QUANTIFICATION OF OPERATIONAL RISKS IN BANKS:

A THEORETICAL ANALYSIS WITH EMPRICAL TESTING

Associate Professor John Evans*, Faculty of Business, UNSW
Associate Professor Robert Womersley, Faculty of Science, UNSW
Danny Wong, Faculty of Science, UNSW

*** corresponding author**

This is a work in progress and is not for citing or reproduction

:
Keywords: Operational risk, Extreme Value Theory

JEL Classification: C15, G21, G28,

This paper will consider the appropriateness of the quantification of operational risk using the Loss Distribution Approach, whereby we adapt already well established statistical and actuarial techniques to our modelling problem. We will consider the use of Extreme Value Theory to account for the heavy tails of the losses and the use of Copulas to measure the dependence between the different operational risk classifications. These models will be applied to a historical data set provided by an Australian bank.

The Nature of Operational Risk & Basel 11

The major difference between operational risk and other types of risk is that it represents a downside risk. In other words, the uncertainty forms a loss distribution with no upside potential. This effectively makes the use of any distribution function with unbounded lower end points (e.g. Normal) questionable. Another difference is that it is mostly idiosyncratic in that the risk is mainly firm-specific, subject to the internal drivers of the institution such as operating environment, processes, human resources, internal system and management.

The lack of quality data acts as a physical barrier that impedes the advancement of operational risk research. Most institutions in the past have neglected to collect any operational risk data as it was generally not needed and the cost incurred in such a task could not be justified.

There is also the issue of the ability to accurately measure operational losses. Direct losses are easily quantified (e.g. fines paid to the regulator for certain breaches). On the other hand, indirect losses such as system errors, which cause delay in transactions, may produce losses which are not readily tangible through damage of the institution's reputation. Only rough estimates can be made in such cases.

Furthermore, the duration of operational loss events can vary significantly depending on the loss type which means that evaluation of the impact of the error could take many years to finalise. Once again, an estimation of the expected present value of the loss amount has to be calculated and then recorded as the loss figure.

Risk Measures

There are many statistical techniques that can be applied to financial risks. Ignoring the very basic representations of expected values and variance, Value-at-Risk (VaR) is a risk measure first formally introduced by JP Morgan in 1995 [17] and it has now become the industry standard for the measurement of market risk and has been used intensively as a risk management tool. It is defined as the "predicted worst-case loss at a specific confidence level (e.g. 95%) over a certain period of time (e.g. 1 day)". The $\mathrm{VaR}(\alpha)$ measure cannot however be applied directly to operational losses for several reasons. First, the VaR approach assumes a Gaussian distribution which is inappropriate for the use with operational losses, and second, VaR assumes a continuous stochastic process while operational losses follow a discrete process through 'frequency of events'. However, by relaxing or modifying these assumptions through techniques used in insurance mathematics, it is possible to adapt the VaR approach to the area of operational risk.

All models exhibit some form of estimation error in the underlying parameters, which translate to prediction errors for the calculation of the risk capital. In addition, the calculated values are forecasts of the future and are obviously subject to uncertainty, hence, any value would have to be an expectation with the probability of achieving this exact expectation being almost surely zero. Not surprisingly, the regulator would not accept a single figure for the risk capital, but instead they would also want a confidence interval for an expected range of values.

This confidence interval can be computed in a number of ways. If a parametric distribution is assumed for the aggregate loss distribution, then we can use Monte Carlo simulation to generate a distribution for the risk capital itself in which we can easily produce a confidence interval. Alternatively, we can use the method of bootstrapping which is a method often used when data is scarce. Further details on the method of bootstrapping can be found in Sprent [25].

Bassel 11: Advanced Measurement Approach for Banks

The Advanced Measurement Approach (AMA) is designed to allow banks the greatest flexibility in implementing their own quantitative risk measurement techniques. The underlying motive for such an approach is that each bank has its own risk profile and it would be impossible to devise a set method which can adequately adapt to the profiles of all banks. A number of guidelines are being developed which detail how the quantitative risk management process should be implemented. The attraction of the AMA over the other approaches resides in the fact that the computed capital reserves will be more tailored to the bank's risk profile. The calculated risk capital from the AMA is however subject to a floor amount which is currently set at 75% of the risk capital calculated using the Standardised Approach (similar to a weighted expected value of each business line with the weights being the risk factors). The usage of the AMA serves as a signal to the market place that the bank has well developed risk management practices as only banks which satisfy a stringent list of conditions are allowed to use this more advanced approach. The intention of the AMA is to provide incentives for banks to invest in operational risk measurement projects, best risk practices and systematic data collection . In addition to the use of internal data for the AMA, three other sources of information are required to supplement the modelling process; these are external data, scenario analysis, and business environment and internal control factors.

Capital-at-Risk under Bassel 11

The Capital-at-Risk (CaR) represents the minimum capital requirement that a bank has to hold in order to satisfy the Pillar 1 requirements of Basel II. It is also commonly referred to as the risk capital, capital charge or reserves. There are three commonly accepted definitions for capital-at-risk given as follows

1. CaR: The capital charge is the α = 99.9% quintile of the total loss distribution,
2. $\mathrm{CaR}^{\mathrm{UL}}$: The capital charge only includes the unexpected losses, that is, $\mathrm{VaR}(\alpha)$ minus the expected losses,
3. CaR^{H}: The capital charge is the α = 99.9% quintile of the aggregate loss distribution where only the losses above the threshold H are considered

The separate calculation of the expected and unexpected losses can be justified by arguing that the expected losses are covered by the bank in several ways such as the use of reserves, provisions, pricing, margin income and budgeting [13], and it is therefore the unexpected losses that need capital reserves.

Loss Distribution Approach (LDA)

One of the most advanced methods of loss assessment taken from insurance modelling is the Loss Distribution Approach (LDA). LDA models separate the severity and frequency probability distribution functions for each business line and event type. Modelling the two distributions separately makes sense intuitively. Many techniques have been developed to combine these two distributions to form the aggregate loss distribution. These include Monte-Carlo simulation, Panjer recursion, and the use of the properties of the characteristic function of the distribution functions [14] [21]. Accuracy and quality of data become important issues as the LDA aims to extract more detailed information from the data. Adjustments need to be made to the data to correct for the bias before the loss data can be used to estimate the severity distribution. This bias is caused during the process of data collection where losses less than a certain about is not recorded. Baud et al. [4] showed using simulated data that the calculated $\mathrm{CaR}(\alpha)$ can be overestimated by over 50% if the bias is ignored.

It is important to distinguish between the following two types of operational losses when estimating the LDA frequency distribution:

1. high frequency and low severity events; and,
2. low frequency and high severity events.

In the first case, any significant errors in estimating the frequency will lead to substantial differences in the $\mathrm{CaR}(\alpha)$ as the frequency plays a major role in determining $\mathrm{CaR}(\alpha)$ due to the losses being not only generated by one extreme loss, but also by many small losses. However, in the second case, the error in estimating the frequency will only affect the $\mathrm{CaR}(\alpha)$ slightly [3].

The Poisson distribution adapted from insurance modelling is more commonly used to model the frequency of loss events [15]. The benefits of using the Poisson distribution are as follows (for more properties, see Klugman et al. [21]):

- a Poisson distribution which fits an entire data set will also fit a truncated data set with a simple change in parameter [20],
- it has the property that $\sum_{i=1}^{n} \mathrm{Poisson}(\lambda_i) = \mathrm{Poisson}\left(\sum_{i=1}^{n} \lambda_i\right)$ [21], and
- it is very simple as it only requires a single parameter to specify the entire distribution, however, this also means a lack of flexibility of the distribution [14].

An alternative to the Poisson distribution is the Negative Binomial distribution [21]. It is likely to give a better fit with two parameters specifying the distribution allowing for greater flexibility in the shape [20]. It is also the preferred alternative to the Poisson distribution when the data is "over dispersed".

Many tests can be used to test the suitability or appropriateness of the selected distribution including graphical diagnostics as well as goodness of fit tests [25].

To form the aggregate loss distribution, the collective model in insurance mathematics can be used in which independence is assumed between the frequency and severity distributions [18]. The aggregate loss distribution is taken as the convolution of the severity distribution given that there are N losses and assuming the individual losses X are independently and identically distributed.

Correlation Issues

The simple summation of individual $\mathrm{CaR}_{ij}(\alpha)$'s to calculate the aggregate $\mathrm{CaR}(\alpha)$ assumes perfect dependence among the risk cells. In other words, this implies that everything will simultaneously go wrong for the bank in each risk cell which of course is highly improbable [16]. Frachot et al. [16] show that the correlation of the aggregate losses between the risk cells is generally around 5%~10% which may lead to large amounts of diversification benefits and a significant reduction in capital requirements. In addition, it is the modelling of the dependence between the tails of the risk cells that is important as it is the tails which will make significant differences in the value of the bankwide $\mathrm{CaR}(\alpha)$ [1].

It is likely that the frequency distribution across risk classes will have some sort of dependence due to the exposure to common factors like economic cycles, size of operations etc. This correlation can be easily measured from the historical data assuming that the dependence is relatively linear. On the other hand, it is conceptually difficult, if not impossible given the amount of data available, to include correlation among severity [16].

Frachot et al. [16] argue that the cheapest way to include correlation within the standard LDA framework is to only use frequency correlation and ignore severity correlation.

Extreme Value Theory

Extreme Value Theory (EVT) is a branch of statistics which attempts to account for the behaviour of long tail risks. The key attraction of this methodology is that it offers a well developed approach to deal with the problematic nature of operational risk analysis. EVT aims to extrapolate from past data and forecast extreme events (e.g. events which are only likely to occur once in a 100 years, corresponding to a confidence level of 99.99% on a VaR basis). There are two ways to proceed with EVT. The block maxima method only uses the largest value from each data set, and hence, takes time into consideration as each data set would correspond to each period of data collection. The main distribution function used under this method is the Generalised Extreme Value (GEV) distribution with location parameter μ, scale parameter σ and shape parameter ξ. On the other hand, the peaks over threshold method which is a more widely used method especially in areas where the data is scarce, ignores time and selects events which are larger than a certain threshold u [7]. However, it is possible to fit the model with time-dependent parameters which will allow for non-stationarity of data. The corresponding distribution function used in this case is the Generalised Pareto distribution (GPD) with scale parameter σ and shape parameter ξ. As a final note, any direct implementation of either EVT methods is highly questionable due to the present data availability and data structure. For more details on the mathematics of EVT or on its implementation please refer to McNeil et al. [23].

Copulas

Joint distributions contain information about the marginal distribution of the random variables as well as information about their dependence structure. A copula is a multivariate cumulative distribution function with uniform marginal distributions. Taking a copula approach to operational losses is useful as copulas provide a way of isolating the dependence structure and help enables us to overcome the limitations of linear dependence. It also provides us with a methodology to account for dependence in the tail of the marginal distributions [12].

Empirical Analysis

The data we will use represents historical data from an Australian bank. This data is obtained under strict confidentiality agreements, and has been altered to mask the identity of the bank. As a caveat, by using historical data to model losses we are implicitly assuming that the past losses represent the risks the bank is currently facing which is obviously not correct especially given banks regularly undergo various forms of restructuring and hopefully improve risk management procedures.

For the purposes of this analysis, we will work in calendar years. The gross loss amounts will be adjusted for inflation to ensure the nominal amounts are comparable with each other. The standard measure of inflation that is used is the Consumer Price Index (CPI) as provided by the Australian Bureau of Statistics (ABS) [2]. The CPI indices are rebased to 100 for the date 1/1/1996.

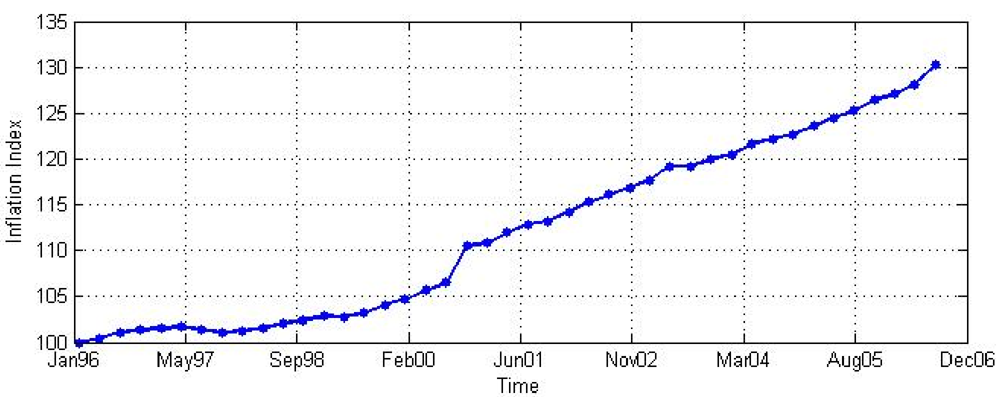

Figure 1

The indices increase approximately at a linear rate apart from the kink in the September quarter of the year 2000. As such, we will fit three linear regression lines corresponding to the interval before the kink, at the kink and after the kink. A continuous compounding rate will then be produced from each of these lines and applied to each loss amount. The reference date used is 01/08/2006, and thus all the losses will be inflated/deflated to this date.

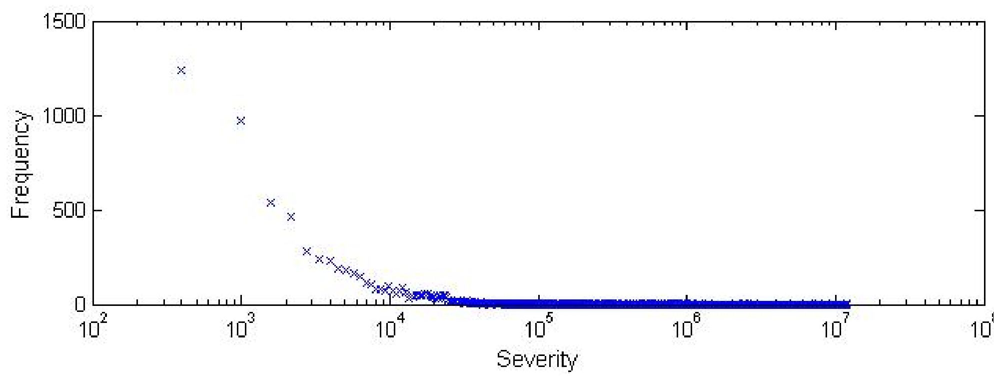

Figure 2

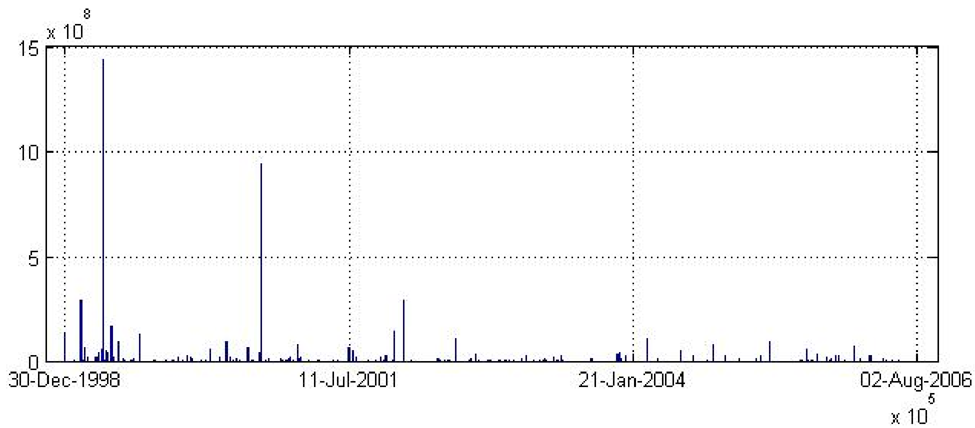

Figure 3

Some of the facts observed are:

- The historical period of the data is relatively short with less than an eight year span. This becomes a problem when we fit the frequency model on an annual basis as it means only seven data points are available.
- There are a large number of small losses combined with small number of large losses as seen in Figure 2.
- The time series plot in Figure 3 reveals clear evidence of extreme values. Figure 4a also supports this view showing the empirical density of the losses while Figure 4b shows the empirical distribution on a logarithmic scale.
- The occurrence of the losses are irregularly spaced in time, suggesting non-stationarity.
- The severity and frequency of losses tend to decrease with time. This contradicts the feature of a reporting bias in some previous studies (e.g. Chavez-Demoulin and Embrechts [6] and Embrechts et al. [11]) where the severity and

frequency tend to increase in time, reflecting the increased awareness and reporting of operational losses. This may reflect improved risk management practices and/or modified the data collection process.

- The data is very skewed and kurtotic as expected. 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.

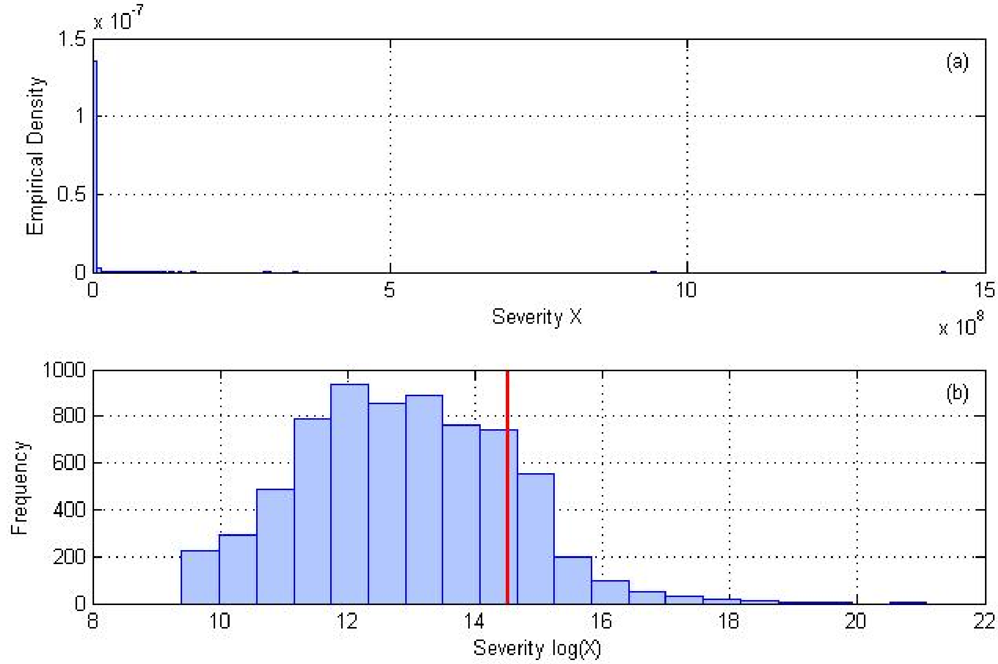

Figure 4

The basic statistical characteristics of the data are:

Statistic	**Value**
Mean	2.0425×10^6
Median	3.6542×10^5
Variance	5.0472×10^{14}
Standard Deviation	2.2466×10^7
Semi-variance	6.0177×10^{14}
Kurtosis	2.8328×10^3
Skewness	49.2578
Minimum	1.2128×10^4
Maximum	1.4334×10^9

Table 1

Table 2 shows the percentage of the number of losses made in each risk cell that we established while Table 3 shows the percentage of the amount of losses made in each risk cell. Similar to the findings made by the Loss Data Collection Exercise (LDCE) conducted by the Basel II committee [13], the majority of the losses are made in Retail Banking business line in the 'Fraud' (Internal and External) and 'Damage to Physical Assets' loss event types. The infrequency of losses in certain risk cells may either reflect their rarity or it may be that these types of losses are often undisclosed or misclassified [8]. In the absence of data over a longer historical period, it is only possible to make the simple assumption that certain losses are more rare than others as suggested by the studies conducted on a consortium of U.S. banks.

	1	**2**	**3**	**4**	**5**	**6**	**7**	**Total**
1	0	0	0	0.01	0	0	0	0.01
2	0	0	0.01	0	0	0.03	0.19	0.23
3	5.53	73.41	1.13	0.04	4.21	0.07	1.75	86.13
4	0.04	0.72	0.29	1.10	0.03	0.14	9.30	11.63
5	0	0.06	0.03	0.07	0	0.12	1.02	1.30
6	0	0	0	0	0	0	0	0
7	0.01	0.04	0.16	0.25	0	0	0.22	0.68
8	0	0.01	0	0	0	0	0	0.01
Total	5.58	74.25	1.62	1.47	4.24	0.36	12.48	100

Table 2

	1	**2**	**3**	**4**	**5**	**6**	**7**	**Total**
1	0	0	0	0.10	0	0	0	0.10
2	0	0	0.09	0	0	0	0.07	0.16
3	5.03	52.52	1.92	0.25	2.33	0.23	9.80	72.09
4	1.05	7.10	0.60	0.24	0.03	0.07	15.80	24.89
5	0	0.02	0	0	0	0.03	0.09	0.14
6	0	0	0	0	0	0	0	0
7	0.93	0.02	0.30	1.19	0	0	0.17	2.60
8	0	0.02	0	0	0	0	0	0.02
Total	7.01	59.68	2.91	1.78	2.36	0.34	25.92	100

Table 3

The following table indicates the proportion of losses falling into each of the years (2006 is not a full year)

	Frequency (%)	**Total Loss (%)**
1999	8.89	26.55
2000	19.58	21.41
2001	14.99	11.41
2002	13.46	11.57
2003	13.17	9.65
2004	11.18	7.79
2005	10.49	7.99

2006	8.24	3.64

Table 4

The distorted nature of the normality plot in Figure 5 clearly supports our hypothesis that operational risk data is not Normally distributed. Both the QQ-plots (left) and PP-plots (right) deviate from the reference line significantly. Significant improvements are made when we model the data with a LogNormal distribution as seen in Figure 6. 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, we deduce that even the heavy tailed LogNormal distribution is unable to properly account for the extreme nature of the data. The results for the other standard insurance loss severity distributions, Weibull and Gamma, are similar. Although there is still significant deviation from linearity, they still perform better than the Normal distribution. The right tail remains poorly accounted for using these distributions. Although the left tail appears to be more curved than the right, this distortion is more likely due to the fact that these plots are made on a logarithmic scale.

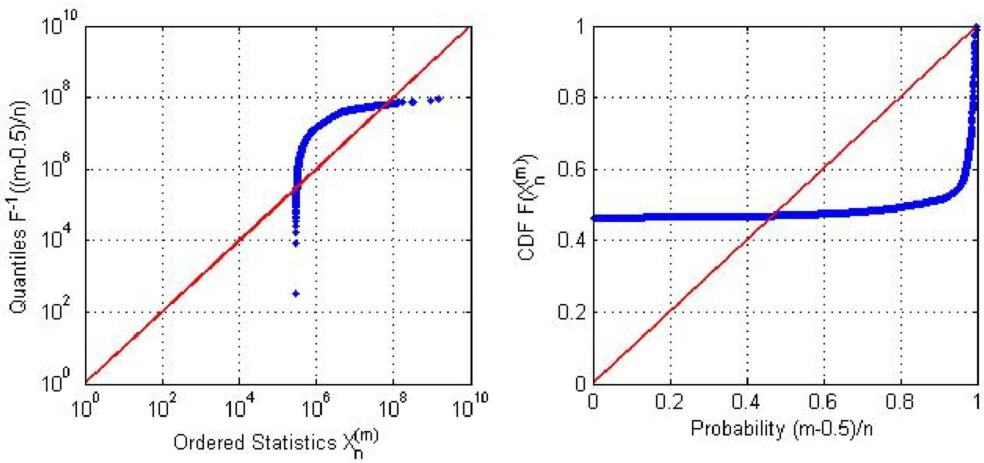

Figure 5

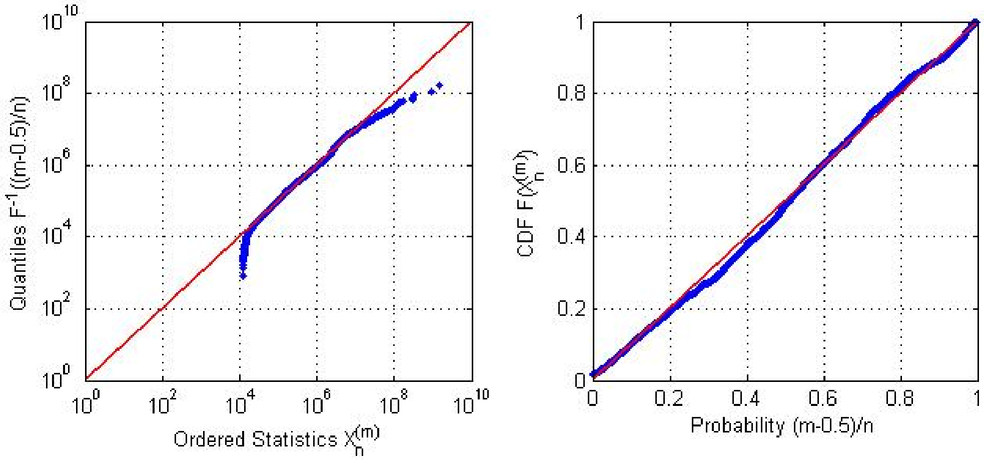

Figure 6

Due to prior knowledge of the data, we know there exists an "optional truncation" threshold in the data. Here, optional means that any losses below the truncation point H_c, are not mandated to be recorded. To adjust for this truncation problem, we first made the assumption that the data is not from a random sample of all operational losses, but instead is a biased sample containing a disproportionate number of large losses. The truncation point for the internal data is known and is constant H_c. This will affect the frequency distribution but the truncation will have little, or no effect, on our analysis of the severity since we expect our GPD threshold u to be greater than H_c.

When the data is combined from the different risk cells into the bankwide level, an implicit assumption that there is zero correlation between each business line and risk type is made. This assumption of no correlation will probably not hold in practice as there are many instances where losses can occur across multiple business lines and event types. When this occurs, the loss is further segregated into separate risk cells. However, it is understood that the resulting correlation will be minimal as the independence of other incurred losses is likely to override any existing correlation. In addition, we assume that there is no serial dependence among both the frequency or severity of losses.

The Poisson and Negative Binomial distributions will be used to model the frequency of losses. The peaks over thresholds method in extreme value theory will then be applied to the severity of the losses. To apply the loss distribution approach, the simplifying assumption that there is independence between frequency and severity is made. Finally, Monte Carlo simulations will be used to form the aggregated loss distribution.

As an extension to the modelling of the severity, we will attempt to apply a mixture distribution where the distribution is divided into two portions and modelled separately. The body will be modelled with the standard heavy tail LogNormal distribution and the tail with the GPD.

The idea of using a mixture distributions is also supported by the fact that there is an abundance of data in the body of the distribution (75% of the data consists of losses less than 10000) allowing us to apply more conventional statistical techniques.

It is impractical, and almost impossible, to measure the dependence structure for all 56 risk cells as specified under the Basel II guidelines. This is due not only to the difficulties in the mathematical modelling but also due to the lack of sufficient data in each of the risk cells. Laker [22] suggests mapping the eight business lines as devised by Basel II into only two – 'Retail and Commercial' and 'All Other Activities', reflecting the fact that Australian banks predominately operate in the Retail and Commercial sector . This not only has the added benefit of increasing the number of data points for modelling but it also reduces the number of dimensions we have to deal with in our analysis which will dampen the effects of the curse of dimensionality problem [5]. Thus, the dependence analysis is conducted by splitting the data up into a bivariate case consisting of two business lines only - 'Retail Banking' (RB) and 'All Others' (AO). We are unable to conduct the analysis in the way suggested by Laker [22] due to the limited amount of data available in the other business lines.

Copula theory will be applied to the aggregate losses of RB and AO. We will be unable to reliably fit the dependence structure of the aggregate losses on an annual basis as our data spans only over seven full years which effectively gives us seven aggregate losses. Hence, we will model our losses on a monthly basis and generate joint distributions which are on a monthly basis. We will apply copulas to the data in a multi-step process. First, the data is split up into their respective categories, then the marginal distributions for the aggregate losses is estimated using the same techniques in the bankwide case. The maximum likelihood estimation technique will then be used to fit the copula to the data. The dependence structure will be modelled empirically by using the empirical estimators for the rank correlations.

Empirical Results – Bankwide

We have assumed that the time aspect beyond inflation adjustments is negligible with no significant structural changes to the data over time [10]. In addition, initial analysis of the data suggests that the data is not affected by the optional truncation policy as over 60% of the data points are less than H_c. Hence, for the purposes of this data set, the analysis performed assumes that there does not exist a truncation bias in the data.

We have applied both the Poisson and Negative Binomial distributions to our data for illustration purposes. It is clearly evident that the Poisson distribution is inappropriate for modelling the frequency of losses. 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 as follows:

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 5

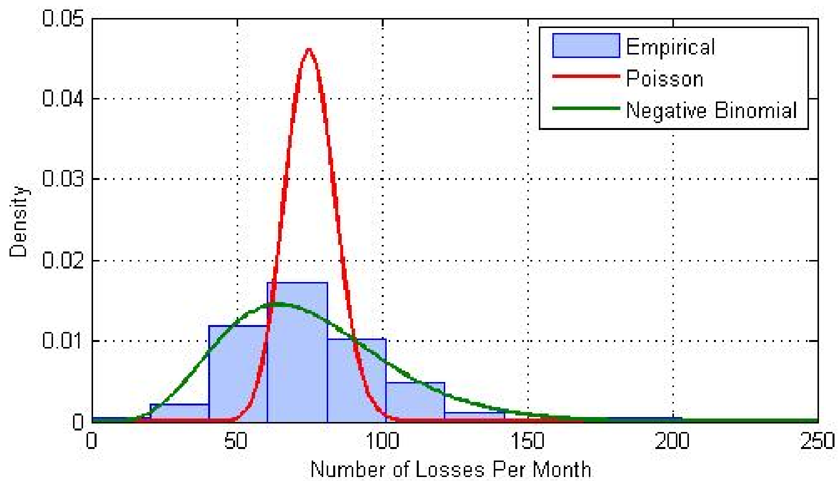

Figure 7

The peaks over thresholds methodology in EVT is used to model the severity of the losses. We begin by estimating the threshold parameter u. Initial analysis suggests that the value of u should be in the interval [1800000,3600000] as the shape parameter exhibits irregular jumps for values outside this interval. The plot of the empirical estimation for the mean excess function (Figure 8) appears to be fairly linear throughout suggesting that the data already conforms to a GPD. Owing to the need to satisfy the theoretical condition in the Pickands-Balkema-de Haan Theorem, we have analysed the mean excess plot for signs of non-linearity. There is an apparent kink at $u \approx 2900000$ after magnification of the plot. Hence, the analysis is based on values $Y = X - u = X - 2900000$, leaving less than 10% of the original data – representing approximately the 90^{th} percentile.

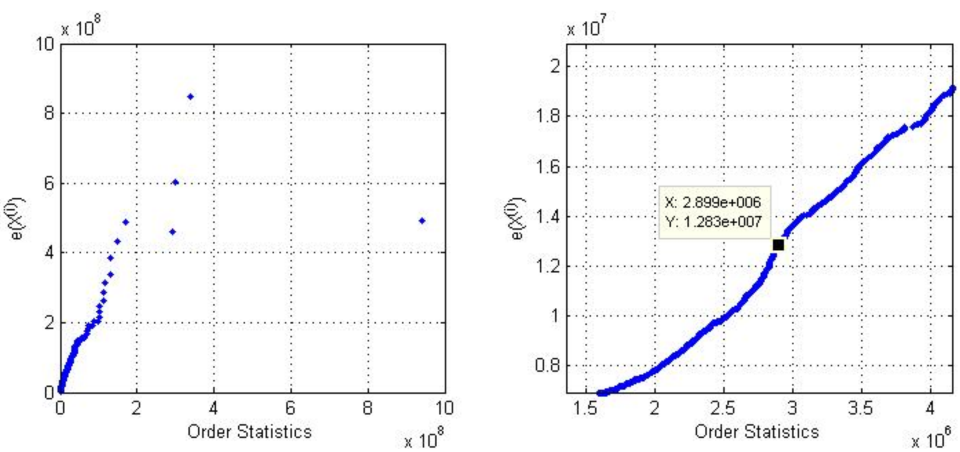

Figure 8

The estimated parameters for the GDP using both maximum likelihood estimation (here after $\mathbf{G}^{\text{MLE(S)}}$) and probability weight moments (here after $\mathbf{G}^{\text{PWM(S)}}$) are:

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

Table 6

From the QQ and PP plots below we can see that using GPD does indeed improve the fit for the loss data. The plots are fairly linearly and coincides well with the 45 degree line.

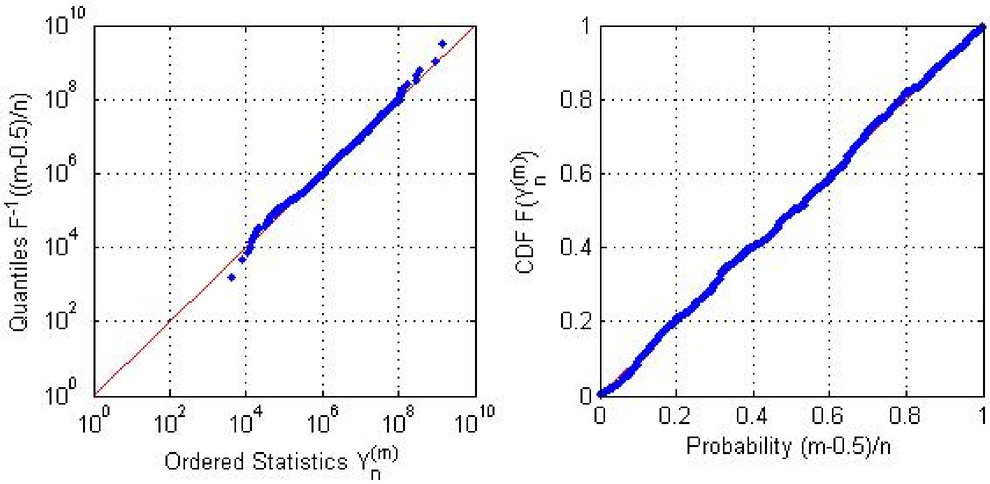

Figure 9: QQ-plot (left) and PP-plot (right) for the truncated severity data using the MLE parameters. The plots using PWM are similar.

We now explore the VaR performance of $\mathbf{G}^{\text{MLE(S)}}$ and $\mathbf{G}^{\text{PWM(S)}}$ along with the LogNormal distribution for comparison. The number of estimated violations from the model are compared with the expected number of violations. A violation occurs when the data value exceeds the calculated $\text{VaR}(\alpha)$ using the model in question. 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 the following table, it is clear that $\mathbf{G}^{\text{MLE(S)}}$ provides the best performance in terms of violations while the number of violations for the LogNormal severely increases as the confidence level increases. Both the GPD models passes 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.

		Number of Violations		
α	**Theoretical**	$\mathbf{G}^{\text{MLE(S)}}$	$\mathbf{G}^{\text{PWM(S)}}$	**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 7

The aggregate loss distribution is formed by simulating annual aggregate losses and then fitting these losses with an appropriate distribution. A hundred thousand simulations are used corresponding to a hundred thousand operation years for the bank. The frequency distribution used is the Negative Binomial with parameters as calculated previously. The severity is simulated using both sets of parameters $\mathbf{G}^{\text{MLE(S)}}$ and $\mathbf{G}^{\text{PWM(S)}}$. The simulations generated from $\mathbf{G}^{\text{MLE(S)}}$ and $\mathbf{G}^{\text{PWM(S)}}$ is denoted as $S(\mathbf{G}^{\text{MLE(S)}})$ and $S(\mathbf{G}^{\text{PWM(S)}})$, respectively. The statistical characteristics of the resulting simulation is shown below. The simulated data continues to shows clear evidence of skewness and kurtosis even on a logarithmic scale.

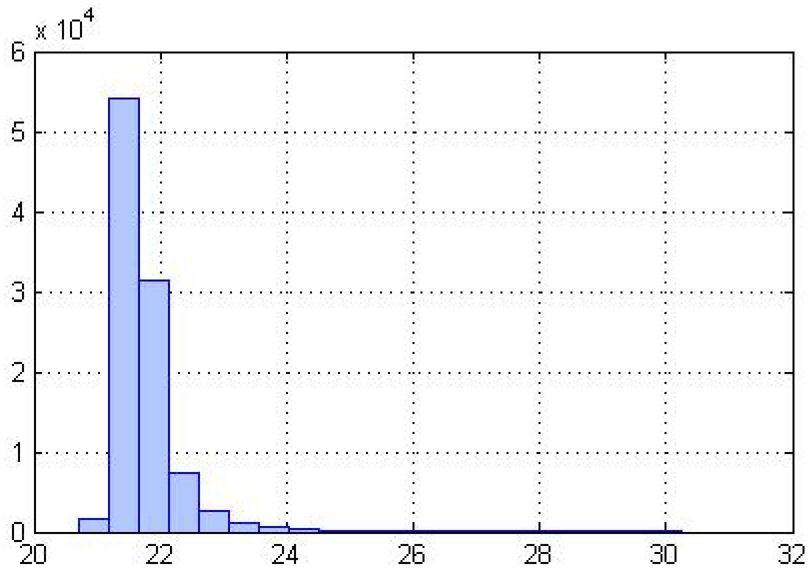

Figure 10: Histogram for the aggregate losses for the simulation using the MLE parameters $\mathbf{G}^{\text{MLE(S)}}$ on a logarithmic scale. The corresponding histogram for $\mathbf{G}^{\text{PWM(S)}}$ is similar.

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 for it to be realistic, especially given the size of the bank we have in question adjusted for scaling. 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. As an impartial sensitivity analysis on the parameters, various combinations of σ and ξ were used to perform the simulation. The size of the losses where particularly sensitive in the changes in ξ (e.g. increments of ± 0.05), on the other hand, even large changes in σ (e.g. ± 100000) 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 8

Both the GEV and GPD are fitted to the simulated loss data. The GEV distribution provides a much better fit than the GPD so we will only consider the GEV fit. Let $\mathbf{G}^{\mathrm{MLE(A)}}_{S(\mathbf{G}^{\mathrm{MLE(S)}})}$ and $\mathbf{G}^{\mathrm{PWM(A)}}_{S(\mathbf{G}^{\mathrm{MLE(S)}})}$ denote the GEV fit using MLE and PWM techniques to the simulated data set $S(\mathbf{G}^{\mathrm{MLE(S)}})$, respectively. Similarly, the notation for $S(\mathbf{G}^{\mathrm{PWM(S)}})$ fits are $\mathbf{G}^{\mathrm{MLE(A)}}_{S(\mathbf{G}^{\mathrm{PWM(S)}})}$ and $\mathbf{G}^{\mathrm{PWM(A)}}_{S(\mathbf{G}^{\mathrm{PWM(S)}})}$. Both MLE and PWM methods give 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 as it adheres to the empirical distribution better. 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 below. We conclude that the PWM method produces a better fit for the tail.

		Number of Violations			
α	Theoretical	$\mathbf{G}^{\mathrm{MLE(A)}}_{S(\mathbf{G}^{\mathrm{MLE(S)}})}$	$\mathbf{G}^{\mathrm{PWM(A)}}_{S(\mathbf{G}^{\mathrm{MLE(S)}})}$	$\mathbf{G}^{\mathrm{MLE(A)}}_{S(\mathbf{G}^{\mathrm{PWM(S)}})}$	$\mathbf{G}^{\mathrm{PWM(A)}}_{S(\mathbf{G}^{\mathrm{PWM(S)}})}$
0.95	05000	6168	3820	4833	3328
0.99	01000	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 9

It is clear that the tail of the aggregate losses is overestimated significantly. This problem is also highlighted by King [19]. The reason for such an occurrence is due to the use of a non-analytical technique for combining the frequency and severity distributions. The frequency for large losses is small but this rarity is not reflected in our simulation process. By not correcting for this we are effectively assuming that the occurrence of a loss is equally likely to be in the body or the tail of the severity distribution. For example, if we simulated a frequency of 100 losses for a particular year, these 100 losses are then randomly distributed over the severity distribution. In fact, as the number of simulation runs is increased, the amount of overestimation becomes more pronounced. In reality, most of these losses should be spread over the body while a small proportion should be placed in the tail. Thus, the ratio of small to large losses is not properly represented in the model.

The $\mathrm{VaR}(\alpha)$ is calculated using the fitted GEV distributions. Although the MLE fitted parameters appear to provide a better fit compared to the PWM fitted parameters, both sets of $\mathrm{VaR}(\alpha)$ figures will be included for the purpose of illustration. These can be seen in the table below. These results show that the $\mathrm{VaR}(\alpha)$ increases with the confidence level. In addition, the MLE parameters $S(\mathbf{G}^{\mathrm{MLE(S)}})$ produces significantly larger $\mathrm{VaR}(\alpha)$ than the corresponding PWM parameters $S(\mathbf{G}^{\mathrm{PWM(S)}})$.

	$\mathrm{VaR}(\alpha)$			
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}	$2.8702\mathrm{E} \times 10^{10}$	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 10

We have seen previously that if we do not make the distinction between the frequency of rare extreme losses and losses in the body of the distribution, the aggregate loss will be overestimated. A possible solution for this is the use of a mixture distribution 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 a mixture model 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.
Many difficulties were encountered in attempting to fit the mixture distribution. In our first attempt, MLE was used to simultaneously maximise the parameters 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. In the second attempt, the process is simplified by taking a multi-step approach. The data is split into two portions - smaller than threshold u_t (body X_b) and larger than u_t (tail X_t). See Figure 11 for a graphical representation.

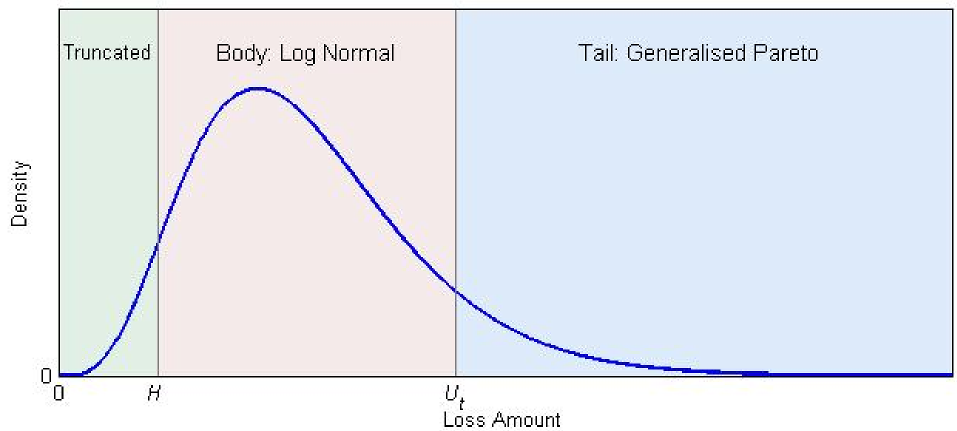

Figure 11

The losses X_b is fitted with the LogNormal distribution using MLE. Extreme value theory techniques applied to X_t. The log-likelihood is obtained for both fits. Let ℓ_m denote the log-likelihood function for the mixture model and let f_{LN} and f_{GPD} be the densities of the fitted LogNormal and GPD, respectively. The weighting factor w is varied to maximise

$$\ell_m = \sum \ln\left(wf_{\text{LN}} + (1-w)\,f_{\text{GPD}}\right).$$

The above steps are repeated for a series of thresholds and one which produces the largest log-likelihood is chosen. However, this method also failed to converge to a solution. The resulting weights were either close to zero or close to one.

As a last resort, an empirical based technique was used to implement the mixture distribution. We choose u to be the truncation point used in the bankwide analysis, that is, $u = 2900000$. The weighting factor will be the empirical estimation for the proportion that the loss will be less than u, that is, $w = \frac{\text{\# losses less than } u}{\text{total \# losses}}$. The LogNormal is estimated using MLE yielding parameters $\mu = 12.5693$ and $\sigma^2 = 1.3522$. The statistical characteristics of the mixed severity distribution are as follows:

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

Maximum	5.5123 x10^{11}	7.6331 x10^{11}

Table 11: Statistics for the aggregate loss simulation produced from the mixture severity distribution.

The use of the mixture distribution produces less extreme losses for the aggregate distribution. This is evident when Table 8 and Table 11 are compared. The maximum losses and variance are significantly smaller for the mixture distribution case. This is especially true for the simulations using $\mathbf{G}^{\text{MLE(S)}}$ parameters where the maximum loss and variance is reduced by a factor of 30 and 1400, respectively. Furthermore, the mean and median remain stable for both methods once again reflecting the reduction in extremes but giving similar aggregate losses on average. The changes for the MLE are more pronounced indicating that $\mathbf{G}^{\text{MLE(S)}}$ generates larger extremes than $\mathbf{G}^{\text{PWM(S)}}$. The GEV distribution once again generates a better fit. Although the mixture model clearly produces different results, the parameter estimates are very similar.

Capital-at-Risk

Under the Basel II requirements, our bankwide $\text{CaR}(\alpha)$ is equal to the bankwide $\text{VaR}(\alpha)$ assuming comonotonicity holds. That is, the $\text{VaR}(\alpha)$ values of individual business lines simply added together. The table below lists these aggregate $\text{VaR}(\alpha)$ values generated by summing over the $\text{VaR}(\alpha)$ produced from the chosen marginal models (GEV_{RB} and GPD_{AO}). The $\text{VaR}(\alpha)$ for the copulas are calculated from simulating the quantiles of the fitted copula and then taking the upper α percentile. At 95% there does not seem to be much difference in the calculated $\text{VaR}(\alpha)$. However, as the confidence level is increased, the differences in $\text{VaR}(\alpha)$ become slightly more apparent. At the 99.9% level, we are able to reduce the $\text{VaR}(\alpha)$ by over $1.3 billion. This represents a major difference between the two calculated values which gives rise to large diversification benefits.

α	$\text{VaR}_{L_{RB}} + \text{VaR}_{L_{AO}}$	$\text{VaR}(L_{RB}+L_{AO})$ **Gaussian**	**t-**	**Frank**
0.95	4.3687 x 10^8	4.4722 x 10^8	4.4840 x 10^8	4.4604 x 10^8
0.99	1.0054 x 10^9	9.3739 x 10^8	9.3826 x 10^8	9.4004 x 10^8
0.995	1.6241 x 10^9	1.3740 x 10^9	1.3796 x 10^9	1.3802 x 10^9
0.997	2.4973 x 10^9	1.9627 x 10^9	1.9829 x 10^9	1.9869 x 10^9
0.999	7.9342 x 10^9	6.3195 x 10^9	6.6082 x 10^9	6.2130 x 10^9

Table 12

Conclusion

The paper describes some well developed modelling approaches from statistics and actuarial mathematics that can be adapted to operational risk data. However, due to the unique nature of the data which is compounded by the lack of a large sample of loss data, a direct application of any technique described is arguably inappropriate. Indeed, caution needs to be exercised in the interpretation of the results on both the statistical and intuitive level.

Extreme value theory has demonstrated enormous potential to account for the heavy tail of operational losses where other conventional methods fall short. 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 while conventional models place emphasis on fitting the central or body of the data, and thus, neglect the extreme percentiles. On the contrary, extreme value theory has delivered promising results which is fruitful for further research.

However, a major limitation in the implementation of extreme value theory is the lack of data which inhibits us from fully capturing the generalised Pareto nature of the excess distributions without sacrificing the majority of our data set. Our ability to model any sort of dependence is also limited by the availability of quality data. Even if we can overcome these limitations, the regulators may not permit the use of models based on such a small sample despite the accuracy of our 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.

Our results demonstrate tremendous differences in comparison to studies conducted overseas in terms of the empirical analysis, features and characteristics of the data.

The value-at-risk amounts are much smaller even when compared to a medium-sized non-internationally active U.S. bank [24]. ADIs in Australia have a tendency to hold higher proportions of residential mortgage loans in their accounts than most overseas banks. Hence, Australian banks are expected to be less risky than equivalent overseas banks [9], and consequently, reduce the need to hold large capital reserves.

Other studies which have been conducted use external databases supplied by vendors such as OpRisk Analytics and OpVantage [8]. These databases contain much larger losses, and therefore, it follows that it would be unreasonable for us to make a comparison with these results. The analysis performed by de Fontnouvelle et al. [8] indicate 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 certainly 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 statistical properties inherent in the Poisson distribution rather than the ability to produce a better fit. Even so, the Poisson has proved to be exceptionally inappropriate for our data set due to the greater spread. The area where most studies and banks agree on is the modelling of the severity with the generalised Pareto distribution. Even though different truncation values are used in each study, the tail index (reciprocal of the shape parameter ξ) is always roughly similar ranging from values of 0.8 to 1.2.

References

1. 2 J. Aparicio and E. Keskiner, *A Review of Operational Risk Quantitative Methodologies within the Basel II Framework*, 2004. http://www.gloriamundi.org/picsresources/jaek.pdf. Last Visited: 13/10/2006.
2. 4 Australian Bureau of Statistics, *Consumer Price Index*. http://www.abs.gov.au. Last visited: 23/09/2006.
3. 17 N. Baud, A. Frachot and T. Roncalli, *Internal Data, External Data and Consortium Data for Operational Risk Measurement: How to Pool Data Properly?*, 2002 http://gro.creditlyonnais.fr/content/fr/home_presentation.htm. Last visited: 13/10/2006.
4. 18 N. Baud, A. Frachot and T. Roncalli, *How to Avoid Over-estimating Capital Charge for Operational Risk*, OperationalRisk – Risks Newsletter, 2003.
5. 21 R. Bellman. *Adaptive Control Processes*, Princeton University Press, Princeton, NJ, 1961.
6. 26 V. Chavez-Demoulin and P. Embrechts, *Advanced Extermal Models for Operational Risk*, 2004. http://www.math.ethz.ch/~baltes/ftp/papers.html. Last visited: 13/10/2006.
7. 34 M. G. Cruz, *Modelling, Measuring and Hedging Operational Risk*, John Wiley & Sons Ltd, England, 2005.
8. 41 P. de Fontnouvelle, V. DeJesus-Rueff, J. Jordan and E. Rosengren, *Using Loss Data to Quantify Operational Risk*, Apr 2003. http://www.bis.org/bcbs/events/wkshop0303/p04deforose.pdf, Last visited: 13/10/2006.
9. 45 B. Egan, *The Basel II Capital Framework in Australia*, APRA Insight, 4th Quarter, 2004.
10. 47 P. Embrechts, H. Furrer and R. Kaufmann, *Quantifying Regulatory Capital for Operational Risk*, Derivative Use, Trading & Regulation, 9, 2003, pp. 217-233.
11. 48 P. Embrechts, R. Kaufmann and G. Samorodnitsky, *Ruin Theory Revisited: Stochastic Models for Operational Risk*, in Risk Management for Central Bank Foreign Reserves, C. Bernadell et al., ed., European Central Bank, Apr 2004, pp. 243-261.
12. 50 P. Embrechts, F. Lindskog and A. McNeil, *Modelling Extremal Events for Insurance and Finance*, Springer, 1997.
13. 56 Federal Reserve System, *The 2002 Loss Data Collection Exercise for Operational Risk: Summary of the Data Collected*, May 2003. http://www.bos.frb.org/bankinfo/qau/papers/pd051295.pdf. Last visited: 13/10/2006.
14. 58 A. Frachot, P. Georges and T. Roncalli, *Loss Distribution Approach for Operational Risk*, Apr 2001. http://gro.creditlyonnais.fr/content/fr/home_presentation.htm. Last visited: 13/10/2006.
15. 59 A. Frachot, O. Moudoulaud and T. Roncalli, *Loss Distribution Approach in Practice*, in The Basel Handbook: A Guide for Financial Practitioners, M. Ong, ed., Risk Books, 2004.
16. 60 A. Frachot, T. Roncalli and E. Salomon, *The Correlation Problem in Operational Risk*,, OperationalRisk – Risks Newsletter, 2004.

17. 73 JP Morgan and Reuters, *RiskMetrics*^{TM} *– Technical document*, New York, Dec 1996.
18. 74 R. Kaas, M. Goovaerts, J. Dhaene and M. Denuit, *Modern Actuarial Risk Theory*, Kluwer Academic Publishers, Boston, 2001.
19. 76 J. L. King, *Measurement and Modelling Operational Risk*, Wiley Finance, May 2001.
20. 77 S. A. Klugman, H. H. Panjer and G. E. Willmot, *Loss Models: From Data to Decisions*, John Wiley & Sons Inc, United States of America, 1997.
21. 78 S. A. Klugman, H. H. Panjer and G. E. Willmot, *Loss Models: From Data to Decisions*, John Wiley & Sons Inc, United States of America, 2004.
22. 85 J. F. Laker, *Basel II – Observations From Down Under*, Apr 2006. http://www.apra.gov.au/Policy/. Last visited: 13/10/2006.
23. 95 A. J. McNeil, R. Frey and P. Embrechts, *Quantitative Risk Management*, Princeton University Press, Princeton, New Jersey, 2005.
24. 97 M. Moscadelli, *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, 2004.
25. 117 P. Sprent, *Data Driven Statistical Methods*, Chapman & Hall, London, 1998.