ISCTE ◈ Business School
Instituto Universitário de Lisboa

OpRisk: The challenges of Basel Advanced Approach

Carla Sofia Santos Carloto

A Dissertation presented in partial fulfillment of the requirements for the Degree of

Master in Finance

Supervisor
Prof. Miguel Ferreira, Prof. Associado,
ISCTE Business School, Finance Department

Setembro 2009

Abstract

Operational Risk is defined by Basel Committee as "the risk of loss resulting from inadequate or failed internal processes, people and systems or from external events."

Since the beginning, all institutions know that operational risk is present in their activities, but just when Basel Committee introduced as mandatory to have regulatory capital requirements, institutions change their focus from Credit Risk and Market Risk to manage operational risk as a way to reduce regulatory capital.

To present some alternative models to be support Advanced Approach, I investigate possible approaches and available statistical distributions that better explain factors/variables like operational risk losses using public data.

From the reports published by ORX and FED, I simulate capital requirements using different distributions for each Business Line and Event type, and compare final results and behaviors.

The important conclusion of this paper is that is critical to consider all four elements to build a soundness model to estimate capital needs with internal models. The model should be suitable for the reality of institutions rather than be evaluated as the best in statistical measures. More than be regulatory requirement these internal models should be considered an important tool for risk management.

Resumo

Risco Operacional é definido pelo Comité de Basileia como o "Risco de perdas em resultado da inadequação ou falha de processos internos, pessoas, sistemas ou eventos externos".

Desde sempre que as Instituições tem conhecimento da existência de Risco Operacional, mas apenas quando o Comité introduziu como requisito obrigatório no cálculo de capital regulamentar, que as instituições alteraram o enfoque da sua gestão de risco do Risco de Crédito e de Mercado para a gestão do Risco Operacional para optimizarem o capital regulamentar.

Para apresentar modelos alternativos de suporte à Abordagem Avançada do ponto de vista quantitativo, investiguei possíveis abordagens e distribuições estatísticas que melhor explicassem os acontecimentos de risco operacional recorrendo a dados públicos. Dos relatórios publicados pela ORX e FED, simulei os requisitos de capitais por cada Linha de Negócio e Tipo de Evento recorrendo a diferentes distribuições e comparando os resultados finais, assim como, o comportamento das mesmas.

A conclusão importante deste estudo é que é crucial considerar os quatro elementos para a construção de um modelo interno robusto para estimar as necessidades de capital. O modelo deve reflectir a realidade das instituições e não apenas obter melhores medidas estatísticas em relação à sua qualidade. Mais do que um requisito regulamentar, os modelos internos devem ser considerados uma importante ferramenta para a gestão de risco nas instituições.

Acknowledgements

I dedicate this study to my family and friends for their support and encouragement.

I would like to thank to my supervisory on this study, Professor Miguel Ferreira for his superior orientation on the work presented

INDEX

1. Introduction

In a time when once again the financial system is going through a crisis, is also the time of Basel II rules are beginning to be mandatory in several financial systems. Operational risk is the newest requirement introduced by the Committee with Basel II Accord and has become a fundamental part of every institution risks management strategy. Basel document by itself has no mandatory enforcement, so it was necessary for European Union to transpose and make minor changes and publish CRD (Capital Requirements Directive) in October 2005.

2008 is year zero for institutions to start to use this new framework, and Operational Risk appears to be the major challenge due to lack of historical sound data and experience in building models the same way that institutions have been doing in the past years to Credit and Market Risk. These models were used only for internal management proposals and now institutions can use them to calculate regulatory capital.

The Committee define major principals for operational risk management framework and advice supervisors how to validate internal models, but has not yet been sensitivity to define values and detail approaches.

The new framework presents several available approaches which allow institutions to adopt the one that best fits their risk profile and risk management strategy. The main objective of three available approaches is to provide incentives for institutions to improve their risk-management practices, with more risk-sensitive risk weightings as institutions adopt more sophisticated approaches to operational risk management.

For institutions the first challenge is to decide their own definition of operational risk and their scope with the question of including legal and reputational risk. They have to plan their strategy for risk management by choosing which approach to adopt and the next steps to implement a more complex framework, allow reducing regulatory capital allocated to operational risk as management process becomes more effective and sound.

The available approaches for operational risk are Basic Indicator (BIA), Standard Approach (STA) and Advanced Measurement Approach (AMA), and in each one are defined quantitative and qualitative requirements that institutions must be compliant with.

With respect to the AMA Approach, the major challenge is sound information with minimum historical of three years. To build a sound internal model is necessary information to allow shaping the better statistical distribution that explains the events and will permit to estimate losses with 99,90% of confidence level. Institutions must implement the loss data collection process and guarantee completeness and timeliness of this process.

This paper explores theoretical models available from other risk management models like Actuarial model for insurance risk provisions. I apply one approach to aggregate loss data collected from public reports. I simulate and build Aggregated Loss Distribution for each business line and event type and estimate Capital at Risk according to distributions results.

After introducing the regulatory principles, this paper presents measurement models for operational risk quantification. An operational risk model should include Internal Data, External Data, Scenario Analysis and Business Environment. Institutions must define their framework for each part of the model and how to integrate the different techniques used.

For the loss data collect, institutions should use a technique that best fits the available loss data. The most common techniques are: Empirical Loss distributions, Parametric Loss Distributions, Actuarial Models and Extreme Value Theory.

Empirical Loss Distributions and Parametric Loss Distributions are Total Loss Techniques meaning that they do not separate severity from frequency. In contrast, Actuarial model is a technique of Aggregated Loss and considers severity and frequency as two separated input factors.

Historically the major operational risk events are characterized by higher severity and low frequency and for statistical modeling of these events, the Extreme Value Theory is very useful. This technique allows modeling the right tail of an Aggregated Loss Distribution without compromise the distribution's principal modeling process.

After choosing the internal data treatment methodology, the institutions must use external data to adjust their estimates. External data can be used only as quality data or be integrated in the modeling process. When used is necessary to make adjustments through data treatment and define which technique to use to integrate both data in the same model.

Adopting the Advanced Approach implies the use of business factors on model calibration. This calibration can be made indirectly through scenario analysis or directly recurring to direct calibration with Q factor build using score functions.

As mentioned before, data integration is very critical and I identify several methodologies to be applied to the process of data integration and Bayesian Integration or Integration by convolution are the ones that best fits the needs of data modeling for Operational Risk.

Other possibility is to integrate output data instead of input data applying either weighting average Capital at Risk estimated in each intermediate steps (e.g. LDA, External data) or considering the first output and then add next results using adjustments factors.

The final step is this framework is to build a matrix of estimated Capital at Risk for each business line and event type, and once again there are available several techniques assuming that losses are independent.

In this paper I apply Loss Data Aggregation methodology to published data from FED and ORX, and it is possible to generate as output Capital at Risk for each business line and event type, and compare results between studies, statistical distributions and input data presented with detail. The methodology and results are presented in sections 3 and 4 of the paper, detailed data, indicators and graphs are presented in the appendix.

It is not possible to compare the three approaches (BIA, STA and AMA) between themselves because it is not available about participants and their historical individual losses, so this paper focuses in how to build the quantitative framework for Advanced Measurement Approach for Operational Risk.

At the same time as this study was prepared, Basel Committee published the 2008 LDCE for Operational Risk, where is presented the set of methodologies implemented by institutions. When comparing both documents is understandable that the methodologies proposed and implemented are similar and also some institutions use the same statistical distributions building models that better fit their reality and their historical data.

The diversity of results obtained demonstrate that soundness information and their completeness is the challenge for all institutions, for building a strong internal model of Operational Risk that aims not just risk management but also helps institutions to mitigate risks and incentive them to implement advanced approaches as a path to adjust regulatory capital to their own risk profile.

The main conclusion from the results is that is necessary to consider all four elements [Data (Internal and external), scenario analysis and business environment] to build a soundness internal and the critical factor of success is the quality of data collected during the previous years.

2. Theoretical framework

2.1 Basel II and Committee's recommendations for Operational Risk

Operational risk is introduced by Basel II as mandatory requirement for capital adequacy of financial institutions and this framework has been adopted by almost all supervisors. Basel II is a challenge for banks and their supervisors, because of the advanced approach to calculate minimum capital requirements.

Operational risk is defined in the document "as the risk of loss resulting from inadequate or failed internal processes, people and systems or from external events. This definition includes legal risk, but excludes strategic and reputational risk." Notice that "legal risk includes, but is not limited to, exposure to fines, penalties, or punitive damages resulting from supervisory actions, as well as private settlements."

To help institutions and their supervisors, BIS also publish a document with recommendations about operational risk management: "Sound Practices for the Management and Supervision of Operational Risk". In this document there are ten best practices to orient institutions to create their own framework and to help supervisors to evaluate presented methodologies.

The document presents a framework with three possible approaches to calculate capital charge and to manage operational risk according with calculation methodology chosen.

Operational risk framework defines the following approaches:
1) Basic Indicator Approach;
2) Standardized Approach (also available "Alternative Standardized Approach");
3) Advanced Measurement Approaches (AMA).

There are two major differences between approaches related with calculations method and mandatory quality requirements associated to those methods. As the complexity of calculations increases the management requirements also increase because internal data becomes fundamental to the framework.

2.1.1 BASIC INDICATOR (BIA)

In the first approach the capital charge is 15% of annual gross income (average of last three years with positive income) and recommends institutions to comply with "Sound Practices for the Management and Supervision of Operational Risk".

The capital requirements are given by:

$$K_{BIA} = \frac{[\Sigma(GI_{1\ldots n}) \times \alpha]}{n}$$

where:

K_{BIA} = capital charge under the Basic Indicator Approach for Operational Risk;

GI = annual gross income of the last three years (if all positive);

N= 3 if all previous years the gross income is positive;

α = 15% which is set by the Committee.

As Gross Income could have several interpretations, the Committee defines GI as the sum of net interest income and net non-interest income. This measure should be gross of any provisions and gross of operating expenses including those related to outsourcing service providers. Any profit or loss related to: sale of securities in the banking book, extraordinary items and related to insurance should be excluded from this measure.

2.1.2 STANDARDIZED APPROACH

The second approach allows institutions to calculate capital charge according with their business lines, because the Committee defines different charge for each business line. The gross income is used as an indicator to determine the scale of business activity and therefore an estimative of operational risk exposure.

The first step is to segment institutions gross income into eight business lines:
1) Corporate Finance;
2) Trading and Sales;

3) Retail Banking;

4) Commercial Banking;

5) Payments and Settlements;

6) Agency Services;

7) Asset Management;

8) Retail Brokerage.

The capital charge is calculated as follows:

$$K_{TSA} = \frac{\left\{\sum_{years\ 1-3} \max\left[\sum\left[(GI_{1-8}] \times \beta_{1-8}\right); 0\right]\right\}}{3}$$

where:

K_{TSA} = capital charge under the Standardized Approach for Operational Risk;

GI_{1-8} = annual gross income of the last three years (as defined above) for each business line;

β_{1-8} = a fixed percentage, set by the Committee, relating the level of required capital for each business line. The values of Betas are:

Business Lines	Beta Factors
Corporate finance (β_1)	18%
Trading and sales (β_2)	18%
Retail banking (β_3)	12%
Commercial banking (β_4)	15%
Payment and settlement (β_5)	18%
Agency services (β_6)	15%
Asset management (β_7)	12%
Retail brokerage (β_8)	12%

Figure 1: Beta coefficients (Standardized Approach)

The proposal values for Beta's is meant to be used as an estimated factor supported in industry-wide operational risk loss experience and the aggregated level of gross income in each business line.

Note that the Committee intends to review the calibration of Betas (and Alpha for Basic Indicator) when more risk data are available to support the calibration process.

The Committee also refers that national supervisors can authorize institutions to adopt Alternative Standardized Approach. The calculations are equal to what is present expected to Retail Banking and Commercial Banking. For these two business lines the formula can be:

$$K_{RB} = \beta_{RB} \times m \times LA_{RB}$$

where:

K_{RB} = capital charge under the Standardized Approach for Retail Banking;

β_{RB} = Beta for Retail Banking business line;

LA_{RB} = last three years average of total outstanding retail loans and advances with any adjustment provided by risk weight or by provisions. In the full document is defined what can be considered to build this indicator.

m = 0,035 (value defined by the Committee)

For a risk management point of view, the Committee is more demanding in this approach and Sound Practices became mandatory to be implemented and supervisors must be more meticulous on acceptance process of institutions choices.

The Committee is clear when defining as mandatory an implementation of a governance model for operational risk management that must include the following topics:
- Top management involvement;
- Organizational structure and Operational Risk Management Processes ;
- Operational Risk Management Policy;
- Definitions and glossary for Operational Risk;
- Criteria for Mapping the Gross Income into eight business lines;
- Incentive systems to sound risk management.

This governance model should guarantee that Operational Risk Management function includes the definition of identification, assessment, monitoring, control and mitigation of operational risks.

Because more advanced methodologies depend on information collected by institutions and their historical data, this approach introduces the process of risk identification and measurement usual known as Loss Data Collection and uses important tools like Scenario Analysis and periodic Risk Self Assessments.

For the Committee and local supervisors is fundamental reporting about real operational risk losses, which in a first level should occur internally to Top Management and business units (if applicable) and then externally to supervisor.

The reporting process should comply with the following principles:
- *Completeness*: comprehensive set of collected losses.
- *Timeliness*: time proximity of the recorded information to the loss event.
- *Accessibility*: cost of getting the information.
- *Quality and quantity of information collected*
    - Date, nature and amount of loss, etc.
    - Exposure information.
    - Rating information.

The Risk Assessment process must guarantee that all relevant operational risk data like material losses occurred are included in assessments and this tool is important to the management process and for the definition of the institution risk profile.

All processes referred before must be documented and be well known by all institution's employees and be reviewed and validated by independent department (internal audit). Even with an internal review process implemented the operational risk management process must be subject to regular review by external auditors and/or supervisors.

2.1.3 ADVANCED MEASUREMENT APPROACHES (AMA)

Under this approach the regulatory capital requirement is calculated by their internal operational risk management system, which needs approval by national supervisor.

The requirements in this approach include all mentioned for the Standard Approach, and add some additional qualitative and quantitative criteria.

The most important is that when an institution adopts a more advanced approach it can not go back if the regulatory capital requirements are bigger than if using Standard approach.

For adopting this approach is necessary for a period of time to be under supervisor monitoring, which will allow the supervisor to determine if the internal approach is appropriated and credible. The most important is that the internal measurement system estimates reasonably well operational risk unexpected losses based on the combined use of data about internal losses calibrated with relevant external losses, and use of scenario analysis adjusted with information from business environment and internal control factors.

The system should be capable of allocating economic capital for operational risk in each business line to be an incentive to improve individual operational risk management of business lines.

In terms of quantitative standards the Committee has not specified the approach or any distributional assumption for the measurement system. But the institution must demonstrate to the local supervisor that the internal model captures severe "tail" loss events and be comparable to one year holding period and a 99.9^{th} percentile confidence interval like it is happens in internal ratings based approach for credit risk.

In the development of an internal measurement system, institutions must have and maintain rigorous procedures that will allow an independent validation. The Committee will review industry practices and the levels of capital requirements estimated by AMA and redefine its proposals if appropriate.

The document presents a series of quantitative standards:
- The internal risk measurement must be consistent with Operational risk definition and the loss events types defined in the document;
- An institution should calculate its regulatory capital requirements as the sum of expected loss and unexpected loss;

- The internal system must be sufficiently granular to capture the drivers that affect the shape of the tail of the loss estimates.

- Risk measures for different operational risk must be added for determining the regulatory capital, but institutions can use correlations determined internally, if the methodology used is sound, implemented with integrity and approved by national supervisor.

- Must have certain features like the use of internal data, relevant external data, scenario analysis, business environment and internal control systems.

- All system must be well documented, must be credible and transparent. All features mentioned before should be weighted in the system and this must be verifiable by independent entities and supervisor.

The Basel document in following sections presents the principles for each key feature to orient institutions and supervisors for what is expected to be an internal system to measure operational risk.

For capital requirement it is allowed to use risk mitigation like insurances. The Committee has limited to 20% of total capital charge calculated the use of insurances and institutions must be complying with several criteria presented.

In practice it allows institutions to adopt partially Advanced Approach to some parts of its operations and a less advanced approach for the balance sheet if the conditions defined in the document are met.

2.2 Measurement Models for Operational Risk

The document published by the Committee only presents general principles for building a sound system for measuring operational risk and in the following sections are available methodologies to accomplish the several features necessary to have an internal system.

2.2.1 INTERNAL DATA

When an institution decides to adopt the AMA approach for the treatment of operational risk, it is need to be compliant by developing their own model based on internal and external data, developing scenarios analysis and integrating information from business environment and internal control.

If institutions collect 3 to 5 years of historical data about operational risk events, they will be able to use a methodology denominated "Loss Distribution Approach". This approach assumes that using the data it is possible to build 2 distributions:
1. Frequency distribution, and
2. Severity distribution.

The next step is to aggregate both distributions using Monte Carlo simulation resulting in one distribution, the Aggregated Loss Distribution (ALD).

This distribution permits to use several statistical techniques to estimate operational risk losses, and estimate the level of capital needed to be compliant for operational risk.

The most common techniques used in this methodology are:
1. Empirical Loss distributions;
2. Parametric Loss Distributions;
3. Actuarial Models; and
4. Extreme Value Theory.

Typically using this approach the results are more reliable from a mathematical point of view. But this affirmation can be questioned, being only valid when the data used permits. It is necessary that we are aware that the event collection is new and that

majority of the institutions are having problems in collecting data needed for the modeling process. This is a newly process and the historical data collect can be insufficient or their quality can be questionable.

EMPIRICAL LOSS DISTRIBUTIONS

The quantitative analysis of Operational Risk can be done using empirical loss distributions, known as empirical simulation technique. An important reminder is that this technique do distinguishes between frequency and severity loss distribution for loss events (Total Loss Approach).

This technique consists in collecting internal data of losses occurred during a period of time, sorted by ascending order of loss amounts and estimating VaR for a particular confidence level; using the Basel Regulation the confidence level is 99, 90%.

According to the Basel Rules, this technique requires:

1) The collection of 5 years of data of losses related with operational risk events;

2) Analyze the data to determine loss distribution, meaning calculating the frequency of each loss level. For improving data quality it is possible if relevant to remove inflation effect or some seasonal effects present in data.

3) The final step is to identify the loss distribution's quantile and corresponding loss value. This quantile is defined by the confidence level, and the amount is the estimative for "Capital at Risk" calculated only with historical data.

In short the most important advantages and disadvantages of this technique are:

Advantages	Disadvantages
Intuitive Method	All estimations are made assuming that historical losses are a good indicator for predicting future losses, meaning that the risk management process keeps equals, and this is difficult to guarantee.
Permits calculations without considering the statistical distribution of the underlying phenomenon, the loss events. This simplifies calculation and reduces time to produce results.	This technique assumes that future losses will have the same behavior of past losses.
	It is necessary a large set of data with a certain number of losses.

PARAMETRIC LOSS DISTRIBUTIONS

In contrast to the empirical loss distribution, this methodology does not uses historical data to estimate operational risk losses. Historical data only provides the theoretical distribution parameters that best fit the underlying phenomenon.

This methodology also is a technique of Total Loss.

If for Market Risk is common to assume returns normality, for Operational Risk that assumption is not valid because of the low frequency of events, the large number of events with a low loss amount, and the reduced number of events with a high loss amount. Given these facts the distribution chosen should be an asymmetric because it fits better this kind of data.

In this methodology the necessary steps are:

1) Data collection and treatment for removing external effects mentioned above. It should be created a set of theoretical distributions, from which is chosen the distribution that best support the modeling process. All distributions considered in this stage are asymmetric and with slight right tail (e.g. Log-Normal, Weibull, Gamma)

2) Next step is to identify a distribution that best fits losses behavior. One of the most popular methods is the quant-quantile graph that using data analysis compares sample's quantiles with distribution's quantiles.

3) For selecting a distribution, non-parametric tests can be used only for the sub-set selected in previous step.

   The test is:

$$H_0: \quad x_i \cap F_0(x), \qquad F_0(x) = Theoretical\ Distribution$$

This test is expressed in terms of distance between the data distribution function and theoretical distribution.

4) And after selecting a distribution is necessary to estimate their parameters using the data collected in step 1. These parameters can be estimated using maximum likelihood method or moment's method. With a full specification of distribution is possible to determine Capital at Risk (level of confidence of 99, 9%)

In summary the most important advantages and disadvantages of this technique are:

Advantages	Disadvantages
Quantiles can be estimated using a minor data sample.	Can be difficult to select only one theoretical distribution for data adjustments.
Maximum Likelihood Estimators have important properties like has lower mean squared error.	Trying to obtain an adjustment for total data can cause bad estimations of right tail quantile (suggestion: Extreme Value Theory).
The data distribution is determinate statically and not only supported by exploratory data analysis.	In practice maximum likelihood estimators are difficult to calculate.

ACTUARIAL MODELS

For Operational Risk estimation is possible to apply actuarial models developed in the Insurance Industry. These kinds of models are built using two key factors/elements:

1) Number of Events – Frequency;
2) Losses Amount – Severity.

This methodology can be described in the following steps/stages:

1) Data collection and treatment, like in previous techniques but now is necessary to replicate data in 2 basis:

    a. Frequency Table: Frequency along a time period;

    b. Severity Table: Presented by ascending order of amounts.

2) Next step is selecting an adequate distribution for each phenomenon modeling. The statistical methods used are equal/ similar to what have been presented early. For Frequency is common use Poisson Distribution and for Severity is more typical choose Distribution like LogNormal, Pareto, Weibull or Gamma.

3) Build an Aggregated Loss Distribution considered aggregated loss a stochastic process $S(t)$, therefore $S(t) = \sum_{i=1}^{K(t)} W_i$, $0 \leq t \leq T$.

The aggregated loss distribution function can be calculated through a convolution operator using both distribution functions to produce a joint distribution function. In practice this operator is too complicated to implement and in alternative, institutions adopt an easier methodology based on Monte Carlo Simulations, ie:

$$S(t) = \sum_{i=1}^{K(t)} W_i, 0 \leq t \leq T \Rightarrow (S|K = k) = \sum_{i=1}^{K} W_i \quad (1)$$

For example, considering that:

- Frequency $= K$ a random variable $\sim$ Poisson Distribution

- Severity $= W$ a random variable $\sim$ LogNormal Distribution

The following procedures are executed:

1) Random extraction of x numbers from Poisson

2) Random extraction of x random numbers from LogNormal

3) Repeating this process R times (R equal a number of Monte Carlo Simulations made).

The final result consists in a data sample with R dimension from where is possible to extract the adequate quantile value and determine CaR.

4) The last step consists in estimate ALD parameters, in case there is some evidence that the data are suppressed or truncated. It is possible that loss data does not exist or are uncompleted if the loss amount is lower than a certain value. For these situations, maximum likelihood estimators calculated without any adjustment can overestimate Capital at Risk.

One available alternative to solve this problem is the EM Algorithm (Dempser et al – 1977[1] and McLachlan and Krishnan - 1996[2]). EM Algorithm consists in an interactive method that calculates maximum likelihood estimators using uncompleted data.

This method creates a hypothetic likelihood function, build on base of replacement of omitted values by values that are expected to follow the initial chosen distribution.

In summary the most important advantages and disadvantages of this technique are:

Advantages	Disadvantages
Based on math theories already proved that allow separated analysis of Frequency and Severity	Can be complicated to determine only one theoretical distribution for modeling the collected data
For identify Frequency and Severity Distributions are used Maximum Likelihood estimators	Assuming that loss amounts are identical distributed and independent from the number of loss events can be limiting,

Advantages	Disadvantages
	and is possible to produce less reliable results
Quantiles' calculation is straight forward after estimating ALD parameters.	Trying to obtain adjustments for all data can cause a bad estimation for extreme values
	The methods that use simulations can lead to not reliable results, mainly about quantiles calculations for aggregated loss distribution (this problem can be minimized by executing more simulations)

EXTREME VALUES THEORY

It is common to find extreme events when measuring any type of financial risk. This methodology has the objective to create models that capture situations whit low frequency but with high impact. This theory (EVT) allows us modeling the right tail of ALD without compromise the modeling of the distribution's principal. By modeling a full distribution we incur in the risk of not modeling correctly not only the frequently events with lower impact but also the rare events with higher impact.

For Operational Risk Framework this method is extremely helpful as extreme losses are rare but with very large impact (the ones that have higher impact in business).

This method is implemented in three steps:
1) Data Collection: there are 3 major factors in this stage:
    a. Time frame: It should be as long as possible for guarantee a major set of data;
    b. Granularity level: It will depend on time frame scale and initial loss detail available;
    c. Quality Data Analysis: Using available tools to evaluate their quality recurring to exploratory data analysis methods.
2) Next step consist in selecting a data sample with extreme values from the initial loss data set. For this it is available 2 methods: Block Maxima and Peaks-over-Threshold (POT).

Consider $X_1, X_2, ..., X_n$ as a set of random variables with the following distribution functions $F_{X_i} = P\{X_i \leq x\}$. Each variable represents an operational loss in a specific time period.

The extreme values of the distribution right tail are those who verify the following condition: $M_n = \max\{(X_1), ..., X_n\}$. This definition is according Block Maxima method, in which observations are divided by several categories being selected the maximum value as the extreme values of each one. With the selected observations we build a table to study according with asymptotic theory.

The alternative methodology "Peaks-over-Threshold" consist on previous definition of cut point for loss amount, and all values higher are considered extreme values (excess).

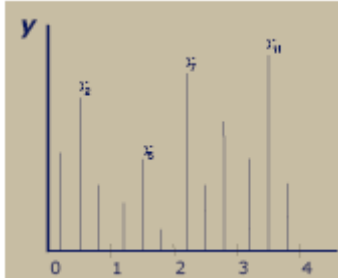Graphically, the selected processes of samples with extreme values are:
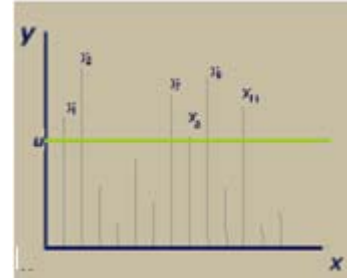

Figure 2: Block Maxima method   Figure 3: Peaks-over-Threshold

3) The following step consists in defining the cutoff point using the POT method. It is necessary to consider a tradeoff between variance and distortion of parameters estimative. A correct method to do this tradeoff evaluation can be graphical analysis of the excess variability for each level of cutoff point.
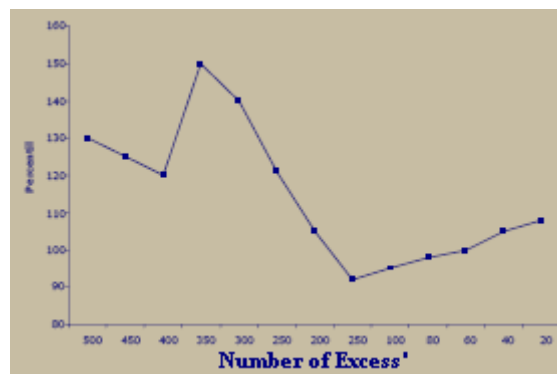
Figure 4: POT – Example of graphical analysis of the excess

4) The final step is estimating parameters for the select distribution for those extreme values and identifies the specific quantile. Once the sample is set the asymptotic distribution of those observations can be one of the following, depending on the method used to build the sample:

a. Generalized Extreme Value Distribution, GEV (Block Maxima Method):

$$y = f(x|\mu, \sigma, \xi) = \left(\frac{1}{\sigma}\right) e^{\left(-\left(1 + k\frac{(x-\mu)}{\sigma}\right)^{-\frac{1}{k}}\right)} \left(1 + k\frac{(x-\mu)}{\sigma}\right)^{-1-\frac{1}{k}}, 1 + k\frac{(x-\mu)}{\sigma} > 0 \quad (2)$$

b. Generalized Pareto Distribution, GPD (Peaks-over-Threshold Method):

$$y = f(x|k, \sigma, \theta) = \left(\frac{1}{\sigma}\right)\left(1 + k\frac{(x-\theta)}{\sigma}\right)^{-1-\frac{1}{k}}, x > \theta \quad (3)$$

In summary the most important advantages and disadvantages of this technique are:

Advantages	Disadvantages
Focus only on estimation of extreme values that are located on tails.	Asymptotic distributions are very sensitive to changes on parameters values.
Asymptotic distributions are available for these extreme cases.	

2.2.2 EXTERNAL DATA

The use of external data is a requirement of the AMA approach in Basel II. This data can be used in two different methods:

- Scenario analysis as quality data (very conservative approach);
- LDA approach as quantitative data.

For now, we are going to present only the second method because the first one will be detailed on a specific section.

Before including external data on the modeling process is necessary to define some elements:

1) Data treatment:

    a. Linear adjustment: a unique coefficient is applied for adjusting data to institution dimension;

    b. No-Linear adjustment: define specific coefficients for each type of event using regressions;

    c. Data Filter: guarantee that only is used data form institutions with similar dimensions;

    d. Not doing any data adjustment and integrate them on the modeling process.

2) How to join internal and external data:

    a. Common usage: Aggregate external data to internal data base , creating a more robust base and apply previous presented methods;

    b. Separated usage: Internal and external data are analyzed separately; creating a specific ALD just for this set of data, and after integrates both using one of following techniques: qualitative aggregation, aggregation throw linear combinatory, Bayesian aggregation or Convolution.

3) If the previous choice is common usage, then define how to treat losses amount:

    a. Aggregation made without consider loss amounts;

    b. Aggregation made using loss amount criteria: internal data are used to model losses below a specific level and extremes for modeling loss above this level.

4) Define external data proportion in the final data base used to estimate CaR.

2.2.3 SCENARIO ANALYSIS

Nowadays one of the major problems that institutions face when try to model Operational Risk is the lack of information, either internal or external. This factor is critical for statistical estimation of event's severity distribution.

According to Basel, institutions are obligated to use scenario analysis to validate or incorporate additional data to their previous results, mainly for extreme events. This is an alternative approach to Extreme Value Theory presented before.

The goal of the scenario analysis is to create fictional events, with the same event's characteristics happened in the past, but due to lack of information are not included in the statistical analysis.

Scenario analysis is an important element for AMA approach in Operational Risk. These scenarios are building using empirical knowledge from institutions experts. This is a new reality for all institutions and is still in development in practical and in theoretical terms, and therefore somehow questionable.

This methodology can be summarized in figure below:
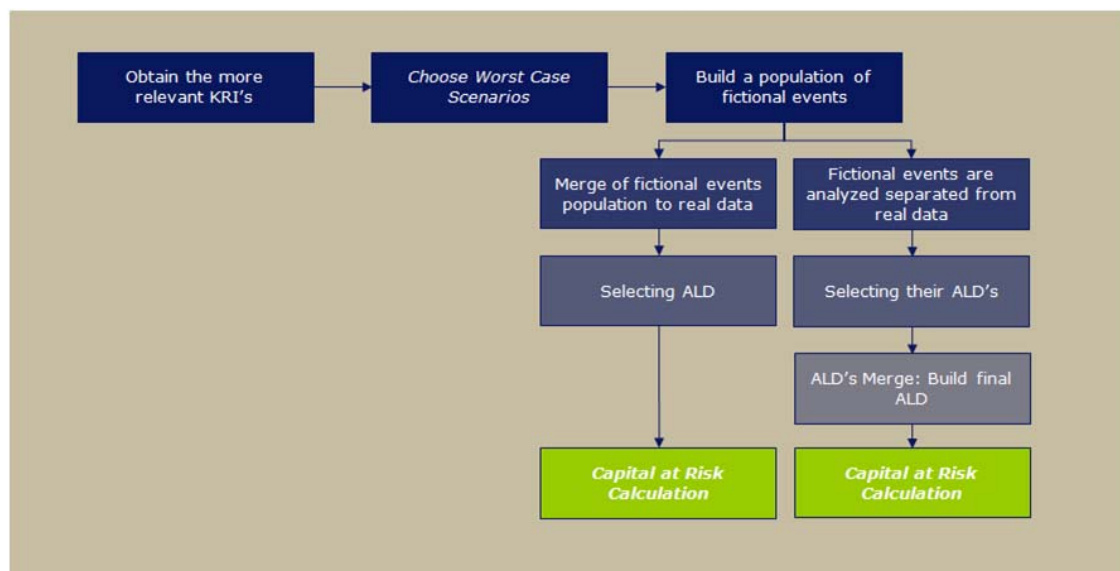
Figure 5: Methodology Diagram

2.2.4 BUSINESS ENVIROMENT

For AMA implementation, a necessary requirement is adjusting results of modeling internal data using institution's risk factors identified as cause of the loss events.

This information can be used in several levels:

- Indirect calibration when implement scenario analysis (methodology presented in above);
- Direct calibration (Q Factor) after aggregating LDA approach with scenario analysis.

For direct calibration, the necessary steps are:

1. Business environment and internal control mechanisms are defined by Top Management with a list of Key Risk Indicators (KRI's);
2. Each KRI is classified considering his impact on Operational Risk using Scorecard;
3. Capital at Risk is obtained in previous stages and adjusted using this new quality information.

This Scorecard is used to evaluate institution's exposure to each KRI. For each exposure level is associated a relative weight. After consolidate all results is built a graph like this:
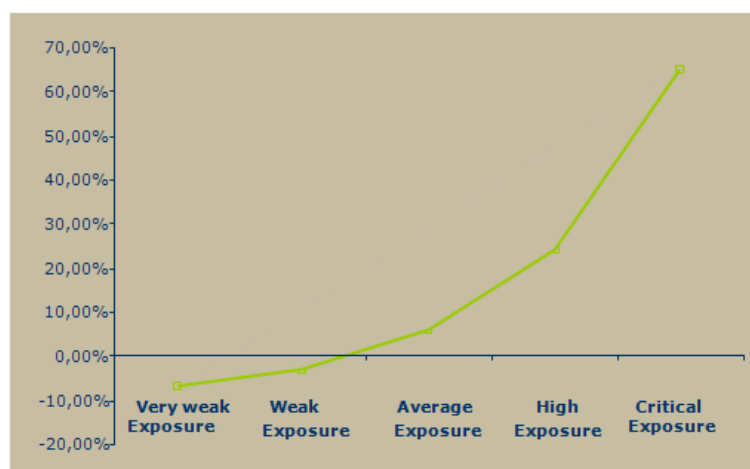
Figure 6: Example of a final graph

For a more conservative approach, weights can be adjusted using an exponential function, which means that factors with critical exposure will have an impact exponentially higher than factors where institutions have a lower exposure.

This coefficient is Q Factor and is used to adjust the CaR obtained before.

This methodology includes the following steps:
- Collect information about the highest number of available KRI's using self-assessments replied by High Management. This information is treated using univariate statistical techniques to determine discriminator capability of variables and to avoid inclusion of KRI correlated (Kruskal- Wallis Test, Discriminator Tree, and correlation analysis, among others).
- In case of collecting a large number of KRI's should be considered the usage of techniques for sample dimension reduction to guarantee only the more relevant KRI's in analysis (Principal components analysis or Factorial Analysis).
- Considered N business lines, and build a Score function for each one using a multiple linear regression:

$$Score_{linha\ de\ negócio\ j} = S_j = \beta_{j0} + \beta_{j1}KRI_{j1} + \beta_{j2}KRI_{j2} + \ldots + \beta_{jt}KRI_{jt} + u_j$$
(4)

Coefficients resulting from regression are each KRI associated weight. The Q-Factor is obtained with the sum up of all Scores in each Business Line:

$$Q\ Factor = \sum_{j=1}^{N} W_j S_j$$
(5)

The previous sum is obtained doing weighted sum considering relative weight of each business line on sample used for analysis.

2.2.5 INTEGRATION TECHNIQUES

- Input Aggregation

It is not possible to apply Aggregated Loss Distribution at all moments of data integration, especially in the input of data analysis.

As shown before, integration of internal and external data can be done by aggregating their distributions using the following techniques:

- Qualitative Aggregation – used generally when external data are also qualitative and heterogeneous (when their presentation is different from internal data format). This technique is based on estimating a parameter using qualitative data that will be used for adjusting quantitative data. The greater advantage is the technique's simplicity, but in the other side is adjusting statistical outputs with parameters estimated based on qualitative data.

- Integration using Linear Combination is applied when both information sources are homogeneous. This technique consist on estimating factors which the sum is equal to one and weight both sources:

$$L(x) = \lambda \, L_1(x) + (1 - \lambda) L_2(x) \quad (6)$$

The main advantage is its linearity and the disadvantage is trying to combine qualitative factors.

- Bayesian Integration is the more valid method from the mathematician point of view, but more difficult to implement. Theoretically is based on Bayes Theorem :

$$P(H|E, c) = \frac{P(H|c) * P(E|H, c)}{P(E|c)} \quad (7)$$

This formula resumes using a priori information on the modeling process. When applied to Operational Risk, external data are a priori information available and using this method these data will adjust internal data results for creating a loss distribution a posteriori:
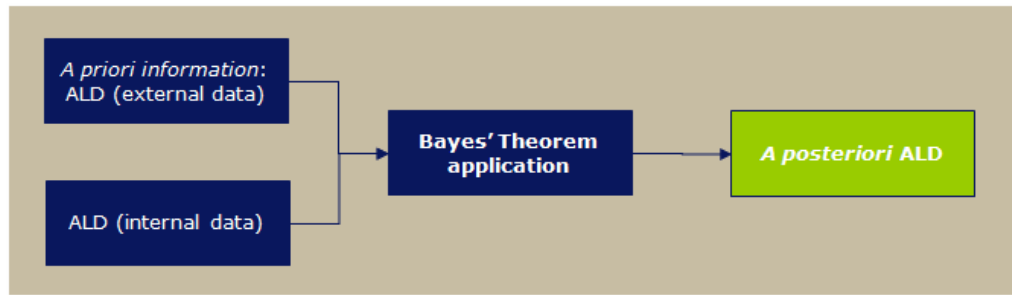
Figure 7: Methodology of Input Aggregation

This methodology can be very important when external data available have identical format of internal data. This method incorporates qualitative data in a statistical methodology, and that became it principal advantage. On other side is the dificult degree of implementing in a real scenario.

Aggregation by convolution was used in actuarial method to create an Aggregated Loss Distribution based on Severity and Frequency Empirical distributions. Because is also complicated to be implemented, in several occasions is replaced by Monte Carlo Simulations.

This methodology can also be applied to two Aggregated Loss Distributions. When decided to use external and internal data, aggregation can be done throw a joint data base. The new data base is more completed and will allow a usage of Loss Distribution approach.

External data (qualitative and/or quantitative) integration provided by scenario analysis or business scorecard can be applied using any of the four methodologies presented.

- Output Aggregation

Aggregating output can be understood as using CaR calculated in the previous stage to adjust the next stage output.

The integration can be applied using two methods:

1. Just considering the previous output as an adjustment factor demonstrated in the next diagram:
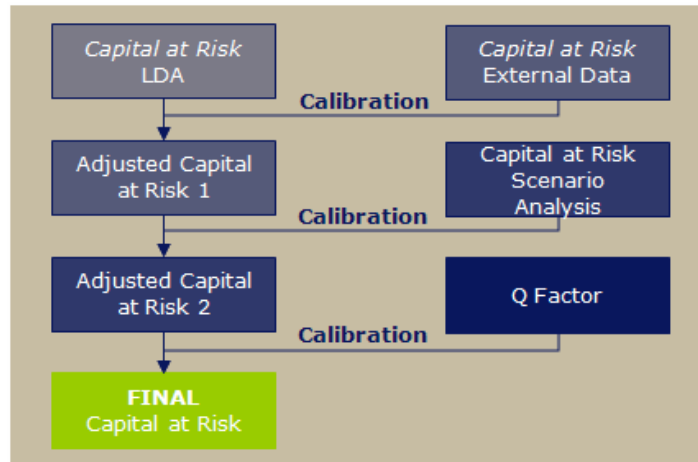
Figure 8: Output Aggregation Framework

2. Incorporating throw a weighted average of CaR represented in the following expression:

$$CaR\ Final = W_{LDA} * CaR_{LDA} + W_{External\ Data} * CaR_{External\ Data} +$$
$$+ W_{Scenario\ Analysis} * CaR_{Scenario\ Analysis} + W_Q * Factor_Q \tag{8}$$

And W_i is the weight attributed to methodology (i)

- Business Line and Event Type Aggregation

All process describe above is made individually by Business Line and Event Type until integration of information provided by Business Environment. Considering (i) Business Lines and (j) Event Type, will obtain $(i \times j)$ values of Car that complete the following AMA matrix:

		Event Type			
		1	2	...	J
Business Line	1	CaR(a;1,1)	CaR(a;1,2)	...	Car(a;1,j)
	...	...	...	...	...
	I	CaR(a;i,1)	CaR(a;i,2)	...	CaR(a;i,1j)

Figure 9: AMA Matrix of CaR

28

To integrate information from Business Environment is necessary previous to aggregate CaR from every Business Line and Event Type in one, which should represent the CaR of the institution. This last step can be accomplished using different methods:

- Sum all Matrix's CaR considering that all have the same weight:

$$CaR(\alpha) = \sum_{i=1}^{i} \sum_{j=1}^{j} CaR\ (\alpha, i, j) \quad (9)$$

- Weight each Business Line CaR based on original Loss Weight of a BL in Total Losses of Institution:

$$CaR(\alpha) = \sum_{i=1}^{i} wi \sum_{j=1}^{j} CaR\ (\alpha, i, j) \quad (10)$$

- Considering each pair Business Line/ Event Type are independent between themselves. The first step is aggregation of Loss Distribution Function by Event Type using Convolution Operator and having as result an aggregated loss distribution (I). The next step is aggregating all I Function in only one ALD using also a convolution operator. The final step is to extract CaR value from the last constructed ALD.

These approaches consider losses as independent phenomenon from BL and Event Type. However, it is possible to assume that loss in a specific business line is not independent from others BL (e.g. losses in Commercial Department and in Products Department). The same assumption can be made to event type's categories (e.g. internal fraud can be related to external fraud). In short, CaR calculated with independency assumption is greater, because considers the same weight to losses occurred in different Business Lines.

A possible solution of this question is the use of Copula Functions, because this type of functions are used to model dependency between 2 or more several marginal distribution functions. In case of linear dependency it is possible to use Gauss' Copula Function.

3. Application Monte Carlo Simulations and Capital at Risk

The methodology adopted can be divided in the following steps:

1. Data Collection from public studies;
2. Choice of distribution based on their own characteristics;
3. Monte Carlo simulation for frequency;
4. Monte Carlo simulation for aggregated losses;
5. Comparing results between distributions.

The first step is to collect public data to test a model for Advanced Approach of Operational Risk.

Since Basel II, Operational Risk has been a subject of many discussions, but there are only 5 important studies with data collections. These studies are:

1) BIS LDCE 2002
2) FED LDCE - 2004
3) Bank of Japan 2006
4) ORX May 2007
5) BIS LDCE 2008 (published only in July of 2009)

For the purpose of our application, we use FED LDCE 2004 (US reality) and ORX 2007 (majority European Institutions).

The first step is to extract data and apply statistical treatment to extract inputs for the selected model.

It is necessary to disaggregate the data by year for each event type and business line. Then I calculate descriptive statistics like average, weight average, standard deviation presented in Appendix 5 and 6 (Tables D and E).

For this type of data we select the Poisson distribution to treat Event Frequency and 5 different asymmetric distributions for Aggregated Losses: Pareto, Beta, Gamma, Weibull and LogNormal.

To build a model it is necessary to have real data of occurred events to study, but it is not possible, so it is necessary to use Monte Carlo Simulation Method to create a population to study. This methodology is used in two different stages: first to simulate frequency using Poisson distribution and then integrate this results and simulating an Aggregated Loss Distribution using also as input Severity's average and Severity's Standard Deviation.

In each stage we made 10.000 simulations and used two different softwares: Crystal Ball for Frequency and Excel for Aggregated Loss, because of Crystal Ball limitations (maximum average <= 1000 and only one step simulations).

To begin the simulation process is necessary to determine inputs for the several distributions. For each distribution is necessary Average and Standard Deviation of Frequency and Severity to calculate parameters that are inputs of each distribution.

For Frequency Distributions we generate a Poisson with λ parameter equal to the annual average of number of events.

And for Aggregated Loss Distribution there are 2 distinct situations:
- For LogNormal, Weibull and Gamma, the theoretical distribution parameters are estimated using simple average and standard deviation of Severity.
- For Beta and Pareto, the parameter estimative does not depend only on average and standard deviation, and is necessary complex calculation with real information not available. So, I assume values for these parameters that from an empirical point of view have proved to adjust to this type of data analysis.

The methodology can be described using the following set of representation:
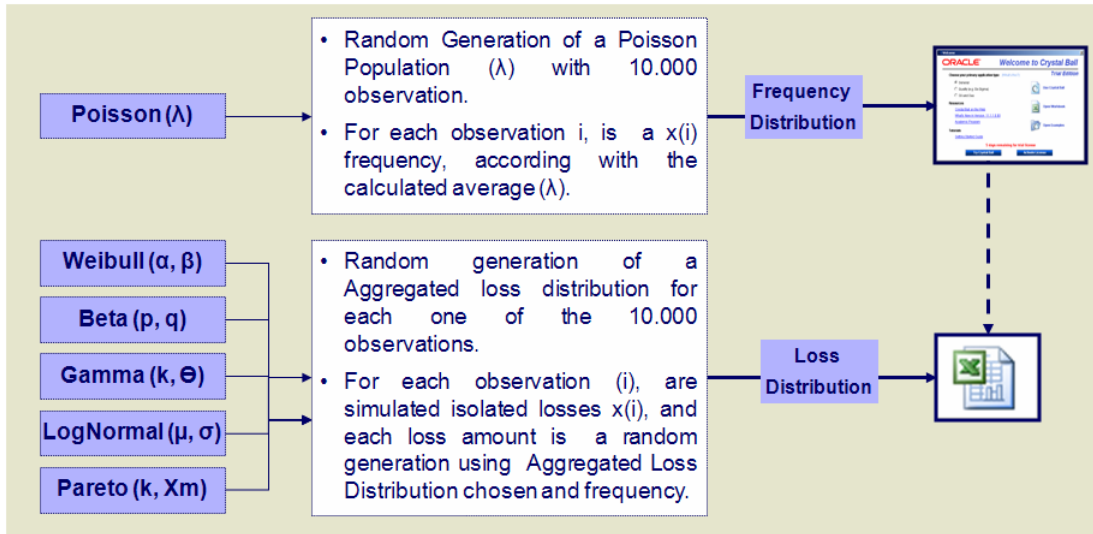
Figure 10: Methodology Framework (1)
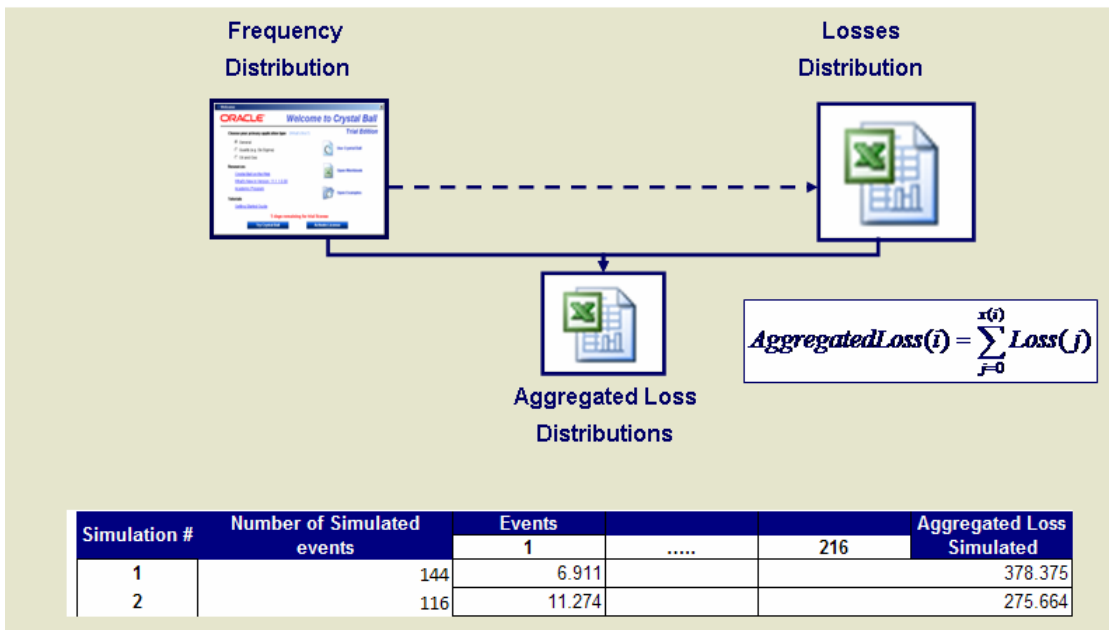
Figure 11: Methodology Framework (2)
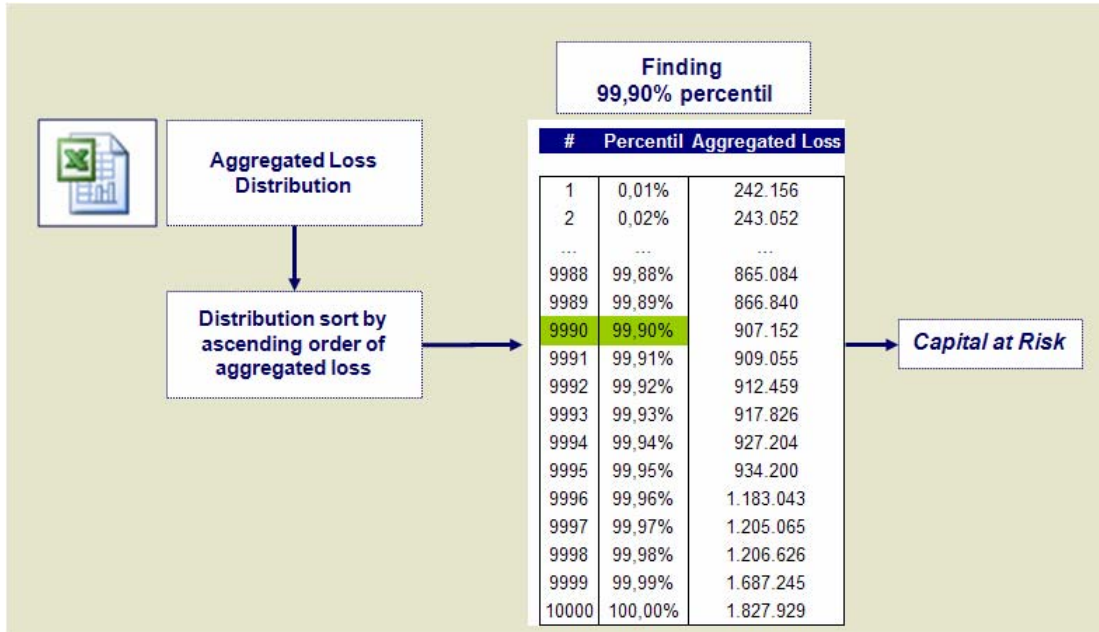
Figure 12: Methodology Framework (3)
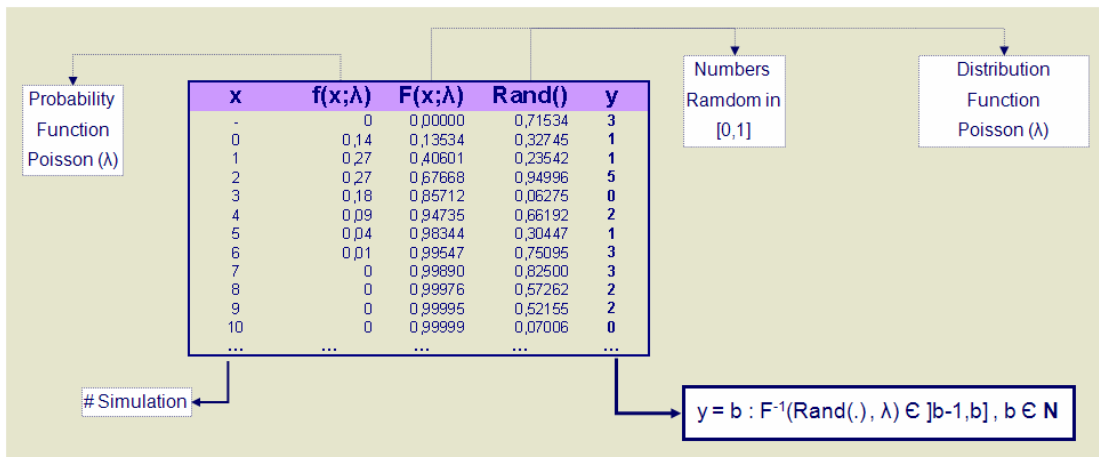
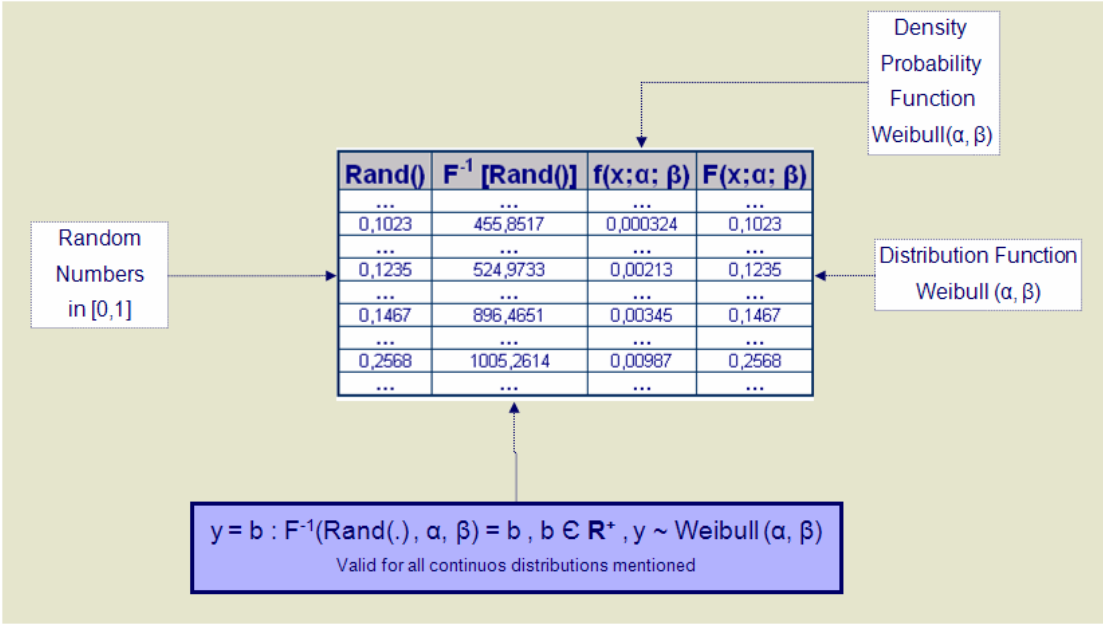Figure 13: Monte Carlo Simulation for Frequency (1)

Figure 14: Monte Carlo Simulation for Aggregated Loss (2)

4. Data Description

4.1. FEDERAL 2004

To evaluate impact of Basel II in minimum capital requirements, FED made a survey called QIS-4 (Quantitative Impact Study – 4) with the voluntary participation of institutions present in the U.S. A total of 27 institutions participated in two exercises proposed, but only 23 reported LDC data, presented in the published document with information summarize.

The collected loss data can be summarized in the following table by threshold and the number of institutions involved:

Loss Data Collection Threshold	Number of Participants	Percentage of Total Losses	Percentage of Total Loss Value
$0 for most or all business lines	6	96.3%	21.8%
Less than or equal to $1,000 (excluding $0)	4	0.8%	1.0%
$5,000 for all business lines	4	1.3%	3.2%
$10,000 for all business lines	6	0.4%	6.2%
More than $10,000 for some or all business lines	3	1.1%	67.8%
Total	23	100.0%	100.0%

Loss Data Collection Thresholds Used by 2004 LDCE Participants

Table 1: Levels of the loss data collection threshold (published in Results of the 2004 Loss Data Collection Exercise for Operational Risk).

From the data presented we select Sample 1 to obtain meaningful results for loss frequency and loss severity. The sample only includes losses greater or equal to $10,000 and occurred during a time of period over which loss frequency appears to be stable. The data is build using three years of historical information related from 2002 to 2004.

This study uses the data presents in Appendix 2 and 3, that allow to built the two crucial indicators to the proposal model: Average of Number of Events and Average Loss by Event (Appendix 4) , for both analytical dimensions: Event Type and Business Line.

For each indicator are calculated simple average, weighted average and standard deviation. The values can be viewed in Appendix 5 and 6.

The following step is to perform Monte Carlo Simulations. First to event frequency (Poisson distribution) and then to Aggregated Losses using the five different distributions early presented. The results are reported in Appendix 7.

The final value for each business line or event type presents a significant range of amounts representing that the chosen distribution has influence in the result. Because the population is simulated using Monte Carlo is not possible to evaluate which distributions best fits to explain Operational Risk losses and has the best estimate for regulatory capital requirements.

4.2. ORX Report 2007

ORX Association is The Operational Riskdata eXchange Association is the world's leading operational risk loss data consortium for the financial industry. ORX was founded with the main goal of creating a sound platform for the secure and anonymised exchange of high-quality operational risk loss data.

ORX currently has 50 members and has over the past four years developed a database of 102,500 operational risk losses and can be very useful for its members as a credible data base for External Data to be used in their internal models of Operational Risk.

As the study presented before is only for US institutions, the ORX data capture realities majority from European markets but has also data from US and Canada. Notice that in terms of dimension the North America members' revenues in average is the double of European members.

The data in this report can be summarized in the following table:

	2002	2003	2004	2005	2006	Total
Total Number of Loss Events	7.814	9.614	12.850	15.606	17.670	63.554
Total Gross Loss Amount (€millions)	5.003	6.668	3.896	3.395	2.831	21.794
Total Gross Income Amount (€millions)	179.707	203.613	260.588	309.963	326.719	1.280.590

Table 2: High level Summary of ORX reported data

The data shows that the major losses took place in 2002 and 2003, but the number of events has increased over the years. The report also show that minor losses have higher frequency and higher losses have lower frequency, as expected in operational risk.

With the data provided by the report about distribution of Loss Frequency and Loss Severity, similar to what is presented for FED, I calculate simple indicator to each Business Line and Event Type presented in Appendix 8.

The next steps are equal to FED and the results of Monte Carlo simulations and aggregated loss distributions estimated are presented in Appendix 9 and 10.

5. Conclusions

Basel II is a challenge for all institutions because it represents an opportunity to define and implement a model that allows managing Operational Risk and also to identify critical sources of Operational risk in their activities. This identification will give the chance to institutions to focus on improving risk management and being able to adjust regulatory capital to their risk profile.

This paper investigates the available methodologies to be applied to create an internal model to estimate regulatory capital using an Aggregated Loss methodology.

The data used in this paper has origin in two different reports, so they are aggregated and treated with different methodologies. This also allows using two different realities: US institutions in FED report and more European institutions in ORX report. Is not correct to simply compare the data but it is useful for testing the proposal model and simulating regulatory capital using different statistical distributions. These distributions are not symmetric. Gamma, Weibull and LogNormal are thin tailed distributions and Pareto is a "fat" tailed distribution, and because of their differences is possible to see different results, meaning that are some distributions that better fit some phenomenons than others, as the outputs graphs in the Appendix demonstrate.

For statistical treatment of frequency was used Poisson, that due to the distribution's characteristics is the one that better fits to explain and estimate the frequency of operational risk events.

The CaR results obtained for both reports can be shown in the next figures divided by business line and event type:

USD	Weibull	LogNormal	Pareto	Beta	Gamma
Corporate Finance	35.645.704,52 USD	219.409.780,31 USD	907.152,28 USD	5.284.170,56 USD	N/A
Trading and Sales	1.735.226.137,32 USD	3.816.218.829,11 USD	3.899.260,98 USD	61.399.713,48 USD	N/A
Retail Banking	359.931.054,73 USD	189.701.665,89 USD	15.622.230,54 USD	323.194.655,21 USD	187.585.333,21 USD
Commercial Banking	48.199.391,30 USD	65.634.867,37 USD	2.735.901,10 USD	40.865.891,31 USD	N/A
Payments & Settlement	15.451.882,32 USD	26.057.592,21 USD	2.585.015,74 USD	36.364.760,05 USD	19.775.772,65 USD
Agency Services	59.580.807,90 USD	55.293.953,44 USD	1.077.566,69 USD	50.687.029,06 USD	53.542.900,63 USD
Asset Mangement	105.731.850,23 USD	619.061.446,79 USD	1.521.698,81 USD	24.121.029,82 USD	N/A
Retail Brokerage	32.259.842,04 USD	26.577.063,59 USD	3.797.130,05 USD	69.801.888,04 USD	26.543.581,28 USD

Legend:　　Maximium Value　　Minimium Value

Figure 15: FED Final Output: Capital Requirement estimative by business line. The values are in US$.

USD	Weibull	LogNormal	Pareto	Beta	Gamma
Internal Fraud	168.816.816,53 USD	453.942.315,76 USD	1.770.391,74 USD	28.853.704,46 USD	N/A
External Fraud	4.745.986.027,42 USD	12.523.682.772,28 USD	12.114.454,63 USD	260.768.294,99 USD	N/A
Employment Practices and workplace safety	31.141.940,19 USD	25.959.959,73 USD	3.829.102,02 USD	57.716.763,84 USD	26.129.797,02 USD
Clientes, Products and Business Practices	3.224.804.912,83 USD	8.028.826.614,66 USD	3.803.948,69 USD	64.427.659,31 USD	390.051.720,95 USD
Damages to Physical Assets	31.253.116,45 USD	160.806.679,15 USD	914.765,78 USD	9.603.988,90 USD	N/A
Business Disruption & System Failures	59.618.855,74 USD	52.347.938,70 USD	900.718,67 USD	9.459.414,52 USD	49.578.888,02 USD
Execution, Delivery and Process Management	114.597.301,17 USD	97.968.650,68 USD	11.667.021,09 USD	236.655.282,43 USD	97.334.198,62 USD
Fraud	31.463.560,50 USD	102.100.224,52 USD	2.184.605,87 USD	30.195.210,78 USD	N/A

Legend:　　Maximium Value　　Minimium Value

Figure 16: FED Final Output: Capital Requirement estimative for each event type. The values are in US$.

Unit: USD	Weibull	LogNormal	Pareto	Beta	Gamma
Corporate Finance	498.919.780,21 USD	3.757.276.681,23 USD	1.081.442,60 USD	9.140.229,23 USD	68.789.208,82 USD
Trading and Sales	3.556.727.855,73 USD	4.716.429.115,07 USD	5.908.592,82 USD	90.545.585,69 USD	1.103.840.992,66 USD
Retail Banking	24.208.069.351,99 USD	21.632.787.340,93 USD	21.049.433,39 USD	444.093.716,18 USD	10.954.655.001,34 USD
Commercial Banking	1.288.629.436,43 USD	1.393.794.805,92 USD	3.788.011,44 USD	67.835.296,46 USD	821.648.888,21 USD
Clearing	42.245.920,99 USD	63.943.731,76 USD	1.454.065,99 USD	21.123.697,27 USD	49.335.933,55 USD
Agency Services	241.841.887,17 USD	802.408.518,74 USD	1.731.190,05 USD	25.257.752,74 USD	N/A
Asset Mangement	1.900.554.098,28 USD	1.284.870.063,88 USD	2.340.608,69 USD	33.073.557,22 USD	N/A
Retail Brokerage	364.471.689,26 USD	264.295.285,32 USD	2.662.003,82 USD	33.796.672,88 USD	154.024.801,07 USD
Private Banking	431.258.174,79 USD	789.048.930,30 USD	3.180.848,06 USD	47.865.287,80 USD	N/A
Corporate Items	302.181.163,86 USD	339.003.436,64 USD	1.929.991,25 USD	22.548.102,55 USD	239.699.156,72 USD

Legend: Maximum Value   Minimum Value
Note: The values calculated using Average quote EUR/USD. Source FED

Figure 17: ORX Final Output: Capital Requirement estimative by business line. The values are in US$ (Exchange rate used in Appendix 1)

Unit: USD	Weibull	LogNormal	Pareto	Beta	Gamma
Internal Fraud	415.136.291,36 USD	484.671.961,70 USD	2.800.194,25 USD	39.246.832,19 USD	342.703.608,93 USD
External Fraud	3.413.073.890,04 USD	8.120.774.255,83 USD	11.653.544,20 USD	217.300.065,15 USD	N/A
Employment Practices and workplace safety	659.469.882,47 USD	341.698.731,35 USD	2.515.482,03 USD	86.485.339,87 USD	N/A
Clientes, Products and Business Practices	1.724.061.371,29 USD	8.498.898.871,70 USD	86.539.319,96 USD	87.009.095,01 USD	1.724.061.371,29 USD
Disasters and Public Safety	53.028.358,97 USD	265.311.791,74 USD	1.143.391,58 USD	11.566.369,73 USD	N/A
Technology and Infrastructure Failures	46.966.945,48 USD	109.123.761,50 USD	1.534.534,11 USD	21.563.858,13 USD	63.412.563,58 USD
Execution, Delivery and Process Management	8.268.753.317,09 USD	8.184.825.090,89 USD	13.465.224,77 USD	215.529.343,99 USD	3.229.016.123,50 USD
Malicious Damage (*)	N/D	N/D	N/D	N/D	N/D

Legend: Maximum Value   Minimum Value
(*) Data Not representative
Note: The values calculated using Average quote EUR/USD. Source FED

Figure 18: ORX Final Output: Capital Requirement estimative for each event type. The values are in US$. (Exchange rate used in Appendix 1)

From the statistical point of view Pareto's estimative present the lowest values. This happens because Pareto's does not depend on severity average and it uses the same parameters (shape and scale).

For ORX, Weibull and LogNormal Distributions return the higher estimative for CaR and for FED data Beta Distribution also presents higher values.

Due to the fact that I do not have real populations of events I can not make any statistical test to determine which one better describes expect behavior in any of the reports used. When looking to final results available in figures presented above, the major CaR appear in business lines and event type that in initial data had more frequency and severity as expected. It is in other business line or event type that the results show a higher variation between the minimum and maximum value estimates.

Is important to remember that ORX report has more years of historical and has been collected more recently, if we want to compare the results obtained.

This last fact may indicate that the data quality is better because institutions have developed in the past years more effective operational risk management policies and they have implemented also some procedures. It is also important to refer that their report is a voluntary decision as it is to be a member of ORX association.

The final values of the proposed model are higher than the initial ones as expected because this model has should be able to evaluate the capital needs to face expected and unexpected losses. The unexpected losses are the crucial values that regulatory capital should be able to cover when an operational risk loss happens, because the expected losses, institutions should be able to mitigate or for example to recover them in pricing or having contracted any type insurances.

The model's principles establish in Basel II framework are very similar to VaR model, a concept introduced to measure market risk and used by all institutions. The important difference is that market risk losses are easy to track and report and operational risk losses are not, what transform the modeling process just the final step and focus the major concern is Loss Data Collection Process. For some opinions, these concerns are reflected in confidence level defined by Basel Committee of 99.90%, a conservative value that can be an indicator of some questions about historical data soundness and lack of knowledge for modeling process this type of factors.

The diversity of results obtain demonstrate that soundness information and their completeness is the challenge for all institutions, for building a strong internal model of Operational Risk that aims not just risk manage but also helps institutions to mitigate

risks and incentive them to adopt advanced approaches as a way to adjust regulatory capital to their own risk profile.

As this study was developed, Basel Committee published the 2008 LDCE for Operational Risk and an additional paper with the observed range of practice in key elements of Advanced Measurement Approaches. This exercise was the first international effort to collect information on all four data elements used in the AMA and with the bigger participation of institutions.

From the range of practice is important to retain the diversity of distributions used to modeling the severity of operational risk losses and the most representative are the same used in the model presented in this paper, as presented in the next figure:

Severity Distributions Applied to the Entire Distribution
Number and percentage of Banks by Region

	All Participating AMA Banks		Australia		Europe		Japan		North America	
Number of AMA banks	42		5		20		7		10	
Lognormal	14	33%	2	40%	8	40%	1	14%	3	30%
Gamma	3	7%	1	20%	2	10%	0	0%	0	0%
Generalised Pareto	2	5%	1	20%	0	0%	0	0%	1	10%
Weibull	7	17%	2	40%	4	20%	0	0%	1	10%
g and h	1	2%	0	0%	0	0%	0	0%	1	10%
Generalised beta	0	0%	0	0%	0	0%	0	0%	0	0%
Mixture of Lognormal-Gamma	2	5%	0	0%	1	5%	0	0%	1	10%
Mixture of Lognormal	3	7%	1	20%	1	5%	0	0%	1	10%
Empirical distribution	6	14%	1	20%	4	20%	1	14%	0	0%
Others	6	14%	1	20%	3	15%	1	14%	1	10%

Note: Banks were able to select more than one answer per question regarding the type of frequency and severity distributions used.

Figure 19: Results presented by Basel Committee from the document "Observed range of practice in key elements of Advanced Measurement Approaches (AMA)"

The above figure show the values published regarding the application of one single distribution model for all data because is the same approach proposed in this paper. In this report of Basel Committee there are available data about other models used in practice.

This is important to support the model developed in this paper because the institutions are using real and detailed information and the majority is using approaches similar to the ones I proposed.

The main conclusion of this paper is that is crucial to consider all four elements to build a soundness model to estimate regulatory capital with internal models and information quality is the critical factor of success. Institutions should have in mind that is not enough to obtain a good fit based on statistical measures, but also that the model developed should be suitable for their reality.

6. References

(1) Maximum Likelihood from Incomplete Data via the EM Algorithm - A. P. Dempster, N. M. Laird and D. B. Rubin *Journal of the Royal Statistical Society. Series B (Methodological),* Vol. 39, No. 1 (1977), pp. 1-38   (article consists of 38 pages)

(2) The EM Algorithm and Extensions, 2E
Series: Wiley Series in Probability and Statistics - Published Online: 30 Apr 2007
Author(s): Geoffrey J. McLachlan, Thriyambakam Krishnan

7. Bibliography

International Convergence of Capital Measurement and Capital Standards (November 2005), Basel Committee on Banking Supervision (BIS)

Operational risk management practices: Feedback from a thematic review (February 2005), Financial Services Authority

Operational Risk in Banks: an analysis of empirical data from an Australian Bank (September 2007), Institute of Actuaries of Australia

ORX Operational Risk Report (2007), ORX Association

Results of the 2004 Loss Data Collection Exercise for Operational Risk (May 12, 2005), Federal Reserve System Office of the Comptroller of the Currency Office of Thrift Supervision Federal Deposit Insurance Corporation

Results of the 2007 Operational Risk Data Collection Exercise (August 10, 2007), Planning and Coordination Bureau, Financial Service Agency Financial Systems and Bank Examination Department, Bank of Japan

Sound Practices for the Management and Supervision of Operational Risk (February 2003), Basel Committee on Banking Supervision (BIS)

The 2002 Loss Data Collection Exercise for Operational Risk: Summary of the Data Collected (March 2003), Basel Committee on Banking Supervision (BIS)

The Quantitative Impact Study for Operational Risk: Overview of Individual Loss Data and Lessons Learned (January 2002), Basel Committee on Banking Supervision (BIS)

Results from the 2008 Loss Data Collection Exercise for Operational Risk (July 2009), Basel Committee on Banking Supervision (BIS)

Observed range of practice in key elements of AMA (July 2009), Basel Committee on Banking Supervision (BIS)

Bühlmann, Hans; Shevchenko, Pavel V. & Wüthrich, Mario V. (February 19, 2007), **A "Toy" Model for Operational Risk Quantification using Credibility Theory**

Dempster, A.P.; Laird, N. M. & Rubin, D. B. (1977), **Maximum Likelihood from Incomplete Data via the EM Algorithm** published in *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 39, No. 1

Haugh, Martin (2004), **The Monte Carlo Framework, Examples from Finance and Generating Correlated Random Variables** (IEOR E4703 – Fall 2004)

Jobst, Andreas A. (2007), **Operational Risk – The Sting is Still in the Tail But the Poison Depends on the Dose** (Working Paper version February 24, 2007), published in Journal of Operational Risk Vol. 2, No. 2 (Summer)

Jobst, Andreas A. (February 7, 2007), **Constraints of Operational Risk Measurement and The Treatment of Operational Risk under the New Basel Framework,** Working Paper

Leippold, Markus & Vanin, Paolo (November 3, 2003), **The Quantification of Operational Risk**

Laviada, Ana Fernández; García, Francisco Javier Martínez & Rodríguez, Francisco Somohano (August 2005), **Operational risk management under Basel II: the case of the Spanish financial services** – published in European Finance Association - 32nd Annual Meeting

McLachlan, Geoffrey J. & Krishnan, Thriyambakam (Second edition 2008), **The EM Algorithm and Extensions**

Valle, Luciana Dalla & Giudici, Paolo (2008) **Bayesian Copulae distributions, with application to Operational Risk**

Web Sites:

Bank for International Settlements, www.bis.org

ORX Association, www.orx.org

Gloria Mundi, www.gloriamundi.org

SAS Software, www.sas.com

Federal Reserve, www.federalreserve.gov

Scholar Google, http://scholar.google.pt/

Wikipedia, www.wikipedia.org

Appendix

Appendix 1 Summary of Exchange Rates used in calculations

EUR/USD

Spot Rates

Exchange Rate	2000	2001	2002	2003	2004	2005	2006	2007	2008
EUR/USD	0,9159	0,8952	0,9416	1,1321	1,2438	1,2449	1,2563	1,3711	1,5294

Annual average of available quotations
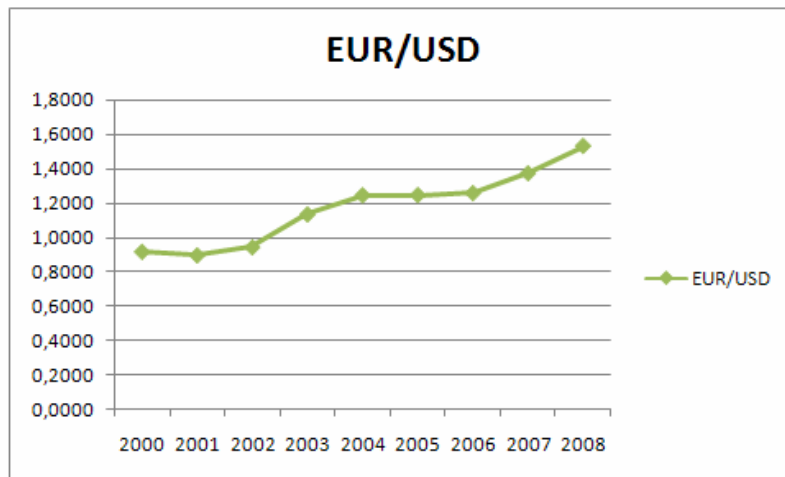
Source: FED HISTORICAL DATA

Appendix 2 FED LDCE DATA

Table 3: Number of Losses, Annualized by Business Line and Event Type Sample 1 (Thousands)

	Internal Fraud	External Fraud	Employment Practices and workplace safety	Clientes, Products and Business Practices	Damages to Physical Assets	Business Disruption & System Failures	Execution, Delivery and Process Management	Other	Fraud	Total	% of Total
Corporate Finance	1,4	2,1	10,9	14,1	0,4	-	22,6	6,0	1,6	59,1	0,3%
Trading and Sales	4,1	2,5	30,4	35,7	5,8	43,7	1.204,2	-	8,5	1.334,9	7,3%
Retail Banking	419,8	6.218,3	690,2	810,5	103,1	39,2	2.256,7	126,3	385,0	11.049,1	60,1%
Commercial Banking	8,5	484,1	31,5	65,4	1,4	5,2	254,2	4,0	80,6	935,0	5,1%
Payments & Settlement	96,3	81,4	32,4	6,8	1,8	9,5	549,3	1,0	41,7	820,3	4,5%
Agency Services	1,4	6,1	7,0	57,1	1,8	25,8	829,6	-	-	928,8	5,1%
Asset Mangement	0,3	47,0	19,2	24,7	0,2	6,6	335,1	-	16,2	449,3	2,4%
Retail Brokerage	11,0	19,0	254,4	606,2	-	1,8	404,1	-	36,4	1.333,0	7,3%
Other	76,3	304,4	321,6	72,6	22,0	3,9	633,9	13,0	13,9	1.461,7	8,0%
Total	619,2	7.164,9	1.397,7	1.693,2	136,5	135,7	6.489,7	150,4	583,9	18.371,1	100%
% of Total	3,4%	39,0%	7,6%	9,2%	0,7%	0,7%	35,3%	0,8%	3,2%	100,0%	

Table A: Data Used in this document Source: LDCE 2004 – FED (Number of Losses, Annualized By Business Line and Event Type **Sample 1: losses ≥ $10,000 Occurring in Years When Data Capture Appears Stable**

Appendix 3 FED LDCE DATA

Table 4: Loss Amount , Annualized by Business Line and Event Type Sample 1 (USD Millions)

	Internal Fraud	External Fraud	Employment Practices and workplace safety	Clientes, Products and Business Practices	Damages to Physical Assets	Business Disruption & System Failures	Execution, Delivery and Process Management	Other	Fraud	Total	% of Total
Corporate Finance	11,9	0,2	2,2	25,5	0,0		4,2	0,6	0,0	44,7	0,5%
Trading and Sales	8,5	100,9	4,3	370,9	0,3	5,0	238,8		12,9	741,6	8,6%
Retail Banking	35,9	237,3	75,3	346,8	8,8	17,8	316,2	5,6	22,8	1.066,4	12,3%
Commercial Banking	0,6	60,4	2,8	67,2	0,1	0,2	24,1	0,2	3,5	159,2	1,8%
Payments & Settlement	7,2	10,9	2,1	1,0	0,2	1,9	24,1	0,0	4,1	51,4	0,6%
Agency Services	1,3	0,8	1,6	5,5	0,7	1,5	85,5			97,0	1,1%
Asset Mangement	0,1	1,6	2,1	181,4	0,0	0,7	32,8		0,5	219,3	2,5%
Retail Brokerage	2,2	1,5	28,4	81,0		0,0	22,0		6,0	141,1	1,6%
Other	8,9	25,5	29,6	5.820,5	110,7	38,1	84,9	3,9	0,5	6.122,5	70,8%
Total	76,6	439,1	148,3	6.899,7	120,8	65,3	832,7	10,2	50,4	8.643,1	100%
% of Total	0,9%	5,1%	1,7%	79,8%	1,4%	0,8%	9,6%	0,1%	0,6%	100,0%	

Table B: Data Used in this document Source: LDCE 2004 – FED (Total Loss Amount ($US Millions), Annualized By Business Line and Event Type
Sample 1: losses ≥ $10,000 Occurring in Years When Data Capture Appears Stable

Appendix 4 FED LDCE DATA

	Internal Fraud	External Fraud	Employment Practices and workplace safety	Clientes, Products and Business Practices	Damages to Physical Assets	Business Disruption & System Failures	Execution, Delivery and Process Management	Other	Fraud	Total	% of Total
Corporate Finance	8,500	0,095	0,203	1,809	0,000	-	0,187	0,102	0,000	10,9	6,2%
Trading and Sales	2,073	40,360	0,141	10,388	0,052	0,115	0,198	-	1,515	54,8	31,1%
Retail Banking	0,085	0,038	0,109	0,428	0,085	0,455	0,140	0,044	0,059	1,4	0,8%
Commercial Banking	0,071	0,125	0,088	1,028	0,079	0,044	0,095	0,052	0,044	1,6	0,9%
Payments & Settlement	0,075	0,134	0,065	0,147	0,106	0,200	0,044	0,000	0,098	0,9	0,5%
Agency Services	0,915	0,131	0,229	0,096	0,400	0,059	0,103	-	-	1,9	1,1%
Asset Mangement	0,303	0,034	0,109	7,344	0,000	0,106	0,098	-	0,033	8,0	4,6%
Retail Brokerage	0,200	0,079	0,112	0,134	-	0,000	0,055	-	0,165	0,7	0,4%
Other	0,117	0,084	0,092	80,171	5,032	9,769	0,134	0,296	0,038	95,7	54,4%
Total	12,339	41,080	1,147	101,545	5,753	10,748	1,053	0,494	1,953	176,112	100%
% of Total	7,0%	23,3%	0,7%	57,7%	3,3%	6,1%	0,6%	0,3%	1,1%	100,0%	

Table C: Data Output for Average Loss by event of Operational Risk using information presented in Appendix 2 and 3. The values are in $US Millions

Appendix 5- 2004 FED LDCE

Bussiness Lines Average Annual (US$ Millions)	Weighted Average	Simple Average	Standard deviation (using weight average)	Standard deviation (using simple average)
Corporate Finance	6,94	1,36	2,96	6,31
Trading and Sales	31,79	6,86	14,45	28,82
Retail Banking	0,31	0,16	0,17	0,22
Commercial Banking	0,68	0,18	0,46	0,68
Payments & Settlement	0,13	0,10	0,05	0,27
Agency Services	0,56	0,28	0,35	1,66
Asset Mangement	6,73	1,00	2,00	6,07
Retail Brokerage	0,14	0,11	0,05	0,06
Other	68,40	10,64	26,74	106,27

Event Type Average Annual (US$ Millions)	Weighted Average	Simple Average	Standard deviation (using weight average)	Standard deviation (using simple average)
Internal Fraud	6,28	1,37	3,33	5,93
External Fraud	39,65	4,56	5,28	35,49
Employment Practices and workplace safety	0,15	0,13	0,06	0,06
Clientes, Products and Business Practices	64,94	11,28	29,55	61,25
Damages to Physical Assets	4,43	0,72	1,58	4,04
Business Disruption & System Failures	8,91	1,34	2,73	1,21
Execution, Delivery and Process Management	0,14	0,12	0,05	0,05
Fraud	1,20	0,24	0,59	1,12
Other	0,21	0,10	0,11	1,25

Table E: Descriptive Statistics for the Average Loss by business line and event type using Appendix 2 and Appendix 3 data

Appendix 6: ORX Report 2007: Original Data

	Internal Fraud	External Fraud	Employment Practices and workplace safety	Clientes, Products and Business Practices	Disasters and Public Safety	Technology and Infrastructure Failures	Execution, Delivery and Process Management	Malicious Damage	Total Number of Losses
Corporate Finance	26,15	91,53	30,51	4897,11	0,00	0,00	50,13	0,00	5095,44
Trading and Sales	311,65	34,87	58,84	1658,52	0,00	56,66	1867,75	0,00	3988,30
Retail Banking	455,49	1610,58	610,23	1357,77	82,82	150,38	1512,50	2,18	5781,95
Commercial Banking	311,65	202,68	47,95	793,30	4,36	26,15	557,93	0,00	1944,02
Clearing	6,54	23,97	6,54	82,82	2,18	10,90	63,20	0,00	196,15
Agency Services	6,54	4,36	17,44	671,26	0,00	4,36	220,12	0,00	924,07
Asset Mangement	15,26	10,90	34,87	867,40	2,18	6,54	185,25	0,00	1122,39
Retail Brokerage	43,59	28,33	39,23	239,73	2,18	0,00	63,20	0,00	416,27
Private Banking	154,74	23,97	28,33	701,77	2,18	4,36	169,99	0,00	1085,34
Corporate Items	8,72	15,26	91,53	422,80	427,16	2,18	272,43	0,00	1240,08
Total Number of Losses	1340,33	2046,46	965,47	11692,48	523,06	261,53	4962,49	2,18	21794,00

Table F: Number of losses distributed by business line and event type

Appendix 7: ORX Report 2007 – Descriptive Statistics

Bussiness Lines Average Annual (€ Thousands)	Weighted Average	Simple Average	Standard deviation (using weight average)	Standard deviation (using simple average)
Corporate Finance	89.373,37	12.088,56	17.745,15	79.295,85
Trading and Sales	30.190,58	9.461,96	9.582,53	22.836,38
Retail Banking	24.075,69	13.717,26	8.511,19	13.406,62
Commercial Banking	10.561,76	4.612,06	4.507,04	7.464,07
Clearing	1.126,01	465,34	517,32	839,11
Agency Services	10.257,63	2.192,28	4.104,18	9.049,54
Asset Mangement	13.330,30	2.662,79	5.814,60	12.149,29
Retail Brokerage	2.996,17	987,56	1.818,80	2.709,71
Private Banking	9.560,51	2.574,89	5.115,93	8.658,62
Corporate Items	6.797,54	2.942,00	2.118,11	4.399,05

Event Type Average Annual (€ Thousands)	Weighted Average	Simple Average	Standard deviation (using weight average)	Standard deviation (using simple average)
Internal Fraud	6070	2544	2419	4276
External Fraud	24549	3884	11591	23693
Employment Practices and workplace safety	7693	1832	5105	5724
Clientes, Products and Business Practices	50553	22192	36584	39160
Disasters and Public Safety	6871	993	2618	4994
Technology and Infrastructure Failures	1939	496	1097	1383
Execution, Delivery and Process Management	24033	9419	12738	17024
Malicious Damage	41	4	0	41

Table G: Descriptive Statistics for the Average Loss by business line and event type using ORX data. The values are in Euros Thousands

Appendix 8: 2007 ORX Report – Results produced by the model (representative graphs for each distribution by Business Line

Exhibit 8.1: Agency Services
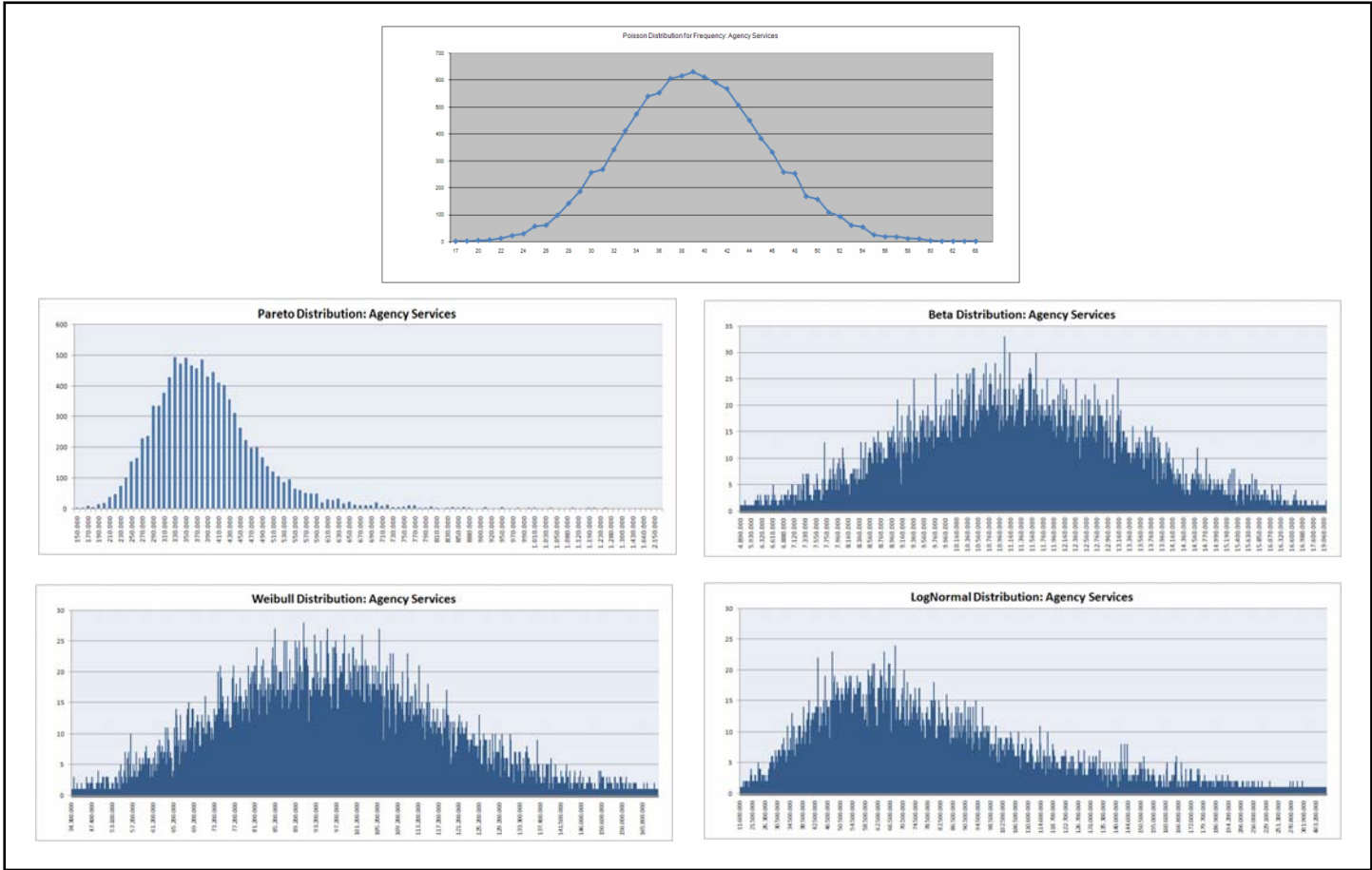
Exhibit 8.2: Asset Management
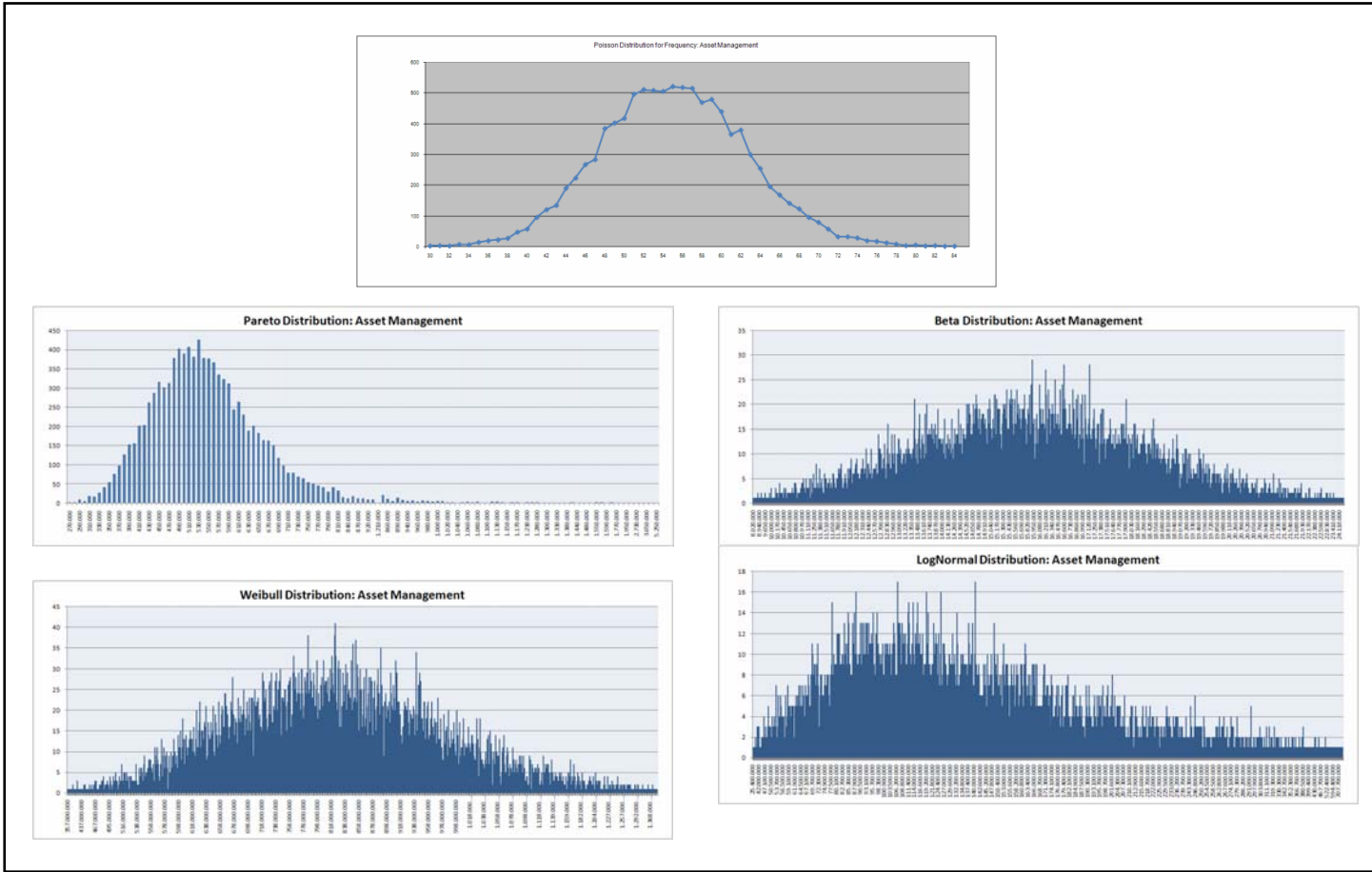
Exhibit 8.3: Clearing
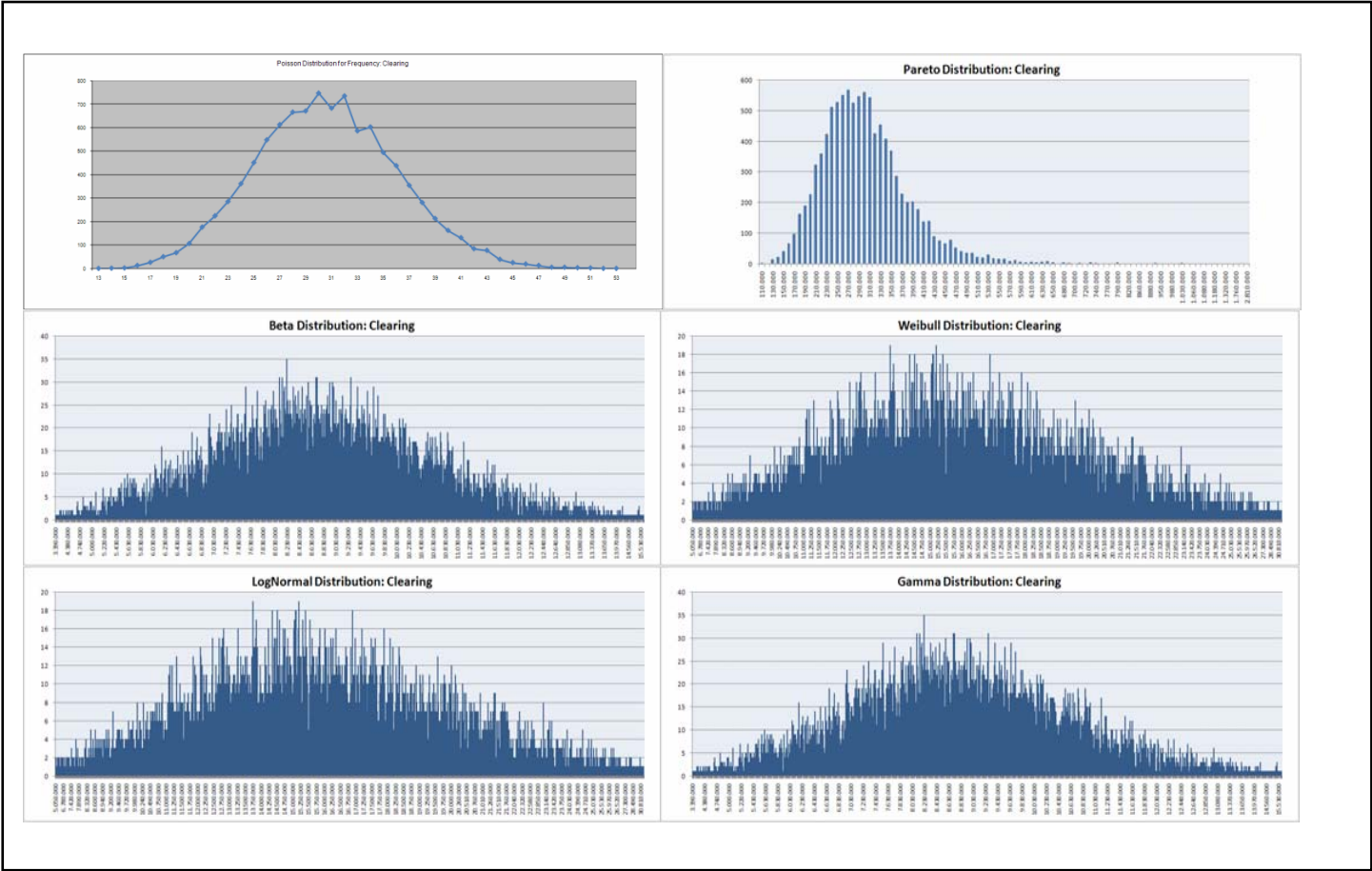
Exhibit 8.4: Commercial Banking
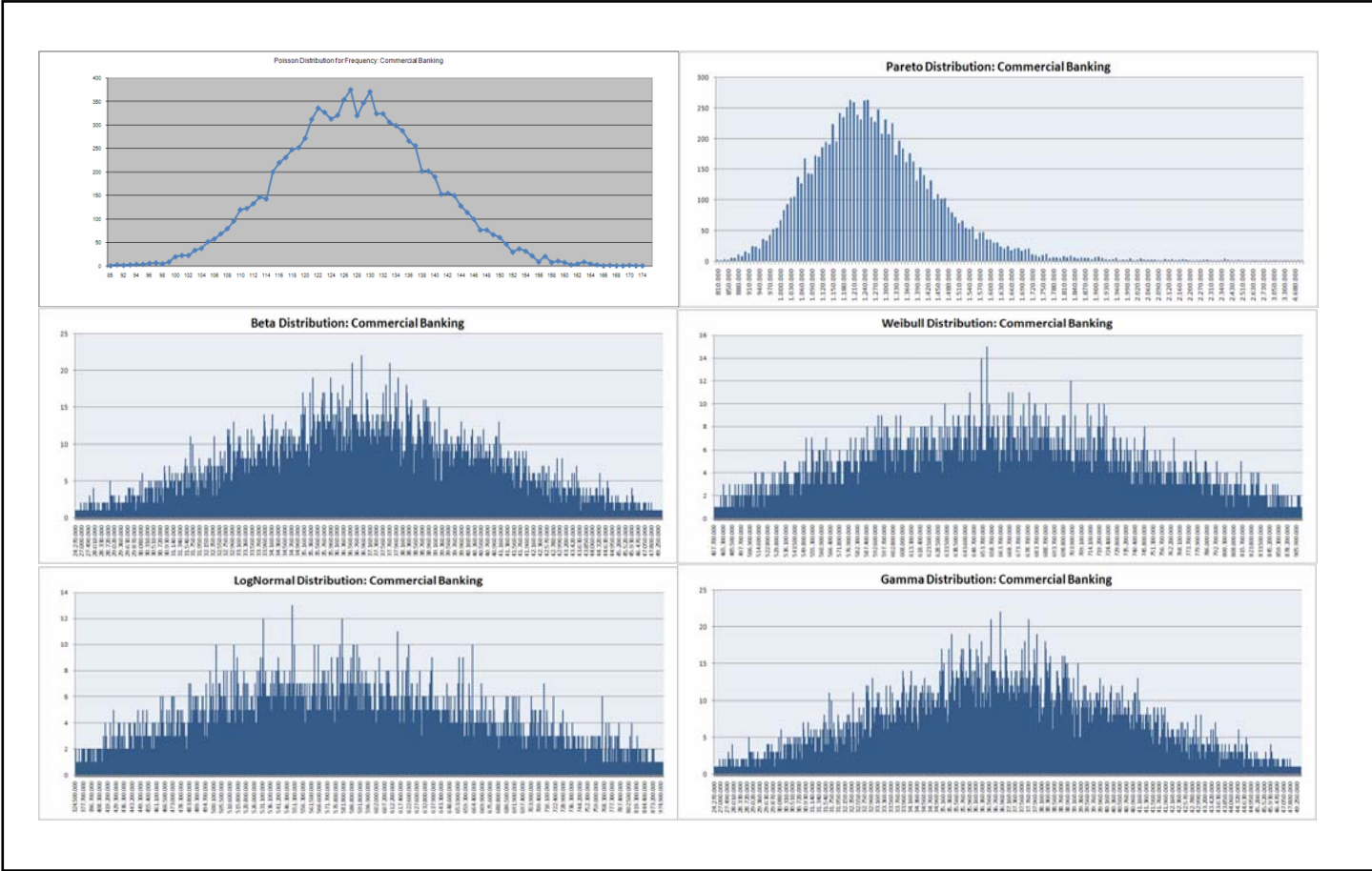
Exhibit 8.5: Corporate Finance
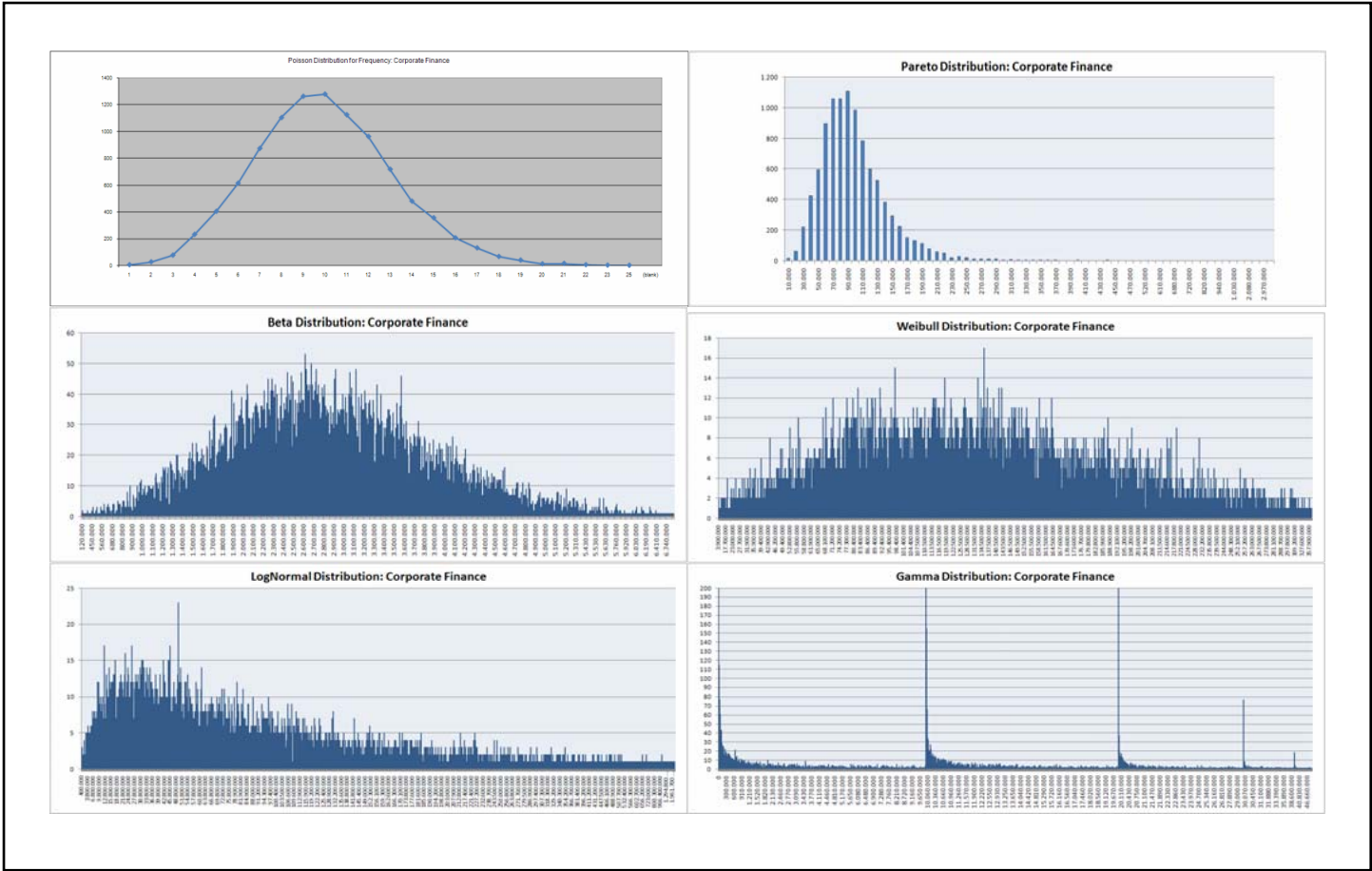
Exhibit 8.6: Corporate Items
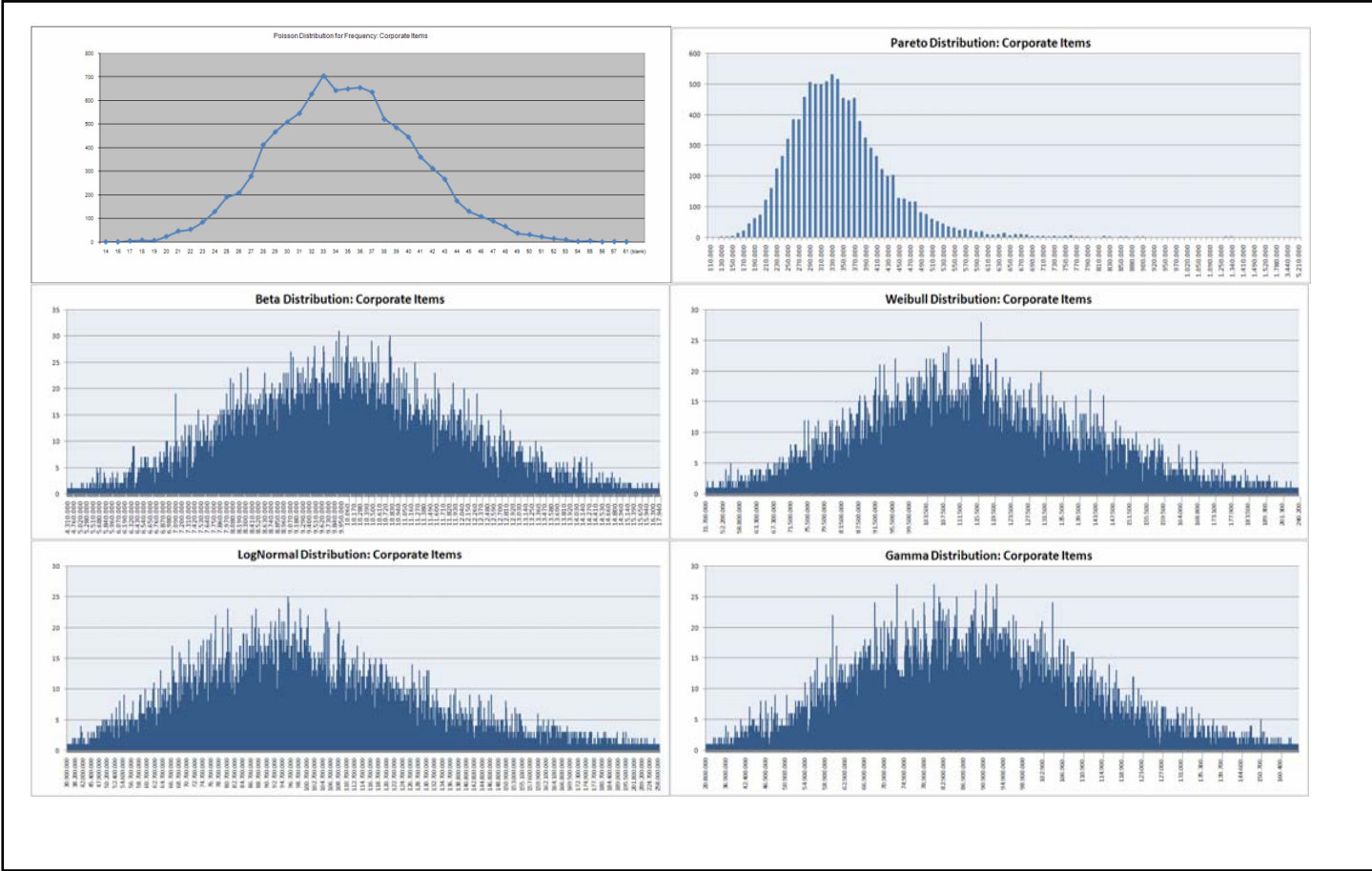
Exhibit 8.7: Private Banking
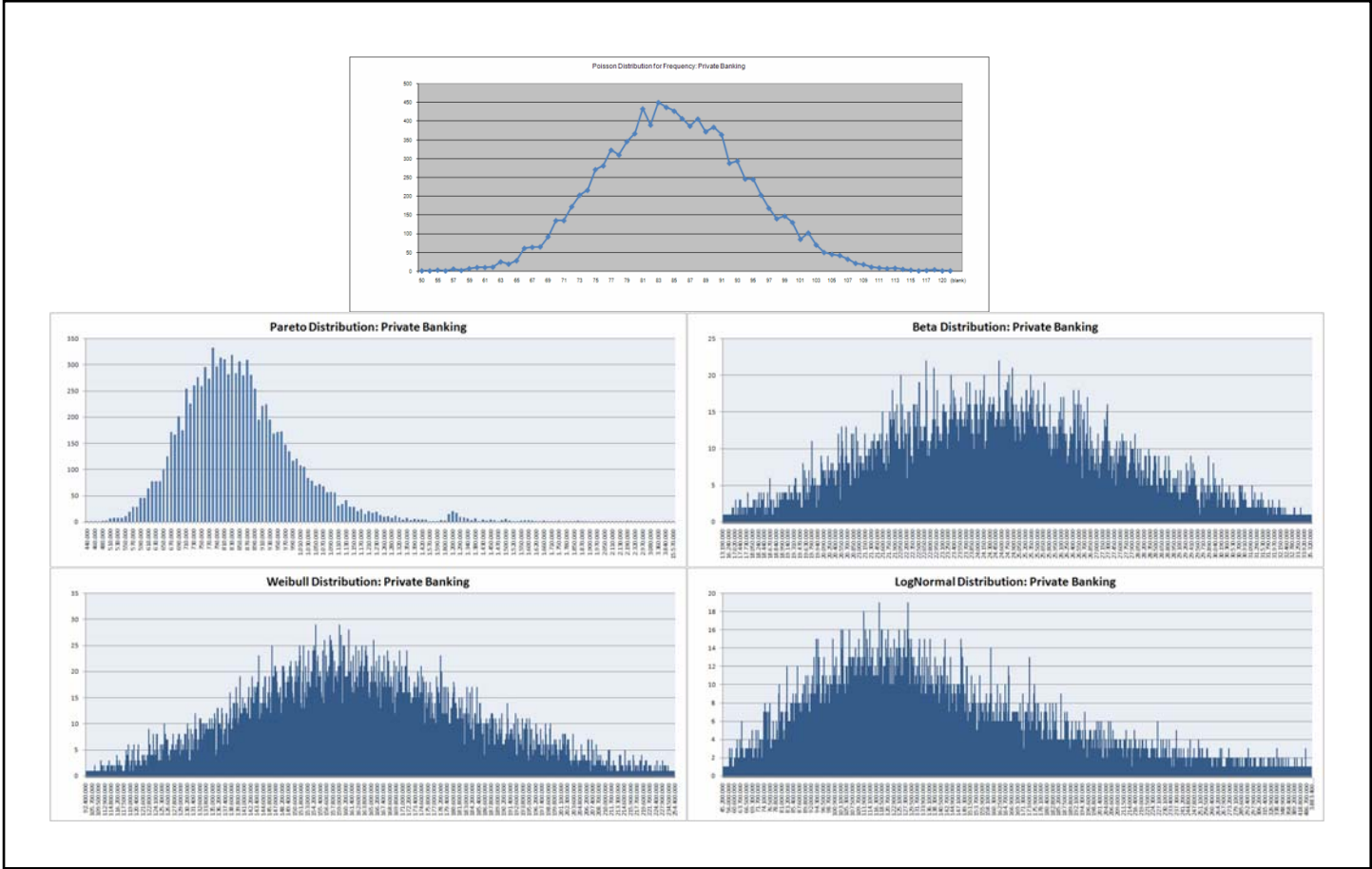
Exhibit 8.8: Retail Brokerage
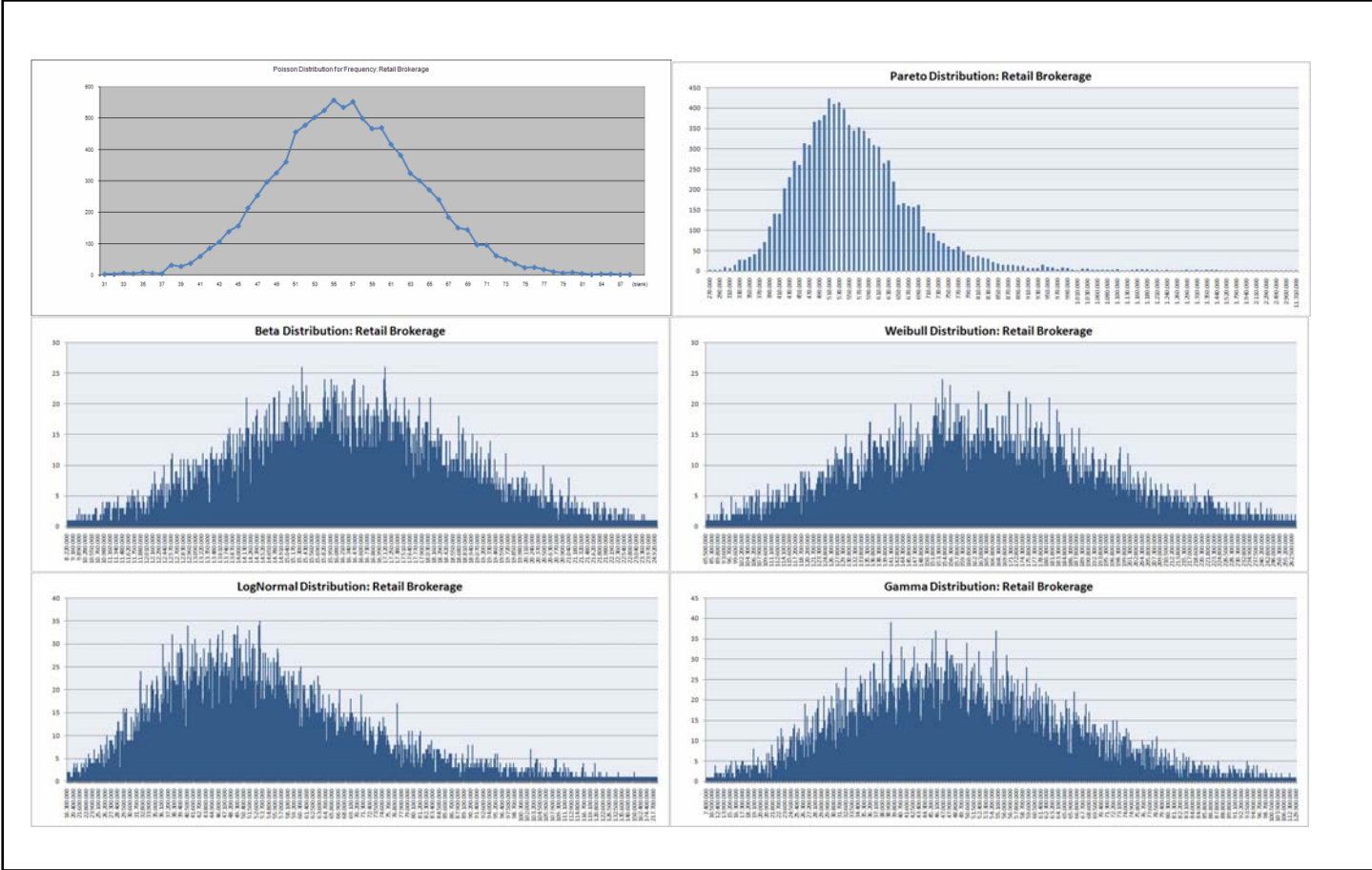
Exhibit 8.9: Trading & Sales
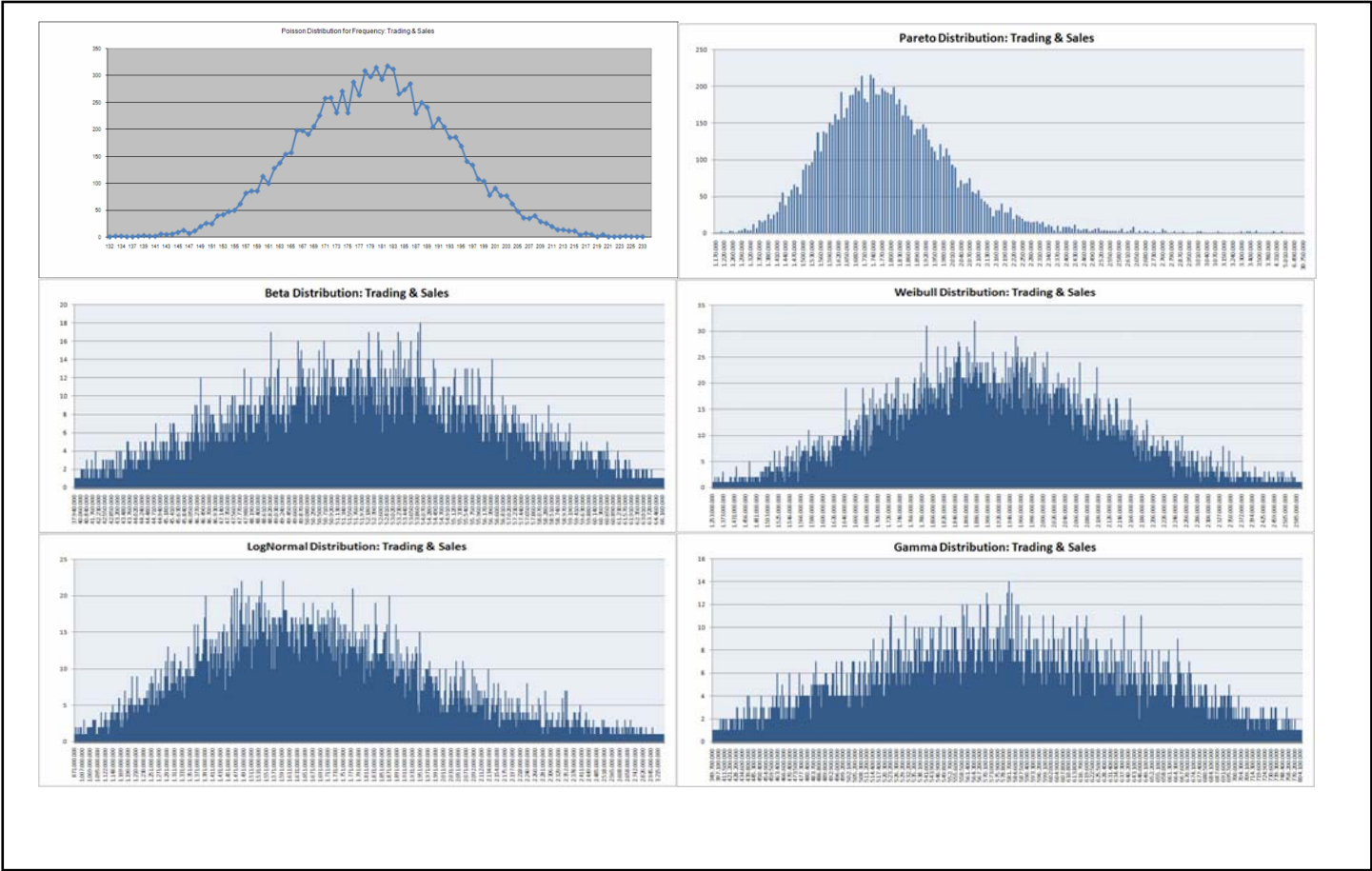
Exhibit 8.10: Retail Banking
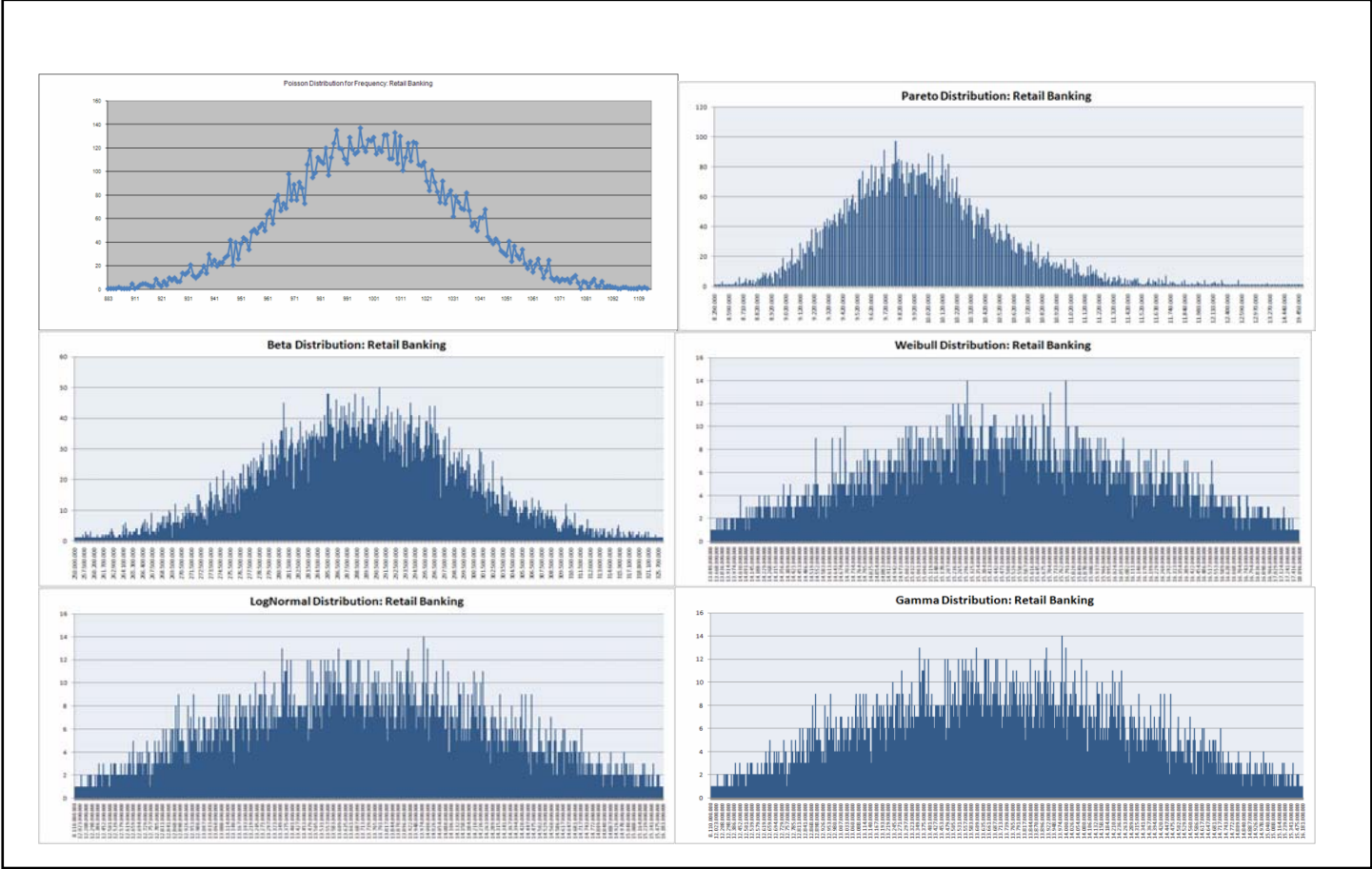
Appendix 9: 2007 ORX Report – Results produced by the model (representative graphs for each distribution by Event Type

Exhibit 9.1: Clients, Products and Business Practices
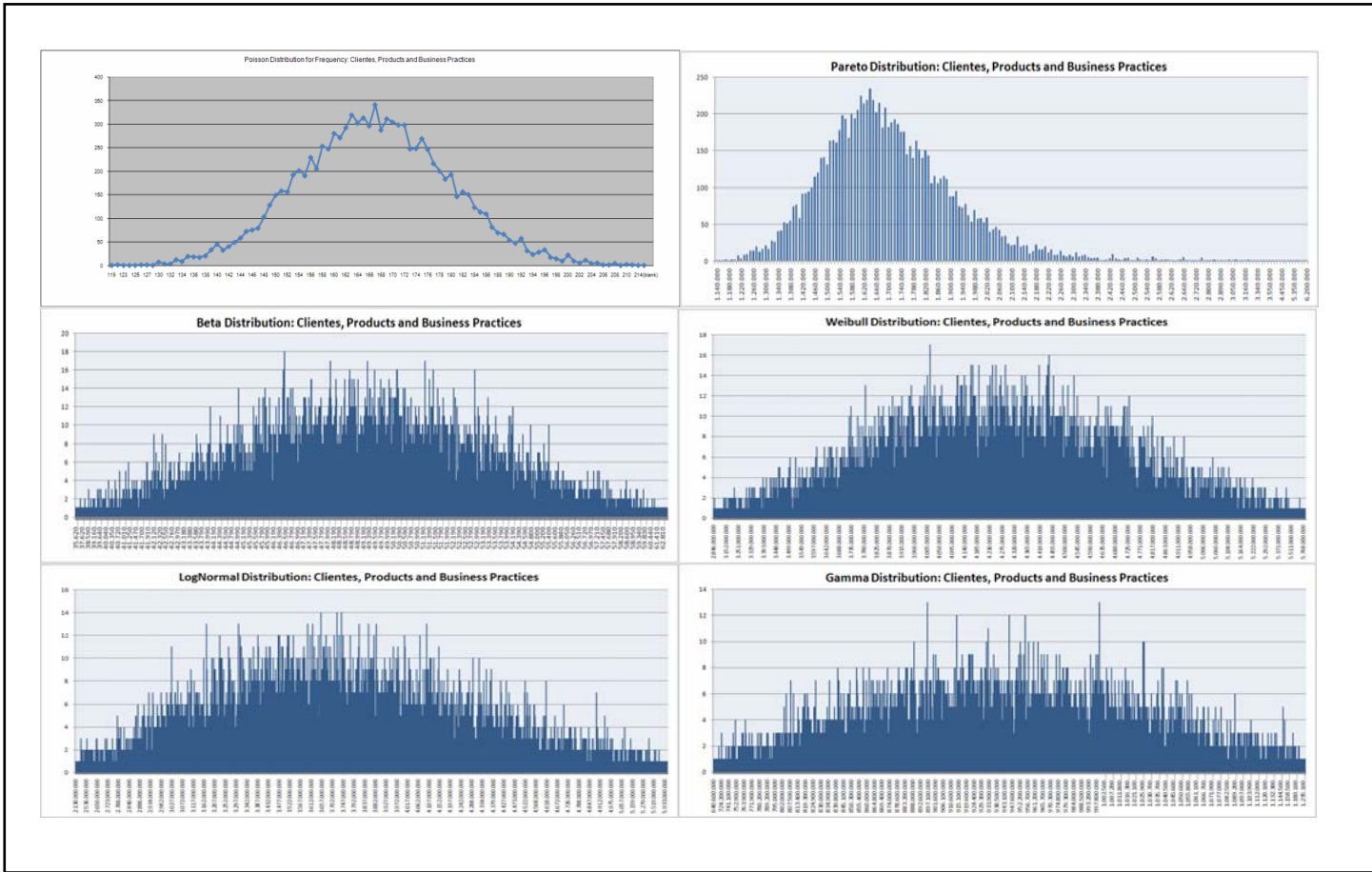
Exhibit 9.2: Disasters and Public Safety
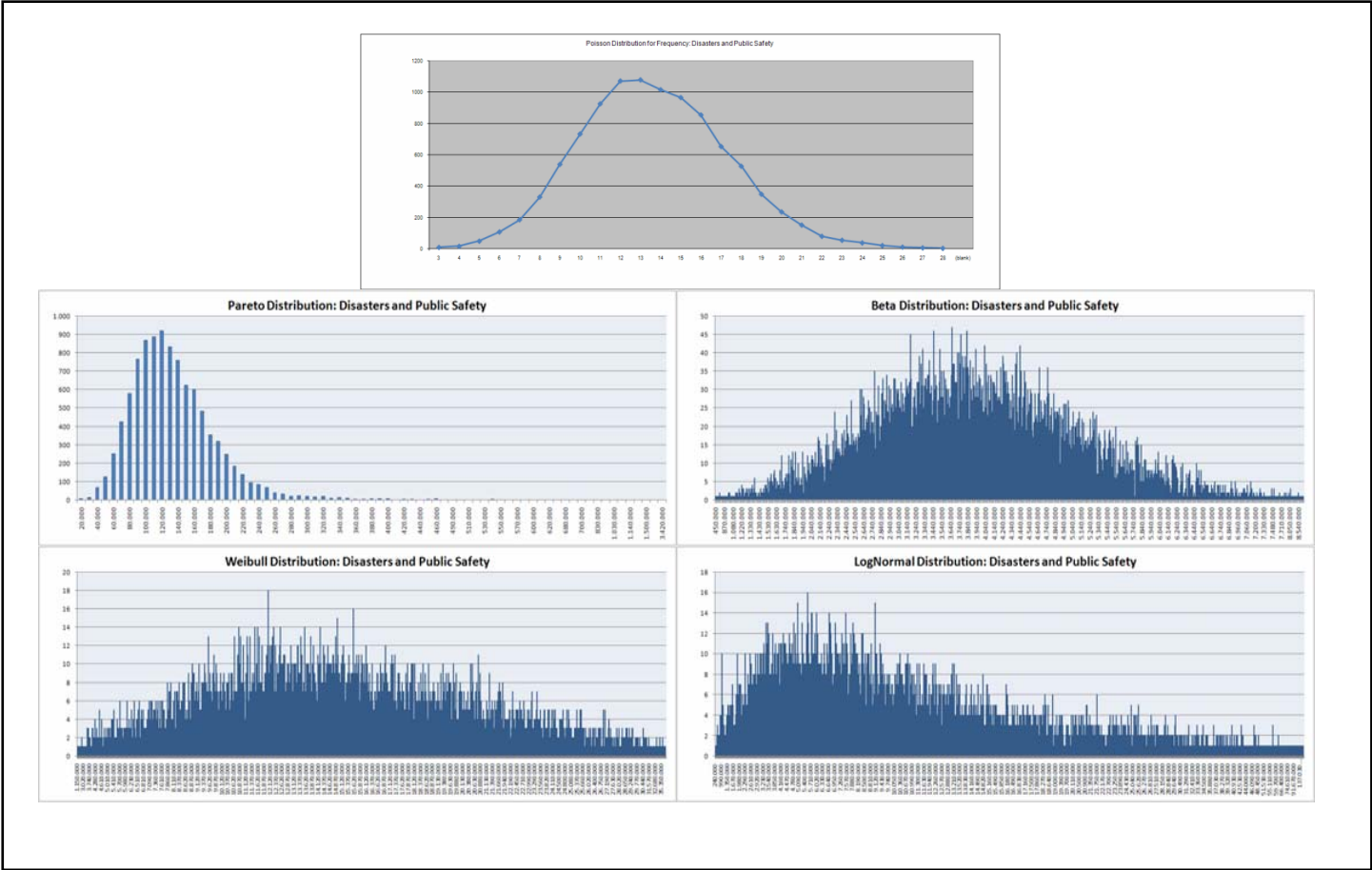
Exhibit 9.3: Employment Practices and workplace Safety
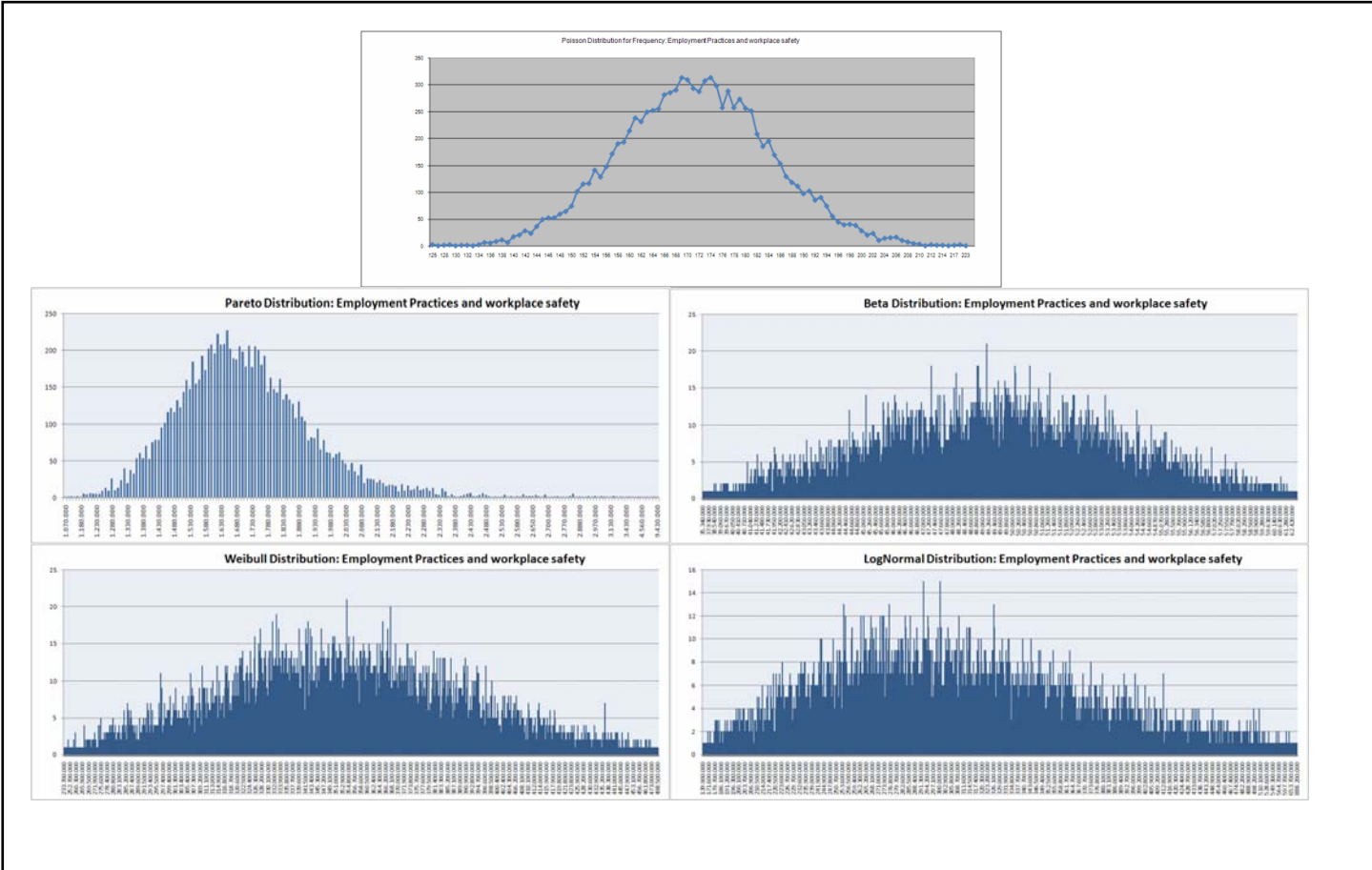
Exhibit 9.4: Technology and Infrastructures Failures
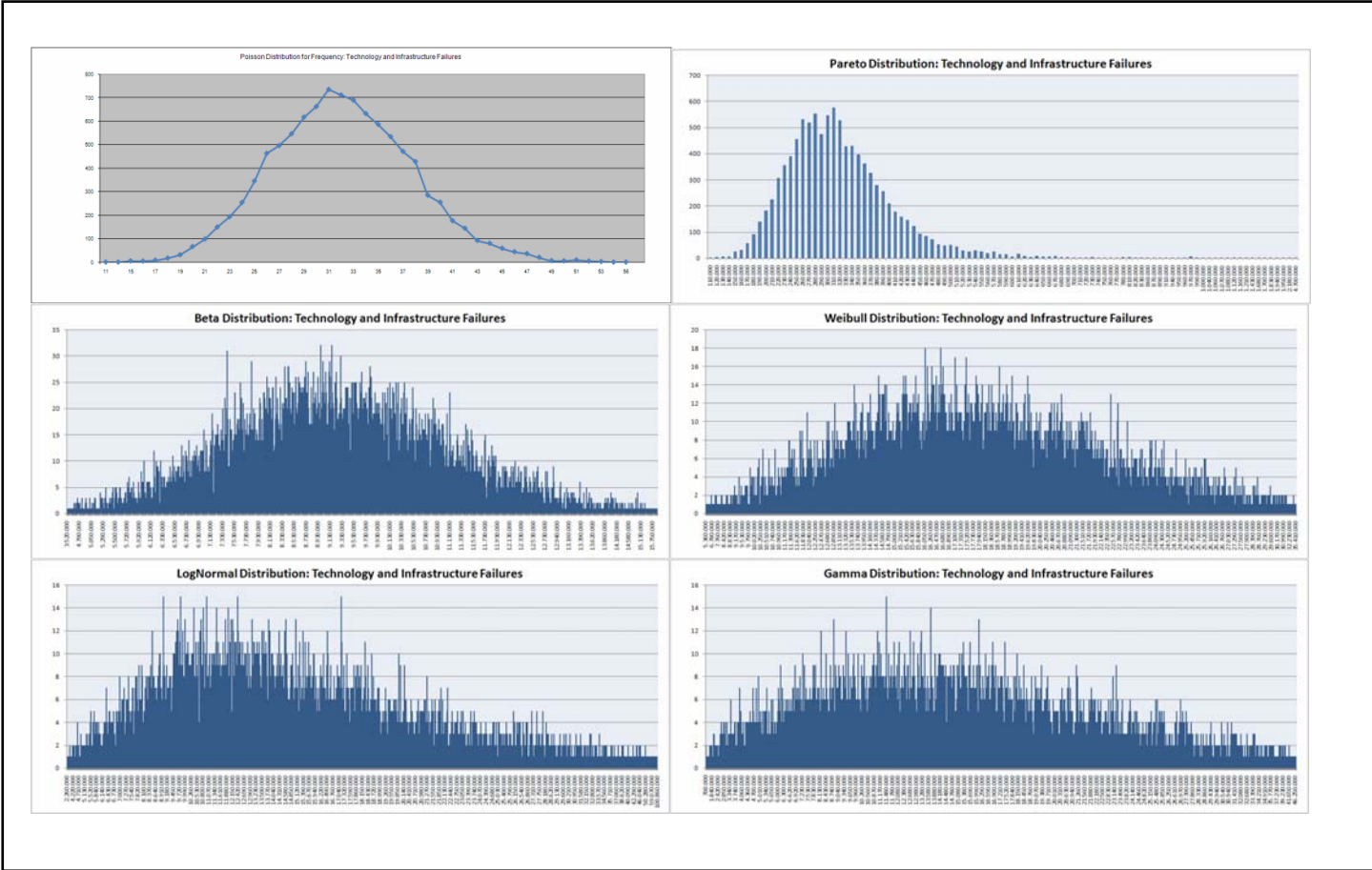
Exhibit 9.5: Execution, Delivery and Process Management
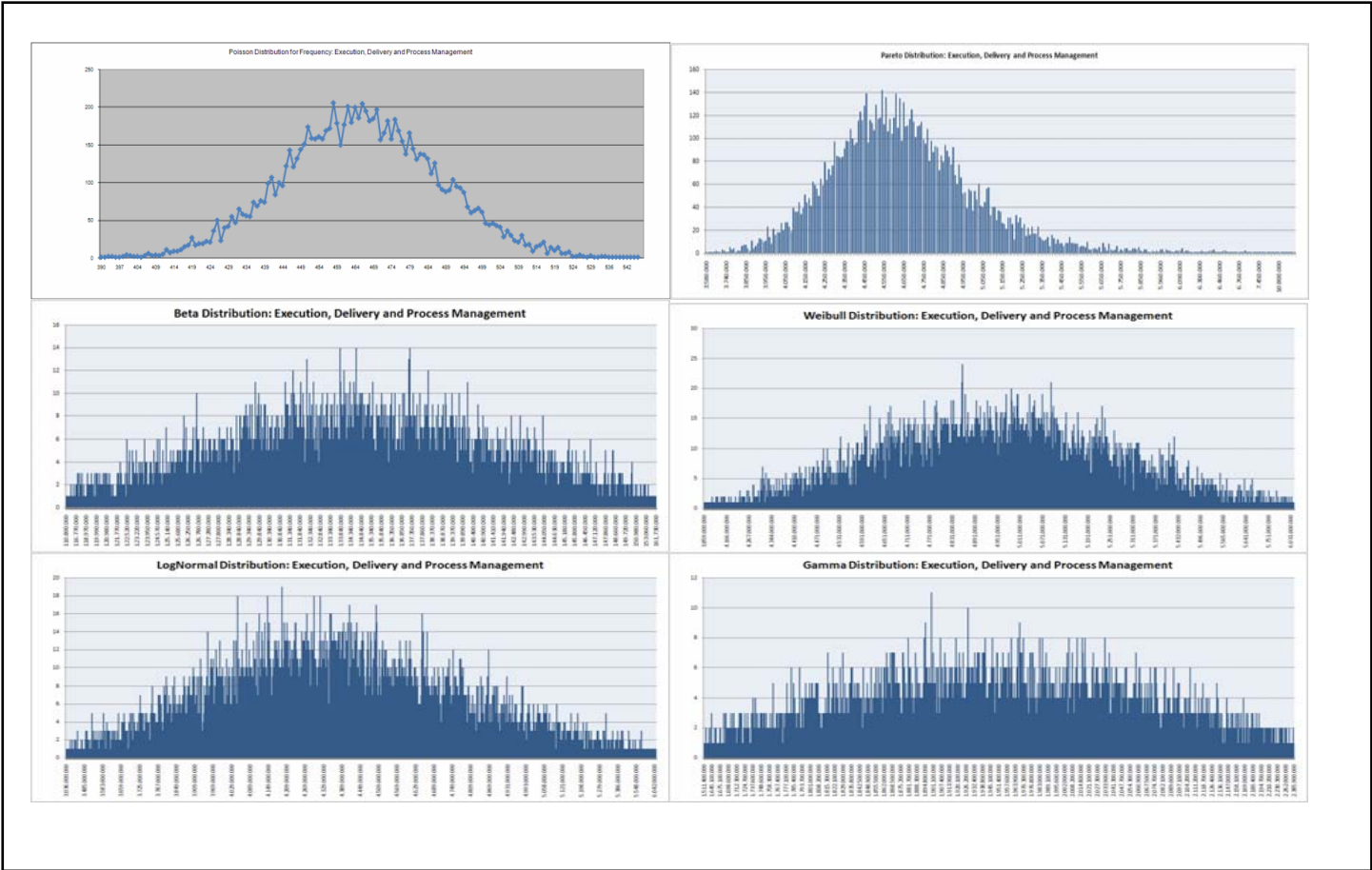
Exhibit 9.6: External Fraud
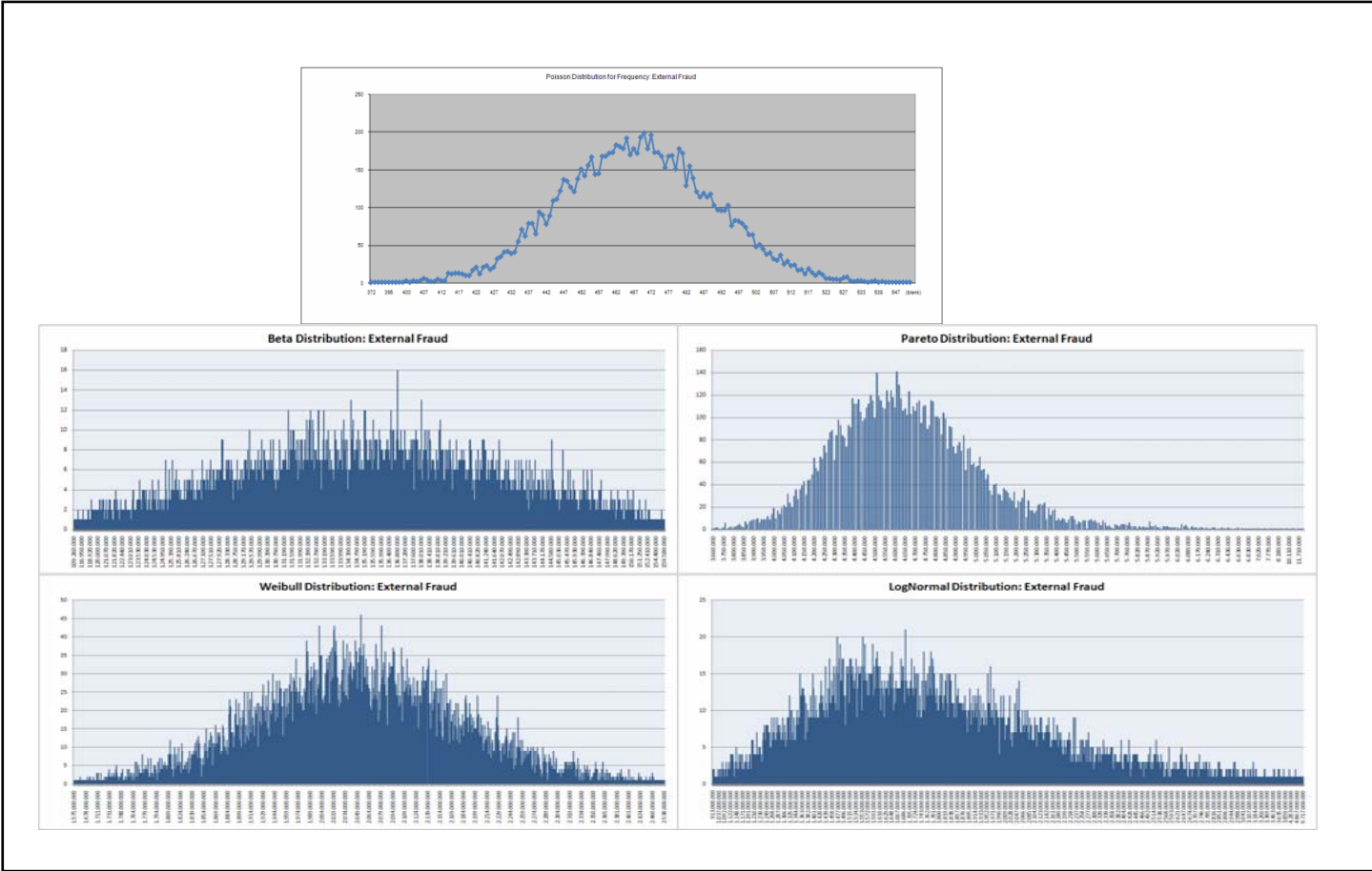
Exhibit 9.7: Internal Fraud
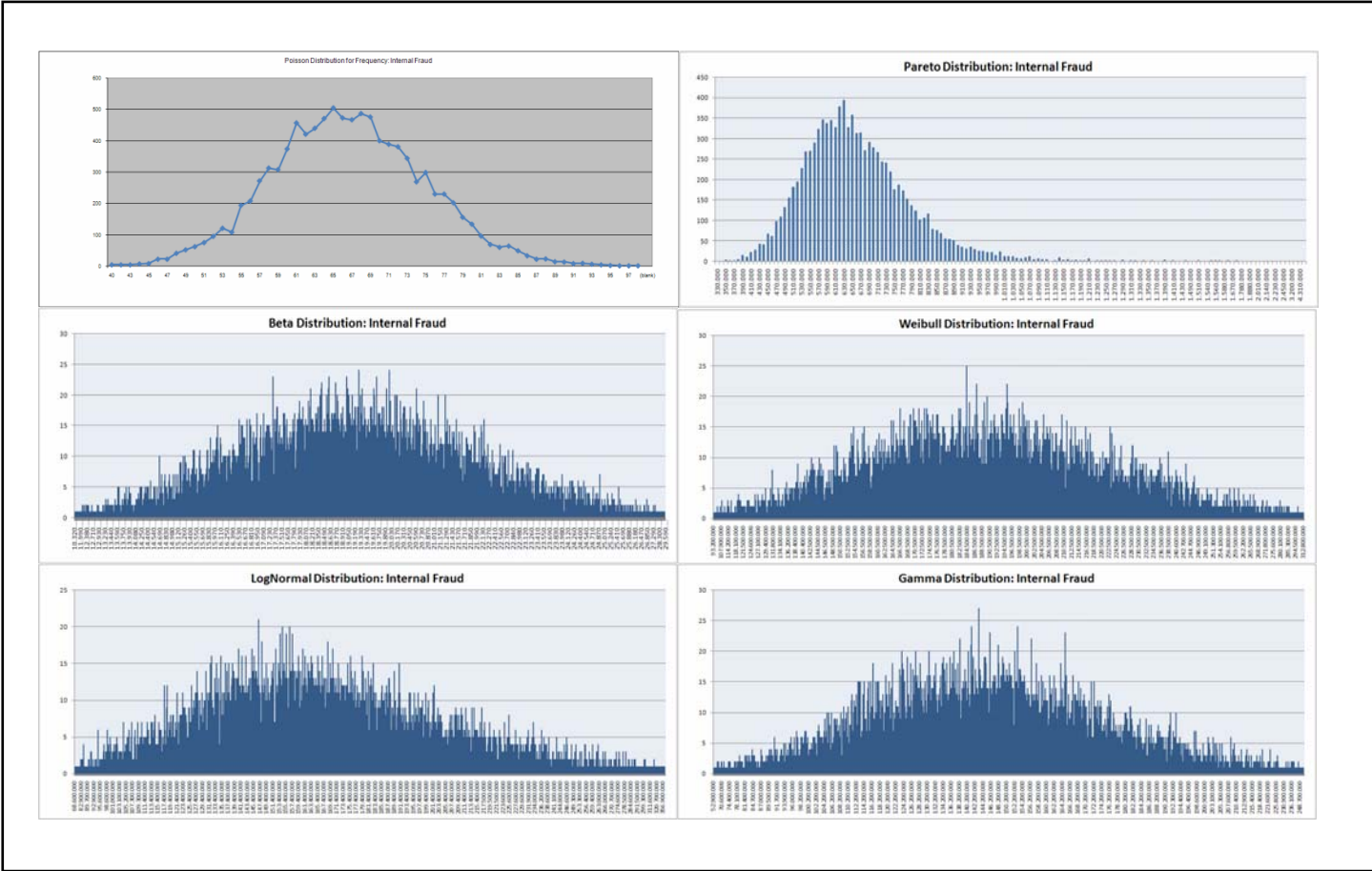
Appendix 10 2007 FED Results produced by the model (representative graphs for each distribution by Business Line

Exhibit10.1 Agency Services:
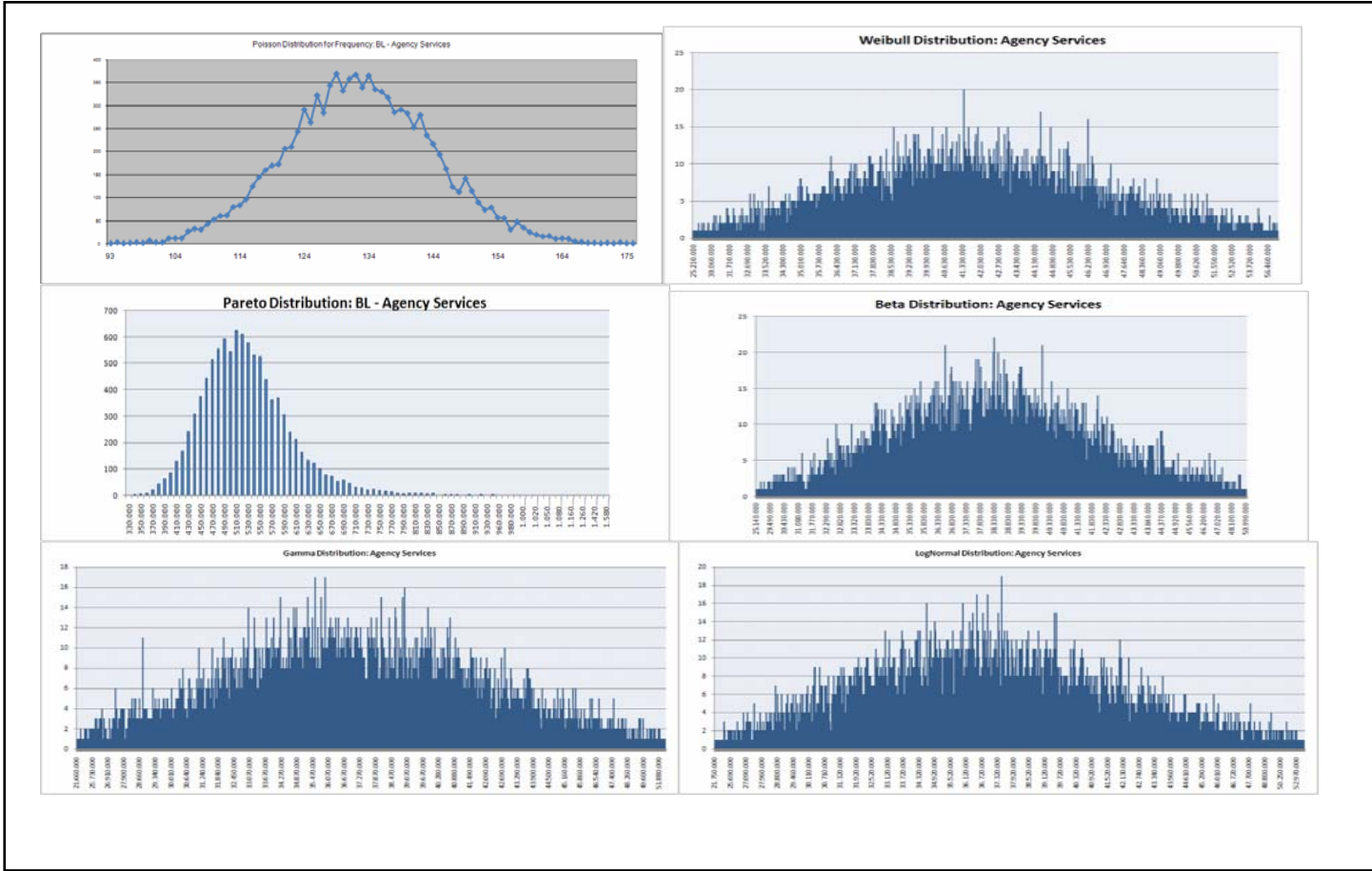
Exhibit 10.2 Asset management
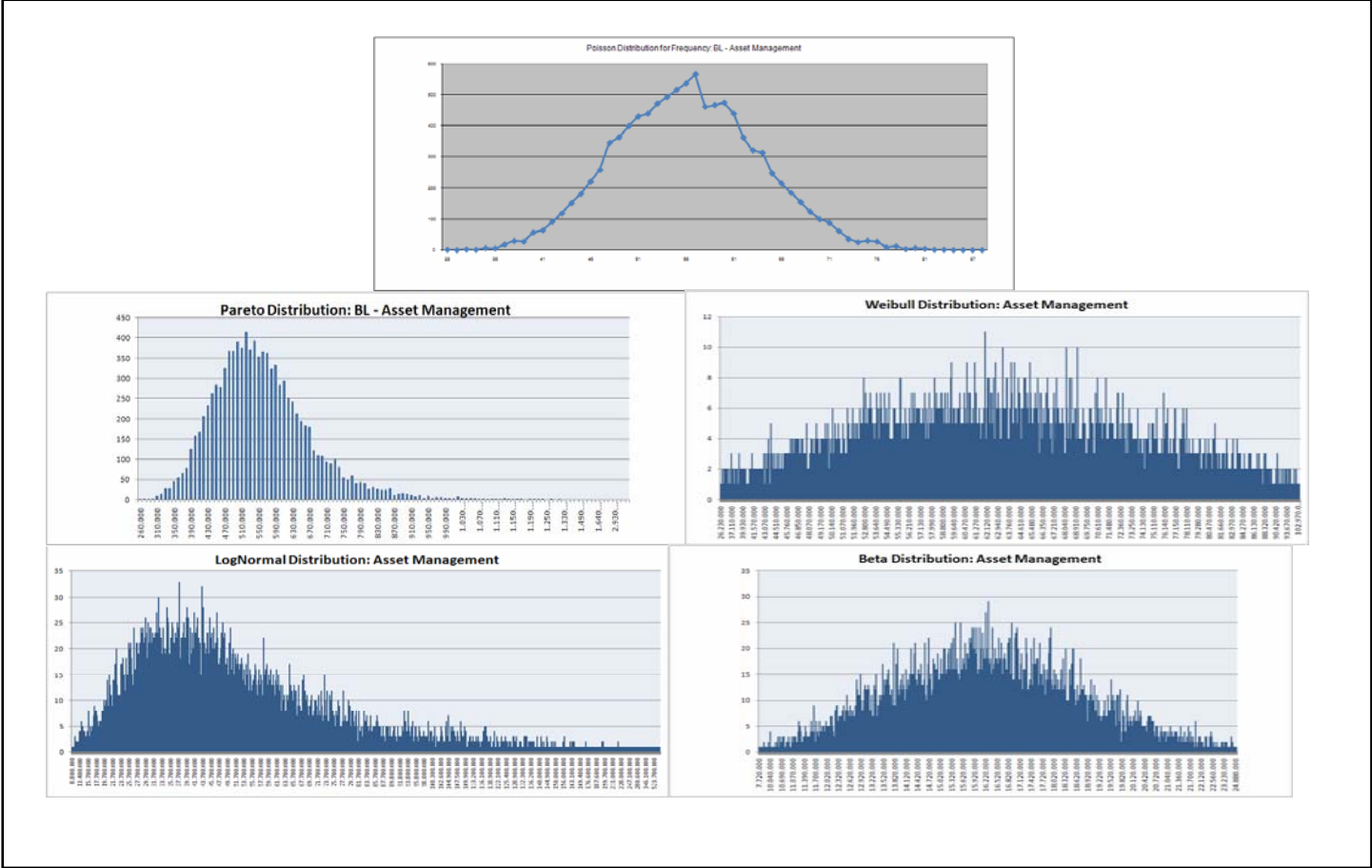
Exhibit 10.3 Commercial Banking
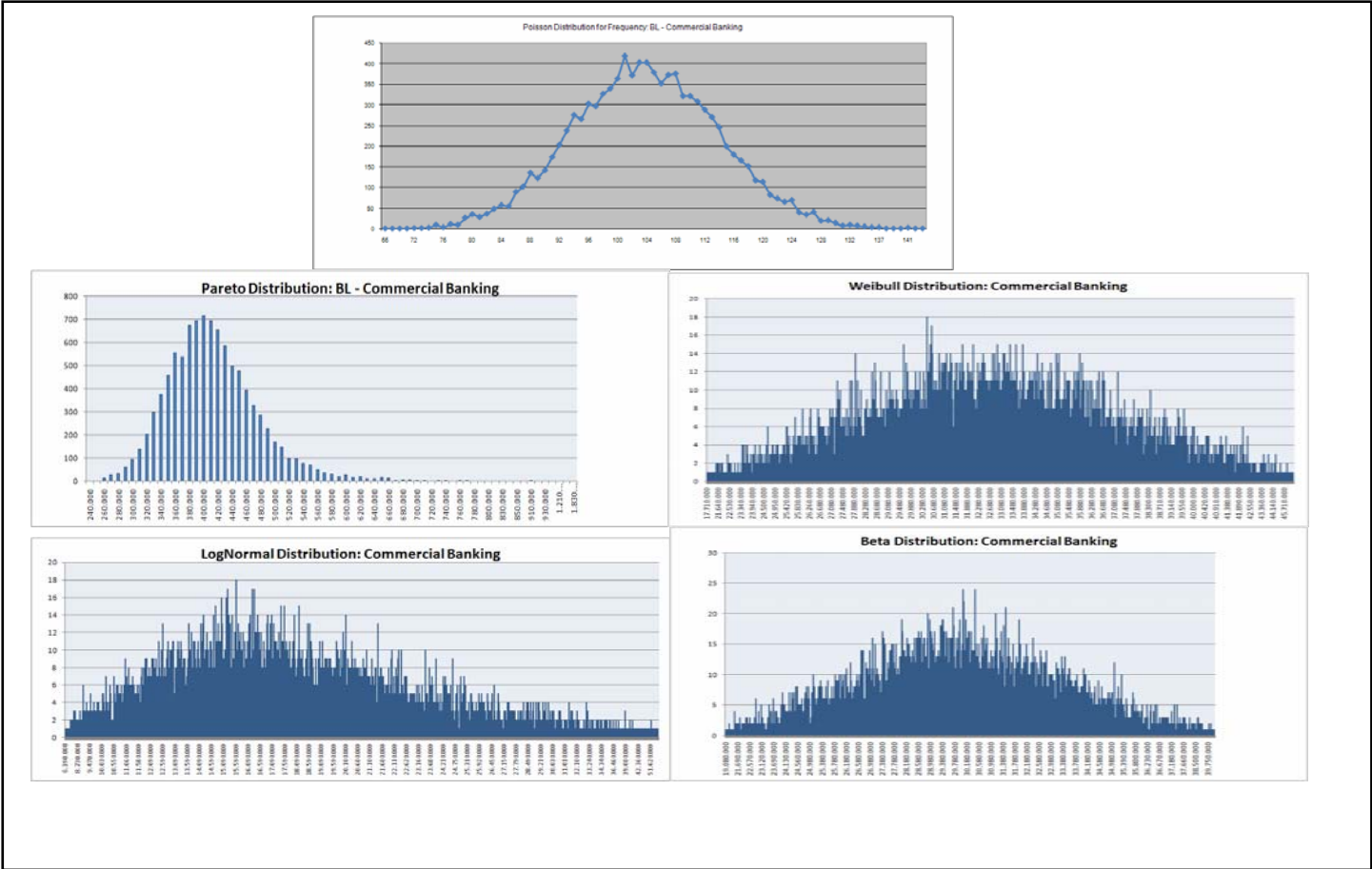
Exhibit 10.4 Corporate Finance
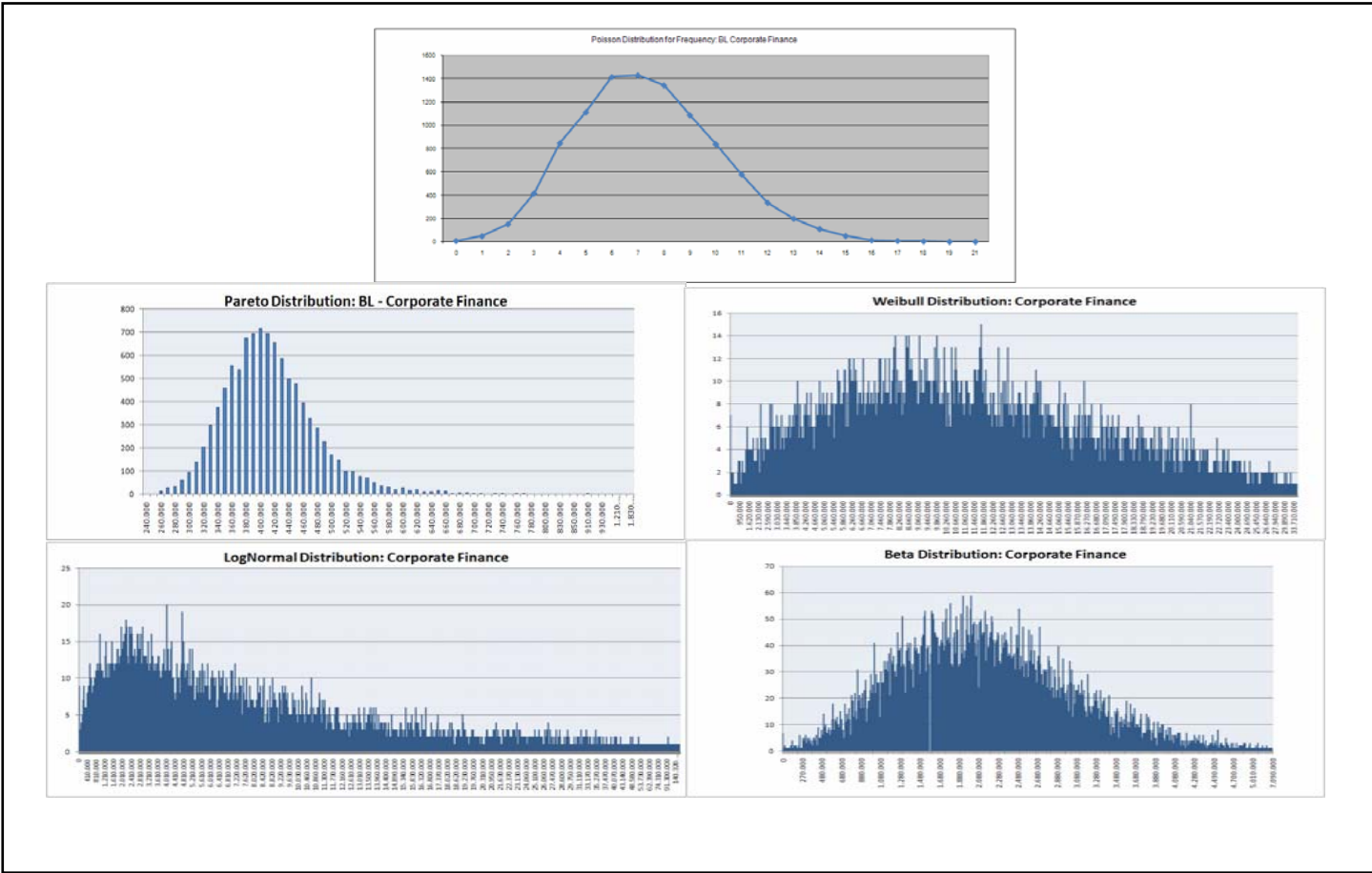
Exhibit 10.5 Payments & Settlements
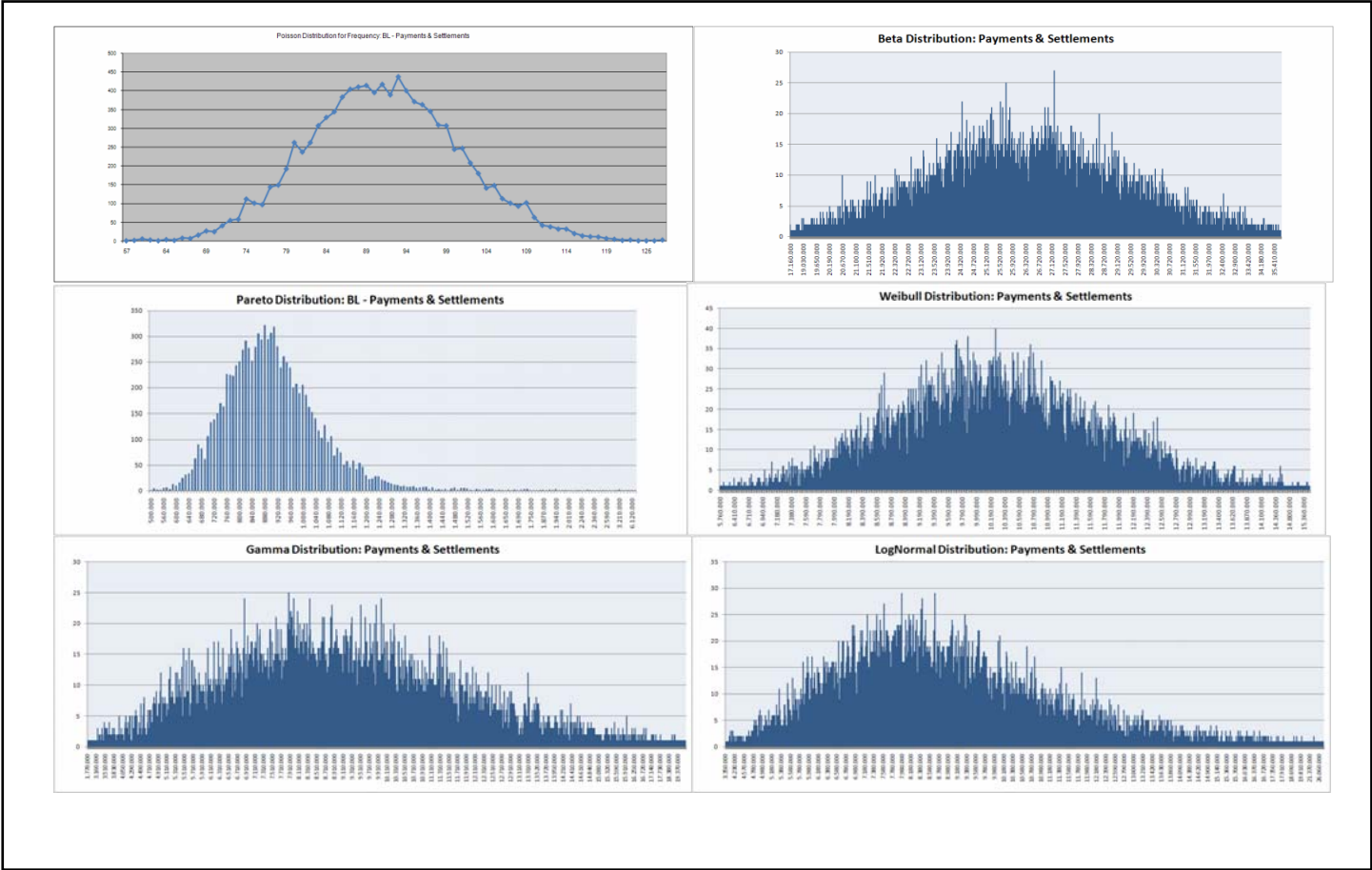
Exhibit 10.6 Retail Banking
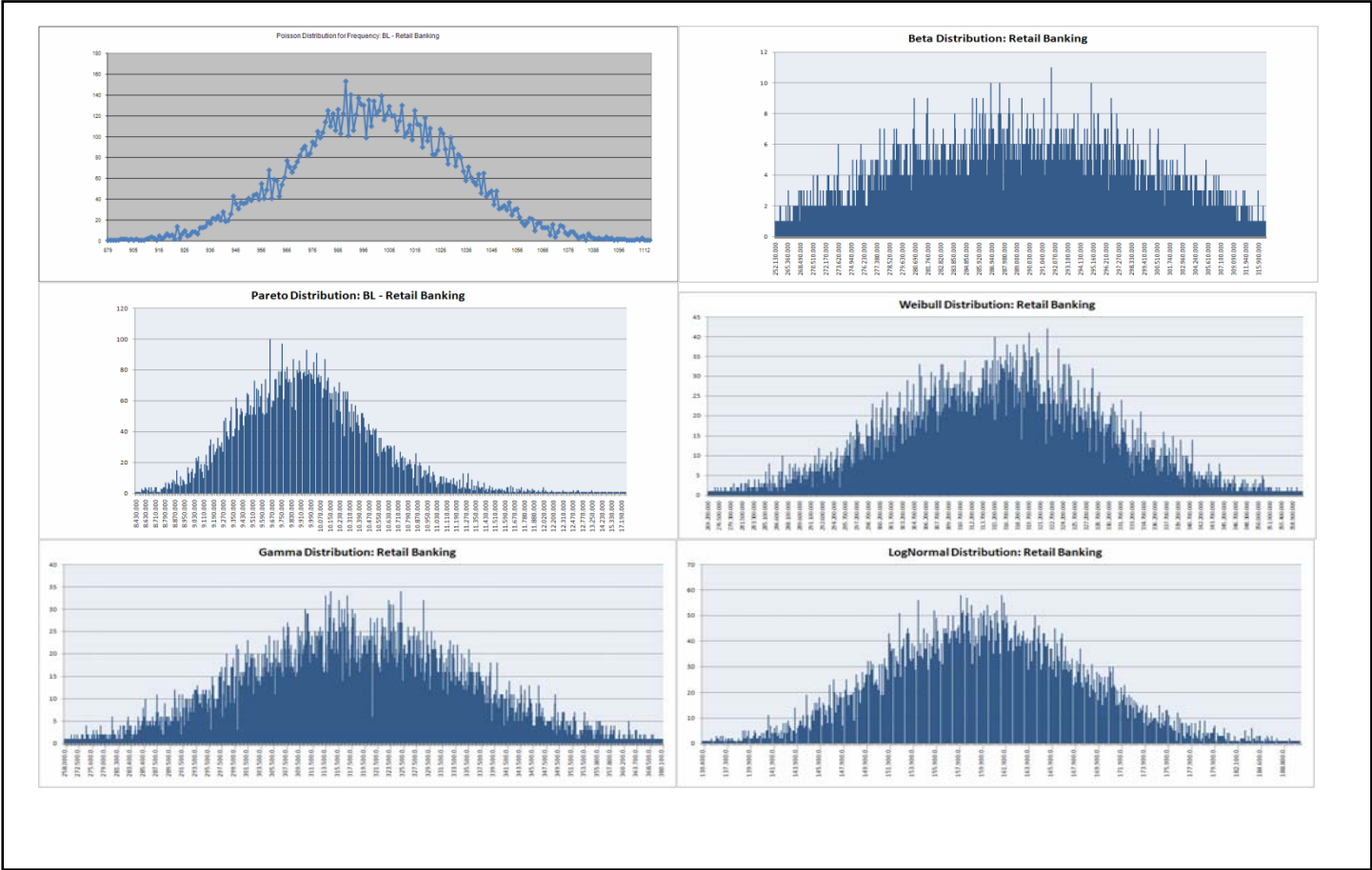
Exhibit 10.7 Retail Brockerage
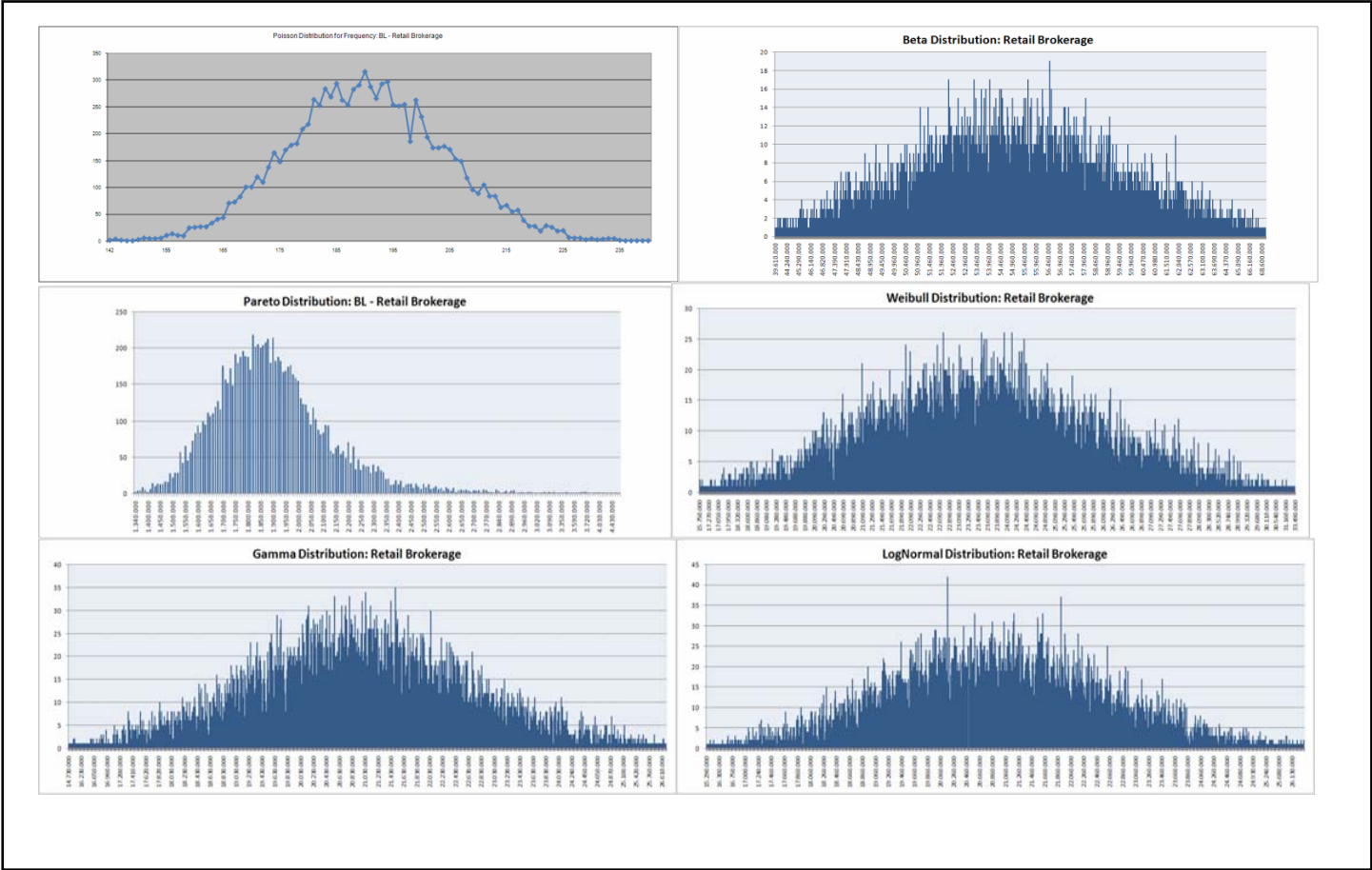
Exhibit 10.6 Trading & Sales
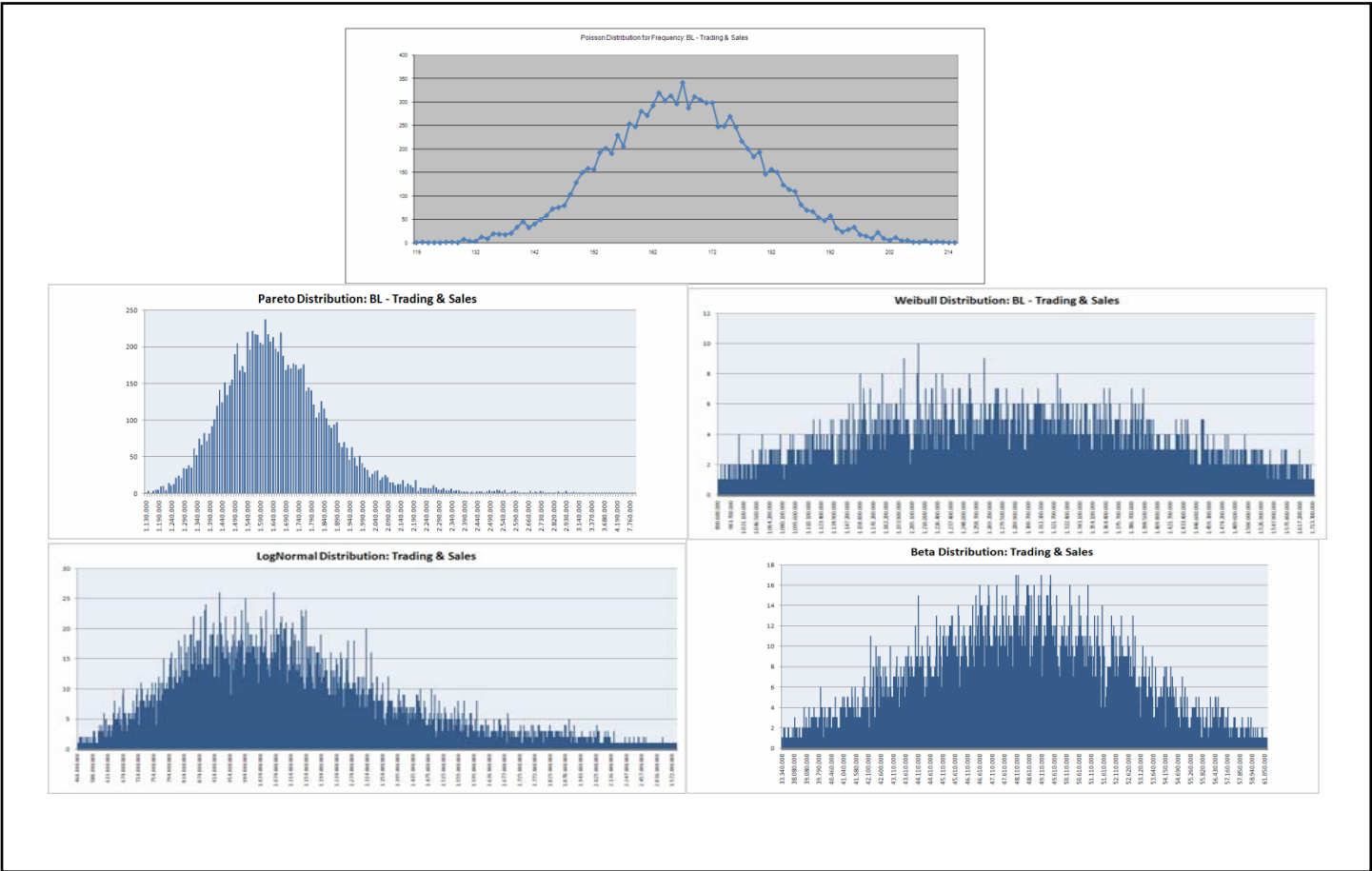
Appendix 11 2007 FED Results produced by the model (representative graphs for each distribution by Event Type

Exhibit 11.1 Business Distribution
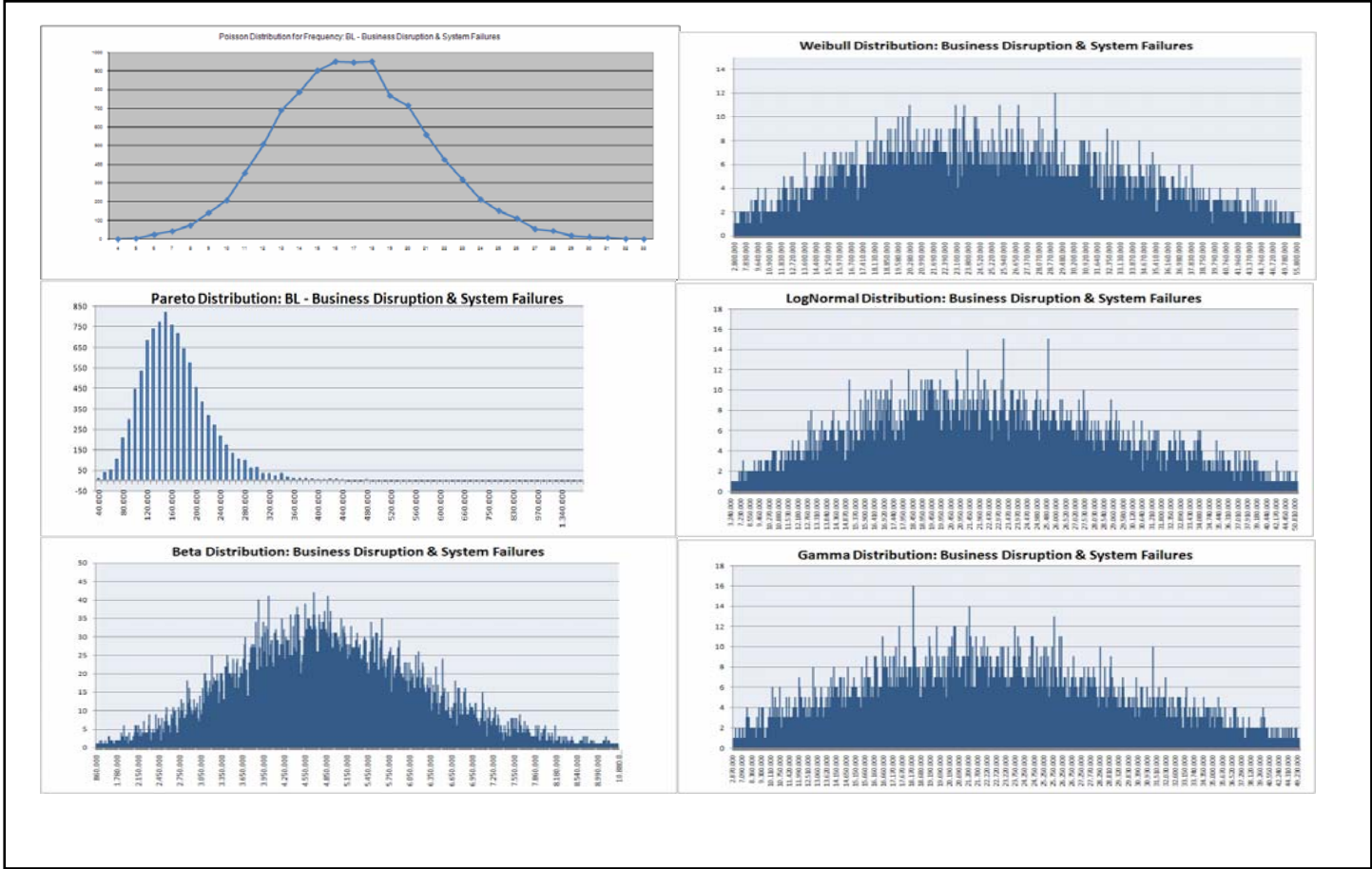
Exhibit 11.2 Clients, Products and Business Practices
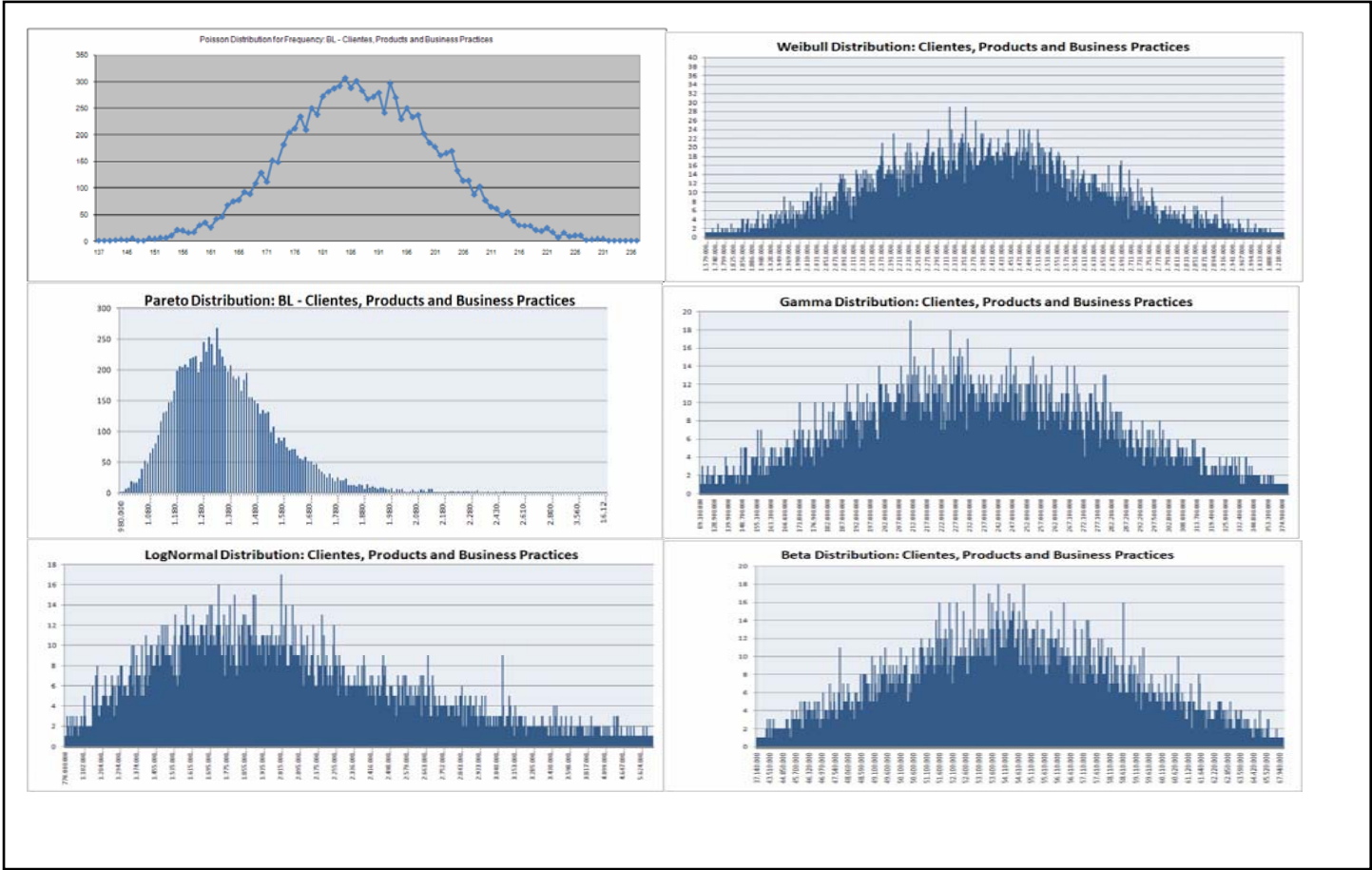
Exhibit 11.3 Damages to Physical Assets
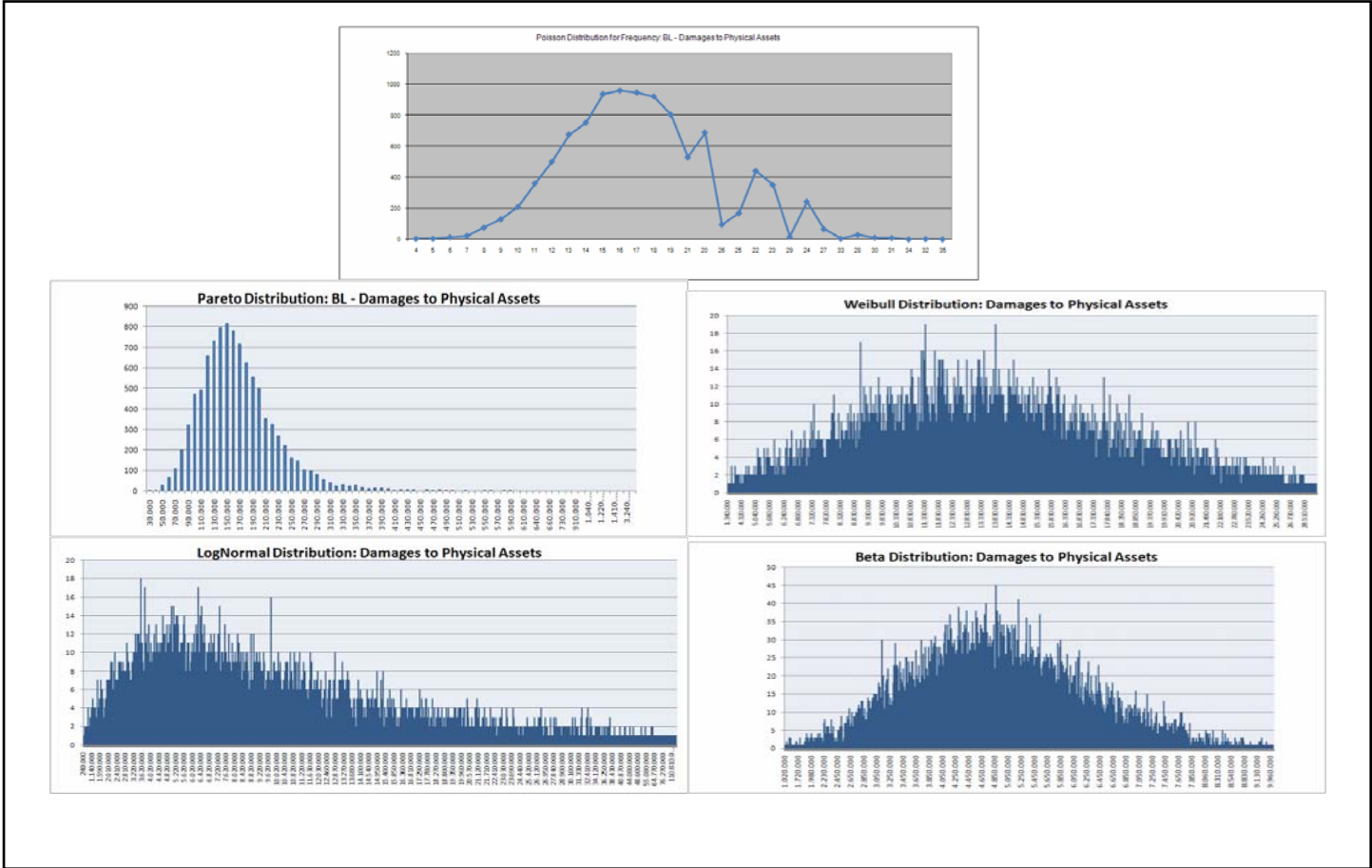
Exhibit 11.4 Employment Practices and Workplace safety
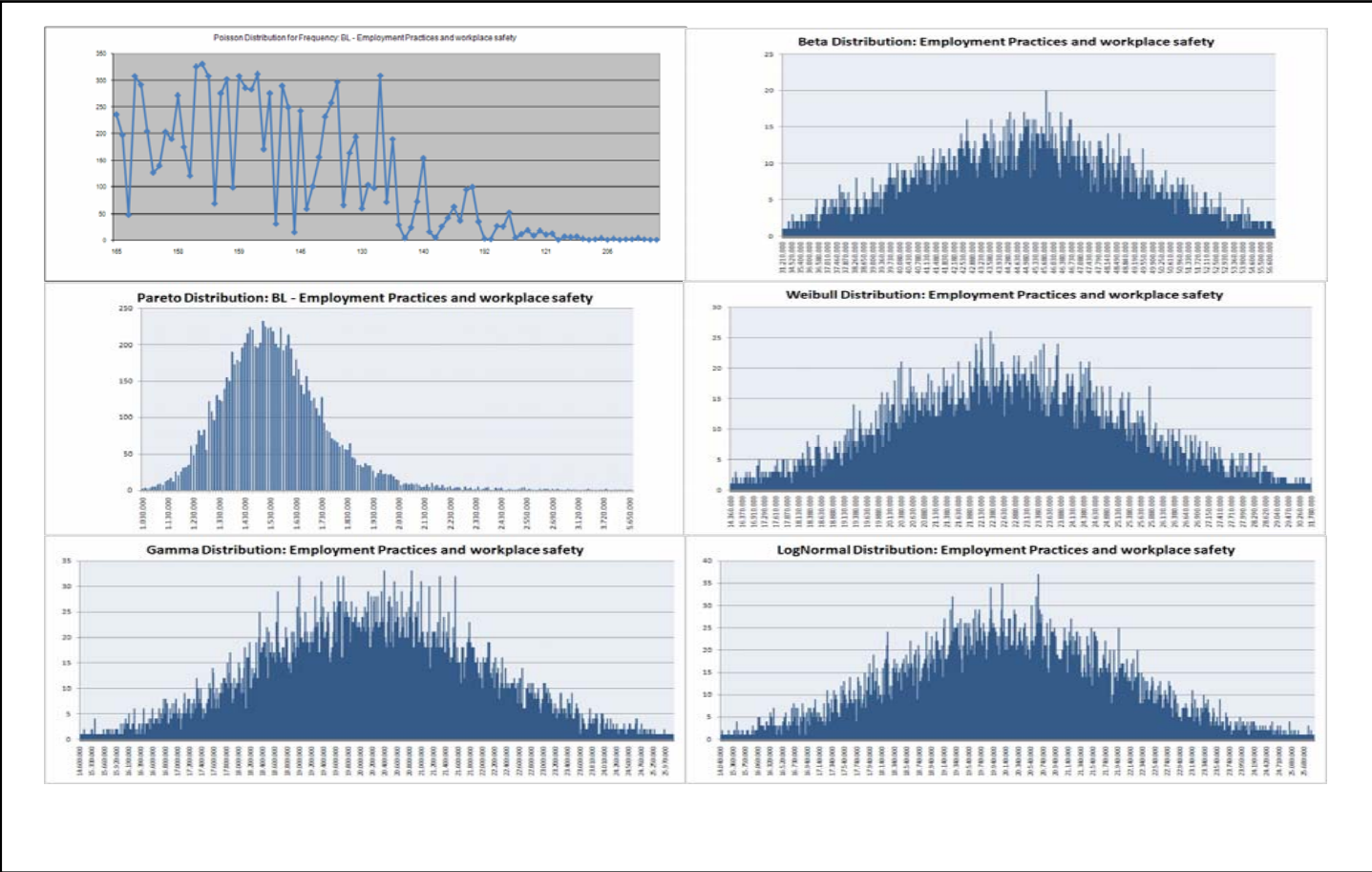
Exhibit 11.5 Execution, Delivery and Process Management
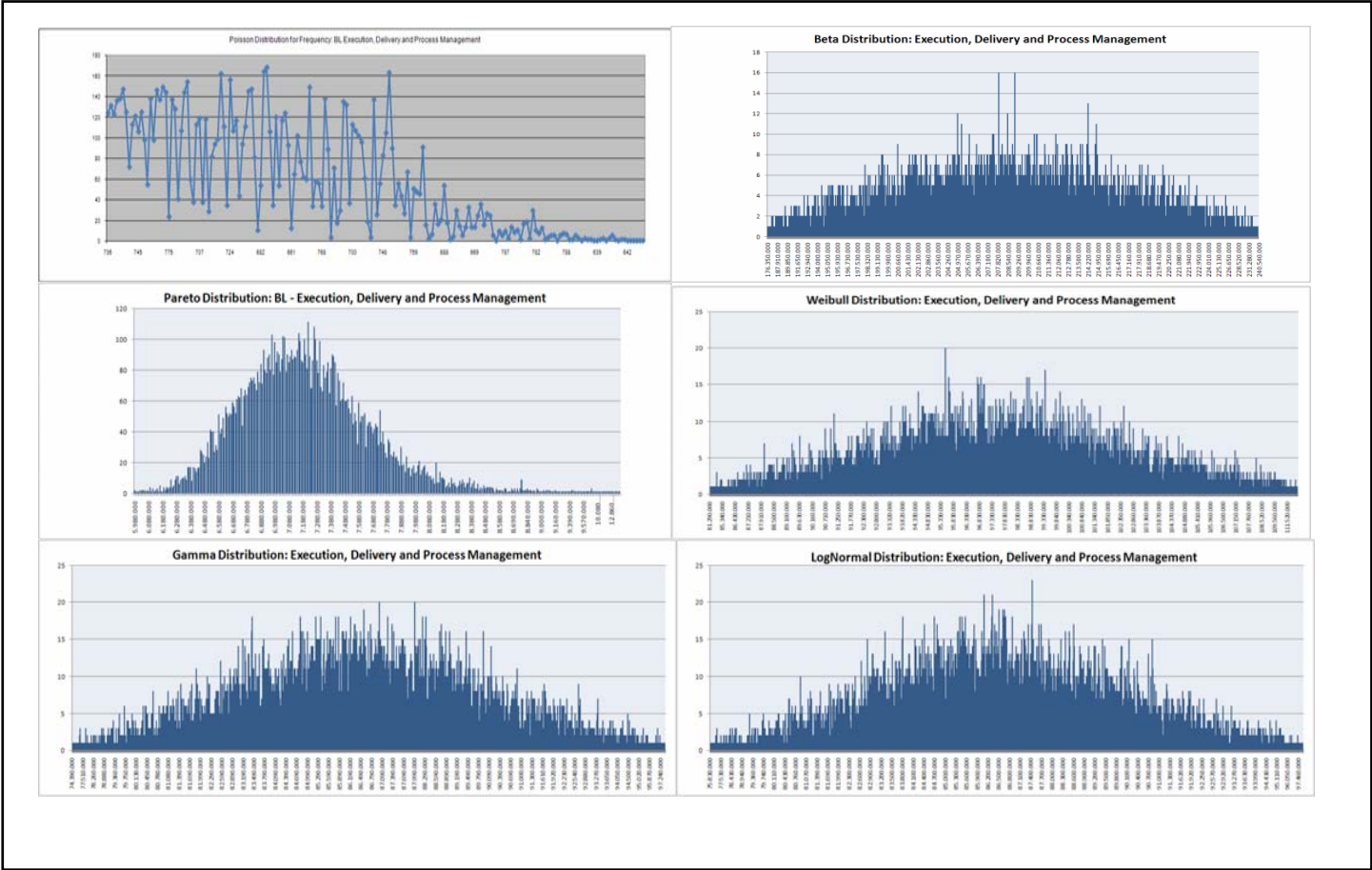
Exhibit 11.6 External Fraud
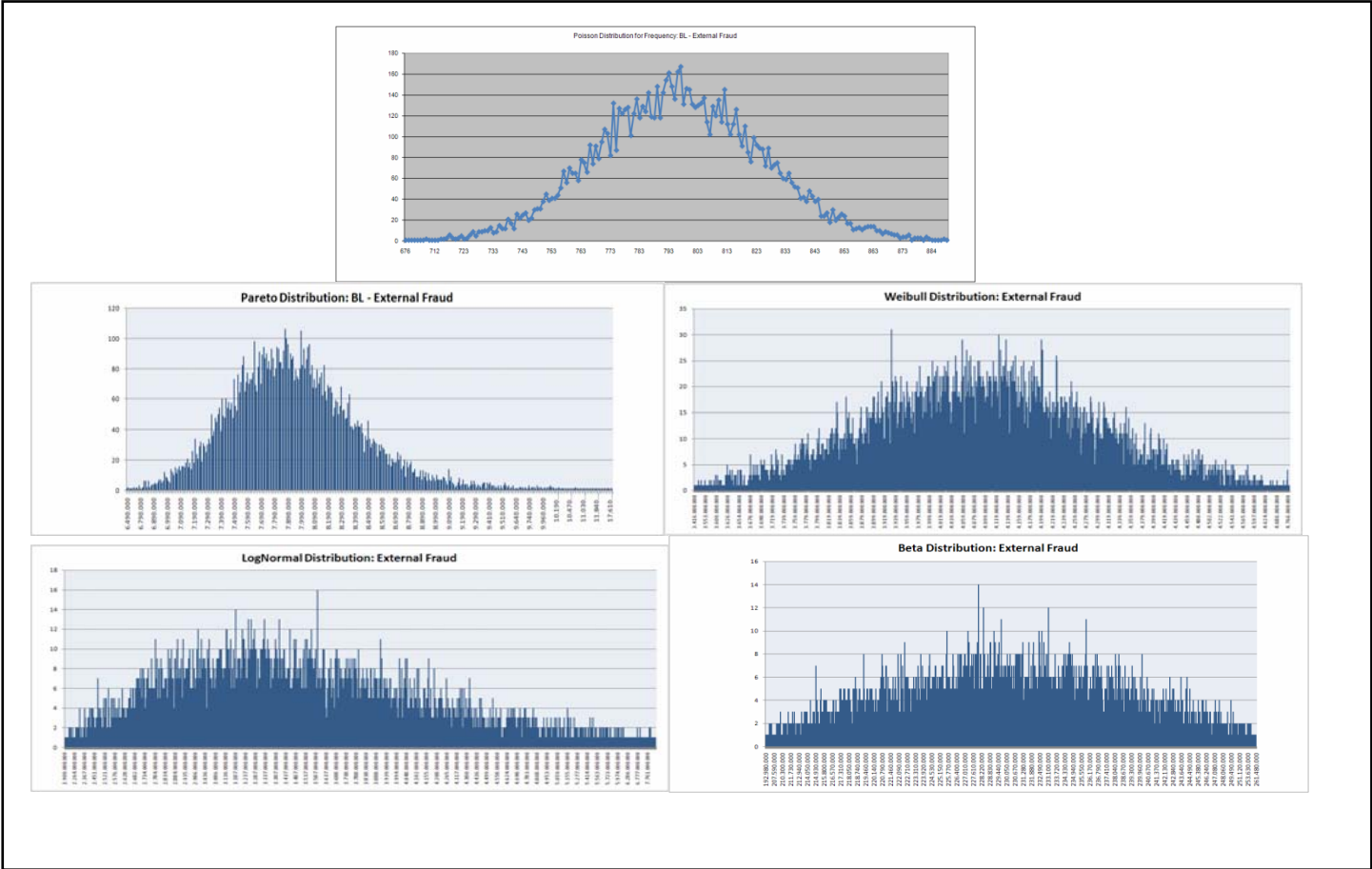
Exhibit 11.7 Fraud
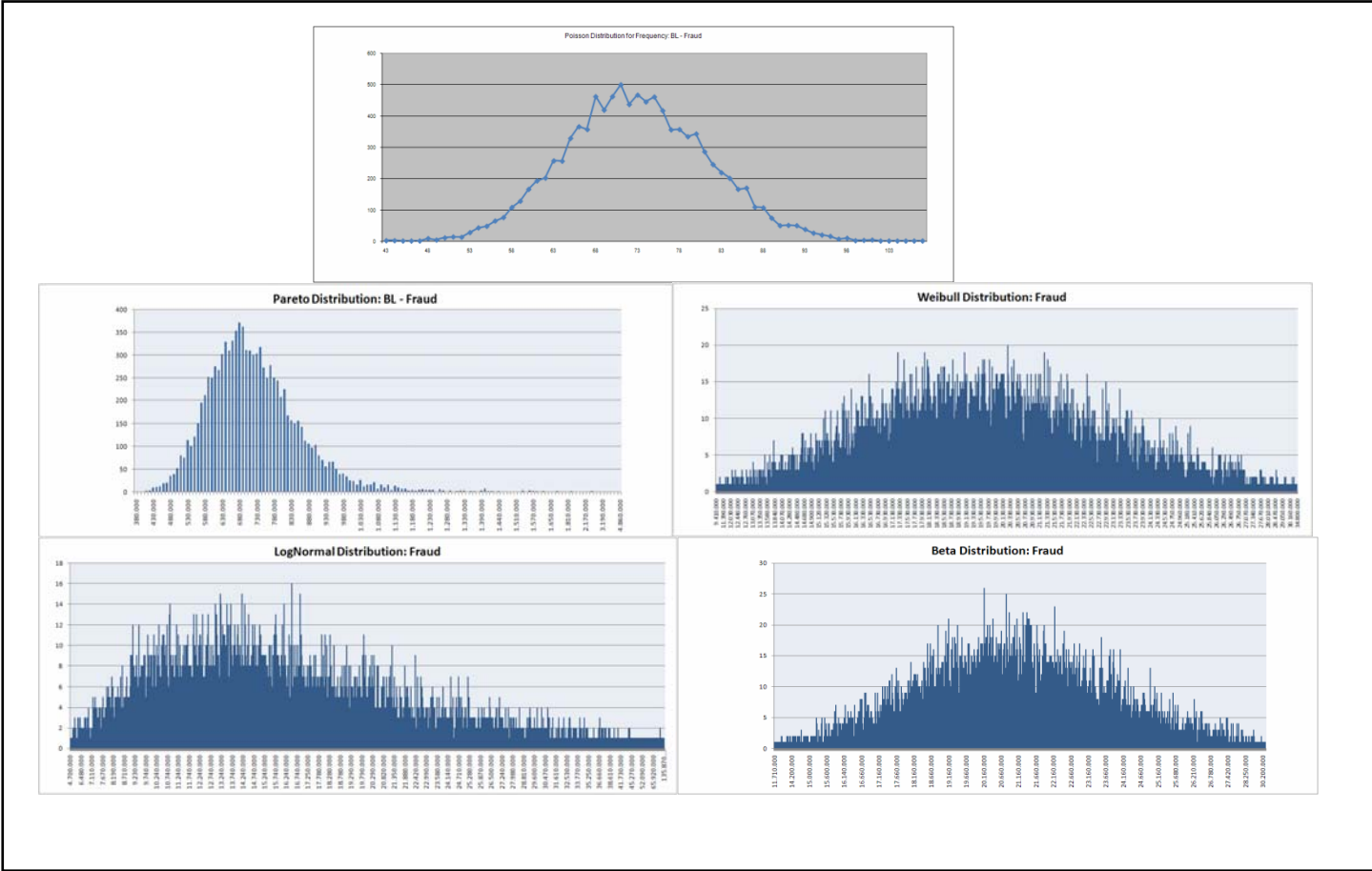
Exhibit 11.8 Internal Fraud
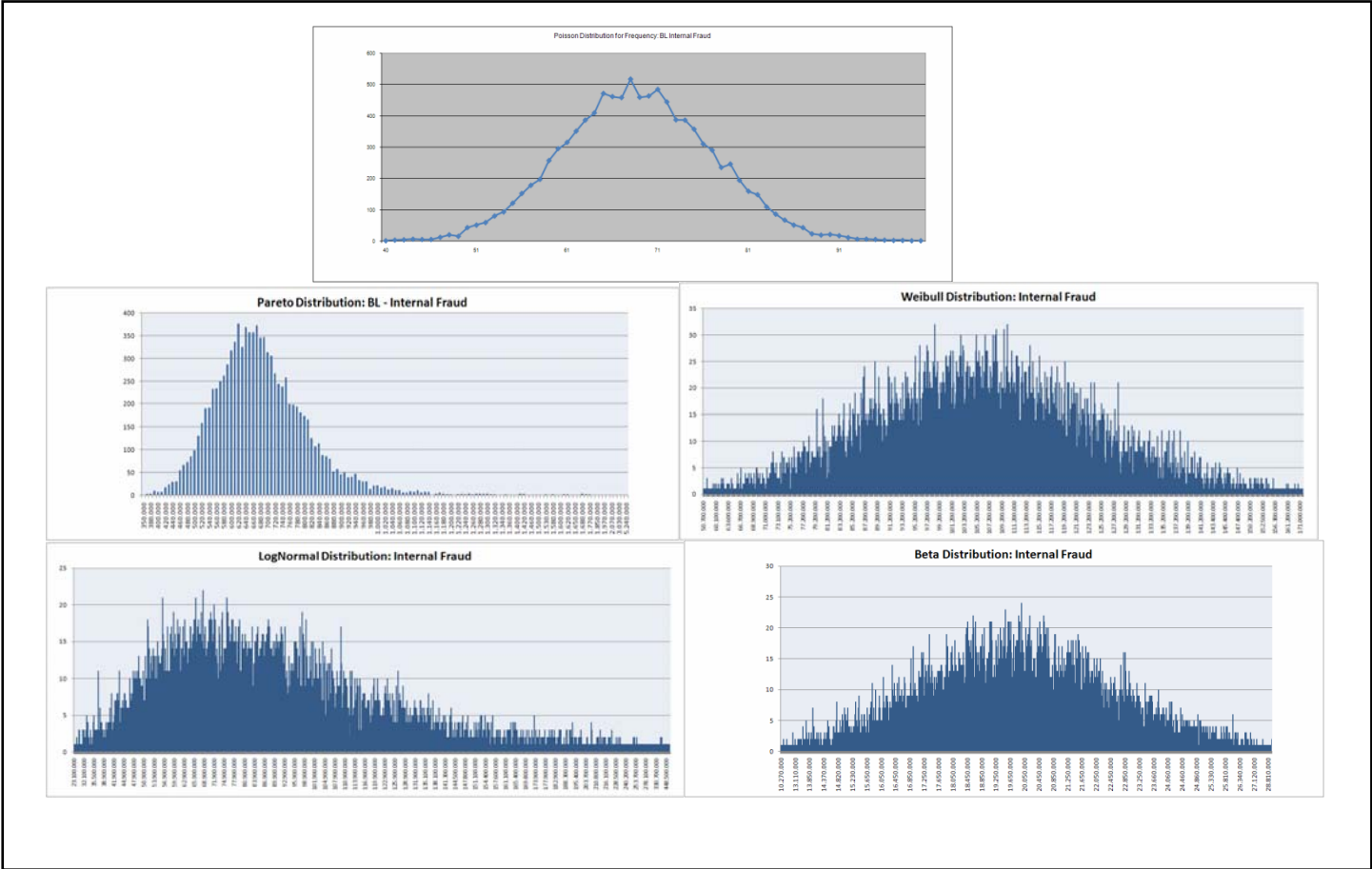
Appendix 12: 2004 LDCE FED – Results produced by the model for frequency distribution (Poisson) using Crystal Ball

Crystal Ball Report - Full
Simulation started on 8/27/2008 at 20:19:53
Simulation stopped on 8/27/2008 at 20:19:54

Run preferences:
Number of trials run — 10.000
Extreme speed
Monte Carlo
Random seed
Precision control on
Confidence level — 95,00%

Run statistics:
Total running time (sec) — 0,56
Trials/second (average) — 17.978
Random numbers per sec — 287.644

Crystal Ball data:
Assumptions — 16
Correlations — 0
Correlated groups — 0
Decision variables — 0
Forecasts — 0

Assumptions

Worksheet: [20080827.SimulacaoPoissonFED.xlsx]Sheet1

Assumption: FED – Agency Services

Poisson distribution with parameters:
Rate — 132,69

Assumption: FED – Asset Management

Poisson distribution with parameters:
Rate — 56,17

Assumption: FED – Business Disruption & System Failures

Poisson distribution with parameters:
Rate — 16,96
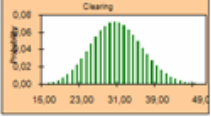
Assumption: FED – Clientes, Products and Business Practices

Poisson distribution with parameters:
Rate — 188,13
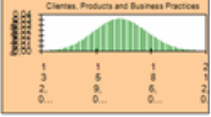
Assumption: FED – Commercial Banking

Poisson distribution with parameters:
Rate — 103,88

Assumption: FED – Corporate Finance

Poisson distribution with parameters:
Rate — 7,39
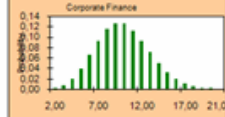
Assumption: FED – Damages to Physical Assets

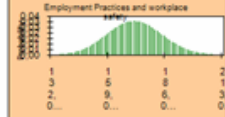Poisson distribution with parameters:
Rate — 17,06
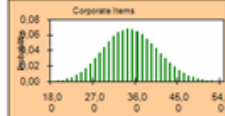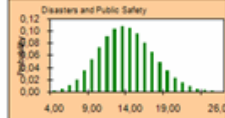
Assumption: FED – Employment Practices and workplace safety
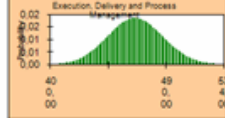
Poisson distribution with parameters:
Rate — 155,30

Assumption: FED – Execution, Delivery and Process Management
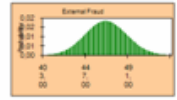
Poisson distribution with parameters:
Rate — 721,08

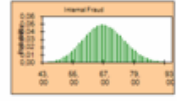**Assumption: FED – External Fraud**

Poisson distribution with parameters:
Rate — 796,10

Assumption: FED – Fraud

Poisson distribution with parameters:
Rate — 72,99

Assumption: FED – Internal Fraud

Poisson distribution with parameters:
Rate — 68,79

Assumption: FED – Payments & Settlements

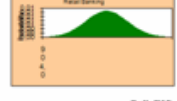Poisson distribution with parameters:
Rate — 91,14
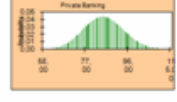
Assumption: FED – Retail Banking

Poisson distribution with parameters:
Rate — 1.000,00

Assumption: FED – Retail Brokerage
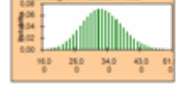
Poisson distribution with parameters:
Rate — 190,42

Assumption: FED – Trading Sales

Poisson distribution with parameters:
Rate — 166,86

End of Assumptions

Appendix 13: 2007 ORX Report – Results produced by the model for frequency distribution (Poisson) using Crystal Ball