Using ELD: The Australian Experience

Harvey Crapp

Australian Prudential Regulation Authority

Bank of Japan

19 March 2008

Agenda

1. **Importance of ELD**

2. **Sources of ELD**
   - Publicly Available
   - Consortium
   - Insurance
3. **Biases**
   - Reporting
   - Control
   - Scale
4. **AMA Incorporation**
   - Integration
   - Model Independently
   - Implicitly Incorporate
   - Uncertainty

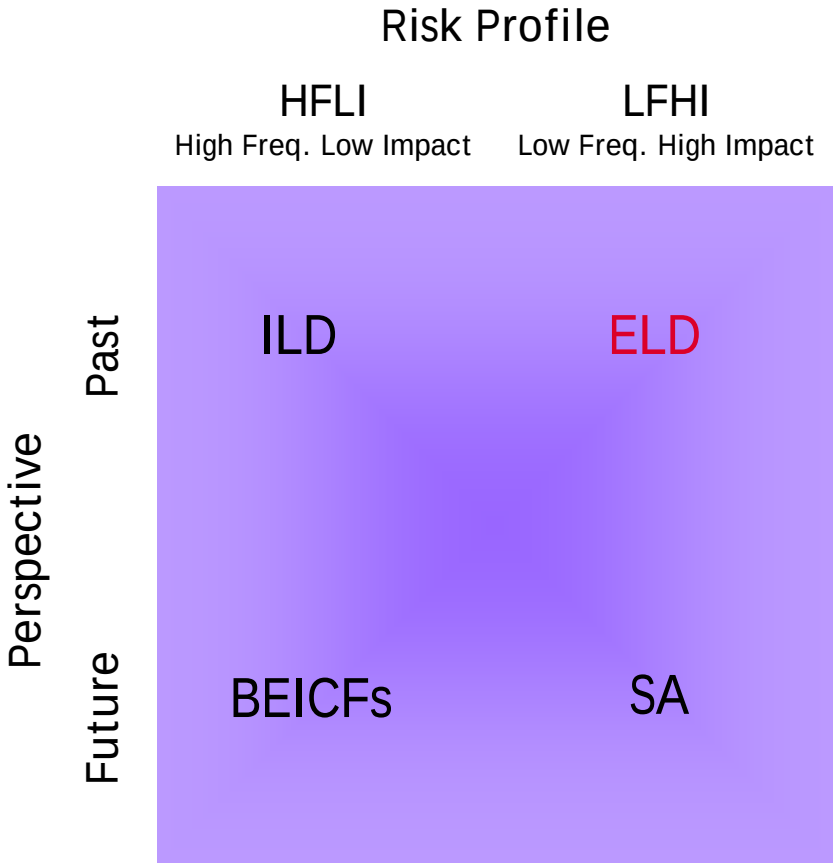
Importance of ELD

- **Management**

  Relevant ELD offers a forward looking perspective to a bank's loss profile. For example, if a bank has a probable risk of an event but is yet to experience a significant loss; then ELD may be useful in developing Key Risk Indicators (KRIs) or other BEICFs.

- **Measurement**

  Due to the paucity of internal databases, ELD is often used to supplement a bank's internal loss experience with the "infrequent yet potentially severe operational risk loss events" not normally experienced in a bank's loss history.



Sources of ELD

- In Australia, AMA accredited banks have predominately used at least one of public, consortium or insurance external loss databases.
  - **Publicly Available Data**: Collected by specialist institutions, these databases contain of operational risk losses reported in newspapers, press releases etc. Data is generally collected above a common threshold.

Pros

- Provides background information about the bank which experienced the loss. This information is invaluable during the modelling process as it allows for meaningful selection of the loss and for scaling.

Cons:

- Public databases generally only contain very large losses which are likely to attract the attention of media and shareholders.
- Certain risk types seem not to be adequately represented in the dataset e.g. Execution, Delivery and Process Management.
- Losses sourced from public data may also be subject to rounding bias due to the imperfect information available in the media.

- **Consortium Data:** These databases contain non-public operational risk loss data sourced from participating financial institutions. Consortium facilitators compile data from member institutions and provide anonymous statistical analysis and basic loss information.

Pros

- Consortium data spans a wider array of loss severities and risk types than other types of external data because it is taken directly from the members' loss database. As such, consortium databases are seen as more complete than other types of ELD.

Cons

- To ensure the anonymity of the consortium members, each loss has had any identifying information removed, which makes meaningful analysis and scaling of the data difficult.
- Some consortiums do not allow event-by-event access to the losses and only provide summary statistics to the participating institutions.

Sources of ELD

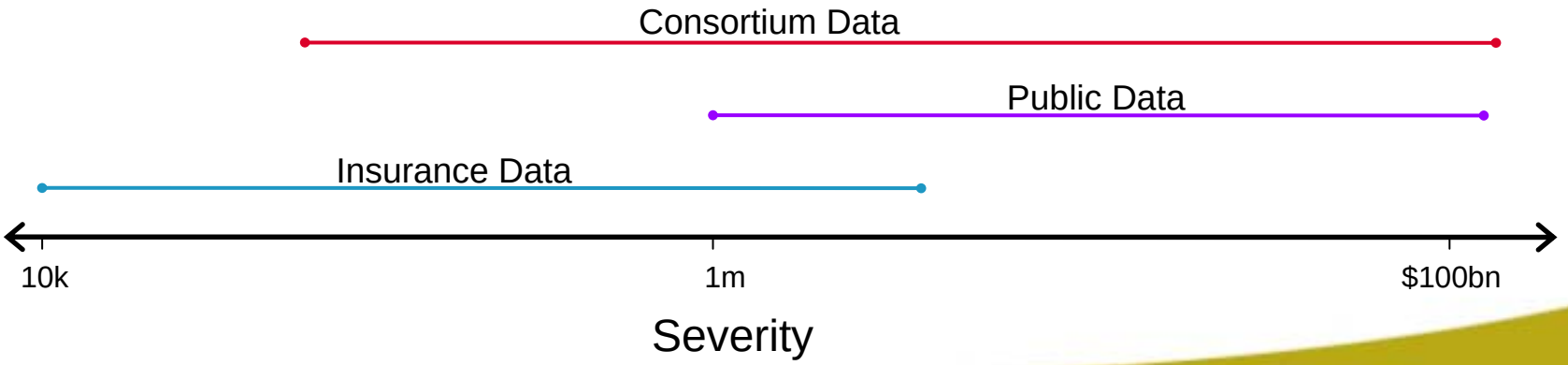
- **Insurance Data:** Provided by insurance brokers, the data originates from operational risk related insurance claims. Most insurance databases are structured similarly to consortium databases.

Pros

- The severity range of insurance data is generally lower than that of other ELD sources because of the limited range of policies available in the market. This can be beneficial to banks who have limited internal loss data.

Cons

- The risk type and severity coverage of insurance data depends on the range of policies held by institutions, the deductibles taken and will only include those risk types which are insurable.



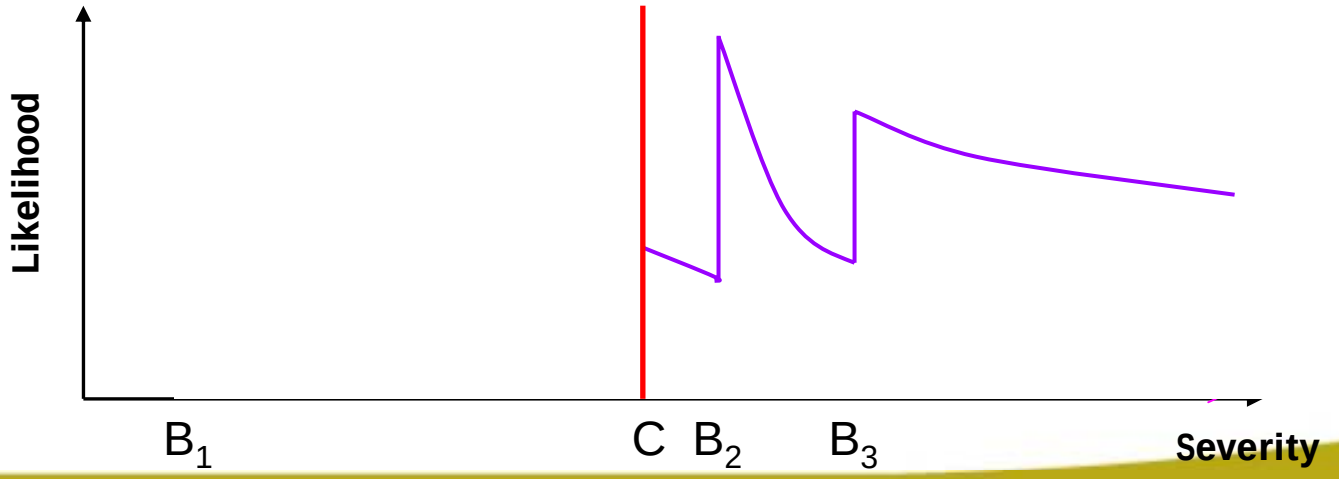
Inherent Biases: Reporting Bias

Due to the nature of its collection, all external databases are subject to inherent bias. These biases hamper the explicit use of ELD in the AMA operational risk measurement process.

- **Reporting Bias**

Affecting all data types, reporting bias occurs when the loss data in the external database is not considered a random sample of the population of data.

- **Consortium Data** – Currently, it is not possible to observe whether an institution has recorded and submitted losses above or below the consortium threshold (C) because individual banks are not identified in the database. Banks with differing thresholds (B_1, B_2, B_3) may distort the quantiles of the aggregate loss profile of the consortium.



Inherent Biases: Reporting Bias

- **Insurance Data** – Reporting bias occurs in insurance data because the probability that the loss is claimed for, and thus included in the data set, depends on the size of the deductible and the type of insurance policy held. As such the issues are similar to consortium databases.

Overcoming Reporting Bias

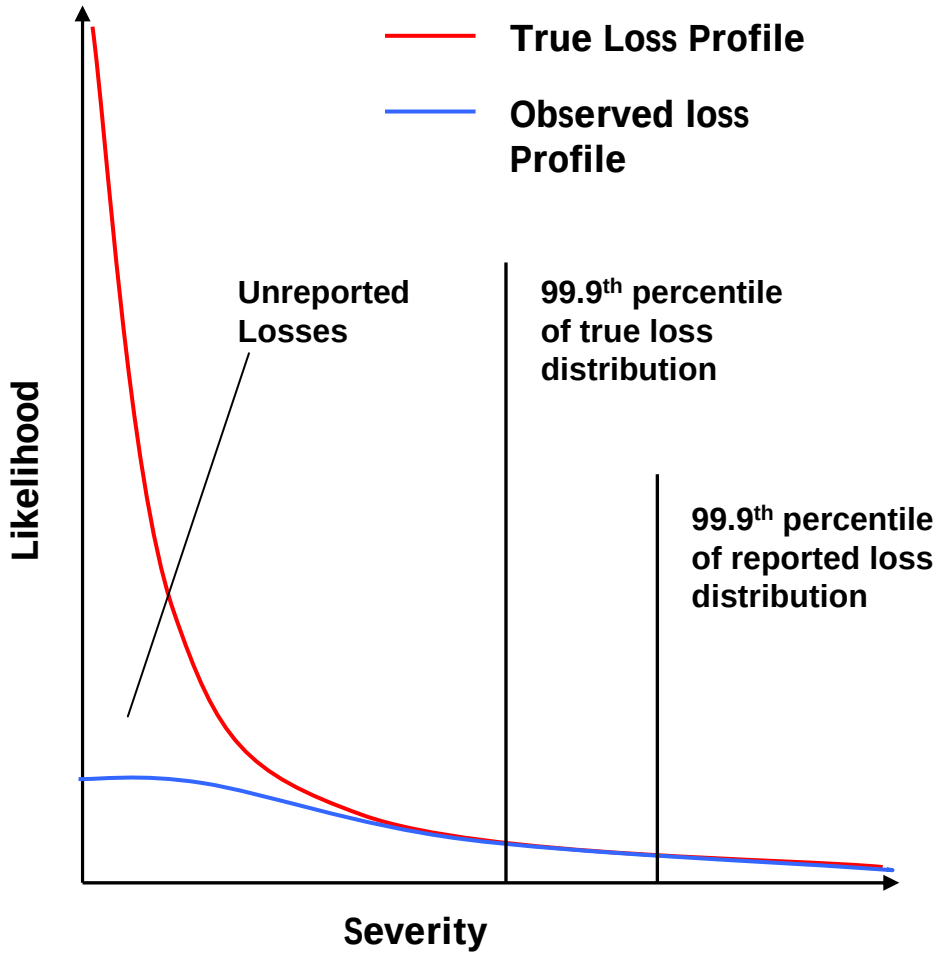
- Baud, Frachot and Rancalli developed a method whereby the threshold of each contributing institution is assumed to be random and can be estimated via a probability distribution.
- The method was initially devised using consortium databases, however it could easily be extended to apply to insurance data sets and even public datasets
- A major drawback in their method is that it assumes ELD is drawn from the same distribution as ILD, but is truncated above some threshold. Because this assumption is not statistically true for all datasets, results become spurious if the assumption is violated.

Inherent Biases: Reporting Bias

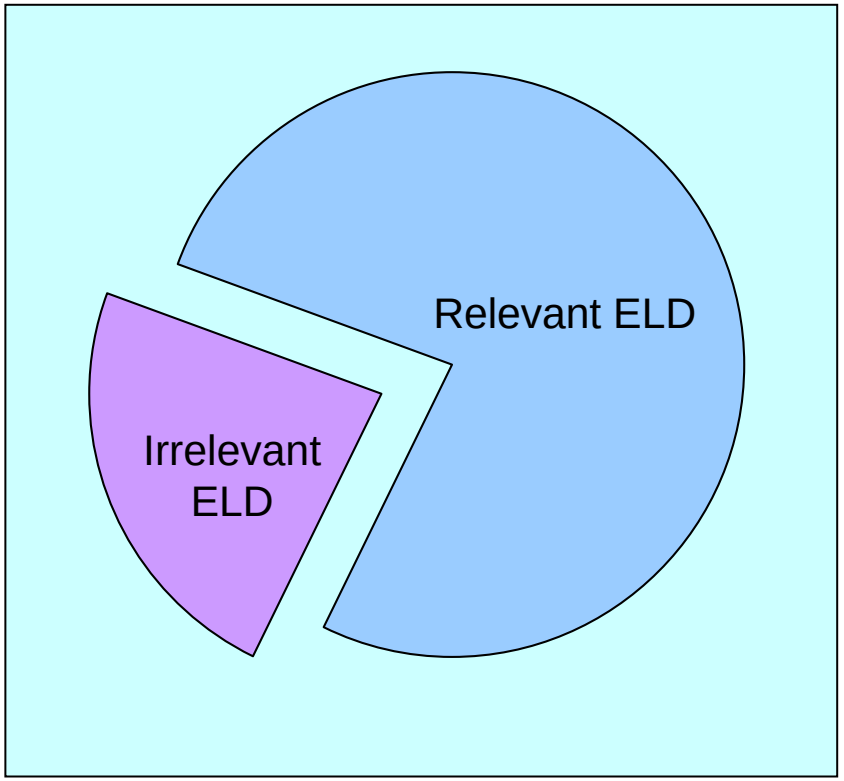
Public Data - Due to factors such as the size and nature of the loss, larger losses tend to be reported in Public Loss databases. The probability of a loss being reported increases with the severity of the loss and as such, the number of larger losses tends to be overstated relative to smaller losses.

Overcoming Reporting Bias

- Correction methods have been proposed whereby subject matter experts are asked to estimate the extent to which reporting bias appears in the dataset.
- Estimating such information is difficult even for the most experienced expert.



Inherent Biases: Control Bias



- **Control Bias**

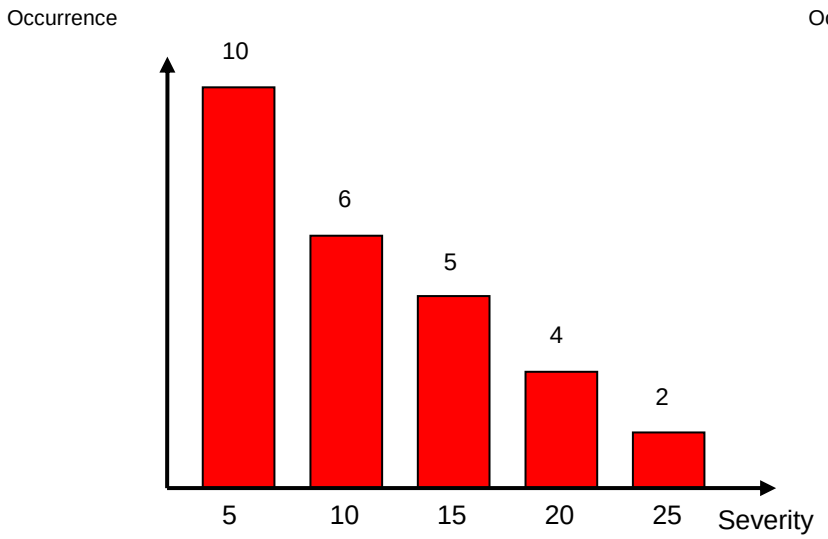
  Control bias refers to the relevance of losses that come from banks with different control mechanisms. All losses arise as a consequence of a specific set of circumstances, due to a lack of, or failure in controls and as such not all losses will be relevant to all banks. Banks can filter the data to obtain a subset of the external data that is relevant to their own operations and control structure.

  To eliminate control bias, banks need to identify which businesses, in which organisations have similar control structures to their own, and eliminate those which do not meet certain criteria. This is generally not possible with current data sources.

Overcoming Control Bias

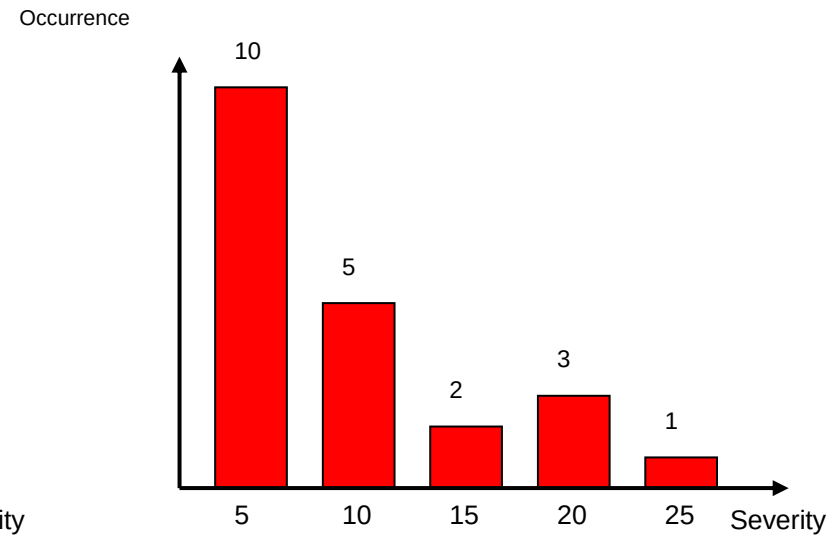
Khan suggests that institutions should not 'cherry pick' losses from relevant subsets of data on the basis that they are deemed more relevant than others, because one cannot know in fact if one loss is really more relevant than another. A better approach would use all relevant external data, so that the data, in the context of a distribution can explain the relative probabilities associated with each loss.

Figure 1



Loss proportions before cherry picking

Figure 2

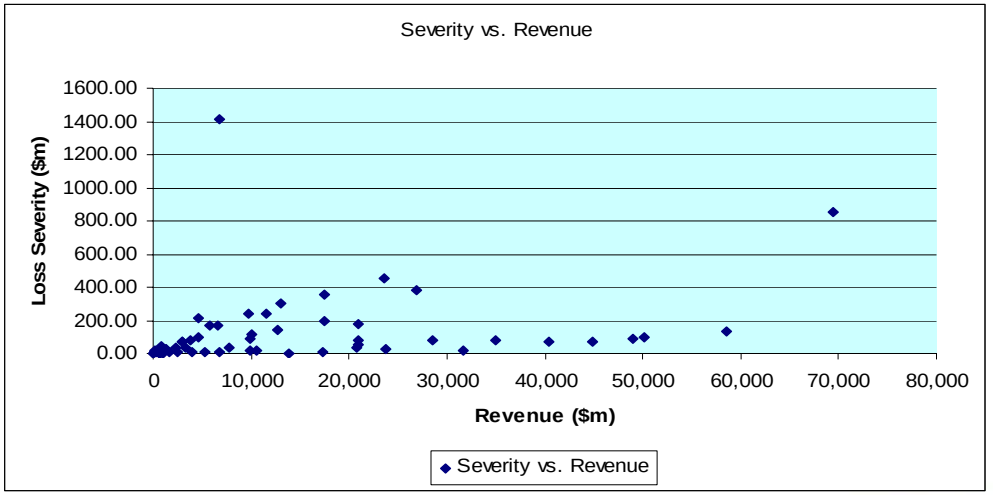


Loss proportions after cherry picking

- **Scale Bias**

  Scale bias exists in all external loss databases, and occurs because the losses recorded within each database come from banks of varying magnitudes (in terms of their operations, assets, number of employees, revenue etc.). As a result many banks use a severity scaling mechanism to scale loss values up or down in relation to some proxy.

  Studies have shown that a relationship exists between the size of a firm and the severity of the loss. However, banks have had difficulty determining a proxy which adequately describes the relationship between the two.



An example of the relationship between loss severity and income in a sample of ELD

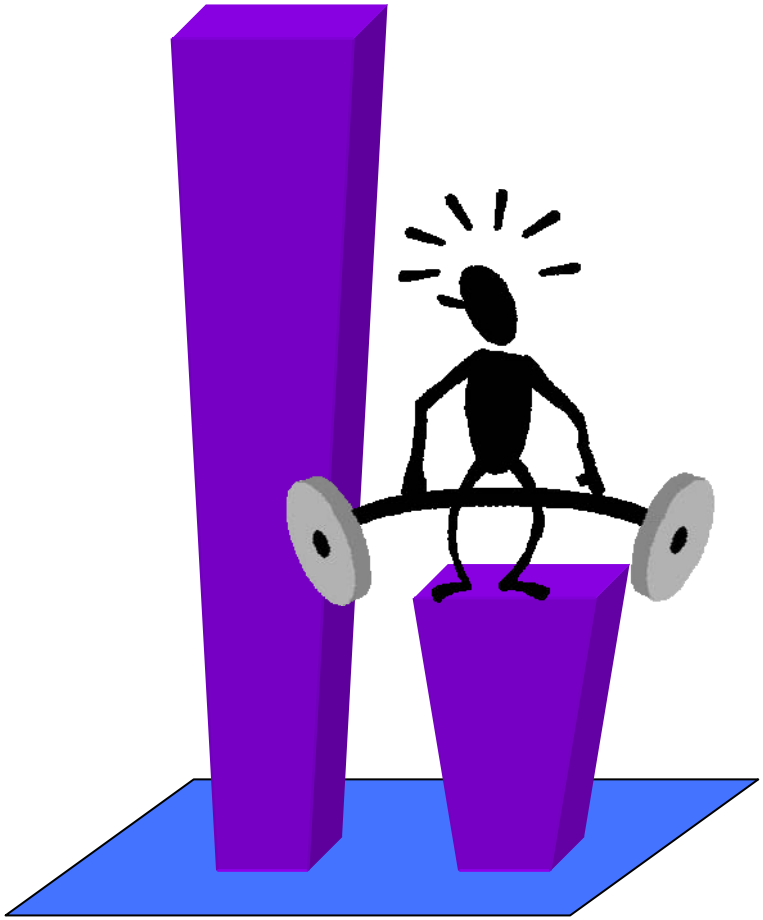
Overcoming Scale Bias

Recent studies have tried to incorporate other factors into the scaling mechanism. Na suggests that an operational risk loss can be broken up into two parts:

- A **common** component which explains the macroeconomic, geopolitical and cultural environment.
- An **idiosyncratic** part which captures the risk and control environment of individual banks.

Using the quotient of some function of their idiosyncratic parts, a suitable scaling proxy can be developed such that a loss occurring in bank X can be scaled to its equivalent in bank Y.

$$Loss_Y = \frac{f(Component_{Idio})_Y}{f(Component_{Idio})_X} \times Loss_X$$

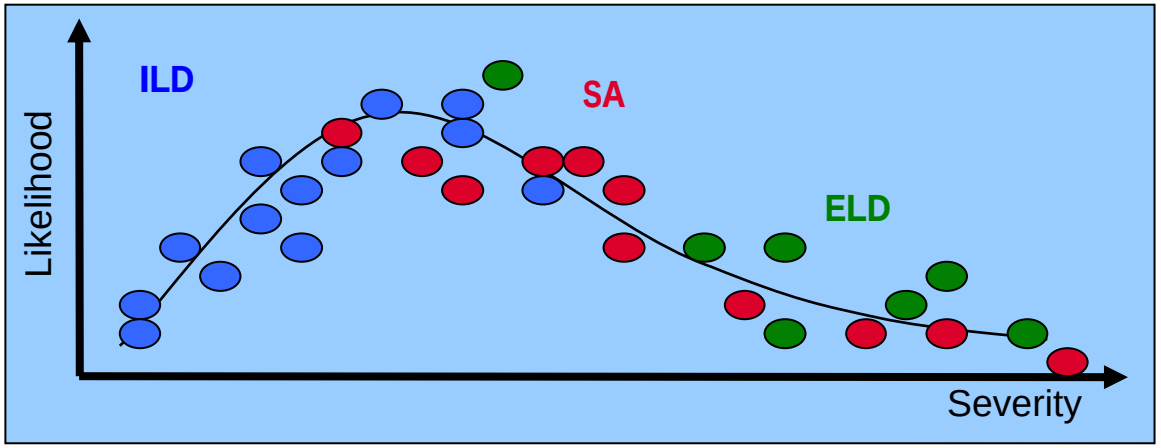


Incorporation into AMA

Accredited AMA Australian Banks have generally used one of three methods to incorporate ELD into the AMA process.

- **Integration**

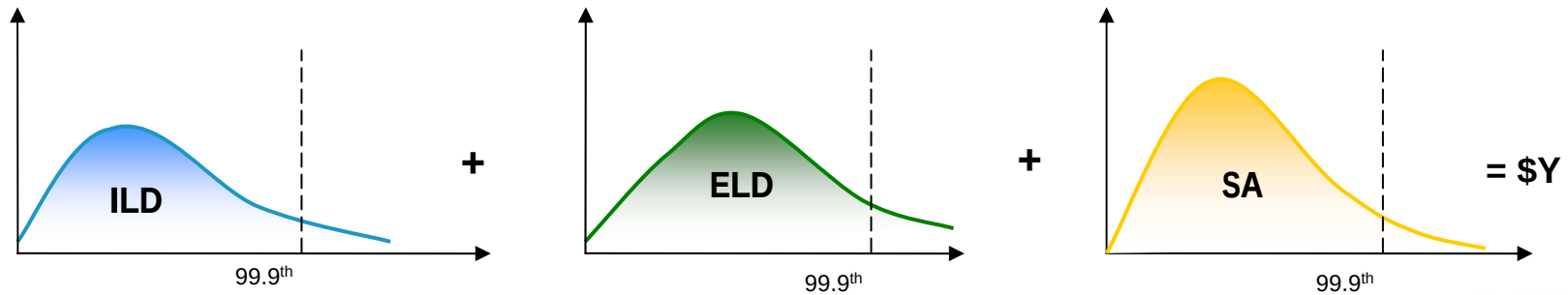
  Integration is the pooling of ELD with one or more other data sources. The pooled data is then used as if the data points had originated from the same source. In theory, this approach would work if the bank had perfect data. Perfect data would mirror the bank's loss profile and give a regulatory capital figure that is commensurate with their risk exposure. In reality, data is not perfect and by combining imperfect data together, banks may compound the problems of the individual data sources and find it difficult to measure and mitigate the uncertainty created by this method.



- **Model Independently**

  ELD can be modelled independently from other data sources to obtain a separate ELD 'component' of regulatory capital. This component is then weighted and combined with the capital derived from other data components to obtain a final capital charge. However, to obtain a robust capital estimate an adequate amount of data points are required for each data source and as such this option may not be viable for all institutions.

  Adjustments to ELD may be needed to ensure that the capital charge from its component is not overestimated. Because ELD tends to be biased towards high severity losses and certain risk types, banks need to ensure that all loss sizes and risk types are adequately represented in the final capital requirement by means of the other data sources.



Incorporation into AMA

- **Implicitly Incorporate**

  Banks not wishing to use ELD explicitly in their operational risk measurement model, perhaps due to the uncertainty created by the inherent biases, have instead used ELD as a reference source in the formation of dependence structures and in scenario analysis workshops.

Scenario Analysis

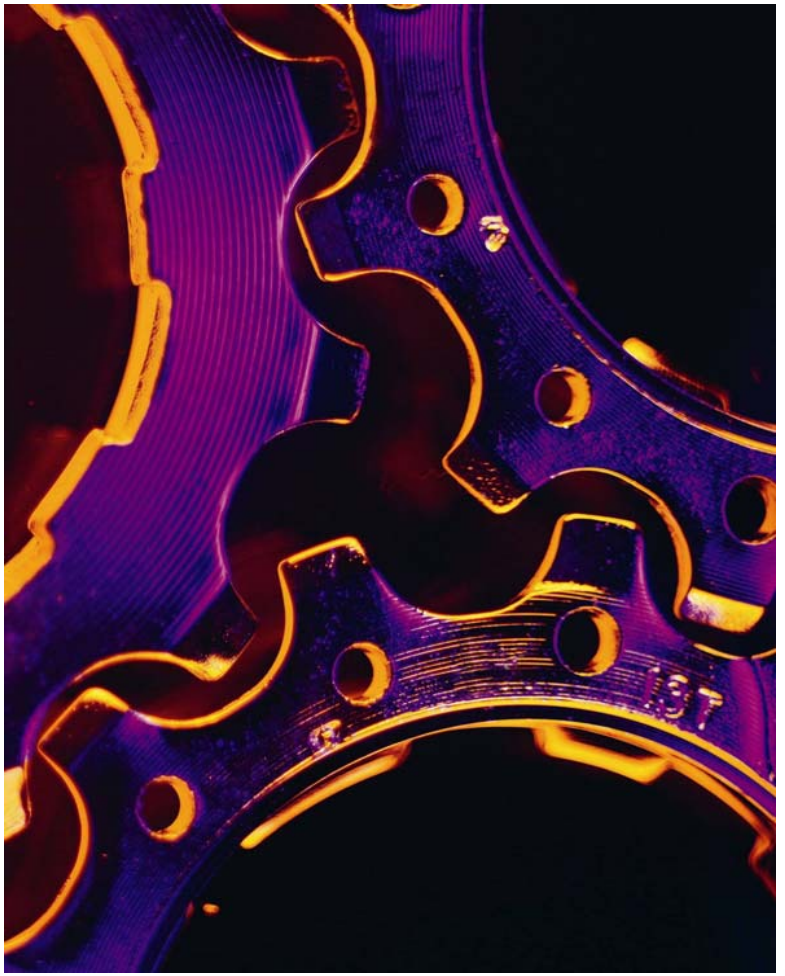
ELD provides indications of both the severity and frequency of extreme losses, for which a bank has little of their own data. Generally, the ELD is cleansed and the relevant losses are incorporated into an information pack (with other relevant ILD and BEICFs) and distributed to subject matter experts to facilitate the scenario elicitation process.

Dependence Structures

Rather than assuming zero correlation between losses (i.e. there is no connection between different risk type and business line combinations), some banks have used ELD to develop dependency structures to better enhance the connection between risk capital and loss profile.

ELD is useful for this process because it generally covers a wider range of risk type/business line combinations than the banks internal data. However, problems may arise if there is no clear mapping between internal and Basel business lines.

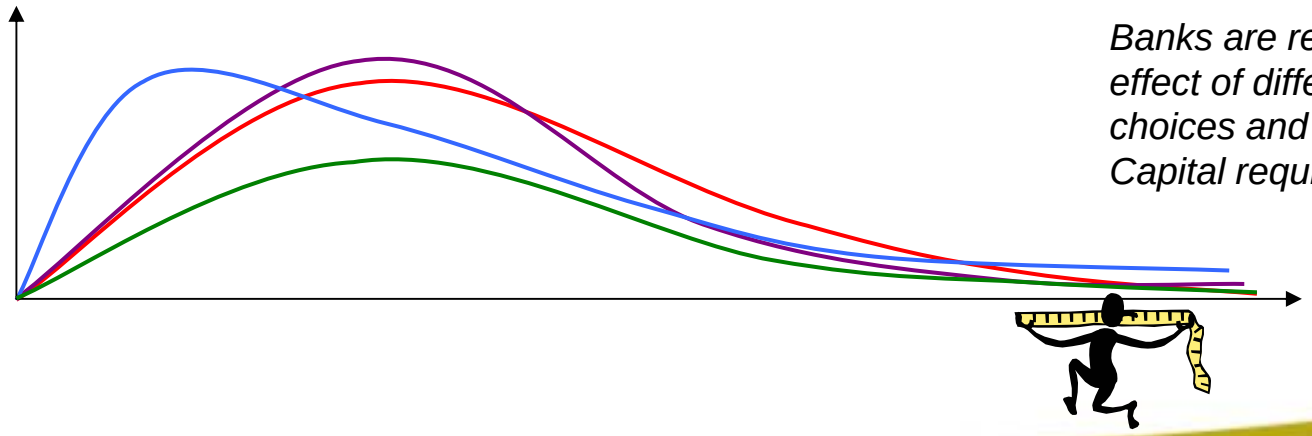
By incorporating the data in an informational capacity only, the inherent biases are not directly transferred to the capital outcome.



- **Uncertainty**

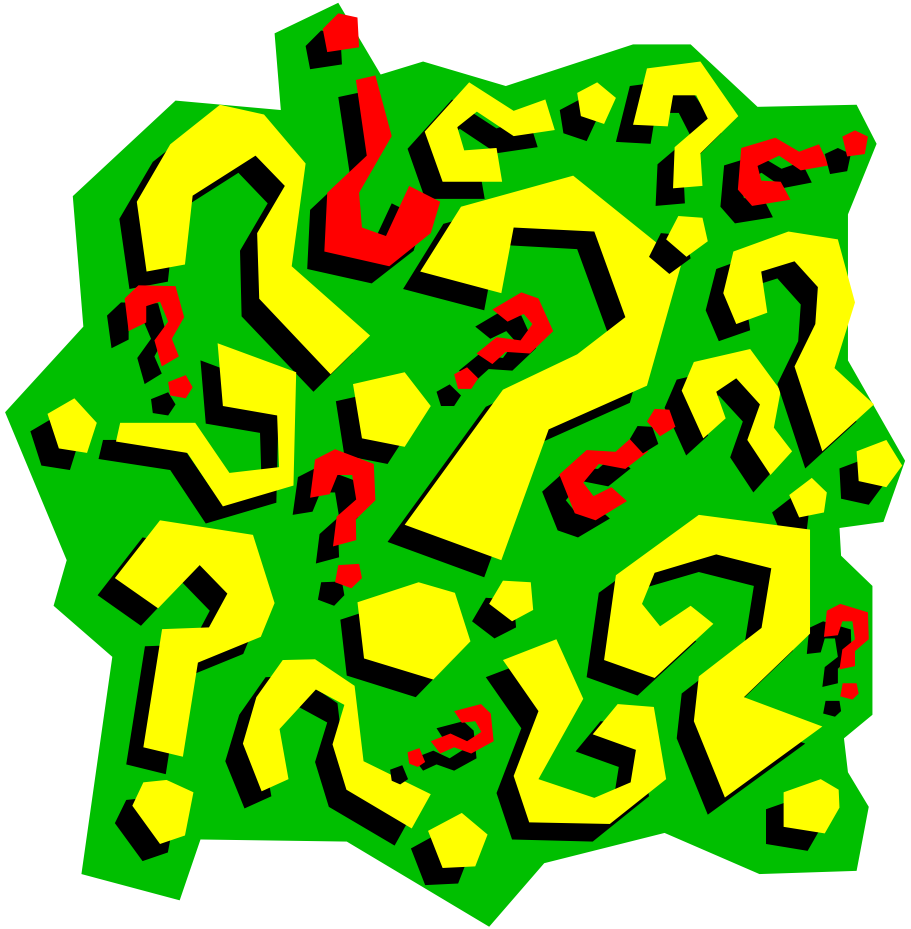
Given the inherent uncertainty created by explicitly including ELD into the operational risk modelling process, banks must ensure that they are conservative in their modelling choices and assumptions and give consideration to the results of sensitivity analysis surrounding external data. Any model uncertainty created through the use of ELD (and other data points) must be identified and mitigated with commensurate conservatism to the model inputs, outputs and/or calculation.

Although using ELD implicitly in the calculation process does not directly transfer the uncertainty created by the biases in ELD, a high degree of rigour is expected from banks when incorporating ELD into their scenario analysis, dependency structures etc.



Banks are required to measure the effect of different modelling choices and assumptions on Capital requirement

Summary of Key Issues



- Australian AMA banks use three sources of external data; Consortium, Publicly Available and Insurance Data.
- ELD is used for both operational risk management and measurement purposes.
- Unless corrected, the biases in ELD transfer to biases in the calculated risk capital.
- Correction techniques exist for some biases in ELD, however research is still being conducted in this area.
- Model uncertainty created through the use of ELD (and other data points) must be identified and mitigated with commensurate conservatism

References

- Australian Prudential Regulation Authority (APRA) (2008) 'Prudential Standard APS 115 Capital Adequacy: Advanced Measurement Approaches to Operational Risk', January 2008.
- Australian Prudential Regulation Authority (APRA) (2008) 'Prudential Practice Guide APG 115 - Advanced Measurement Approaches to Operational Risk', January 2008.
- Baud, N., Frachot, A and Roncalli, T. (2002) 'Internal data, external data and consortium data for operational risk measurement: How to pool data properly?' Groupe de Recherche Opérationnelle, Crédit Lyonnais, France, June 2002.
- Cagen, P. (2005) 'External Data: Reaching for the Truth' Algorithmics Incorporated, December 2005.
- Dahen, H. and Dionne, G. (2007) 'Scaling Methods for Severity and Frequency of External Loss Data', Working Paper 07-01- Canada Research Chair in Finance, January 2007.
- De Fontnouvelle, P., De Jesus-Rueff, V., Jordan, J. and Rosengren E. (2003) 'Using Loss Data to Quantify Operational Risk', Federal Reserve of Boston, April 2003
- Frachot, A. and Roncalli, T. (2002) 'Mixing internal and external data for managing operational risk', Groupe de Recherche Opérationnelle, Crédit Lyonnais, France, January 2002.
- Frachot, A., Roncalli, T. and Salomon, E. (2004) 'The Correlation Problem in Operational Risk' Groupe de Recherche Opérationnelle, Crédit Lyonnais, France, January 2004.

References

- Gustafsson, J., Nielsen, J.P., Pritchard, P. and Roberts, D. (2006) 'Quantifying operational risk guided by kernel smoothing and continuous credibility: A practitioner's view', The Journal of Operational Risk 1(1): 43-57.
- Gustafsson, J., Guillen, M., Perch Nielsen, J. and Pritchard, P (2007) 'Using External Data in Operational Risk', Available at SSRN: http://ssrn.com/abstract=871181, January 2007
- Khan, A., Moncelet, B. and Pinch, T. (2006) 'Uses and Misuses of Loss Data' Global Association of Risk Professionals, May/June 2006.
- Na H., Miranda, L., van den Berg, J. and Leipoldt, M. (2005) Data Scaling for Operational Risk Modelling', ERS-2005-092-LIS - Erasmus Research Institute of Management- Report in Research Management
- Pezier, J. 2002, 'A Constructive Review of Basel's Proposals in Operational Risk', ISMA Discussion Paper in Finance 2002-20, University of Rebankbankng, September 2002
- Shih. J, A. Samad-Khan, and P. Medapa (2000), 'Is the Size of an Operational Loss Related to Firm Size?' *Operational Risk Magazine* **2**, 1.
- Yap, J. (2006) 'External Data and Operational Risk: Practical Issues and Solutions', Actuarial Research Essay – University of Melbourne, October 2006
- Wilson, S (2007) ' A Review of Correction Techniques for Inherent Biases in External Operational Risk Loss Data', Working Paper - Australian Prudential Regulation Authority, November 2007