GIRO Conference and Exhibition 2012
Juggling uncertainty the actuary’s part to play

GIRO Conference and Exhibition 2012
Statistical Modelling of Operational Risk Severity Distributions with Insurance Applications
Jens Perch Nielsen
Four published papers on operational risk


AGENDA
AGENDA

- Painting the picture
- Collecting the materials
AGENDA

- Painting the picture
- Collecting the materials
- Behind the scenes
- Mixing the colours
AGENDA

- Painting the picture
- Collecting the materials
- Behind the scenes
- Mixing the colours
- Painting by numbers

AGENDA

- Painting the picture
- Collecting the materials
- Behind the scenes
- Mixing the colours
- Painting by numbers
- Framing the canvas
Painting the picture

Collecting the material
Behind the scenes
Mixing the colours
Painting by numbers
Framing the canvas
Up to $100 Billion

$7.2 Billion

9/11 (2001), New York, USA

Trading Loss

Société Générale, Paris

Painting by numbers

Mixing the colours

Framing the canvas

Collecting the material

Behind the scenes

Painting the picture

Festina Lente
$100 Billion (1995), London, Barings Bank
Trading Loss

$7.2 Billion

$1.3 Billion

$42 Billion (2004), Moscow, Yukos Oil Corporation
Improper Business

Mixing the colours
Painting by numbers
Framing the canvas
Behind the scenes
Collecting the material
Painting the picture
More than 100 losses exceeding $100 Million over the last decade in the financial industry
Definition

Operational risk is the risk of economic loss resulting from inadequate or failed internal process and methodologies, people, systems, or from external events.

Operational Risk Universe

- Crime
- Employment Practices
- External Requirements
- External Events
- Internal Changes
- Systems
- Business Processing
- Relationship with Counterparties

From ? to !

Mixing the colours
Painting by numbers
Framing the canvas
Collecting the materials

Operational Risk Quantification
Operational Risk Quantification

Scenario 1

Frequency

Severity

Festina Lente
Operational Risk Quantification

Scenario 1

Severity

Scenario 2

Scenario 3

Operational Risk Quantification

Scenario 1

Severity

Scenario 2

Scenario 3
Operational Risk Quantification

Scenario 1

Scenario 2

Scenario 3

Scenario $n$

Festina Lente
Operational Risk Quantification

Scenario 1

Scenario 2

Scenario 3

Scenario n

Operational Risk Quantification

Scenario 1

Scenario 2

Scenario 3

Scenario n
Operational Risk Quantification

Scenario 1

Scenario 2

Scenario 3

Scenario n

What type of frequency and severity model should a company use to predict a correct operational risk exposure?
A Comprehensive overview

Mixing the colours

Painting by numbers

Framing the canvas

Collecting the material

Behind the scenes

Painting the picture

Data

Internal Data, Consortium, Publ. Reported Data
A Comprehensive overview
A Comprehensive overview

<table>
<thead>
<tr>
<th>Frequency Distribution</th>
<th>Severity Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative Binomial</td>
<td>Credibility Theory</td>
</tr>
</tbody>
</table>

Monte Carlo Simulation

Internal Data, Consortium, Publ. Reported

A Comprehensive overview

<table>
<thead>
<tr>
<th>Frequency Distribution</th>
<th>Severity Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative Binomial</td>
<td>Credibility Theory</td>
</tr>
<tr>
<td>EVT / GPD</td>
<td></td>
</tr>
</tbody>
</table>

Monte Carlo Simulation

Internal Data, Consortium, Publ. Reported
## A Comprehensive overview

<table>
<thead>
<tr>
<th>Frequency Distribution</th>
<th>Severity Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative Binomial</td>
<td>Credibility Theory</td>
</tr>
<tr>
<td>EVT / GPD</td>
<td>EVT / GPD</td>
</tr>
<tr>
<td>Semi-Parametric, Transformation Approach</td>
<td>-</td>
</tr>
<tr>
<td>Polynomial</td>
<td>Polynomial</td>
</tr>
<tr>
<td>Underreporting</td>
<td>Overreporting</td>
</tr>
<tr>
<td>Overreporting</td>
<td>Overreporting</td>
</tr>
<tr>
<td>Credibility Theory or Bayesian Theory</td>
<td>-</td>
</tr>
</tbody>
</table>

### Data

- Internal Data, Consortium, Published Reported

### Monte Carlo Simulation

- A Comprehensive overview
- Mixing the colours
- Painting by numbers
- Framing the canvas

### Sources

- 3 sources
- Credibility Theory or Bayesian Theory
### A Comprehensive overview

<table>
<thead>
<tr>
<th>Data</th>
<th>Frequency Distribution</th>
<th>Severity Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Negative Binomial</td>
<td>Credibility Theory</td>
</tr>
<tr>
<td></td>
<td>EVT / GPD</td>
<td>EVT / GPD</td>
</tr>
<tr>
<td></td>
<td>Semi-Parametric</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Polynomial</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Overreporting</td>
<td>Overreporting</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Credibility Theory or Bayesian Theory</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 sources</td>
</tr>
</tbody>
</table>

#### Monte Carlo Simulation

- Laplace transform, Fast fourier Transform
- n-fold convolution, Recursive method

### Behind the scenes
Semiparametric Transformation Estimator

- When data are limited, the model is close to a parametric model.
- As the number of losses increases, the model becomes more non-parametric.

**Distribution Assumption.** Estimate the distribution parameters on the data \((X_i)_{i=1}^{100}\)

\[ Z_i = F_{\hat{\theta}}(X_i) \]

Transform losses to the bounded support \([0, 1]\) with

\[ \hat{g}(z) = \frac{1}{100} \sum_{i=1}^{100} K_{h_i}(z - Z_i) \]

with \( K(z) = \frac{3}{4} \left( 1 - z^2 \right) I_{[0,1]}(z) \) and \( K_h(z) = \frac{1}{h} K \left( \frac{z}{h} \right) \)

**KDE:**

\[ \hat{f}(x) = f_{\hat{\theta}}(x) \hat{g}(F_{\hat{\theta}}(x)) \]

Semi-parametric density,
Semiparametric Transformation Estimator

Kernel Density Estimation

Density vs Transformed Data

Z = 0.7
Semiparametric Transformation Estimator

Kernel Density Estimation

0.0 0.2 0.4 0.6 0.8 1.0
0.0 0.5 1.0 1.5 2.0

Density

0.0 0.5 1.0 1.5 2.0
0.0 0.2 0.4 0.6 0.8 1.0

Transformed Data

Semiparametric Transformation Estimator

Kernel Density Estimation

0.0 0.2 0.4 0.6 0.8 1.0
0.0 0.5 1.0 1.5 2.0

Density

0.0 0.5 1.0 1.5 2.0
0.0 0.2 0.4 0.6 0.8 1.0

Transformed Data
Semiparametric Transformation Estimator

Kernel Density Estimation

Transformed Data

Density

0.0 0.2 0.4 0.6 0.8 1.0

0.0 0.5 1.0 1.5 2.0

Semiparametric Transformation Estimator

Kernel Density Estimation

Transformed Data

Density

0.0 0.2 0.4 0.6 0.8 1.0

0.0 0.5 1.0 1.5 2.0
Semiparametric Transformation Estimator

Kernel Density Estimation

0.0 0.2 0.4 0.6 0.8 1.0
0.0 0.5 1.0 1.5 2.0
Density

Transformed Data

0.0 0.2 0.4 0.6 0.8 1.0
0.0 0.5 1.0 1.5 2.0
Density

Transformed Data
Semiparametric Transformation Estimator

Kernel Density Estimation

Transformed Data

Density

0.0 0.2 0.4 0.6 0.8 1.0

0.0 0.5 1.0 1.5 2.0

Kernel Density Estimation

Transformed Data

Density

0.0 0.2 0.4 0.6 0.8 1.0

0.0 0.5 1.0 1.5 2.0

Semiparametric Transformation Estimator
Semiparametric Transformation Estimator

Mixing the colours
Underreporting Model

Underreporting means that not all losses in the company are reported.

Occurred losses $(X_i)_{i\in\mathbb{N}}$  

Indicator function $I(i) = \begin{cases} 1 & \text{if } X_i \text{ is reported} \\ 0 & \text{otherwise} \end{cases}$
Underreporting Model

Losses could be hidden, ignored, unrecorded or reported elsewhere

Reported losses $\{(Y_i)_{1 \leq i \leq N}\}$

Total number of reported: $N = \sum_{i=1}^{N} I(i)$

Underreporting Model

An underreporting function encodes the likelihood that a loss of a particular size is being reported
**Underreporting Model**

An underreporting function encodes the likelihood that a loss of a particular size is being reported.

- Define $g$ as a function of $f$ and an underreporting function $u$:
  \[ g(x) = \frac{f(x)u(x)}{\int f(w)u(w)dw} \]

- The probability of observing an Operational risk loss:
  \[ P = \int g(w)u(w)dw \]

- Assume Poisson distribution, then occurred / reported losses:
  \[ M \sim \text{Po}(\lambda), \quad N \sim \text{Po}(\mu - \lambda) \]

---

**Mixing Model**

- Include prior knowledge from external data.
- Correct the external global start with internal observed data.
Mixing Model

- Include prior knowledge from external data.
- Correct the external global start with internal observed data.

Distribution Assumption. Estimate the distribution parameters on the data $(Y)_{i=1}^{M}$

Transform losses to the bounded support $[0,1]$ with $Z_i = F_{\hat{h}}(X_i)$

KDE: $\hat{g}(z) = \frac{1}{100} \sum_{i=1}^{100} K_i \mathbf{c}_i \mathbf{z} \mathbf{z}$ with $K(z) = \frac{3}{4} \left(1 - \frac{1}{2} z^2 \right) 1_{|z| \leq 1}$ and $K_{\hat{h}}(\cdot) = \frac{1}{h} K\left(\frac{\cdot}{h}\right)$

Semi-parametric density, $f_{\hat{h}}(x) = f_{\hat{g}}(x) \hat{g}(F_{\hat{h}}(x))$

Mixing Model

Density Estimation (Main Body)

Mixing the colours
Painting by numbers
Framing the canvas
Collecting the material
Behind the scenes
Painting the picture
Festina Lente

Density Estimation (Tail)

Internal Lognormal
External Lognormal

Probability
Size

0.000
0.001
0.002
0.003
0.004
0.005
0.006
0.007
0.008
0.009
0.010

0.000
0.010
0.020
0.030
0.040
0.050
0.060
0.070
0.080

0.000
0.010
0.020
0.030
0.040
0.050
0.060
0.070
0.080

0.000
0.010
0.020
0.030
0.040
0.050
0.060
0.070
0.080

0.000
0.010
0.020
0.030
0.040
0.050
0.060
0.070
0.080

Festina Lente

Density Estimation (Main Body)

Mixing the colours
Painting by numbers
Framing the canvas
Collecting the material
Behind the scenes
Painting the picture
Festina Lente

Density Estimation (Tail)

Internal Lognormal
External Lognormal

Probability
Size

0.000
0.001
0.002
0.003
0.004
0.005
0.006
0.007
0.008
0.009
0.010

0.000
0.010
0.020
0.030
0.040
0.050
0.060
0.070
0.080

0.000
0.010
0.020
0.030
0.040
0.050
0.060
0.070
0.080

0.000
0.010
0.020
0.030
0.040
0.050
0.060
0.070
0.080

0.000
0.010
0.020
0.030
0.040
0.050
0.060
0.070
0.080

Festina Lente
Mixing Model

Density Estimation (Main Body)

- Internal Lognormal
- External Lognormal
- Semi-parametric Internal

Density Estimation (Tail)

- Internal Lognormal
- External Lognormal
- Semi-parametric Internal

Mixing Model

- Mixing the colours
- Painting by numbers
- Framing the canvas
- Collecting the material
- Behind the scenes
- Painting the picture

Density Estimation (Main Body)

- Internal Lognormal
- External Lognormal
- Semi-parametric Internal

Density Estimation (Tail)

- Internal Lognormal
- External Lognormal
- Semi-parametric Internal

Density Estimation (Tail)

- Internal Lognormal
- External Lognormal
- Semi-parametric Internal
Asymptotic Theory: Semiparametric Estimator

Let the transformation function $T$ be two times differentiable known function. Assume that $f$ is also two times differentiable. Then the bias of $\hat{f}$ is given by

$$E\hat{f}(x) - f(x) = \mu_{\mu(K)} B, b^2 + o(b^2)$$

and the variance is given by

$$\sqrt{\text{Var}(\hat{f}(x))} = (nb)^{-1} R(K) T'(x) f(x) + o((nb)^{-1})$$

where the asymptotics is given for $n \to \infty$.

Asymptotic Theory: Semiparametric Estimator corrected for UR

Let the transformation function $T$ and the underreporting function $u$ be two times differentiable known function. Assume that $g$ is also two times differentiable. Then the bias of $\hat{g}$ is given by

$$E\hat{g}(x) - g(x) = \mu_{\mu(K)} \left[ \frac{B}{u(x)} A - \frac{A}{B} \int \frac{B}{u(w)} dw \right] b^2 + o(b^2)$$

where

$$A = \frac{f(x)}{u(x)}, B = \int f(w) u(w) dw$$

and the variance is given by

$$\sqrt{\text{Var}(\hat{g}(x))} = \frac{R(K) T'(x) g(x)}{u(x) B} m b + o((mb)^{-1})$$

where the asymptotics is given for $m \to \infty, b \to 0$ and $n = P - m$. 
Painting by numbers

Results Paper A (20,000 Simulations)
Results Paper A (20,000 Simulations)

Mixing the colours

Painting by numbers

Framing the canvas

Collecting the material

Behind the scenes

Painting the picture
Key Findings Paper A

- Distribution dependent results
- Correction for underreporting increases the capital
- A small stabilizing effect by introducing a correction for underreporting

Results Paper B (10,000 Simulations)
Results Paper B (10,000 Simulations)

Mixing the colours
Painting by numbers
Framing the canvas
Behind the scenes
Collecting the material
Painting the picture
Results Paper B (10,000 Simulations)

Mixing the colours
Painting by numbers
Framing the canvas
Behind the scenes
Collecting the material
Painting the picture

Results Paper B (10,000 Simulations)

49%
33%
4%
3%

Festina Lente
Key Findings Paper B

- Consistency with the findings in Paper A
- A stabilizing effect by introducing kernel smoothing
- Even more stabilizing effect by combining correction for UR and incorporate KS

Results Paper C (10,000 Simulations)
Key Findings Paper C

- For lower quantiles, the mixing model resembles the internal model.
- For higher quantiles, more prior knowledge from the external source is present.

Results Paper D (20,000 Simulations)
Results Paper D (20,000 Simulations)

[Graph showing data]

Results Paper D (20,000 Simulations)

[Graph showing data]
Results Paper D (20,000 Simulations)

Key Findings Paper D

- Consistency with the findings in paper C:
  - Mixing model resembles the internal model for lower quantities. For higher quantities, more information from external data is present.

- Consistency with findings in paper A and B:
  - Correction for underreporting increases the capital.
Main Findings and Conclusions

➢ The developed method can
  • Pooling internal and external data to produce more credible capital estimates
  • Incorporate correction for truncation (unrecorded losses) and random censoring (unreported losses)
  • Robust capital method, not distribution dependent
Considerations and future research

- Improve boundary correction, bandwidth selection and correction function
- Investigate and implement possible correlation between event risk categories
- Scaling and filtration (external data)
- Take into account insurance cover
- Compare against other mixing models

Sponsorship / Consultancy proposal

- We are looking for two sponsors of 50K£ each to sponsor our upcoming book on operational risk
- The money is used to buy off the four authors (academics as well as practitioners) from other duties
- The sponsors receive two VERY VALUABLE opportunities:
  1. Will play a key role when the book is launched at an event at Cass Business School
  2. One week exclusive course in the book in Copenhagen, copy of the book and the software. Accomodation paid for (up to four people).
THANK YOU!

Questions or comments?

Expressions of individual views by members of The Actuarial Profession and its staff are encouraged. The views expressed in this presentation are those of the presenter.