Institute of Actuaries of Australia

OPERATIONAL RISKS IN BANKS: AN ANALYSIS OF EMPRICAL DATA FROM AN AUSTRALIAN BANK

Prepared by
Associate Professor John Evans, Australian School of Business, UNSW
Associate Professor Robert Womersley, Faculty of Science, UNSW
Danny Wong, Faculty of Science, UNSW

Presented to the Institute of Actuaries of Australia
Biennial Convention 23-26 September 2007
Christchurch, New Zealand

The Institute of Actuaries of Australia
Level 7 Challis House 4 Martin Place
Sydney NSW Australia 2000
Telephone: +61 2 9233 3466 Facsimile: +61 2 9233 3446
Email: actuaries@actuaries.asn.au Website: www.actuaries.asn.au

Abstract

This paper reports the results of an empirical analysis of operational risk in an Australian Bank and derives a model to represent the distribution of losses. Comparisons are made with models traditionally used to model operational risk. The paper concentrates more on the severity issue than the frequency issue.

Keywords: Operational Risk, Extreme Value Theory, Multi Distributions

The Nature of Operational Risk

Operational risk is usually defined as the risk resulting from inadequate or failed internal processes, people, and systems or from external events.

Banks in Australia are required by the regulator to model their operational risks and to satisfy the regulator of not only the suitability of the modelling process, but also of the appropriateness of the resultant capital required to manage the residual risk remaining to give an overall probability of survival of the bank of 99.9%, which is a very high order of implied accuracy of the modelling. Naturally, banks want the smallest estimate of expected losses that satisfies this confidence limit to minimise capital required.

One of the impediments to meeting this requirement is that most banks have not collected any operational risk data as it was generally not needed and the cost was deemed not justifiable. Even if the data had been collected, accuracy would have been an issue as indirect losses such as system errors, which cause delay in transactions, may produce losses which are not readily quantifiable and the duration of operational loss events can vary significantly.

There is also the problem of “truncation” which refers to the minimum loss for reporting purposes, and this usually changes over time, and varies between banks, making inter bank comparisons difficult.

There are two separate issues that need to be considered when evaluating operational risk, namely, the severity of losses, i.e. the amount, and the frequency of losses, i.e. the number of occasions the loss occurs. In this paper we will concentrate more on the severity issue.

Data

Data was obtained from an Australian bank for the period 1988 through 1996 and adjusted for inflation.

The following table shows the summary statistics of the operational losses reported:

Statistic	Value
Mean	\$2.0 x 10^{6}
Median	\$3.6 x 10^{5}
Variance	\$5.0 x 10^{14}
Standard Deviation	\$2.2 x 10^{7}
Minimum	\$1.2 x 10^{4}
Maximum	\$1.4 x 10^{9}

Table 1: Summary Statistics

The following graph shows the frequency and severity of the operational losses:

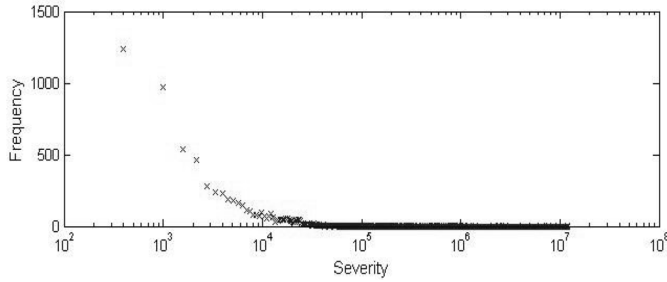

Figure 1: Frequency v Severity of Operational Risk Losses

Observations on the Loss Distribution

Some of the main observations are:

- There are a large number of small losses combined with small number of large losses as indicated in the frequency versus severity plot in Figure 1.
- The time series plot in Figure 2 reveals clear evidence of extreme values. Figure 3 showing the empirical density of the losses and the empirical distribution on a logarithmic scale also supports this view.
- The occurrences of the losses are irregularly spaced in time, suggesting non stationarity.
- The severity and frequency of losses tend to decrease with time. This does not support the hypothesis that a reporting bias exists as suggested in some previous studies (e.g. Chavez-Demoulin and Embrechts [1] and Embrechts et al. [3]) where the severity and frequency tend to increase in time, reflecting the increased awareness and reporting of operational losses. Our results may reflect improved risk management practices and/or a modified data collection process.
- The data is very skewed and kurtotic. The kurtosis stems from the concentration of data points in the lower losses and the skewness is due to the extreme data points with the largest loss being approximately 64 standard deviations away from the mean.

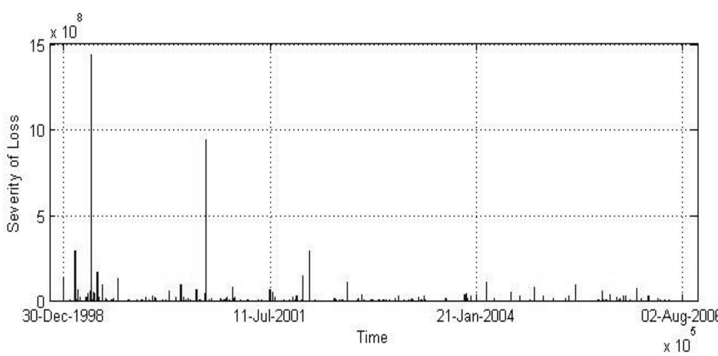

Figure 2: Time Series plot of Operational Risk Losses

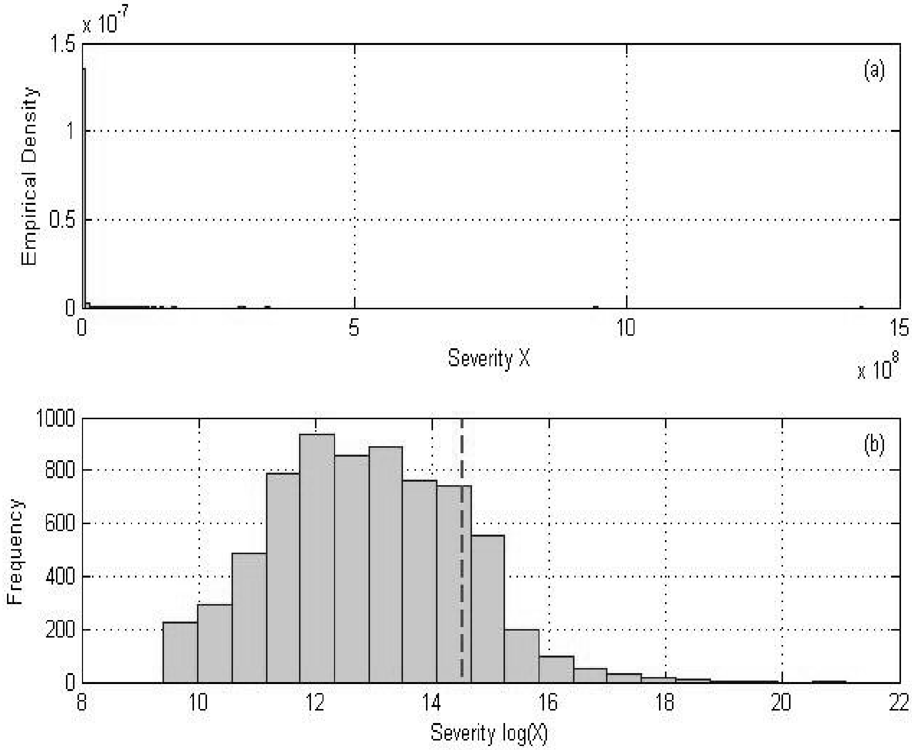

Figure 3: (Top) Empirical Distribution of Losses (Bottom) Empirical Distribution of Losses on a Logarithmic Scale (the dotted line represents the mean)

Normality Analysis

The distorted nature of the normality plot in Figure 4 clearly supports the hypothesis that operational risk data is not normally distributed, with both the QQ-plots and PP-plots deviating from the 45 degree linear reference line significantly.

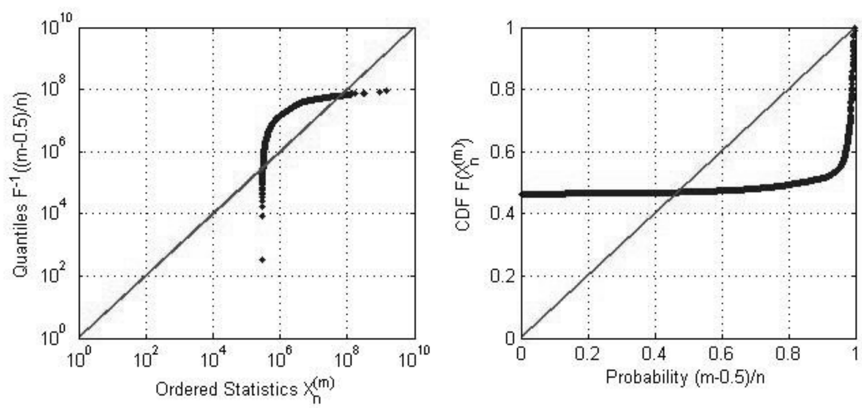

Figure 4: Normality Plot of the Severity of Losses (the diagonal line represents the 45 degree reference line)

Lognormal Analysis

Significant improvements are made when we model the data with a lognormal distribution as seen in Figure 5. The PP-plot almost coincides with the reference line and the majority of the QQ-plot is linear.

However, due to the curvature at the tails, even the lognormal distribution is unable to properly account for the extreme nature of the data.

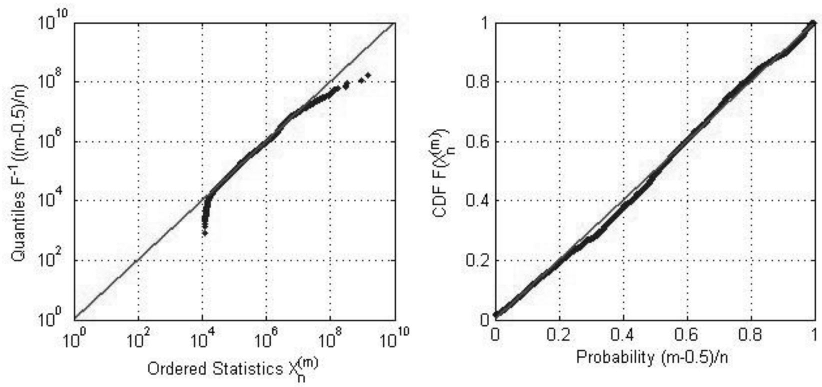

Figure 5: QQ-plot using the Lognormal Distribution of the Severity of Losses (the diagonal line represents the 45 degree reference line)

Single Distribution Modeling

It was impractical to measure the dependence structure for all 56 risk cells as specified under the Basel II guidelines. Thus, the dependence analysis was conducted by splitting the data up into a bivariate case consisting of two business lines only - ‘Retail Banking’ (RB) and ‘All Others’ (AO).

Both the Poisson and Negative Binomial distributions were applied to the data. The Poisson distribution is inappropriate for modelling the frequency of losses as the ratio between the sample variance and sample mean should be approximately equal to 1 for the Poisson distribution to be suitable and in our case this ratio is 10.8. The estimated parameters and the corresponding 95% confidence levels are shown in Table 2, and comparisons of the expected results using the Poisson and Negative Binomial distributions and the actual results are shown in Figure 6.

Distribution	**Parameters**	**95% Confidence Level**	
Poisson	$\lambda = 75.4176$	73.6333	77.2019
Negative Binomial	$r = 7.2923$	4.8808	9.7038
	$\frac{1}{1+\beta} = 0.0882$	0.0608	0.1155

Table 2: Parameter Estimation for the Frequency of Losses.

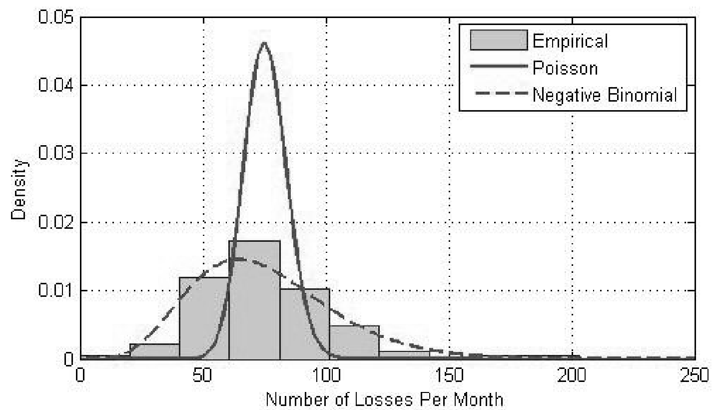

Figure 6: Plot of the Empirical Frequency Distribution fitted with the Poisson and Negative Binomial Distribution

The peaks over thresholds methodology in Extreme Value Theory was used to model the severity of the losses. Two different parameter estimation methods were used, namely maximum likelihood estimation (MLE) and probability weighted moments (PWM). The results can be seen in Table 3 where $\mathbf{G}^{\mathrm{MLE(S)}}$ and $\mathbf{G}^{\mathrm{PWM(S)}}$ denote the severity models corresponding to MLE and PWM estimated parameters, respectively.

	Parameter	**Estimate**	**95% CI**	
MLE	σ	1966911.1422	1680562.7830	2302049.9327
($\mathbf{G}^{\mathrm{MLE(S)}}$)	ξ	1.0397	0.8810	1.1986
PWM	σ	2438920.8784	2116459.8983	2903196.9420
($\mathbf{G}^{\mathrm{PWM(S)}}$)	ξ	0.8099	0.6928	0.8704

Table 3: Parameter Estimation for the Severity of Losses using the GDP Distribution

From the QQ and PP plots in Figure 7 it is evident that using GDP does improve the fit for the loss data as the plots are fairly linear and coincide well with the 45 degree line.

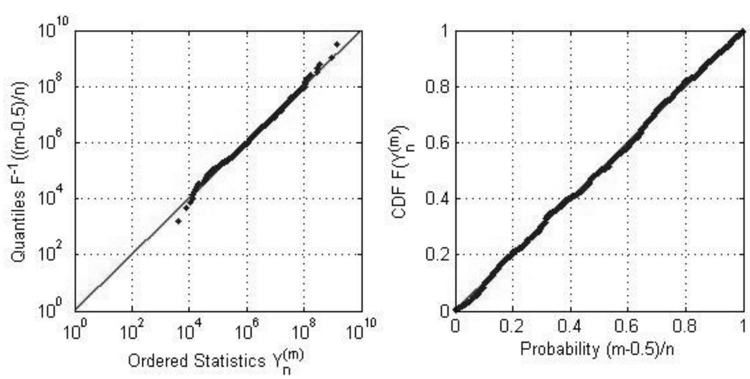

Figure 7: QQ-Plot (Left) and PP-Plot (Right) for the Truncated Severity Data using the MLE Parameters

We considered the effect on VaR of $\mathbf{G}^{\mathrm{MLE(S)}}$ and $\mathbf{G}^{\mathrm{PWM(S)}}$ along with the lognormal distribution for comparison, and compared the number of violations from the model with the expected number of violations, where a violation occurred when the data value exceeded the calculated VaR(α). For a confidence level α with n observations, the expected number of violations will be $n(1-\alpha)$. If the number of violations is higher than the expected number of violations, then the model underestimates the extreme risk. From Table 4 it is clear that $\mathbf{G}^{\mathrm{MLE(S)}}$ provides the best fit in terms of least violations while the number of violations for the Lognormal distribution significantly increases as the confidence level increases. Both the GPD models pass the kupiec test in that they coincide with the null hypothesis that the data conforms with the selected model, whereas the Lognormal rejects the null hypothesis at all α levels.

α	Number of Violations			
	Theoretical	$\mathbf{G}^{\mathrm{MLE(S)}}$	$\mathbf{G}^{\mathrm{PWM(S)}}$	**Lognormal**
0.950	347	346	321	286
0.990	69	72	83	90
0.995	35	32	43	61
0.997	21	21	28	45
0.999	7	5	9	25

Table 4: Number of Violations calculated using the Fitted Distributions

An aggregate loss distribution was formed by simulating annual aggregate losses and then fitting those losses to an appropriate distribution. One hundred thousand simulations were used. The frequency distribution used was the Negative Binomial with parameters as in Table 2. The severity was simulated using both sets of parameters $\mathbf{G}^{\mathrm{MLE(S)}}$ and $\mathbf{G}^{\mathrm{PWM(S)}}$. The simulations generated from $\mathbf{G}^{\mathrm{MLE(S)}}$ and $\mathbf{G}^{\mathrm{PWM(S)}}$ is denoted as $S(\mathbf{G}^{\mathrm{MLE(S)}})$ and $S(\mathbf{G}^{\mathrm{PWM(S)}})$, respectively. The statistical characteristic of the resulting simulation is shown Figure 8. The simulated data continues to shows clear evidence of skewness and kurtosis even on a logarithmic scale.

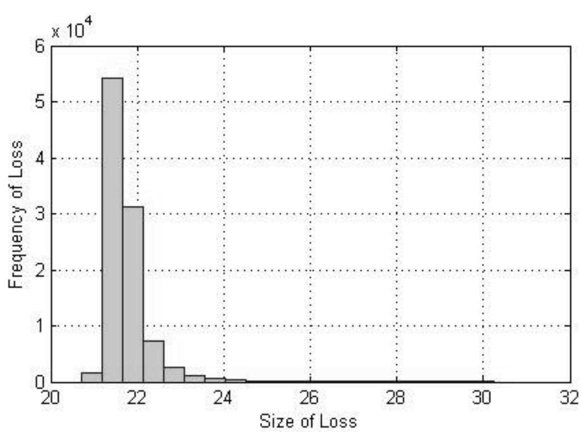

Figure 8: Histogram of the Aggregate Losses for the Simulation using the MLE Parameters $\mathbf{G}^{\mathrm{MLE(S)}}$ on a Logarithmic Scale (the corresponding histogram for $\mathbf{G}^{\mathrm{PWM(S)}}$ is similar, but slightly less extreme in nature)

The distribution is again very kurtotic and right-skewed as expected. The largest loss in $S(\mathbf{G}^{\text{MLE(S)}})$ is 1.3599×10^{13} and in $S(\mathbf{G}^{\text{PWM(S)}})$ is 7.1276×10^{11}. Intuition suggests that both these amounts are far too large to be realistic, especially given the size of the bank for which we have data. The MLE parameters clearly give much larger losses as can be seen in the statistics. However, the simulations for PWM are more kurtotic and skewed suggesting that there is a greater tendency for smaller losses. To test sensitivity of the parameters, various combinations of σ and ξ were used to perform the simulation. It was found the size of the losses where particularly sensitive to changes in ξ but even large changes in σ did not have any noticeable effect on the size of losses.

Statistic	Value (MLE)	Value (PWM)
Mean	4.6333×10^{9}	1.8301×10^{9}
Median	2.2529×10^{9}	1.5372×10^{9}
Variance	9.8909×10^{21}	3.1668×10^{19}
Standard Deviation	9.9453×10^{10}	5.6275×10^{9}
Semi-variance	1.0796×10^{22}	4.1584×10^{19}
Kurtosis	1.7138×10^{4}	2.0631×10^{4}
Skewness	1.2180×10^{2}	1.2985×10^{2}
Minimum	8.7975×10^{8}	6.7782×10^{8}
Maximum	1.6930×10^{13}	9.9939×10^{11}

Table 5: Summary Statistics of the Aggregate Losses under MLE and PWM

Both the GEV and GPD were fitted to the simulated loss data. The GEV distribution provided a much better fit than the GPD so we only considered the GEV fit. We let $\mathbf{G}^{\text{MLE(A)}}_{S(\mathbf{G}^{\text{MLE(S)}})}$ and $\mathbf{G}^{\text{PWM(A)}}_{S(\mathbf{G}^{\text{MLE(S)}})}$ denote the GEV fit using MLE and PWM techniques to the simulated data $S(\mathbf{G}^{\text{MLE(S)}})$, respectively, and similarly, the notation for $S(\mathbf{G}^{\text{PWM(S)}})$ fits are $\mathbf{G}^{\text{MLE(A)}}_{S(\mathbf{G}^{\text{PWM(S)}})}$ and $\mathbf{G}^{\text{PWM(A)}}_{S(\mathbf{G}^{\text{PWM(S)}})}$. Both MLE and PWM methods gave roughly the same fit for the GEV distribution. The MLE approach seems to provide a slightly more accurate fit to the body of data when compared to the PWM approach. This can possibly be explained by the fact that we have enough simulated data points for the MLE approach to reach asymptotic convergence. The PWM, on the other hand, gives a heavier tail to the distribution, and as a result, performs better in the violations analysis shown in Table 6.

α	Number of Violations				
	Theoretical	$\mathbf{G}^{\text{MLE(A)}}_{S(\mathbf{G}^{\text{MLE(S)}})}$	$\mathbf{G}^{\text{PWM(A)}}_{S(\mathbf{G}^{\text{MLE(S)}})}$	$\mathbf{G}^{\text{MLE(A)}}_{S(\mathbf{G}^{\text{PWM(S)}})}$	$\mathbf{G}^{\text{PWM(A)}}_{S(\mathbf{G}^{\text{PWM(S)}})}$
0.950	5000	6168	3820	4833	3328
0.990	1000	2491	909	1839	843
0.995	500	1749	515	1309	512
0.997	300	1396	323	1019	346
0.999	100	817	124	660	150

Table 6: Number of Violations calculated using the Fitted Distributions

The VaR(α) was calculated using the fitted GEV distributions. These results show that the VaR(α) increases with the confidence level. In addition, the MLE parameters $S(\mathbf{G}^{\mathrm{MLE(S)}})$ produce significantly larger VaR(α) than the corresponding PWM parameters $S(\mathbf{G}^{\mathrm{PWM(S)}})$.

α	VaR(α)			
	$\mathbf{G}^{\mathrm{MLE(A)}}_{S(\mathbf{G}^{\mathrm{MLE(S)}})}$	$\mathbf{G}^{\mathrm{PWM(A)}}_{S(\mathbf{G}^{\mathrm{MLE(S)}})}$	$\mathbf{G}^{\mathrm{MLE(A)}}_{S(\mathbf{G}^{\mathrm{PWM(S)}})}$	$\mathbf{G}^{\mathrm{PWM(A)}}_{S(\mathbf{G}^{\mathrm{PWM(S)}})}$
0.95	6.2016×10^{9}	8.5478×10^{9}	3.0752×10^{9}	3.5329×10^{9}
0.99	1.1906×10^{10}	2.8702E $\times 10^{10}$	4.6355×10^{9}	7.1614×10^{9}
0.995	1.5905×10^{10}	4.9976×10^{10}	5.5387×10^{9}	9.9851×10^{9}
0.997	1.9740×10^{10}	7.5671×10^{10}	6.3192×10^{9}	1.2854×10^{9}
0.999	3.1591×10^{10}	1.8651×10^{11}	8.4054×10^{9}	2.2479×10^{10}

Table 7: Value-at-Risk Values

It is clear that this methodology overestimates the tail of the aggregate losses significantly. This problem was also highlighted by King [4].

Multi Distribution Modelling

A possible solution for this tail estimation problem is the use of multiple distributions which will dampen the overestimating of the tails as it will place more weight on smaller losses and less weight on the tail. The use of multiple models essentially restricts the number of larger losses that can occur, thus giving a much more reliable estimate of the aggregate distribution as it takes into account the rarity of the extreme losses in the frequency.

Difficulties were encountered in attempting to fit multiple distributions. We used MLE to simultaneously maximise the likelihood in both distributions as well as the weighting factor. The algorithm used was based on trying to maximise the log-likelihood of the mixture solution but unfortunately no optimal solution was found. The process was simplified by taking a multi-step approach. The data was split into two portions - smaller than threshold u_t (body X_b) and larger than u_t (tail X_t) as represented in Figure 9.

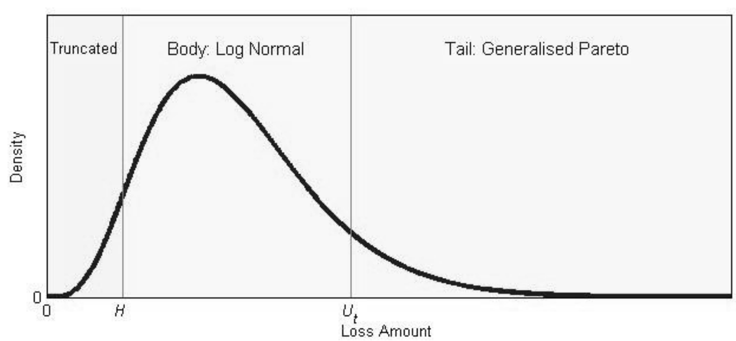

Figure 9: Illustration of the Mixed Distribution Concept (*H* is the truncation point and U_t is the value were the tail is believed to begin)

The losses X_b were fitted with the lognormal distribution using MLE and extreme value theory techniques applied to X_t. If l_m denotes the log-likelihood function for the mixture model and f_{LN} and f_{GPD} are the densities of the fitted Lognormal and GPD, the weighting factor w was varied between zero and one to maximise

$$\mathrm{l}_m = \sum \ln\left(w f_{\mathrm{LN}} + (1-w) f_{\mathrm{GPD}} \right).$$

This method also failed to converge to a solution. The resulting weights were either close to zero or close to one.

An empirical based technique was then used to implement the mixture distribution. We chose u to be the truncation point used in the bank wide analysis, and the weighting factor was the empirical estimation for the proportion of losses less than u, that is, $w = \frac{\text{\# losses less than } u}{\text{total \# losses}}$. The Lognormal distribution was estimated using MLE yielding parameters $\mu = 12.5693$ and $\sigma^2 = 1.3522$. The statistical characteristics of the mixed severity distribution are shown in Table 8.

Statistic	Value (MLE)	Value (PWM)
Mean	4.4979×10^9	1.6299×10^9
Median	2.3252×10^9	1.3954×10^9
Variance	6.7723×10^{18}	1.2010×10^{19}
Standard Deviation	2.6024×10^9	3.4656×10^9
Semi-variance	9.4711×10^{18}	1.6111×10^{19}
Kurtosis	2.1616×10^4	2.4960×10^4
Skewness	1.1967×10^2	1.2918×10^2
Minimum	6.1717×10^8	6.4101×10^8
Maximum	5.5123×10^{11}	7.6331×10^{11}

Table 8: Statistics for the Aggregate Loss Simulation produced from the Mixture Severity Distribution.

The use of the multiple distributions produced less extreme losses for the aggregate distribution. This is evident when Table 5 and Table 8 are compared. The maximum losses and variance are significantly smaller for the multiple distribution case. This is especially true for the simulations using $\mathbf{G}^{\mathrm{MLE(S)}}$ parameters where the maximum loss and variance was reduced by a factor of 30 and 1400, respectively. Furthermore, the mean and median remained stable for both methods. The GEV distribution generated a better fit.

Comparisons with non Australian Analysis

Our results demonstrate major differences with studies conducted overseas in terms of the empirical analysis, features and characteristics of the data. Our value-at-risk amounts are much smaller even when compared to a medium-sized non-internationally active U.S. bank [5]. It may be that since Australian banks tend to hold higher proportions of residential mortgage loans in their accounts than most overseas banks, Australian banks are expected to be less risky than equivalent overseas banks , and consequently the need to hold large capital reserves is reduced.

The analysis performed by de Fontnouvelle et al. [2] indicates that non-U.S. operational losses are significantly larger than U.S. losses. The percentiles for the non-U.S. losses are approximately double the equivalent percentiles for U.S. losses at both the aggregate and business line level. This is inconsistent with our data set, yielding capital reserves in the order of hundreds of millions rather than billions.

Another inconsistency is the modelling of the frequency of losses where most banks have used the Poisson distribution. This is most likely due to the greater number of favourable statistical properties inherent in the Poisson distribution rather than the ability to produce a better fit.

Conclusion

Extreme value theory has demonstrated significant potential to account for the heavy tail of operational losses where other conventional methods fail. We have shown statistically that the use of conventional methods to model severity is inadequate because the operational loss data exhibits kurtotic and right-skewed behaviour whilst conventional models place emphasis on fitting the central or body of the data, and thus, neglect the extreme percentiles. It is the extreme losses that are important in the type of analysis required under Basel 11.

However, a major limitation in the implementation of extreme value theory is the lack of data which inhibits capturing the generalised Pareto nature of the excess distributions without sacrificing the majority of our data set. The ability to model any sort of dependence is also limited by the availability of quality data. Even if we could overcome these limitations, the regulators may not permit the use of models based on such a small sample despite the accuracy of the dependence models. As such, it may take many years before any banks can convincingly justify the use of any sophisticated dependence structure between the various risk cells and reap the benefits of diversification.

In our view, the Poisson distribution, which is the most common distribution used to model operational risks has proved to be inappropriate for our data set.

References

1. V. Chavez-Demoulin and P. Embrechts, 2004, *Advanced External Models for Operational Risk*, Dept of Mathematics ETH-Zentrum
2. P. de Fontnouvelle, V. DeJesus-Rueff, J. Jordan and E. Rosengren, 2003, *Using Loss Data to Quantify Operational Risk*, Federal Reserve Bank of Boston
3. P. Embrechts, H. Furrer and R. Kaufmann, 2003, *Quantifying Regulatory Capital for Operational Risk*, *Derivative Use, Trading & Regulation*, Dept of Mathematics ETH-Zentrum
4. J. L. King, May 2001, *Measurement and Modelling Operational Risk*, Wiley Finance
5. M. Moscadelli, 2004, *The Modelling of Operational Risk: Experience with the Analysis of the Data Collected by the Basel Committee*, Banca d'Italia, Temi di discussione del Servizio Studi