Milliman Research Report

Prepared by:
Joshua Corrigan
Principal

Paola Luraschi
Principal

Reviewed by:
Neil Cantle
Principal

February 2013

Operational risk modelling framework

Milliman

Milliman is among the world's largest providers of actuarial and related products and services. The firm has consulting practices in healthcare, property & casualty insurance, life insurance and financial services, and employee benefits. Founded in 1947, Milliman is an independent firm with offices in major cities around the globe. For further information, visit milliman.com.

ACKNOWLEDGEMENTS

The authors would like to thank the many people who contributed to the production of this research report. Without their help much of it would not have been possible.

In particular we would like to thank:

Adrian Allott, Senior Consultant, Romania

Neil Allan, Systemic Consult, London

Ashleigh Hickey, Analyst, London

Marcin Krzykowski , Principal, Poland

Camilo Salazar, Principal, Latin America

Paul Sinnott, Principal, Hong Kong

Marc Slutzky, Principal, New York

Fred Vosvenieks, Consultant, London

Caroline Wang, Consultant, Hong Kong

Wing Wong, Principal, Hong Kong

Masaaki Yoshimura, Managing Director, Tokyo

TABLE OF CONTENTS

ACKNOWLEDGEMENTS 1

1 EXECUTIVE SUMMARY 4

1.1 Background and objectives 4
1.2 Scope and structure of the report 4
1.3 Key research findings 5
1.4 Feedback 6

2 INTRODUCTION 7

2.1 Definition of operational risk 7
2.2 Why should we care about operational risk management? 8
2.3 Requirements of and use cases for an operational risk management framework 8
2.4 A complex systems science approach 8

3 NATURE OF OPERATIONAL RISK EVENTS 11

3.1 Loss frequency and severity distributions 11
3.2 Operational risk event examples 12
3.3 Implications 14

4 BASIC INDICATORS AND STANDARD FORMULA APPROACHES 15

5 QUANTITATIVE RISK ASSESSMENT OR SCENARIO ANALYSIS 17

6 STATISTICAL MODELS: LOSS DISTRIBUTION APPROACH 20

6.1 Data sources for calibration 20
6.2 Event type segmentation 20
6.3 Loss severity 21
6.4 Loss frequency 25
6.5 Loss interrelationships 26
6.6 Aggregation and capital assessment 31
6.7 Advantages and limitations of an LDA 33

7 STRUCTURAL OR CAUSAL APPROACHES 35

7.1 Introduction 35
7.2 Eliciting the system structure 35
7.3 Identifying the most important components of the system 37
7.4 Bayesian networks: A quantitative framework for integrating information 38
7.5 Determining appropriate states of business drivers and risk factors 40
7.6 Defining causal relationships 41
7.7 Aggregation 46
7.8 Risk assessment 48
7.9 Capital assessment 49
7.10 Sensitivity analysis 50
7.11 Stress testing 50
7.12 Reverse stress testing 52
7.13 Risk appetite and risk limit setting 53
7.14 Advantages and limitations of a structural framework 54
7.15 Choice of technique 55

8 REGULATORY REQUIREMENTS 56

8.1 Overview of Basel II/III framework for operational risk 56
8.2 Overview of insurance operational risk regulatory requirements 58

9 EMERGING OPERATIONAL RISK ASSESSMENT AND BLACK SWANS 64

9.1 Operational risk dynamics and evolution 64
9.2 Phylogenetics 64
9.3 Derivative trading loss example 65
9.4 Implications 70

10 LOSS DATA COLLECTION 71

10.1 What is loss data collection? 71
10.2 Methodology 71
10.3 Important design elements 73
10.4 Risk measurement and management use cases 76
10.5 Possible future developments 76

11 CASE STUDIES 78

11.1 Basic indicators and standard formula approaches 78
11.2 Loss distribution approach 79
11.3 Structural or causal model 85

12 REFERENCES 94

1 EXECUTIVE SUMMARY

1.1 Background and objectives

All organisations bear operational risk in order to achieve their objectives. For organisations motivated by profit, operational profit is the return required by the capital owners for bearing the operational risk associated with the production process. Hence operational risk is of primary concern to internal stakeholders–including the capital owners of the business, management who control operational risk, and employees who are part of the operational production process–and external parties such as creditors, customers, and regulators who are impacted by operational risk failures.

As it is such a fundamental risk, most organisations are very conscious of operational risk, and many of them are very good at managing and mitigating operational risk. Despite this, however, the field of operational risk assessment is still relatively new, particularly when it comes to its inclusion in capital frameworks. The global banking industry has developed regulatory capital frameworks for operational risk that have been used for the last five to 10 years, and some countries are currently following suit with respect to equivalent insurance regulatory capital standards. But outside of these two industries, there is relatively little operational risk assessment activity from regulators, industry bodies, or individual companies.

The authors thus felt it a worthwhile area of research to investigate the current state of play in operational risk assessment frameworks, and how they are likely to evolve over the coming years. The objectives of this research report are therefore to provide readers with an overview of existing operational risk assessment frameworks around the world, along with a detailed investigation into the current methods and emerging practice in this area.

1.2 Scope and structure of the report

We start this report with an introduction to operational risk, in Section 2, which is designed to put operational risk in context for readers who are new to this area. We then delve into the nature of operational risk events in Section 3, highlighting the highly nonsymmetric nature of operational risk, which sees operational event numbers dominated by small loss sizes while overall operational losses are dominated by a very small number of extremely large events. This is one of the defining features of operational risk that makes it so unique relative to other risk types, and consequently difficult to assess.

We then work through a detailed assessment of the four main approaches used to assess operational risk. In increasing order of complexity they are:

- Basic indicators and standard formulas (Section 4)
- Quantitative risk assessment or scenario approach (Section 5)
- Statistical or loss distribution approach (LDA) (Section 6)
- Structural or causal approach (Section 7)

For each of these approaches we outline the methodology and discuss the advantages and limitations. Examples are used to illustrate methods, with additional case studies provided in Section 11.

After investigating the various approaches, we then turn our attention, in Section 8, to the various regulatory requirements for operational risk that exist in the financial services sector around the world. Operational risk is one of the areas where such frameworks exhibit the most divergence, with some jurisdictions clearly more sophisticated in their approaches relative to others.

Section 9 discusses emerging operational risks, for which we explore the use of phylogenetic methods to understand the underlying evolutionary process of characteristics that drive operational risk losses. Phylogenetic methods enable the relationships between characteristics of 'things' (organisms, companies, risk events, etc.) to be determined in the most parsimonious and objective way. This also has important implications for the loss data collection process, which is the subject of

Section 10. This is the process through which operational losses are recorded, which is used as a key resource to calibrate many operational risk assessment models.

Given the significant length of this research report and the fact that many sections are relatively independent of one another, we encourage readers to skip ahead to those sections of the report that are of particular interest to them.

1.3 Key research findings

There are a number of key findings from this research paper, which we have synthesised below.

1. Operational risk is a material risk and is one of the major causes of organisational failure and the destruction of shareholder value

Recent history bears witness to some of the most significant operational risk events that have caused not only complete organisational failure and the destruction of significant shareholder value, but perhaps most unfortunately also the loss of many lives. Significant and increasing effort is expended by most organisations to mitigate the likelihood and severity of operational risk events given their consequences, although the science of operational risk assessment is still in its relative infancy. It has only been in the last 10 years that Basel II has upped the ante in this area, requiring banks to methodically quantify their operational risk, with other industries such as insurance not too far behind. This discipline is likely to develop significantly over coming years as the value of an effective operational risk management framework in creating a resilient organisation becomes increasingly accepted.

2. The scope of operational risk needs to consider all factors of production

Operational risk is typically defined with respect to people, process, and systems. However, we feel that this is somewhat limiting, and that it is important to consider all the factors of production. Certain production factors will be more or less important for different industries and organisations, and the relativities between the different forms of labour input are changing rapidly.

3. Operational losses are highly skewed and dominated by high-severity, low frequency events

Whilst the numbers of operational loss events are dominated by those with low severity, total operational losses are dominated by high-severity, low frequency events. This interdependency and nonlinearity is a key feature that must influence the entire operational risk assessment and management framework. This means that any assessment of operational catastrophe losses, such as what is used to set required capital, effectively boils down to an assessment of these high-severity, low-frequency events.

4. Basic indicators and standard formula approaches are a simple but ultimately very blunt tool

Whilst appealing through their simplicity, basic indicator and standard formula approaches can be very misleading, as they are insensitive to the actual operational state of the organisation (as demonstrated by the case study in Section 11). Such methods should not be relied upon as an accurate or informative measure of actual operational risk.

5. Scenario and statistical modelling approaches such as LDA are useful, but they have significant limitations

Current practice for the leading global banks is dominated by the use of the combination of scenario and statistical modelling approaches such as LDA. These are significantly better than standard formula approaches as they integrate multiple sources of information: internal and external loss data, scenario information, and business environment and internal control factors. However, they do have significant limitations, most of which are a result of the disjoint between modelled loss outcomes and the evolving state of the organisation's operational risk drivers. Abstract statistical methods such as correlations and copulas mean that modelled outcomes can be difficult to calibrate and explain, and lack robustness for high severity loss estimates given the paucity of historical loss data and difficulties people have in estimating a 1-in-200 or -1,000-year event. Such models in the banking sector also failed to respond to clearly changing operational risk levels over the period of the recent global financial crisis. There is increasing acceptance within the industry that these approaches need improvement.

6. Structural or causal-based models are leading emerging best practice in this field

Latest emerging practice is in the field of structural models, which link outcomes directly to the primary causal drivers of operational risk. This is typically achieved via a Bayesian network approach that conditions operational outcomes (both financial and nonfinancial) upon the range of causal drivers directly accounting for their complex interdependencies. The strength of this approach is that it is also able to flow information in both directions through the network via Bayesian inference, which enables the robust determination of operational risk limits to be derived consistent with operational risk appetite levels. The limitation with this approach is that it can be more complex and involved to implement, and thus needs to be judiciously applied.

7. There is wide divergence in regulatory maturity levels for operational risk across regions, countries, industries, and companies

Regulatory frameworks for operational risk are evolving rapidly, particularly in the banking sector and increasingly in other financial services sectors such as insurance. However, there is still significant divergence in practice by region, with many countries including the largest insurance markets of the United States and Japan still only using a basic indicator approach. Outside of the financial services industry, there is a general lack of regulatory capital frameworks for operational risk, although there is increasing recognition that quantification of operational risk is becoming an increasingly important part of an enterprise risk management framework.

8. Phylogenetic techniques are now being used as a method of assessing emerging operational risks

The assessment of emerging operational risk is a very new area. Existing techniques largely centre around scenario analysis. However, phylogenetic techniques are showing some promise and are being considered by a number of organisations. These techniques enable the objective assessment of the evolutionary relationships of the characteristics of operational risks to be determined, thus enabling a structured way in which emerging or evolving risks might occur in the future.

9. The development of a loss data collection process can add significant value to operational risk management

A loss data collection (LDC) process is a central element in an operational risk assessment and management framework that can add significant value by enabling the analysis of actual losses to be used to condition future expected losses and to identify appropriate risk mitigation controls. One important aspect of these frameworks is that they should be constructed such that they do not lose information in the data capture process, and such that the underlying causal drivers and characteristics can be identified, thus enabling advanced analytical techniques such as phylogenetic methods to extract as much insight as possible from them.

1.4 Feedback

The authors welcome any feedback or comments that readers may have on this report, and we sincerely hope that you find it a useful resource for professional use over many years. Please feel free to contact the authors at joshua.corrigan@milliman.com or paola.luraschi@milliman.com with your questions regarding the report or to further discuss its findings.

2 INTRODUCTION

2.1 Definition of operational risk

The primary objective of most business organisations is to generate operating profit. This profit is effectively the return to the owners of the business for undertaking the operating activities of the firm, from which they bear operational risk in the process. Operational risk is thus the risk of loss resulting from failed operational activities. Such failures arise from application of the firm's productive inputs, such as natural resources, labour, and capital, to the process of the production of output goods and services. Understanding this production process is the key to understanding operational risk.

Most regulatory frameworks used in financial services define operational risk as *the risk of loss resulting from inadequate or failed internal processes, people, and systems, or from external events*. Whilst this definition is clear and pragmatic, we think it is valuable to consider the full range of potential productive inputs that constitute operational activities, and how these activities interact with exogenous environmental factors. In many cases it is often a failure in an endogenous productive input, which creates an inherent sensitivity to some exogenous factor, which becomes the *final nail in the coffin*. So perhaps a more fundamental definition would be *the risk of loss resulting from inadequate or failed productive inputs used in an operational activity*.

A firm's productive inputs, or factors of production, should ideally be considered at a granular level, as these give us a detailed insight into the nature of operational activities and thus operational risk. For example, when we talk about people risk, what exactly are we referring to? Is it the number of people, their levels of operational expertise as embodied in their intellectual capital, the application of their physical labour, or their social capital in contributing to operational culture? We feel that it is important for the operational risk management framework to be cognisant of these different elements in order for it to be complete. Figure 1 identifies and describes these factors of production at a high level.

FIGURE 1: FACTORS OF PRODUCTION RELEVANT FOR OPERATIONAL RISK

TYPE	PRODUCTIVE INPUT	DESCRIPTION
Natural Resources	Land	The physical space used to carry out the production process that may be owned, rented, or otherwise utilised
Natural Resources	Raw materials	Naturally occurring goods such as water, air, minerals, flora, and fauna
Labour	Physical labour	Physical work performed by people
Labour	Human capital	The value that employees provide through the application of their personal skills that are not owned by an organisation
Labour	Intellectual capital	The supportive infrastructure, brand, patents, philosophies, processes, and databases that enable human capital to function
Labour	Social capital	The stock of trust, mutual understanding, shared values, and socially held knowledge, commonly transmitted throughout an organisation as part of its culture
Capital	Working capital	The stock of intermediate goods and services used in the production process, such as parts, machines, and buildings
Capital	Public capital	The stock of public goods and services used but not owned by the organisation, such as roads and the Internet

Assessing and understanding how these input factors are used in the operational production process is the first step in operational risk management. Different industries and companies within each are characterised by different relative quantities of inputs, and the different methods in which they are combined to produce outputs. For example, financial services firms use relatively little direct natural resources relative to a manufacturing company, but will use proportionately greater intellectual and social capital. The relative quantities of inputs will also likely change over time, which means that an operational risk framework needs to be adaptive. As a result of the numerous sources of production inputs, operational risks are very broad and heterogeneous across industries and companies.

2.2 Why should we care about operational risk management?

Operational risk management is the heart of almost every business, because a majority of the productive inputs of a company are used in operational activities. Whether people realise it or not, many are already experts at operational risk management as it relates to their segment of a specific operational activity. However, for any company larger than a sole trader, it is almost impossible for any one individual to be an expert in operational risk management at an enterprise level. Many companies fail as a result of operational risk, such as Barings Bank when it became bankrupt in 1995, which was due to the actions of a rogue trader, Nick Leeson. Hence systematically managing operational risk at an enterprise level is a critical management activity for complex organisations.

When undertaken properly, operational risk management can add significant value to a business by:

- Identifying the most critical operational risk drivers requiring management, thus reducing the risk of operational failures

- Increasing the likelihood of strategic business objectives being met

- Optimising operational risk exposures within a dynamic and evolving business environment

- Facilitating better management decisions through a framework that is engaging to business managers because it is framed in terms of real business drivers and language

- Integrating multiple risk constructs, including risk assessment, capital assessment, risk appetite, setting operational risk limits, emerging risk assessment, risk culture, risk monitoring, and reporting

2.3 Requirements of and use cases for an operational risk management framework

An effective operational risk management framework needs to be able to integrate all of the various risk constructs, including:

- Risk identification
- Risk assessment
- Risk capital assessment
- Risk monitoring
- Risk mitigation
- Risk appetite and risk limit setting
- Risk sensitivity analysis
- Emerging risk assessment
- Risk culture assessment
- Risk reporting, distribution, and communication

All of these components are important. Traditional frameworks have tended to treat each of them as separate tasks, with relatively loose points of integration between them. For example, assessment of operational risk capital using simple techniques such as standard formulas, results in a disconnect to the setting of risk appetite and risk limits based upon sources of risk and business risk drivers. Unless these elements are effectively integrated into one framework, they will lack the integrity needed to manage the business effectively and demonstrate value add.

2.4 A complex systems science approach

Milliman's approach to operational risk management is founded on the work that has been developed over recent decades in the field of complex systems science. This field has developed important findings that we believe play a central role in the assessment of risks for organisations, which themselves can be viewed as complex adaptive systems. These findings are based upon the premise that risk frameworks and tools need to embrace the following properties to be effective:

- **Holism**: A system is more than just the sum of its parts, and in order to understand its behaviour a holistic rather than a reductionist approach is needed.
- **Nonlinearity**: All complex systems are characterised by nonlinear behaviour, which needs to be a core feature of the interrelationships between various states of system drivers.
- **Human bias**: Companies are comprised of people that are subject to a range of biases, which risk frameworks need to incorporate.
- **Emergence**: Risk can be viewed as the unintended emergent property of a complex adaptive system that is subject to environmental adaption forces.

These properties are the defining characteristics of a complex adaptive system, of which large businesses are prime examples. Such systems exhibit very interesting and important behaviours for risk management, including:

- **Self-organisation**: They have structure and hierarchy, but few leverage points.
- **Interacting feedback loops**: They can cause highly nonlinear behaviour, resulting in potentially catastrophic outcomes.
- **Tipping points**: Once the system reaches a critical threshold or level of complexity, they can collapse.
- **Path-dependencies**: Historical evolutionary paths are important as they influence behaviour.

Figure 2 illustrates the relationships between system structure and crisis events.

FIGURE 2: UNDERSTANDING HOW SYSTEM STRUCTURE RESULTS IN CRISIS EVENTS

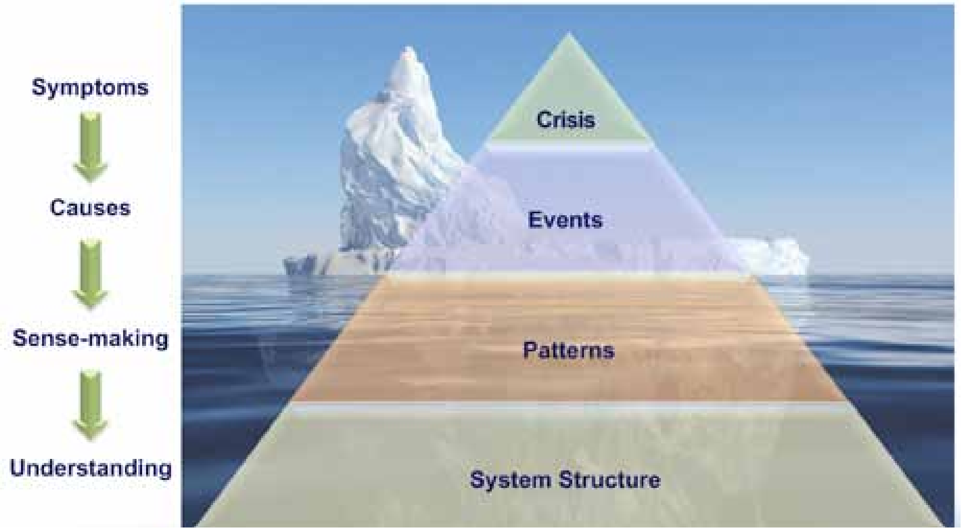


Figure 2 illustrates that, by understanding the somewhat opaque underlying system structure, we are better able to make sense of the patterns of outcomes that the system is characterised by, which in turn lead to risk events and occasionally crises. Traditional operational risk management frameworks focus on the obvious components that are above the water line, which results in backward-looking, relatively static approaches to risk management which don't provide any explanatory power. However, by focusing on the underlying system structure, a much richer set of information is then available from which to populate all the operational risk constructs, resulting in insight being generated earlier and with better predictive capability.

Before turning our attention to the practical aspects of a risk framework based upon such a philosophy, we first outline in the following sections the nature of operational risk events, and the traditional approaches that are used for operational risk modelling and management. For advanced readers who are familiar with such material and wish to focus on complex system science approaches, we invite you to skip ahead to Section 7.

3 NATURE OF OPERATIONAL RISK EVENTS

3.1 Loss frequency and severity distributions

Unlike many other types of risk, operational risks and their associated losses are commonly characterised as having heavy-tailed distributions: a large number of very small losses and a very small number of extremely large losses. This is clearly evident in the global banking industry, as shown by Figures 3 and 4, produced by ORX.[1]

FIGURE 3: DISTRIBUTION OF NUMBER OF EVENTS BY SIZE

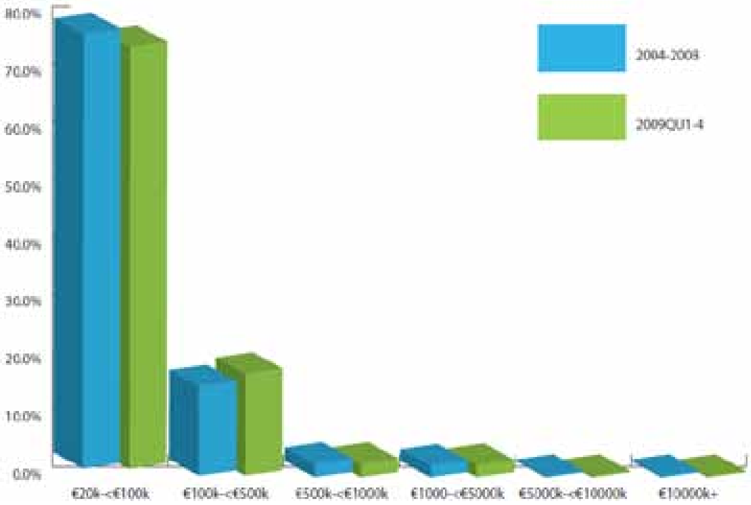
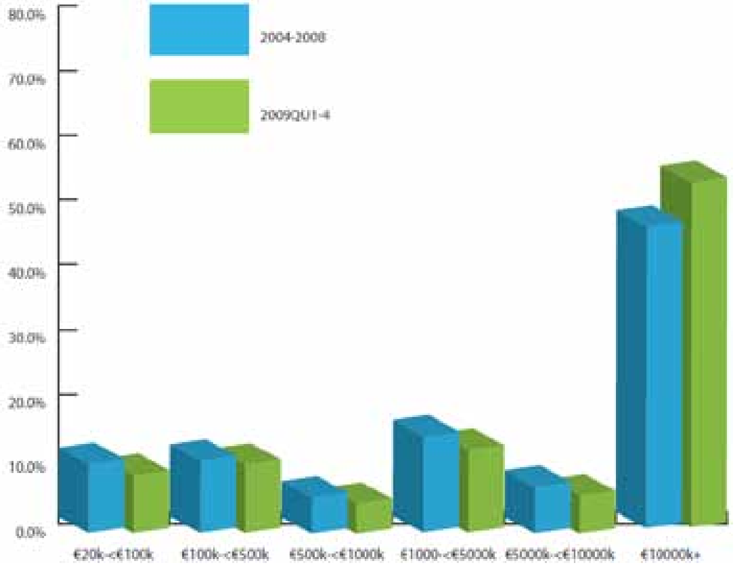


FIGURE 4: DISTRIBUTION OF TOTAL GROSS LOSS BY SIZE



[1] Refer ORX (2010).

Figure 3 shows that almost 80% of operational risk events are in the smallest loss size category[2] of EUR 20,000-100,000, whilst Figure 4 shows that almost half of the total gross losses are in the highest loss size category of EUR 100,000+. These distributions tend to be even more skewed when considered at more granular levels. For example, according to ORX,[3] out of a set of 6,593 operational losses in the retail banking business line for the *Clients, Products, and Business Practices* event type, the 10 largest losses account for 44% of the total losses, and the 100 largest losses account for nearly 75% of the total loss value.

These findings present difficulties to those interested in assessing operational risk. The fact that total losses are dominated by a handful of extreme tail events means that distribution parameter estimates and statistical characteristics such as means and VaR/CTE are likely to be unstable, as they can be materially influenced by single observations. Thus capital measurements which are dependent upon correct modelling of the tail of such distributions are highly sensitive to single large loss events and are thus quite uncertain.

It is highly unlikely that a single organisation has a statistically significant number of large loss events to properly assess the tail of the distribution. Even when external loss data is available, such as in banking, there is still a general paucity of data in the tails, particularly when it is segmented by business line and event type. Cope et al. (2009) conclude that it would take many thousands of years of loss data to be collected in order to have sufficient confidence in the assessment of large severity operational risk losses.

Within the financial services industry, operational risk losses are measured in terms of financial loss. In other industries, however, it's just one of the ways of measuring operational loss, in many cases not even the primary unit of measurement. The number of lives lost is perhaps the most important way of comparing many operational risk events, particularly in highly industrialised sectors such as mining and energy. Clearly, there are varying degrees of severity below this, relating to physical, mental, and emotional damage that can occur to people experiencing an operational event. Similarly, varying degrees of loss can occur to any of the factors of production used in a production process, whether it be natural resources, labour, or capital, and the challenge for operational risk management is to understand, assess, and manage all these important dimensions of operational risk.

3.2 Operational risk event examples

Given the above discussion on the statistical distributions that operational losses tend to follow, it is typical to consider segmenting the distribution into a few alternative components. The most common method of segmentation is into three categories: low-severity and high-frequency events, medium-severity and medium-frequency events, and high-severity and low-frequency events. These components are broadly meant to capture different underlying loss generation mechanisms or operational processes, although this is a very broad approach and not tied to any specific activities.

[2] Note that this database has a minimum reporting size for any loss event of EUR 20,000.

[3] Refer ORX (2009).

Figure 5 provides some examples of operational events for each of these categories in a range of industries.

FIGURE 5: OPERATIONAL EVENT EXAMPLES

	INSURANCE	BANKING	MINING	ENERGY
Low Severity, High Frequency	Claims processing, data errors	ATM failures	Transport service interruption	Meter reading errors
Medium Severity, Medium Frequency	Fraud, regulatory compliance failure	Online security breach	Environmental contamination	Environmental contamination
High Severity, Low Frequency	Mis-selling, mis-pricing	Rogue trader	Mine collapse	Oil spill, gas plant fire

Clearly, there are many types of events that could be categorised in this way. Some specific examples of very high severity and publicly noteworthy operational risk events across these industries include the following:

- **Insurance: Mis-selling**
  - Equitable Life, the world's oldest mutual insurer, almost became bankrupt in early 2000, after it mis-priced embedded guarantees in its with-profits products.
  - Endowment mortgages sold to 11 million people in the UK in the late 1980s in which the risks of the product were not adequately explained. The Treasury Select Committee estimated in 2004 that total shortfalls were circa GBP 40 billion, with the industry paying compensation of around GBP 5 billion.[4]
  - HIH is Australia's largest corporate collapse, having losses totalling AUD 5.3 billion, as a result of a range of problems including pricing and reserving errors, and internal fraud by management.
- **Banking: Rogue traders**
  - Barings Bank, 1995, Nick Leeson lost GBP 827 million, which caused the bank to become insolvent.
  - Society Generale, 2008, Jerome Kerviel lost EUR 4.9 billion.
  - UBS, 2011, Kweku Adoboli lost USD 2.3 billion.
- **Mining**
  - Benxihu Colliery, 1942, in China the world's worst known mine disaster killed 1,549 miners.
  - Ulyanovskaya mine, 2007, in Russia, which killed 106 miners.
  - Pike River mine, 2010, in New Zealand, which killed 29 miners.

[4] Refer Corrigan et al. (2011).

- **Energy**

  — The Gulf War oil spill and associated Kuwaiti oil fires, 1991, in Kuwait, as a result of the defeat of the Iraqi army in Kuwait.

  — Deepwater Horizon oil spill, operated by BP in 2010, which spilt 560 tons of crude oil into the Gulf of Mexico over a three-month period and killed 11 men working on the platform.

  — Exxon Valdex oil spill in 1989, which spilt up to 750,000 barrels of crude oil into Prince William Sound in Alaska.

  — Chernobyl disaster in Ukraine in 1986, which released large quantities of radioactive contamination into the atmosphere, spreading across much of Western USSR and Europe. Thirty-one deaths are directly attributed to the accident, although many groups estimate deaths that are due to radiation could reach into the tens of thousands.

Whilst operational risk and capital assessment are important risk management activities, the primary goal of operational risk management is to mitigate the risk of these high severity types of events from happening. However, unfortunately, they seem to occur with worrying regularity, which clearly points to a failing in operational risk management within these companies.

3.3 Implications

The above discussion serves as important context for consideration of operational risk frameworks. One of the most critical implications is that the underlying loss generation mechanism needs to be considered very clearly. We define the *loss-generation mechanism* as the operational activity or production process with which an operational risk event is associated. Note that we use the term *associated*, rather than *caused by*, as more than one operational activity or production process may be the cause of a specific operational risk event (e.g., rogue trading).

When identifying, understanding, and assessing an operational risk, it is important to consider whether or not the loss generation mechanism is stable over time, or if it changes. If it isn't stable and exhibits change, historical loss data becomes less relevant as time passes. All aspects of change, including process, regulation, legal environment, political environment, technology, etc., need to be considered here. As organisations pay more attention to operational risk management, governance, and compliance activities, they are able to exert better control over loss events through prevention and mitigation actions, which would also change the assessment of loss likelihood and severity. Companies also change over time. They grow in size, merge or divest business units, enter new markets, and restructure their organisations. The occurrence of an extreme loss, either directly or to a close competitor, will usually bring about large-scale changes in the way a company conducts and controls its business, such that the occurrence of a similar loss of equal magnitude arising again is greatly diminished.

In the case where the loss generation mechanism is stable, then the application of statistical techniques is much more robust. This is more likely to be the case for operational risks characterised in the low and medium frequency categories. An example of this would be highly repeatable activities such as payment processing errors, which tend to be stable for a given suite of systems and processes. The implications for risk mitigation and management would tend to focus on general principles and practices (and risk culture) rather than on refining detailed operational systems and processes, for which the marginal cost could exceed the marginal benefit.

In the case where the loss generation mechanism is unstable, then the application of statistical techniques can be significantly misleading. These events are much more likely to be in the medium- to high-severity categories. The implication for these operational risk events is that risk assessment becomes increasingly more difficult and uncertain. Risk mitigation needs to focus less on general principles and practices, and much more on the specific risk drivers of an operational activity.

4 BASIC INDICATORS AND STANDARD FORMULA APPROACHES

The simplest method for calculating operational risk capital is the use of a single indicator as a proxy for operational risk. For example, a fixed percentage of gross operational income would be a universally applicable way of calculating operational risk capital across a wide range of companies. However, the limitations of such a simple approach are significant, the chief one being that it does not directly relate to the operational risk characteristics of each unique business, and hence is effectively useless in the management of operational risk.

In order to capture more of the differing risk characteristics across each firm within an industry, some regulatory environments introduce additional indicators and vary the capital quantity by each. Such approaches are called standardised approaches. Figure 6 shows examples of such basic indicator and standardised approaches for a range of industries and geographies.

FIGURE 6: SUMMARY OF BASIC INDICATOR AND STANDARD FORMULA APPROACHES FOR OPERATIONAL RISK CAPITAL

GEOGRAPHY	INDUSTRY	REGULATORY SCHEME	INDICATOR	PERCENTAGE
Global	Banking	Basel II, basic indicator	Gross income over those of the previous three years in which it had been positive.	15%.
Global	Banking	Basel II, standardised	Gross income by business lines[5] over the previous three years (positive part). Loans and advances amount for the retail banking and commercial banking business lines can be used alternatively.	Range from 12% to 18% by business line. If loans and advances are used, 3.5% of them should be used instead of gross income (that is, before multiplication by the 12% to 18% factor).
European Union	Insurance	Solvency II	BSCR, the total of all other risk capital amounts. Earned premiums (EP). Technical provisions (TP). Expenses.	Complex function subject to various maximums:[6] • 30% of BSCR • 4% life premiums • 3% nonlife premiums • 0.45% life TP • 3% nonlife TPs • 25% expenses
Australia	Insurance	APRA, LAGIC	Premium income. Policy liabilities. Policy payments.	Various depending upon the class of insurance.[7]
Japan	Insurance	Statutory Solvency Requirement	Total risk (insurance + interest crediting + asset + guarantee lines).	3% if P&L is negative, 2% if P&L is positive.
Republic of South Africa (RSA)	Insurance	Solvency Assessment and Management	BSCR, the total of all other risk capital amounts. Earned premiums (EP). Last year of EP growth if over 10% (GEP). Technical provisions (TP). Expenses for unit-linked business (other than commission).	25% of UL expenses plus minimum of 30% of BSCR and maximum of: • Sum of: 3% of nonlife TP and 0.45% of life TP • Sum of 3% of nonlife EP and GEP and 4% of life EP and GEP
Taiwan	Insurance	Risk-Based Capital	Premium income. Assets under management (AUM).	0.5% for life business. 1% for annuity business. 1.5% for all other business. 0.25% for AUM.
Others (USA, Europe except EU, Other Asia, Russia, New Zealand)	Insurance		None, no explicit formula exists. However, it is quite common to have requirements for operational risk management and internal control.	None.

[5] Refer to Section 8 for the categorisation of these business lines.
[6] Refer to Section 8 for the detailed operational risk formula.
[7] Refer to Section 8 for the detailed operational risk formula.

The main benefit of this approach is that it is very simple to understand, transparent, efficient, and cheap to implement. However, this approach has a serious flaw in that the risk capital is not directly linked to actual risk exposure and mitigation drivers in the business. As a result of the numerous sources of production inputs, operational risks are very broad and heterogeneous across industries and companies, in contrast to other risk categories, such as market risk, which are significantly more homogeneous. This fact means that *one-size-fits-all* approaches to operational risk assessment such as the use of basic indicators and standard formulas are likely to be relatively ineffective. These methods also do not provide any information on the types of operational risk events, their risk profiles, or how to control them, and hence it does not help or incentivise good risk management processes.

Because of these limitations, regulators of industries such as banks allow their regulated entities to use more sophisticated methods such as the advanced measurement approach (AMA) allowed under Basel II. These approaches include the loss distribution approach (LDA), the scenario approach, and the structural modelling approach, which enable firms to tailor the models to better reflect their actual risk profiles. These approaches are the subject of the following sections of this paper.

5 QUANTITATIVE RISK ASSESSMENT OR SCENARIO ANALYSIS

Quantitative risk assessment (QRA), also known as scenario analysis (SA), is a fusion of the expert knowledge of the current and future states of the business and the environment in which it operates, in order to assess operational risks. The integrity and consistency of the outputs produced under this approach must be ensured through the use of a strong governance framework. The elements of this framework are:

- **Governance**
  Effective governance is essential to ensure support for the process from participants and stakeholders, such that it has integrity, consistency, and longevity.

- **Preparation**
  A clear understanding of the objectives, framework, and input requirements is needed in order for participants to effectively contribute to the process. Potential areas for scenarios and any relevant data should be sourced and prepared in advance.

- **Assessment**
  This is typically undertaken through a series of workshops, interviews, or questionnaires involving a combination of subject matter experts, executives, risk managers, and group functions. Quantitative input parameters on risk severity and frequency of each scenario are required.

- **Validation**
  Given the biases that may enter into the process, scenarios need to be reviewed holistically and challenged to ensure consistency in their framing, depth, breadth, and calibration.

- **Reporting**
  Scenario results must be reported to a variety of stakeholders across various levels of the business, from business units all the way through to the board.

- **Evolution**
  Scenarios must be updated to reflect the ongoing evolutionary process through which the business dynamically changes to respond to changes in the environment in which it operates

The assessment process is typically undertaken by senior managers considering a range of large to catastrophic scenarios that the organisation could suffer that are mostly unrelated to one another. Examples of such scenarios include:

- Failure to deliver a significant operational process change program
- Failure to recruit and retain staff of appropriate quality and in sufficient numbers
- Failure of the loss of key infrastructure assets
- Large fraud event

For each scenario, the relevant experts provide their opinions on the likely size of each loss, and how likely the losses are to occur in the next year. Ideally and where possible, they would provide these estimates for a few different points on the likelihood and severity distributions, in order for the distributions to be calibrated. In undertaking this, they would need to consider questions such as what controls would have to fail for a loss event to occur, and whether such failures have happened before and during their working lifetimes, as well as the lifetime of the organisation.

An alternative scorecard approach could also be used, such that participants rate or rank each scenario representing a defined loss severity and frequency and its key risk indicators. Weights would then be determined based upon historical data or expert opinion to aggregate the scores, which would then be reflected in actual distributions.

Consideration of business environment and internal control factors (BEICFs) can assist the assessment process in making estimates that are more relevant to the current residual risk profile of the organisation, as reflected by their inclusion in specific scenarios. These scenarios can also be expressed directly in terms of likelihood and severity outcomes. The use of such information provides a mechanism to link the top-down requirements to measure operational risk and capital with bottom-up business drivers, thus providing a framework for linking risk measurement with risk management.

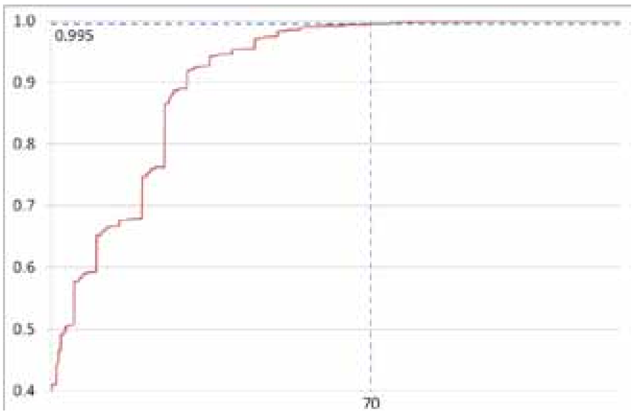
Consider the following example of 26 independent scenarios, each with their own severity and likelihood estimates, which is illustrated in Figure 7.

FIGURE 7: EXAMPLE SCENARIOS WITH SEVERITY AND LIKELIHOOD ESTIMATES

SCENARIO	SEVERITY (M)	LIKELIHOOD (P.A.)	SCENARIO	SEVERITY (M)	LIKELIHOOD (P.A.)
1	5	5.00%	14	75	0.25%
2	10	1.00%	15	30	0.50%
3	1	3.00%	16	20	5.00%
4	10	1.00%	17	20	5.00%
5	10	1.00%	18	10	5.00%
6	10	5.00%	19	3	2.00%
7	20	5.00%	20	5	1.00%
8	5	5.00%	21	2	1.00%
9	5	5.00%	22	2	5.00%
10	30	0.50%	23	1	5.00%
11	25	0.25%	24	1.5	5.00%
12	75	0.10%	25	25	10.00%
13	10	0.10%	26	25	10.00%

Assuming that all these scenarios are independent, then we can use a generalised binomial distribution to estimate the annual loss distribution for any and all combinations of the above events occurring. Figure 8 shows the cumulative distribution function of annual losses.

FIGURE 8: CUMULATIVE DISTRIBUTION FUNCTION OF ANNUAL LOSSES



From the above distribution it is possible to calculate the tail loss amount, such as 70m at the 99.5th percentile level. It is also possible to introduce scenario interdependence through the application of correlation and copula techniques.[8]

The benefit of this approach is that it is forward-looking and hence likely to be adapted to the evolving dynamics of an organisation. It is also transparent and thus easy to understand in that risk exposures are tied to the qualitative information provided by management.

However, there are a number of challenges and limitations with this approach. These include:

- Bias in scenario selection: It is difficult to know where to draw the line in judging how many and what types of scenarios should be included in the assessment.

- Quality of information gathered: Humans are typically very poor at assessing the severity and frequency of risks, particularly high severity low frequency risks which they may never have experienced. This is particularly acute for operational risk capital for banks, which needs to be set at a level to withstand a 1-in-1,000-year event—how is anyone ever meant to be able to understand what such an event might look like given 40-year average working lifetimes?

- Allowance for complexity: Humans are also generally very poor at resolving the complexity of the interactions that characterise complex scenarios and hence may miss sources of operational loss, such as those arising from negatively reinforcing feedback loops.

- Difficulty in aggregation: Assessing the interdependencies between the scenarios can be very challenging, as they are typically framed individually.

- Lack of allowance for uncertainty: Use of point estimates overstates the degree of uncertainty associated with the scenarios.

[8] Refer to Section 6 for more details.

6 STATISTICAL MODELS: LOSS DISTRIBUTION APPROACH

The loss distribution approach (LDA) is a popular statistical approach for computing aggregate loss distributions. In this section, we define the underlying mathematical modelling framework and provide some of the typical algorithms to compute loss distributions. We also discuss the pros and cons of this method.

LDA is a commonly used approach to quantify operational risk, where the severity and frequency of operational risk losses are considered and modelled separately. Once they have been calculated, the aggregate loss distribution can be generated by combining them through Monte Carlo simulation techniques.

6.1 Data sources for calibration

Sourcing data in order to calibrate the distributions using an LDA approach is the first step in the assessment process. There are broadly four categories of data that can be used for this purpose. The Basel II banking accord requires all four of these data sources to be utilised for banks wishing to obtain regulatory approval for use of an advanced measurement approach (AMA) for operational risk capital assessment, of which LDA is an approved method. The four data categories are:

- **Internal loss data (ILD)**: This is data on an organisation's actual operational loss experience. Because ILD is usually very limited for high-severity, low-frequency loss events, it is necessary in order to assess the tail of the distributions to consider the alternative data sources, which produces meaningful estimates of capital requirements. Consideration should be given to the length of history for which ILD is appropriate to use, given potential changes in an organisation's operational process, risk control process, and risk appetite levels.

- **External loss data (ELD)**: This is data on the actual operational loss experience of other organisations. ELD is an essential data source for high-severity, low-frequency events which the organisation under consideration may not have experienced. Such databases are selectively available by industry, such as the ORX database, which contains operational risk losses for global banks.[9] ELD may also have additional uses, for example in assessing the riskiness of new business lines, in benchmarking analysis, or in estimating competitors'/industry loss experience. Whilst ELD can be very useful, there is a risk that it may not appropriately reflect the risk profile of a given organisation because of reporting bias or uniqueness, and hence particular care must be used with this data source.

- **Business environment and internal control factors (BEICFs)**: BEICFs are factors that capture a specific aspect of the risk associated with an organisation's operational process. They may be operational risk management indicators that provide forward-looking assessments of internal and external operational risk factors as well as risk controls. Incorporating BEICFs directly into models has traditionally presented challenges given their subjectivity, although techniques such as Bayesian methods have increasingly been used for this purpose successfully.

- **Scenarios analysis (SA)**: As outlined in Section 5, SA is a fusion of the expert knowledge of the current and future states of the business and the environment in which it operates, in order to assess operational risks.

6.2 Event type segmentation

Operational loss assessment is typically undertaken on a more granular basis than at the total level. This is because most organisations have a multitude of operational activities which involve many different types of input resources, business lines, processes, and product types, each of which can give rise to different types of operational events. Hence it is common for the framework to segment losses into as many homogeneous categories as is practical, given the constraints on data availability, data quality and the resources available to undertake the assessment process.

[9] See the ORX Association website at http://www.orx.org.

By way of example, Figure 9 outlines the risk event types in the ORX database for the global banking industry. Such event types are used to identify the nature of the cause of the loss, for each separate business line.

FIGURE 9: OPERATIONAL RISK EVENT TYPES

RISK TYPE	EXAMPLES
Internal fraud	Unauthorised activity; internal theft and fraud; system security internal – wilful damage
External fraud	External theft and fraud; system security internal – wilful damage
Employment practices and workplace safety	Employee relations; safe workplace environment; employment diversity and discrimination
Clients, products, and business practices	Suitability, disclosure, and fiduciary; improper business or market practices; product flaws; selection, sponsorship, and exposure; advisory activities
Disasters and public safety	Natural disasters and other events; accidents and public safety; wilful damage and terrorism
Technology and infrastructure failures	Power outage, network failure
Execution, delivery, and process management	Transaction capture, execution and maintenance; monitoring and reporting; customer intake and documentation; customer/client account management
Malicious damage	Vandalism

Another analogy is also pertinent here. In the health industry, the cause of death is often classified by what finally killed someone. However, unless the medical notes are very complete, most of the time this isn't very informative. The taxonomy of classifiers is meant to be a cause of death, but in reality it is only the very final straw and it makes no real sense to classify according to that. For example, a person with HIV who dies because of influenza. It would make infinitely more sense to focus more on symptoms, underlying conditions, or previous ailments that lead to death rather than the final nail.

Loss severity and frequency distributions would then be derived for each of the above risk event type categories.

6.3 Loss severity

6.3.1 Choice of distribution

Loss severity is the variable that measures the economic impact of a risk when it occurs. It is assumed that losses are independent and identically distributed according to some statistical distribution. Dutta and Perry (2007) suggest that the criteria that should be used when assessing the suitability of a particular distribution include:[10]

- **Realistic**: The distribution must reflect the actual nature of the underlying loss process. If the underlying process isn't stable or is adaptive in nature, then the LDA approach will be highly flawed from the outset.

- **Well-specified**: The distribution must be able to be meaningfully calibrated.

- **Flexible**: The distribution must be sufficiently flexible to deal with a number of alternative shapes.

- **Simple**: The distribution must be simple to use and not be more complex than is necessary.

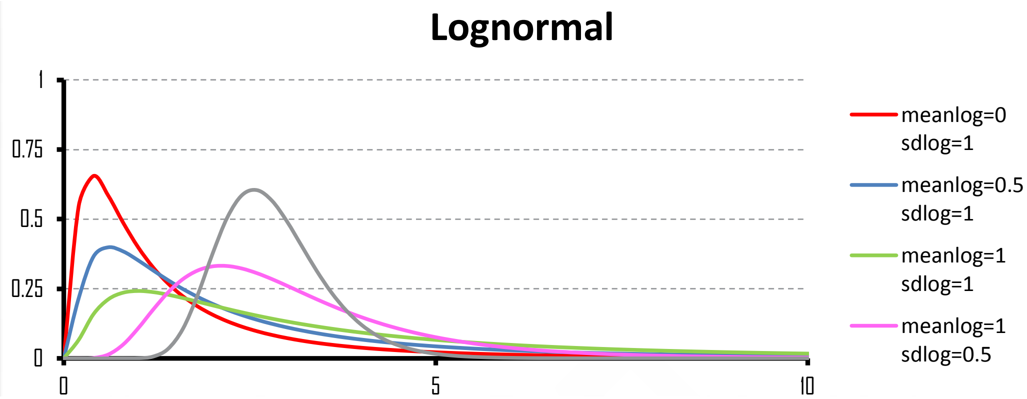
[10] Some general principles on appropriate choice of severity distributions can be found in the Basel Operational Risk – Supervisory Guidelines for the AMA.

The loss severity distribution should not be arbitrarily chosen based only on data fitting, but rather it should be derived using robust statistical theory. For example, Peak Over Threshold theory[11] and Extreme Values theory[12] show that under certain hypotheses variables converge to one particular distribution: the generalised Pareto distribution (GPD).

In general, loss severity can be modelled as a variable with positive real values, with a probability concentrated on zero for events that don't generate losses. The biggest concern when selecting the distribution is whether to use a light-, medium-, or heavy-tailed distribution, because this choice affects the final outcome dramatically. The most frequently used distributions to model severity are the lognormal, Pareto, gamma, and Weibull distributions. In addition, empirical distributions or splines are sometimes used.[13] Figures 10-13 outline the distribution probability density function (pdf) and graphs of calibrations of each of these distributions.

$$f(x;\mu,\sigma) = \frac{1}{x\sigma\sqrt{2\pi}}\,e^{\frac{(lnx-\mu)^2}{2\sigma^2}}$$

FIGURE 10: LOGNORMAL DISTRIBUTION PDF AND SAMPLE DISTRIBUTIONS



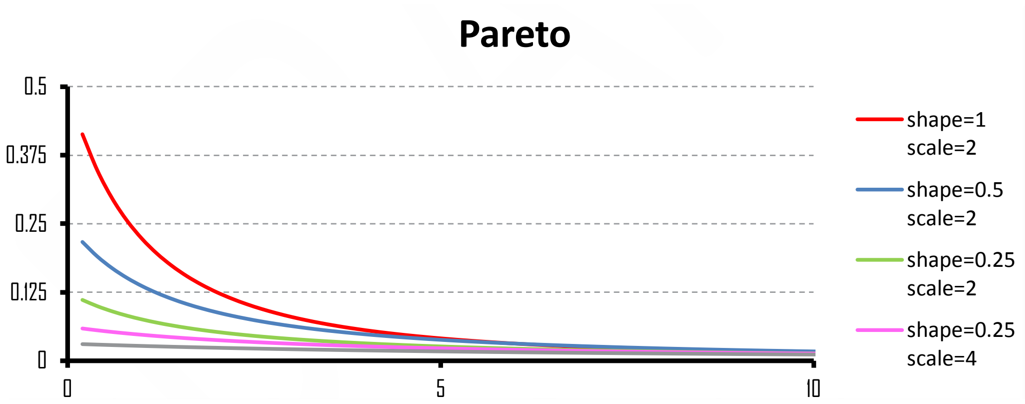
[11] Refer McNeil et al. (1997).

[12] Refer Embrechts et al. (1997).

[13] Refer Operational Risk – Supervisory Guidelines for the AMA, BIS, June 2011, paragraphs: 194 and 193-208.

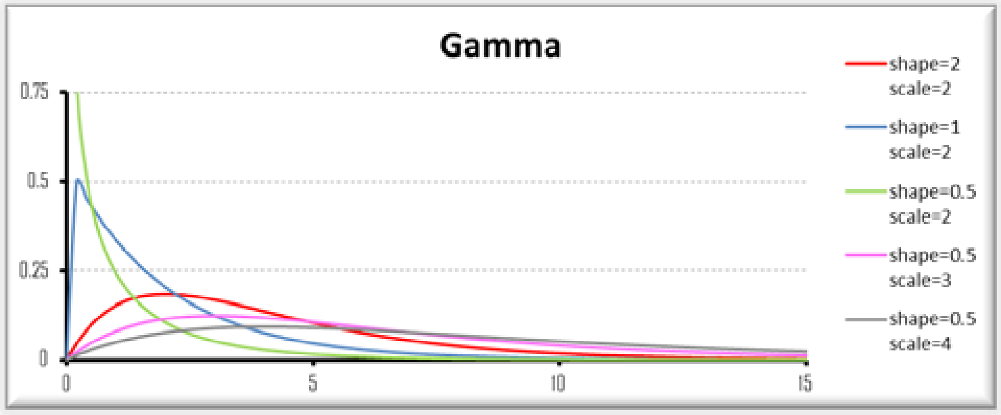
$$f(x) = \alpha \frac{k^{\alpha}}{x^{\alpha+1}} \quad \text{where } \alpha = \text{shape, } k = \text{scale}$$

FIGURE 11: PARETO DISTRIBUTION PDF AND SAMPLE DISTRIBUTIONS



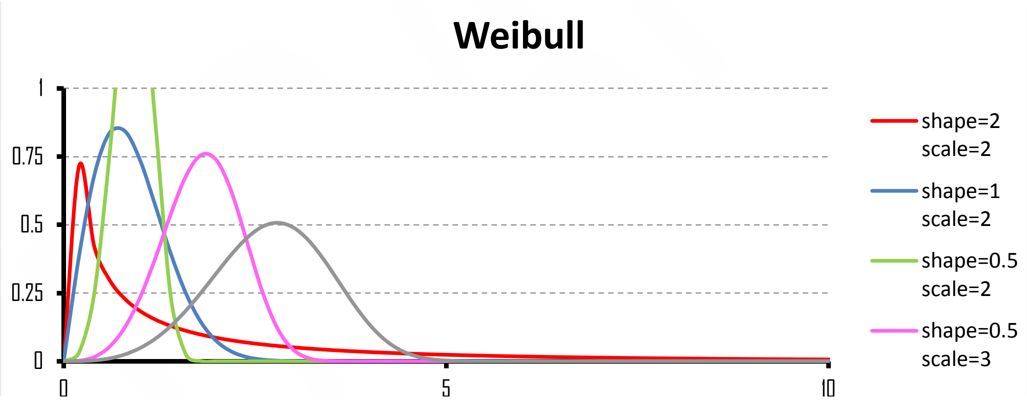
$$f(x; k, \theta) = \frac{1}{\theta^k} \frac{1}{\Gamma(k)} x^{k-1} e^{\frac{x}{\theta}} \quad \text{where } k = \text{shape; } \theta = \text{scale}$$

FIGURE 12: GAMMA DISTRIBUTION PDF AND SAMPLE DISTRIBUTIONS



$$f(x;\lambda,k) = \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k} \quad \text{where } k = \text{shape};\ \lambda = \text{scale}$$

FIGURE 13: WEIBULL DISTRIBUTION PDF AND SAMPLE DISTRIBUTIONS



6.3.2 Distribution segmentation

Distribution fitting to derive the loss severity distribution can be made using two or sometimes three different distributions to cover the low-, medium-, and high-severity losses. In this case, one or two thresholds are chosen and the distribution is defined and calibrated uniquely in each segment. This can enable more accurate fitting, but the final outcome should not be too sensitive to the choice of thresholds as that choice can be quite subjective. This technique is based upon the assumption that events are more appropriately modelled if they are grouped into multiple severity buckets.

This approach can be useful in capturing differences in the underlying nature of the operational loss process we are attempting to model. Low-severity, high-frequency events will be typical of an operational process that is relatively stable and involves a large number of independent homogeneous events. For this type of risk, high-quality loss data is typically available, and the use of normal and lognormal severity distributions can be appropriate.

At the other extreme, high-severity, low-frequency events will be typical of an operational process that is not stable, as it may be quite unique or new to the organisation, and that may involve a number of different types of control events simultaneously failing across multiple interconnected activities. Derivative losses by rogue traders are a good example. In this case, the use of a generalised Pareto distribution can be appropriate given the large uncertainties involved.

In between the two extremes, medium-severity, medium-frequency events will be typical of operational processes that, whilst they may not be highly repeatable, are relatively well understood and for which data in one form or another is available if judiciously used. In this case, the use of lognormal, gamma, and Weibull distributions is appropriate.

According to the Bank of International Settlements, which conducted a survey published in June 2011, the range of practice for AMA banks in operational loss severity modelling varies significantly: 31% of AMA banks used a single distribution for the entire curve, and nearly 50% used two separate distributions for the body–or high-frequency, low-impact (HFLI) region–and the tail, or low-frequency, high-impact (LFHI) region.

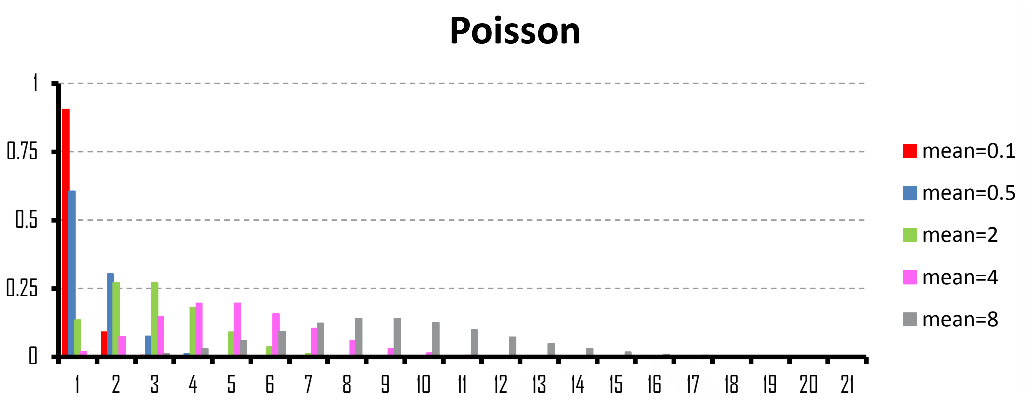
Loss severity distribution estimation is a key element in the LDA framework, and particular care must be taken in the truncation threshold if a piecewise segmentation approach is used. The threshold of truncation should be considered and estimated at the same time with the other parameters of the distributions, using statistical data fitting methods.

6.4 Loss frequency

Loss frequency (or likelihood) is the variable that measures how many times in a certain period a risk occurs. The time period typically considered is one year. The standard assumption is that the frequency of losses are independent and identically distributed according to a given distribution. By its nature this is a discrete variable, requiring a discrete distribution to estimate it. In general, every discrete distribution with values greater than or equal to zero could be used. However, common practice shows that the most popular distributions used are the Poisson and the negative binomial. Modifications of these distributions can also be used, for example compound and mixed distributions, as well as zero-modified and zero-truncated ones.[14] Figures 14 and 15 outline the distribution probability density function and graphs of calibrations of each of these distributions.

$$p_x = \frac{e^{-\lambda}\lambda^x}{x!} \text{ where } \lambda = \text{mean}$$

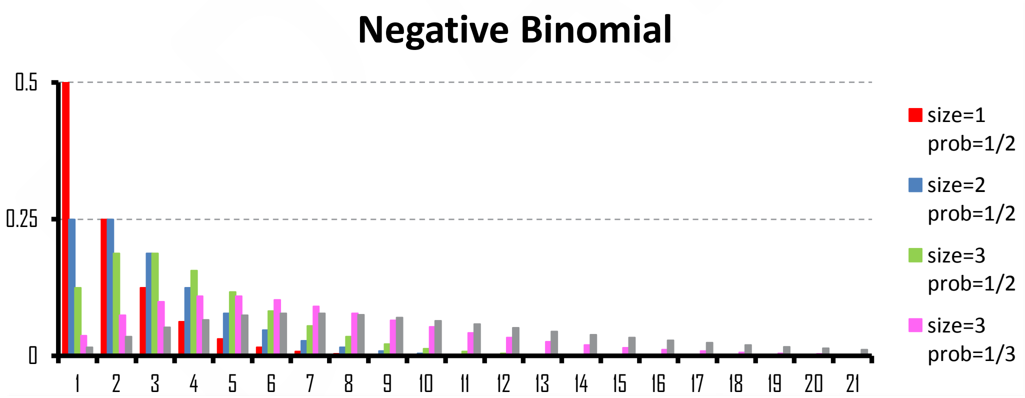
FIGURE 14: POISSON DISTRIBUTION PDF AND SAMPLE DISTRIBUTIONS



[14] Refer Chapter 6 of Klugman et al. (2008).

$$p_x = \binom{x+r-1}{x}\left(\frac{1}{1+\beta}\right)^r\left(\frac{\beta}{1+\beta}\right)^x \text{ where } \beta = \text{probability and } r = \text{size}$$

FIGURE 15: NEGATIVE BINOMIAL DISTRIBUTION PDF AND SAMPLE DISTRIBUTIONS



The fitting of the loss frequency distribution proceeds similarly to that of the loss severity distribution, based upon whatever data sources are appropriate and available.

6.5 Loss interrelationships

Interrelationships and dependence across the various elements of the loss assessment process may arise from exposure to common operational process elements or structural factors (e.g., people, businesses, processes, IT systems) or from environmental factors (e.g., a change in legal risk associated with certain business practices), both of which affect multiple areas of the firm. These factors can influence the observed frequency or severity of losses. However, the choice of approach for incorporating dependent interrelationships can have a significant impact on the results of the assessment process.

Dependence assumptions should be supported to the greatest extent possible by an appropriate combination of empirical data analysis and expert judgment. It is important to recognise that using internal and external data to model dependence presents challenges, as data limitations observed in the univariate context are likely to be more significant in the multivariate context. Using judgment to model dependence presents its own challenges, as eliciting accurate but subjective estimates can be more difficult in the multivariate context than in the univariate context. As such, the specification of dependence structures represents one of the most significant challenges in modelling operational risk.

Assumptions regarding dependence should preferably err on the conservative side, given the significant uncertainties. Consequently, the dependence methods considered should not be limited to those based on normal or normal-like distributions (e.g., Student's t-distributions with many degrees of freedom), as normality may underestimate the amount of dependence between tail events. The two most common methods for modelling dependence are correlations and copulas.

6.5.1 Correlations

Correlation is a measure of the degree of the linear relationship between two variables.[15] Note that this relationship is not directional in nature, and hence cannot be interpreted as a causal relationship: *Correlation does not imply causation!*

Under this approach, a decision needs to be made as to the structure of aggregation via correlation. Aggregation of the severity and frequency distributions for each risk event type and business line segment occurs via Monte Carlo simulation. This is the current approach typically used by AMA banking organisations, making the simplifying assumption of independence between severity and frequency distributions. The operational loss distribution for the entire organisation is then derived based upon the application of a correlation matrix covering each segment's loss distribution.

Alternative aggregation approaches could also be used whereby correlations are applied separately to severity and frequency distributions across the risk event or business line dimensions first, before then aggregating frequency and severity distributions together at a total level. The choice of structural method to use will depend upon the quality and sufficiency of data, the homogeneity of risk event types across business lines, and the nature of dependence between these elements. For example, there will be cases where the independence assumption between loss severity and frequency clearly does not hold, such as where poor compliance practices may lead to an increase in both the likelihood of fraud occurring, as well as the potential size of fraudulent losses. If this is considered material, then it should be reflected in the use of a correlation assumption and the aggregation structure.

6.5.2 Copulas

6.5.2.1 Overview

In cases where two variables exhibit more complex interrelationships than simple linear dependence, then a different approach is warranted. The use of copulas is one method that can account for nonlinear dependence in the tails of the marginal distributions of two variables.

From a mathematical point of view, a copula can be defined as a cumulative n-dimensional density function C, with standard uniform marginal distributions. Sklar's theorem states the following:

Let $F(x_1, \ldots, x_n)$ be a joint distribution with marginal functions $F(x_1), \ldots., F(x_n)$. Then there exists a function $C:[0, 1]n \rightarrow [0, 1]$ such that:

$$F(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n)$$

All continuous multivariate distribution functions contain a unique copula and that copula may be used in conjunction with univariate distribution functions to construct multivariate distribution functions. The following three properties must hold for this:

1. $C(u_1, \ldots, u_n)$ is increasing in each component u_i
2. $C(1, \ldots, 1, u_1, 1, \ldots, 1) = u_i$ is the marginal function of C_i for each $u_i \in [0,1]$
3. C is always nonnegative

A useful corollary derived from the uniqueness has the representation:

$$C(u_1, \ldots, u_n) = F(F_1^{-1}(u_1), \ldots, F_n^{-1}(u_n))$$

Copulas provide the user with flexibility in that they allow the fitting of any marginal distribution to different random variables, and these distributions may differ from one to another. Any joint distribution function can then be fitted (i.e., the copula across the marginal ones). Traditional methods

[15] Technically, this refers to the Pearson correlation coefficient.

of multivariate distributions require that all random variables have the same marginal distributions. If a multivariate normal distribution is fitted across a set of random variables, then we are forced to fit univariate normalised distributions to each of the marginal distributions. This yields to the definition of implicit copulas, such as the elliptical Gaussian copula, where the marginal distributions and multivariate distribution all follow the normal distribution:

$$C_\rho^{Ga} = \phi_\rho(\phi^{-1}(u_1), \ldots, \phi^{-1}(u_n)$$

where:
ϕ is Normal distribution
ρ is the correlation matrix

This is unsatisfactory when aggregating across different risks. Instead, copulas allow a much wider range of possible dependence structures. This allows the building of non-elliptical models using the implicit copulas mentioned above for the dependence structure and combining them with arbitrary margin distribution:

$$F^{meta\text{-}Ga}(x_1, \ldots, x_n) = C^{Ga}(F_1(x_1), \ldots, F_n(x_n))$$

If $x_1, x_2, \ldots, x_n$ represent losses that are due to different types of risk, then $F(x_1), F(x_2), \ldots., F(x_n)$ are arbitrary marginal functions (i.e., they are not limited to the normal distribution functions). The simulation of a copula dependence structure is best suited to aggregate different risk types. The shape of the marginal distributions of risks is not restricted to a specific form but can assume whichever one best represents the behaviour of the risk factors.

6.5.2.2 Types of copulas

A wide choice of copulas exists. The most commonly used are the Gaussian copula for linear correlation, the t-copula, the Archimedean copula for dependence in the tail, and the Gumbel copula for extreme distributions. These copulas are also known as parametric copulas as they have distributional assumptions. Nonparametric copulas have the advantage of not making any assumptions about the nature of empirical distribution functions.

However, the downside is that they require a large amount of data to calibrate, and are extremely computationally intensive to process. There are dimensionality problems inherent in these types of approaches and a smoothing method is required for the hypercube. The robustness of these techniques needs to be tested in each situation. There can also be poor behaviour in the tails of the distribution in situations where not much data exists, such as for high severity operational loss events.

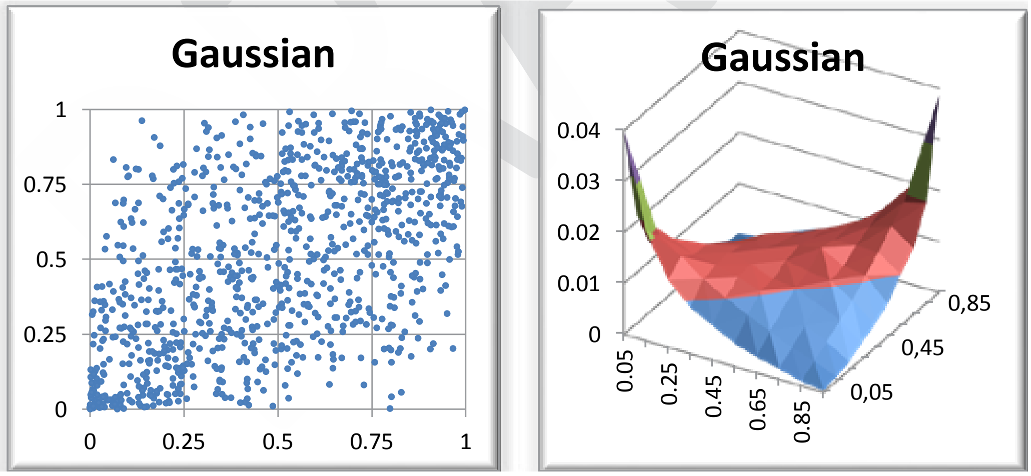
6.5.2.2.1 Gaussian copula

The process of obtaining the aggregate distributions through a Gaussian copula is described below:

- Generate independent normal random numbers Z
- These random numbers are correlated through the correlation matrix, using the Cholesky decomposition, A, of matrix R, i.e., set $x = Az$
- Calculate the normal cumulative probabilities . The vector $u_i = \phi(x_i), i= 1, \ldots., n$ is a random variate from the n-dimensional Gaussian copula $C_\rho^{Ga}(u)$
- The x_i (i.e., the loss of risk i) is determined by inverting the marginal distributions F_i: $x_i = F_i^{-1}(u_i)$
- Iterating this process and adding the losses x_i the whole integrated distribution can be calculated

Figure 16 illustrates the distribution of a Gaussian copula.

FIGURE 16: BIVARIATE GAUSSIAN COPULA DISTRIBUTIONS WITH RHO = 0.6



The advantage of the Gaussian copula is that, because of its familiarity, it is relatively simple to model. But the disadvantage is that operational loss datasets typically do not exhibit symmetry or exponentially decaying tail behaviour. Hence the normal distribution approximation may severely underestimate the loss experience in the tails.

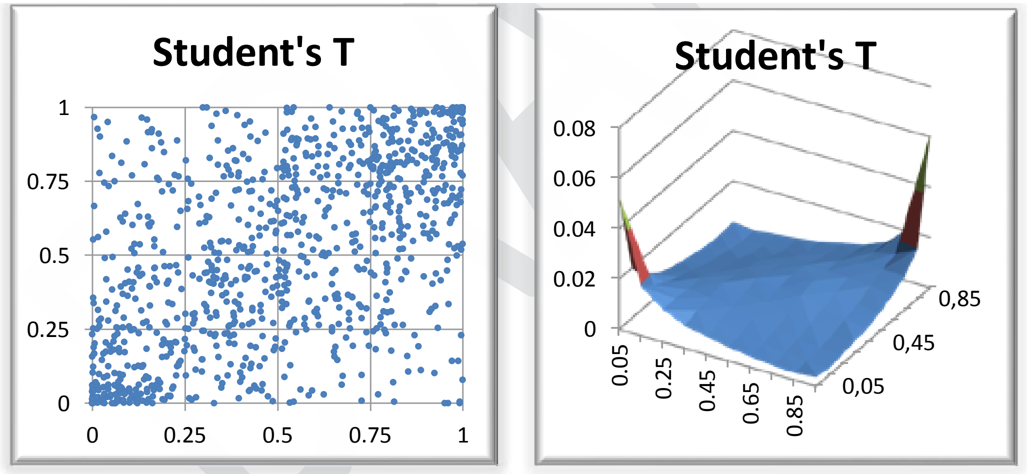
6.5.2.2.2 Student's t-copula

The t-copula displays heavier tails in comparison to the normal copula, with polynomial decay at the tails which are better able to capture fat-tailed operational loss events. It has the advantage that it is similar to the familiar Gaussian copula, but it has an extra parameter to control the tail dependency, which is the degree of freedom. Unlike the Gaussian structure where dependence is captured only via correlation, the t-dependence structure retains the use of correlation while introducing a single additional parameter, the degrees of freedom that determines the potential for extreme co-movements. In particular, increasing the value of the degrees of freedom decreases the tendency of the underlying distributions to exhibit extreme co-movements. A low value of the degrees of freedom gives high dependency at the tails and independence in the middle, whereas a t-copula with high degrees of freedom will behave like a Gaussian copula. However, the downside is that the t-copula still maintains the hypothesis of symmetry, even when the data may be asymmetric.

To simulate from a meta-t distribution, with v degrees of freedom, the procedure is similar to the meta-Gaussian copula described above. In addition, we need to simulate an additional random variable y from the distribution, which is independent from vector X, and use the transformation $Z\ y\ X = v$ to derive t-distributed random numbers. The arguments of $C_{\rho}^{Students's\ t}\ (u)$ are thus obtained calculating the t-cumulative probabilities $t(Z)$.

Figure 17 illustrates the distribution of a Student’s t-copula.

FIGURE 17: BIVARIATE STUDENT’S T-COPULA DISTRIBUTION WITH RHO = 0.6 AND NI=4



6.5.2.2.3 Archimedean copula

Archimedean copulas are a popular type of copula as they are able to model dependence in arbitrarily high dimensions through the use of only one parameter, which governs the strength of the dependence. An Archimedean copula can be written in the following form:

$$C(u_1, \ldots, u_n) = \psi^{-1}(\psi(u_1), \ldots, \psi(u_n)$$

for all $0 \leq u_1, \ldots, u_n \leq 1$ and where ψ is a function called the generator, satisfying the following properties:

$\psi(1) = 0$
for all $t \in [0,1]$, $\psi'(t) < 0$, i.e., ψ has a negative first order derivative , it is decreasing
for all $t \in [0,1]$, $\psi''(t) \geq 0$, i.e., ψ has a non-negative second order derivative, it is convex

The Gumbel copula described below is an example of an Archimedean bivariate copula:

$$\psi^{(t)} = (-\ln t)^{\alpha},\ \alpha \geq 1\text{: } C(u_1, u_2) = exp\{-[(-\ln u_1)^{\alpha} + (-\ln u_2)^{\alpha}]^{1/\alpha}\}$$

The Gumbel copula is the most popular copula for extreme distributions. However, although Archimedean copulas are simple to model for bivariate cases, the modelling requirements are complex for multivariate cases. Its difficulty is also made worse as the inverse generators do not exist for certain Archimedean copulas.

6.5.2.3 Choosing a copula

Where a copula approach can be justified because of the existence of material nonlinear dependence, the general approach would be to use the simplest copula that is able to appropriately capture the nonlinearity. This is because the more complex the copula, the greater the challenge in obtaining an accurate and robust calibration.

There is some evidence to suggest that the use of Student's t-copulas are appropriate for modelling correlated extreme events for economic capital purposes,[16] but significant effort may be required to assess the alternatives, which may be unique to the organisation and data available.

The main challenge with using a copula is that a very large amount of data is required in order to determine a robust calibration. This can be orders of magnitude larger that the datasets that are currently available for operational risk.

6.6 Aggregation and capital assessment

6.6.1 Aggregation

With an LDA, the total operational losses can be modelled as the sum of individual losses generated by all operational loss events over a given time period (typically one year). Formally:

$$S_i = \sum_{j=1}^{N_i} X_{i,j}$$

where:

S_i are the aggregated losses for the ith segmented risk category (e.g., by event type and/or business line)

$X_{i,j}$ is the amount of the losses of a single event of the ith segmented risk (i.e., severity)

N_i is the number of events that occurred over the time period (i.e., frequency)

Total operational losses S can therefore be expressed as:

$$S = \sum_{i=1}^{k} S_i$$

where: k is the total number of segmented risk categories/groups

To calculate the distribution for total losses for a single segmented risk category S_i, the following assumptions must be made:

1. Conditional on $N_i=n_i$, the random variables $X_1, X_2, \ldots, X_{ni}$ are independent and identically distributed.
2. Conditional on $N_i=n_i$, the common distribution of the variables $X_1, X_2, \ldots, X_{ni}$ does not depend on n_i.
3. The distribution of N_i does not depend in any way on the values of $X_1, X_2, \ldots, X_{ni}$.

Generally, an analytical formula can't be derived for S_i, so computational techniques are required in order to undertake the convolution of the distributions. Several methods are illustrated in the literature, but the most widely used is Monte Carlo simulation. The simulation process can be summarised in the following four steps:

1. Build a model for S that depends only on random variables X_{ij}, N_i where their distributions and any dependencies are known.

2. For $k=0,\ldots,m$, generate pseudorandom values of frequency n_i, and then for $j=1, \ldots, n_i$ generate severity values x_{ij} and compute S_i using the model of Step 1.

[16] Refer Cornaglia et al. (2007).

3. The cumulative distribution function of S_i may be approximated by $F_{n(i)}(s_i)$, the empirical cumulated distribution function based on the pseudorandom samples $s_1,\ldots,s_{n(i)}$.

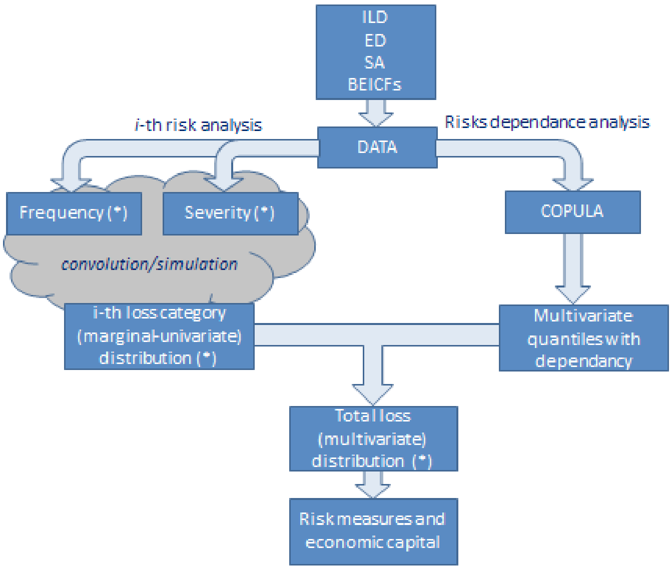
4. Compute the statistical results of interests using the empirical cumulative distribution function.

To compute the total losses, the distribution for all the risk category quintiles are generated from the correlation/copula and transformed by inverting the marginal distribution (i.e., single risk distribution).

The final outcome of the LDA is a set of pseudorandom data from the compound distribution of the total losses, obtained by the estimated marginal distributions with estimated dependence modelled by the copula (i.e., for each simulation, a vector with generic element equal to the total loss over the time horizon; e.g., one year) of the ith risk category and length equal to the number of risk categories. From this data an empirical cumulated distribution function is obtained based on the empirical frequency of data.

From a process perspective, the main elements of implementing an LDA approach are summarised in Figure 18.

FIGURE 18: LDA IMPLEMENTATION PROCESS


ILD
ED
SA
BEICFs
i-th risk analysis
DATA
Risks dependance analysis
Frequency (*)
Severity (*)
COPULA
convolution/simulation
i-th loss category (marginal-univariate) distribution (*)
Multivariate quantiles with dependancy
Total loss (multivariate) distribution (*)
Risk measures and economic capital

The calibration of each distribution needs to be validated with a statistical test. The items marked with an asterisk (*) in Figure 18 need to pass a validation process. This will ensure that the distributions used are appropriately calibrated to the available data.

6.6.2 Capital assessment

The cumulative distribution function calculated using the above process can be used to determine a number of risk statistics, which can be considered as appropriate for the purposes of economic capital.

In both Basel II and Solvency II regimes, economic capital is a risk measure based on a Value at Risk (VaR) statistic. Under Basel II this is set at the 99.9th percentile level of the total annual loss distribution (i.e., 1-in-1,000-year event), whilst under Solvency II it is set at the 99.5th[17] percentile level of the total annual loss distribution (i.e., 1-in-200-year event).

Whilst VaR measures have had much use as risk measures in finance, it is important to note that it is not a coherent risk measure. Coherent risk measures must have the following desirable properties:

- **Monotonicity**: If A always has better outcomes than B, then the risk of A should be less than B.

- **Sub-additivity**: If two risks are added together, their combined risk cannot be any greater than the sum of the two risks separately. This is the principle of diversification.

- **Homogeneity**: The application of a scalar factor to the underlying variable should similarly scale the risk measure.

- **Translational invariance**: Adding a risk-free element to the variable reduces risk by the same amount

If a risk measure isn't coherent, and when distributions are highly skewed, heavy-tailed, or far from normality, problems can arise such as negative diversification benefits. For example, say the VaR from each independent source A and B is 100 and 50, respectively. But when combined into A + B, the VaR could be 200. The non-sub-additivity effect is also emphasised by some distributions with heavy right tail asymmetries, such as the lognormal (discussed above), which is also one of the most commonly used distributions to model severity.

In the ORX database of global operational banking losses, Cope et al. (2009) observed that losses in different business lines are often quite variable in terms of both frequency and severity. For example, in retail banking, losses are more frequent but loss amounts rarely exceed EUR 1 million, whilst in corporate banking, losses are very rare but can be as high as hundreds of millions or billions of euros. So a high severity low frequency loss category will be modelled using few data points, resulting in extremely large expected losses that are also extremely uncertain. As economic capital measures are high quantile measures (e.g., 99.9th percentile VaR), they are effectively dominated by the highly uncertain measurements of high severity low frequency types of events.

The main reason for the adoption of VaR measures is that no perfect alternative exists. Whilst the conditional tail expectation (CTE, also known as TailVaR) is a sub-additive measure, it is highly dependent upon the very extremes of the tail distribution, and cannot be easily interpreted as a 1-in-n-year event, unlike VaR. Use of CTE measures requires a high degree of modelling certainty in the extremes of the tail, which is typically not the case with operational risk, where there is a paucity of data in tails to calibrate to.

6.7 Advantages and limitations of an LDA

The advantages of an LDA approach include:

- The use of well-known statistical distributions can aid the calibration process.

- Splitting losses into frequency and severity components can enable more accurate assessments of these drivers, which are typically characterised by different statistical processes. It can also give a better understanding of the nature of the losses.

- The impact of risk control mechanisms, such as deductibles and insurance, can be readily included.

- Building, implementing, validating, and operating an LDA model does not require large computational resources.

[17] Based on QIS 5, which is current as at the date of this publication.

- It is reasonably flexible as it can adapt to new and evolving business operations, and it can be extended to multiyear periods.

The limitations of an LDA approach include:

- It requires stable underlying processes and relationship structures that are simple and relatively linear. Where processes are adaptive and interrelationships complex and nonlinear, then statistical approaches can be highly flawed.

- It is heavily dependent upon sufficiently credible calibration data. In the absence of this, the credibility of results deteriorates significantly. This is particularly relevant for the material high severity low frequency events, and for new organisations with little to no history of data, or with new types of operational processes and controls.

- Statistical approaches rely on the underlying assumption that the nature of the underlying process is stable. This may not be the case for organisations where operational activities are changing, becoming more or less risky in the process.

- The integration of internal data, external data, scenario analysis, and expert judgment in the framework can be challenging.

- Results can be very sensitive to the range of difficult subjective choices that need to be made around:

  — Segmentation of data into homogeneous risk categories
  — Segmentation of the loss curves into high/medium/low-severity/frequency classes
  — Methods to capture and model dependencies

- The typical assumption of independence between frequency and severity distributions is a major limitation. The use of copulas and/or correlations between different segments of the curves can be equally difficult to calibrate.

- It is difficult to relate results back to the states of business drivers, which creates challenges in the communication and engagement of the business with the model.

There is a range of research on the use of LDA to the modelling of operational risk in the banking industry. Moscadelli (2004) and Evans et al. (2007) found that, because of the highly skewed nature of operational loss data, conventional frequency and severity models used in LDA are unable to provide adequate results in describing the loss data, particularly in the extreme percentiles. Cope et al. (2009) concluded that actual operational risk capital estimates should be handled with care, which is due to the instability of estimates and the sensitivity of results to extreme data points.

From an experience perspective, many banks had AMA approval for LDA models over the period of the global financial crisis of 2008 and beyond. What was found was that the level of operational risk capital was relatively insensitive to these conditions as they occurred, which clearly didn't make sense when many banks suffered significant operational failures.

7 STRUCTURAL OR CAUSAL APPROACHES

7.1 Introduction

A structural modelling framework conditions loss outcomes upon the states of underlying business drivers that cause those events to occur. Hence they are also synonymously referred to as causal models.

This framework synthesises a range of methods that have been developed and successfully used in the science of complex adaptive systems. The underlying structure of the framework is based upon an assessment of the *system* to which it relates. By *system*, we don't mean an IT system, but rather the complex system that relates operational outcomes to the important business drivers and risk factors upon which they depend. By incorporating directly the business drivers into the modelling framework, it is possible to clearly explain how loss outcomes are driven by the states in which business drivers can exist.

The broad process that is used to construct a structural operational modelling framework includes:

1. Eliciting the system structure that drives operational risk events.
2. Identifying the most critical business drivers and risk factors within this system.
3. Determining the appropriate states of business drivers and risk factors.
4. Defining the causal relationships across business drivers, risk factors, and outcomes.
5. Aggregation of operational risk elements.
6. Analysis and usage.

We will step through each of these processes in turn in the following subsections. Case studies are provided in the appendices which outline tangible examples of this approach.

7.2 Eliciting the system structure

Traditional operational risk management frameworks have been highly abstract, which has caused problems engaging with the people who spend most of their time working in an operational environment. Such nonspecialists in risk typically find risk management concepts very difficult to understand as they are very different from the concepts they deal with on a daily basis. In order to gain traction and engagement with these key groups so that we can leverage their expertise and knowledge, it is critical that the language that operational risk frameworks and risk models are couched in is that of the business.

In order to achieve this, we start holistically assessing the systems structure of a business and its operations through the use of cognitive mapping techniques. Cognitive mapping was developed in the early 1990s based upon George Kelly's theory of personal constructs. Essentially this theory suggests that people attempt to make sense of the world in order to predict what the future state of the world will be, and therefore decide how to act now in order to achieve the outcomes they desire. Early pioneers of the cognitive mapping technique were Fran Ackermann, Colin Eden, and Steve Cropper.

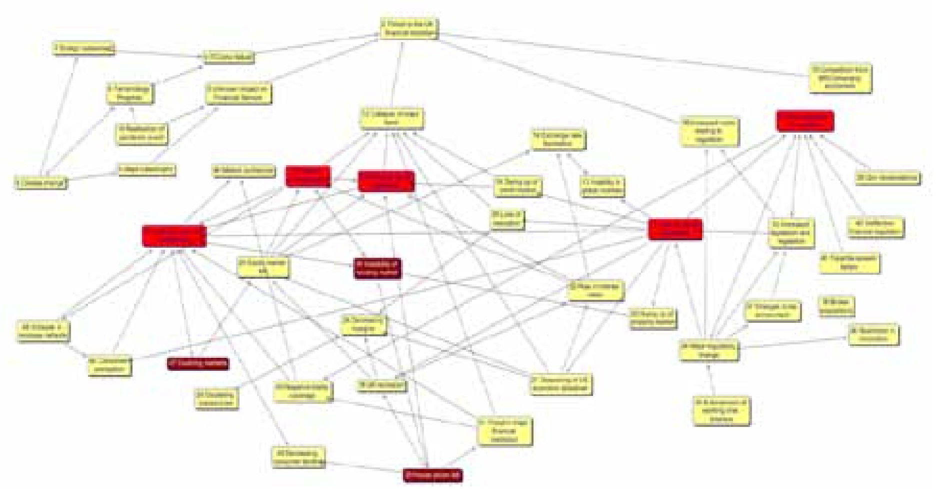
Essentially, cognitive mapping involves breaking down an account of a problem into its constituent elements. These are the distinct concepts which are connected to visually represent the narrative of the problem. This nonlinear representation of the narrative is the first step in making the underlying complexity of a business or an operational system transparent. The sense of hierarchy is important in constructing these stories and involves identifying which parts of the story lead to other parts of the story.

If you imagine a group of friends going to watch a film at the cinema and then comparing notes afterwards to discuss interesting parts of the film, each person will remember slightly different details. If you use a cognitive map to represent each piece that a friend recalls and the manner in which that piece connects to the others, then you will rapidly be able to reconstruct the entire movie, with all its subtle features, as a cognitive map.

One key feature of a good cognitive map is the ability to identify goals. These provide a natural end for the stories and give us an indicator about whether a particular element of the story has reached its conclusion or whether some of the detail has been missed.

Although the picture in Figure 19 is intentionally too small to read the individual elements, it shows how a cognitive map can become quite complex as a story is told.[18]

FIGURE 19: EXAMPLE OF COGNITIVE MAP



In this format we are able to apply mathematical techniques from graph theory to identify the parts of the map which seem to be particularly *connected* to the rest of the picture. This might mean that lots of other concepts are instantly linked to them, or that they connect over a greater distance. In either case we can be sure that, for some reason, these concepts *matter* to the story being told.

It is also possible to identify which story beginnings most often lead to these key story elements. Such concepts are clearly important as spotting evidence of their occurrence would give us a clue that one of the *important* parts of the story might happen later. In the context of an operational risk scenario, these events are the emerging risk signals that we want to watch for.

We can also spot *holes* in stories. If something jumps uncomfortably from an early stage in a story to a later stage, we would find that incompleteness and ask whether we can fill in the missing details. This could represent a lack of understanding about how a particular scenario might emerge or missing information about how different factors interact to produce a particular outcome. Either way it suggests an area for further investigation.

It is also possible to identify some of the dynamic behaviours in a story. Where one thing leads to another, which leads to another, which leads back to the first thing, we end up with forces working to keep something in balance or tending to push things out of control. Understanding such dynamics is critical to understanding how risks may propagate and potentially be controlled or avoided.

[18] Cognitive maps produced by Milliman using Banxia Decision Explorer software.

The use of this technique can be particularly powerful when undertaken across a number of individuals and business silos. This is because concepts and causal links which are correct are more likely to be reinforced by others than those that are not. Hence, as the process is repeated, different perspectives are integrated and the map will tend to reflect the common set of perspectives more than those that are wrong or incomplete. This is the mechanism through which individual bias is reduced dramatically. Like any model, it is trying to capture what is most important rather than everything, and the 80/20, or 90/10, rule strongly applies in this process.

Cognitive mapping therefore presents a powerful tool for combining the insights of a diverse group of experts into a holistic overview of a complex situation such as a business operation. Careful analysis permits the removal of significant amounts of cognitive bias or other diversionary contributions in the interviews so that the final outcome is close to the underlying reality of the system dynamics as understood by the experts.

Using this technique enables all the relevant concepts that influence an operational activity, production process, or operational risk scenario to be elicited. These concepts would include all of the following:

- **Business objectives** such as return on equity (ROE) and growth target, customer satisfaction levels, social responsibility

- **Business drivers (or business environment factors)** such as the quality of processes, effectiveness of systems, culture, the availability of quality people and capital resources

- **Outcomes and impacts** such as sources of financial loss, damage to capital resources, accidents to staff, damage to societal resources such as the environment

- **Risk factors** such as economic and capital market variables, physical environmental conditions, safety standards and practices

- **Internal control factors** such as quality of audit process, effectiveness and maturity of governance, appropriateness of systems, process controls, adequacy of staff training, segregation of duties, documentation requirements

7.3 Identifying the most important components of the system

Most cognitive maps that are developed are highly complex, containing a large number of concepts and a tangled web of causal relationships. Most people, when faced with this information, look to simplify this complexity. But such reductionist approaches tend to destroy most of the information value within a map. Instead, the challenge is to elicit the most important structures and dynamics of the cognitive map, retaining the complexity and information content in the process. In order to do this, we use graph theory to analyse the map as a network.

Like any network, we want to understand its structure and, most importantly, identify the most connected concepts. Highly connected concepts are central to the story. From an operational risk perspective, they represent business drivers that are central to the functioning of an operational activity. An example of this might be the various dimensions to the effectiveness of communication (i.e., information transfer), whether it is between systems, people, or businesses. Graph theory is readily able to identify these concepts objectively using a range of measures of connectedness, including the number of direct immediate connections, to a weighted number of connections dependent upon the degree of separation. These concepts have been coloured bright red in Figure 19 above.

Once the critical concepts (business drivers, risk factors, etc.) have been identified, then we can also identify what other concepts are most connected to them. They are the most potent drivers/risk factors within the business or operational activity that influence multiple critical drivers. Again, graph theory can objectively identify them.

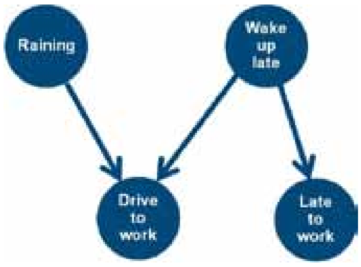
The objective analysis of the cognitive map is thus capable of identifying the subset of concepts that represents the skeletal framework of a business or operational activity through which both valuable outcomes are generated and risks are propagated. This analysis will typically result in 80%+ of the answer. With the application of expert judgment from both a risk management and subject matter expert perspective, the remaining 20% can typically be quickly identified.

It is upon this subset of concepts that a practical quantitative modelling framework can initially be built.

7.4 Bayesian networks: A quantitative framework for integrating information

A Bayesian network is a visual description (formally, a directed acyclical graph) of the relationships between causes and effects. It is made up of nodes and arcs such as those as shown in Figure 20.

FIGURE 20: EXAMPLE OF SIMPLE BAYESIAN NETWORK



Each node in the network represents a *variable* and the arcs represent the causal relationships between the variables. Bayesian networks permit the modeller to capture their reasoning about how a particular outcome may arise. The model does not say that a particular outcome will occur but is able to express how likely that outcome is, conditional upon any supporting events having taken place.

Bayesian networks use Bayes' theorem to compute the probabilities in the model. Bayes' theorem was developed over 250 years ago by Thomas Bayes. Formally it is written as follows:

$$p(A|B) = \frac{p(B|A)}{p(B)} \times p(A)$$

where:

$p(A|B)$ is the *posterior*, the probability of event A occurring given that event B has occurred

$p(A)$ is the *prior*, the probability of event A occurring

$p(B|A)/p(B)$ is the *evidence*, the probability of event B occurring given that event A has occurred, divided by the probability of event B occurring

Note that this is a discrete probability statement. Knowing any two of these components means that the third can be determined.

Let us consider a simple example based upon the network above. Assume that each node can be Yes or No, and assume that we have the probability of outcomes for *late to work* based upon the states of *wake up late*, as outlined in Figure 21.

FIGURE 21: CONDITIONAL PROBABILITY DISTRIBUTION OF *WAKE UP LATE*

WAKE UP LATE	NO	YES
LATE TO WORK		
NO	0.9	0.3
YES	0.1	0.7

The distribution for *drive to work* is slightly more complicated as it depends upon two factors rather than one, as outlined in Figure 22.

FIGURE 22: CONDITIONAL PROBABILITY DISTRIBUTION OF *DRIVE TO WORK*

RAINING	NO		YES	
WAKE UP LATE	NO	YES	NO	YES
DRIVE TO WORK				
NO	0.8	0.4	0.3	0.1
YES	0.2	0.6	0.7	0.9

Let us also suppose that there is a 30% chance of it raining and a 40% chance of waking up late. We can now ask basic questions such as whether it is more or less likely that this person woke up late if we know the person was late to work.

We start with:

$$p(oversleep)=0.4.$$

From Bayes' theorem we know that:

$$p(oversleep \mid late) = p(late \mid oversleep) \;.\; p(oversleep) / p(late)$$

$$p(oversleep \mid late) = 0.7 \times 0.4 / p(late)$$

We can indirectly determine,

$$p(late) = p(late|oversleep)\ p(oversleep){+}p(late|not\ oversleep)\ p(not\ oversleep)$$

$$p(late) = 0.7 \times 0.4 + 0.1 \times 0.6 = 0.34$$

So, $p(oversleep \mid late)$ = 0.82. This means that, if we know the person was late, then the chance the person overslept is significantly higher than the original estimate of 0.4 for oversleeping.

We can also consider what we can infer about the chance of driving to work. This is more complicated as it depends on more factors. The original probability of seeing the person drive to work is:

$$
\begin{aligned}
&p(drive) = p(drive|rain,\ oversleep)\ p(rain)\ p(oversleep)\\
&+ p(drive|rain,\ not\ oversleep)\ p(rain)\ p(not\ oversleep)\\
&+ p(drive|not\ rain,\ oversleep)\ p(not\ rain)\ p(oversleep)\\
&+ p(drive|not\ rain,\ not\ oversleep)\ p(not\ rain)\ p(not\ oversleep)\\
&= 0.9 \times 0.3 \times 0.4 + 0.7 \times 0.3 \times 0.6 + 0.6 \times 0.7 \times 0.4 + 0.2 \times 0.7 \times 0.6\\
&= 0.486
\end{aligned}
$$

But our revised posterior estimate $p(oversleep) = 0.82$ and $p(not\ oversleep) = 0.18$, and so $p(drive) = 0.6288$. Hence we see that the evidence of this person being late also increases our estimate of that person's chance of driving to work.

Over the past decade or so, computer algorithms have been developed which make the propagation of evidence through a complex Bayesian network easy to achieve. Evidence can be propagated both up and down a network to explain a consequence or to examine the consequence of a cause. It is now possible to implement quite large Bayesian network models which can propagate evidence virtually instantly. This provides a significant advantage over Monte Carlo methods for the purposes of sensitivity testing, stress testing, reverse stress testing, and *what if* analyses.

Armed with this technique we are now ready to construct an operational risk model.

7.5 Determining appropriate states of business drivers and risk factors

The outcome of the cognitive mapping analysis has identified the causal drivers upon which business and operational outcomes are conditional. That is, outcome A is dependent upon drivers B and C. The next step is to define the states of each of these concepts.

When most people communicate a story about how a business or operational activity works, they use natural language, whether in verbal or written form. Such language almost always uses discrete states to describe a driver or risk factor of a business. Consider Figure 23, where discrete states can be derived from very common statements that business people use to communicate.

FIGURE 23: EXAMPLES OF STATE DEFINITIONS BASED UPON COMMON BUSINESS COMMUNICATIONS

STATEMENT	IMPLIED DRIVER STATES
Credit market conditions are extremely stressed.	Stressed, normal
The outlook for economic growth is negative and highly uncertain.	Depression, recession, low, normal, strong
Market liquidity has dried up.	Illiquid, constrained, liquid
We are currently in a wet season, which has broken the recent drought conditions.	Wet, normal, dry
I just don't have sufficient quality resources to get the job done.	Sufficient/insufficient numbers; adequate/inadequate quality
We are at systems capacity and can't handle any more throughput.	At capacity, nearing capacity, below capacity
If Tom, Dick, and Harry all resign or became unavailable, we are in deep trouble as they have 90% of our operational intellectual capital.	Intellectual capital available or unavailable
Our modelling assumptions are inappropriate.	Appropriate, minimally sufficient, flawed
The accuracy of our data is poor.	High/medium/low quality
Capital requirements are too high for this product, it won't provide a sufficient return on shareholder equity.	Too high/OK but constrained/acceptable

In almost all of these cases, the unit of measurement of the driver under consideration is a continuous distribution. For example, credit market conditions could be measured by credit spread levels or systems utilisation levels as a percentage of capacity. Whilst it would be possible to define these drivers directly in terms of these continuous units of measurement, it is potentially misleading to do so. This is because humans are not very good at statistics. We all have a natural tendency to segment detail into discrete states and attach a label to each one of them. The point of this is that each state of a driver influences other drivers in a direct way. For example, it doesn't really matter if actual GDP is -8% or -8.5%, but rather that we are in a depression, and the way in which the components of the economy and capital markets behave has fundamentally shifted relative to being in a normal growth phase.

By defining drivers using a minimally sufficient number of states, we are able to relate value and loss outcomes to the states in which the business finds easiest to communicate. Thus it is possible to leverage to the fullest extent possible the information value contained across the spectrum of business communications. By being couched in natural business language it enables business and detailed operational people the maximum opportunity to be engaged in the framework.

7.6 Defining causal relationships

The central question that goes to the heart of building a risk management and modelling framework is:

Do I have any information upon which to condition an outcome that I am interested in, and what is its quality level?

If no information exists, or if the information that is available has no value in it, that is, it is of poor quality, then there is no point in adding anything more to the modelling framework. The process of building a model can continue up until this point is reached. Of course, other constraints also play an important role in defining the appropriate place to stop (e.g., computational limitations, information collection constraints) but in this section we are only considering the pure value of information.

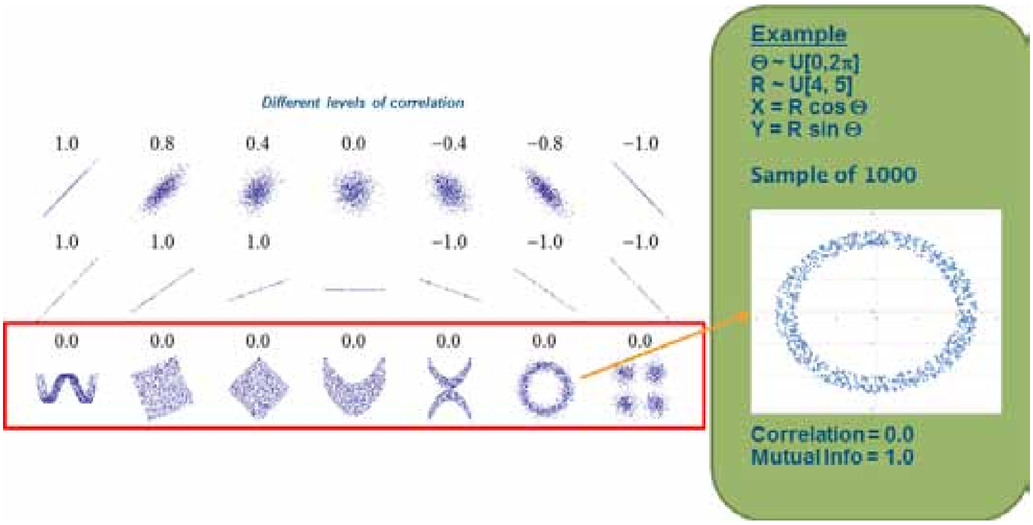
Perhaps not surprisingly, the science of information theory provides some guidance on how to answer this question. The two important concepts are:

1. **Mutual information**: A measure which quantifies the amount of information in common between two random variables.

2. **Entropy**: A measure which quantifies the uncertainty involved in predicting the value of a random variable.

Importantly, mutual information captures the full range of nonlinear dependence, as opposed to a correlation measure which only captures linear dependence. Figure 24 on page 42 provides some examples of the limitations of correlation measures in capturing nonlinear dependence structures.

FIGURE 24: MUTUAL INFORMATION VERSUS CORRELATION



Entropy is a statistical measure of uncertainty, and hence is akin to the quality associated with information. An important property of entropy is that it is maximised when there is no information available on the likelihood of an outcome.

Both mutual information and entropy measures are higher-order concepts compared to their better known and simpler, yet more limited, counterparts of correlation and variance. With these concepts and a sufficiently creditable dataset, we can use these methods to determine the distributions of business drivers and outcomes such as loss severity and loss likelihood. Mutual information enables insight into what types of risk, control, or performance indicators will be most useful in explaining outcomes, thus alleviating the need to include all potential indicators.

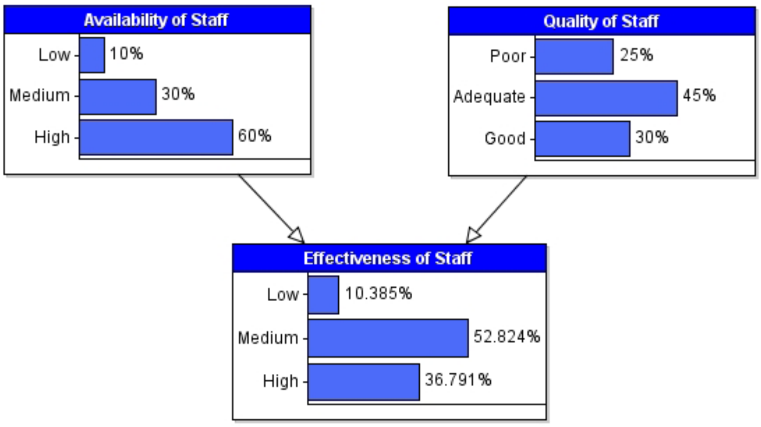
However, what are we to do when we don't have access to a credible dataset? In that case, it is necessary to use any available information that has at least some quality from any of the following sources:

- Historical internal data
- Historical external data
- Construction and collection of new data that is appropriate
- Direct input from management, risk analysts, and/or subject matter experts (SMEs)
- Constraint propagation using Bayesian inference techniques

Where direct input from the business is used, it should be focused on capturing the relationships around the inflection points in the behaviour of drivers and risk factors. For example, *The effectiveness of staff will be in a poor state when either the availability of staff is low, or the quality of staff is poor. Otherwise staff effectiveness will be in a good state.* These conditional statements tend to be much easier for people to understand, and the nature of the relationship can be identified (in this case it is weighted more heavily toward the bad outcomes). Figure 25 shows this simple dependent relationship structure.[19]

[19] Bayesian Network analysis produced by Milliman using AgenaRisk software.

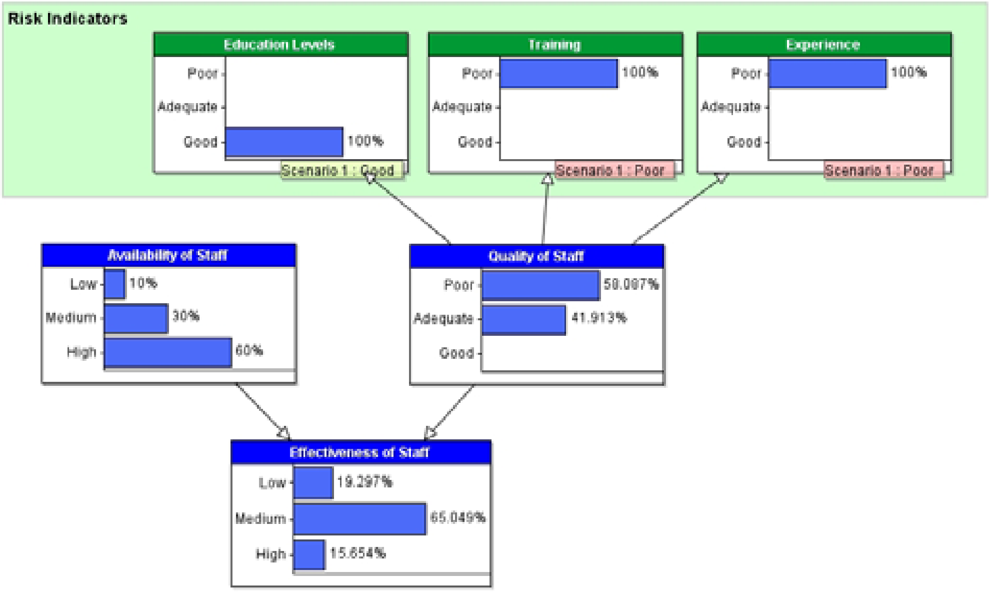
FIGURE 25: EXAMPLE OF A BAYESIAN RELATIONSHIP STRUCTURE FOR STAFF EFFECTIVENESS



In this example, the effectiveness of staff is modelled as a truncated normal distribution with a mean that is a weighted function of the means of the independent factors, and with a variance that reflects the level of uncertainty related to an assessment of the quality of the relationship. Example distributions are shown for the independent factors, which combine through this relationship function to determine the distribution of the dependent variable, effectiveness of staff.

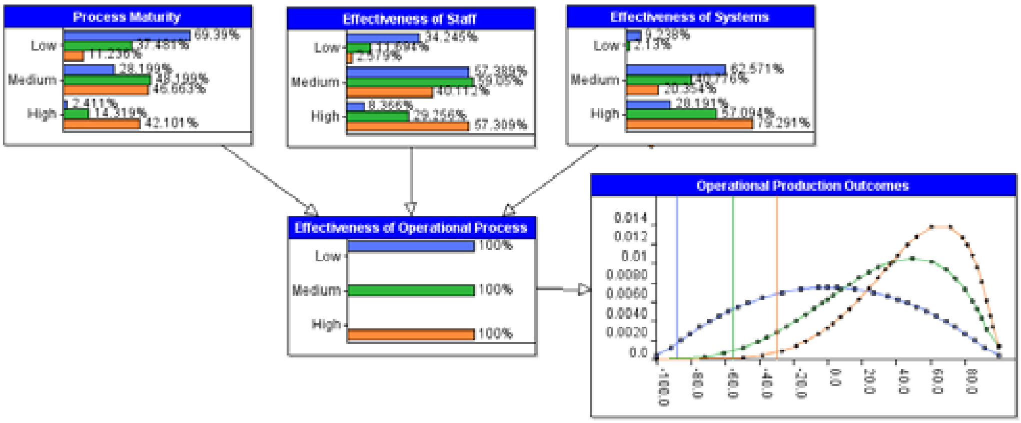
The independent risk drivers of *availability of staff* and *quality of staff* can then be relatively objectively assessed using a range of observable risk indicators. Figure 26 shows the inferred distribution results when education, training, and experience are added in as indicators of staff quality along with a simple dependency relationship.

FIGURE 26: INTRODUCTION OF MEASUREABLE KEY RISK INDICATORS



In this example, although education levels might be assessed as being in a good state, if training and experience levels are poor, then the inferred distributions of staff quality and hence the effectiveness of staff deteriorate significantly. We can continue this process by bringing in the other parts of the framework, ultimately resulting in a conditioned set of outcomes. In Figure 27, we have used various beta distributions to model operational production outcomes as a function of the effectiveness of the production process.

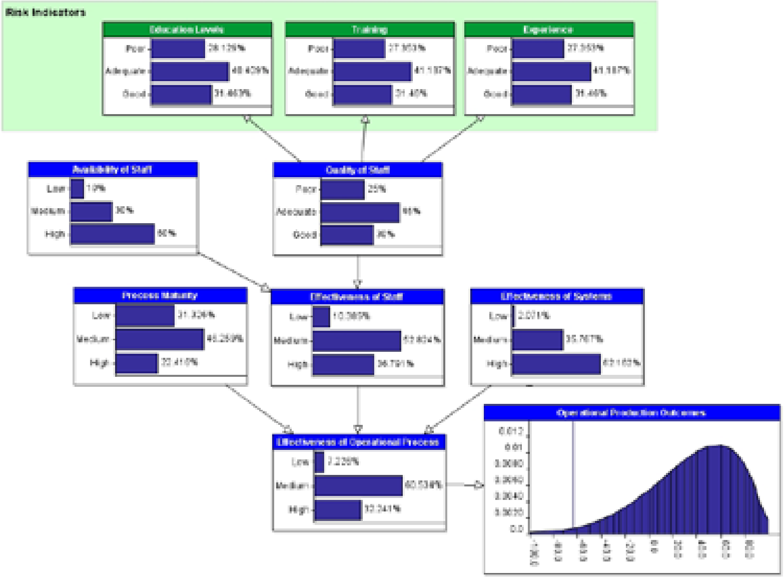
FIGURE 27: LINKING DRIVERS TO OUTCOMES



This figure shows the relationships between drivers and outcomes for three scenarios relating to three levels of the Effectiveness of Operational Process.

The distribution and 99th percentiles are shown for three scenarios, each relating to the states of the effectiveness of operational processes. Figure 28 shows the entire connected model in its base state with no evidence present.

FIGURE 28: FRAMEWORK OF OUTCOMES



This Bayesian framework now links operational outcomes to business drivers and ultimately all the way through to individual risk indicators. The calibration process reflects the full range of uncertainty inherent in the underlying operational business risk drivers. From this it is now possible to estimate operational risk capital at a range of lower percentile levels (e.g., -63.1 at the 99th percentile level).

The above simplified example effectively combines a range of different sources of information and conditions loss outcomes upon them. Implicit within this is a definition of time to which both the business drivers and loss outcomes relate, for example the definition of the *availability of staff* could be a forward-looking assessment over the next 12 months. However, it would be just as easy to define this over a shorter or longer period, and then introduce the traditional concepts of likelihood and severity to define the temporal and scalar dimensions separately. One of the values of doing this is that it then enables common causal drivers of each to be identified and conditioned upon. For example, nearing or reaching systems capacity levels might have an impact on both the likelihood of problems occurring, which might lead to increased pressure and strain on staff effectiveness, which leads to high severity losses. These causal links can be extremely effective in identifying, understanding, and explaining the sources of interdependence between likelihood and severity distributions.

A Bayesian framework also has the ability to encompass the full range of data, distributions, and simulations that are required to undertake operational risk capital assessments. Where available, both internal and external data can be used to not only calibrate the end loss distribution (or likelihood and severity equivalents), but it can also be used to calibrate the distributions of the underlying drivers. Many companies are collecting large databases of such information, and analytic and predictive modelling techniques are being used successfully to define these distributions.

The underlying business drivers can be developed from an assessment of:

- **Business environment and internal control factors (BEICFs)**, which are characteristics of the internal and external operating environment that bear an exposure to operational risk. Business environment factors are the inherent risk, and internal control factors act to mitigate the risk, resulting in residual risk. Residual risks may be measured using qualitative expert judgment gathered via balanced scorecards, or quantitatively using data on key performance indicators (KPIs) and key risk indicators (KRIs).

- **Scenario analysis**, which includes specific states across a range of internal and external factors.

Ultimately the above two sources should be self-consistent. This is somewhat difficult if they are considered independently. However, the use of a Bayesian framework enables them to be integrated. That's because scenario analysis typically represents the state of the system BEICFs when they are at extremes. Ensuring that the conditional relationship structures of the BEICFs are consistent between the two sets of information is important.

When considering sources of information, it is critical that the level of quality or uncertainty associated with it is also assessed. Sometimes information will be of a very high degree of certainty, perhaps because it is derived from a team of people with a lot of experience handling a stable operational process, but in other cases it will be highly uncertain. In all cases, the level of quality or uncertainty in the underlying information should be reflected in the distribution parameter estimates, rather than assuming that they are perfect. Again, this is very straightforward within a Bayesian framework.

Where data exists on the state of the business drivers of the operational system, Bayesian inference and learning techniques can be used to refine the assessments of the distributional parameters and conditional relationships of each driver.

7.7 Aggregation

Aggregation is the process by which multiple risk elements are brought together to assess total risk. This is necessary because traditional methods of assessing total or operational risk are based upon a reductionist approach, which attempts to simplify complexity into discrete risk elements. As detailed in Section 6, traditional approaches to aggregating components of operational risk, whether via LDA models or scenario approaches, require the use of statistical methods such as correlations and/or copulas. Aggregation in these instances typically occurs at a relatively high level, for example at the scenario level or to total severity/frequency distribution level.

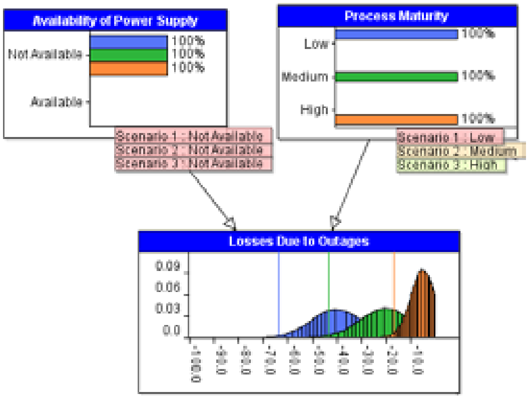
In contrast to this, structural approaches do not use such statistical methods. This is because the interrelationships between different elements of an operational risk framework are accounted for directly through the causal relationship structure. This is one of the strong advantages of structural models, whereby the nature of the dependency across multiple types of operational risk scenarios is modelled and can be explained directly through the interrelationships between the causal factors.

If two operational processes, business units, or scenarios are truly independent, then this can be shown to be the case as the operational drivers in the system structure (both the cognitive map and Bayesian network) of each will be unique. In this case, their total outcome distributions can simply be added together. However, in the event there are operational drivers that are common to both operational system structures, then clearly there will be some element of dependence. In this case, we can integrate the two systems by structurally combining them such that they are both connected to the common driver. Once this is done, total outcomes and risks can again be simply added together at the top level.

The dependence between the two operational systems is now explicitly related to the set of common drivers. In this framework, we have thus achieved not only aggregating complex interdependent risks, but also have created a mechanism for explaining exactly where and how the interdependency arises.

Building upon the previous example, suppose we now introduce the operational risk that an outage occurs, which is due to an interruption in the power supply to an important operational centre and has a likelihood of occurring of 10%. Operational losses will be incurred as a result, but the extent of them will be dependent upon the degree of process maturity. Where process maturity is high, such as where an extensively tested business continuity plan (BCP) is in place, workarounds such as transitioning to offsite facilities mean that production losses are minimised. However, if process maturity is low, such as where a BCP is not even in place, these losses will be much more significant. Figure 29 illustrates these dynamics for three power failure scenarios relating to different process maturity levels.

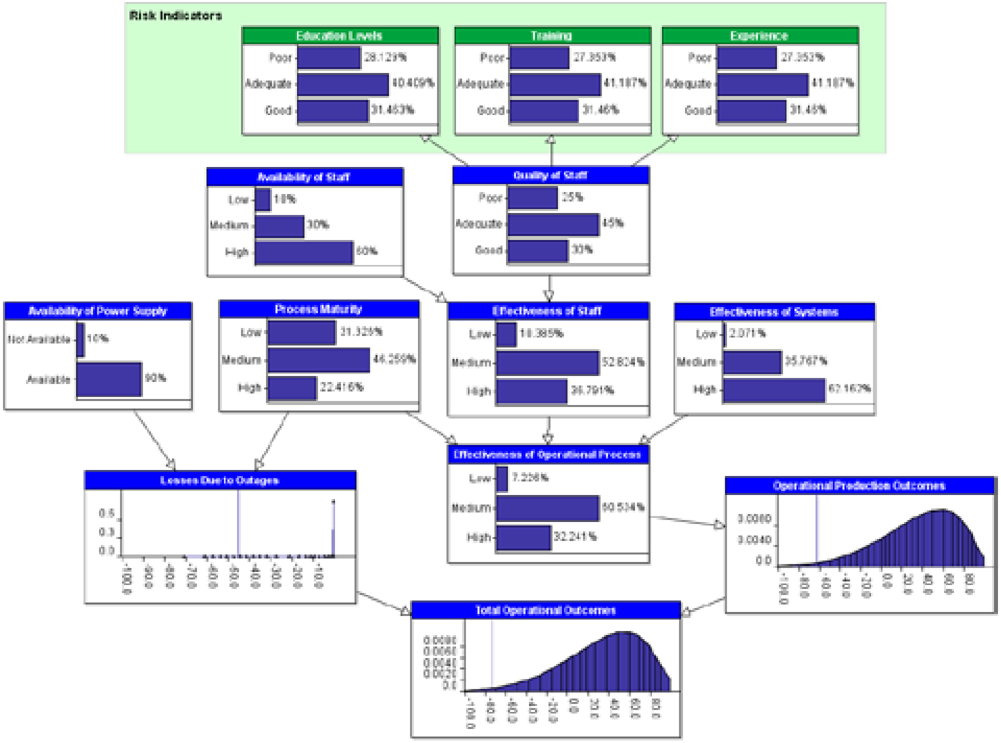
FIGURE 29: LOSSES DUE TO POWER OUTAGES FOR THREE PROCESS MATURITY SCENARIOS



The 99th percentile outcome for *losses due to outages*, conditioned upon the fact that the power supply is not available, is -63.5, -43.4, and -17.0 for the low, medium, and high scenarios for process maturity. When the state of the power supply is not known (i.e., a likelihood of 10% being unavailable), then the 99th percentile outcomes reduce to -52.8, -32.9, and -11.9 respectively. When the state of process maturity is also not known with certainty, then the 99th percentile outcome is -44.9.

Given that process maturity is also an important driver in the main example, we can now link the two structures together, to obtain an assessment of the distribution of the aggregate set of outcomes. This is shown in Figure 30.

FIGURE 30: AGGREGATE OPERATIONAL OUTCOMES



Total operational outcomes are now a complex function of process maturity, which impacts outcomes through two pathways. The 99th percentile outcome of -63.1 relates to the first pathway through operational production outcomes, whilst the 99th percentile outcome of -44.9 relates to the second pathway through losses that are due to outages. These combine to give a 99th percentile outcome of -73.6 for total operational outcomes.

Using traditional techniques such as LDA and scenario analysis, the two elements to this structural framework, operational production outcomes and losses that are due to outages, would have been modelled separately. Aggregation would then have occurred through a general correlation assumption. However, in a Bayesian framework, we can explicitly identify the what, where, why, and how of the interdependencies between the two operational elements. They are completely independent of one another across all underlying drivers, except for whenever process maturity is impacted. In a traditional statistical framework, we would say that they have a correlation of zero, except for when process maturity is impacted. Figure 31 illustrates the interdependent impact when process maturity is in its different states.

FIGURE 31: IMPACT OF THREE DIFFERENT STATES OF PROCESS MATURITY

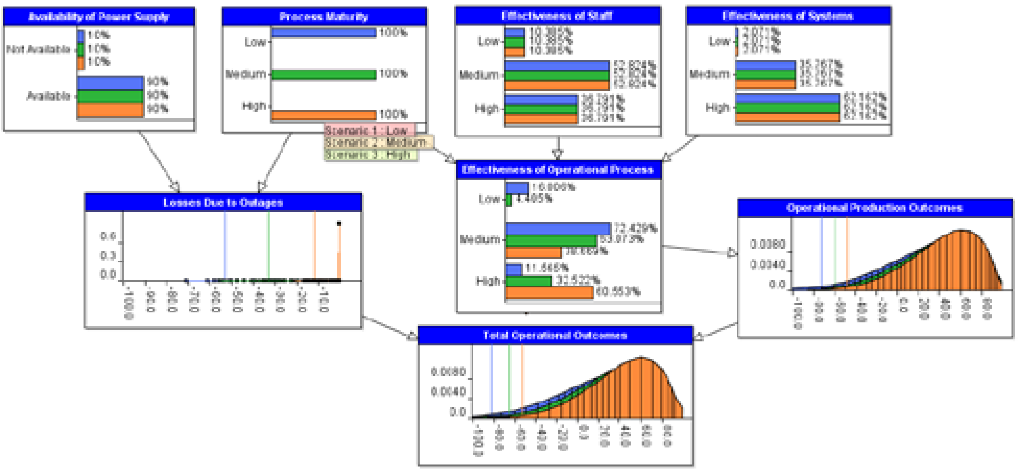


Figure 32 shows the 99th percentile results for each of the scenarios as well as for the base scenario where the state of process maturity is uncertain.

FIGURE 32: IMPACT ON 99TH PERCENTILE OUTCOMES OF THREE DIFFERENT STATES OF PROCESS MATURITY

PROCESS MATURITY	LOW	MEDIUM	HIGH	BASE
OPERATIONAL PRODUCTION OUTCOMES	-72.4	-59.1	-47.7	-63.1
LOSSES DUE TO OUTAGES	-52.8	-32.9	-11.9	-44.9
TOTAL OPERATIONAL OUTCOMES	-82.9	-65.3	-51.9	-73.6
IMPLIED CORRELATION	-0.15	-0.08	0.24	-0.10

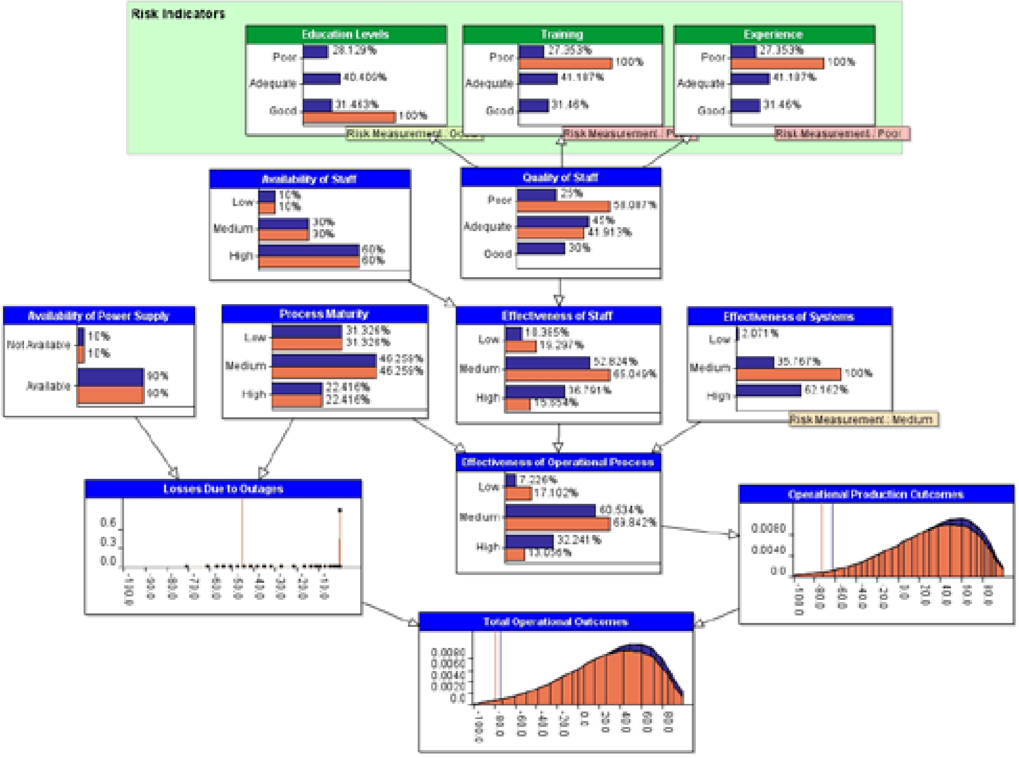
From Figure 32, it can be seen that both risk elements are impacted by the various states of process maturity. It is also possible to infer the implied statistical correlation between these elements for each of the scenarios, as shown in the final row. What can be seen here is that the nature of the interdependency changes from a negative to a positive one as process maturity moves from a low to a high state. Nonlinearity is clearly evident here, and it can be quantified explicitly. This shows that, as process maturity improves, overall risk reduces, but the size of the diversification benefit also reduces.

7.8 Risk assessment

Once the system structure of an operational process has been fleshed out, it can readily be used to actively and dynamically monitor risk levels. As management information is collected on the risk indicators and other business drivers of the model, the model will then update the state of the dependent outcomes. For example, Figure 33 shows the state of the system under the constrained set of outcomes given the following measurements:

- Education levels = Good
- Training = Poor
- Experience = Poor
- Effectiveness of systems = Medium
- Availability of staff, availability of power supply, and process maturity remain unmeasured

FIGURE 33: IMPACT OF MEASUREMENTS ON OUTCOMES VERSUS BASE SCENARIO



The impact of this information can be clearly seen compared to the base scenario: the quality of staff, the effectiveness of staff, and the effectiveness of operational process have all deteriorated. Risk can then be measured at the total operational outcome level, using the typical statistical concepts such as variance and tail risk measures such as Value at Risk (VaR) and CTE (otherwise known as tail VaR). In this example, the 99th percentile total operational outcome, or VaR, has increased in magnitude from -73.6 to -79.2.

7.9 Capital assessment

Bayesian network frameworks are ideally suited to assessing capital requirements. This is because they can simultaneously account for the full range of operational outcomes, both positive and negative, and common and extreme. The above examples could be readily interpreted as a calculation of operational risk capital, whereby the quantity of capital held is calculated using a 99th percentile VaR on total operational outcomes over a defined time horizon such as one year (i.e., 73.6).

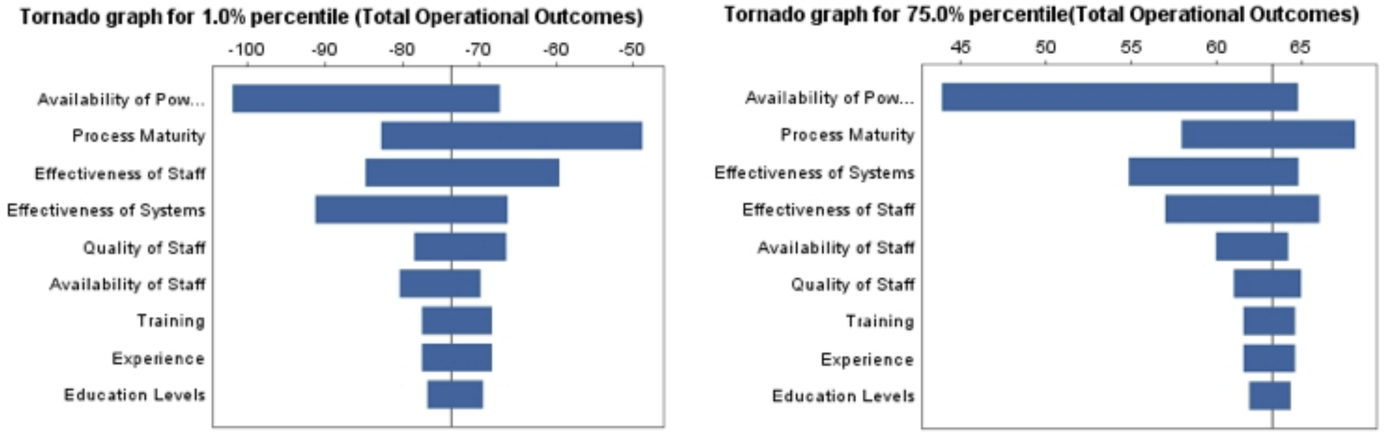
The benefit of a Bayesian approach for capital assessment is that we are now able to directly link capital requirements to the observed states of important business drivers, and thus can dynamically manage capital in response to the evolving business environment.

From a banking perspective, Bayesian models are able to integrate the four data source elements directly into the framework: internal and external historical loss data, scenario analysis, and business environment and internal control factors.

7.10 Sensitivity analysis

An important question that both business and risk managers want to know the answer to is: What business drivers have the most material impact on my profitability and capital? If we measure capital as a 1st percentile outcome (i.e., a 99th percentile VaR), and profitability as a 75th percentile outcome, then it is possible to measure the sensitivity of these variables to the full range of possible states of the underlying business drivers. These are shown in Figure 34.

FIGURE 34: SENSITIVITY OF CAPITAL AND PROFIT MEASURES TO UNDERLYING BUSINESS DRIVERS



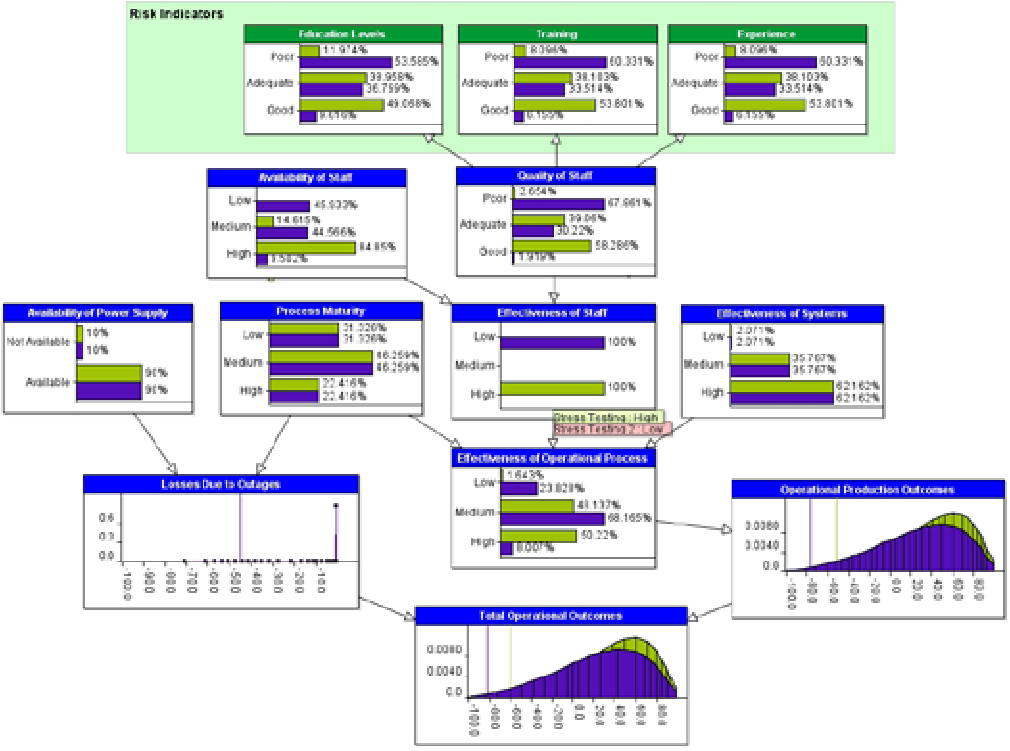
What the above analysis tells us is that capital is the most sensitive to the availability of power supply. If this deteriorates, then capital measures will also significantly deteriorate. The effectiveness of systems and staff are the next most important drivers of potential deteriorations in capital. In contrast to this, capital is highly leveraged to process maturity, which presents the biggest opportunity to improve capital requirements. In terms of profitability, the availability of power supply similarly has the highest downside sensitivity, whilst process maturity has the highest upside sensitivity.

Critically, these results reflect the full level of interdependencies between all the dependent business drivers. Sensitivity analyses undertaken using traditional LDA or scenario methods typically do not do this as they have no mechanism to resolve these important dependencies. This is very important for operational models that have a lot of embedded complexity and interdependencies, unlike the current example, which is simplified for the purposes of this paper.

7.11 Stress testing

Stress testing is a useful mechanism for understanding the impact that a change to a business driver has upon operational outcomes. A Bayesian approach is not only able to do this with ease, but by using Bayesian inference, it is able to infer the states of all business drivers that are causally related to the business driver(s) under consideration. Figure 35 shows the impact when the effectiveness of staff is set to both low and high states.

FIGURE 35: STRESS TESTS FOR EFFECTIVENESS OF STAFF SET TO LOW (PURPLE) AND HIGH (GREEN) STATES



The impact upon upstream business drivers and risk indicators is able to be inferred by the Bayesian network. This shows the range of state outcomes that are consistent with each of the scenarios. This shows that the most likely source of a low measure of staff effectiveness is a poor level of staff quality, whilst the most likely source of a high measure of staff effectiveness is a high measure of staff availability. This provides important information to business managers on where to spend marginal resources (i.e., staff quantity or quality) in order to enhance business outcomes.

The impact of this stress test on profitability and capital measures is summarised in Figure 36.

FIGURE 36: STRESS TESTS FOR EFFECTIVENESS OF STAFF, IMPACT ON CAPITAL AND PROFITABILITY

EFFECTIVENESS OF STAFF	LOW	BASE	HIGH
CAPITAL (1ST PERCENTILE LEVEL)	-82.6	-73.6	-61.4
PROFITABILITY (75TH PERCENTILE LEVEL)	57.0	63.2	66.0

This example shows that staff effectiveness has a proportionately greater impact on capital than it does on profitability.

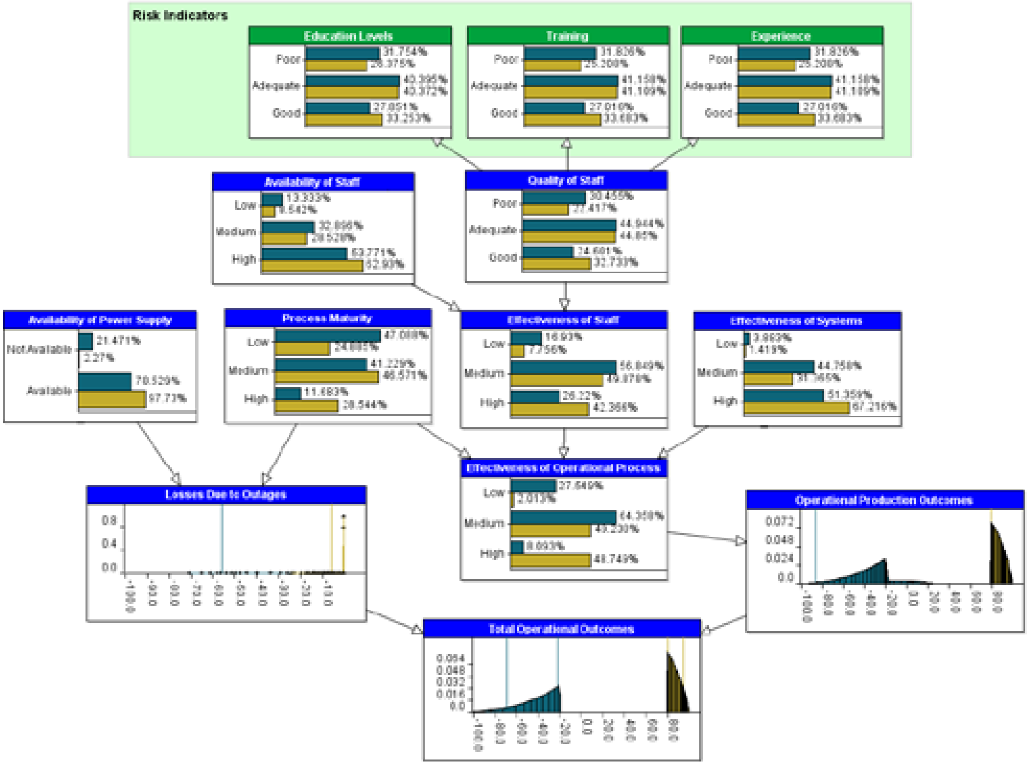
The example shown above on risk assessment could also be interpreted as an example of a combination stress whereby multiple business drivers are impacted. Combination stress tests can be extremely informative as they most closely capture the dynamics of real-world stresses, which are almost always multivariate rather than univariate in nature.

7.12 Reverse stress testing

Reverse stress testing is a powerful mechanism for addressing the important question of: *What does my business look like if it is doing really well compared to if it is failing?*

The only way to answer this question is to use inference techniques. As Bayesian networks are built using Bayesian inference techniques, they are able to handle such questions with ease. Figure 37 shows the states of the business drivers when total operational outcomes are preconditioned to be within the bottom 10% of outcomes, and then separately the top 10% of outcomes.

FIGURE 37: REVERSE STRESS TESTS FOR BOTTOM AND TOP 10% OF OUTCOMES



From this analysis we are able to quantify the states of each underlying business driver that are consistent with either a good or bad outcome. From this analysis it is clear that the biggest differences between these two opposing business states are:

- Availability of power supply, with almost a tenfold difference in the risk of power being unavailable between the two extremes
- Process maturity, with a significant increase in both high and low states
- Moderate differences in staff and systems effectiveness

In this example, it is interesting to note that the staff level risk indicators of education, training, and experience are actually very consistent between these two extremes. Findings like these might at first seem counterintuitive, but as we can explain why this is the case, it might be an important insight that we shouldn't overly focus on these indicators as critical drivers of capital and profitability.

7.13 Risk appetite and risk limit setting

The risk appetite statement is the primary mechanism the board uses to set the degree and type of risk that an organisation is willing to accept in pursuit of its strategic goals. The risk appetite statement contains a series of specifications of the range of outcomes or risks across a number of dimensions. Examples of risk appetite statements include:

- Not bearing a loss of greater than x more than every n years (e.g., 100m every 10 years)
- Not suffering a workplace fatality
- Having a low tolerance to reputation risk
- Ensuring full compliance with all relevant laws and professional standards

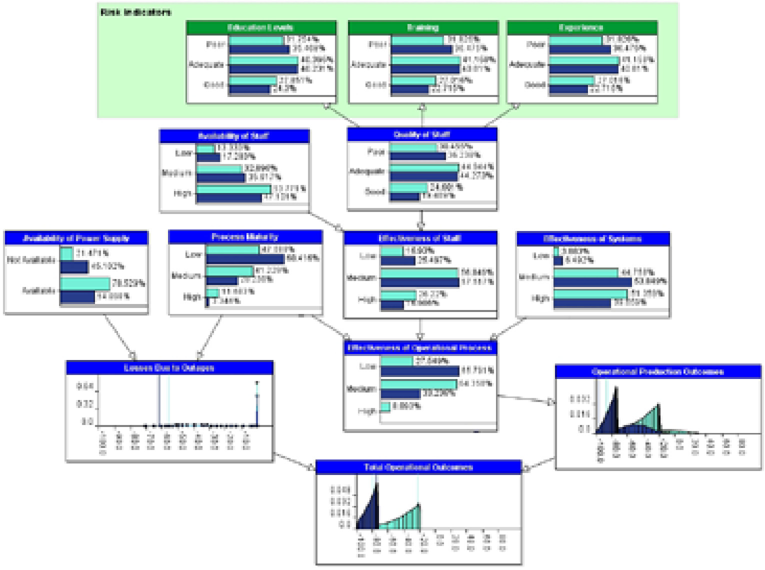
Risk appetite statements such as these are defined at a high level with respect to tangible outcomes that the board, shareholders, and executive management care about. One of the key challenges that managers must address is to take these statements, which act as constraints on the organisation, and to understand what they mean at a more granular level. This requires translating them into consistent states across a multitude of business drivers, risk categories, and risk factors. This can be extremely difficult, as there is a significant amount of complexity and interrelationships which need to be resolved in this process, which humans are not naturally well equipped to deal with.

However, a Bayesian framework is ideally suited to resolving this problem. By way of example, consider two risk appetite statements that relate to the above operational system:

1. Total operational outcomes must not be less than the 1st percentile level of possible outcomes
2. Total operational outcomes must not be less than the 10th percentile level of possible outcomes

What do these two statements imply about the states of the underlying operational drivers, and how do we set risk limits consistent with these outcomes? Figure 38 illustrates two examples.

FIGURE 38: RISK APPETITE AND RISK LIMIT SETTING FOR THE 1ST AND 10TH PERCENTILE OUTCOMES



What can be seen here is the range of states for each driver that are consistent with the two conditioned outcomes. For example, power supply must not be available at least 21.5% and 45.1% of the time, respectively, in order for total operational outcomes to fall within the bottom 10th and 1st percentile. This is significantly worse than the base estimate of 10% that power supply is not available. In effect, these two probability outcomes reflect the two risk limits that are consistent with both the risk appetite statement constraints, as well as all of the other possible states that other drivers could be in. Likewise the probabilities for the states of the other business drivers reflect their appropriate risk limits that are consistent with the risk appetite statement.

This is quite a powerful framework because it now gives management a guide as to how to dynamically update risk limits as actual risk levels change in business whilst still being consistent with the overall risk appetite statement.

7.14 Advantages and limitations of a structural framework

The advantages of a structural framework are that it:

- Provides a meaningful explanation of how operational loss outcomes are directly related to the states of business drivers
- Provides engagement with the business as it is framed directly in terms of the fuzzy discrete states and relationships that most business language is couched in
- Provides a single consistent framework to undertake risk assessment, capital assessment, risk appetite, the setting of operational risk limits, stress testing, reverse stress testing, and complex scenario testing
- Identifies and captures where and how risk mitigation actions can reduce risk likelihood and severity
- Incorporates a wide variety of types of information used across the business, including quantitative loss data and the intellectual property/expert judgment of individuals and teams, directly accounting for the quality of this information
- Does not rely on knowing/hypothesising the prior distributions of loss severity or frequency outcomes, which could be highly nonlinear
- Captures full range of complex and nonlinear interdependencies
- Can evolve dynamically over time to reflect the changing operational and risk management structure of a company
- Can degenerate into deterministic and stochastic frameworks, which makes them ideal for extensions of existing frameworks to deal with the marginal questions that other frameworks find difficult to answer

The limitations of a structural framework include:

- It can be time-consuming to undertake, and requires expertise in the overall operational system as well as Bayesian methods and experience in their applications.
- Eliciting the marginal dependent relationships can be challenging. It can require significant data analysis, and also relies upon subjective opinion, which may be of poor quality (notwithstanding the potential value of this information).
- A higher level of modelling expertise and capability is generally required for Bayesian approaches relative to deterministic and stochastic approaches.

7.15 Choice of technique

It is important to note that not all circumstances would warrant a Bayesian approach, given the additional effort that is typically required. For organisations with very linear and simple operational models, the use of traditional simple and statistical approaches may be sufficient to meet requirements. In this case a Bayesian framework could still be used, but it would essentially degenerate into the simpler deterministic and stochastic methods and may provide relatively little marginal value.

Structural frameworks are particularly valuable wherever significant complexity exists in an organisation, and in particular for the high-severity, low-likelihood risks where a deep dive into the causal drivers can be justified. They can also be combined with and leverage scenario analysis and LDA techniques to reflect the structure and information that these approaches provide.

Finally, it is worth noting that a combination of methodologies might be the best framework for an operationally complex organisation. Simple standard formulas might be sufficiently accurate and robust for very well established and stable operational processes that give rise to low-severity, high-frequency losses, whereby the underlying processes are highly correlated to a few well-chosen key risk indicators. Concept mapping techniques can be used to elicit the system structure when adopting a scenario analysis approach, and then integrated into a holistic LDA or Bayesian framework, which can leverage BEICFs and historical loss data to parameterise the interconnected severity and likelihood distributions that drive a large range of operational loss sources. Lastly, for very material and low likelihood events such as rogue traders, a deep dive might be justified using cognitive mapping and Bayesian techniques based upon both internal business drivers and leveraging the findings from an assessment of the few external events that do exist.

8 REGULATORY REQUIREMENTS

8.1 Overview of Basel II/III framework for operational risk

Operational risk is mentioned as one of the main reasons to move from the 1988 Basel Accord (sometimes referred to as Basel I), which was concentrated mainly on credit risk, to the Basel II directive. More direct reasons include events such as the fall of the Barings Bank (caused by unauthorised transactions on derivatives) and the terrorist attacks of 11 September 2001.

Basel II is divided into the three so-called *pillars*:

- Minimum capital requirements, defining the amount of surplus capital that needs to be held by an entity to cover three major quantifiable risks: credit, operational, and market

- Supervisory review process, stating the principles of national supervision and describing the tools that should be used to assure proper capitalisation of banks

- Market discipline, defining the information about the risk profile that an entity should publish so that market participants can properly assess the financial situation of a company.

Each of these pillars outlines requirements and/or advice concerning operational risk.

Pillar I describes methods of determining the amount of capital required to support the operational risk of a company, defined as the risk resulting from *inadequate or failed internal processes, people and systems or from external events*. There are three methods of calculation proposed in the document: the basic indicator approach, the standardised approach, and the advanced measurement approach (AMA). Each is more complex than the previous one, and requires that management be more aware of the characteristics of risk specific to the company. Industry expectations are that using more complex methods should yield a lower capital requirement, and this is supported by the Basel II document itself, stating that banks should be encouraged to develop more sophisticated operational risk measures.

The basic indicator and standardised approaches both use the assumption that the amount of operational risk is proportional to a positive part of gross income (specifically, its average annual value over the last three years). The former uses a simple formula:

$$Capital\ charge = 15\% \times \sum_{i=1}^{3} max(0, GI_i)/n$$

where:
GI_i is the annual gross income over the ith previous year
n is the number of the previous three years for which gross income is positive

The standardised approach uses only a slightly more complex definition, dividing the total gross income between the eight business lines:

$$Capital\ charge = \sum_{j=1}^{8} \beta_j \times \sum_{i=1}^{3} max(0, GI_{i,j})/3$$

where:
β_j is a fixed percentage set for the jth business line
$GI_{i,j}$ is the annual gross income over the ith previous year for the jth business line

The β_j coefficients are set to:

- 18% for the corporate finance, trading, sales, payment, and settlements
- 15% for the commercial banking and agency services
- 12% for the retail banking, asset management, and retail brokerage business lines

Additionally, at a national supervisor's discretion, the bank may be allowed to use the alternative standardised approach, which bases its calculations for the retail and commercial banking on 3.5% of total outstanding loans and advances, instead of the gross income.

AMA allows the exact formula for required capital to be specified by the financial institution itself, and as such is subject to supervisory approval. It should aim at holding enough capital to cover the operational losses in the one-year horizon with 99.9% probability. There is an expectation that, once adopted, risk allocation techniques will be constantly developed. Using AMA only for some part of a company's operation and simpler approaches for other parts is also possible, under certain conditions.

While any bank can use the basic indicator approach, there are requirements that a company must fulfil in order to qualify to use the standardised approach or AMA. For both, the basic expectation is that the financial institution has generally well-implemented operational risk management systems. To use AMA, however, the bank must fulfil much more restrictive criteria, specifying detailed qualitative standards (aimed at integrating the risk measurement and the risk management system) and quantitative standards (aimed at capturing the most severe events from the tail of the distribution).

The criteria include:

- Utilisation of internal loss data (for at least a five-year period)
- Use of relevant external loss data for the measurement of the probability of rare events
- Analysis of scenarios for an evaluation of exposure to severe events
- Capture of key risk factors that can change the risk profile (related to business environment or internal control processes)

Additionally, before the bank can use AMA for regulatory capital calculation purposes, the model is subject to initial monitoring of its appropriateness by the supervisor.

Pillar II in the Basel II document itself is quite sparing with advice about operational risk management supervision–it contains only a single, rough guide about considering whether gross income is a proper base for the operational risk capital calculation in the case of every bank. However, more detailed guidance was developed and published by a committee in a follow-up document, regarding the principles of sound operational risk management, which should be followed by a bank using at least the standardised approach, and supervisory guidelines for AMA.

Sound risk management principles should underlie every risk management system employed in a bank. That framework is based on the three *lines of defence*:

1. Business line management (responsible for the identification and assessment of specific risks)

2. Independent corporate risk management function (responsible for framework development and maintenance)

3. Independent review (responsible for verification of the effectiveness of the framework and validation of its methodology)

Processes based on this structure should exhibit certain characteristics: integration with organisational culture, involvement of the highest management and supervision level of the company, defining risk appetite, identifying and assessing risk among existing and new products, constant monitoring, reporting, control and mitigation of risk, loss reduction, and public disclosures allowing assessment of the bank's approach to risk.

Guidelines for supervising AMA are very general and based on principles, so that they allow for flexibility and don't hinder the evolution of new methods of operational risk assessment. However, some practices that emerged were discussed by the regulators, who subsequently released additional detailed advice. Examples include the calculation of loss value (i.e., what should be included in the recorded amount) and the date at which a loss should be recorded–date of discovery, of accounting, etc.

Pillar III specifies the information regarding the operational risk management system that should be publicly disclosed by a company. It is limited to the operational risk valuation approach type used by a bank, and, if it's an AMA, the description of the approach, including key internal and external risk factors used.

Basel II, which was induced by, among other events, significant operational losses, contains substantial amounts of guidelines regarding operational risk. The Basel III accord that followed it, on the other hand, concentrated on the risks that caused the global financial crisis in the late 2000s and as such did not contribute much to the operational risk framework founded by the previous accord.

From the insurance market point of view, it might be interesting to note that Basel II forbids the use of capital invested in insurance entities owned by a bank to satisfy its capital needs, with a possible exception of surplus capital of majority-owned companies (at the national regulator's discretion). This means that if there is a regulatory requirement for operational risk capital in the insurance subsidiary, it is not possible to allow the same capital cover both bank and the insurer operational risk.

It should be also noted that companies using AMA are allowed, under certain conditions, to consider an insurance company as an operational risk mitigation tool and include it to lower capital requirements. Moreover, in the follow-up documents by the Basel committee, banks are encouraged to implement the use of this tool to their risk management on a continuous basis. Such an approach has also benefited the insurance industry, as it induced the creation of new kinds of insurance products specifically designed to satisfy the requirements of Basel II.

8.2 Overview of insurance operational risk regulatory requirements

8.2.1 Europe

8.2.1.1 Solvency II

Solvency II, the directive codifying EU insurance regulation, was issued in 2009 and aimed at implementing various rules regarding insurance companies by 2012 (which has subsequently slipped a number of years beyond this). Because it arrived later than the Basel II regulations, Solvency II obviously was strongly influenced by the bank supervision arrangements; Solvency II is even sometimes called *Basel for insurers*.

One such influence was the recognition of the need to include operational risk requirements as a part of the regulations. As this type of risk was not well recognised among the insurers, the advice included in the directive is not as detailed as in the Basel accord. Operational risk is defined exactly the same way as in Basel II, and the only other guidelines concerning the topic are the (very general) requirements on the indicators used to calculate the capital requirements for operational risk: annual expenses for unit-linked life business, premium income and technical provisions for other lines (with a cap specified based on the other risks' capital).

The detailed implementation measures for the directive, including the formulae for capital requirements' calculations, are still to be published. They are likely to be based on five (so far) quantitative impact studies (QISs) that were carried out by EIOPA, the EU-wide insurance regulator.

Each of those studies required insurers to calculate their capital requirements using a specified formula and to answer some qualitative questions (contrary to the name of the study).

Following the QISs, we can see how the operational risk requirements have evolved:

- In QIS 1 there were no questions or ways to calculate operational risk's capital requirement.

- In QIS 2 operational risk was one of the main four risk drivers (along with market, credit, and underwriting risk). It included questions to insurers regarding risk management structure and processes, including the use of scenarios and/or statistical data to quantify operational risk. The formula for capital requirement was based on a maximum of some percentage of premiums and reserves (different for life, nonlife, and health lines of business). Unit-linked (UL) life business was treated preferentially, using only 10% of the respective values.

- QIS 3 was heavily impacted by a document published by a Basel committee regarding sound risk management; most of the questions asked related to the use of risk management techniques covered in the above guidelines. Capital charges have been bounded by 30% of the capital requirement for other risks, and unit-linked business relief was excluded from calculations. Additionally, risk resulting from risk mitigation failures was recognised as a part of operational risk.

- QIS 4 changed the approach to operational risk for groups, requiring them to report not only the sum of the amounts of operational risk capital in their subsidiaries, but also the consolidated operational risk for the whole group; they were also encouraged to present the evidence that the operational risks of their different subsidiaries have a correlation of less than 1. The formula for capital charge was changed regarding the unit-linked business, basing their operational risk on the amount of annual expenses, instead of premiums and provisions.

- Finally, QIS 5 has changed only the formula for the operational risk capital charge, to a much more sophisticated equation:

$$SCR_{Op} = min(max(Op_{prem}, Op_{prov}), 0.3 \times BSCR) + 0.25 \times Exp_{ul}$$

where:
$BSCR$ is the aggregated capital requirements for the other risks of non-UL business
Exp_{ul} are the annual expenses in UL business in the previous 12 months

Op_{prem} and Op_{prov} are the charges calculated on a base of last year premiums and current technical provisions, according to the formulae:

$$Op_{prov} = 0.045 \times max(0, TP_{life-nonUL}) + 0.03 \times max(0, TP_{non\text{-}life}),$$

$$\begin{aligned} Op_{prem} = {} & 0.004 \times (Earn_{life-nonUL}) + 0.03 \times Earn_{non\text{-}life} \\ & + max(0, 0.04 \times (Earn_{life-nonUL} - 1.1 \times pEarn_{life-nonUL})) \\ & + max(0, 0.03 \times (Earn_{non-life} - 1.1 \times pEarn_{non-life}) \end{aligned}$$

where:
TP_{xxx} are the technical provisions for business unit xxx
$Earn_{xxx}$ are the earned premiums during the last 12 months for business unit xxx
$pEarn_{xxx}$ are the earned premiums during the previous12 months for business unit xxx

These formulae are for life (excluding unit-linked) and nonlife business respectively. Health business was integrated into either life or nonlife lines, depending on its nature.

Generally, EIPOA's opinion is that it's hard to find a formula that suits every entity. Instead, its idea is to use the formula to measure operational risk for most of the companies and allow others, which don't find that suitable, to apply for a partial internal model. Also, a standard formula should not be too sophisticated, in order to maintain its ease of use on a common basis. However, some *ladder factors* are currently considered to allow companies a discount in capital charge reflecting the degree of progress in the operational risk management framework.

8.2.1.2 Current status

The Solvency II final implementation date has been postponed, and there are some concerns whether the current date is realistic. By that time and whilst predicting future regulations from the QIS documents and other EIOPA publications, companies must set up some operational risk management and quantification systems according to national regulations. From a practical point of view in most of the EU countries, no consolidated and structured approach to operational risk management exists. This is quite often also true for those insurance companies which are part of bancassurance groups, which tend to adopt the standard banking approach. Only in a few cases have they started to undertake a loss data collection exercise to develop operational loss models tailored more to their businesses.

For example, most Central and Eastern Europe (CEE) countries have not yet developed any operational risk requirements or recommendations. In Romania, there are some regulations governing insurance company risk management in general, which require the company to assess some stress tests' impacts, but this does not result in a capital charge. In Poland, the national supervisor controls exposure to operational risk annually, based on expert judgment; results are disclosed only to the examined entity and no capital charges are applied. However, detailed guidelines for banks are in the process of consultations, and as they are supervised by the same institution, there is a possibility for future quick development in this area.

8.2.2 Australia

The Australian Prudential Regulatory Authority (APRA) is responsible for setting operational risk requirements for regulated financial services entities operating in Australia.

8.2.2.1 Life and General Insurance Capital (LAGIC) Prudential Standards

Prudential Standard LPS 118 outlines the calculation methodology of the operational risk charge that life insurance companies must hold to cover operational risks. The operational risk charge (ORC) is calculated as the sum of that charge for risk business (ORCR), for investment-linked business (ORCI), and for other business (ORCO):

$$ORC = ORCR + ORCI + ORCO$$

The ORCR is calculated as follows:

$$ORCR = A \times [max(GP_1, NL_1) + max(0, |GP_1 - GP_0| - 0.2 \times GP_0)]$$

where:
A is 2% for statutory funds of specialist reinsurer and 3% for other funds
GP_t is gross premium income for the 12 months ending on the reporting date at time t
NL_1 is the net adjusted policy liabilities at the reporting date

The ORCI and ORCO are calculated as follows:

$$ORCI\ or\ ORCO = B \times [NL_1 + max(0, GP_1 - 0.2 \times GL_0) + max(0, C_1 - 0.2 \times GL_0)]$$

where:
B is 0.15% for statutory funds of specialist reinsurer and 0.25% for other funds
NL_1 is the net adjusted policy liabilities at the reporting date
GP_1 is gross premium income for the 12 months ending on the reporting date at time 1
GL_0 is gross adjusted policy liabilities for the 12 months ending on the reporting date at time 0
C_1 is all gross payments to meet liabilities to policy owners for the 12 months ending on the reporting date at time 1

Prudential Standard GPS 118 outlines the calculation methodology of the operational risk charge that general insurance companies must hold to cover operational risks. The operational risk charge (ORC) for general insurance companies is calculated as the sum of that for inwards reinsurance business (ORCI) and for not inwards reinsurance business (ORCNI):

$$ORC = ORCI + ORCNI$$

The ORCI and ORCNI are calculated respectively as follows:

$$ORCI = 2\% \times [max(GP_1, NL_1) + max(0, |GP_1 - GP_0| - 0.2 \times GP_0)]$$

$$ORCNI = 3\% \times [max(GP_1, NL_1) + max(0, |GP_1 - GP_0| - 0.2 \times GP_0)]$$

where:
GP_t is gross written premium revenue for the 12 months ending on the reporting date at time t
NL is the central estimate of insurance liabilities at the reporting date

8.2.2.2 Superannuation funds

Prudential Standard SPS 114 outlines the calculation methodology of the operational risk financial requirement (ORFR) that superannuation entities must hold to cover operational risks. Unlike that for insurance companies, the ORFR is not calculated using a standard formula approach. Instead, the ORFR must be determined by the entity and it must reflect the size, business mix, and complexity of the operations, at a minimum including:

- An amount that the entity determines is necessary to address operational risks that it has identified in its risk management framework, having taken into account appropriate risk mitigations and controls. This amount must reflect any uncertainty in the scale of losses.

- An additional amount that the registrable superannuation entity (RSE) licensee determines is required for operational risk losses that have not been specifically identified in its risk management framework.

A tolerance limit below the ORFR must then be set, such that, if financial resources fall below it, the entity must take action to replenish resources to meet the ORFR.

8.2.3 United States and Canada

There is no explicit formula for operational risk in either the United States or Canadian insurance regulatory capital frameworks. However, both of those frameworks do have a broad component that could be considered to cover at least an element of operational risk.

In the US risk-based capital formula, this is the business risk component. Business risk for life insurers is based on premium income, annuity considerations, and separate account liabilities. Also included in business risk exposures are litigation, expenses relating to certain accident and health coverage, and expenses. Business risk for health insurers consists of the following subcomponents: administrative expense risk (variability of operating expenses), non-underwritten and limited risk (collectability of payments for administering third-party programs), guaranty fund assessment risk, and excessive growth. These subcomponents recognise that instability can result from poor controls on administrative expenses as well as from instability in medical expenses.

In the Canadian life insurance regulatory standards, operational risk is covered implicitly, through the regulator requiring all companies to be at a minimum of 150% of the minimum continuing capital and surplus requirements (MCCSR). This buffer is due in part to cover operational risk. However, at the time of this writing, the Canadian regulator was in the process of a series of quantitative impact studies similar to Solvency II, and had announced the intention of introducing a specific element for operational risk capital, although the proposed method had not yet been released.

8.2.4 Japan

Under Japan's statutory solvency requirement for life insurers, operational risk is effectively captured by the component of capital called the management-related risk capital (MRC). This is calculated as:

$$MRC = (R1 + R2 + R3 + R7 + R8) \times (Risk\ Factor)$$

where: $Risk\ Factor$ = 3% in case that profit for the year is a negative value
2% in case that profit for the year is a positive value or zero

and where:
$R1$ is the risk capital for insurance risk
$R2$ is the risk capital for interest-crediting risk capital
$R3$ is the risk capital for asset risk
$R7$ is the risk capital for products with minimum guarantee benefits
$R8$ is the risk capital for insurance risk relating to third-sector products

Full details of how $R1$ through $R8$ elements are calculated are beyond the scope of this document, but they can be found on the Japanese Financial Services Authority (FSA) website.

8.2.5 China, Hong Kong, and Taiwan

China and Hong Kong do not have specific capital charges for operational risk. As they operate under the old EU Solvency I rules of total capital held equal to 4% of reserves plus 0.3% of sum at risk, operational risk capital is implicitly captured within this.

Taiwan's risk-based capital structure specifies an operational risk charge ($C4$) calculated as:

$$C4 = x\%\ of\ premium\ income + 0.25\%\ of\ assets\ under\ management$$

where x = 0.5% for life business, 1% for annuity business, and 1.5% for all other business

8.2.6 Republic of South Africa

The Republic of South Africa (RSA) currently undergoes the process of creating regulations regarding insurance companies' solvency, called *solvency assessment and management* (SAM). The Financial Services Board (the national insurance supervisor) derives local rules from the European Solvency II directive, and performs similar quantitative impact studies. The final implementation date is set for the end of 2015.

Currently, QIS 2 is being carried out; it corresponds to the European QIS 5, but participation is voluntary. As for the operational risk, it uses exactly the same definition and formulae as in Solvency II; however, unit-linked portfolios' expenses are defined in a more detailed way than in the EU QIS 5.

8.2.7 Russia

No detailed requirements on operational risk are available in Russia. However, operational risk management frameworks do exist, which state that operational risk should be controlled by the company's headquarters. This control should include the collection of loss data, the testing of operational systems (at least every six months), and the use of operational risk transfer to third parties if required. Contrary to usual practice, legal risk is considered separately from operational risk (along with strategic and reputational risks).

9 EMERGING OPERATIONAL RISK ASSESSMENT AND BLACK SWANS

9.1 Operational risk dynamics and evolution

Organisations change. In market-based economies, they either adapt to the changing environment or they eventually die. Such changes are reflected in the operational production processes of the organisation, and as such, the associated operational risks which they face also evolve. Hence any effective operational risk management framework needs to be able to evolve along with it. Static frameworks that never change their inputs will eventually become irrelevant at best, and very misleading and potentially dangerous at worst.

Instead, the only aspects that should remain static in an operational risk framework are the central questions that it seeks to answer:

- How much operational risk is the company currently exposed to, and what operational processes and factors of production is it sourced from?
- What is the company's appetite for operational risk, and how does it set consistent risk limits for each granular operational production process across its business?
- How can the company most effectively mitigate excess operational risk that it does not wish to bear?
- How can the company tactically optimise its operational risk budget as actual operational risk changes over time?
- What are the company's emerging operational risks and how might it be able to manage them?

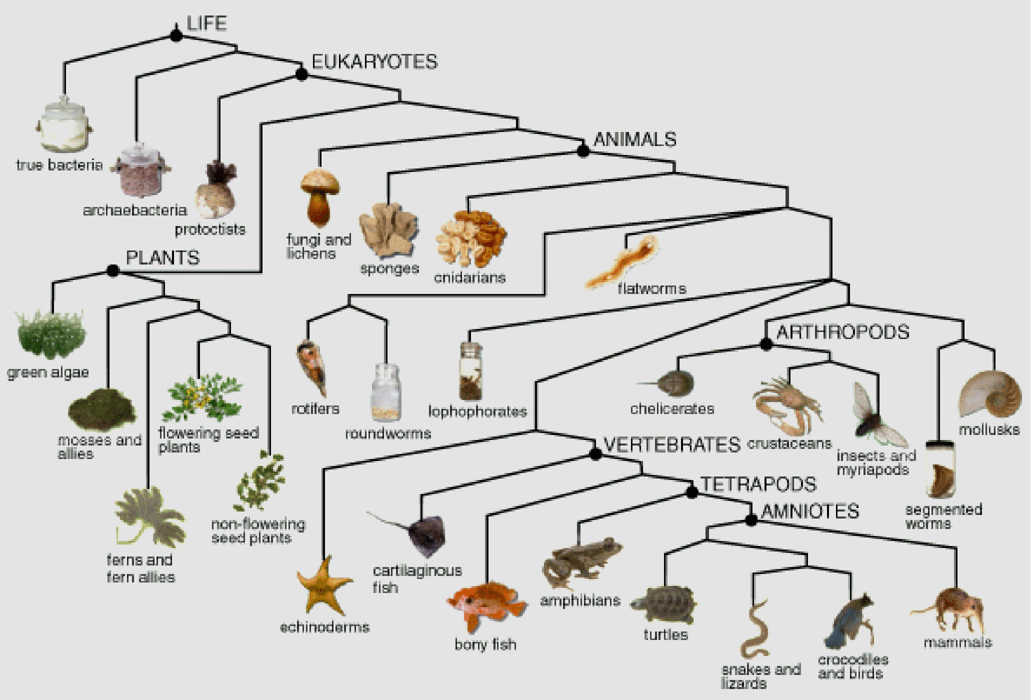
As operational activities and the factors of production evolve over time, the sources and dynamics of operational risk also change. A static framework of risk drivers and key risk indicators will eventually become insufficient to properly measure and manage operational risk. For example, a car manufacturer that moves from a labour-intensive to an automated robotic-operated production process has completely different operational risk drivers as the relative sources of production inputs have changed. Static measurements of operational risk based on such factors as the number of cars produced, types of cars produced, and the length of time it takes to complete each process are independent of production inputs. Instead, different measurements are needed based upon the characteristics of the production process itself, and its interaction with its environment.

Understanding the relationship between these characteristics can give us insight into the emergent properties of operational risk systems. This is the domain of phylogenetics.

9.2 Phylogenetics

Phylogenetics is the study of the evolutionary relationships between both living and nonliving things. In particular, it is based upon analysis of the characteristics that define each thing, and it seeks to draw on one-to-many relationships between them that represent the simplest way to structure the relationships. This can be objectively determined, which is one of its strengths. Figure 39 shows a high-level cladogram of the tree of life which is based upon this technique.

FIGURE 39: CLADOGRAM OF THE TREE OF LIFE

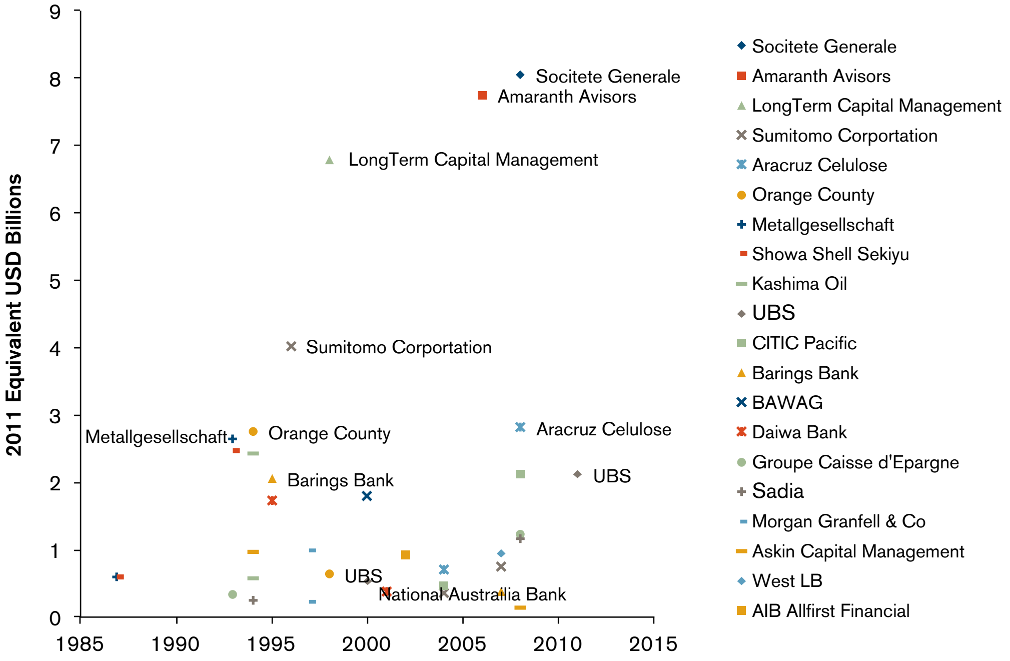


We can apply this to the study of operational risk by substituting risk events for the *things* in Figure 39. We can then explore the relationship of the characteristics that define these events, in order to understand how certain characteristics are evolving over time to generate new emergent risks.

9.3 Derivative trading loss example

In order to demonstrate this technique, we have applied it to operational losses associated with derivatives. We have leveraged the work produced by Coleman (2011), who mapped a range of relevant characteristics to a number of major derivative loss events. The loss events are shown in Figure 40 on page 66.

FIGURE 40: SELECTION OF LARGE DERIVATIVE TRADING LOSSES (2011 USD EQUIVALENT FIGURES)



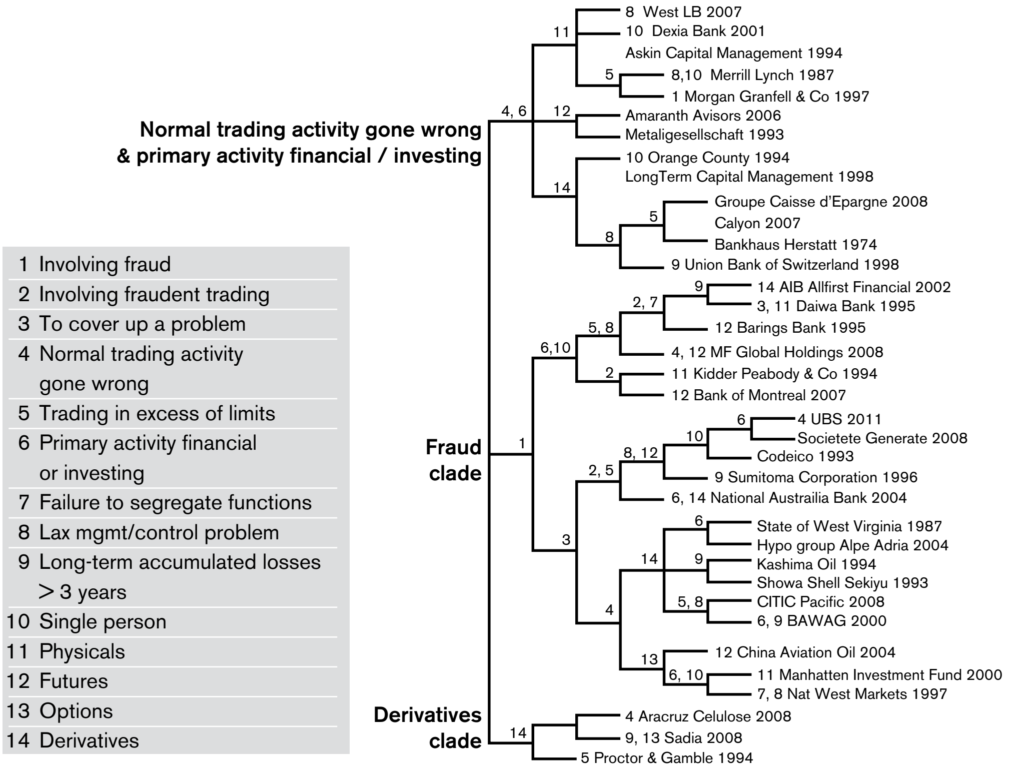
The characteristics these risk events have been mapped to are:

1. Involving fraud
2. Involving fraudulent trading
3. To cover up a problem
4. Normal trading activity gone wrong
5. Trading in excess of limits
6. Primary activity financial or investing
7. Failure to segregate functions
8. Lax management/risk control problem
9. Long-term accumulated losses in excess of three years
10. Single person
11. Physicals
12. Futures
13. Options
14. Derivatives

We have taken this mapping data at face value from Coleman (2011), with the exception of aggregating some of the finer levels of granularity on the security type. These characteristics are somewhat subjective, and clearly it would be possible to define additional characteristics, but they are sufficient for our purposes to demonstrate this technique.

Figure 41 shows the cladogram of this mapping.[20]

FIGURE 41: CLADOGRAM OF LARGE DERIVATIVE LOSS EVENTS AND CHARACTERISTICS


Normal trading activity gone wrong & primary activity financial / investing
Fraud clade
Derivatives clade
1 Involving fraud
2 Involving fraudent trading
3 To cover up a problem
4 Normal trading activity gone wrong
5 Trading in excess of limits
6 Primary activity financial or investing
7 Failure to segregate functions
8 Lax mgmt/control problem
9 Long-term accumulated losses > 3 years
10 Single person
11 Physicals
12 Futures
13 Options
14 Derivatives
4, 6
11
8 West LB 2007
10 Dexia Bank 2001
Askin Capital Management 1994
5
8,10 Merrill Lynch 1987
1 Morgan Granfell & Co 1997
12
Amaranth Avisors 2006
Metaligesellschaft 1993
14
10 Orange County 1994
LongTerm Capital Management 1998
8
5
Groupe Caisse d'Epargne 2008
Calyon 2007
Bankhaus Herstatt 1974
9 Union Bank of Switzerland 1998
1
6,10
5, 8
2, 7
9
14 AIB Allfirst Financial 2002
3, 11 Daiwa Bank 1995
12 Barings Bank 1995
4, 12 MF Global Holdings 2008
2
11 Kidder Peabody & Co 1994
12 Bank of Montreal 2007
3
2, 5
8, 12
10
6
4 UBS 2011
Societete Generate 2008
Codeico 1993
9 Sumitoma Corporation 1996
6, 14 National Austrailia Bank 2004
4
14
6
State of West Virginia 1987
Hypo group Alpe Adria 2004
9
Kashima Oil 1994
Showa Shell Sekiyu 1993
5, 8
CITIC Pacific 2008
6, 9 BAWAG 2000
13
12 China Aviation Oil 2004
6, 10
11 Manhatten Investment Fund 2000
7, 8 Nat West Markets 1997
14
4 Aracruz Celulose 2008
9, 13 Sadia 2008
5 Proctor & Gamble 1994

Each branch in the above cladogram ends in a specific event. Each branching point is defined by a split in the characteristics as identified by the numbers that are common to all members of the subbranches. The length of the branch represents the number of characteristics that *evolved* to define that branch, with more characteristics leading the longer branches.

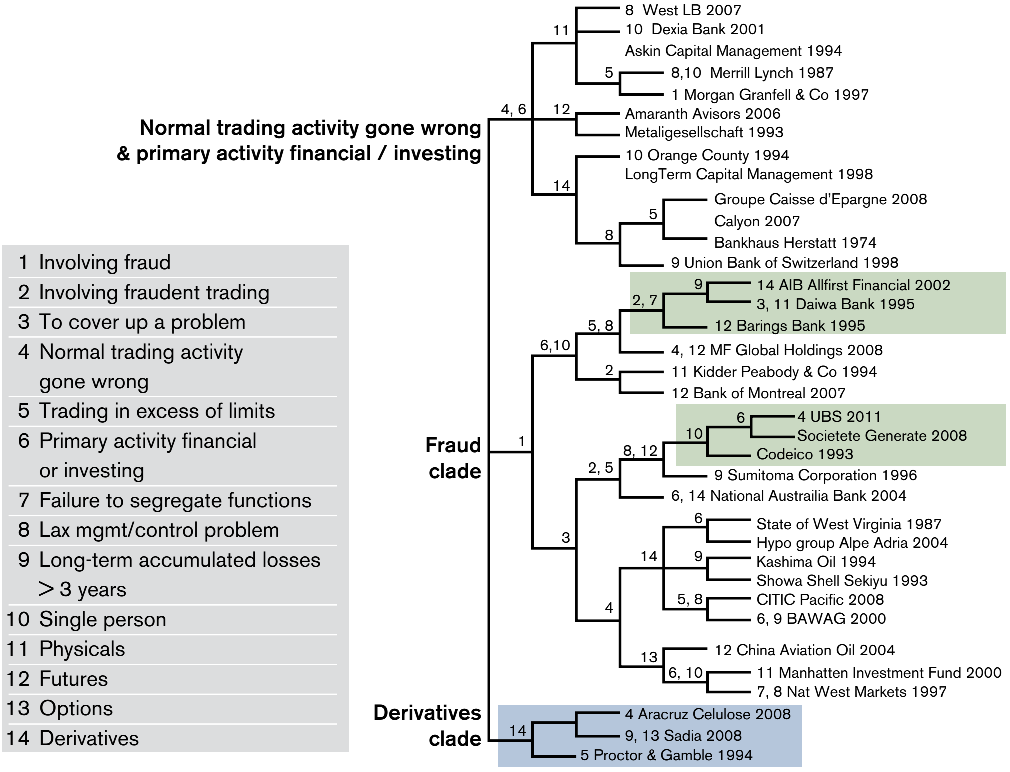
These diagrams are very useful in helping to visually identify patterns of interest. The first thing that is noticeable in this cladogram is the division into three major clades or groups:

- Normal activity gone wrong
- Fraudulent activity
- Collection of *simple* events characterised by the use of a range of derivatives

These can be considered the fundamental risk elements. Essentially, the presence or absence of fraud defines the first major break in lineage. We can then analyse which event types are more evolved than others by analysing the branch length as shown in Figure 42.

[20] Cladograms produced by Neil Allan from Systemic Consult using Evolutionary Risk Analysis software.

FIGURE 42: CLADOGRAM OF LARGE DERIVATIVE LOSS EVENTS AND CHARACTERISTICS: EVOLUTIONARY EVENTS

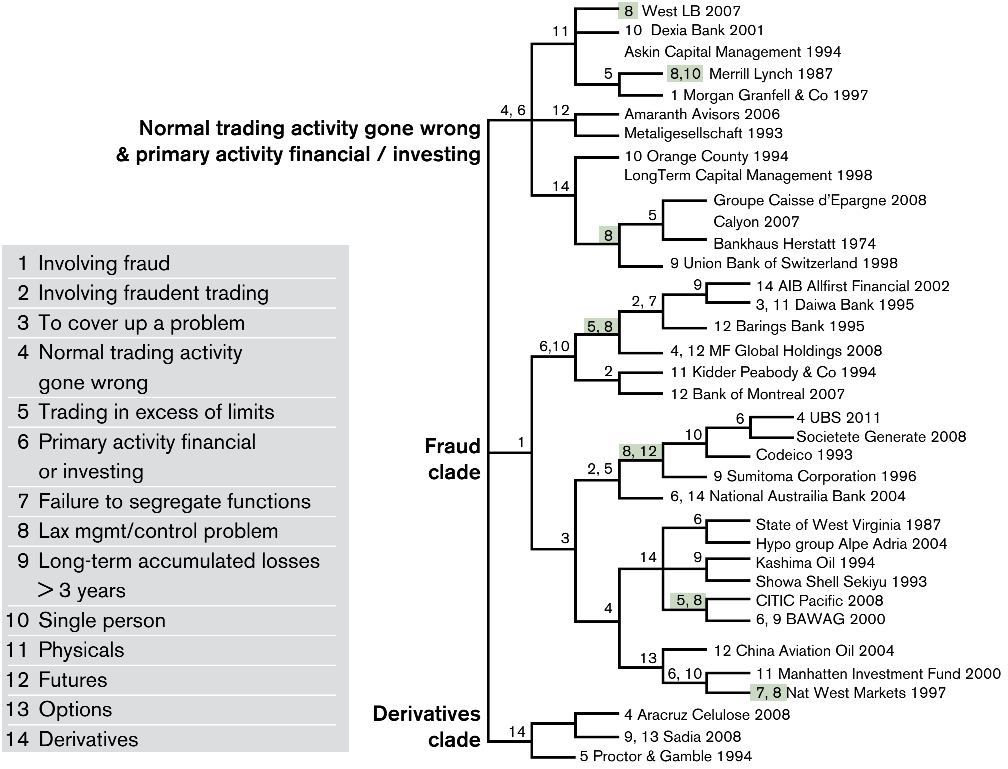

Normal trading activity gone wrong & primary activity financial / investing
Fraud clade
Derivatives clade
1 Involving fraud
2 Involving fraudent trading
3 To cover up a problem
4 Normal trading activity gone wrong
5 Trading in excess of limits
6 Primary activity financial or investing
7 Failure to segregate functions
8 Lax mgmt/control problem
9 Long-term accumulated losses > 3 years
10 Single person
11 Physicals
12 Futures
13 Options
14 Derivatives
4, 6
11
8 West LB 2007
10 Dexia Bank 2001
Askin Capital Management 1994
5
8,10 Merrill Lynch 1987
1 Morgan Granfell & Co 1997
12
Amaranth Avisors 2006
Metaligesellschaft 1993
14
10 Orange County 1994
LongTerm Capital Management 1998
8
5
Groupe Caisse d'Epargne 2008
Calyon 2007
Bankhaus Herstatt 1974
9 Union Bank of Switzerland 1998
1
6,10
5, 8
2, 7
9
14 AIB Allfirst Financial 2002
3, 11 Daiwa Bank 1995
12 Barings Bank 1995
4, 12 MF Global Holdings 2008
2
11 Kidder Peabody & Co 1994
12 Bank of Montreal 2007
3
2, 5
8, 12
10
6
4 UBS 2011
Societete Generate 2008
Codeico 1993
9 Sumitoma Corporation 1996
6, 14 National Austrailia Bank 2004
4
14
6
State of West Virginia 1987
Hypo group Alpe Adria 2004
9
Kashima Oil 1994
Showa Shell Sekiyu 1993
5, 8
CITIC Pacific 2008
6, 9 BAWAG 2000
13
12 China Aviation Oil 2004
6, 10
11 Manhatten Investment Fund 2000
7, 8 Nat West Markets 1997
14
4 Aracruz Celulose 2008
9, 13 Sadia 2008
5 Proctor & Gamble 1994

The bottom highlighted group, the derivatives clade, shows very little evolutionary process. These events can be considered to be relatively stable and unchanging in nature. They are the crocodiles of the risk world–they have reached their evolutionary peak and show little sign of emergent behaviour. Hence companies with similar operational activities to the sequence of characteristics of the events may be less likely to be exposed to emerging risks.

In contrast to these events, the two most evolved groups in the fraud clade show significant evolution through a large number of bifurcations in characteristics. They can be considered to be highly evolved risk events, essentially derivatives of earlier risk events that occur further back along the branch path. These types of events should be studied in detail, as they are likely to give us greater insight into the types of events that are more likely to be subject to evolutionary forces in the future. Companies with similar operational activities to the sequence of characteristics of the events may be more likely to be exposed to emerging risks. Generally, we see an increased complexity in the new risks that evolve in these highly active areas.

Figure 43 looks at the characteristics that are defining the evolutionary process.

FIGURE 43: CLADOGRAM OF LARGE DERIVATIVE LOSS EVENTS AND CHARACTERISTICS: CHARACTERISTICS


1 Involving fraud
2 Involving fraudent trading
3 To cover up a problem
4 Normal trading activity gone wrong
5 Trading in excess of limits
6 Primary activity financial or investing
7 Failure to segregate functions
8 Lax mgmt/control problem
9 Long-term accumulated losses > 3 years
10 Single person
11 Physicals
12 Futures
13 Options
14 Derivatives
Normal trading activity gone wrong & primary activity financial / investing
Fraud clade
Derivatives clade
4, 6
11
8 West LB 2007
10 Dexia Bank 2001
Askin Capital Management 1994
5
8,10 Merrill Lynch 1987
1 Morgan Granfell & Co 1997
12
Amaranth Avisors 2006
Metaligesellschaft 1993
14
10 Orange County 1994
LongTerm Capital Management 1998
8
5
Groupe Caisse d'Epargne 2008
Calyon 2007
Bankhaus Herstatt 1974
9 Union Bank of Switzerland 1998
1
6,10
5, 8
2, 7
9
14 AIB Allfirst Financial 2002
3, 11 Daiwa Bank 1995
12 Barings Bank 1995
4, 12 MF Global Holdings 2008
2
11 Kidder Peabody & Co 1994
12 Bank of Montreal 2007
3
2, 5
8, 12
10
6
4 UBS 2011
Societete Generate 2008
Codeico 1993
9 Sumitoma Corporation 1996
6, 14 National Austrailia Bank 2004
4
14
6
State of West Virginia 1987
Hypo group Alpe Adria 2004
9
Kashima Oil 1994
Showa Shell Sekiyu 1993
5, 8
CITIC Pacific 2008
6, 9 BAWAG 2000
13
12 China Aviation Oil 2004
6, 10
11 Manhatten Investment Fund 2000
7, 8 Nat West Markets 1997
14
4 Aracruz Celulose 2008
9, 13 Sadia 2008
5 Proctor & Gamble 1994

Characteristics that appear frequently are more likely to appear in the future. The sequence of characteristics can also be important, as some tend to occur toward the end of branches rather than at the beginning. For example, characteristic 9 (long-term accumulated losses in excess of three years) always occurs at the end of a branch structure, indicating that it could readily jump across to another branch to define a new emerging risk characteristic.

We have highlighted bifurcations involving characteristic 8 (lax management/risk control problem). This is a very common characteristic as it is evident in almost all branches/events. In many cases, it is also evolving jointly along with a number of other characteristics such as:

- 10: Single person
- 5: Trading in excess of limits
- 7: Failure to segregate functions

Characteristics 8 (lax management/risk control problem) and 5 (trading in excess of limits) seem in particular to be very closely related in evolutionary terms as they tend to appear together in sub-clades. Note that this seems somewhat logical in hindsight, but we arrived at this conclusion through an objective analysis based purely upon a rich classification dataset. This could be very important information as it provides clues as to what characteristics emerging risk events might have in the future. From this we can then ask more focused questions of ourselves such as:

- If we see one of the jointly evolving characteristics in one branch but not the other, we can ask why, as it is breaking an evolutionary trend.

- What would events such as the next West LB (very top) or NatWest Markets (near bottom) look like, if they evolved to contain a characteristic 5 (trading in excess of limits) as they already have a characteristic 8?

- What would this event possibly look like if it happened at my organisation?

9.4 Implications

This emerging operational risk framework has a number of implications.

The first is that risk can be seen to be an evolutionary process, which gives rise to emerging risks. This will be the case whenever the underlying system is a complex adaptive one, rather than a static or chaotic one. Investigating the evolving characteristics of system events in the past can provide insight into our understanding of how emerging risks might occur in the future.

The second is that it is important to capture multiple characteristics of operational risk events, both in terms of realised historical events, as well as forward-looking events. Valuable information may be lost if risks are forced into only single categories or characteristics, which may be the case if risk register software constraints exist, if a prescriptive risk classification framework is narrowly defined, or if the emerging risk identification approach is biased from the outset to focus on single operational processes or risk silos. The quality and completeness of loss data collection and the classification process become critical activities in the emerging risk process.

The third is that the risk taxonomy for operational risk can be determined objectively from the data, rather than being defined prescriptively in an ex ante sense. Risk taxonomies are almost always defined on the latter basis, resulting in linear structures, which is appropriate whenever system complexity is low. However, humans tend to overly simplify situations where there is complexity, losing valuable information in the process. By defining the risk taxonomy objectively through this framework, we are able to map the interrelationships and connectivity between different risk branches, to gain insight into how risk events are truly related.

This is closely related to the discussion on the boundary between operational risk and other risk types. Whilst it is a natural human response to try to carve everything up neatly into independent risk silos, with operational risk it is not quite as appropriate to do so because of the high degree of interaction with other risk types. It is also much more of a process rather than a single event, so its nature can and will tend to change over time. The example above of derivative losses is a good one as there are clearly elements of market risk, operational risk, and liquidity risk involved. We would suggest that we need to move beyond the traditional silo view to understand and ultimately to manage risks that span multiple silos.

The final implication is that the above framework provides a structured way of addressing emerging risk. It is another lens through which we can possibly gain insight into future emerging risk events that we haven't yet seen, when we are not sure yet exactly what we should be looking for.

10 LOSS DATA COLLECTION

10.1 What is loss data collection?

One of the key elements to identify, measure, and manage operational risk is the availability of comprehensive loss data information. The process of loss data collection (LDC) is not just a question of the collection of data, but also the framework of how the data is specified. As previously noted in Section 6, the availability of homogeneous, complete, and reliable data is critical to the development, calibration, and use of operational risk measurement and management models, which support the identification and planning of operational risk mitigation actions.

Operational risk is usually defined by the cause of loss rather than the effect of the risk. However, both are equally important in understanding operational risks and hence both risk causes and effects should be identified when recording loss data.

In this context, the main objective of the LDC process is the assessment of the cause and effect generated by an operational event (i.e., the quantification of the loss). Effects should capture both direct financial/economic impacts, as well as non-economic impacts (e.g., reputational damage, loss of/damage to lives) and opportunity costs. A loss event is defined as a specific detrimental state of the business that generates one or more harmful effects (i.e., losses). Such events may also lead to further loss events.

The collection process should include potential losses and should allow the analysis of recorded losses depending on their causes (i.e., risk factors/risk drivers). The collection of information solely on the effect of the risk is generally not sufficient, and each loss should ideally be analysed across various dimensions.

The data recorded for each loss can generally be split into two main categories:

- **Quantitative information**, such as the severity of loss, both before and after any recoveries (e.g., insurance)

- **Qualitative information**, which enables a more detailed assessment of the nature of the loss such as the causal risk factor drivers and segmentation information such as geography, business line, event type, and activity type

10.2 Methodology

In this subsection we discuss the theoretical and practical aspects of a best practice loss data collection process. LDC is, or at least it should be, much more than just a ritual process of recording loss amounts into a database. Instead, a mature and valuable process involves several steps or phases, with each presenting various challenges. These steps include:

1. Definition and identification
2. Loss data elements
3. Validation
4. Analysis
5. Reporting

10.2.1 Step 1: Definition and identification of the loss

The first step is to define how the losses will be identified, which should be included in the database. This should cover the following aspects:

- Who is responsible for the identification of losses? For example, should it be from the top down (e.g., senior management), or from the bottom up with validation from senior managers?
- Whether the loss should be identified by the business area impacted by the loss or by the business area related to the cause of the loss.
- The loss data repository system.
- The level of automation of loss collection.
- The thresholds used to define the level of detail required to be captured, including any minimum amount for reported losses.

The main challenges that need to be addressed with this process are the possible biases that could result from people being reluctant to record losses caused by their business function or by themselves. This is largely a function of the risk culture within the organisation. The boundary between automated versus human loss recording also needs to be clearly defined and understood.

10.2.2 Step 2: Loss data elements

The next step is to identify the quantitative and qualitative dimensions of the loss to be captured. The typical quantitative factors that need to be included are:

- The magnitude of the loss.
- Whether the loss is known with certainty or not. If a loss is uncertain, then what criteria are used to evaluate the probable loss? In some cases the maximum potential loss amount could be used, which is a conservative estimate, or alternatively the mean or mode of the loss could be appropriate, where a probability distribution effectively exists.
- The identification of any recoveries that might be known or expected, and the general criteria used to estimate them. In most of the cases recoveries are uncertain when the loss produces its effect (i.e., when it is included in the database). Moreover, in the case of more than one recovery, it could be that the various recoveries are not simultaneous, which means the estimation at the time of initial recording should take this effect into account. Rules should also be specified as to when to update the loss database when recovery information becomes available.

The typical qualitative factors to be included need to take account of:

- Identification of a criteria to register the so-called *linked event* (i.e., a single event which impacts more than one business line) or when the responsibility for an event is factually unclear. In some cases where a linked event occurs, the loss is assigned to the business in which the event began and when the responsibility is unclear the event can be split to more than one business line/process/risk factor (different criteria could be used, such as the owner of the transaction, the business process out of which the event arose, etc.).
- Identification of robust criteria to register losses which are spread in time and/or which could be classified using multiple approaches (e.g., they could be associated to more than one event, etc.). The criteria should minimise the risk of double-counting losses and/or omissions (i.e., losses can be registered twice or more when associated to different events and/or not registered at all–this risk is higher in a bottom-up process).

- Identification of the key parameters to be collected for each loss. It's crucial to identify the minimum set of qualitative parameters which should be collected. Additional data could be recorded for the *large loss event* (i.e., an event which causes losses higher than a predetermined threshold for large loss events). The required information should be a balance between what is useful to generate significant clusters of the data versus what can be efficiently collected given the collection process.

10.2.3 Step 3: Validation

The third step involves validating the data collected. The purpose of validation of internal loss data is to ensure that data capture is complete, accurate, and appropriately classified. Given the nature of the LDC activity, the validation should be an ongoing process. As for any other check, validation should be properly documented and performed by an independent function/person with respect to the one who collected the information (the so-called *second line of defence*).

Aspects which need to be addressed in this step include:

- Identification/creation/allocation of the role(s) within the organisation to do the validation. This will depend on the LDC approach and organisational hierarchy.

- Timing of the validation sign-off process.

- The level of validation required. This should balance the cost of validation with the materiality of the benefit.

- The appropriate level of documentation required for the validation process, which can be important because of the multiple temporal dimensions to loss recording.

10.2.4 Step 4: Analysis

This step involves analysis of the data which has been collected. Analysis of the data involves quantifying the loss frequency and severity distributions by segment/cluster. The means, medians, variances, and tail quantities of this distribution are used to understand not only actual operational performance, but are also used to calibrate and validate the forward-looking operational risk models. One of the key issues that must be dealt with here is the level of granularity used to define each segment. More granularity leads to increased homogeneity, but at the likely cost of reduced data sufficiency, particularly for low likelihood high severity risks.

10.2.5 Step 5: Reporting

The final step involves the development, production, and distribution of suitable reports for each type of stakeholder, both internal and external. This typically requires some form of reporting system which can access the data and produce reports in real time, ideally with functionality that enables the user to drill down into each segment of the data. As with any management information process, the scope, methodology, validation, and frequency of reports generated need to be well defined, and the reporting process needs adequate resources to be effective.

10.3 Important design elements

10.3.1 Recoveries management

One of the important design elements of the LDC process is the ability to record information on recoveries. A recovery is a dependent event, separate in time from the original loss event but intimately related to it, which reduces the size of the loss. A recovery arising from insurance policies, settlements, and/or repayments could either be direct or indirect. An indirect recovery relates to a recovery paid for in advance, such as a recovery of loss through an insurance contract, whereas a direct recovery relates to a recovery of loss that is obtained after the event.

Recoveries should be captured as a separate field in the database and reports, and should be able to capture the information on losses both gross and net of all recoveries. A decision should also be taken as to whether to include or exclude expected recoveries on potential losses.

Any severity thresholds used are typically applied to gross losses, rather than losses net of recoveries, although there is information value in both measures.

The LDC process should allow for material durations between the date of an event occurrence and the dates of loss estimation and related recoveries. The LDC process should therefore allow for initial estimation of the losses/related recoveries, which are refined in various subsequent steps. A quite common practice is to record information on recoveries on a dedicated and separate session of each loss, where detailed information on each recovery can be recorded and managed, for example:

- Recovery type (direct, indirect)
- State of the recovery amount (estimated, actual)
- State of the recovery (closed or open)
- Expected amount
- Actual amount received
- Additional information to identify the recovery (e.g., insurance company and policy number in the case of insurance recoveries)

10.3.2 Choice of granularity

Another key design element of the LDC process is the choice of granularity. This should be struck at a level that balances:

- The need for detailed information to calibrate operational risk models and to inform the effectiveness of actual operational risk mitigation and management activities.

- The need for data sufficiency in order to have meaningful summary statistics for the above.

- Appropriate breadth to enable analysis from different and multiple perspectives. Ideally, data should be recorded at the most granular level possible, enabling the risk structure/segmentation to be determined objectively on an ex post basis.

The Working Paper Basel Committee[21] is considered as the starting reference point for setting up the classification of events and losses in an LDC. This could be a guide for both banks and insurance companies as well as for any other companies that can just apply appropriate adjustments to fit the indication to the operational risk type/organisation specificity of the relevant business.

The model proposed by the Basel Committee breaks down operational risk exposures and losses into the following categories:

- **Time**: Dates of occurrence, discovery, loss severity estimation, loss severity finalisation, and profit and loss (P&L) recognition.

- **Geography**: Country code where the loss occurred, and where it was booked.

- **Business line**: The functional business unit(s) involved in the loss.

- **Product/service type**: The products and services that form the sources of revenue to which the operational loss relates.

[21] Refer Working Paper Basel Committee (2001).

- **Process type**: The operational activities to which the loss relates.

- **Event type**: The nature of the underlying event.

The classification of loss event type represents the heart of a loss data collection and the Basel Committee suggests the use of three levels of granularity:

- Level 3 category, which identifies the single activity which caused a loss (e.g., transactions not reported–intentional, misappropriation of assets, malicious destruction of assets, etc.)

- Level 2 category, which identifies the *activity type* that caused the loss (e.g., unauthorised activity, theft and fraud, etc.)

- Level 1 category, which is the least granular and identifies seven macro activity types (i.e., internal fraud; external fraud; employment practices and workplace safety; clients products and business practices; damage to physical assets, business disruption, and system failures; execution; delivery and process management).

10.3.3 Data quality

One of the most important aspects of the LDC process is to ensure that data captured is of sufficient quality for the uses to which it will be put. This can be quite challenging given the potentially large amount of data that could be or is collected in the LDC process, as well as the nature of the loss assessment process, which can involve expert judgment, certain data, and quantitative estimates. For this reason, the data validation process should cover a number of dimensions such as loss definition, identification of loss sources, estimation methods, completeness of data capture, consistency of categorisation, etc.

An adequate validation process should therefore comprise all the dimensions of the LDC process, including both sample tests and regular reconciliations.

To stress the relevance of the validation process, the global operational loss database manager for the banking sector ORX[22] supports data quality through processes operated by the Quality Assurance Working Group (QAWG),[23] which specified a number of tests to be applied to data as well as an *annual data attestation exercise* and *periodic portfolio reviews*. Companies sending data to ORX are expected to conduct an annual data quality review involving an independent party with appropriate expertise.

10.3.4 Data sufficiency

The availability of complete, consistent, and meaningful data is a critical issue in operational risk management and this is true both from operational risk modelling and management points of view.

Bias can be introduced in a number of ways through the data collection process. Firstly, in some instances, data may be completely missing for particular business lines, event types, or time periods.

Data might also be biased because of the use of minimum thresholds in order to control the cost of recording high-frequency, low-severity losses. This approach, which is necessary and reasonable from a practical point of view, can lead to significant distortion of statistical analytic results. An example of this is ORX,[24] whereby only gross losses equal to or greater than EUR 20,000 (as an individual event or an aggregate event amount) are recorded.

[22] Refer ORX (2011)
[23] See QA Working Group (QAWG) Home Page, available at http://www.w3.org/QA/WG/.
[24] Refer ORX (2011).

For certain event types and business lines, an organisation might have no known or recorded operational losses. However, this does not mean that their likelihood of occurring is zero. In these cases, access to a database of external loss events for similar organisations could provide a valuable source of information in which to address such biases.

10.4 Risk measurement and management use cases

A valid and complete loss data collection process, both qualitatively and quantitatively, represents a vital source of empirical information that enables the assessment of operational risk through the application of statistical models. Although this is an important use case from the perspective of this report, an LDC process has various other use cases that are just as important, if not more so. These include:

- An LDC process can help organisations which do not have sophisticated statistical models to quickly estimate the order of magnitude of particular operational risks. This can facilitate a quick comparison of current operational risk exposure against historical levels, risk appetite levels, external regulatory requirements, and competitors.

- An LDC process can strengthen risk management culture and force a higher level of risk consciousness within the organisation.

- An LDC process can be used to monitor and manage current operational risk exposure, providing useful information to implement risk mitigation strategies. For example, an organisation can identify the processes and business functions where more operational losses have been registered, and decide to change the focus of current risk mitigation efforts in order to reduce losses. Information captured on actual costs of the operational risks enables the cost-benefit analysis of possible mitigation strategies to be completed.

10.5 Possible future developments

One of the main challenges in designing and implementing an LDC process is the lack of low frequency high severity losses. Outside of this, it can also be challenging for a single organisation to generate statistically adequate and robust datasets for each risk segment.

For these reasons, it can be beneficial and, in some cases, necessary to create a database which registers loss information coming from various organisations in an industry. A primary example of this is the ORX[25] database for the global banking industry, covering 62 members from 18 countries (as at November 2012).

The use of external data varies considerably by industry and organisation. Some companies use external data to fill gaps in internal data, and some use it to inform scenario analysis, whilst others compare the frequency of internal losses to external losses. Having access to an external database can really only be a positive thing, because internal databases by definition only register incidents that have already occurred in the company, which is not a complete picture of forward-looking events. In order to obtain a more realistic basis of measurement, sampling of loss data obtained from other like organisations can add significant value. The use of this data, however, requires additional effort in order to analyse it and make any necessary adjustments such that it becomes appropriate to the specifics of the primary organisation.

One possible future development could be the creation of external loss databases for organisations in industries other than banking. This would enable contributing members access to aggregate data on a no-name basis, and provide a valuable resource for academics and other researchers to use in order to generate insightful research to the benefit of the industry. There is potentially much to be gained from leveraging the framework provided by ORX to expand it to other industries. Whilst there

[25] See the ORX Association website at http://www.orx.org/.

would be significant value to be derived from this, there are also a number of challenges that would need to be overcome, including:

- A minimum number of organisations are required at outset to contribute data in order for a sufficient aggregation of data to occur such that any one organisation's data is not able to be inferred.

- A central processing function or organisation needs to be established (such as ORX) which can support an external LDC process, which might also require an element of funding. This organisational function should provide not only an efficient mechanism to collect and report data, but it should also provide relevant warranties to all parties in term of data protection. It could be necessary to gain the support of industry, regulatory, or governmental associations.

11 CASE STUDIES

11.1 Basic indicators and standard formula approaches

The main purpose of using basic indicators and standard formulas for calculating capital requirements is to keep the calculations as simple and as widely applicable as possible. As such, they are unlikely to appropriately assess operational risk for all possible financial products, either in the market or that may possibly be conceived. It is thus important to be aware of their main drawbacks and weaknesses. In order to help illustrate them, we provide some case study examples below of situations in which products with arguably comparable risks result in significantly different operational risk capital requirements according to these methods.

Let us for example consider two life insurance products: an endowment and a level term policy. The difference between them is that while the latter pays the sum assured only in case of death within the specified term, the former also benefits policyholders who survive the whole period. Because the probability of survival is much higher than the probability of death for the vast majority of policyholder ages, the premiums for the endowment are usually much greater than those for the level term policy; and a larger proportion of them need to be kept as a reserve to fund the payouts.

To illustrate this, consider a European insurance company subject to the Solvency II regulatory framework, with a mixed portfolio consisting of savings, unit-linked products, and risk products, with the number of policies comparable between these categories.

Let's assume that in 2011 the company had EUR 124 million of operational risk (SCR_{Op}). EUR 6.6 million of this was generated by EUR 655,000 unit-linked product expenses, whilst the rest related to traditional products. The average annual premium of the 839,000 pure risk policies is EUR 838, whilst on 467,000 non-unit-linked savings products it is EUR 4,726.

Suppose that, while keeping average amounts and the total number of policies constant, the composition of the company portfolio changes significantly, and all policies become non-unit-linked savings products. In such a case, Op_{prem} would increase to EUR 370.9 million—that is, almost triple the original amount. On the other hand, if the portfolio comprised only pure risk policies, Op_{prem} would reduce to EUR 68.9 million—almost by 50%. To highlight an even more extreme event, if all those policies were unit-linked, operational risk capital would reduce to less than EUR 20 million—which is slightly more than 5% of the amount in the first of our cases.

Details of these calculations can be found in Figure 44 (number of policies and averages are for traditional business).

FIGURE 44: EXAMPLES OF CAPITAL REQUIREMENTS: INSURANCE STANDARD FORMULA

	BASE CASE	CASE RISK	CASE SAVINGS	CASE UNIT-LINKED
NUMBER OF POLICIES	1,306,940	1,962,214	1,962,214	-
AVERAGE PREMIUM	2,254	878	4,726	-
OP_{PREM}	117,854,099	68,915,310	370,943,999	-
AVERAGE RESERVE	14,063	383	39,318	-
OP_{PROV}	82,709,502	3,384,348	347,180,017	-
UL EXPENSES	26,210,936	-	-	-78,488,546
SCR_{OP}	124,406,833	68,915,310	370,943,999	19,622,136
% OF BASE CASE	100%	55%	298%	16%

It is hard to find any justification of such large differences in operational risk across these scenarios. Operational risk in the different kinds of policies doesn't seem to exhibit a different profile; this is particularly true in the case of traditional products with equal sum assured. Pure risk products can also be more prone to operational losses than saving products, which is not captured in this example.

The actual historical growth of the company also highlights a different problem with this approach. Because of the favourable economic conditions, premiums of traditional products in 2011 have grown by 72% (from almost EUR 3 billion to about EUR 5 billion), while the number of policies has been relatively unchanged. This caused the SCROp to grow by 127% (to EUR 270 million). It is doubtful whether this increase represented growth in operational risk exposure as the company is unlikely to have changed operational processes or activities, and had even experienced this situation before. Clearly the use of such simple methods can be very misleading.

11.2 Loss distribution approach

This section provides an example of the Solvency II solvency capital requirement (SCR) simplified calculation, for a medium-size Italian insurer. This case study example will highlight the main steps and aspects of the loss distribution approach (LDA).

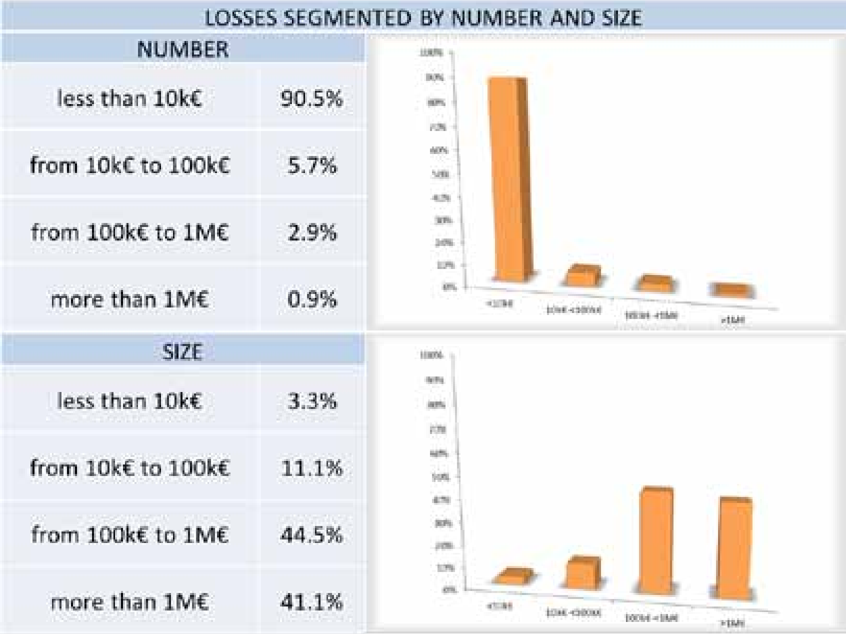
11.2.1 Company context

This case study is based on an assumed medium-size Italian bancassurance company, with a retail business selling unit-linked, with-profit, and term products through the bank's agency distribution network.

Assume that the LDC database used is well structured, but with just three years of history, so it needs to be integrated with external data and expert opinions to build a sound model. For these reasons, a simplified LDA model will be used, grouping data into seven event type risk categories.

Analysing the gross data, total losses amount to EUR 9 million, with 90% of loss events producing losses that are less than EUR 10,000, and less than 1% being greater than EUR 1 million, as shown in Figure 45.

FIGURE 45: EXAMPLES OF LOSS SEGMENTATION BY LOSS SEVERITY BUCKETS



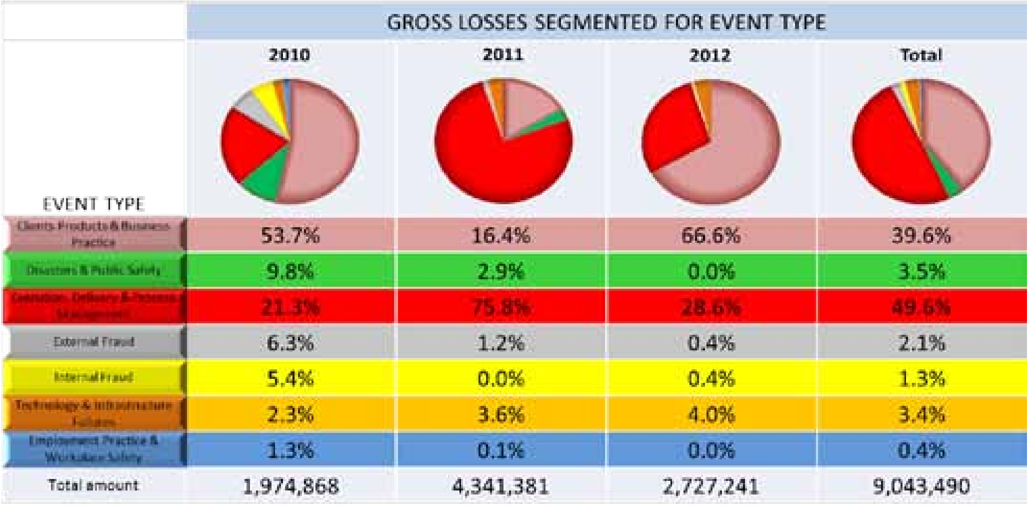
LOSSES SEGMENTED BY NUMBER AND SIZE

NUMBER	
less than 10k€	90.5%
from 10k€ to 100k€	5.7%
from 100k€ to 1M€	2.9%
more than 1M€	0.9%
SIZE	
less than 10k€	3.3%
from 10k€ to 100k€	11.1%
from 100k€ to 1M€	44.5%
more than 1M€	41.1%

Notice that 90% of events produce just 3.3% of the total losses, whilst losses over EUR 1 million are only 0.9% of the total number but collectively account for 41.1% of total losses. This is the numerical evidence of the dominance of low frequency high severity events on total losses.

If we segment data by event type, it is clear that it is dominated in size by two event types: *Clients, products, and business practice* and *Execution, delivery, and process management*. They cover 90% of the losses generated over the three years; and each event type also exhibits volatility across the different years. Figure 46 summarises these observations.

FIGURE 46: EXAMPLES OF LOSS SEGMENTATION BY EVENT TYPE



	GROSS LOSSES SEGMENTED FOR EVENT TYPE			
EVENT TYPE	2010	2011	2012	Total
Clients Products & Business Practice	53.7%	16.4%	66.6%	39.6%
Disasters & Public Safety	9.8%	2.9%	0.0%	3.5%
Execution, Delivery & Process Management	21.3%	75.8%	28.6%	49.6%
External Fraud	6.3%	1.2%	0.4%	2.1%
Internal Fraud	5.4%	0.0%	0.4%	1.3%
Technology & Infrastructure Failures	2.3%	3.6%	4.0%	3.4%
Employment Practice & Workplace Safety	1.3%	0.1%	0.0%	0.4%
Total amount	1,974,868	4,341,381	2,727,241	9,043,490

If we look at the data in terms of frequency and severity, we can see how the different event types have different characteristics, as shown in Figure 47.

FIGURE 47: EXAMPLES OF LOSS SEGMENTATION BY LOSS SEVERITY BUCKETS AND EVENT TYPES

YEARLY AVERAGES OF GROSS DATA

EVENT TYPE	TOTAL LOSS (€)	NUMBER OF EVENTS	LOSS PER EVENT (€)
CLIENTS PRODUCTS AND BUSINESS PRACTICE	1,196,640	9	128,211
DISASTERS AND PUBLIC SAFETY	106,673	2	64,004
EXECUTION, DELIVERY, AND PROCESS MANAGEMENT	1,496,192	363	4,122
EXTERNAL FRAUD	63,386	2	38,032
INTERNAL FRAUD	38,867	2	23,320
TECHNOLOGY AND INFRASTRUCTURE FAILURES	103,721	10	10,372
EMPLOYMENT PRACTICE AND WORKPLACE SAFETY	9,018	2	3,865
TOTAL	3,014,497	390	7,736

This data exhibits some consistency with the ORX data[26] and analysis, which states that … *the top three event types [by size] are 'Execution, Delivery & Process Management', 'External Fraud' and 'Clients Products & Business Practice'* … The one difference is that, in our case study, external fraud is not in the top three.

If we focus on the *Execution, Delivery and Process Management* event type, it look like a high-frequency, low/medium-severity event type class, but if a threshold of EUR 1,000 is set, then this changes somewhat, as evidenced by Figure 48.

FIGURE 48: EXAMPLES OF LOSSES CONCENTRATION

YEARLY AVERAGES OF GROSS DATA

EVENT TYPE	TOTAL LOSS (€)	NUMBER OF EVENTS	LOSS PER EVENT (€)
EXECUTION, DELIVERY, AND PROCESS MANAGEMENT (EVENTS THAT PRODUCE A LOSS BIGGER THAN €1,000)	1,482,628	19	79,426
EXECUTION, DELIVERY, AND PROCESS MANAGEMENT (ALL EVENTS RECORDED)	1,496,192	363	4,122

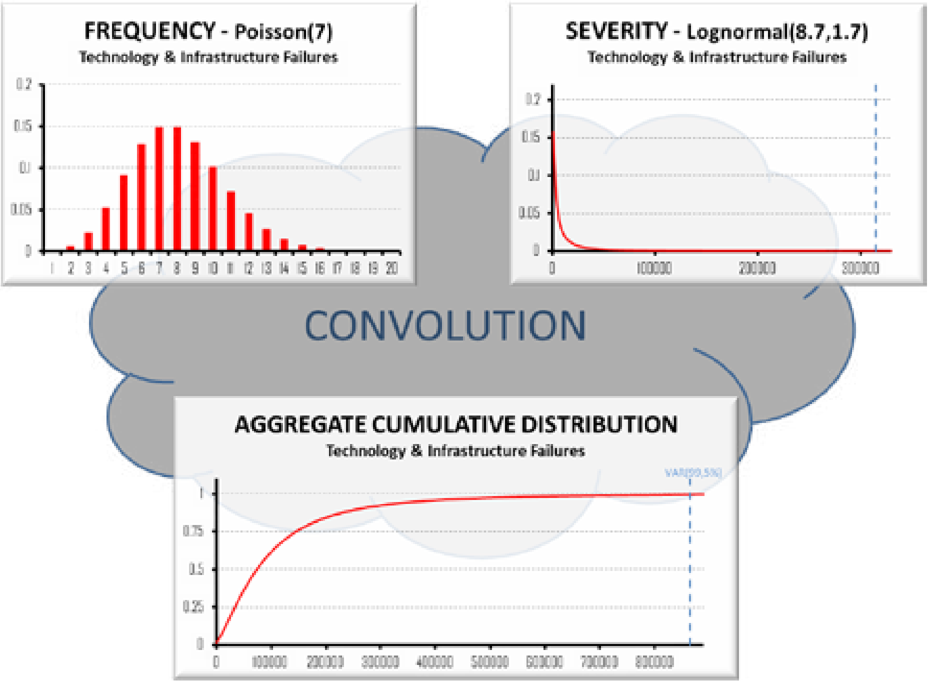
It is clear that almost all the losses are generated by just a few events, so once again the dominance of low frequency high severity events can be observed in a single risk category. This is an easy example that underlines the importance of an adequate segmentation of data for frequency and severity curve fitting, in order to capture skew effects.

11.2.2 Distribution and correlation fitting

Because this case study is intended to be a simplified LDA model, for all the events types, the frequency distributions are modelled with a Poisson process and the severities are modelled with a lognormal distribution (we assume that the reader is familiar with the calibration of these distributions to a given dataset, and do not describe the parameter calibration process). For each risk category, an aggregate cumulative loss distribution is obtained through the Monte Carlo simulation of the convolution of the frequency and severity distributions. For simplicity we have assumed that these distributions are independent of one another. Figure 49 on page 82 shows graphically the shape of the distribution for one event type.

[26] Refer ORX (2011).

FIGURE 49: SUMMARY OF THE MAIN STEPS OF AN LDA



Once this process is complete for each risk category, a simple assessment of model fit can be made by the comparison between the mean of the raw data and the mean of the modelled aggregate distribution. In addition to this, it is possible to calculate the $VaR_{99.5\%}$ statistic for each event type, and the sum of all the VaR numbers will provide a rough benchmark to compare with the total losses distribution VaR (which is due to the effect of diversification, notwithstanding the limitation that VaR is not a coherent risk statistic).

FIGURE 50: EXAMPLES OF CAPITAL REQUIREMENT BY EVENT TYPE

	DATA	MODEL	
EVENT TYPE	TOTAL LOSS (YEARLY AVERAGE)	MEAN	$\text{VaR}_{99.5\%}$
CLIENTS PRODUCTS AND BUSINESS PRACTICE	1,196,640	1,105,944	8,855,135
DISASTERS AND PUBLIC SAFETY	106,673	252,916	9,760,624
EXECUTION, DELIVERY, AND PROCESS MANAGEMENT	1,496,192	1,700,701	5,901,388
EXTERNAL FRAUD	63,386	35,382	502,020
INTERNAL FRAUD	38,867	24,039	587,671
TECHNOLOGY AND INFRASTRUCTURE FAILURES	103,721	117,554	865,331
EMPLOYMENT PRACTICE AND WORKPLACE SAFETY	9,018	10,723	280,751
TOTAL (SUM)	3,014,497	3,247,259	26,752,919*

* This is just the sum of the elements, it is not the VaR for the total losses.

Because of the limited availability of only three years of data, it is not possible to undertake a quantitative estimation of the correlation matrix for the copula method. Instead, a qualitative analysis of data and risk categories compared to the dependence structure in the ORX database was made, which led to a qualitative assessment of the dependencies with the matrix shown in Figure 51.

FIGURE 51: EXAMPLES OF DEPENDENCIES BETWEEN EVENT TYPES

	EVENT TYPES DEPENDENCE HEAT MAP					
	CPBP	DPS	EDPM	EF	IF	TIF
Clients Products and Business Practice (CPBP)						
Disasters and Public Safety (DPS)						
Execution, Delivery, and Process Management (EDPM)	med					
External Fraud (EF)	low		low			
Internal Fraud (IF)				very low		
Technology and Infrastructure Failures (TIF)	med	very low	very low	med	very low	
Employment Practice and Workplace Safety (EPWS)	very low		very low		Low	

Note that the two dominant events (CPBP and EDPM) are correlated, so it will be expected that dependencies are relevant in the calculation of operational risk capital. This is discussed in the following subsection, where an uncorrelated Gaussian copula is used in comparison to a Gaussian copula. In addition to this, a correlated Gaussian copula will also be compared to Student's t-copula with different degrees of freedom, in order to show the impact of the copula used to aggregate risks.

11.2.3 Operational risk capital assessment

The risk measure adopted in Solvency II to calculate the SCR for operational risk is the $\text{VaR}_{99.5\%}$ statistic.

The first step is to choose the type of copula that is the most appropriate to use. The two most used are the Gaussian and the Student's t-copulas. This choice is not arbitrary as it depends on the data, and it must be supported by statistical analysis. The benefit of using the Gaussian copula is that we reduce the impact of the tails. In other words, if we have aggregate fat-tailed distributions, with a Gaussian copula we obtain an aggregate distribution with thinner tails. If we are aggregating heavy-tailed distributions (as it is in most LDA models) a Student's t-copula would be preferred because it can avoid this effect. The Student's t-copula has one parameter, degrees of freedom. The closer it is to zero, the more the heavy tail of the marginal distributions (e.g., the distribution of each event type) are preserved or even enhanced in the aggregation process. The further away it is from zero, the more the copula tends toward becoming Gaussian.

Figure 52 compares the SCR, mean, and median of the total loss distribution obtained using the Gaussian copula and the Student's t-copula with different degrees of freedom (based upon 1,000,000 simulations).

FIGURE 52: EXAMPLES OF CAPITAL REQUIREMENT BY LOSS DISTRIBUTION CURVE ASSUMPTIONS

	GAUSSIAN	STUDENT'S T DOF=100	STUDENT'S T DOF=50	STUDENT'S T DOF=25	STUDENT'S T DOF=10	STUDENT'S T DOF=4	STUDENT'S T DOF=3
MEDIAN	2,693,730	2,612,042	2,615,644	2,630,381	2,668,094	2,776,721	2,836,758
MEAN	3,285,306	3,296,856	3,349,237	3,463,551	3,897,981	6,102,504	8,321,521
VAR99.5%	13,121,395	16,128,774	18,024,158	22,004,621	54,073,759	122,230,574	193,101,285

As expected, moving from a Gaussian to a Student's t-copula increases the SCR. The issue of the VaR statistic not being coherent is also evident here. Whenever the degrees of freedom are 10 or below, the $VaR_{99.5\%}$ of the total loss is bigger than the sum of the VaR of each single risk (circa 26 million), resulting in a non-sensible negative diversification effect.

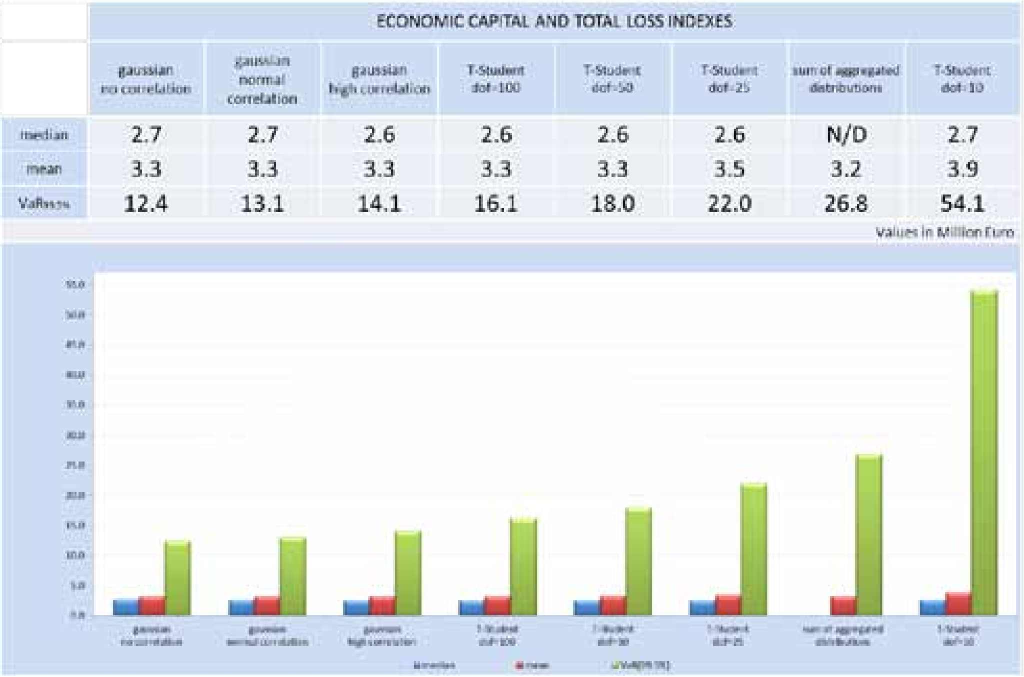
The other key element in this calculation is the effect of the correlation matrix on the SCR. In the previous section we defined a qualitative correlation matrix for the copula. Figure 53 compares the results if we consider normal or high correlation parameters to having zero correlations (i.e., independence).

FIGURE 53: EXAMPLES OF CAPITAL REQUIREMENT BY CORRELATION ASSUMPTIONS

		GAUSSIAN	
	NO CORRELATION	NORMAL CORRELATION	HIGH CORRELATION
MEDIAN	2,745,900	2,693,730	2,635,674
MEAN	3,282,048	3,285,306	3,284,674
VAR99.5%	12,374,408	13,121,395	14,060,921

As expected, the correlation parameters have no material impact on the mean or mediate results, but they do have a material impact on the tails and hence capital. Figure 54 presents the SCR, mean, and median for all the aggregation types.

FIGURE 54: EXAMPLES OF CAPITAL REQUIREMENT: SUMMARY OF RESULTS



	ECONOMIC CAPITAL AND TOTAL LOSS INDEXES							
	gaussian no correlation	gaussian normal correlation	gaussian high correlation	T-Student dof=100	T-Student dof=50	T-Student dof=25	sum of aggregated distributions	T-Student dof=10
median	2.7	2.7	2.6	2.6	2.6	2.6	N/D	2.7
mean	3.3	3.3	3.3	3.3	3.3	3.5	3.2	3.9
VaR$_{99.5\%}$	12.4	13.1	14.1	16.1	18.0	22.0	26.8	54.1

Values in Million Euro

From this case study we conclude that the issues of data segmentation (as shown in Figure 47 above), skewness of distributions, and the choice and calibration of the aggregation measure can all have a material impact on the generation of an operational risk evaluation. In particular, this case study shows the material impact that the choice and calibration of both the severity distributions and the copula have on capital requirements.

With respect to the choice and calibration of the copula it is important to note the key role that expert judgment plays. A robust calibration of the parameters of a copula would ideally require the availability of a significant sample of data. However, available data is typically quite limited, so it is important to undertake ex post coherence analysis to check the sensibility and consistency of the calibration results.

11.3 Structural or causal model

11.3.1 The challenge

Perhaps the most significant operational risk that can and has caused banks to fail is rogue traders. Rogue traders are authorised employees that make unauthorised trades on behalf of their employers. Such traders typically use derivative financial instruments in the context of a proprietary trading operation, where large financial exposures against a wide array of market risks can be created relatively quickly.

Nick Leeson is perhaps the quintessential rogue trader, whose losses in 1995 were sufficient to bankrupt Barings Bank. The table in Figure 55 on page 86 shows a brief selection of some of the largest rogue trader losses that have occurred over the last two decades.

FIGURE 55: EXAMPLES OF ROGUE TRADER LOSSES

NAME	DATE	INSTITUTION	COUNTRY	INSTRUMENTS	LOSS (IN MILLIONS)
Nick Leeson	1995	Barings Bank	UK	Nikkei index futures	GBP 827
Toshihide Iguchi	1995	Daiwa Bank	Japan, United States	U.S. T-bonds	USD 1,100
Yasuo Hamanaka	1996	Sumitomo Corporation	Japan	Copper	USD 2,600
John Rusnak	2002	Allied Irish Banks	United States	Foreign exchange options	USD 691
Gianni Gray, David Bullen, Vince Ficarra, Luke Duffy	2003-2004	National Australia Bank	Australia	Foreign exchange options	AUD 360
Chen Jiulin	2005	China Aviation Oil	Singapore	Jet fuel futures	USD 550
Brian Hunter	2006	Amaranth Advisors	United States	Natural gas futures	USD 6,500
Jerome Kerviel	2006-2008	Societe Generale	France	Euro stock index futures	EUR 4,900
Boris Picano-Nacci	2008	Groupe Caisse d'Epargne	France	Equity derivatives	EUR 751
Kweku Adoboli	2011	UBS	Switzerland	S&P 500, DAX, EuroStoxx index futures	USD 2,300

For risk managers and executives, such events are the stuff of nightmares. Many executives in these businesses at the time of these events describe themselves as being dumbstruck as to how such severe losses could appear to be completely out of the blue.

So how can we as risk managers make these seemingly unpredictable risk events predictable? This is a critical question for any company that operates in the financial markets, particularly those using derivatives, and even more so those who are operating proprietary rather than hedging operations. For banks, this risk is a material part of the operational risk capital that they must hold under Basel II, and for regulators it is a critical risk that has the potential to destabilise the financial system if it occurs to a systemically critical institution.

As risk managers, the key questions we want answers to are:

- How can we identify the conditions that are likely to cause such events?
- How can we assess the likelihood, severity, total expected losses, and capital measures, in order to determine an appropriate level of operational risk capital to hold?
- How can we link risk assessment with the unique and observed state of the business?
- How can we determine what business drivers are the most critical factors in driving this risk?
- Are actual risk exposures within our risk appetite?
- How can we determine what the most effective risk mitigation actions are to reduce this risk?
- How can we continue to do the above whilst our business and operational processes keep evolving?

These questions set a very high standard for a risk framework to live up to, particularly so for such an extreme risk in terms of severity and frequency.

11.3.2 Traditional risk management approach

The traditional approach to managing this risk involves two main avenues:

- Mitigation via operational process controls and risk reporting
- Capital assessment via LDA or scenario approaches

Significant resources are allocated to the management of operational process controls and risk reporting to help companies control the likelihood of these events occurring. Typical risk mitigation activities include:

- The setting of risk exposure limits
- The separation of back, middle, and front office operations
- Frequent risk and audit reporting

No doubt they have been effective in mitigating many potential trading errors and breaches, although clearly they were ineffective in the examples in Figure 55. So the question is how can we ensure that risk mitigation activities are structured and completed in a way that they are most effective in reducing this risk?

From an assessment perspective, both the likelihood and severity of these events present particular problems. For the majority of companies that have not experienced a rogue trader, no internal loss data exists from which to draw upon. Furthermore, the variability in the range of instruments exploited, as detailed in Figure 55, and the unique and differing nature of each one's operations and risk control framework to the business under question, do not engender a high degree of confidence in the use of external loss data for this risk.

Assessment of the frequency of such events is also problematic, particularly when it is necessary to have sufficient accuracy in order to derive a 99.5th or 99.9th percentile result for capital calculation purposes. Humans are naturally pretty poor at assessing the properties of statistical distributions, and the degree of value placed in any one person's judgment as to how often such an event is likely to occur over the span of two to 10 times their own life expectancy must be naturally very low. This also severely flaws the use of scenario analysis, which is heavily dependent upon the unconditioned expectations of individuals hypothesising what such losses might look like.

Such unconditioned statistical approaches can be more harmful than helpful in these situations, as the structural dynamics of the operational process which the statistical process is trying to capture are highly unlikely to be stationary (i.e., static or stable).

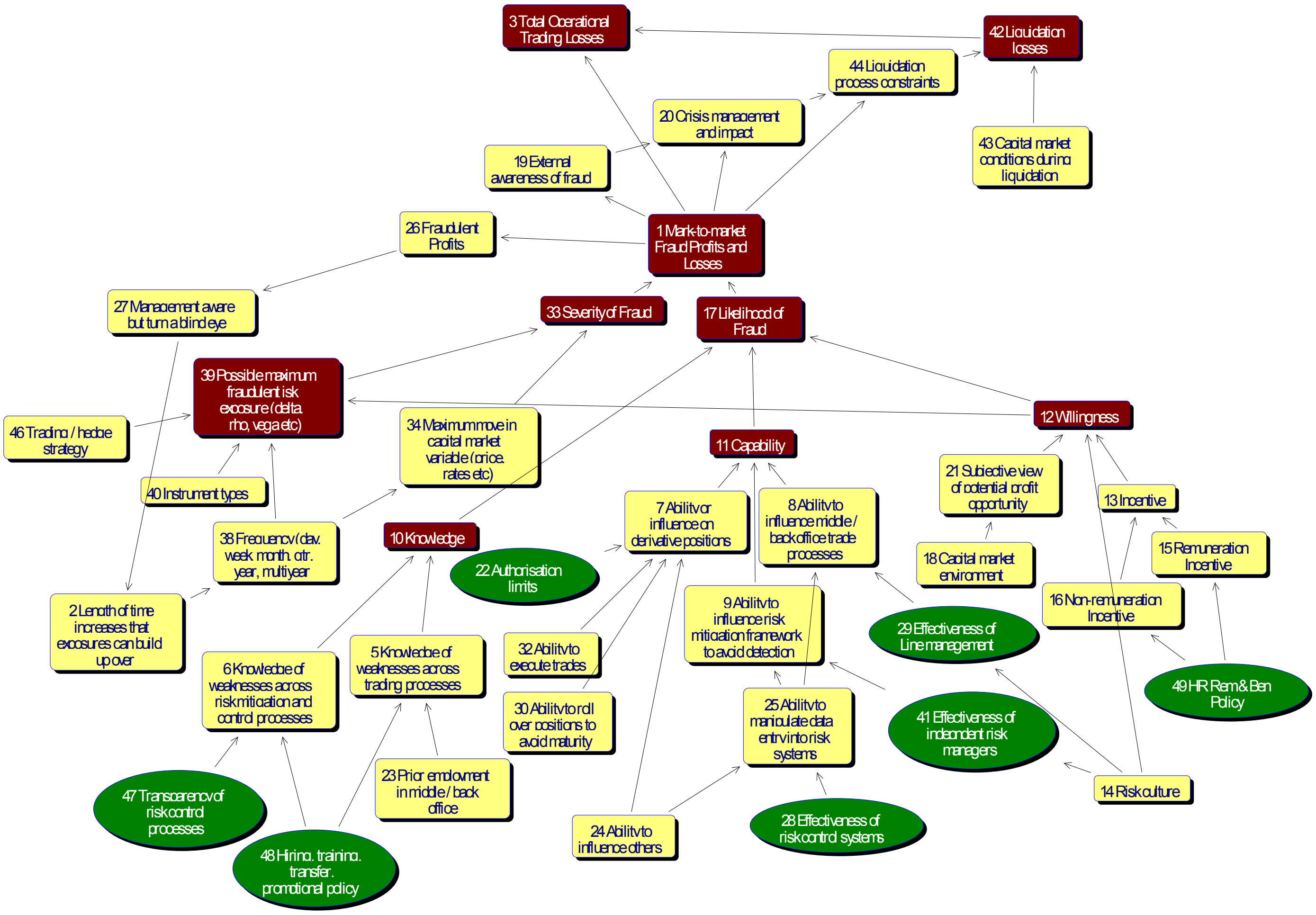
11.3.3 Eliciting the structure of the risk system with cognitive mapping

We believe structural frameworks are the most appropriate way to address the challenges this risk poses. The starting point for such an approach is to elicit the system structure of this risk using cognitive mapping techniques. This technique elicits the system structure through a qualitative assessment of the drivers of the business. This can be done using a variety of sources, including reports of similar internal or external events or through workshops with the relevant subject matter experts involved in all aspects of the business operation.

For the purposes of this case study, we have constructed such a cognitive map based upon the investigation into the Society General event, overlaid with our own subjective judgment leveraging our own derivative hedge management operational expertise in this field. The diagram in Figure 56 shows this cognitive map.

In practice, these cognitive maps would be derived from operational business experts. Multiple views would be taken into account in order to minimise the natural human bias and gaps that any one individual would be subject to.

FIGURE 56: COGNITIVE MAP OF ROGUE TRADING


3 Total Operational Trading Losses
42 Liquidation losses
44 Liquidation process constraints
20 Crisis management and impact
43 Capital market conditions during liquidation
19 External awareness of fraud
26 Fraudulent Profits
1 Mark-to-market Fraud Profits and Losses
27 Management aware but turn a blind eye
33 Severity of Fraud
17 Likelihood of Fraud
39 Possible maximum fraudulent risk exposure (delta, rho, vega etc)
46 Trading / hedge strategy
34 Maximum move in capital market variable (price, rates etc)
12 Willingness
11 Capability
21 Subjective view of potential profit opportunity
13 Incentive
40 Instrument types
7 Ability or influence on derivative positions
8 Ability to influence middle / back office trade processes
38 Frequency (day, week, month, qtr, year, multiyear
10 Knowledge
22 Authorisation limits
18 Capital market environment
15 Remuneration Incentive
2 Length of time increases that exposures can build up over
9 Ability to influence risk mitigation framework to avoid detection
16 Non-remuneration Incentive
29 Effectiveness of Line management
6 Knowledge of weaknesses across risk mitigation and control processes
5 Knowledge of weaknesses across trading processes
32 Ability to execute trades
49 HR Rem & Ben Policy
30 Ability to roll over positions to avoid maturity
25 Ability to manipulate data entry into risk systems
41 Effectiveness of indendent risk managers
23 Prior employment in middle / back office
47 Transparency of risk control processes
14 Risk culture
24 Ability to influence others
28 Effectiveness of risk control systems
48 Hiring, training, transfer, promotional policy

The map in Figure 56 is intended to be indicative only for the purposes of illustrating this case study, rather than a complete description of Society General or any other company's particular risk system structure. Whilst broad similarity would be expected in the structural components of these maps between companies, the more detailed drivers and key risk indicators (KRIs) that accompany them will tend to becoming increasingly unique.

The cognitive map in Figure 56 can be broken down into the following components:

- **The likelihood of fraudulent trading**: This is a function of the knowledge, capability, and willingness of individuals to undertake fraudulent trading activity.

- **The severity of fraudulent trading**: This is a function of the degree to which risk exposures could potentially be increased as well as the state of the market to impact the profit and loss.

- **The crisis management process**: This describes the process by which management resolves the crisis, which can have a very large effect on the final losses incurred.

We now discuss each of these in further detail.

11.3.4 Likelihood of fraud

The likelihood of fraud represents the probability that a fraud event will occur. Three conditions are necessary for trading fraud to occur:

1. **Knowledge**: The person(s) must have the knowledge of the entire trading process, as well as that of the risk control processes.

2. **Capability**: The person(s) must have the capability to influence the trading process and systems.

3. **Willingness**: The person(s) must have the willingness to act to commit fraud.

If only one or two of these conditions are met, then fraud will not occur. Hence the probability of fraud needs to be conditioned upon the state that these dependent drivers are met. For example, Figure 57 outlines a possible joint distribution probability state space reflecting the mean and uncertainty estimates for various combinations of dependent states. These would be determined based upon a combination of expert judgment, historical data, and current measures. The analysis could be segmented to apply to a specific group, team, or individual as necessary.

FIGURE 57: INDICATIVE STATE-BASED PROBABILITY MEASURES OF LIKELIHOOD

JOINT PROBABILITY STATE	KNOWLEDGE	CAPABILITY	WILLINGNESS	MEAN LIKELIHOOD OF FRAUD	UNCERTAINTY LIKELIHOOD OF FRAUD
STATE 1	NO	NO	NO	0.01	0.01
STATE 2	NO	NO	YES	0.05	0.05
STATE 3	NO	YES	NO	0.05	0.05
STATE 4	NO	YES	YES	0.10	0.10
STATE 5	YES	NO	NO	0.05	0.05
STATE 6	YES	NO	YES	0.10	0.10
STATE 7	YES	YES	NO	0.10	0.10
STATE 8	YES	YES	YES	0.95	0.05

Note that the mean and uncertainty estimates refer to the mean and standard deviations of a truncated normal distribution with a range of 0 to 1. These means and uncertainty estimates are highly conditional upon their underlying state variables, and there is a huge degree of nonlinearity involved.

Whether the group/team/individual has the knowledge or not is likely to be (at least) dependent upon:

- Their knowledge of the weaknesses in the front, middle, and back office trading processes. This in turn could be influenced by whether they have held a role in more than one of these positions.

- Their knowledge of the weaknesses in the risk mitigation and control processes.

By identifying these drivers, we can immediately see where risk mitigation actions can be inserted to influence the likelihood of these outcomes:

- Ensure that people have not held roles in all positions within the trading process within the organisation, and potentially not with other similar organisations

- Ensure that the knowledge of the risk mitigation and control processes are kept at least mostly from those of involved in trading

Figures 58-60 show indicative Bayesian models for each of the three elements of likelihood. They have been parameterised using indicative Boolean distributions. Each element would be assessed reflecting both the quantity and quality of any information available that is relevant. For some drivers, there might be a significant amount of high-quality information that can be objectively assessed, such as *prior employment in back office*, whilst other drivers might have a lower degree of data quality, such as *effectiveness of risk managers*. The scenario in blue shows a typical risk measurement scenario, whilst the scenario in green is a reverse stress scenario where the outcome is conditioned to be 100% likely, and the underlying risks are resolved using Bayesian inference.

FIGURE 58: BAYESIAN NETWORK FOR KNOWLEDGE OF WEAKNESSES

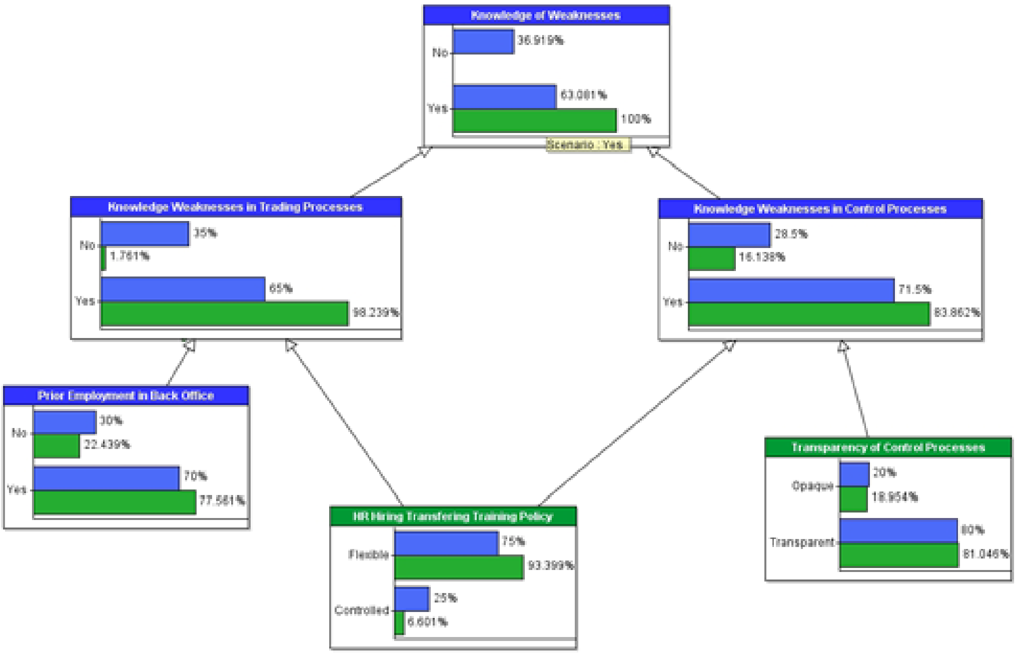


FIGURE 59: BAYESIAN NETWORK FOR CAPABILITY

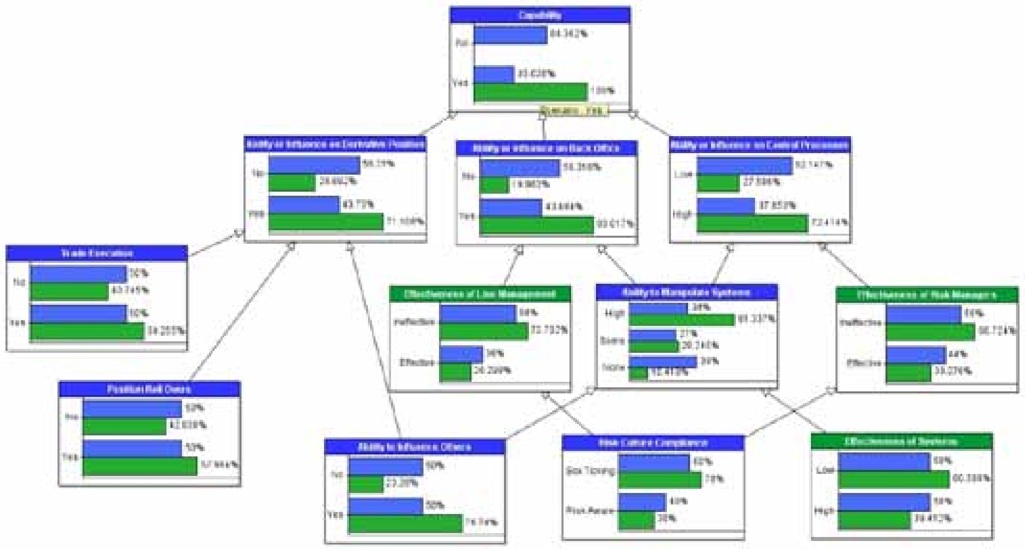
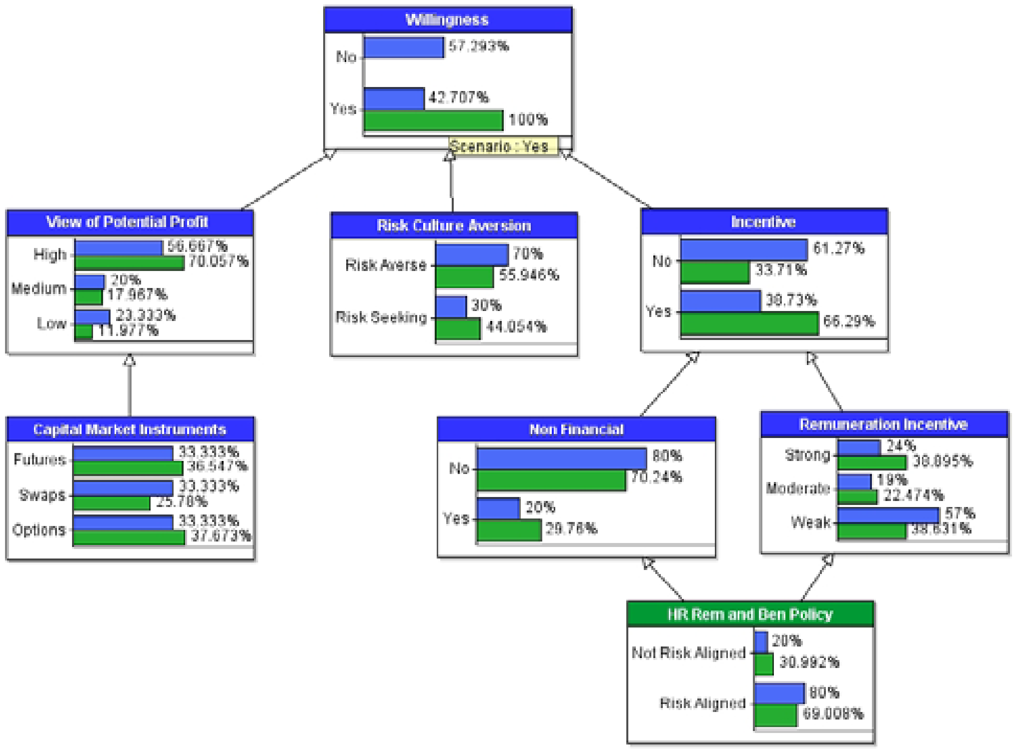


FIGURE 60: BAYESIAN NETWORK FOR WILLINGNESS



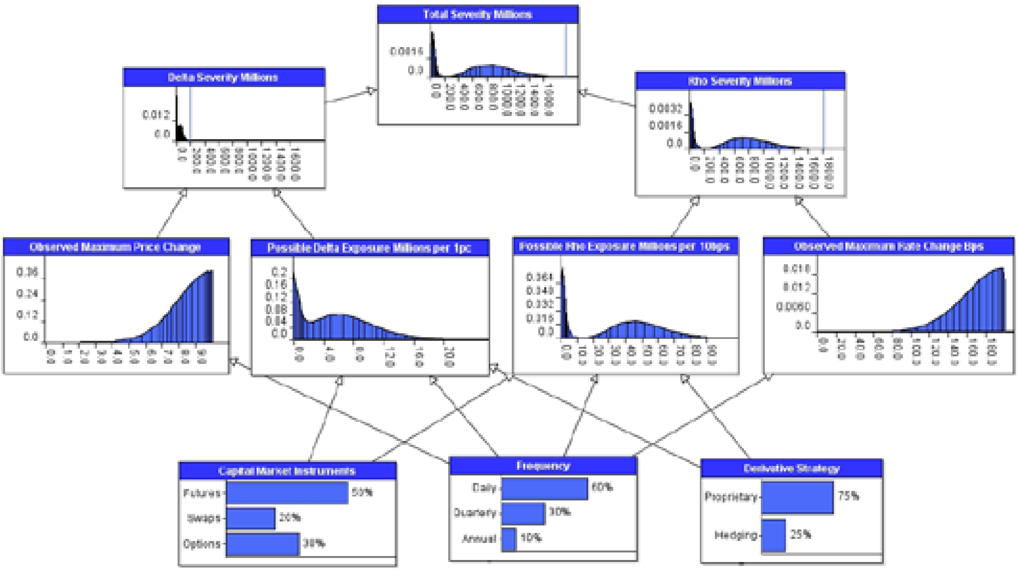
In the reverse stress scenarios above, the source of deterioration relative to the current risk-measured scenario can be clearly identified. This provides important information on where to focus risk control efforts.

11.3.5 Potential fraud severity

Potential fraud severity is dependent upon the types of instruments traded, the frequency with which they are traded, and the nature of the trading strategy (hedging vs. proprietary). These risk factors determine exposures to market levels (delta) and interest rates (rho) and, combined with the exogenous risk of market levels and interest rates, result in potential severity levels.

Figure 61 shows a Bayesian network for these factors using an indicative parameterisation.

FIGURE 61: BAYESIAN NETWORK FOR FRAUD SEVERITY



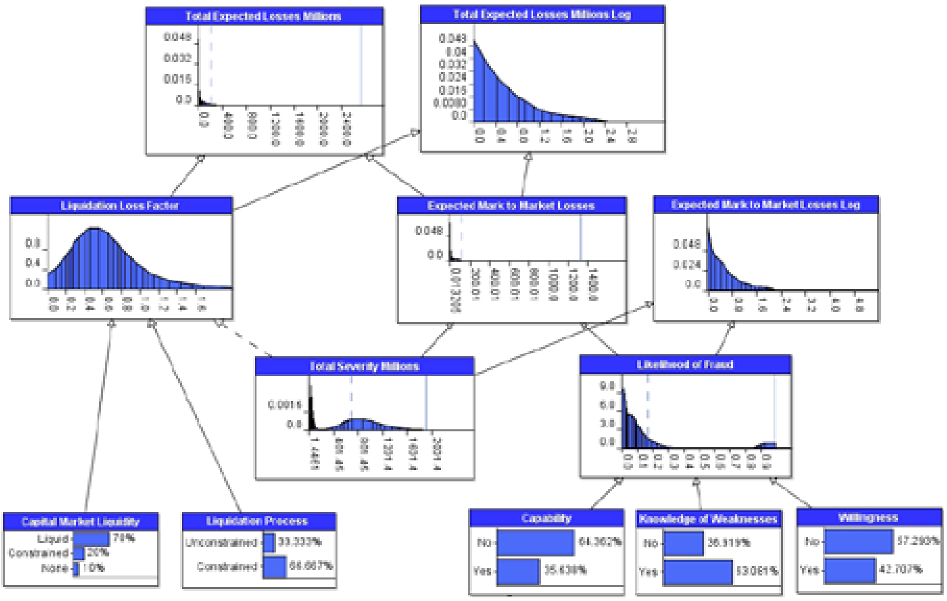
Note the high degree of nonlinearity in the exposure levels, which drive highly nonlinear severity estimates. Also note that the mode of loss severity is very low, but with a significant right-hand tail.

11.3.6 Expected loss

Severity is also potentially impacted by the liquidation process. If this is constrained, in the sense that exposures need to be unwound immediately regardless of capital market conditions, then losses can be significantly greater than if flexibility exists to be able to manage the process. This is because trade volumes can be significant proportions of market turnover. Furthermore, if liquidity in the capital markets is constrained, then losses can also be significantly higher.

Figure 62 shows the derivation of expected losses, based upon the likelihood, severity, and liquidation process. Note that the expected loss curve is highly skewed, and we have shown this in log form to show the distribution in greater detail.

FIGURE 62: BAYESIAN NETWORK FOR EXPECTED LOSSES



From this it is possible to derive the risk measures such as a mean loss of $200 million, whilst the 99.5th percentile loss is $2.7 billion. This framework thus is able to model a full range of modelled outcomes as a complex function of risk drivers of the business.

12 REFERENCES

APRA (2008). *Prudential Standard APS 115 – Capital Adequacy: Advanced Measurement Approaches to Operational Risk*. Australian Prudential Regulation Authority.

APRA (2012). *Prudential Standard SPS 114 – Operational Risk Financial Requirement*. Australian Prudential Regulation Authority.

APRA (2013) *Prudential Standard LPS 118 – Capital Adequacy: Operational Risk Charge*. Australian Prudential Regulation Authority.

Bank of International Settlements (June 2011). *Operational Risk – Supervisory Guidelines for the Advanced Measurement Approach*.

Bank of International Settlements (June 2011). Principles for the Sound Management of Operational Risk.

Basel Committee on Banking Supervision (September 2001). Working Paper on the Regulatory Treatment of Operational Risk.

Cantle, N., Orros, G., Puchy, R. & Wang, H. (2010). ERM for Strategic and Emerging Risks. The Actuarial Profession Risk and Investment Conference 2010.

Chapelle, A. (2012) *The Paths to Rogue Trading*. University of Brussels.

Chapelle, A., Crama, Y., Hubner, G. & Peters, J.P. (2008). Practical methods for measuring and managing operational risk in the financial sector: A clinical study. *Journal of Banking and Finance* Volume 32, pp 1049-1061.

Coleman, T. (2011). *A Practical Guide to Risk Management*. Research Foundation of the CFA Institute.

Cope, E. & Antonini, G. (2008). *Observed Correlations and Dependencies Among Operational Losses in the ORX Consortium Database*. IBM Zurich Research Lab and ORX Analytics Working Group.

Cope, E. & Labbi, A. (2008). *Operational Loss Scaling by Exposure Indicators: Evidence from the ORX Database*. IBM Zurich Research Lab and ORX Analytics Working Group.

Cope, E., Antonini, G., Mignola, G. & Ugoccioni, R. (2009). *Challenges in Measuring Operational Risk from Loss Data*.

Cornaglia, Mignola & Morone (2007). *Economic Capital Assessment via copulas: Aggregation and allocation of different risk types*. Intesa-Sanpaolo.

Corrigan, J., Graham, S., Hague, A., Highham, A., Holt, J. & Mowbray, P. (2011). *Transforming Consumer Information*. Consumer Information Working Party, Institute and Faculty of Actuaries.

Delasey, M. & Sakalo, T. (2011). A framework for uncertainty modelling in operational risk. *Journal of Operational Risk* Vol 6, no 4.

Dutta, K. & Perry, J. (2007). A Tale of Tails: An Empirical Analysis of Loss Distribution Models for Estimating Operational Risk Capital. Federal Reserve Bank of Boston Working Paper No 06-13.

Embrechts, P., Klüppelberg, C. & Mikosch, T. (1997). *Modelling Extremal Events: for Insurance and Finance*. Berlin.

Evans, J., Womersley, R., Wong, D. & Woodbury, G. (2007) Operational risks in banks. *The Finsia Journal of Applied Finance* Issue 2, pp9-16.

Hoffman, I., Jenkins, T. & Wilkinson, P. (2011). *A Systems Dynamics Analysis of the Barrier-Based Systems for Preventing Major Operational Events*.

Institute of Operational Risk (2010). *Operational Risk Sound Practice Guidance: Key Risk Indicators*.

Institute of Operational Risk (2010). *Operational Risk Sound Practice Guidance: Operational Risk Governance*.

Klugman, S., Panjer, H., & Willmot, G. (2008) *Loss models: from data to decisions*, 3rd edition.

Levy, A. & Watchorn, E. (2008). *Information Paper – Developing Business Environment and Internal Control Factors for Operational Risk Measurement and Management*. Australian Prudential Regulation Authority.

Moscadelli, M. (2004). *The modelling of operational risk: experience with the analysis of the data collected by the Basel Committee*. In Temi di discussione (economic working papers), Bank of Italy, Economic Research Department.

McNeil, A. J. & Saladin, T. (1997). *The peaks over thresholds method for estimating high quantiles of loss distributions*. Proceedings of 28th International ASTIN Colloquium.

ORX (2010). *ORX Operational Risk Report*.

ORX (2009). *Operational Riskdata eXchange database*.

ORX (2011). *Operational Risk Reporting Standards*, Edition 2011.

Milliman

Joshua Corrigan
joshua.corrigan@milliman.com

Paola Luraschi
paola.luraschi@milliman.com

milliman.com