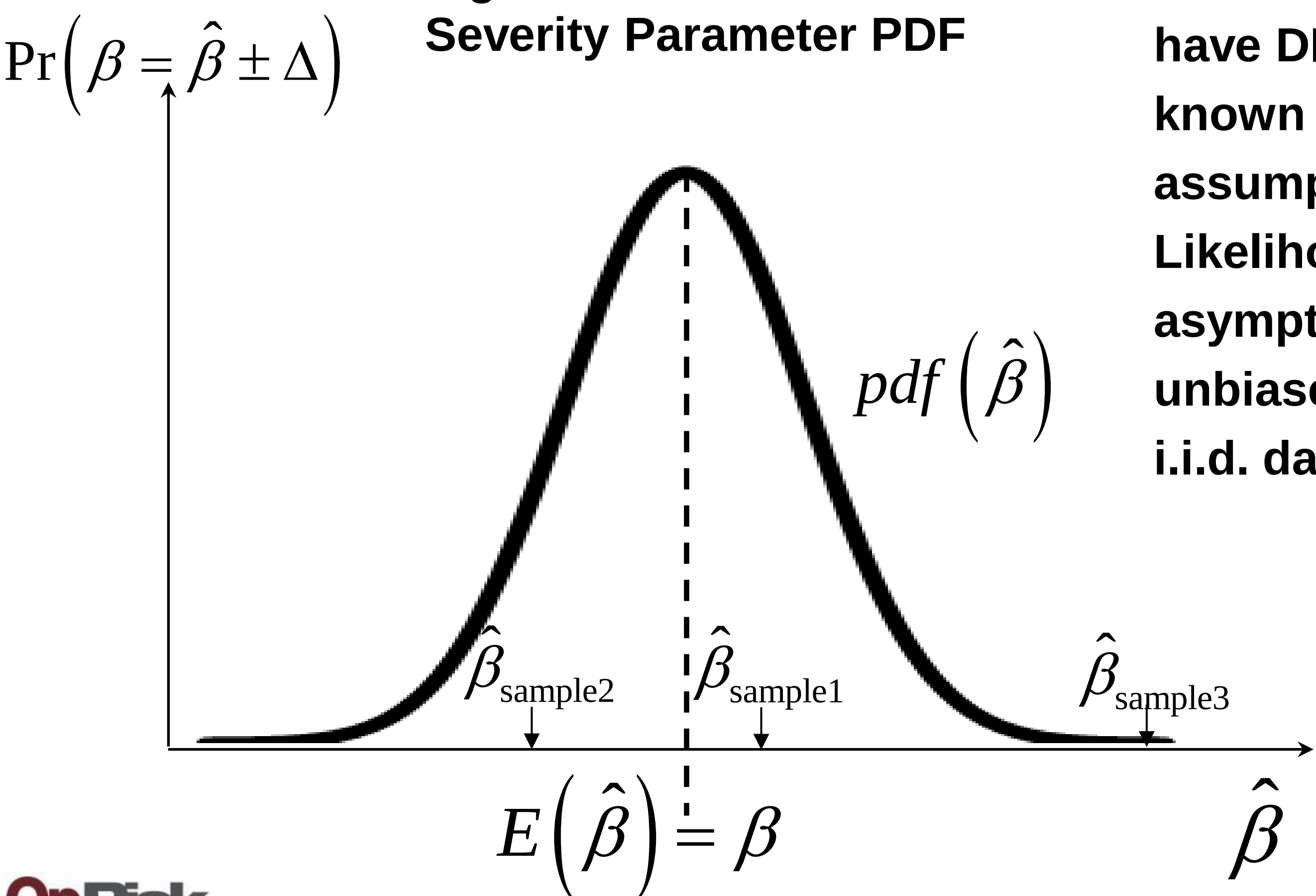
- **NOTE: Severity VaR is much more extreme than ALD VaR, because for, say, *λ* = *30*, and *α* = 0.999 and *α* = 0.9997, $p=\left[1-\left(1-\alpha\right)/\lambda\right]=$ 0.999967 and 0.99999, respectively.**

OpRisk NorthAmerica

3.a. Apparent Convexity of Severity VaR ⇒ Inflated Capital

- **Operational Risk Loss Event Data = a Sample, NOT a Population**
- **Therefore, true severity parameters, β , will never be known.**

Figure 1:
Severity Parameter PDF



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)

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

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

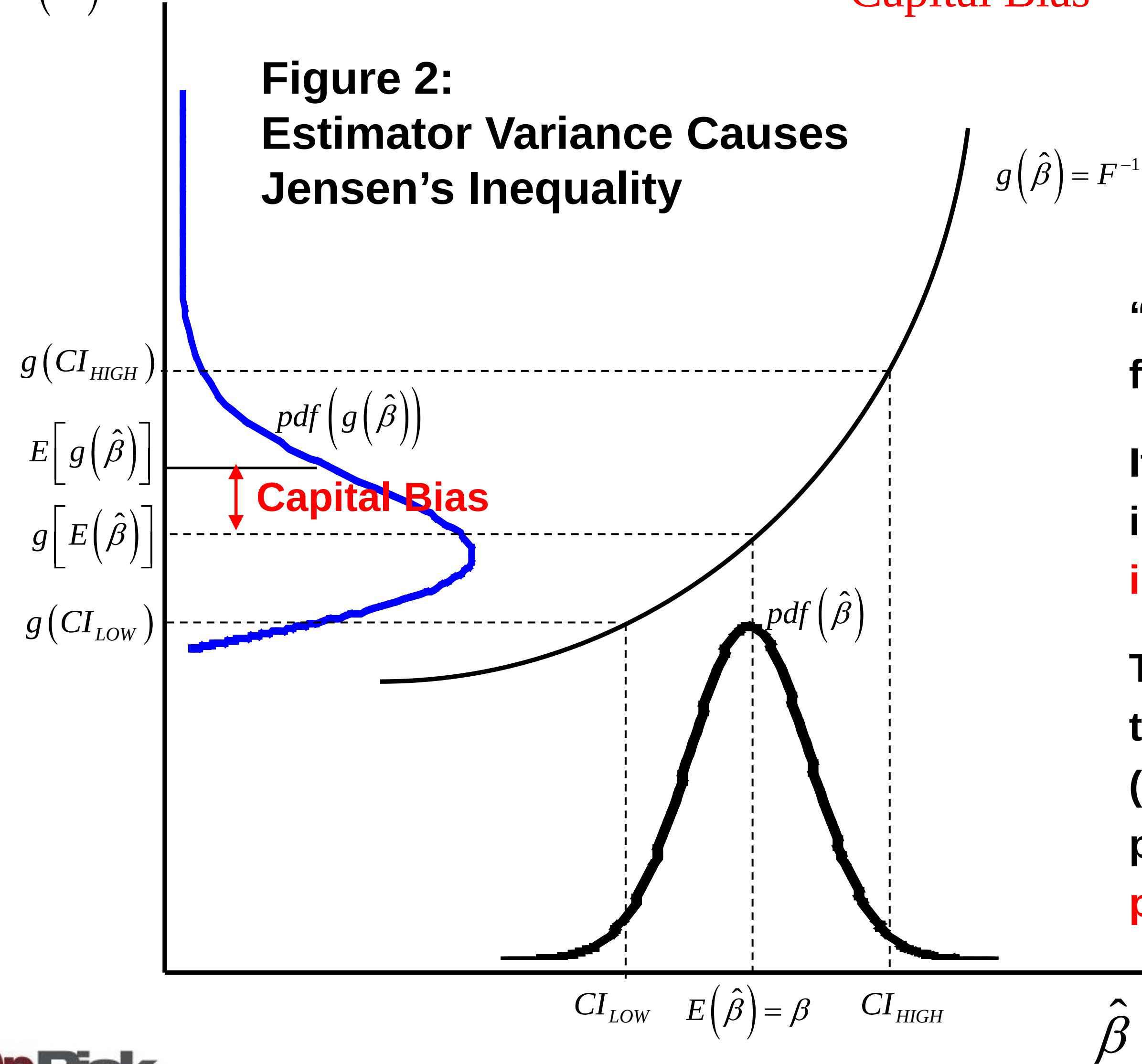


Figure 2:
Estimator Variance Causes Jensen's Inequality

"Jensen's inequality" first proved in 1906.

If VaR, g(), is strictly 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.**

OpRisk NorthAmerica

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

- **Of course, this is convexity with respect to estimated severity parameters. This is explicitly stated in Opdyke and Cavallo (2012a) on p.68, and again in Opdyke (2014) on p.12, respectively, as below:**

- **"This is illustrated in Figure 20 (from Kennedy (1992, p. 37)). This applies to quantile estimation of all commonly used severity distributions: if *β* is a random variable (here, our severity distribution parameter estimates) and $g(\cdot)$ is a (strictly) convex function (here, the inverse of our severity distribution CDF), then $g\left(E\left[\hat{\beta}\right]\right) < E\left[g\left(\hat{\beta}\right)\right]$, and our quantile estimate (capital estimate) is biased upward."**

- **"under these conditions, VaR appears to always be a convex function, like $g(\cdot)$, of the parameters of the severity distribution, which here is the vector *β* (we can visualize *β* as a single parameter without loss of generality as the multivariate case for Jensen's inequality is well established (see Schaefer 1976)). Consequently, the capital estimation, $\hat{V} = g\left(\hat{\beta}\right)$ will be biased upward."**

OpRisk NorthAmerica

3.a. Apparent Convexity of Severity VaR ⇒Inflated Capital

- **Unfortunately, there is a little confusion on this point in an unpublished paper (see Larsen, 2015):**
- **"This mean bias is a central object of study in Opdyke and Cavallo (2012), where they claim that MLE results in capital overestimation. The meaning of this statistic for modeling decisions, however, is not completely clear. … Opdyke and Cavallo (2012) write that the mean OpVaR bias is a consequence of Jensen's inequality, but no further details are given. This would follow if the CDF $F(x|\theta)$ for a heavy-tailed distribution were a convex function. There is no mention whether convexity is with respect to the loss variable x or with respect to the parameters *θ*. For the Jensen's inequality argument of Opdyke and Cavallo (2012) to be valid, convexity must be shown with respect to the parameters *θ*, not the loss amount x.[fn3] Specifically, we would have to show that, for all loss amounts x in a neighborhood of the true OpVaR, the Hessian of $F(x|\theta)$ with respect to *θ* is negative definite (and hence the Hessian of the quantile function of $F(x|\theta)$ would be positive definite). This property is trivial to verify for the Pareto distribution considered here as depending only on one variable, but is less than straightforward for more complicated distributions. That there is still something to prove before invoking Jensen's inequality is mentioned in a subsequent paper (Opdyke, 2014)."**
- **In footnote 3 Larsen (2015) examines potential convexity of VaR with respect to "x," the variable representing the size of the loss events. But these are not being ESTIMATED – they are the data points themselves! Jensen's inequality is fundamentally about ESTIMATION, not data per se, so the point of the footnote is unclear, if not misguided. We encourage (re)reading Opdyke and Cavallo (2012a) and Opdyke (2014) above to avoid any confusion regarding the relevance Jensen's inequality in this setting. Finally, Mayorov, Opdyke, and Balakrishnan (forthcoming, 2016) ANALYTICALLY demonstrate that examining the positive vs. negative definiteness of the Hessian alone is not enough to verify VaR's local convexity here, and they establish more rigorous conditions for this to hold.**
- **The 2nd confusion in Larsen (2015), this time regarding bias, is addressed below.**

3.a. Apparent Convexity of Severity VaR ⇒Inflated Capital

- It is critical to note here that even though capital estimates will be, on average, high 50% of the time and low 50% of the time even under Jensen's inequality, the AMOUNTS that they are high vs. low are very different: **when high, they are often much higher than true capital, but when low, they often are not much lower than true capital. Would you/your bank bet on a nickel gain vs. a dollar loss with equal probability?!**

- When comparing capital estimates to true capital, **probability alone is not sufficient here – the absolute DISTANCE from true capital matters too**. And it is the mean (expected value), rather than specific quantiles like the median, that is determined by BOTH the probability, AND the absolute distance from true capital, associated with specific capital estimates.

- The capital estimate distribution, and all of its relevant characteristics, are examined throughout this presentation. The specific issue of the distance of true capital from specific quantiles of the distribution (e.g. the median) is examined in great detail in Appendix C herein, as well as in footnote 67, p.59, of Opdyke (2014), where it is shown that **so-called "median bias" is an essentially irrelevant artifice in this setting**.

OpRisk NorthAmerica

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 A), and the very strong empirical evidence is exactly consistent with the effects of convexity in that we observe Jensen's Inequality empirically.**

- **Still, we would like to PROVE Jensen's inequality for a) Severity VaR under these conditions, and b) Severity VaR for all relevant severities [a) and b) would be proven asymptotically: ultimately we would like to prove Jensen's inequality for c) arbitrary finite sample size.]**

OpRisk NorthAmerica

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.**

- **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,

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

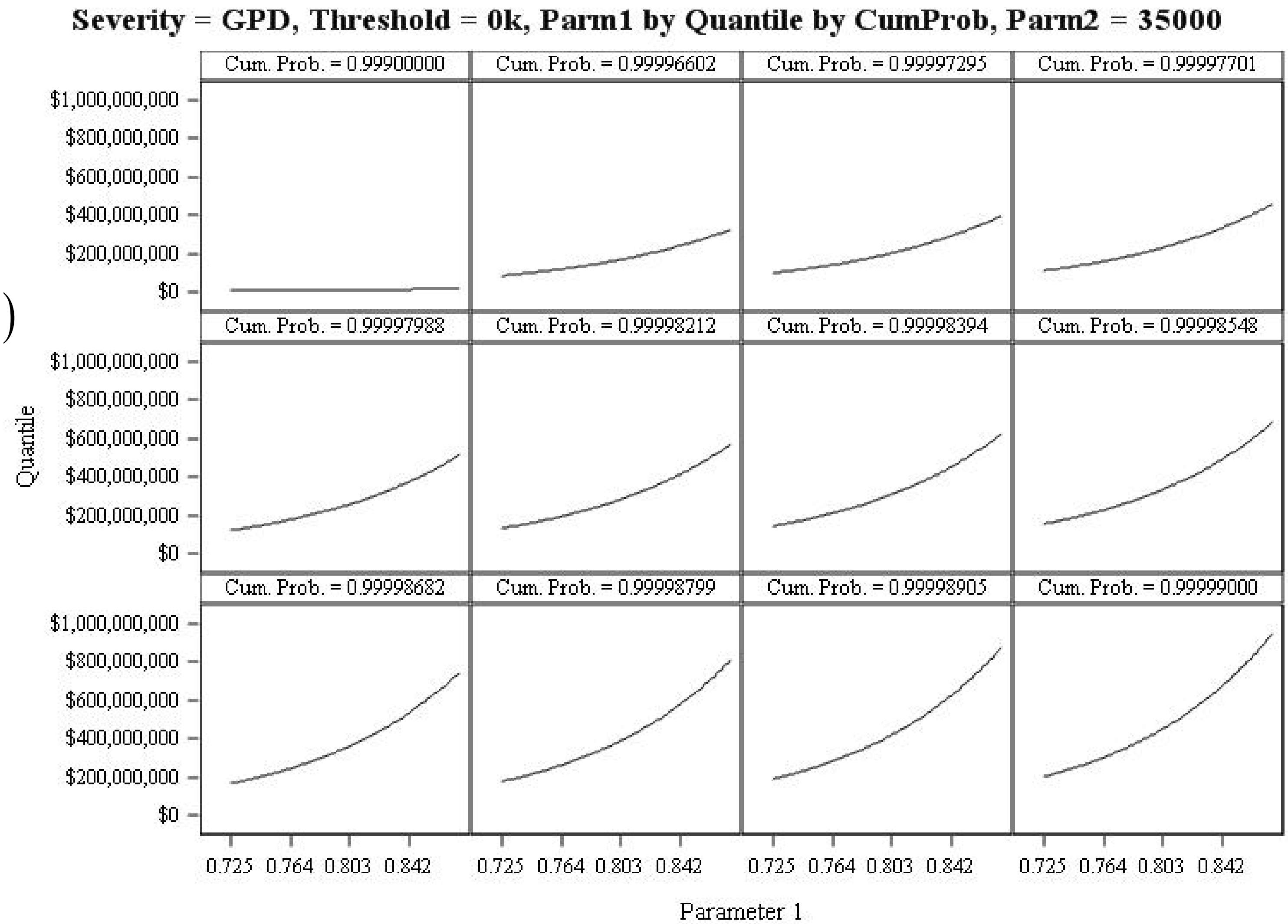
$$\partial^2 VaR / \partial\mu^2 = VaR > 0$$

$$\partial^2 VaR / \partial\sigma^2 = VaR \cdot \left[\Phi^{-1}\left(p\right)\right]^2 > 0$$

  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).

OpRisk NorthAmerica

Figure 3a:

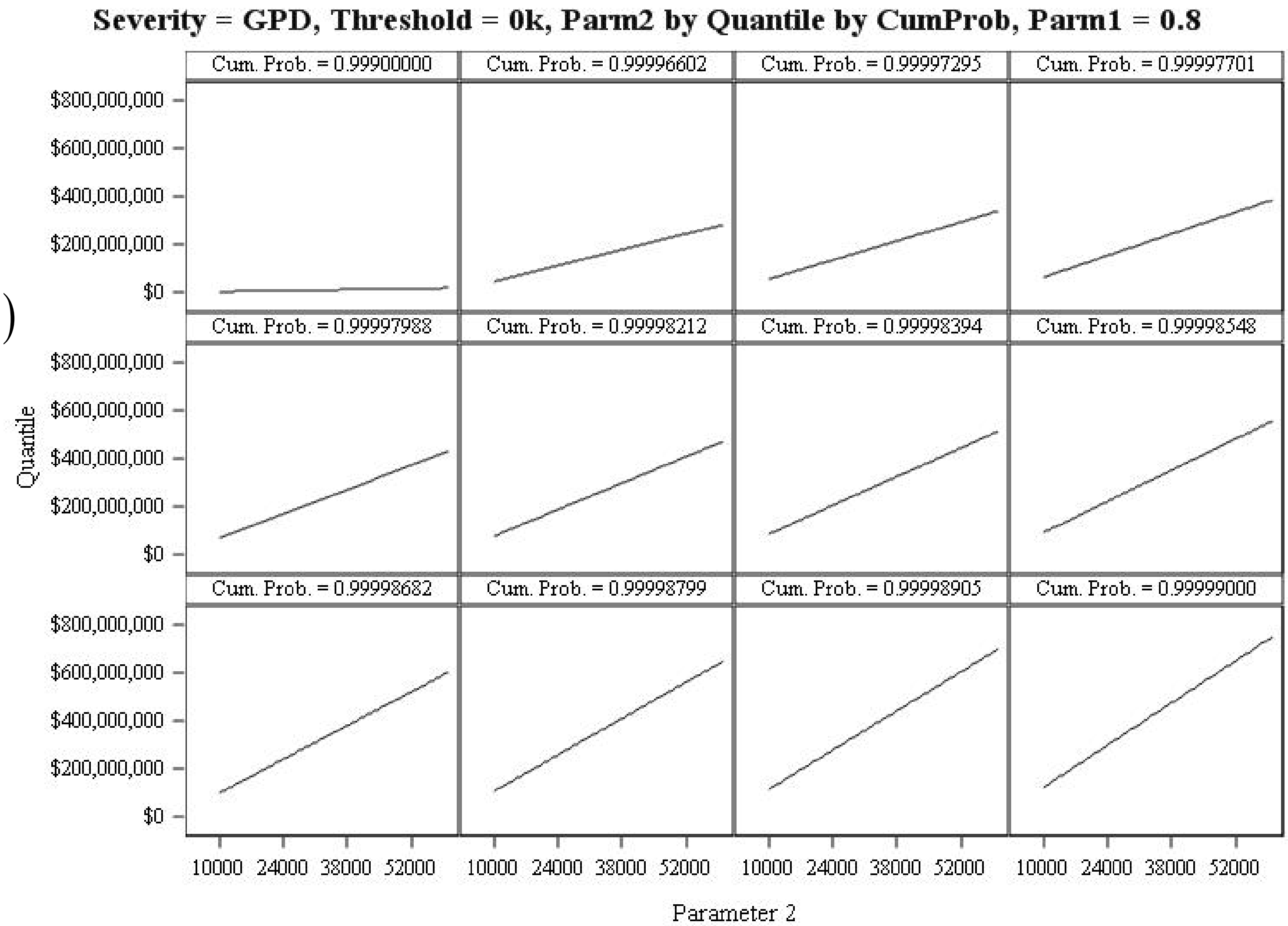


$F^{-1}(p;\xi)$

3.b. Multiple Checks for Convexity of Severity VaR

Figure 3b:

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



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.

***NOTE*:** Although in the past spliced and mixed-distribution severities have been used by a number of banks, the most recent Interagency Guidance (June, 2014) indicated strong preference for single-density severity estimation with fewer parameters, both to avoid potential for overfitting the loss event data. Specifically, the LogNormal, LogGamma, GPD, and Burr Type XII severities were mentioned.

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).

3.b. Multiple Checks for Convexity of Severity VaR

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

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

In other words, the cumulative weight of the evidence – even in the absence of a "smoking-gun" absolute mathematical proof – is very strong here. An apt analogy is the relationship between smoking and cancer: no one study definitively proves the now-known and widely accepted relationship between the two – it was the weight of cumulative evidence from disparate sources that eventually became accepted wisdom and scientific fact.

All strong and consistent evidence here points to Jensen's Inequality as the source of bias, so we should not delay in allowing this assumption to guide the design of solutions to it.

It is also crucial to note that a strictly convex VaR surface is not necessary for Jensen's inequality to be true, and this is a widely proven result: the surface need only be sufficiently convex.

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).

3.b. Multiple Checks for Convexity of Severity VaR

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).

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)

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 sufficient VaR convexity for Jensen's Inequality to hold, and design a capital estimator accordingly to mitigate the actual capital bias/inflation of which it is the presumed source…

OpRisk NorthAmerica

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.**

- **As noted above, Mayorov, Opdyke, and Balakrishnan (forthcoming, 2016) establish strong ANALYTICAL support for VaR's local convexity here.**

- **But for the time being 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.

4. When is Capital Bias (Inflation) Material?

Convexity in Severity VaR $\Rightarrow$ Capital Bias is upwards*

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, AMA–LDA OpRisk capital estimation is squarely in the bias zone!

*Again, convexity applies to the pool of relevant severities here, which are medium- to heavy-tailed.

OpRisk NorthAmerica

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 (arguably contrived) 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.

OpRisk NorthAmerica

5. RCE – Reduced-bias Capital Estimator

RCE (Reduced-bias Capital Estimator) is the only published 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.

RCE Conceptually Defined:

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

Step 2: Using 1), simulate *K* iid data samples and estimate parameters of each

Step 3: 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

Step 4:
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).

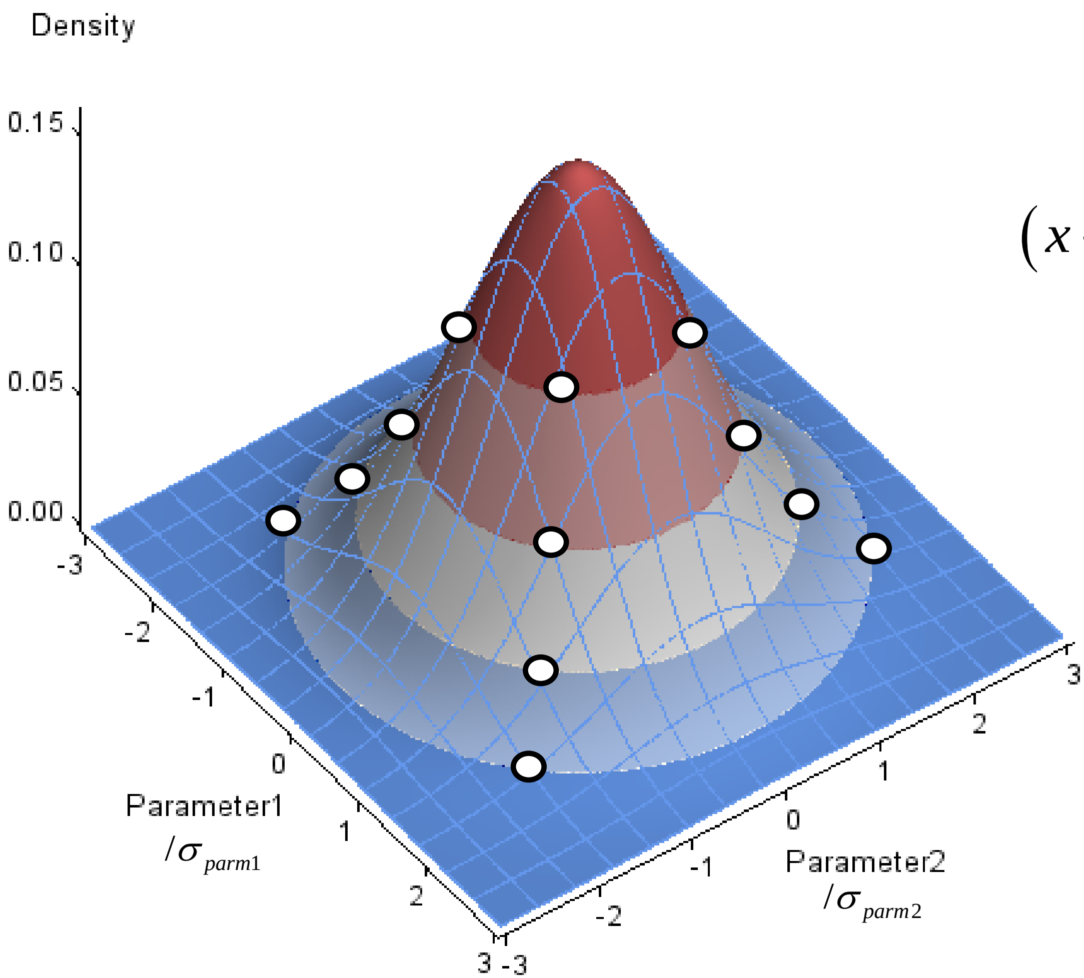
Step 2: Using 1), generate *K* parameter vectors based on the Var-Cov matrix using iso-density sampling (see Figure 4 below): use iso-density ellipses to select parameter values associated with a given probability, and change parameter values to reach these ellipses via the decrease-decrease, decrease-increase, 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*

OpRisk NorthAmerica

FIGURE 4: Iso-density Perturbation of the Joint Severity Parameter Distribution



For multivariate normal (e.g. all M-class estimators), ellipses are given by:

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

where *x* is a *k*- (2-) dimensional vector, *μ* is the known *k* -dimensional mean vector (the parameter estimates), ∑ is the known covariance matrix (the inverse of the Fisher information of the given severity), and $\chi_k^2(p)$ 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 may 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 B.

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 Step 2 of RCE's implementation; and then the 99%tile of THIS Fisher information is obtained in Step 3, based on the joint parameter distribution of the Step 2 values. Capital calculated in Step 3 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.

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

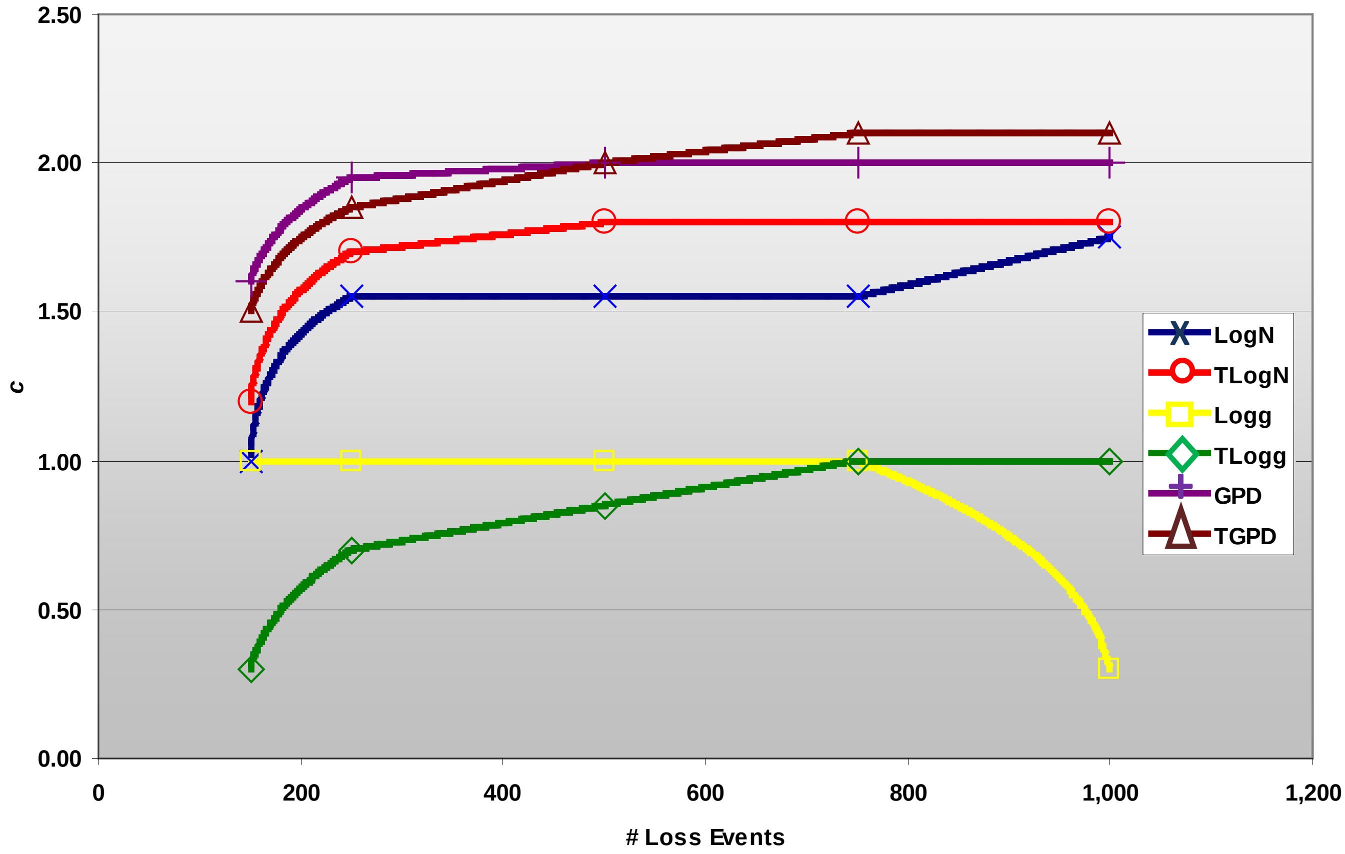
Method 1: Conduct a simulation study to empirically determine the value of *c* for the relevant sample sizes and severities (both known ex ante within the LDA framework) using three sets of parameter values: the original estimates, and those corresponding to the 2.5%tile and the 97.5%tile of the joint 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).

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: 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).

OpRisk NorthAmerica

6. Simulation Study: RCE vs. MLE

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

- **λ = 25 (Poisson-distributed average annual losses … so n = 250, on average, over 10 years)**
- **α = 0.999 and 0.9997 for Regulatory and Economic Capital, respectively (so [1 – (1- α) / λ] = 0.99996 and 0.999988, respectively).**

Selected Results of RCE capital vs. MLE capital:

- LogNormal
- LogGamma
- GPD
- Truncated LogNormal
- Truncated LogGamma
- Truncated GPD

***Note** that true bias is probably far greater than that associated with MLE-based capital below, since under the i.i.d. presumption MLE is maximally efficient.**

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

OpRisk NorthAmerica

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

OpRisk NorthAmerica

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

** $\lambda = 25$; $\alpha = 0.999$ RC; $\alpha = 0.9997$ EC

6. Simulation Study: RCE vs. MLE

Table 4e:
RCE vs. LDA-MLE for GPD Severity (ξ = 0.875, θ = 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

OpRisk NorthAmerica

Table 4f:
RCE vs. LDA-MLE for Truncated GPD Severity (ξ = 0.8675, θ = 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

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) <u>RCE is Dramatically More Accurate</u>: LDA-MLE Bias can be ENORMOUS: $Billion+ just for one uom!

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

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

7. Alternate Estimators

1. An alternate form of RCE is to simply use estimated capital as the first term, and then scale it based on the perturbation of its frequency and severity parameters:

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

Modified RCE:
MRCE = estimated capital * [median(K medians) / weighted mean(*K* medians)]^*c*.

This approach has the advantage of simply being a scalar of existing capital, but requires re-estimation of the values of "*c*" for some combinations of severity distribution + sample size. However, with respect to the variance of capital estimate, RCE maintains the distinct advantage (i.e. RCE decreases it).

2. A non-published paper by Zhou, Durfee, and Fabozzi (2015) presents a modification of the RCE approach. Curiously, even though Zhou et al. (2015) follows Opdyke (2014), in both timing and methodology, changes made to the RCE estimator appear to worsen not only its performance in terms of bias, speed of execution, and stability, but also increase its likelihood of regulatory rejection due to its reliance on "trimming" (which RCE avoids). See Appendix C for further details.

8. 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.**

Contents

IV. Summary and Conclusions

I. **Regression-Based (AMA/LDA) Capital Estimation using "granular" KRIs is the only way to scientifically, objectively, and accurately manage and mitigate operational risk using statistically causal risk drivers.** "Red-Light , Green-Light" Aggregated "Rolled Up" KRIs cannot do this, and worse, damage estimation and management with i) misleading inferences that fail to take into account covariance between risk drivers (we live in a multivariate world!), and ii) a false sense of security (which increases risk!) based on little more than gut feels and thoughtful guesswork.

II. **SLA-based Capital Approximations are systematically upwardly biased.** ISLA2 (Opdyke and Mayorov, 2016) eliminates this bias, remains straightforward, accurate, and fast, and for banks/sifi's already using SLA, is readily implemented.

III. **Jensen's Inequality systematically, upwardly biases AMA/LDA Capital ESTIMATION, and the magnitude of this bias can be enormous (e.g. beyond $1b for a single UoM; see Opdyke, 2014, and Mayorov, Opdyke, and Balakrishnan, 2016).** The Reduced-bias Capital Estimator (RCE) of Opdyke (2014) dramatically mitigates this bias, while simultaneously increasing capital precision and robustness. RCE is completely compatible with and consistent with Regression-Based AMA/LDA Capital Estimation.

IV. Summary and Conclusions

"Measurement is the first step that leads to control and eventually to improvement. If you can't measure something, you can't understand it. If you can't understand it, you can't control it. If you can't control it, you can't improve it."

- **H.J. Harrington**

Measurement may not be everything, but without reasonably accurate, precise, and robust measurement, its very hard to argue that you can do OpRisk management and mitigation right (or even in a way that doesn't do more harm than good).

Contents

V. References

- 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.
- Chapelle, A., (2012), "Selecting Good KRIs," PRMIA – Publication Weblogs, August 31, 2012.
- Chapelle, A., (2013), "Four Candidates for Preventive KRIs," PRMIA – Publication Weblogs, March 19, 2013.
- 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.
- Cruz, Marcello, (2012), "Latest Techniques in Modeling OR," presentation at American Bankers Association Operational Risk Forum, July 19, 2012, Washington,DC.
- 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.
- Hernandez, L., Tejero, J., Suarez, A., and Carrillo-Menedez, S., (2013), "Closed-form Approximations for Operational Value-at-Risk," *Journal of Operational Risk*, 8(4), 39-54.
- Interagency Guidance: Supervisory Guidance for Data, Modeling, and Model Risk Management Under the Operational Risk Advanced Measurement Approaches, Board of Governors of the Federal Reserve System, June 30, 2014.
- 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.
- Larsen, P. (2015), unpublished paper, "Operational Risk Models and Maximum Likelihood Estimation Error for Small Sample-Sizes"
- Mayorov, K., (2014), Email correspondence with the author, September, 2014.

V. References

- Mayorov, K., Opdyke, J.D., and Balakrishnan, N., (2016, forthcoming), "Jensen's Inequality, Extreme Quantiles of Compound Loss Distributions, and Laplace Approximations," *forthcoming*, 2016.
- 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 Operational Risk: New Frontiers Explored, ed. E. Davis, Risk Books, Incisive Media, London.
- Opdyke, J.D. (2014), "Estimating Operational Risk Capital with Greater Accuracy, Precision, and Robustness," *The Journal of Operational Risk*, Vol. 9, No. 4, December.
- Opdyke, J.D., and Mayorov, K. (2016, forthcoming), "Fast, Accurate, Straightforward Extreme Quantiles of Compound Loss Distributions," *forthcoming*, 2016.
- Panjer, H. (2006), Operational Risk Modeling Analytics, Wiley Series in Probability and Statistics, John Wiley & Sons, Inc., Hoboken, NJ.
- Rigby RA, Stasinopoulos DM (2001). "The GAMLSS project: a Flexible Approach to Statistical Modelling." In B Klein, L Korsholm (eds.), New Trends in Statistical Modelling: Proceedings of the 16th International Workshop on Statistical Modelling, pp. 249-256. Odense, Denmark.
- Roehr, A., (2002), "Modelling Operational Losses", *Algo Research Quarterly*, Vol.5, No.2, Summer 2002.
- Shevchenko, P. (2014), "Loss Distribution Approach for Operational Risk: Australian Experience," OpRisk North America, Risk conference, New York 26-27 March 2014.
- 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.
- Zhou, X., Durfee, A., and Fabozzi, F., "On Stability of Operational Risk Estimates by LDA: From Causes to Approaches," Non-published working paper, 2015.

V. Appendix A: VaR's Empirical Convexity Over Relevant Domain (p>0.999)

- **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**
	Parameter 1	Parameter 2	Parameter 3	**Parameters**
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.

V. Appendix B: Severity PDFs, CDFs, & Means for Capital Approximation

- **PDF and CDF of LogNormal:**

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

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

- **Inverse Fisher information of LogNormal:**

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

OpRisk NorthAmerica

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

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

$H < x < \infty,\ 0 < \sigma < \infty$

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

- **Mean of Truncated LogNormal:**

$$E(X)=e^{\mu+\sigma^2/2}\cdot\Phi\left(\frac{\mu+\sigma^2-\ln(H)}{\sigma}\right)\cdot\frac{1}{\left[1-F(H)\right]}$$

where $\Phi(\ )$ is the standard normal CDF.

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

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

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

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.

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

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

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

- **Inverse Fisher information of GPD:**

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

From Smith (1987)

OpRisk NorthAmerica

- **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:** $E(X)=\frac{\theta}{\xi}\cdot\left(\frac{\left[1-F(H)\right]^{-\xi}}{1-\xi}-1\right)$ for $\xi<1$ $(=\infty$ for $\xi \geq 1)$

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

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

From Roehr (2002)

OpRisk NorthAmerica

- **PDF and CDF of LogGamma*:**

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

$1\le x<\infty;\ 0<a;\ 0<b$

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

- **Mean of LogGamma:** $E(X)=\left(\frac{b}{b-1}\right)^a$ 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.**

- **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:**

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

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

From Opdyke (2014)

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

V. Appendix B: Severity PDFs, CDFs, & Means for Capital Approximation

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

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

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

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

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

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 $\text{trigamma}(z) = \partial^2/\partial z^2 \ln\left[\Gamma(z)\right]$.

OpRisk NorthAmerica

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

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

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

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

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

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

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

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

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

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

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

OpRisk NorthAmerica

V. Appendix C: Rejection of "Trimming" Methods

In a non-published paper, Zhou et al. (2015) present a modification of RCE. The approach follow's Opdyke (2014) in both timing and methodology by using a median/mean ratio of estimated capital combined with an adjustment factor.

Adjusted capital = capital * [median of simulated capital / mean of simulated capital]

Unfortunately, in attempting to compensate for greater instability due to its reliance on simple parameter simulation (as opposed to a far more stable approach based on the median-of-median of parameter estimates), their adjustment factor relies on data "trimming." Estimation methods like "trimming" that rely on systematically discarding a percentage of observed loss data (or simulated data based on parameter estimates which are based on observed loss data) have not been well received by regulators. In addition, the more simple approach of Zhou et al. (2015) approach has the following disadvantages relative to RCE:

1. **It appears to be far less stable than RCE, which is designed specifically to avoid these instability issues (see above)**
2. **It is tested far less extensively on fewer severities**
3. **It appears to have greater capital bias compared to RCE, and the authors state that further "tuning" of the amount of "trimming" required is needed for its application to additional severities**
4. **Its execution time is slower, sometimes by orders of magnitude (RCE typically is implemented within one or two seconds)**
5. **The authors themselves conclude that their alternate method provides " 'limited' improvement" and is not sufficient to use within a loss distribution approach for operational risk capital estimation.**

V. Appendix C: Rejection of "Trimming" Methods

In addition, Zhou et al.'s (2015) focus on so-called "median bias" is at odds with their own estimator, the statistical literature, and the primary goals of the operational risk capital estimation setting.

1) For nearly a century, statistical "bias" has been defined with respect to the mean of an estimator, not one of its quantiles (such as the median).

2) To the extent that researchers would like to design an estimator centered on a particular quantile (such as a median), the (highly) skewed nature of the operational risk capital distribution (under the loss distribution approach) means that the capital estimator cannot be unbiased simultaneously with respect to both the mean and the median. Zhou et al. (2015) acknowledge this, but then proceed to follow Opdyke (2014) and attempt to design a capital estimator (actually, to modify RCE) in a manner that is "unbiased" in the traditional sense (i.e. vis-à-vis the mean) while ignoring so-called "median bias".

3) Exploring the possibility of estimators that are unbiased with respect to a particular quantile is arguably the wrong approach here. Far more relevant is the question of how close to ALL estimator quantiles is the true value of capital, on average? Or even more pertinent, given the extreme right-skewness of the capital distribution (based on ANY of the widely used frequency and severity estimators), is how close is the true value of capital, on average, to the quantiles in the right tail of the (estimator's) capital distribution? Stated differently, how well does the estimator "pull in" and eliminate extremes in the right tail? The most established and widely used statistic that at least indirectly addresses the first question is, simply, the RMSE. And Opdyke (2014) shows RCE-based capital to always have smaller – and often dramatically smaller – RMSE compared to MLE-based capital. Regarding the second question, specifically with reference to RCE, Opdyke (2014) showed empirically that the right tail of the capital distribution (even as close to the body as the 60%tile) was far closer to true capital than that based on MLE. In other words, Opdyke (2014) showed that the RCE-based capital distribution is far less skewed than that based on MLE (by both traditional measures of skew and quantile-based measures). And skewness is the far more important question to address in this setting compared to so-called "median bias": wildly inflated capital estimates in the right tail, due to instability of the estimator (as happens to Zhou et al. (2015) in the absence of "trimming"), are exactly what researchers and regulators are most concerned with and seeking to avoid, not whether the median of the estimator is close(r) to true capital.

Thus does Opdyke (2014) show that the two most established and widely used metrics – skewness and RMSE – that also happen to matter most in this setting are those by which RCE-based capital has been rigorously tested and is vastly superior to MLE-based capital. So-called "median bias" is an irrelevant artifice in this setting.

CONTACT

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