Operational Risk Aggregation Based on Business Line Dependence: A Mutual Information Approach

Wenzhou Wang, 1 Limeng Shi, 1 and Xiaqian Zhu 2

1 School of Economics and Business Administration, Beijing Normal University, Beijing 100875, China
2 Institute of Policy and Management, Chinese Academy of Sciences, Beijing 100190, China

Correspondence should be addressed to Wenzhou Wang; wangwenzhou2007@126.com

Received 14 January 2016; Accepted 31 March 2016

Academic Editor: Francisco R. Villatoro

Copyright © 2016 Wenzhou Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The dependencies between different business lines of banks have serious effects on the accuracy of operational risk estimation. Furthermore, the dependencies are far more complicated than simple linear correlation. While Pearson correlation coefficient is constructed based on the hypothesis of a linear association, the mutual information that measures all the information of a random variable contained in another random variable is a powerful alternative. Based on mutual information, the generalized correlation coefficient which can capture both linear and nonlinear correlation can be derived. This paper models the correlation between business lines by mutual information and normal copula. The experiment on a real-world Chinese bank operational risk data set shows that using mutual information to model the dependencies between business lines is more reasonable than linear correlation.

1. Introduction

As one of the major methods for operational risk estimation, loss distribution approach (LDA) is favored by many regulators and practitioners [1, 2]. However, in the framework of LDA, the aggregation of operational loss across business lines or event types (or both) is still an important issue [3, 4]. Traditionally, the total capital is calculated by summing capital of each business line, which implies perfect linear dependence between business lines. Thus, diversification effects or compounding effects might be ignored, leading to the overestimation or underestimation of operational risk capital [1]. However, Basel II is vague about the correlation that is supposed to be considered between business lines and does not give a specific method due to the difficulty of assessing the dependence [5].

Copula method is now gaining popularity in capturing dependence [6, 7]. In the framework of LDA, Böcker and Klüppelberg used Levy-copula to capture the dependence structure between the operational losses of different business lines [4]. Carla et al. applied the expectation maximization algorithm to estimate the parameters of the left truncated frequency and severity distributions and used copula to calculate the risk capital [8]. Annalisa and Claudio applied copula to model the dependence and calculated the risk capital using different types of operational loss data from the earthquake insurance [9]. When referred to copula, especially elliptical copula, the input parameter is a matrix of linear correlation. However, Frachot et al. discussed the limitations of operational risk estimation based on linear correlation [2].

The linear correlation coefficient is a normalized covariance which only accounts for linear (or linearly transformed) relationships between variables [10]. Therefore, this statistic cannot capture dependence when there are nonlinearities. This can be shown in a simple case of two variables \((X, Y)\) where \(Y = X^2\) and \(X\) follows standard normal distribution. Although variable \(Y\) is completely determined once \(X\) is known, the covariance and the linear correlation of them are zero. For another example, consider the case of \(X\sim\) Lognormal \((0, 1)\) and \(Y = X^\sigma\). Even though \(X\) and \(Y\) have an obvious relationship, the linear correlation coefficient between them approaches zero as \(\sigma\) increases to infinity. Generally speaking, linear independence is not synonymous with independence. However, there is always a misconception of the weak correlation between two variables when linear correlation coefficient is zero.
In practice, linear correlation may not be an ideal measure of dependence because nonlinear relationship exists widely between variables, especially in complex financial activities [10–12]. Besides, it is well known that the linear correlation coefficient is very sensitive to extreme data and sample size. This provides another important reason for discarding linear correlation as a reliable measure of dependence. Therefore, one of the key challenges that lies in operational risk aggregation is whether the dependence measures can capture the generalized dependence including both linear and nonlinear correlation relationship [4,13].

Mutual information, an important part of Shannon’s information theory, measures the information of a random variable contained in another random variable [14,15]. In other words, it measures how much the uncertainty of one variable can be reduced if the other one is known [16]. If two variables are independent, mutual information between them is zero. On the contrary, if two variables are identical, then mutual information is equal to the entropy of either one of them. Mutual information has been used as a criterion to test independence between stochastic processes by Fernandes and Néri [17] and to estimate lag in time series by Granger and Lin [18]. It also has been applied in the spatial case to register images [19]. In feature extraction and parameter selection, the applicability of mutual information has been described by Brillinger [16]. Particularly, in the bivariate case, mutual information is the Kulback-Liebler distance between joint distribution and the product of its marginals [20]. In contrast to the linear correlation coefficient, mutual information is a better measure of correlation relationship for it can capture generalized dependence between random variables [21].

The objective of this paper is to use mutual information to model the dependence between business lines, so as to overcome the weaknesses of the commonly used traditional linear approaches. To be specific, firstly, a generalized correlation coefficient can be derived based on the mutual information. Then, this type of correlation coefficient is used as the input parameter of copula function. Finally, by using copula to capture the dependence between the frequencies of different business lines, the operational risk can be calculated in the framework of LDA. In the empirical analysis, the proposed method is applied to operational risk aggregation based on a real-world Chinese bank operational risk data set. The remainder of this paper is organized as follows. In Section 2, the concepts of mutual information and generalized correlation coefficient (normalized mutual information), the bootstrap procedure for estimating the mutual information, the hypothesis testing on independence, the choice of copula, and the capital computation process are presented. In Section 3, the generalized correlation coefficient based mutual information and the linear correlation coefficient are compared based on a real-word Chinese bank operational risk data set. Section 4 summarizes the conclusions.

2. Operational Risk Aggregation Based on Mutual Information

Many previous studies consider that the loss frequencies of different business lines are correlated with each other [1, 4, 22]. Empirically, loss frequency correlation could be examined and measured by computing the historical correlation between the past frequencies of operational risk events. Furthermore, incorporating correlation between frequencies does not destroy the nature of LDA model, so this type of correlation can be taken into account at minimal cost [2]. In contrast, severity correlation is difficult to tackle in LDA model because it changes the basic foundations of the standard LDA model [2]. It requires building an entirely new family of models and such an extension is out of the scope of this paper. Therefore, in this paper, we choose to consider frequency correlation by using mutual information.

2.1. Mutual Information. Mutual information $I(X,Y)$ between variables $X$ and $Y$ measures the information of $X$ (or $Y$) contained in $Y$ (or $X$) or the reduction of the uncertainty of $X$ (or $Y$) due to the knowledge of $Y$ (or $X$) [16, 18, 19]. Let $P(X = x_i), i = 1,\ldots,N_x$, denote the probability distribution of $X$, where $N_x$ is the number of possible values of $X$. Let $P(Y = y_j), i = 1,\ldots,N_y$, denote the probability distribution of $Y$, where $N_y$ is the number of possible values of $Y$. Then, the specific mathematical expression of $I(X,Y)$ for discrete distribution is given by the following [23–26]:

$$
I(X,Y) = H(Y) - H(X|Y) = H(X) - H(Y|X)
$$

$$
= \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} P(X = x_i, Y = y_j) \ln \frac{P(X = x_i, Y = y_j)}{P(X = x_i)P(Y = y_j)},
$$

where $H(X)$ and $H(Y)$ denote the entropy of $X$ and $Y$, respectively, $H(Y|X)$ and $H(X|Y)$ denote the conditional entropy of $Y$ given $X$ and the conditional entropy of $X$ given $Y$, respectively, and $H(X,Y)$ denotes the joint entropy of $X$ and $Y$.

Some important properties of mutual information are:

1. $I(X,Y) = I(Y,X)$,
2. $I(X,Y) = H(X)$,
3. $I(X,Y) \leq H(X), I(X,Y) \leq H(Y)$.

The mutual information is always nonnegative and symmetric. It is equal to zero only if $X$ and $Y$ are independent. Granger et al. consider that a measure of functional dependence for a pair of random variables $X$ and $Y$ is required to satisfy the following six properties [27]:

1. It is well defined for both continuous and discrete variables.
2. It is normalized to zero if $X$ and $Y$ are independent and lie in $(0,1)$.
3. The absolute value of the measure should be equal to 1 if there is an exact nonlinear relationship between the variables.
4. It is equal to or has a simple relationship with the linear correlation coefficient in the case of a bivariate normal distribution.
(5) It is metric; that is, it is a true measure of distance and not just of divergence.

(6) The measure is invariant under continuous and strictly increasing transformations function.

Mutual information can satisfy most of the desirable properties of a good dependence measure [11, 15, 21, 28–30]. Specifically, it satisfies (1) and (6) easily [15] and after some transformations it also satisfies properties (2)–(4). Since $H(X) \geq H(X | Y)$, we have $I(X | Y) \geq 0$. Besides, $I(X | Y) = 0$ only if $X$ and $Y$ are statistically independent. The definition of mutual information decides that $0 \leq I(X, Y) \leq \infty$, which makes the comparisons between different samples difficult. To obtain a statistic that satisfies property (4) and does not lose properties (1) to (3), Granger and Lin used a standard measure for mutual information as the following [18]:

$$\lambda(X, Y) = \sqrt{1 - \exp(-2I(x, y))},$$  \hspace{1cm} (2)

where $\lambda(X, Y)$ is often called generalized correlation coefficient or global correlation coefficient. Generalized correlation coefficient varies between 0 and 1, and so it is directly comparable to the linear correlation coefficient. For a special case of $(X, Y)$ following bivariate normal distribution, the mutual information can be calculated by

$$I(X, Y) = -\frac{1}{2} \ln (1 - \rho^2).$$  \hspace{1cm} (3)

In this case, we have $\lambda(X, Y) = |\rho(X, Y)|$. The generalized correlation coefficient does not conform to the triangle inequality and so it cannot satisfy property (5). However, compared with the linear correlation coefficient, it could satisfy most of the properties. In this sense, it is a natural measure of the correlation between variables.

2.2. Mutual Information in a Simple Nonlinear Case. In order to illustrate that mutual information can detect nonlinearity well, we consider a simple case of two variables $(X, Y)$ for simplicity.

Let $X \sim N(0, 1)$ and $Y: y_i = x_i^2, i = 1, 2, \ldots, n$. Then, $Y$ follows the distribution as

$$f(y) = \frac{1}{2^{k/2} \Gamma(k/2)} y^{k/2-1} e^{-y/2}$$

$$= \frac{1}{2^{k/2} \Gamma(1/2)} y^{1/2} e^{-y/2} \quad (y \geq 0),$$  \hspace{1cm} (4)

where $k$ denotes a positive integer that specifies the number of degrees of freedom and $\Gamma(k/2)$ denotes the gamma function that has closed-form values at the half integers; for example, $\Gamma(1/2) = \sqrt{\pi}$.

Therefore, $H(Y)$ can be obtained by

$$H(Y) = \int_{-\infty}^{\infty} f(y; k) \ln f(y; k) \, dy$$

$$= \frac{k}{2} + \ln \left(2\Gamma\left(\frac{k}{2}\right)\right) + \left(1 - \frac{k}{2}\right)\psi\left(\frac{k}{2}\right)$$  \hspace{1cm} (5)

where $\psi(x)$ denotes Digamma function, which is defined as the logarithmic derivative of the gamma function, and $\psi(1/2) = -2 \ln 2 - \gamma$ and $\gamma$ denotes the Euler-Mascheroni constant.

According to (1), it is easy to obtain that $I(X, Y)$ is 0.7838 and $\lambda(X, Y)$ is 0.8897. The two variables $X$ and $Y$ are obviously correlated with each other in this case; however, it is easy to find out that linear correlation coefficient is zero.

2.3. Numerical Solution to the Mutual Information. In general, it is difficult to obtain the theoretical value of mutual information directly because probability density functions are always unknown. Therefore, several frequently used numerical methods have emerged to estimate mutual information [12, 23–26]:

(a) Histogram-based estimators.

(b) Kernel density estimator (KDE).

(c) K-nearest neighbor (KNN) samples.

(d) Adaptive partitioning of the $XY$ plane.

The most straightforward and widespread method for estimating mutual information is to approximate the probability densities using histograms [12]. It is often employed due to its computational benefit. This estimation method based on histograms consists in partitioning $X$ and $Y$ into bins of finite size and approximating $I(X, Y)$ by the finite sum as the following [23]:

$$I(X, Y) = I_{\text{binned}}(X, Y) = \sum_{ij} p(i, j) \log \frac{p(i, j)}{p_X(i) p_Y(j)}.$$  \hspace{1cm} (6)

An estimator of $I_{\text{binned}}(X, Y)$ is obtained by counting the number of points falling into the bins. Let $n_x(i)$ denote the number of points falling into the $i$th bin of $X$, $n_y(j)$ the number of points falling into the $j$th bin of $Y$, and $n(i, j)$ is the number of points in their intersection; then, we approximate $p_x(i) = n_x(i)/N$, $p_y(j) = n_y(j)/N$, and $p(i, j) = n(i, j)/N$. It is easily seen that (6) converges to $I(X, Y)$ if $N \to \infty$ and all bin sizes tend to zero. The bin sizes used in (6) do not need to be the same for all bins.

In practice, the operational loss data are hard to collect and so the sample size $N$ is always very small. In order to get a more accurate estimate, we resort to the bootstrap technique. Bootstrap can overcome the bias for small set of data [31]. Let $Z = \{z_1, z_2, \ldots, z_N\} = \{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\}$. At each iteration of bootstrap, $N$ random numbers are drawn to form a new sample $Z^* = \{z_1^*, z_2^*, \ldots, z_N^*\}$ and the mutual information is computed based on the new sample. The procedure of basic bootstrap method for estimating mutual information is summarized as follows [24]:

(1) Draw a random sample $Z^*$ from $Z$.

(2) Calculate the estimated mutual information $\overline{I}_n$ based on the sample $Z^*$.

(3) Repeat step (1) and step (2) $B$ times; that is, $n = 1, 2, \ldots, B$.
(4) Sort the bootstrap estimates $T_i$ in ascending order to obtain $T_{k_1}, T_{k_2}, \ldots, T_{k_8}$.

(5) The $(1 - \alpha) \times 100\%$ confidence interval is $(T_{k_p}, T_{k_q})$, where $p = \lfloor \alpha / 2 \rfloor$ and $q = B - p + 1$.

(6) The final estimated mutual information $\hat{I}$ is the average value of the mutual information in the interval $(T_{k_p}, T_{k_q})$.

2.4. Independence Test. Because the definition of mutual information is relatively complicated, it is difficult to judge whether the dependence exists or not. Therefore, Dionisio et al. proposed independence test for mutual information [30]. According to the properties of mutual information, we can construct an independence test based on the hypotheses as

\[ H_0: p_{X,Y}(x, y) = p_X(x)p_Y(y); \]
\[ H_1: p_{X,Y}(x, y) \neq p_X(x)p_Y(y). \]  

If $H_0$ holds, then $I(X, Y) = 0$ and we conclude that the variables $X$ and $Y$ are independent. If $H_1$ holds, then $I(X, Y) > 0$ and we reject the null hypothesis of independence. The above hypotheses can be reformulated as

\[ H_0: I(X, Y) = 0; \]
\[ H_1: I(X, Y) > 0. \]  

In order to implement the independence test between the variables, we need to compute the critical value of the mutual information. In this paper, the critical values are calculated and sorted in ascending order. Finally, the generalized correlation coefficient of these samples are simulated from the white noise series. The mutual information is relatively complicated, it is difficult to judge whether the dependence exists or not. Therefore, Dionisio et al. proposed independence test for mutual information [30].

2.5. Choice of Copula. Copula function is a useful tool for constructing and simulating multivariate distributions [4, 9]. There are many types of copulas, such as normal copula, t copula, Gumbel copula, Clayton copula, and Frank copula. For a detailed introduction of copulas, please refer to Li et al. [10]. In this study, for simplicity and ease of use, we use the most common normal copula to estimate the aggregated loss distribution for operational risk.

The multivariate normal copula $C^N_R(u_1, \ldots, u_n)$ is defined as the following [10]:

\[ C^N_R(u_1, \ldots, u_n) = \Phi_R(\phi^{-1}(u_1), \ldots, \phi^{-1}(u_n)), \]  

where $\Phi_R$ denotes the standardized multivariate normal distribution with correlation matrix $R$ and $\phi^{-1}$ denotes the inverse function of standard univariate normal distribution. As shown in (9), the correlation matrix is the only parameter of normal copula. In this study, the generalized correlation coefficient instead of linear correlation matrix is used as the parameter of normal copula.

2.6. Capital Computation. LDA is an actuarial technique that separately estimates the frequency distribution and severity distribution of operational risk loss and then combines them by convolution to derive annual operational risk distribution [13, 32]. For the procedure of a standard LDA, please refer to Frachot et al. [2] and Li et al. [33]. Because the convolution is always difficult to implement, Monte Carlo simulation is often used instead in practice [33–35]. The normal copula with generalized correlation coefficient models the dependence structure between loss frequencies of different business lines. Besides, perfect dependence and linear dependence are also employed for comparison. The empirical aggregated losses and VaR are derived by the following steps, respectively.

2.6.1. Generalized Dependence

Step 1. Generate a multivariate random vector $(v_1, v_2, \ldots, v_k)$ from normal copula $C$ with the correlation matrix of generalized correlation coefficient.

Step 2. Transform the vector $(v_1, v_2, \ldots, v_k)$ into a sample of frequencies $(n_1, n_2, \ldots, n_k)$ for each business line by using corresponding marginal frequencies distribution.

Step 3. Simulate $n_1$ losses, $n_2$ losses, \ldots, and $n_k$ losses for each business line from corresponding loss severity distribution.

Step 4. Obtain aggregate loss by summing all the losses in Step 3.

Step 5. Repeat Step 1 to Step 4 S times and obtain S aggregate losses.

Step 6. VaR is calculated as the corresponding percentile of aggregate losses in ascending order.

2.6.2. Linear Dependence. Replace the generalized correlation coefficient with the linear coefficient in Step 1 of generalized dependence and repeat Step 2 to Step 6.

2.6.3. Perfect Dependence

Step 1. Generate a random number $n$ from the frequency probability distribution.

Step 2. Generate $n$ severity values from corresponding loss severity distribution.

Step 3. Calculate the aggregation loss by summing the $n$ severity values.

Step 4. Repeat Step 1 to Step 3 S times to obtain S aggregate losses.

Step 5. Individual VaR for each business line is calculated as the corresponding percentile of aggregate losses in ascending order.

Step 6. Repeat Step 1 to Step 5 for each business line and sum the individual VaRs of different business lines to obtain the final VaR of operational risk.
3. An Application to Chinese Banking

In this section, the model presented in Section 2 is applied to the operational risk aggregation of Chinese banks. The data set in this application consists of a total of 860 operational risk loss events of Chinese banks, spanning from 1994 to 2006. BCBS divides banks’ activities into eight business lines [5]. This data set shows that most loss events occur in trading and sales (TS), retailing banking (RB), and commercial banking (CB), so other business lines are not considered in this experiment because their data do not support reliable parameter estimation.

3.1. Distribution Fitting Result. Operational risk is characterized by “leptokurtosis and fat tail” [22]. Many distributions such as lognormal distribution, exponential distribution, Pareto distribution, gamma distribution, Weibull distribution, generalized hyperbolic distribution, and generalized error distribution (GHD) have been used to fit loss severity [1, 3, 13]. Besides, Poisson distribution, negative binomial distribution, and geometric distribution are always used to fit loss frequency [36, 37]. Among these types of distributions, Basel Committee on Banking Supervision (BCBS) points out that it is common for banks to use Poisson distribution for estimating frequency and use lognormal distribution for modeling severity [22]. Therefore, in line with BCBS and many other studies [2, 33–36], we also assume that loss severity follows lognormal distribution and the loss frequency follows Poisson distribution in this study.

The parameters of frequency and severity distributions are estimated by maximum likelihood method. Besides, Kolmogorov-Smirnov test (KS test) is also used to examine whether these distributions fit frequencies and severities well or not. For KS test, the larger the p value is, the better the distribution fits the data and threshold is usually set as 0.05. Tables 1 and 2 show the results of parameter estimation and KS test. All the p values of KS test are larger than 0.05, which means that using these distributions to fit the severities and frequencies of the three business lines is proper.

3.2. Mutual Information Estimation Result. Mutual information between loss frequencies of different business lines is estimated by using the histogram method and bootstrap technique described in Section 2.3. In this experiment, $\alpha = 0.05$ and $B = 5000$ in the simulations of bootstrap. The results of mutual information are shown in Table 3. Based on the mutual information, the generalized correlation coefficient between business lines can be calculated, which is shown in Table 4. Linear correlation coefficients are also presented for comparison.

As shown in Table 4, it is easy to find that all generalized correlation coefficients are larger than the corresponding linear correlation coefficients, which implies the existence of nonlinearities between business lines. Linear relationship between different variables is based on linear regression techniques and estimated by ordinary least squares. Therefore, possible nonlinear component is omitted. Besides, the estimated coefficient may suffer severe biases since the residuals hardly behave as white noises. In contrast to linear correlation coefficient, mutual information can capture, in a quite global way, the generalized relationship between variables. Furthermore, it does not have any requirement for a theoretical probability distribution or a specific model of dependency.

3.3. Independence Test Result. In order to test whether the generalized correlation coefficient in Section 3.2 is significant or not, the independence test described in Section 2.4 is implemented. Firstly, 10000 samples are simulated from the standard bivariate normal distribution with zero correlation coefficient. Then, mutual information of each sample is calculated. Finally, the 90%, 95%, and 99% percentiles of the mutual information values are the critical values at corresponding confidence levels. As shown in Table 5, at 90% confidence level, the critical value for mutual information is 0.15 and for generalized correlation coefficient is 0.51. The mutual information values in Table 3 are larger than 0.15 and generalized correlation coefficients in Table 4 are larger than 0.51, which means that they are statistically significant at 90% confidence level. This further demonstrates the existence of nonlinear dependence between business lines.

### Table 1: The parameters of loss severity distribution for different business lines.

<table>
<thead>
<tr>
<th>Business line</th>
<th>Distribution</th>
<th>Parameters</th>
<th>KS test (p value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TS</td>
<td>Lognormal</td>
<td>5.15 2.69</td>
<td>0.18</td>
</tr>
<tr>
<td>RB</td>
<td>Lognormal</td>
<td>7.83 2.31</td>
<td>0.43</td>
</tr>
<tr>
<td>CB</td>
<td>Lognormal</td>
<td>6.54 2.51</td>
<td>0.85</td>
</tr>
</tbody>
</table>

### Table 2: The parameters of loss frequency distribution for different business lines.

<table>
<thead>
<tr>
<th>Business line</th>
<th>Distribution</th>
<th>Parameter</th>
<th>KS test (p value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TS</td>
<td>Poisson</td>
<td>36</td>
<td>0.07</td>
</tr>
<tr>
<td>RB</td>
<td>Poisson</td>
<td>17</td>
<td>0.19</td>
</tr>
<tr>
<td>CB</td>
<td>Poisson</td>
<td>10</td>
<td>0.14</td>
</tr>
</tbody>
</table>

### Table 3: Mutual information between different business lines.

<table>
<thead>
<tr>
<th></th>
<th>TS</th>
<th>RB</th>
<th>CB</th>
</tr>
</thead>
<tbody>
<tr>
<td>TS</td>
<td>2.95</td>
<td>0.22</td>
<td>0.20</td>
</tr>
<tr>
<td>RB</td>
<td>0.22</td>
<td>2.95</td>
<td>0.17</td>
</tr>
<tr>
<td>CB</td>
<td>0.20</td>
<td>0.17</td>
<td>2.95</td>
</tr>
</tbody>
</table>

### Table 4: Linear and generalized correlation coefficients between different business lines.

<table>
<thead>
<tr>
<th>Linear correlation coefficient</th>
<th>Generalized correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>TS</td>
<td>1</td>
</tr>
<tr>
<td>RB</td>
<td>0.44</td>
</tr>
<tr>
<td>CB</td>
<td>0.55</td>
</tr>
</tbody>
</table>
3.4. Aggregation Result Analysis. After obtaining the generalized correlation coefficients between business lines, the capital requirement for operational risk at different confidence levels can be calculated by using Monte Carlo simulation described in Section 2.6. In order to balance the accuracy and time cost, the number of simulations is set as 100,000. For the purpose of an overall comparison, the results of linear dependence and perfect dependence are also presented in Table 6.

At all confidence levels, the VaRs with perfect dependence assumption are the largest and the VaRs with linear dependence assumption are the smallest. There is mounting evidence of nonlinear dependence between financial risks [10]; however, linear dependence ignores possible nonlinearity between business lines. Therefore, it is considered to underestimate the operational risk. On the contrary, the perfect dependence assumes that the loss frequencies of different business lines are perfectly correlated with each other and simply adds VaRs of different business lines up. This assumption is generally considered to overestimate the VaR because it ignores possible diversification benefits.

Generalized dependence gives VaR between linear dependence and perfect dependence. It is more reasonable and realistic because it can capture both linear and nonlinear dependence. As the confidence level increases, VaR becomes larger. Conventionally, the economic capital requirement is set to protect against losses over 1 year at 99.9% level. BCBS also recommends 99.9% as a proper confidence level [5]. Because the generalized correlation assumption is more reasonable than the other two assumptions, we conclude that the operational risk for Chinese banking is 248 billion CNY.

4. Conclusion

In this paper, mutual information is used to model the frequency dependence in operational risk. Firstly, the generalized correlation coefficient is calculated based on mutual information. Then, the normal copula with generalized correlation coefficient is used to model the dependence between the loss frequencies of different business lines. Finally, operational risk is calculated in the framework of loss distribution approach. In the empirical analysis, the proposed method is applied to Chinese banking based on a data set consisting of 860 operational risk loss events. The results show that the generalized correlation coefficient is more rational than the linear correlation coefficient in modeling dependence.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

This research is partially supported by the National Natural Science Foundation of China under Grants nos. 71173018 and 71433001 and by the Fundamental Research Funds for the Central Universities under Grant no. SKZZY2013014.

References
