International Journal of Information Technology & Decision Making
Vol. 8, No. 4 (2009) 727–747

A PIECEWISE-DEFINED SEVERITY DISTRIBUTION-BASED LOSS DISTRIBUTION APPROACH TO ESTIMATE OPERATIONAL RISK: EVIDENCE FROM CHINESE NATIONAL COMMERCIAL BANKS

JIANPING LI[*,†,§], JICHUANG FENG[‡,¶] and JIANMING CHEN[†,‖]

[†]*Institute of Policy & Management*
Chinese Academy of Sciences
Beijing 100190, P. R. China

[‡]*School of Management*
University of Science and Technology of China
Hefei, Anhui 230026, P. R. China
[§]*ljp@casipm.ac.cn*
[¶]*fengjc@mail.ustc.edu.cn*
[‖]*jmchen@casipm.ac.cn*

Following the Basel II Accord, with the increased focus on operational risk as an aspect distinct from credit and market risk, quantification of operational risk has been a major challenge for banks. This paper analyzes implications of the advanced measurement approach to estimate the operational risk. When modeling the severity of losses in a realistic manner, our preliminary tests indicate that classic distributions are unable to fit the entire range of operational risk data samples (collected from public information sources) well. Then, we propose a piecewise-defined severity distribution (PSD) that combines a parameter form for ordinary losses and a generalized Pareto distribution (GPD) for large losses, and estimate operational risk by the loss distribution approach (LDA) with Monte Carlo simulation. We compare the operational risk measured with piecewise-defined severity distribution based LDA (PSD-LDA) with those obtained from the basic indicator approach (BIA), and the ratios of operational risk regulatory capital of some major international banks with those of Chinese commercial banks. The empirical results reveal the rationality and promise of application of the PSD-LDA for Chinese national commercial banks.

Keywords: Operational risk; LDA; frequency distribution; piecewise-defined severity distribution; Monte Carlo simulation.

1. Introduction

Although financial institutions have always been exposed to operational risk, throughout their history, management of operational risk has attracted serious attention from regulators, managers, and investors only during the last one decade

*Corresponding author.

or so. Operational risk events have, indiscernibly, caused banks considerable financial damage. Take the instance of unauthorized trading at Singapore office of the Barings Bank as an example. Nick Leeson's trading activities resulted in a loss of US$1.4 billion to Barings Bank, eventually resulting in the bank's bankruptcy. The detrimental consequences of exposure to operational risk cannot be overstated. While market risk has traditionally received the attention of financial institutions, operational losses also are an important source of risk for banks, and the capital charge for operational risk will often exceed the charge for market risk.[1,2] Basel II, recommendations on banking laws and regulations, issued by the Basel Committee on Banking Supervision, seeks to assess operational risk, an area previously not clearly defined in Basel I. The supervisory norms seek to promote operational risk management as an integral part of risk management.

Quantifying operational risk is the precondition for formulation of an effective economic capital framework.[3] The only feasible way to manage operational risk successfully is by identifying and minimizing it, which requires the development of adequate quantification techniques.[4] There seems to be a consensus that quantification is needed to better understand operational risk profiles and to estimate the capital required for meeting operational risk.

Unlike market and credit risk, measurement of operational risk faces the challenge of limited availability of data. Furthermore, due to the sensitive nature of operational loss data, banks are not likely to freely share them. Anyone venturing into operational risk management has quickly learned that the process is doomed without robust data.[5,6]

The Basel II Accord suggests three methods for calculating operational risk capital charges: (i) the basic indicator approach (BIA), (ii) the standardized approach (STA) and (iii) the advanced measurement approach (AMA). BIA and STA are simple approaches, and they might be suitable for small banks with a simple range of business activities, while AMA is more sophisticated and much more risk sensitive than the former two approaches. AMA offers maximum elasticity for banks to measure risk. The Basel Committee expects internationally active banks, and banks with significant operational risk, to use the AMA within the overall operational risk management framework.

While considering the different methodologies for operational risk management, we focus on the AMA. In implementing AMA, there is a choice of "top-down" versus "bottom-up" solutions.[7] Top-down operational risk solutions tend to be of very little use in designing procedures to reduce operational risk in any particularly exposed area of the organization. They are inadequate analytical tools as they ignore the organization's processes and procedures. The bottom-up approach analyzes operational risks from the point of view of individual business activities in different areas. Bottom-up models use two different approaches to estimate operational risk: (i) the process approach; and (ii) the actuarial approach.

Although the application of AMA is open to any proprietary model, in principle, by far the most popular methodology is the LDA,[8] which is the actuarial approach.

Recently, the measurement of operational risk has moved towards a data-driven LDA, and many financial institutions have started collecting operational loss data; some have already accumulated several years' data.

LDA is a parametric technique that consists of estimating a frequency distribution of the occurrence of operational losses, and a severity distribution of the economic impact of individual instances of losses. In order to obtain the total distribution of operational losses, these two distributions are combined through η-convolution of the severity distribution with itself, where η is a random variable that follows the frequency distribution. The output of the LDA methodology is a full characterization of the distribution of annual operational losses of the bank. This distribution contains all relevant information for computation of the capital charge, as this requires knowledge of the expected loss, and of the 99.9% percentile of the distribution, as well as necessary inputs for assessing the efficiency of operational risk management procedures. Several researchers have applied LDA to measure operational risk over the past few years.[9,10]

Variations in severity distributions have a great effect on operational risk VaR.[11] In classical LDA, log-normal, weibull and exponential distributions are usually used to model severity losses, but while these distributions can usually fit the high-frequency low-severity data well, they fit badly for low-frequency high-severity data. However low-frequency high-severity events significantly impact operational risk; therefore, these distributions may not have optimal fit in both the body and the tail.

The Extreme Value Theory (EVT) is a well-founded theory that has widely been applied in physical sciences. As an alternative and promising approach, it is used as severity distribution to measure operational risk nowadays. Embrechts has done some pioneering work using the EVT in operational risk measurement.[12–19] EVT uses generalized Pareto distribution (GPD) for measuring severity distribution, and it fits distribution based on low-frequency high-severity data only. Besides, it ignores ordinary losses, which is also a shortcoming; it has been proved that EVT produces a lower estimate of VaR than variance–covariance, historical simulation and conditional historical simulation methods.[20]

In this research, our preliminary tests (Sec. 5) indicate that classic distributions are unable to fit the entire range of operational risk data samples, when modeling the severity of losses. Then, we suppose the use of different severity distributions for high-frequency low-severity data and low-frequency high-severity data may improve the accuracy of measuring operational risk. We propose piecewise-defined distribution, which involves estimating two different distributions, one above a certain threshold, and one below it. Then we combine the two distributions into a third one. We apply GPD to low-frequency high-severity operational loss data, and log-normal, weibull and exponential distributions to high-frequency low-severity loss data. We apply this piecewise-defined severity distribution to operational risk loss database of Chinese national commercial banks, to quantify their operational risks. Empirical results reveal the rationality and the promising nature of this application of the piecewise-defined severity distribution based loss distribution approach (PSD-LDA).

The remaining of this paper is organized as follows. Section 2 describes the basic methodology. Section 3 describes PSD-LDA. In Sec. 4 we outline the characteristics of the data we have collected. Section 5 tests the risk measurement model on real data (conducted with the R software, 2008). Section 6 analyzes the results and compare them. Section 7 presents some conclusions.

2. The Basic Methodology

2.1. *Basic indicator approach*

The BIA uses a single indicator to calculate the capital reserve. In BIA, operational risk capital charge equals to the average over the previous three years, of a fixed percentage of positive annual gross income. Figures for any year in which annual gross income is negative or zero should be excluded from both numerator and denominator when calculating the average. The fixed percentage α is typically 15% of the annual gross income. Hence

$$K = \frac{\alpha \sum_{i=1}^{n} \mathrm{GI}_i}{m} \tag{1}$$

where K is the capital charge, GI is the positive gross income over the previous three years, and m is the number of the previous three years for which gross income is positive.

2.2. *Extreme value theory*

EVT provides a theoretical framework for studying rare events by focusing on tails of probability distributions. Statisticians have used EVT techniques for a long time but they have only recently been proposed in operational risk management.[12] EVT includes two categories of models-Block Maxima Model (BMM) and Peaks over Threshold method (POT) and has a natural way to fit extreme data. The major defect in BMM is that it is very wasteful of data, so it has been largely superseded in practice by methods based on threshold exceedance. The POT method can get credible outcome with relatively small data sets in measuring extreme values via prearranged thresholds. So the POT method is used in this research.

Suppose the following $Z_1, \ldots, Z_n$ are n observations, which are all independently and identically distributed sequences of losses with distribution function $F_Z(z) = P\{Z \leq z\}$, and the corresponding $Y_1, \ldots, Y_n$ are the excess over the threshold μ. We are interested in understanding distribution function F, particularly on its lower tail. Firstly, we describe the distribution over a certain threshold μ, using the GPD, which is the main distribution model for excess over the threshold. The excess over threshold occurs when $Z_i > \mu$. Let F_μ denote the distribution of the excess over the threshold, and F_μ be called the conditional excess distribution function,

$$\begin{aligned}
F_\mu(y) &= P(Z-\mu \le y | Z > \mu), \quad y \ge 0 \\
F_\mu(y) &= \frac{F(\mu+y)-F(\mu)}{1-F(\mu)} \\
&= \frac{F(z)-F(\mu)}{1-F(\mu)} \Rightarrow F(z) = F_\mu(y)(1-F(\mu)) + F(\mu), \quad z \ge \mu
\end{aligned}$$

For a large number of underlying distribution functions F, the conditional excess distribution function $F_\mu(y)$ (for large μ) is well approximated by:

$$\lim_{\mu \to Z_F} \sup_{0 = < y < Z_F - \mu} |F_\mu(y) - G_{\xi,\beta(\mu)}(y)| = 0.$$

For a certain μ, the excess distribution above the threshold may be taken to be exactly the same as the GPD, for parameters β and ξ.

$$F_u(y) \approx G'_{\xi,\beta}(y) = \begin{cases} 1 - \left(1 + \dfrac{\xi}{\beta} y\right)^{-1/\xi}, & \xi \ne 0 \\ 1 - e^{-y/\beta}, & \xi = 0 \end{cases}, \quad u \to \infty \tag{2}$$

where

$$\xi = \begin{cases} \ge 0, & y \in [0, +\infty), \\ < 0, & y \in [0, -\beta/\xi]. \end{cases}$$

An additional parameter $\beta > 0$ has been introduced as the scale parameter, and ξ is the shape parameter (tail index). Both β and ξ are estimated via real data of excess returns. Tail index ξ gives an indication of the heaviness of the tail; the larger ξ is, the heavier is the tail.

Another crucial issue is how to choose a threshold μ, which is the precondition for estimating β and ξ. If threshold μ is too high, there will be a fewer data over threshold μ and estimations of parameters will have big variance. Meanwhile, if threshold μ is too low, $F_\mu(y)$ and $G'_{\beta,\xi}(y)$ will have a big variance.

3. Loss Distribution Approach Based on Piecewise-Defined Severity Distribution (PSD-LDA)

The main procedures of PSD-LDA are similar to procedures of LDA, but the severity distribution in PSD-LDA has two distinct parts associated with ordinary losses and large losses, respectively. Three assumptions are used in the PSD-LDA model: (i) frequency and severity are independent; (ii) two different losses in the same class are independently and identically distributed; and (iii) ordinary and large losses are independent. The amount of operational risk is estimated by processes outlined hereinbelow.

3.1. *Confirming the threshold*

In PSD-LDA, in order to distinguish large losses from ordinary losses, and to estimate β and ξ in GPD, we should choose an appropriate threshold before frequency and severity distributions are estimated. The mean excess plot is adopted to choose threshold μ in this paper. Let $X_{(1)} > X_{(2)} > \cdots > X_{(n)}$, then the mean excess function of sample data is:

$$e(u) = \frac{\sum_{i=k}^{n}(X_i - u)}{n-k-1}, \quad k = \min\{i|X_i > u\}.$$

The mean excess function of a random vector X with finite mean is given by:

$$e(\mu) = E(X - \mu|X > \mu).$$

The mean excess function of the GPD is easily calculated to be:

$$e(\mu) = E(X - \mu|X > \mu) = \frac{\beta + \xi\mu}{1-\xi}. \tag{3}$$

It may be observed that the mean excess function is linear in threshold μ, which is a characterizing property of the GPD. The value for μ can be chosen as the value at which the plotted curve becomes linear.

3.2. *Estimating the parameters of operational risk frequency distribution*

Availability of operational risk frequency distribution is a precondition for using severity distribution to simulate operational risk. The distribution of η, which is the number of loss events during the year of risk measurement, should be estimated. Although any distribution on a set of non-negative integers can be chosen as frequency distribution, the following three distribution families are used most frequently in LDA models: Poisson distribution, negative binomial distribution, and binomial distribution.[9]

Poisson distribution has a natural way to describe the number of unusual incidents in a unit of time or space. Therefore, we decide to exclusively use Poisson distribution in capital calculation, and this decision reduces the complexity of the model since no statistical test for frequency distributions is required. It means that η follows a Poisson distribution, and the probability function of the Poisson distribution is $f(x) = e^{-\lambda}\lambda^{-x}/x!$. The expected value is λ, which is estimated by equating this with the annual average number of loss events.

In addition to reducing the complexity of the model, the Poisson distribution has another characteristic: $\text{Poisson}(\lambda_1) + \text{Poisson}(\lambda_2) = \text{Poisson}(\lambda_1 + \lambda_2)$. It means both ordinary loss and large loss events follow Poisson distribution, and the characteristic implies convenience of simulation. Parameters of the frequency distribution are estimated using the maximum likelihood method. Parameter λ_1 is estimated by the annual average number of loss events for which loss is greater than μ; λ_2 is not greater than μ. So operational risk frequency distribution is Poisson $(\lambda_1 + \lambda_2)$.

3.3. *Estimating the parameters of piecewise-defined severity distribution*

Severity distribution, which specifies the loss size, is the most important component in quantitative operational risk models. In LDA models, the choice of severity distribution usually has a much more severe impact on capital than the choice of frequency distribution.[9] We use a PSD to model severity losses, as follows.

Step 1: uses the parameter model for high-frequency low-severity losses.
Exponential, weibull and lognormal distributions are used to model the body of severity distribution, and goodness of fit is used to choose the best distribution. The maximum likelihood method is used to estimate parameters of ordinary severity loss distribution.

Step 2: uses the EVT model for low-frequency high-severity losses.
The GPD, which is chosen as the severity distribution for the tail, is estimated by using losses that exceed the threshold. For parameters estimation of the tail, maximum likelihood method (MLE), method of moments (MM), and biased and unbiased probability weighted moments (PWMB and PWMU)[21] are used to increase the goodness of fit.

Step 3: combines the two distributions using the threshold value for large losses.
Different severity distributions are constructed at the threshold μ. The cumulative distribution F_1 is based on loss data in the range from 0 to threshold μ. The conditional distribution function F_2 is based on loss data over the threshold μ. Distributions estimated above are combined at the threshold, to derive a piecewise-defined distribution. This is done to improve the goodness of fit, which tends to be poor when using a single distribution of severity. The piecewise-defined distribution is defined as:

$$F(x) = \begin{cases} F_1(x), & 0 < x \leq \mu \\ 1 - (1 - F_1(\mu))(1 - F_2(x)), & x > \mu \end{cases}. \tag{4}$$

3.4. *Simulating the loss amount*

We use the loss frequency distribution estimated in Sec. 3.2, and the number of annual loss instances N is derived. Then the severity of losses of N occurrences X_i $(i = 1, 2, \ldots, N)$ is derived from the severity distribution estimated in Sec. 3.3. The total amount S of operational risk losses in one year equals to:

$$S = \sum_{i=1}^{N} X_i \tag{5}$$

3.5. *Calculating the amount of risk*

In this research, value at risk (VaR), expected shortfall (ES), and economic capital (EC) are used as the amount of risk. Unlike VaR, ES satisfies properties of a coherent risk measure,[22] and captures tail events better. A sound standard, similar to the standard adopted for credit risk, is mandatory. This standard is set to a confidence level of 99.9% for a one-year holding period, but we also choose 99% and 99.98% as confidence intervals. Banks are required to include both the expected loss (EL) and unexpected loss (UL) in their regulatory capital requirements for operational risk, with the UL defined for a one-year holding period at the 99.9% confidence interval.[23] The UL, which equals to VaR minus EL, is economic capital, and regulatory capital, in AMA.

Repeat the procedure proposed in Sec. 3.4 K times and arrange the results in ascending order, as S_j. The greater is the number of times we simulate, the more stable are the results we get, though we also need more time for that. In order to balance the time and stability, we choose to simulate 100,000 times to quantify the operational risk. The amount of risk is:

$$\mathrm{VaR}(\alpha) = S_i, \quad \frac{i-1}{K} \le \alpha < \frac{i}{K}, \quad i = 1, 2, \ldots, K \tag{6}$$

and

$$\mathrm{ES}(\alpha) = \frac{\sum_{i=j}^{K} S_j}{K - j + 1}, \quad \frac{j-1}{K} \le \alpha < \frac{j}{K}, \; j = 1, 2, \ldots, K \tag{7}$$

$$\mathrm{EC} = \mathrm{UL} = \mathrm{VaR} - \mathrm{EL} \tag{8}$$

4. Data

As far as we know, no reliable publicly available data source for operational risk exists in China. This paper utilizes newly available data collected from publicly available information sources, such as newspapers, Internet and court filings. We extract operational risk attributes such as the bank, start time, end time, exposed time, causal type, business line type, loss event type, operational risk loss, province, the key person, and descriptive information concerning each loss, from descriptions of operational risk events. Eight-hundred and sixty pieces of operational risk loss data of Chinese commercial banks were obtained, for the period 1995–2006, and we classify these loss events into eight categories, for individual banks – Bank of Communications (BCM), Bank of China (BOC), China Construction Bank (CCB), Industrial and Commercial Bank of China (ICBC), and Agricultural Bank of China (ABC), some other banks are clubbed together as several commercial banks (SCB), anonymous commercial banks (ACB), and non-national commercial banks (NNCB). SCB means the operational risk loss occurred in several commercial banks but at least one of them is a national commercial bank. ACB denotes operational risk loss occurred but we do not know in which bank it occurred. NNCB denotes operational risk loss occurred in other banks except national commercial banks. Details of

Table 1. Operational risk loss events frequency in each bank.

Bank	BCM	BOC	CCB	ICBC	ABC	SCB	ACB	NNCB
Times	28	84	97	106	192	17	196	171
Percent	3.1%	9.4%	10.9%	11.9%	21.5%	1.91%	22%	19.2%

Table 2. Operational risk loss events frequency in each year.

Year	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
Times	29	43	47	57	59	103	74	67	94	136	97	74

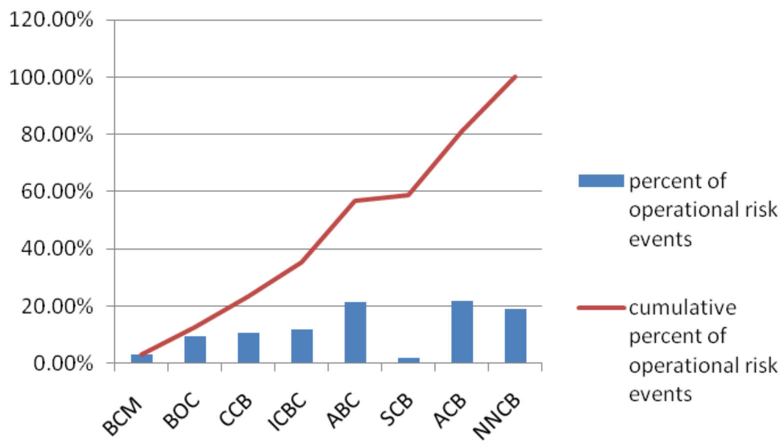


Fig. 1. Operational loss events in percent by banks.

the number of loss occurrences in each commercial bank, over the entire period (1995–2006) are shown in Table 1 and the total number of loss events in each year, in all banks taken together, during this period, is shown in Table 2. In addition, we depict the frequency percent of banks in Fig. 1 and loss frequency per year in Fig. 2.

Total assets of non-national commercial banks were about half of total assets of financial institutions as at the end of 2006, in China. However, only 19.2% instances of operational loss occurred in non-national commercial banks. Since usually operational loss is related with total assets, 19.2% of occurrences cannot denote the non-national commercial banks operational losses. Hence, from the viewpoint of national commercial banks, the sample data can be considered data comprising loss data of other banks, in addition to their own internal loss data.

Figure 2 shows that the frequency of loss events after the year 2000 is obviously higher than before 2000. The reason of this phenomenon may be that few people understood operational risk before 2000; that operational risk is one of the most

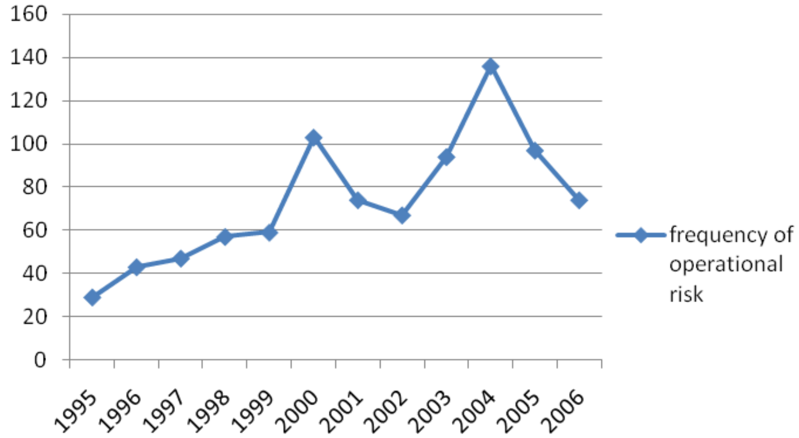


Fig. 2. Operational loss events by year.

crucial risks may not have been realized. The Basel Committee suggests that banks need at least 7 years operational loss data if AMA is used to calculate operational risk capital. So we choose the frequency of operational risk loss per year from 2000 to 2006 for estimating parameters of loss frequency distribution. However, all loss events that occurred from 1995 to 2006 are used in estimating parameters of loss severity distribution.

To determine the characteristics of the sample data set used for operational risk measurement, we evaluate the tail heaviness of the sample distribution. As shown in Fig. 3, distribution of the sample data set exhibits much a heavier tail than that of normal distribution.

We also calculate statistical characteristics of the sample data; as shown in Table 3, operational risk losses obviously have large skewness and kurtosis.

5. Empirical results

5.1. *Confirming the threshold*

As mentioned in Sec. 3, we have specified the mean excess plot to choose the threshold, in Fig. 4. We can see there is a structural change at 20,000; mean excess is linear in the threshold beyond 20,000, which is a characterizing property of the GPD. So in mean excess plot of the sample data, we suppose 20,000 is an appropriate threshold.

5.2. *Estimating the parameters of operational risk frequency distribution*

Using the maximum likelihood method and frequency of operational risk loss per year (larger than 20,000 ten thousand CNY from 2000 to 2006), we can get the

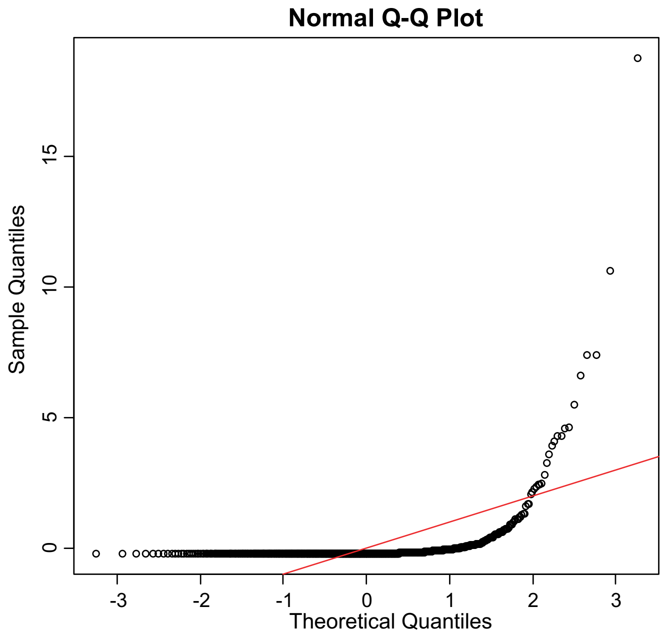


Fig. 3. Q–Q normal plot for data sample.

Table 3. Statistical characteristics of the sample data set.

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Skewness	Kurtosis
3.000e−02	6.400e+01	4.390e+02	1.200e+04	3.728e+03	1.000e+06	10.72768	156.8218

Unit: ten thousand CNY.

parameter $\lambda_1 = 8$ of large loss frequency distribution. Using the frequency of operational risk loss per year which is not larger than 20,000 ten thousand CNY, from 2000 to 2006, we can get the parameter $\lambda_2 = 84$ of ordinary loss frequency distribution. So the frequency distribution is Poisson $(\lambda_1 + \lambda_2) = \text{Poisson}(92)$.

5.3. *Estimating the parameters of operational risk severity distribution*

5.3.1. *Ordinary loss severity distribution*

The ordinary loss severity distribution is estimated by using the full data set. We apply three different distributions (exponential, weibull and lognormal) for ordinary loss severity distribution. The estimated parameters are shown in Table 4.

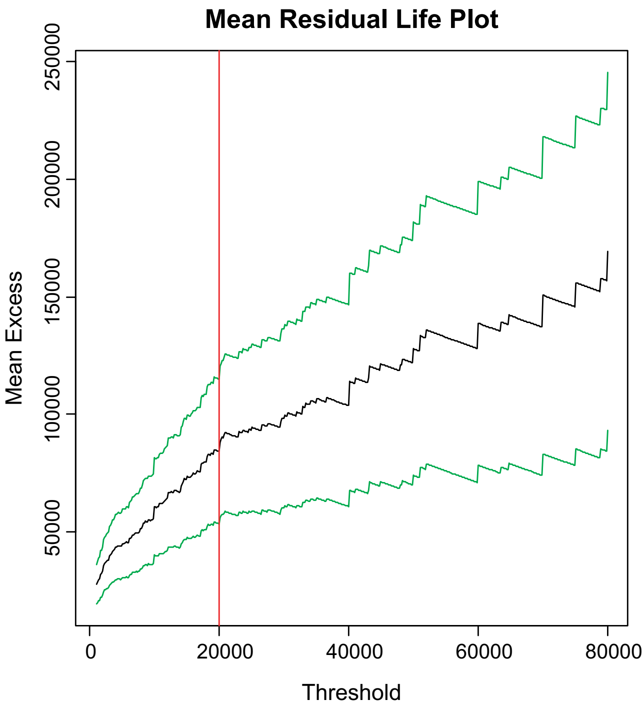


Fig. 4. Mean excess plot; unit: ten thousand CNY.

Table 4. Results of parameter estimation.

	Exponential	Weibull		Lognormal	
MLE	Rate	Shape	Scale	Meanlog	sdlog
	8.331e–05	0.365	1.874e+03	6.178	2.846

Table 5. p-value of the hypothesis test.

Distribution	p-value (chi-square test)	p-value (K–S test)
Exponential	0	2.2e−16
Weibull	6.47e−05	0.001
Lognormal	0.051	0.267

p-values of chi-square and K–S tests, of lognormal, weibull and exponential distributions fitting the loss data, are given in Table 5. Only lognormal chi-square test p-value is over 5%; so the data follows lognormal distribution with meanlog = 6.178, and sdlog = 2.846 is accepted. From the K–S p-value, we can get the same conclusion.

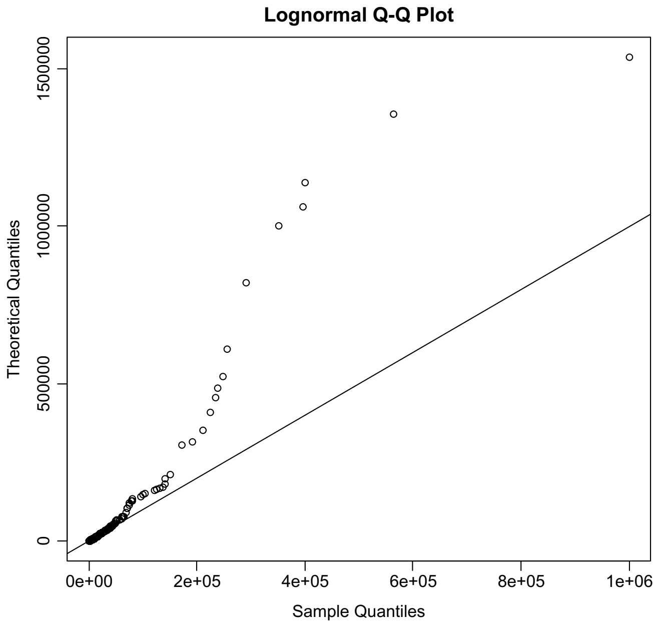


Fig. 5. QQ-plots of data sample.

5.3.2. *Large loss severity distribution*

Although the lognormal p-value is greater than when the significance level is fixed at least at 5%, it is only marginally higher, i.e. 5.1%. The p-value is so small because the lognormal distribution fits the tail badly, as shown in Fig. 5. It indicates that classical distributions are unable to fit the entire range of operational risk data sample well for severity distribution. As a consequence, if lognormal distribution is used as severity distribution, it will over estimate operational risk capital.

The lognormal distribution fits the body well, but is not suitable for severity distribution. So we choose lognormal distribution as the ordinary loss severity distribution, but not as large loss severity distribution.

The large loss severity distribution is estimated by using loss amounts that exceed the threshold 20,000 (ten thousand CNY). We apply GPD for large loss data, and for parameters estimation, MLE, MM, PWMB and PWMU are used. Results of estimations are shown in Table 6.

Table 6. Results of parameter estimations.

MLE		MM		PWMU		PWMB	
Scale	Shape	Scale	Shape	Scale	Shape	Scale	Shape
88500	0.2291	60660	0.3146	45510	0.4857	46446.15	0.4752

We provide PP-plots and QQ-plots of severity losses, together with confidence intervals constructed for $p = 0.95$, in Fig. 6. If points lay outside of these confidence bands, we would have to reject the estimation approach. So PWMU and PWMB are better estimation approaches than MLE and MM, because some observations are outside of these confidence bands in PP-plots of MLE and MM. Confidence intervals associated with the largest observations may seem too wide. But considering the heavy-tail of the distribution we deal with, it is natural that the largest losses have a wide admissible range in terms of theoretical quantiles. It is demonstrated that POT has a natural way to fit low-frequency high-severity losses. It is better to use PWMU and PWMB to estimate GPD based on our data sample. So large loss severity distribution follows Pareto (20000, 45510, 0.4857) or Pareto (20000, 46446.15, 0.4752).

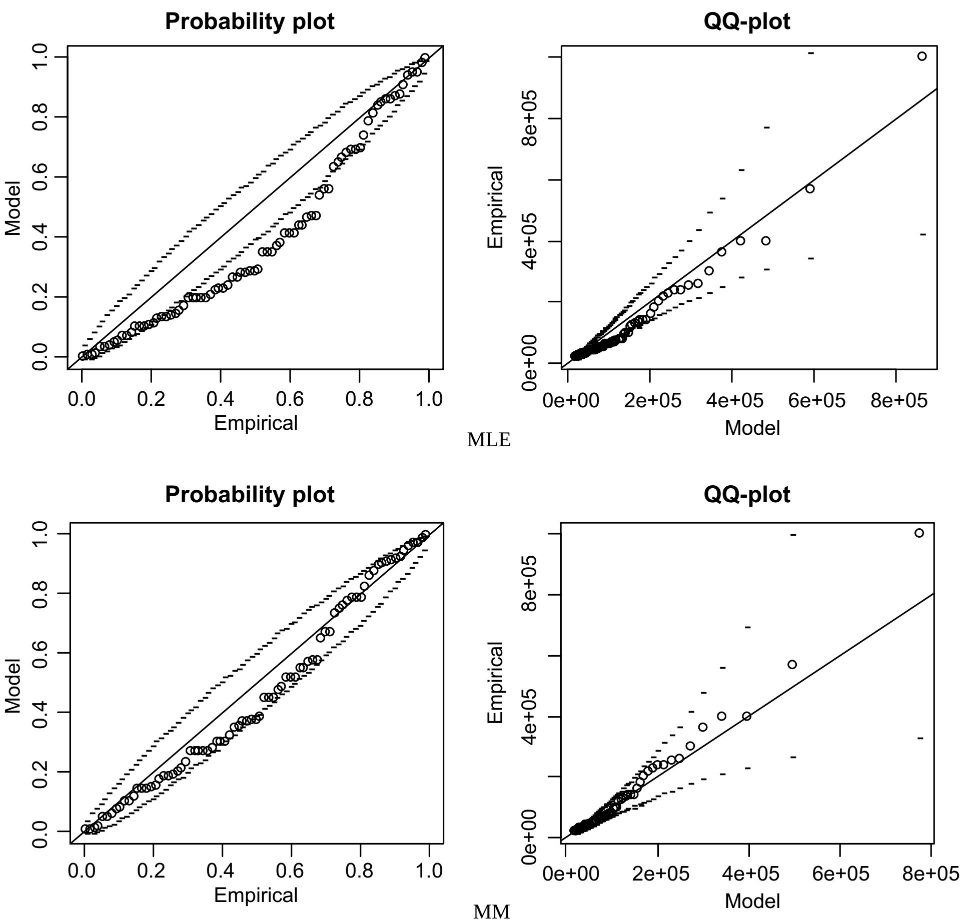
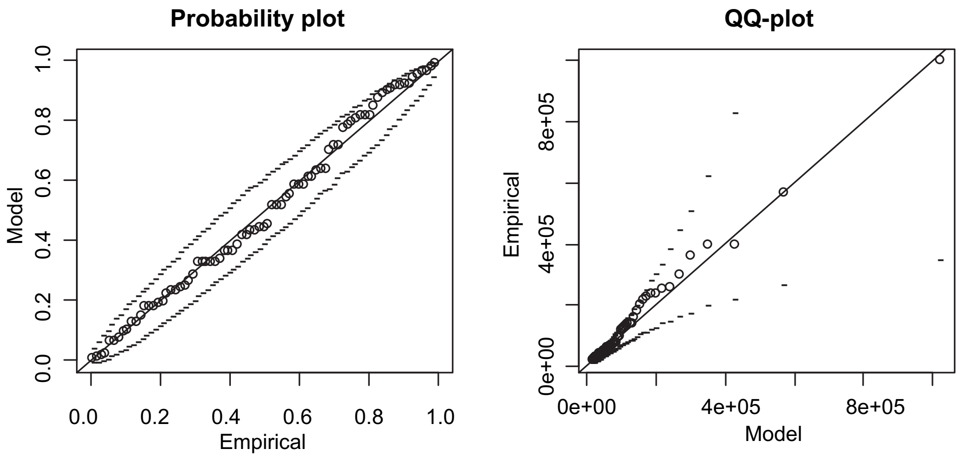
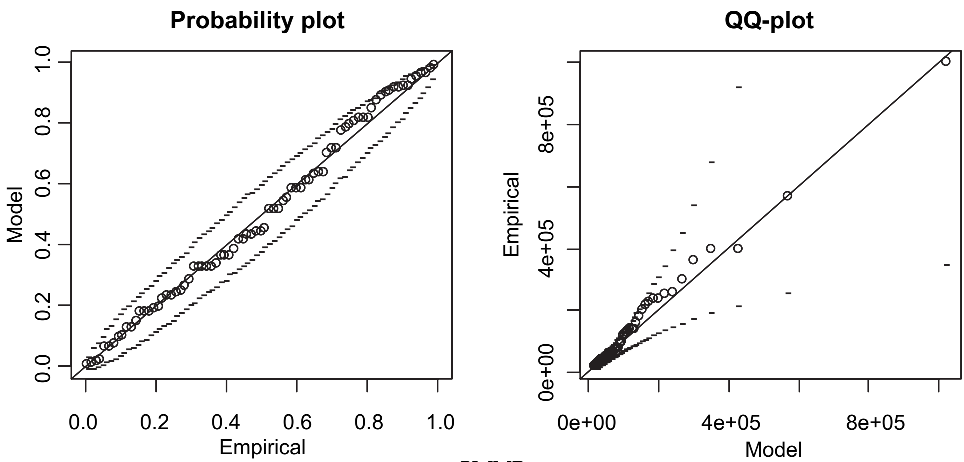


Fig. 6. PP-plots and QQ-plots of different parameters estimation.



PWMU



PWMB

Fig. 6. (*Continued*)

5.3.3. *Piecewise-defined severity distribution*

Combining the lognormal distribution for ordinary loss, and GPD for large loss, and using the threshold μ, the piecewise-defined distribution can be defined as:

$$F(x) = \begin{cases} \Phi\left(\dfrac{\log x - u}{\sigma}\right), & 0 < x \le \mu \\ 1 - \left(1 - \Phi\left(\dfrac{\log \mu - u}{\sigma}\right)\right)\left(1 - \dfrac{N_\mu}{N}\left(1 + \xi\dfrac{x-\mu}{\beta}\right)^{-\frac{1}{\xi}}\right), & x > \mu \end{cases} \tag{9}$$

Considering the parameters estimated, the piecewise-defined distribution can be expressed as:

$$F(x) = \begin{cases} \Phi\left(\dfrac{\log x - 6.18}{2.85}\right), & 0 < x \leq 20000 \\ 1 - \left(1 - \Phi\left(\dfrac{\log x - 6.18}{2.85}\right)\right) & \\ \quad \times \left(1 - \dfrac{8}{92}\left(1 + 0.4857\dfrac{x - 20000}{45510}\right)^{-\frac{1}{0.4857}}\right), & x > 20000 \end{cases}$$

When PWMU is used; or

$$F(x) = \begin{cases} \Phi\left(\dfrac{\log x - 6.18}{2.85}\right), & 0 < x \leq 20000 \\ 1 - \left(1 - \Phi\left(\dfrac{\log x - 6.18}{2.85}\right)\right) & \\ \quad \times \left(1 - \dfrac{8}{92}\left(1 + 0.4752\dfrac{x - 20000}{46446}\right)^{-\frac{1}{0.4752}}\right), & x > 20000 \end{cases} \tag{10}$$

When PWMB is used.

5.4. *Results of PSD-LDA*

When PSD-LDA is used, we simulate risk losses for a year 100,000 times, to quantify the operational risk of Chinese national commercial banks. The amount of operational risk for Chinese national commercial banks is shown in Table 7.

As shown in Table 7, different estimation approaches lead to different results, but the difference is so small that stable results could be obtained using the two

Table 7. Amount of operational risk for Chinese national commercial banks.

		VaR	EL	RC/EC	ES
99%	PWMB	34,223	10,285	23,937	56,753
	PWMU	34,839	10,285	24,553	59,015
	Mean*	34,531	10,285	24,245	57,884
99.9%	PWMB	89,403	10,285	79,118	151,490
	PWMU	94,259	10,285	83,974	161,815
	Mean*	91,831	10,285	81,546	156,652.5
99.98%	PWMB	185,350	10,285	175,064	282,614
	PWMU	198,812	10,285	188,527	306,239
	Mean*	192,081	10,285	181,795.5	294,426.5

Unit: million CNY.

estimation approaches. We use the mean of the results from PWMU and PWMB in this research to analyze operational risk. So according to the requirements of Basel II, under the 99.9% confidence interval, the overall operational risk of Chinese national commercial banks is 9.18×10^{10} CNY; the overall operational risk regulatory (economic) capital, which equals to VaR minus EL, is 8.15×10^{10} CNY every year; and the ES under the 99.9% confidence interval is 1.57×10^{11} CNY. If all national commercial banks in China want to get higher ratings (for example, AA), 1.82×10^{11} CNY operational risk regulatory (economic) capital is needed every year; VaR and ES under the 99.98% confidence interval are 1.92×10^{11} and 2.94×10^{11} CNY, respectively. The data under 99.9% confidence interval can also be analyzed using the same analysis approach. The ES which allows better capture of tail events is much larger than VaR under the same confidence interval. So if the ES as required in Basel II is to be achieved, commercial banks need much more regulatory capital.

6. Results Analysis and Comparison

6.1. *Results analysis*

In Table 8, we present the gross income of Chinese national commercial banks in 2005, 2006 and 2007, and the operational risk estimated by using the BIA. So the total operational risk of Chinese national commercial banks is 96,311.8 million CNY, according to formula (1).

We have measured the operational risk capital charge of Chinese national commercial banks using PSD-LDA and BIA. The operational risk capital charge, estimated by using different approaches, is shown in Table 9. The Basel II Accord suggests three methods for calculating operational risk capital charge, and results of the three approaches should be in conformity with each other. So comparing the two values in Table 9, result of PSD-LDA has its rationality. The BCBS (2004) says that if banks move from BIA along a continuum toward AMA, they will be rewarded with a lower capital charge. Our empirical results show that capital charges obtained with AMA are low when compared to BIA; if AMA is used,

Table 8. Gross income of Chinese national commercial banks.

Year	ICBC	ABC	CCB	BOC	BCM	Chinese National Commercial Banks
2005	162,378	53,893	127,268	116,028	35,214	494,781
2006	180,705	87,499	150,212	137,628	43,203	599,247
2007	254,157	114,830	219,459	182,712	61,050	832,208
Op-risk using BIA	29,862	12,811.1	24,846.95	21,818.4	6973.35	96,311.8

Unit: million CNY.
Notes: Gross incomes of Chinese national commercial banks come from annual reports of BCM, BOC, CCB, ICBC and ABC of 2005, 2006 and 2007.

Table 9. Operational risk capital charges using different approaches.

Approach	BIA	PSD-LDA
Operational risk	9.63×10^{10}	8.15×10^{10}

Unit: CNY.

banks need less regulatory capital charges than charges estimated using BIA, which means banks save costs. So it is worthwhile adopting AMA to replace BIA.

6.2. *Results comparison with some other major banks' capital charges*

We compare the results with some foreign banks' operational risk capital charges. Table 10 shows comparison of some foreign banks operational risk capital charges, including Citigroup, Bank of America Corp., JP Morgan and AIG.

From column 6, we can see under BIA, the ratio of operational risk regulatory capital to total regulatory capital is from 4.45% to 10.53%, and the ratio of Chinese national commercial banks is 8.24%. Column 7 shows under AMA, the ratio of operational risk regulatory capital to total regulatory capital is from 4.65% to 8.03%, and the ratio of Chinese national commercial banks is 6.97%. The above analysis also indicates that PSD-LDA offers reliable estimates for Chinese national commercial banks. Although the two ratios are within the common intervals, operational risk management is relatively a new phenomenon in China; Chinese commercial banks should refer to international best practices in the banking sector to strengthen their competitiveness in the context of financial globalization and financial innovation.

Table 10. Detailed data of different banks.

Name of Bank	Total Assets	Total Regulatory Capital: Basel II Approach	Op-risk Regulatory Capital: Basel II AMA	Op-risk Regulatory Capital: Basel II BIA	The Ratio of Op-risk Regulatory Capital: Basel II BIA	The Ratio of Op-risk Regulatory Capital: Basel II AMA
Citigroup*	$1,484,101	$100,899	$8100	$10,621	10.53%	8.03%
Bank of America Corp.*	$1,044,660	$92,266	NA	$6065	6.57%	NA
JP Morgan*	$1,157,248	$96,807	$4500	$4305	4.45%	4.65%
AIG*	$1,157,248	$73,317	$5388	$4875	6.65%	7.35%
Chinese national commercial banks	¥22,539,040[b]	¥1,169,430[b]	¥81,546	¥96,311.8	8.24%	6.97%

Unit: million US$ for foreign banks, million CNY for Chinese banks.
[a]For details refer to Ref. 24.
[b]Details are computed using data in annual reports of BCM, BOC, CCB, ICBC and ABC of 2006.

7. Conclusion

Banks have made limited progress in quantifying their operational risk, despite many banks having reported operational losses over the past decade in China. The main purpose of this paper is to introduce a new methodology for measuring operational risk and estimating the required capital.

This paper analyzes available data of Chinese commercial banks, and proposes a piecewise-defined severity distribution based loss distribution approach (PSD-LDA) to estimate operational risk of banks. PSD involves estimating two different distributions, one above a certain threshold and one below it. Then we combine the two distributions into a third one. Loss distribution approach (LDA) with Monte Carlo simulation is used to estimate the operational risk; GPD is applied to low-frequency high-severity operational loss data, and log-normal, weibull and exponential distributions are applied to high-frequency low-severity loss data.

In order to evaluate the performance of the proposed method, we implement it to measure commercial banks annual operational risk and capital charges of Chinese national commercial banks, and then compare the results with from BIA. We also compare the risk regulatory capital needed, as estimated by PSD-LDA and BIA of Chinese commercial banks, with corresponding figures for some major international banks. The empirical results reveal the rationality of application of the PSD-LDA for estimating capital requirements of Chinese national commercial banks.

One of the possible future developments of this research would be to apply our methodology to other databases, in order to find out the advantages of application of PSD-LDA. Some other methods such as simulation optimization,[25] and methods for choosing "optimal" distributions among the set of distributions[26] may be other promising tools to measure operational risk. At last, we emphasize that this paper has to be primarily considered in the context of methodological aspects; though our data is real data, it may be biased data. The full potential of this very promising area of research will be realized only after complete operational data of banks are available.

Acknowledgments

We would like to thank professor Cheng-Few Lee, Rutgers Business School, Rutgers University, the state University of New Jersey, for his helpful comments and constructive suggestions.

The research has been partially supported by a grant from the National Science Foundation of China (#70701033, #70531040).

References

1. P. D. Fontnouvelle, V. Jesus-Rueff, J. Jordan and E. Rosengren, *Using Loss Data to Quantify Operational Risk*, Technical Report (Federal Reserve of Boston, April 2003).

2. P. D. Fontnouvelle, V. Jesus-Rueff, J. Jordan and E. Rosengren, Capital and risk: new evidence on implications of large operational losses, *J. Money Credit Bank.* **38** (2006) 1819–1846.
3. K. Fujii, Building Scenarios, in *Operational Risk: Practical Approaches to Implementation*, in eds. E. Davias, (Risk Books, London, 2005), pp. 169–178.
4. K. Bocker and C. Kluppelberg, Operational VAR: A closed-form approximation, *RISK* (Dec. 2005) 90–93.
5. L. Muzzy, The pitfalls of gathering operational risk data, *RMA J.* **85** (2003) 58–62.
6. J. V. Rosenberg and T. Schuermann, A general approach to integrated risk management with skewed, fat-tailed risks, *J. Financial Econ.* **79** (2006) 569–614.
7. C. Cornalba and P. Giudic, Statistical models for operational risk management, *Physica A* **338** (2004) 166–172.
8. A. Chapelle, Y. Crama, G. Huebner and J. P. Peters, Practical methods for measuring and managing operational risk in the financial sector: A clinical study, *J. Bank. Finance* **32** (2008) 1049–1061.
9. F. Aue and M. Kalkbrenner, LDA at work, Deutsche Bank's approach to quantifying operational risk, *J. Oper. Risk* **1** (2006) 49–93.
10. K. Dutta and J. Perry, A tale of tails: An empirical analysis of loss distribution models for estimating operational risk capital, Technical Report, 06-13 (Federal Reserve Bank of Boston, 2007).
11. Mori, T. Kimata and T. Nagafuji, The effect of the choice of the loss severity distribution and the parameter estimation method on operational risk measurement, Technical Report (Bank of Japan, 2007).
12. P. Embrechts, C. Klüppelberg and T. Mikosch, *Modeling Extremal Events for Insurance and Finance* (Springer, Berlin, 1997).
13. P. Embrechts, H. Furrer and R. Kaufmann, Quantifying regulatory capital for operational risk, *Deriv. Use Trad. Regul.* **9** (2003) 217–233.
14. P. Embrechts and V. Chavez-Demoulin, Smooth extremal models in finance and insurance, *J. Risk Insur.* **71** (2004) 183–199.
15. P. Embrechts and G. Puccetti, Aggregating risk capital, with an application to operational risk, *Geneva Risk Insur. Rev.* **31** (2006) 71–90.
16. P. Embrechts and G. Puccetti, Aggregating risk across matrix structured loss data: The case of operational risk, *J. Oper. Risk* **3**(2) (2008) 29–44.
17. V. Chavez-Demoulin, P. Embrechts and J. Neslehova, Quantitative models for operational risk: Extremes, dependence and aggregation, *J. Bank. Finance* **30** (2006) 2635–2658.
18. M. Degen, P. Embrechts and D. Lambrigger, The quantitative modeling of operational risk: Between g-and-h and EVT, *Astin Bull.* **37** (2007), 265–291.
19. J. Neslehova, P. Embrechts and V. Chavez-Demoulin, Infinite mean models and the LDA for operational risk. *J. Oper. Risk* **1** (2006) 3–25.
20. L. Kalyvas and A. Sfetsos, Does the application of innovative internal models diminish regulatory capital? *Int. J. Theore. Appl. Finance* **9** (2006) 217–226.
21. S. Coles, *An Introduction to Statistical Modelling of Extreme Value* (Springer Series in Statistics, London, 2001).
22. P. Artzner, F. Delbaen, J.-M. Eber and D. Heath, Coherent measures of risk, *Math. Finance* **9** (1999) 203–228.
23. BCBS, *International Convergence of Capital Measurement and Capital Standards: A Revised Framework, Basel* (Bank for International Settlements, Basel, Switzerland, 2005).
24. D. Ingram and K. S. Tan, Operational risk and reputational risk: An important part of ERM, *Risk Manage. Brief* **2** (2006).

25. M. Better, F. Glover, G. Kochenberger and H. B. Wang, Simulation optimization: Application in risk management, *Int. J. Inform. Technol. Decision Making* **4** (2008) 571–587.
26. L. V. Utkin, Risk analysis under partial prior information and nonmonotone utility functions, *Int. J. Inform. Technol. Decision Making* **4** (2007) 625–647.
27. R Development Core Team, R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria (2008), http://www.R-project.org.