Scenario Analysis in the Measurement of Operational Risk Capital: A Change of Measure Approach\(^1\)

Kabir K. Dutta\(^2\)
David F. Babbel\(^3\)

First Version: March 25, 2010; This Version: September 24, 2010

Abstract

Operational risk is now increasingly being considered an important financial risk and has been gaining importance similar to market and credit risk. In particular, in the banking regulation for large financial institutions it is required that operational risk be separately measured. The capital being held to safeguard against such risk is very significant at a large financial institution. As our understanding of such risk is evolving so are the methodologies for measuring such risk. While scenario analysis is an important tool for financial risk measurement, its use in the measurement of operational risk capital has been quite arbitrary and often inaccurate. The importance of scenario analysis cannot be overstated. The Federal Reserve System used scenarios to stress test the risk exposures of a financial institution during a recent financial crisis. We propose a method for the measurement of operational risk exposure of an institution using scenario analysis and internal loss data. Using the Change of Measure approach used for asset pricing in financial economics we evaluate the impact of each scenario in the total estimate of the operational risk capital. We show that the proposed method can be used in many situations, such as the calculation of operational risk capital, stress testing, and what-if assessment for scenario analysis. By using this method one could also generate a key ingredient that is a precursor toward creating a catastrophe bond on various segments of operational risk exposures of an institution. Although the method described here is in the context of operational loss, it can be used in modeling scenarios in many other contexts, such as insurance pricing, marketing forecast, and credit evaluations.

Key Words: Scenario Analysis, Operational Risk Capital, Stress Testing, Change of Measure, Internal Loss data Modeling, Basel Capital Accord.

JEL CODES: G10, G20, G21, D81

---

\(^1\) We are grateful to Ravi Reddy and Dave Schramm for valuable help and support at every stage of this work. Ravi Reddy provided several valuable insights and help with the implementation of the methods discussed here. We would also like to thank Ken Swenson who at an early stage of this work provided us with guidance from practical and implementation points of view. We are grateful to Karl Chernak for many useful suggestions on an earlier draft. Many other people enriched the development of the method discussed here by making valuable suggestions and rendering help. We have made every effort to incorporate all the advice we received. The methodology discussed here was developed three and half years ago and was presented for the first time at a US regulatory review held at a financial institution in May of 2008 and was documented in the institution’s operational risk measurement methodology document. The core method discussed here has not changed since then. However, it has gone through many reviews and thorough testing with actual scenario data from several US-based financial institutions. During this time, we did not consider writing this paper without first having tested the ideas expressed here with actual data and thinking the process through from many possible angles. Any remaining errors are ours. In the subsequent version, we would list with gratitude all the names of those individuals from whom we have received support and advice. We sincerely regret that due to the time pressure for releasing this version, we could not compile all the names.

The methodology discussed in this paper, particularly under Section 3.2, in several paragraphs of Section 3.3, and in the Appendix is freely available for use with proper citation. © 2010 by Kabir K. Dutta and David F. Babbel

\(^2\) Kabir Dutta is a Senior Consultant at Charles River Associates International in Boston. Contact email: kabir.dutta.wg97@wharton.upenn.edu

\(^3\) David F. Babbel is a Fellow of the Wharton Financial Institutions Center, Professor at the Wharton School of the University of Pennsylvania and Senior Advisor to Charles River Associates.
Introduction

Scenario analysis is an important tool in decision making. It has been in existence for several decades and has been used in various disciplines, including management, engineering, defense, medicine, finance and economics. Mulvey and Erkan (2003) is an example of how scenario data can be modeled for the risk management of a property/casualty insurance company. It is a tool that, when properly and systematically used, can bring to light many important aspects of a situation that would otherwise be missed. Scenario analysis tries to navigate the possible situations and events that can impact an entity in the future, with respect to the characteristics we are trying to measure and given the state of the entity at the present time. Therefore, by definition, scenario analysis has two important pieces:

1. Evaluation of future possibilities (future states) with respect to a certain characteristic.
2. What we know now (current states) with regard to that characteristic for an entity.

Given those two important aspects of scenario analysis, use of the scenario has to be for a meaningful duration of time, for our knowledge about the current state will change with the passage of time, making the scenarios obsolete. Also, the current state of an entity and the environment in which the entity operates, together, give rise to various possibilities in the future.

In the management of market risk, scenarios play an important role too. Many of the scenarios on the future state of an asset are actively traded in the market, which could be used for the purpose of risk management. Derivatives such as call (or put) options on asset prices are in fact instruments linked to the future possible price of an asset. Suppose for example, Cisco (CSCO) is trading today at $23 in the spot (NASDAQ) market. Yet, in the option market we find many different prices are available as future possibilities. Each of these is, in fact, a scenario for the future state of CSCO. The price for each option is a reflection of the probability that the market attaches of CSCO attaining more (or less) than a particular price on (or before) a certain date in the future. As the market obtains more information with the passage of time, derivatives prices change and our knowledge of the future state enriches. In the language of asset pricing, more information on the future state is revealed.

Scenario analysis rightfully should play an important role in the measurement of operational risk. Banking regulatory requirements stress the need to use scenario analysis in the determination of operational risk capital. Early on, many financial institutions subjected to banking regulatory requirements adopted scenario analysis as a prime component for their operational risk capital calculations. They allocated a significant amount of time and resources to that purpose. However, soon after they encountered many roadblocks. Notably among them was the inability to use scenario data as a direct input into the internal data-driven operational risk capital model. Many attempts in that direction failed miserably, as the combined effect resulted in an unrealistic capital number (e.g., 1,000 times more than the total value of the firm). Such outcomes were typical rather than rare. As a result, bank regulators relaxed some of the requirements with respect to the direct use of scenario data. Instead, regulators suggested using external loss data to replace scenario data as a direct input to the model. External loss events are historical losses that have occurred outside the institution. In many cases, such losses are very different from the loss experience of the institution. In that process the importance of scenarios, in our opinion, was reduced in measuring operational risk capital. Before that, and as a current practice also, external loss data were used in the generation of scenarios.

We believe that the attempts to use scenario data directly in the capital model have failed because of incorrect interpretation and implementation of such data. This work is an attempt to address and resolve
such problems. Our preliminary research shows that the use of external data (i.e., loss experience data from other similar organizations) for measuring operational risk without proper adjustment through scenarios may serve to increase systemic risk. We plan to address this issue in detail in a later publication. Given the use of and success with scenarios in many other disciplines, we would like to think that scenario analysis should be treated as importantly as any other data that an institution may consider for its risk assessments. Some may question, justifiably, the current state of the quality of scenario data and whether such data can be believable. We would like to argue that such is the nature of the beast in every discipline. As we will show, the value in scenario data outweighs the inherent weaknesses it may have. Also, through the process of systematic use we will be able to enhance the quality of the data.

The major contribution of this work is in the correct interpretation of scenario data, consistent with the loss experience of an institution, with regard to both the frequency and severity of the loss. Then, using the correct interpretation, we show how one could effectively use scenario data together with historical data to measure operational risk exposure and, using the Change of Measure concept, evaluate each scenario in terms of its effect on the operational risk measure.

In the next section we discuss why some of the earlier attempts at interpreting scenario data did not succeed and the weaknesses of current practices. We then discuss the nature and type of scenario data we will be using for our models. Following that, we discuss the method we use to model scenario data and show how to evaluate the scenarios using the Change of Measure approach. We conclude the paper with a discussion on some of the issues one may encounter in implementing the method and use of this method in other domains.

1. **Description of the Problem**

To clarify, we will use operational risk measurement methods commonly used at financial institutions subject to Basel banking regulations. Institutions are required to use four data elements in their operational risk capital model: internal loss data, which are collected over a period of time and represent actual losses suffered by the institution; external loss data, which are loss events that have occurred at other institutions and are provided to the institution via a third-party vendor or from a data consortium; scenario data based on the assessment of possible losses an institution may experience in the future; and business environment score, created on the basis of qualitative assessment of the business environment and internal control factors. Using these four data elements, an institution is required to measure its operational risk. The regulatory rule is not prescriptive in terms of how the data elements should be used. However, given the similarity of operational losses to property/casualty losses in the insurance world, the measurement approach predominantly followed is the loss distribution approach that actuaries use for pricing property/casualty insurance.

Of the four data elements, internal loss data is used primarily in a loss distribution approach to arrive at a base model. In a loss distribution approach, one tries to fit two distributions: the severity distribution, which is derived from the loss amounts of all the losses experienced by the institution; and the frequency distribution, which is derived by using number of losses that have occurred at the institution over a predetermined time horizon (typically one year). With respect to the frequency distribution, the Poisson distribution is the choice of nearly all financial institutions. Typically the Poisson parameter is determined by calculating the average number of losses on an annual basis. The loss distribution is obtained by the convolution of the severity distribution with the frequency distribution. In other words, a loss event $L_i$ (also known as the loss severity) is an incident for which an entity suffers damages that can be measured with a monetary value. An aggregate loss over a specified period of time can be expressed as the sum:
where $N$ is a random variable that represents the frequency of losses over the period. We assume that the $L_i$ are independent and identically distributed, and each $L_i$ is independent from $N$. The distribution of the $L_i$ is called the severity distribution, the distribution of $N$ over each period is called the frequency distribution, and the distribution of $S$ is called the aggregate loss distribution. This framework is known as the Loss Distribution Approach (LDA). The risk exposure can be measured as a quantile of $S$.

Given the characteristics and challenges of the data, we can resolve many issues by using an LDA approach. The sum $S$ can be calculated by either Fourier or Laplace transforms as suggested in Klugman et al. (2004), by Monte Carlo simulation, or by an analytical approximation. We have used the simulation method as well as the analytical approximations method here. The LDA has been exhaustively studied by actuaries, mathematicians, and statisticians well before the concept of operational risk came into existence. Regulatory or economic capital for the operational risk exposure of an institution is typically defined as the 99.9 percentile or 99.97 percentile of the loss distribution (the distribution of $S$). The choice of a severity distribution is not trivial. Dutta and Perry (2007) is a good source for the discussion on the severity distributions based on internal loss data.

Many financial institutions have been collecting internal loss data for several years. These internal loss data can be considered the current state for operational risk exposure for a financial institution. Additionally, there are many losses of various types and magnitudes that have occurred at other financial institutions but may not have happened at a particular institution. A financial institution may choose to evaluate such external loss data in order to understand the potential impact such losses might have on its own risk profile. Typically, an institution will analyze those losses based on the appropriate magnitude and probability of happening, given the current state of the institution’s risk profile.

Suppose institution A observes that institution B has incurred a $50 million loss due to external fraud, a type of operational loss. Institution A is also aware of the circumstances that were present at institution B in order for such a loss to occur. After evaluating its own circumstances (current state), institution A determines that it is likely to experience a similar loss at the rate of once every ten years and that such an event would result in a $20 million loss. These are, in fact, the frequency and severity of an event in the future state. Alternatively, the institution could have come up with a range, such as $15 million to $25 million, for the magnitude instead of one number, $20 million. We will discuss this issue further in a later section of the paper.

1.1 Inaccurate Data Pooling

Suppose an institution has collected internal loss data for the last five years. It also generates a scenario for a certain operational loss event that has the likelihood of occurring once in ten years, resulting in a $20 million loss. Interpreting this as just another data point that should be added into the pool of all the internal loss data the institution has observed thus far will be inaccurate. In so doing, the institution will be increasing the scenario data point’s frequency to once in five years from once in ten years. This key insight of our research led to our development of a method to appropriately integrate scenario data with internal loss data. The most common problem reported with integrating scenario data with internal loss data is that the result is an unrealistic capital estimate. This is because the integration process failed to consider the frequency of the scenario data; those unrealistic capital numbers are in fact the consequence of making the adverse scenario data more frequent than it should be. This works the other way too. Suppose the scenario is defined as a $20 million loss with a possibility of occurring once in 2.5 years. By adding this scenario to the pool of internal loss data collected over five years, the institution is diluting the effect of the scenario data. In such a case, simplistically speaking, two such scenario data points should be
added to the pool of five years worth of internal data to approximately calculate the effect of the scenario. Frequency must also be considered when interpreting scenarios across financial institutions.

Unit of measure is the level or degree of granularity at which an institution calculates its operational risk capital. The least granular unit of measure is enterprise-wide. It is more typical, however, for institutions to calculate operational risk capital for several units of measure and aggregate the capital estimates. Units of measure are often determined by business line or loss event type. Smaller business lines and/or less common loss event types are frequently pooled together to create one unit of measure.

Suppose that for the same unit of measure two different institutions have the exact same scenario of a $20 million loss occurring once in 10 years. One institution has an annual loss frequency of 20 and the other has 50. For the institution with 20 losses per year the scenario is much more intense than the institution with 50 losses per year. The method we discuss later will use this insight to properly align the frequency of the scenario data with the time horizon of the internal loss experience. The method we describe is essentially a Change of Measure approach similar in concept to the process used in asset pricing in the sense that current measure in terms of probability of the internal loss data is changed using the severity and frequency given in the scenario data. We would like to note here that the Change of Measure is an evaluation metric. As is done in asset pricing, we will use this metric to validate the usefulness of the scenarios and values in risk measurement. Remember, for regulatory purposes, valuation of risk for operational risk is equal to 99.9% or 99.97% of the aggregate loss distribution. Alternatively we can call it Capital or Price for the risk.

Continuing with our example above of a $20 million loss occurring once in ten years, in order to merge this scenario with the existing internal data, we will have to consistently recreate internal data for a period of ten years from the five years of data that the institution observed. Only then can we merge the $20 million scenario loss with the internal data, and only if such a loss has not already been observed with sufficient frequency in the internal data. In other words, we are using the current state of five years of observed internal loss data to generate enough data to merge the severity of the scenario data at the frequency given in the scenario data.

To summarize, we are dealing with two different states – current and future. The future state has not been revealed to us and we are guessing the future with respect to our knowledge of the current state and the estimated possibility of the future.

For our purpose, the severity and frequency distributions for a unit of measure estimated using internal loss data summarized the current state. The severity distribution in essence assigns a probability for each loss event. Internal loss data are ideal for estimating an institution’s current state. For this method to work effectively, it is important that this current state (severity and frequency distribution) is estimated accurately. We may have many choices for the severity distribution fitting internal loss data. Naturally, one would like to consider the best fit based on model selection criteria as the candidate for current measure. More often than not, it is not absolutely clear which model to choose or it may happen that one model is marginally better than the other. In that case it will be advisable to use all those which could be a possible good fit. This is particularly true when an institution did not experience many losses for a particular unit of measure. This will ensure more stability in the measurement of the current state, which is an important cornerstone for this method. The frequency distribution, on the other hand, tells us how many losses would need to occur before one would likely observe a loss of a given magnitude. One perhaps needs to repeat the draw many thousands of times in order to reduce sampling error.

Given this current state, we would like to predict the probability of an event happening in the future. On the other hand, scenario analysis data also implicitly predict the probability of the same event happening. Here event is defined as a subset of real line $\mathbb{R}$, more specifically $\mathbb{R}^+$. 
More than likely, these two probabilities will not match. Therefore, the probability of the current state must be adjusted in such a way that it accounts for the scenario probability. This is necessary in order to calculate the equivalent distribution if the scenario loss actually happened with frequency specified in the scenario, proportional for the specified period of time. If scenario data lower the probability compared to the internal loss data, then for practical reasons to be discussed later we do not alter the probability predicted by the internal loss data.

Borrowing the terminology from the Black-Scholes option pricing concept, we can call probability implied by the scenario as implied probability, whereas we call the probability implied by the current state estimated using historical losses historical probability. We will explore this further in the methodology section.\footnote{Cochrane (2001) is good source for understanding the pricing of contingent assets and valuation using Change of Measure.}

### 1.2 Measurement Practices using Scenario Data

The Bank of Japan hosted a workshop on operational risk scenario analysis in 2006.\footnote{One can get more information on the presentations and discussions held in that workshop at: http://www.boj.or.jp/en/type/release/zuiji_new/6scr0608a_add.htm. We found it to be a very valuable workshop in addressing the issues related to the use of scenario analysis in measuring and managing operational risk.} The presentations of Nagafuji (2006), Oyama (2006), and Rosengren (2006) adequately capture and summarize the problems with and the art of using scenario analysis for operational risk assessment. The issues discussed in those presentations are still very valid, four years later. In fact, we would argue that since then, there has been very little, if any, focus on the development of scenario-based methodology for operational risk assessment. While much research has focused on finding a severity distribution for fitting internal or external loss event severity data, we are aware of no work that has systematically studied the problems related to integrating scenario analysis data into an institution’s operational risk capital calculation. Yet scenario data, despite their problems, are the essential elements of information that should be taken seriously in the measurement of operation risk.

We could not find in the literature a good and credible method to which we could compare the method we are presenting here. Broadly speaking, we did find only a few methods for integrating scenario analysis data into the operational risk capital calculation currently in use. All of these methods are essentially very ad hoc and mostly integrated with internal or external data without sound theoretical justifications.

The scope of this work is non Bayesian in nature. We do not know of any financial institution that has either a viable Bayesian implementation for internal loss data modeling or is collecting scenario data suitable for Bayesian implementation. The Bayesian concept may be very useful and appears very promising in this context. We would like to see more research in this area. However, without observing a practical implementation of this concept, it is not possible for us to comment on the suitability of such an approach. Therefore, we decided not to include a Bayesian approach in our comparative analysis. We will revisit this in the next section in the context of scenario data generation and will discuss some practical problems for the scenario data generation for the Bayesian implementation.

In one method\footnote{The methods described are not published but observed in practice. Financial institutions have implemented similar methods.} the severities from all the scenarios are pooled together and then a set of uniformly distributed random numbers in the interval of zero to one are drawn several times. The number of uniform random numbers drawn in each trial is equal to the number of units of measure the institution is using for internal or external loss data modeling. If the random number drawn matches with the probability assigned to the scenario for that unit of measure, then the corresponding severity is chosen and added to other severities chosen in that particular trial. Otherwise, zero is added for that unit of measure in that trial. The severity amounts for each trial are summed and a 99.9% or 99.97% level quantile is chosen.
This number is then compared with the 99.9% or 99.97% number from the loss distribution obtained using internal or external loss data and then the institution must decide which number to use for regulatory capital. Typically the scenario-based number will be much higher than the internal or external data-based number. In such cases, a number in between the range is chosen as the 99.9% or 99.97% quantile. Rarely, the scenario-based 99.9% or 99.97% level number would be added to the corresponding number obtained using internal or external loss data to provide an estimate of extreme loss. This method suffers from the drawback that the universe of potential severe loss amounts is limited to the severity values assigned to the scenarios. This is very similar to fitting loss data to an empirical distribution. Dutta and Perry (2007) highlights some of the problems related to using an empirical distribution. In this method, scenario data are completely isolated from internal and external loss data.

In another method, two types of severity numbers are derived from the one scenario created per unit of measure. The first figure is the most likely severity outcome for the scenario and the other represents the worst severity outcome. Then a purely qualitative judgment is made to interpret these two severity values. The worst-case severity outcome is put at the 99th percentile (or higher) of the severity distribution obtained from internal or external data for that unit of measure, and the most likely severity outcome is put at the 50th percentile level of the severity distribution. The 99.9% or 99.97% level number is obtained from the loss distribution after recalibrating the severity distribution with these two numbers. As in the previous method, the 99.9% or 99.97% level number obtained is compared with the corresponding percentile in the distribution based on internal or external loss data. Typically the institution uses purely qualitative judgment to choose an operational risk capital amount in between the two figures.

All other methods of which we are aware are, in essence, some variation or combination of these two methods. Using those methods, institutions adopt some type of ad hoc and often arbitrary weighting of the 99.9% or 99.97% level numbers from the loss distributions derived from both internal loss event data (and also sometimes include external loss event data) and the scenario data to arrive at a final model-based regulatory or economic capital number.

In the next section, we will first discuss the information contained in scenario data, then explore what information should be contained in scenario data in order for it to be more useful for risk management, and finally discuss the issues one needs to be mindful of when integrating scenario data into operational risk measurement.

2. Scenario Data

Scenarios are generated by using an institution’s internal loss experience (internal loss data), the loss experiences of other similar institutions (external loss data), and other available information for the business environment. Most of the US-based financial institutions subjected to Basel regulation use internal loss data (ILD), external loss data (ELD), and other factors such as Business Environment and Internal Control Factor (BEICF). It should be noted that there is no regulatory prescription on which and how information should be used for the scenario data generation. Under the Basel regulatory requirements, all three data elements (ILD, ELD and BEICF), along with the scenario data, are required to be used directly or indirectly in the capital model for the purpose of calculating regulatory or economic capital.

Of these, external loss data predominantly are used as the primary guidance factor for scenario generation at every financial institution. There are several sources of external data for the purpose of scenario generation. The external data being used for scenario generation include the magnitude of the loss amount.

---

---

---
and a description of the loss, including the name of the institution, the business line where the loss happened, and the loss type. Basel regulatory requirements have categorized the operational losses into seven types: Internal Fraud, External Fraud, Employment Practices and Work Place Safety, Client Product and Business Practices, Damage to Physical Assets, Business Disruptions and System Failures, Execution Delivery and Process Management.

The scenarios are generated within a unit of measure. A unit of measure is often across different loss types or across loss types within business lines. It could also be across business or sub-business lines. A unit of measure gets predetermined based on risk management decisions before the scenario data generation happens. For some units of measure, there may not be enough internal loss experience to make any meaningful analysis. Sometimes those units of measure are supplemented with external data. Due to the current state of the availability of external data and based on the preliminary research we have undertaken on external data, we are not comfortable using our approach on those units of measure where there are insufficient internal loss data to develop a stable model. The method we will describe in the next section does not explicitly depend on which data were used to model the “base” model for a unit of measure. However, it will be clear later that an unstable base model will result in a poor estimation of the scenario effect.

In our application we often combine those units of measure for which there has not been enough loss experience to form a separate unit of measure. We could call this the “other” category for the purpose of our calculation. Hopefully the “other” category will have enough internal loss data to have a meaningful and stable model. The issue of the selection of a unit of measure is as important as finding a distribution that will fit a set of data. The suggestion made here is purely for convenience. Having said that, we strongly advocate that the risk management requirements should be the primary criteria for the selection of units of measure.

A scenario workshop is conducted for the generation of scenario data. The participants in the workshops are business line managers, business risk managers, and people with significant knowledge and understanding of their business and the environments within which it operates. Scenarios are generated within each unit of measure, typically with a corporate risk manager or an independent facilitator conducting it. The workshop participants discuss the business environments, current business practices, and take guidance and help from external data such as the following:

**Event A**

*At bank XYZ, when selling convertibles to clients an employee makes inappropriate promises to buy them back at a certain price. The market condition moves in the wrong direction and the bank is required to honor the commitment. As result, the bank suffers a loss of $200,000,000.*

*Question for Workshop: Could a similar situation occur at our institution? If so, what is the potential magnitude of the loss and how frequently might this happen?*

This scenario will be generated, typically, under the unit of measure of Clients, Products and Business Practices. After considering a realm of possibilities, participants at the workshop agree on a scenario related to this. Here we are assuming one scenario per incidence type. If there is more than one scenario, it can be very effectively combined without sacrificing the value of that scenario. We should remember that the scenario is not loss data. It is an impact and sensitivity study of the current risk management environment.

---

9 This example was supplied to us by a client. Our understanding is that it is an adaptation of an event from an external database.
The data it generates have two components – severity and frequency. The severity can be a point estimate (e.g. $2 million) or a range estimate (e.g. between $1 million and $3 million). For our purpose, we would like to work with range estimates, as we believe that in such a hypothetical situation the possibility of a loss occurring within a range is more probable than the possibility of a loss of a specific dollar amount occurring. This is also consistent with continuous severity distributions we use for modeling the severity of internal loss data. For a continuous distribution, the probability of the corresponding random variable taking a particular value is by definition zero. We can artificially convert the point estimate to a range estimate, by taking it as the midpoint of the range of say ±p%. We will revisit this in the Discussion section. The frequency comes in the form of \( m/t \) where \( m \) is the number of times this event will occur in \( t \) years. This we will interpret as \( m \) number of events will occur in a sample size of \( t n \), where \( n \) is the number of losses observed annually distributed by the internal frequency distribution of the loss data for that particular unit of measure. We are assuming that the capital calculation is on an annual basis. We express the frequency denominator on the scale of number of years to express the sample size as multiple of the annualized count of the internal losses at an institution. Like the severity, one could also create the frequency in a range such as \( m/[t_1 \ t_2] \). This we will interpret as \( m \) such events can occur between \( t_1 \) and \( t_2 \) years. In that case we will denote the calculation at the two end points of time \( t_1 \) and \( t_2 \) as the worst case and best case estimates. Alternatively one could also take a number in between \( t_1 \) and \( t_2 \), such as the average. Here we are making one very subtle assumption that we will use throughout our remaining analysis.

**Assumption 1**: During a short and reasonably specified period of time, such as one year or less, the frequency and severity distributions based on the internal loss data for a unit of measure do not change.

This is an important assumption in the sense that the methodology to be discussed in the next section is conditional on given severity and frequency distributions (in this case, distributions based upon internal loss data). If one would like to interpret this on the basis of time and state of the riskiness of the unit of measure, then we would say that at time zero (today) we have full knowledge of the loss events for a unit of measure. Based on this knowledge, we forecast the future for a reasonable period of time in which we can safely assume that the assumption made here is valid.

Once an event is expressed as a severity and frequency, we find severity is an interval on the real line of the form \([a,b]\). Therefore, from now on we will use interchangeably the words severity and event for \([a,b]\). In the next section we will also use some other variations of the representation of \([a,b]\) as an event.

Before concluding the data section we would like to stress the fact that the scenario data are not the loss experience of an institution. In our analysis we do not use scenario data as a substitute for internal loss data. Scenario data represent the possibility of a loss; we are suggesting a method to study its impact. Therefore, we make another vital assumption for the implementation of our method.

**Assumption 2**: The number of scenarios generated for a unit of measure is significantly less than the number of internal loss events observed in that unit of measure.

Subjectivity and possible biases will always be inherent characteristics of scenario data. Methods for interpreting scenarios have to take this into account. As Kahneman, Slovic and Tversky (1982) put it: “A scenario is especially satisfying when the path that leads from the initial to terminal state is not immediately apparent, so that the introduction of intermediate stages actually raises the subjective probability of the target event.” We have undertaken research which seeks to explain how one could control and reduce the biases and subjectivity in scenario data in the context of operational risk. Very preliminary results show that the scenario data generated in the format discussed above are less subjective and therefore more suitable than data produced in other possible formats. We would also like to think that this is perhaps the most natural format in which workshop participants can interact and express their beliefs. Early indications show that the Change of Measure evaluation of scenarios can also serve as a good filtering tool. In
the Bayesian context, discussed earlier, expert opinion (through the scenario generation process) will be sought for the estimation of the “prior” distributions. We chose to use the terminology from the Bayesian framework without any explanation. Discussion on the Bayesian approach is not within the scope of this work. Sivia (1998) is a very good reference for an introduction on the Bayesian method. Unless the data sought in that context are very similar in format to the data we will be using here, the scenario data generation for the Bayesian method will be a big roadblock for useful implementation of such a method.

Earlier we discussed that scenario data are predominantly influenced by external data. We have several concerns with respect to that. We need to understand that the external data are historical events that have occurred at other institutions. Scenario data generated solely by using or over-relying on external data and without taking many other hypothetical situations into consideration may defeat the purpose of the scenario in the first place. We will discuss this issue in later publications, as detailed discussions of these issues are beyond the scope of this paper.

For the methodology we discuss in the next section, any form of scenario can be used if its severity can be mapped as an interval on the real line and its frequency can expressed as a count expressed on the scale of number of years explained earlier.

3. Methodology

In our world, we will be dealing with two types of probability measures for an event. One is the historical probability measure based on internal loss data and the other is the implied probability measure, based on historical data and scenarios. One could argue that the historical probability measure could be based on internal and external loss data as well. In our experiment we have found measurement based on the internal loss data to be much more stable than that based on both internal and external loss data. As we will demonstrate later, the historical measure is an important cornerstone of this method; therefore, we conduct our experiment based upon internal data only.

Using the methodology discussed in this section, we systematically extend the frequency of the underlying historical data to match the frequency for each event on a normalized scale in the set of scenarios. Thus, historical severity and frequency distributions get transformed to the corresponding implied distribution. We then evaluate the effect of each scenario in the ratio of Change of Measure.

Based on our discussion in the section 2 and without the loss of generality, we denote an event by an interval on the positive real line $R^+$. For all practical purposes, our event is a bounded interval $[a,b)$, $(a,b]$, $[a,b)$, or $(a,b)$. Here we are following the notation and nomenclature similar to the one given in section 5 of chapter 2 of Royden (1988). For measurability we use the definition used in Chapter 11 of Royden (1988).

We define the range of an event as $b - a$, when $a$ and $b$ are finite. However, by definition these are not the only events. We can have an event (though it may not be useful for our purpose) of the form $[a,\infty)$, or $(0,\infty)$ so that the event space is a $\sigma$-finite measure space. We say an event $[a,b]$ has occurred when we see in our sample an observation $x$ such that $a \leq x \leq b$. For an event with another format we can use the same definition for the occurrence of an event with suitable adjustment of the boundary points.

3.1 Calculation of Historical Probability Measure

We start with the internal loss data (ILD). As we discussed earlier, They are the most important piece of data conveying an institution’s risk profile. Let $f(x)$ be the probability density function of the severity of
the loss data of an institution for a particular unit of measure \( U \). Here unit of measure could be a subsection of a business line comprising one particular Basel event type, such as external fraud or clients, products and business practices.

Therefore the probability of an event for \([a, b]\) (or the probability of a loss occurring between two loss amounts) is:

\[
\mu = \int_a^b f(x)\,dx
\]  

(1)

We call \( \mu \) a probability measure of the event for \([a, b]\) and denote it by \( \mu([a, b]) \). A probability measure is a measure in a general measure theoretic sense. Please refer to Chapter 12 of Royden (1988). In our subsequent use, we will drop the word “probability” in the phrase “probability measure” and simply call it a measure of \([a, b]\).

Please note that when \( f(x) \) is continuous \( \mu([a, b]) = \mu((a, b)) = \mu([a, b]) = \mu((a, b)) \). Also we should note here that when \( a \to b \) , \( \mu \to 0 \). In other words, when the range of an event goes to zero, its probability of occurring (frequency) also goes to zero. We will be using this important fact when discussing the sensitivity analysis due to the range of an event.

(2)

We can interpret the above measure (1) under the frequency interpretation as well. If we randomly draw \( n \) numbers from the distribution with \( f(x) \) as the probability density function and repeat the draw a sufficient number of times, we would expect to see \( n\mu \) number of observations between \( a \leq x \leq b \).

(3)

Since \( \mu \) is fixed by the severity distribution in (1), the variability in the frequency count will be due to \( n \). In order to understand the impact of \( n \) and the role it plays, we need to understand the frequency distribution of the same unit of measure.

Let \( \lambda(n) \) be the probability density function of the frequency of the internal loss data for the same unit of measure \( U \) as before. Given that it is a frequency distribution, \( \lambda(n) \) is a discrete distribution. In other words, \( n \) is a positive integer occurring with a probability \( \lambda(n) \).

(4)

Combining (3) and (4), one can see that with probability \( \lambda(n) \), \( n\mu \) number of observations will be observed between \( a \leq x \leq b \). For our purposes here, \( f(x) \) and \( \lambda(n) \) are the severity and the frequency distributions for a unit of measure estimated using internal loss data. We will refer to these severity and the frequency distributions as the current or inherent or historical probability measure and \( n\mu \) as the number of observations that will be observed between \( a \leq x \leq b \) in a draw of \( n \) observations from \( f(x) \) under the current or inherent or historical measure of \( f(x) \) and \( \lambda(n) \). The internal loss data reflect the loss experience of an institution, its risk profile and other factors that are particular to it and the environment in which it operates. We refer to a measure based solely on internal data interchangeably as a measure of current state, a current measure, or a historical measure.

On the other hand, the data from scenario analysis will most likely attach a different probability for the same event for \([a, b]\) and therefore imply a different \( f(x) \) and \( \lambda(n) \) from the one obtained under current or inherent measures. In theory, the scenario-based \( f(x) \) and \( \lambda(n) \) can be completely different from the
one obtained using ILD. However, as we will discuss later in our application, the scenario data based \( f(x) \) and \( \lambda(n) \) will differ from their ILD based counterparts only in terms of parameter values.

As discussed in section 2, scenario data are hypothetical data that an institution has not yet experienced. Such losses may or may not be observed in the future. Therefore, we will refer to the scenario-adjusted severity and frequency of the loss events for a unit of measure as the \textit{implied probability measure} for severity and frequency implied by the scenario. Here we would like to draw a parallel between the terminology in financial economics and asset pricing and the one we introduced. In the Black-Scholes world of option pricing, one derives the volatility (the \( \sigma \) of the lognormal distribution) from the price of an option and calls it an implied volatility. Similarly, the probability measure we obtain from the scenario-based loss data will be an implied measure due to a particular set of scenarios.

In the following subsection, we will show how we can calculate this implied probability measure for a set of scenarios.

\subsection*{3.2 Calculation of Implied Probability Measure}

In this section we discuss a method by which one can calculate implied probability measures for a set of scenarios. At the end of this section we have worked with an example to illustrate this method. In our data, severity for an event is defined as \([a,b]\) and frequency comes in the format of \( m/t \), which means that \( m \) of those events are likely or expected to occur in \( t \) number of years. Roughly speaking, in deriving the probability measure due to the scenario, we ask, all else being equal, to what extent does one need to change the historical probability measure so that the frequency count for an event for \([a,b]\) is equal to \( m/t \). If the historical frequency is \( n \) distributed as \( \lambda(n) \), then in \( t \) years we expect to see \( y = \sum_{i=1}^{t} n_i \) number of events, where each of the \( n_i \) is independent and identically distributed as \( \lambda(n) \). (5)

Suppose for an event for \([a,b]\) and for a draw of a sample amount equal to \( y \) from the historical severity distribution we observe \( k \) number of this event occurring (we expect to see \( y \mu \) observations occurring between events for \([a,b]\)). If \( k \) is less than \( m \), then we generate \( m-k \) samples from the historical severity distribution within the range of \([a,b]\) and combine with the \( y \) samples we had drawn from the historical severity. We find a severity distribution using the combined data set where \( y \) numbers of observations are drawn from historical severity distribution and \( m-k \) observations are drawn from the same historical severity distribution within the range of \([a,b]\). This severity distribution is called the \textit{implied severity distribution} due to the scenario for \([a,b]\) which occurs \( m \) times in \( t \) years.

From a practical point of view, \( y \) will be significantly larger than \( m-k \). Therefore it makes no sense to explore a completely different family of distributions from the one we found to characterize the historical severity distribution. In that case, we re-estimate the parameters of the historical severity distribution using the combined data set. We will revisit this issue in the discussion section.

One should also note that the implied severity distribution is sensitive to the value of \( y \) in step (5), which is a random variable distributed as \( \Theta, \lambda(n) \), a convolution of \( t \) number of \( \lambda(n) \)'s. Therefore, the implied severity distribution we discussed earlier is an implied distribution at the instance \( y \). We can now introduce a notation to describe the calculation we have made thus far. By \( I_y^S(x) \), we mean an implied severity distribution due to scenario \( S \) at the instance \( y \).
More generally, if \( \psi = \{ s_1, s_2, s_3, \ldots, s_r \} \) is the set of scenarios, then \( J^y_x(\psi) \) denotes an implied severity distribution due to the combined effect of all the scenarios in set \( \psi \) at the instance of \( y \). The key to appropriately deriving the combined effect is to preserve the frequencies of the scenario events.

Calculation of \( J^y_x(\psi) \) is very similar to and essentially the same as the case of one scenario. Scenario \( s_i \) in \( \psi \) has a severity range \([a_i, b_i]\) and frequency \( m_i / t_i \). Therefore, each scenario is expressed in terms of a different time span. We take \( T \) to be the \( \max \{ t_i \} \). If \( m_i \) of event \( s_i \) occurs in \( t_i \) years then in \( T \) years we should have \( T (m_i / t_i) \) events of \( s_i \). Here we are making a simple linear extrapolation. One could use a stochastic extrapolation on the basis of the frequency distribution. Using Poisson as the frequency distribution, we find that the linear extrapolation is a very close approximation of the stochastic extrapolation. In other words, we have normalized each frequency on a common basis. This may not be a whole number, in which case we use the largest integer less than or equal to \( T (m_i / t_i) + 1 \) and denote it by \( k_i \).

After normalizing each scenario on the dimension of frequency, we have for each scenario \( s_i \), \([a_i, b_i]\) as the severity range and \( k_i \) as the unadjusted frequency, which needs to be adjusted for the cumulative effect of the overlap of events with other scenarios. In the case of more than one scenario, it is necessary to keep an account of the frequency of each scenario. For example, if we have two scenarios with exactly the same frequency and severity, then in our sample of \( y \) observations we need for the number of occurrences of loss events of that severity to be equal to or greater than the sum of the frequencies as dictated by the two scenarios. It is also possible that the severity of a scenario overlaps with the severity of another scenario. In such cases, each of the \( k_i \) needs to be adjusted based on one of the following two situations:

1. The severity range of the scenario is completely disjoint from all other scenarios, or
2. The severity range of the scenario overlaps, completely or partially, with other scenarios in the set of scenarios \( \psi \).

**Case 1**

When the severity range related to a scenario is completely disjoint from all other scenarios, its frequency remains unchanged.

**Case 2**

When the severity range of a scenario overlaps fully or partially with the severity range of other scenarios, we need a systematic way to account for the cumulative frequency for each scenario due to the overlaps. We use that using a cumulative matrix. A cumulative matrix in our case is a lower triangular \( r \) by \( r \) matrix for the \( r \) number of scenarios in the set \( \psi : R = (R_{ij}) \) where \( 1 \leq i, j \leq r \).

We sort the set of scenarios with respect to the lower bound of their severity range. In our application, the value of the lower bound is never less than zero. Then we go through each event. Let the \( i^{th} \) scenario be with severity range \([a_i, b_i]\) and frequency \( k_i \), and let the \( j^{th} \) scenario be with severity range \([a_j, b_j]\) and frequency \( k_j \), where \( i < j \).

We define \( R_{ii} = 1 \) where \( 1 \leq i \leq r \).

For \( 1 \leq i < j \leq r \),

\[
R_{ij} = \begin{cases} 
(b_j - a_j)(b_i - a_i) & \text{when the } j^{th} \text{ scenario overlaps with the } i^{th} \text{ scenario} \\
0 & \text{otherwise.}
\end{cases}
\]
We denote by \( K = (k_i) \) where \( 1 \leq i \leq r \) be the column vector of \( k_i \), the unadjusted but normalized frequencies for each scenario. Therefore, the adjusted cumulative frequency for each scenario is given by the column vector:

\[
\tilde{F} = RK = (f_i) \text{ where } 1 \leq i \leq r.
\]

It should be noted that the construction of \( \tilde{F} \) is essentially for the purpose of an algorithmic implementation to account for the cumulative frequency due to partial or total overlaps of events given by the scenarios in \( \Psi \). In our algorithm, we consider each scenario in order of increasing lower bound severity for calculating cumulative frequency.

Now we have for each scenario \( s_i, [a_i, b_i] \) as the severity range and \( f_i \) as the (cumulative) number of times an event related to this scenario must occur in our sample space of \( y = \sum_{i=1}^{y} n_i \) observations, where each of the \( n_i \) is independent but identically distributed as \( \lambda(n) \). Similar to the case of one scenario discussed earlier, we now need to find \( f_i \) events occurring for each scenario in \( y \) observations. If we find less than \( f_i \) events in \( y \) observations, then we need to augment the set of observations by drawing from the corresponding severity distribution for the difference.

We form the set \( S = \{I^y_x(x) : y\} \) for different draws of \( y \) distributed as \( \Theta_y \lambda(n) \). The size of \( S \) is typically at least 10,000. For our application, for each \( I^y_x(x) \) in \( S \) we denote by \( \eta(I^y_x(x)) \) the 99.9% or 99.97% level value we obtain from the loss distribution by using \( I^y_x(x) \) as the severity and the historical \( \lambda(n) \) as the frequency. It is not necessary that \( I^y_x(x) \) be used for the 99.9% or 99.97% level calculation purpose only. It could be used for any other measure as well. Although we are using the historical frequency count for the purpose of our frequency, the frequency does change also. However, since we assume very little data augmentation compared to total size of the of the sample space, the average value of the frequency will not change. If that is not the case, then we need to readjust the frequency as well. In the case of Poisson as the frequency distribution, this readjustment is trivial. The adjusted Poisson parameter is equal to \((y + \text{total number of draws from the severity distribution for all scenarios combined})/T\). It is advised, for the purpose of saving time that the 99.9% or 99.97% be calculated using the single loss approximation formula given by Böcker and Klüppelberg (2005) instead of a Monte Carlo simulation of one million trials. Typically we take the average of all the \( \eta(I^y_x(x)) \) to arrive at the final capital number due to scenarios in set \( \Psi \) conditional on the current state, which is nothing but the historical loss experienced by an institution. The corresponding \( I^y_x(x) \) is the implied probability distribution that will be used for capital calculation. The reason we chose to take the average is to make it similar to the practice in asset pricing theory. While pricing an asset we take future expected cash flow from an asset under a probability measure (typically risk-neutral) and then discount it the with appropriate discount factor (typically the risk-free rate). Here also we are trying to find the average implied distribution due to the combined effect of a set of scenarios under certain criteria, such as the 99.9% level of the aggregate loss distribution resulting from the implied distribution. Alternatively one can form a 95% band of the estimates by ignoring the bottom and top 2.5% of estimates, or take the median of all the estimates as a final number. If we take the 95% band, then there will be two corresponding distributions for this band. In the case of taking an average, there may not be an exactly corresponding \( I^y_x(x) \). In that case, we take the nearest one to represent the implied distribution.
An Illustration of the Methodology\(^\text{10}\)

The following example illustrates the steps outlined in the Section 3 (Methodology). Table A1 presents a sample of scenario analysis data. The frequency of the scenario should be interpreted as once in the stated frequency number of years. The financial institution obtained range estimates for the severity of each event in scenario workshops; therefore severity is represented by upper and lower severity bounds. Had the workshops yielded a point estimate, we would have calculated a 20% interval around it as has been explained in the methodology section.

**Table A1: Data**

<table>
<thead>
<tr>
<th>No</th>
<th>Frequency</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>25</td>
<td>$29,412</td>
<td>$147,059</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>$73,529</td>
<td>$294,118</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>$88,235</td>
<td>$588,235</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>$500,000</td>
<td>$1,200,000</td>
</tr>
<tr>
<td>5</td>
<td>62</td>
<td>$1,168,500</td>
<td>$1,291,500</td>
</tr>
</tbody>
</table>

**Normalization of the Scenario Frequencies**
The scenario frequencies are normalized such that each normalized frequency value represents the number of times a scenario would occur in a number of years equal to the maximum frequency, which in this example is 62 years (from Table A1).

**Table A2: Scenario Data with Normalized Frequencies**

<table>
<thead>
<tr>
<th>No</th>
<th>Normalized Frequency</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.48</td>
<td>$29,412</td>
<td>$147,059</td>
</tr>
<tr>
<td>2</td>
<td>3.1</td>
<td>$73,529</td>
<td>$294,118</td>
</tr>
<tr>
<td>3</td>
<td>12.4</td>
<td>$88,235</td>
<td>$588,235</td>
</tr>
<tr>
<td>4</td>
<td>6.2</td>
<td>$500,000</td>
<td>$1,200,000</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>$1,168,500</td>
<td>$1,291,500</td>
</tr>
</tbody>
</table>

**Calculation of Cumulative Frequency for Overlap Scenarios**
If the severity ranges of any two scenarios overlap, then each scenario’s frequency needs to be adjusted to account for the overlap. For our approach to accurately incorporate scenario data into the measurement model, we must consider the cumulative effect of scenarios. We want to determine whether the model specified by internal data can satisfy all scenarios within a given severity range, not just each scenario within the range individually. The rationale for our approach is based upon the assumption that the scenarios are independent. Not assuming independence would complicate the model because one would need to account for the dependencies among scenarios, which are difficult to justify given a small amount of data.

\(^{10}\) The example given here is based on an actual implementation of the Change of Measure method at a financial institution. To protect the proprietary information of the institution, the real data have been altered. Therefore, the example here should be used purely for understanding the steps outlined under Methodology earlier and not for any other interpretation of either data or results.
To test this, we calculate the percentage of overlap for scenarios with intersecting severity ranges. We first sort the scenario data by increasing lower bound severity. The overlap between any two scenarios is calculated using the following formula:

\[
\text{Percentage Overlap} = \frac{(b_1 - a_2)}{(b_1 - a_1)}
\]

where \(b_1\) is the upper severity bound for the scenario with the lesser lower severity bound, \(a_2\) (if \(a_2 < b_1\)) is the lower severity bound for the scenario with the greater (or equal) lower severity bound, and \(a_1\) is the lower severity bound for the scenario with the lesser lower severity bound.

The percentage overlap between scenario 1 and 2 using the formula above is 63%. We use the percentage overlap calculations to create a cumulative normalized frequency for each severity range. The normalized frequency in scenario 2 is adjusted so that it takes into account the normalized frequency in scenario 1 based on the percentage overlap using the following formula:

\[
\text{Cumulative Common Frequency} = m_{hi} + (\text{Percentage Overlap}) \times m_{lo},
\]

where \(m_{hi}\) is the normalized frequency for the scenario with the greater lower severity bound and \(m_{lo}\) is the normalized frequency for the scenario with the lesser lower severity bound.

Note that a scenario could have overlapped with multiple scenarios and the percentage overlap is always calculated similarly. Because scenario 3 intersects with multiple scenarios (1 and 2), the cumulative normalized frequency would be determined first by calculating the percentage overlap with scenario 1, and then calculating the percentage overlap with scenario 2, and summing the results from each step. Since we have sorted the data by increasing severity lower bound, at each step of the process, we consider only scenarios above the scenario in question when calculating the cumulative normalized frequency. Table A3 shows the cumulative normalized frequencies based on the above formula and discussion. The cumulative normalized frequencies are implied by the scenarios and not necessarily the same as those implied by the historical or internal loss data.

The result of the calculation above is a lower triangular matrix containing the percentage overlap between the individual scenarios. Applying this formula to the normalized frequencies provided in Table A2 results in the following severity range overlap matrix:

Table A3: Scenario Data Overlap Matrix and Implied Cumulative Frequency

<table>
<thead>
<tr>
<th>No</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
<th>Normalized Frequency</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Cumulative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$29,412</td>
<td>$147,059</td>
<td>2.48</td>
<td>100%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>$73,529</td>
<td>$294,118</td>
<td>3.1</td>
<td>63%</td>
<td>100%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>$88,235</td>
<td>$588,235</td>
<td>12.4</td>
<td>50%</td>
<td>93%</td>
<td>100%</td>
<td>0%</td>
<td>0%</td>
<td>17</td>
</tr>
<tr>
<td>4</td>
<td>$500,000</td>
<td>$1,200,000</td>
<td>6.2</td>
<td>0%</td>
<td>0%</td>
<td>18%</td>
<td>100%</td>
<td>0%</td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>$1,168,500</td>
<td>$1,291,500</td>
<td>1</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>5%</td>
<td>100%</td>
<td>1</td>
</tr>
</tbody>
</table>

Table A3 shows that the severity range for scenario 2 overlaps the severity range for scenario 1 by 63%. Similarly, 5% of the severity range of scenario 4 overlaps with the severity range of scenario 5. By multiplying the scenario severity range overlap matrix by the normalized frequencies found in Table A2, one can calculate the cumulative frequency for each scenario. This is same as the equation given in step 1.
Calculation of Implied Distribution

We use the annual frequency and severity distribution determined from internal data to simulate sample loss data (or sample data) over a number of years. The number of years is determined by the maximum frequency for a particular unit of measure. For example, maximum frequency in this example is 62. The annual number of losses for each of the 62 years of generated data is determined by a draw from the Poisson distribution using the average annual frequency computed from internal data. The severity for each loss is drawn from severity distribution best fitting the internal loss event data.

Now that the scenario data and the internal data are on a common time horizon, the scenario data can be added to the internal data. As noted in the methodology section, the time horizon is the unit in which the total number of data generated is expressed. However, we want to add only those events not sufficiently found in the sample data. If the expected number of events (cumulative frequency) specified by the scenario data is the same or less than the number of events in the sample loss data over the same severity range, then the scenario is satisfied by the sample data. If the cumulative frequency specified by scenario data is greater than the number given by sample loss data, then the scenario is not satisfied by sample data and additional events must be added to the sample data. We calculate the difference between the observed frequency from the sample data, \( k \), and the expected scenario cumulative frequency, \( m \). If \( k \) is less than \( m \), we add \( m \) minus \( k \) events, drawn from the severity range for the scenario from the historical severity distribution.

For example, if scenario data specify that eight events (cumulative frequency) in the range of $500,000 to $1,200,500 occur but the sample loss data include only two events over the same range, then six events (drawn in the range of $500,000 to $1,200,500 from the historical severity distribution) must be added to the sample loss data. Conversely, if the sample loss data had eight or more events over the same range, then this scenario would already be satisfied by the internal loss data and neither any additional loss events would be added to the sample data nor subtracted from the sample data to match with the cumulative frequency for that range.

We re-estimate the parameters of the loss severity distribution that we used to generate the sample data. The augmented sample represents both internal and scenario data. We would expect that most losses in this vector are generated internal data points, only a few are augmented due the adjustments from the scenario considered. As discussed in the Section 4 (Discussion), if this condition does not hold then this method may not be suitable for the set of scenarios considered.

We repeat this process ten thousand times to minimize the simulation variance of the sample data generation process and estimate a range of parameters of the loss severity distribution. We also recalculate the frequency parameter, accounting for the “uncovered” events which are added to the sample data. At each step of the simulation process we could also calculate the 99.9% or 99.97% level value of the aggregate loss distribution using the re-estimated parameters of the frequency and severity distributions. One can use Monte-Carlo simulation of one million trials for that or, in order to save time significantly, one could use the Single Loss Approximation formula given in Böcker and Klüppelberg (2005). Out of all the losses reaching the 99.9% or 99.97% threshold, in each trial generated we use the mean for the purpose of estimating operational risk given the internal loss experience and the scenario generated at an institution.

Table A4 shows the severity ranges for the scenarios considered in this example together with the cumulative frequencies from Table A3. The fifth column of Table A4 shows the historical frequency count, which is the number of loss occurrences generated by internal data within the severity range given and using the severity distribution used to model the severity of the internal loss event data.
### Evaluation of Scenarios: The Change of Measure

Each scenario will change the historical measure for a given severity range and the Change of Measure (COM) associated with each scenario will be given by:

\[
\text{Change of Measure} = \frac{\text{Implied probability measure}}{\text{Historical probability measure}}
\]

Here the Change of Measure is a one-step conditional Change of Measure and only for the severity distribution. In general, one can do the same thing for the frequency as well, in conjunction with the severity. However, as we discussed before, in our application such effort will have very negligible impact in overall computation and scenario studies. The step is from time zero (or current state of loss, i.e., ILD) to time one, the period for which we are trying to estimate the operational loss exposure. It should be noted here this time period should not be long, heuristically speaking (no more than a year). If this time period is long, then one may have to reevaluate the current state. Once again this same issue exists in the financial economics of asset prices. It is extremely difficult, if not impossible, to price an option of very long maturity given the current state of the economy. Borrowing the language of the financial economics, here we are trying to determine the states by the possible future values of the losses. In the case of tail events we don’t know what that loss will be. Therefore it is impossible to know the future values of the tail event losses. What we are trying to estimate is a range of values for likely outcomes. We need to think in terms of the possibilities for changes regarding events that will cause the tail losses.

In other words, based on the combined effect of all the scenarios in the set \( \psi \), what will be the marginal effect of a scenario event relative to historical data in terms of the likelihood of its occurrence? All else being equal, for one unit of change in the historical measure for an event, by how much will the implied measure for the same event change in relation to the set of scenarios in \( \psi \)? The Change of Measure representation here is equivalent to the Radon-Nikodym derivative of the implied measure with respect to the historical measure. When dealing with a continuous measure such as the severity distribution, care should be taken to verify that conditions leading to Radon-Nikodym are satisfied.\(^\text{11}\) There are some cases when the Radon-Nikodym derivative may not exist – for example, when the historical measure is close to zero, or when implied and historical measures are both close to zero. In either of these situations, we need to augment the measure appropriately so that it is significantly different from zero.

---

\(^{11}\) Chapter 11 in Royden (1988) has a good discussion on the Radon-Nikodym derivative.
The COM is a unit-free measurement, which makes it useful for evaluating the internal data derived loss distribution. Suppose we have two different distributions that adequately fit the internal loss data. For one, the COM is 20 and for the other it is 5. We can say that the second distribution is a better predictor of the given scenario than the first one, without knowing anything about any of those distributions.

The purpose of the COM here is to understand the impact of a particular scenario and its marginal contribution in risk valuation. An extremely high Change of Measure can be due to any of these three situations:

1. The historical measure is inaccurate and inconsistent with the risk profile of the institution;
2. The scenario is nearly impossible given the current state of the institution;
3. The risk of the institution is uninsurable (self retention or through risk transfer).

On the other hand, a low Change of Measure may indicate the redundancy of a scenario. We would like to stress the fact again that before we can put any faith in the Change of Measure number, we need to ensure that the numerator and denominator in calculating the COM are calculated correctly. If a density function involving scenarios is not calculated in relation to the historical measure then it will not be a good implied measure and the corresponding value of the COM will be incorrect or misleading. Needless to say, if the historical density is not evaluated correctly in the first place then the implied measure related to it and the corresponding Change of Measure will also be incorrect. The method we discussed earlier will be one such method that systematically (in the sense of a probability measure) updates the historical density to calculate the implied density. COM can also be used to evaluate our valuation of current state. If we had several choices to fit our internal loss severity data, COM can be a criterion for model selection using all those scenarios that we believe are a good representation of the future state of the loss profile.

Therefore, using scenario analysis and its corresponding COM, one can both validate the current state and the usefulness of the scenario under consideration. The method suggested here can be used to evaluate one scenario at a time or any subset of the set of all scenarios. Once again, in the area of asset pricing one often uses price of an option to predict the future distribution. For example, Jackwerth and Rubinstein (1996) is a seminal work where probability distributions of the future price of underlying asset were estimated from the price of set of options. Here, using the same line of thought and using the COM, we could reassess our choice of a severity distribution in terms of its predictive power.

Table 1 shows the goodness-of-fit test results for five different distributions used to model internal loss data for a particular unit of measure at a financial institution. We show two different types of goodness-of-fit tests: Anderson-Darling (AD) and Kolmogorov-Smirnov (KS). On the basis of the goodness-of-fit alone, all five distributions could be used to model the internal loss data. However, based on the p-value of the AD test Burr is the best fit and GB2 comes very close to it. We should note here that GB2 is the parent distribution of Burr. Lognormal was the worst fit, even when the null hypothesis with respect to the lognormal could not be rejected. Based on the p-value of the KS test, GPD is the best and Burr is the second best fit. If we like to fit the tail better than the body then we should give priority to the AD test over the KS test and in that case Burr will be the distribution of choice for this unit of measure. All five distributions could not be rejected by standard statistical goodness-of-fit tests. All five distributions resulted in a reasonable 99.9% and 99.97% level capital with the internal loss data.
Table 1: Goodness-of-Fit Results

<table>
<thead>
<tr>
<th>Distributions</th>
<th>Anderson-Darling</th>
<th></th>
<th></th>
<th></th>
<th>Kolmogorov-Smirnov</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Stat</td>
<td>CV</td>
<td>P</td>
<td>H</td>
<td>Stat</td>
<td>CV</td>
<td>P</td>
<td>H</td>
</tr>
<tr>
<td>Lognormal</td>
<td>0.701652</td>
<td>0.770949</td>
<td>0.078856</td>
<td>0</td>
<td>0.059954</td>
<td>0.065463</td>
<td>0.093988</td>
<td>0</td>
</tr>
<tr>
<td>GPD</td>
<td>0.434731</td>
<td>0.79368</td>
<td>0.308131</td>
<td>0</td>
<td>0.038805</td>
<td>0.062156</td>
<td>0.624382</td>
<td>0</td>
</tr>
<tr>
<td>Loglogistic</td>
<td>0.390558</td>
<td>0.682233</td>
<td>0.323106</td>
<td>0</td>
<td>0.048933</td>
<td>0.059827</td>
<td>0.201647</td>
<td>0</td>
</tr>
<tr>
<td>Burr</td>
<td>0.26832</td>
<td>0.49766</td>
<td>0.43737</td>
<td>0</td>
<td>0.040082</td>
<td>0.056151</td>
<td>0.446828</td>
<td>0</td>
</tr>
<tr>
<td>GB2</td>
<td>0.272043</td>
<td>0.467749</td>
<td>0.316249</td>
<td>0</td>
<td>0.041333</td>
<td>0.057279</td>
<td>0.279376</td>
<td>0</td>
</tr>
</tbody>
</table>

Column headings are defined as follows: Stat : test statistics, CV : critical value, P: p-value, H = 0 : null hypothesis, i.e., a distribution fits the data, could not be rejected; H = 1: alternative hypothesis, i.e., the distribution does not fit the data, could not be rejected.

Table 2: Scenario Data

<table>
<thead>
<tr>
<th>No</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
<th>Cumulative Frequency</th>
<th>GPD Measure</th>
<th>Loglogistic Measure</th>
<th>Lognormal Measure</th>
<th>Burr Measure</th>
<th>GB2 Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>a</td>
<td>b</td>
<td>7</td>
<td>0.88</td>
<td>0.96</td>
<td>0.92</td>
<td>0.88</td>
<td>0.95</td>
</tr>
<tr>
<td>2</td>
<td>2.40a</td>
<td>3.69b</td>
<td>14</td>
<td>1.01</td>
<td>1.13</td>
<td>1.13</td>
<td>1.05</td>
<td>0.97</td>
</tr>
<tr>
<td>3</td>
<td>12.75a</td>
<td>5.10b</td>
<td>18</td>
<td>1.41</td>
<td>1.39</td>
<td>1.56</td>
<td>1.53</td>
<td>1.37</td>
</tr>
<tr>
<td>4</td>
<td>17.50a</td>
<td>6.50b</td>
<td>16</td>
<td>1.55</td>
<td>1.44</td>
<td>1.72</td>
<td>1.68</td>
<td>1.55</td>
</tr>
<tr>
<td>5</td>
<td>60.45a</td>
<td>22.09b</td>
<td>10</td>
<td>2.34</td>
<td>1.69</td>
<td>2.81</td>
<td>2.50</td>
<td>2.79</td>
</tr>
<tr>
<td>6</td>
<td>75.00a</td>
<td>40.00b</td>
<td>20</td>
<td>2.66</td>
<td>1.76</td>
<td>3.29</td>
<td>2.80</td>
<td>3.35</td>
</tr>
<tr>
<td>7</td>
<td>75.00a</td>
<td>45.00b</td>
<td>34</td>
<td>2.70</td>
<td>1.77</td>
<td>3.35</td>
<td>2.83</td>
<td>3.43</td>
</tr>
<tr>
<td>8</td>
<td>75.00a</td>
<td>50.00b</td>
<td>54</td>
<td>2.73</td>
<td>1.78</td>
<td>3.39</td>
<td>2.87</td>
<td>3.49</td>
</tr>
<tr>
<td>9</td>
<td>75.50a</td>
<td>67.20b</td>
<td>60</td>
<td>2.82</td>
<td>1.80</td>
<td>3.51</td>
<td>2.95</td>
<td>3.66</td>
</tr>
<tr>
<td>10</td>
<td>103.00a</td>
<td>41.20b</td>
<td>60</td>
<td>2.86</td>
<td>1.81</td>
<td>3.66</td>
<td>2.99</td>
<td>3.73</td>
</tr>
<tr>
<td>11</td>
<td>106.00a</td>
<td>50.45b</td>
<td>63</td>
<td>2.75</td>
<td>1.79</td>
<td>3.49</td>
<td>2.89</td>
<td>3.53</td>
</tr>
<tr>
<td>12</td>
<td>119.20a</td>
<td>37.17b</td>
<td>58</td>
<td>2.90</td>
<td>1.82</td>
<td>3.76</td>
<td>3.03</td>
<td>3.82</td>
</tr>
<tr>
<td>13</td>
<td>400.00a</td>
<td>120.00b</td>
<td>3</td>
<td>4.42</td>
<td>2.08</td>
<td>7.19</td>
<td>4.37</td>
<td>7.11</td>
</tr>
<tr>
<td>14</td>
<td>1160.00a</td>
<td>387.00b</td>
<td>20</td>
<td>6.54</td>
<td>2.35</td>
<td>14.49</td>
<td>6.11</td>
<td>12.77</td>
</tr>
<tr>
<td>15</td>
<td>1697.36a</td>
<td>395.83b</td>
<td>8</td>
<td>7.07</td>
<td>2.40</td>
<td>17.08</td>
<td>6.53</td>
<td>14.34</td>
</tr>
<tr>
<td>16</td>
<td>3500.00a</td>
<td>1500.00b</td>
<td>7</td>
<td>10.02</td>
<td>2.67</td>
<td>33.97</td>
<td>8.76</td>
<td>24.17</td>
</tr>
</tbody>
</table>

Table 2 depicts sixteen scenarios used at the same financial institution and for the same unit of measure depicted in Table 1. Please note individual scenarios are the prediction on various tail events within a unit of measure. We have disguised the actual data and expressed each scenario as a multiple of the lower and upper bounds of the first scenario, which we have chosen to represent as [a,b]. Table 2 also shows the Change of Measure with respect to the five different distributions that appear in Table 1.

In Chart 1 we show the comparative COM for each scenario using all five distributions shown in Table 1. Although we show COM for each scenario, we should make note that all sixteen scenarios were used to derive the probability distribution. Therefore, in the COM number for each scenario, we have two components: group effect and individual effect. Of these, for scenarios with very high severity values (extreme tail events) the individual effect will be dominant and for scenarios with comparatively low severity
values, the group effect may be dominant. If we would like to understand the absolute effect of a particular scenario (such as in a what-if analysis) we need to evaluate the scenario alone without mixing it with any other scenarios. The analysis with respect to the group effect and individual effect for each scenario will be of value to understand the risk sensitivity of various factors that constitute the risk for a unit of measure. For our purpose, we would use the COM to reflect the individual effect more often than the group effect as we will work with extreme tail events in a scenario.

In Chart 1 we observe that for the first twelve scenarios, the COM is very similar for all distributions. Even though we have not done the analysis to accurately infer this, we believe for the first twelve scenarios it is the group effect that dominates in the COM number. When it comes to the last four scenarios it is clearly the individual effect that dominates in the COM number. All five distributions behave differently with respect to the COM for the last four scenarios. Of these, loglogistic shows the least amount of variance with respect to COM. In that sense, loglogistic has the best predictive power for the given set of scenarios, conditional on the internal loss data; on the other hand, lognormal and GB2 performed the worst in this respect. Please also note that the GB2 distribution has a fatter tail than the loglogistic distribution. Also GB2 is a four-parameter distribution whereas loglogistic is a two-parameter distribution. Here we can also see that the predictive power of an extreme event of a distribution is not necessarily determined by either its tail shape or number of parameters it has. Therefore, all else being equal, with respect to this set of scenarios and the internal loss experience of the institution for this unit of measure, loglogistic will be the best choice for modeling the severity of the data.

Here, one could extend the analysis further in the context of the three situations discussed at the beginning of this section. Our conclusion of loglogistic as the best choice (and not the Burr) with respect to the scenarios is essentially believing that situation 1 (noted above) applies – i.e., that the historical measure is inaccurate and inconsistent with the risk profile of the institution. On the other hand, if we believe that the Burr distribution accurately models the risk profile of this unit of measure then the last three scenario events will probably be less likely to happen given the current state and risk profile of the institution with respect that unit of measure, i.e., believing that situation 2 applies.

The third possibility is when we trust the severity model of the internal loss event data using the Burr distribution and we also believe that the last three scenario events are equally possible. If in that case we arrive at an unrealistic capital estimate then we would like to call the last three scenarios uninsurable as discussed in situation 3 above. In that case, we need to revisit the risk controls and address the risk management problem through risk control mechanisms instead of through additional capital holdings. This analysis is very similar to what one does with respect to managing the market risk and portfolio selection using COM. See Cerny (2009) for details. The application of COM we are showing here can easily be adopted in insurance pricing for an individual or group risk.
Table 3 shows a standard error of the parameter estimates of 1,000 trials for the simulation used to merge the scenario data with the internal loss data across various unit of measures. Here we are showing the standard error for the best fitted distribution for a particular unit of measure. From the standard error estimate we can see that the method of estimation proposed here is very stable. We have tested this with the data from many different financial institutions and found this always to be the case.
4. Discussion

Internal loss data are the core of the methods discussed here. If the internal-loss-data-based model for either severity or frequency is not estimated correctly, then the method suggested here will suffer from instability and inaccuracy. For the purpose of severity and frequency model evaluation, we strongly suggest that along with goodness-of-fit tests for each model, one should measure model performance by the following criteria used in Dutta and Perry (2007).12

1. Realistic - If a method fits well in a statistical sense, does it generate a loss distribution with a realistic capital estimate?
2. Well-Specified - Are the characteristics of the fitted data similar to the loss data and logically consistent?
3. Flexible - How well is the method able to reasonably accommodate a wide variety of empirical loss data?
4. Simple - Is the method easy to apply in practice, and is it easy to generate random numbers for the purposes of loss simulation?

In addition, one should stress-test each distribution in terms of coherence, consistency, and robustness. In the past, institutions often liked to use scenarios either in the absence of internal loss data or when very little internal loss data had been observed. The method we present here cannot be used with scenario data alone.

This method is flexible enough that an institution can use more than one scenario for each unit of measure. We would like to think that an institution may need to generate several scenarios per unit of measure. However, at the same time implementing too many scenarios per unit of measure will be a practical limitation of our method. Heuristically speaking, if the number of scenarios an institution has per unit of measure is greater than 5-10% of the annual frequency of the internal loss event for that unit of measure, then our method may start showing instability with respect to the probability measure being calculated using the scenarios. Following the heuristics, one will not be using very many scenario data points to augment the internal loss data based severity distribution. Our experience has shown if the scenario set has only short term events (between 20 and 25 years), then we typically augment the historical sample by less than 4% percent of the sample size. On the other hand if the scenario set has several long term events (more than 50 years), then we needed to augment the sample by less than 1% of the sample size. For example, for a long term (100 year) event with an annual loss count of , say, 250 for a unit of measure, we will have size on the order of 25,000. As noted in D’Agostino and Stephens (1986), for such large data sets, goodness-of-fit tests are expected to fail for any distribution unless by stroke of luck every data point happens to be generated by the random seed of a distribution. Also, as we insisted, the robust fit of internal data will ensure that the fit is still statistically non-rejectable with the augmented sample. If any of the above conditions are not met, then the data have to be retested for implied probability. In fact, in those situations it will be better not to use this method than go through data fitting of an impractical amount of data such as 25,000.

In theory, our method does not prove that for the augmented data set the historical severity distribution is still a good fit. To be pedagogically correct, one perhaps needs to do a fresh assessment for the fit of the distribution with the augmented data. However, from a practical point of view refitting the distribution typically will have only a negligible effect on the distribution’s shape. On top of that, we are making an assumption here that the internal-loss-data-based severity and frequency distributions are good predictors

12 Page 58 of the earlier release of the Observed Range of Practice in the Key Elements of Advanced Measurement Approaches (AMA) issued in July 2009 by the Basel Committee on Banking Supervision, made incorrect reference. The error was fixed in August 2009.
of the future, as long as the forecast horizon is not too long. Scenario data typically pertain to tail events that have not been observed to a reasonable extent in the severity distribution (based on historical data alone) to take appropriate tail shape. The number of scenario data points is expected to be much less compared to the number of historical data points. The nature and purpose of such data is to forecast the frequency of an event within a specified period. The severity distribution also forecasts the frequency of such an event within the same period. In the COM approach we make sure that the forecast of the frequency of an event from the historical data driven severity distribution is at least as often as the forecast of the frequency given in the scenario, if we believe such a scenario may be a possibility. Therefore, in this method we have recalculated the distribution parameters after augmenting the sample with the appropriate amount of scenario data.

Another reason for not doing data refitting is due to the evolving nature of scenario data generation. In our approach we place more value on internal loss data, especially since many institutions have been collecting internal loss data for a considerable period of time.

Given that there are many practical negative implications of prioritizing scenario data over internal data, the method discussed here has been designed to be conditional on the internal loss data. The loss distribution derived from internal data is taken to be the base case. Accordingly, in our approach we never let current or historical probabilities of an event given in the scenario be adjusted downward based on the scenario frequency. Scenario data for all events given allowed for probability to be adjusted only upward. In that process the mass of the distribution always shifts towards right and making the tail of the distribution fatter, if scenarios are used to predict the tail event as opposed to the body event, which is typically the situation. In the world of market risk and especially in a complete market, the implied price (e.g., future or forward price) cannot be less than the current price due to arbitrage. In generating scenarios for operational risk, no such guarantee can be made. This practical assumption enables us to ensure such a guarantee.

For our method, we need the severity of the scenario event to be in a range such as \([a,b]\), where \(b\) can be finite or infinite. We would like to believe that this is a very natural assumption. For people generating scenario data, this will be a convenient way to express the severity. If severity data comes in another form and we can map such data into a severity interval, then we could use our method for those types of data as well. We are aware of instances where scenario data are generated as a point estimate instead of a range (e.g. a $25 million loss occurring once in 25 years). Based on empirical work not discussed here, we found that calculating a bound of \(\pm 15\%-20\%\) around the point estimate will yield a severity range that is very comparable to data that are generated in a range format. Instead of using \(\pm 15\%-20\%\) across all scenarios, a better approach will be to use a tighter bound (say 15%) for the near term event and progressively increase the bound for the longer term event. This will be consistent with the subjective nature of the scenario. In some sense we are accounting for the subjectivity as the number of years increases for the occurrence of the scenario.

One should be careful in arbitrarily enlarging or reducing the interval range. At Step (2) of Section 3.1, we observed that when the range goes to zero, the corresponding probability (equivalently, the frequency) for a continuous severity distribution also goes to zero. It works the other way too. If we arbitrarily increase the range such that it is too large, then the probability (frequency) also becomes high. If we keep the frequency constant and then arbitrarily narrow the range, we will be getting a high but incorrect estimate. On the other hand if we increase the range arbitrarily without correspondingly increasing the frequency, we will get a low but incorrect estimate. This simple fact is very important for conducting a sensitivity study of this method and also for interpreting the scenario data.

The method discussed here can be used for studying the impact of one scenario at a time or the combined effects of a group of scenarios. Therefore, this method can be used for stress testing the internal-data-
based model, for what-if analysis under a set of scenarios, and for measuring the total impact of all scenarios in conjunction with internal loss data for estimating operational risk. An important use of COM will be for communication purposes as well. Using COM, we can express the net effect of a scenario in relation to the current loss experience at an institution and for a particular unit of measure.

The COM evaluation is important in itself. On evaluating scenarios based on the COM, we get a clear understanding of the likelihood of the scenario in relation to all other events and conditional on the current loss experience. This probability study of the event in a scenario can help one to create a financial product for risk transfer contingent on the occurrence of that event. A financial product traded in the market is essentially a promise of a future cash flow with an objective probability. The price (at time zero) of the financial product traded in the market is calculated by converting the objective probability to a risk-neutral probability and then discounting it with a risk-free rate. One could do the same here for the creation of a catastrophe bond (cat bond) or other insurance products using the probability obtained for the events underlying each scenario. A further study in this domain will be useful for possible risk transfer of some of the operational risks classified by each scenario.

The method discussed here assumes independence of the occurrences of each scenario. If we have a set of scenarios that are correlated, then this method can be very easily extended to account for the correlation among scenarios within a unit of measure.

5. Conclusion

In the study of market risk of financial products, the Change of Measure is used as an important evaluation criterion. Using COM, one can properly evaluate the possibility of a future cash flow of the instrument. This principle was the motivation behind our work. We asked, “Why can’t a scenario be evaluated using the Change of Measure approach?” After all, in a sense future cash flows are nothing but a scenario for some economic process. For us, it is the loss generation process driving the severity distribution calibrated by historical loss experience. If operational risk scenario data were of the quality of market risk scenario data (such as option, future, and forward contracts with high liquidity), then we could have used scenario data to calibrate our loss generating process as is done in many market risk modeling approaches such as in Jackwerth and Rubinstein (1996). In the case of operational risk, we are proposing to use COM in validating and updating the loss generating process. Cerny (2009) in chapter 9 discusses COM in the context of market risk and shows how it is used for dynamic allocation of wealth. An extension of this work will be to study the dynamic allocation of risk management resources using COM. As problematic as operational risk scenario analysis data are, the use of scenario data for purposes of operational risk measurement should be given the same priority as internal loss data. MacMinn (2005) differentiated corporate risk exposure between pure and speculative. In MacMinn’s framework operational risk will be classified as a pure risk. Following similar analysis and using COM we can differentiate among various tails risks and their impacts on the value of the firm based on their insurability. An interesting extension of this work will be to analyze using COM various tail risks in the MacMinn’s framework.

Through the construction of implied distributions we have shown how to merge together internal and scenario data in a systematic way. We thoroughly tested our method using actual data from several financial institutions. Based on those implementations and tests we see the use and further development of this method through research for many other applications beyond the measurement of operational loss exposure of an institution. Through this method we could find a very useful and good use of scenario data which were thought to be practically impossible to model in a systematic way. There are some suggestions to call the scenario data as the forward-looking data. We do not like to characterize it this way. We believe in the context of the operational risk measurement all the data elements that the regulatory world has identified are equally important. One needs to find a systematic (with strong theoretical underpinning)
way to extract the information contained in the data elements. The method we presented here is an effort in that direction to combine the internal loss event data and scenario data. It is unfortunate that in the practice of operational risk measurement there are some tendencies to blame the data if we cannot model them with our available knowledge as if the data “have” to be good for the model to work and not the other way.

If we can draw any inference from the experiences reported in other disciplines where scenarios are used as a cornerstone for measuring and managing uncertainty, then we can say that in due course of time the data will get slightly (but not a whole lot) better. We should search for a method to compensate us for the inherent problem of data quality. We believe our approach, which we have presented here for evaluating scenarios using a Change of Measure approach, is a modest step in that direction.

So far the research for estimating operational risk has been heavily skewed in the direction of statistical research in search of the “best” severity distribution. In our opinion, it should also have been thought of as a financial economics problem. If in the end institutions are going to hold billions of dollars as operational risk capital, its measurement should involve some tools and techniques from financial economics. We hope this work will create enough research interest in studying operational risk in the context of economics rather than the current context, which is that of purely (and very often “impurely”) a data fitting exercise. Also, we hope that this work will encourage effective use of scenarios in measuring “Pure” risk in the categorization of MacMinn (2005) such as operational risk and other types of risk covered by an insurance contract.

Although our method was illustrated using the operational risk data, we would like to believe that our approach is of general nature. We would like to think this approach will be applicable in the management decision making process where scenarios are used to adjust or assess the predictive power of a statistical distribution of an underlying event.
References


