Operational Risk Measurement in Commercial Banks from the Joint Perspective of Multivariate Modeling and Theory

[1st] Lijuan Gai[1], [2nd] Jun Tang[2*]

{glj571113173@163.com[1], tj2001766@sina.com[2*]}

Inner Mongolia university of science and technology, Baotou 014000, China[*]

Abstract. For commercial banks, operational risk has become as important as market and credit risks. By dividing the operational risk units, establishing the EVT loss intensity POT model, determining the optimal thresholds, and constructing the correlation structure of the operational risk units through the EVT-Pair-Copula model, the total VaR of the operational risk units is calculated to be about 3.36% less than that of the traditional EVT-based POT model where the VaR of each operational risk unit is summed up. Commercial banks can use this model to accurately measure operational risk and formulate a compliant risk management system for risk control.

Keywords: operational risk, extreme value theory, EVT-Pair-Copula, VaR, Monte Carlo

1 Introduction

The Twentieth CPC National Congress report requires that risk management of financial institutions be strengthened and that strict precautions be taken to ensure that the bottom line of systemic risk is not touched. Commercial banks, as the cornerstone of the financial system, assume essential risk transmission and bearing functions. In the operation process, operational risk has gradually come to the fore, originating from business complexity, information asymmetry, and other factors. It poses a potential challenge to financial stability.

Regarding the research on measuring commercial banks' operational risk, the industry mainly adopts the advanced econometric method. Foreign research on commercial bank operational risk metrics is relatively early. Perry and Dutta improved EVT based on Mc Neil[1], constructed different parametric models, and finally concluded that different model settings would cause inaccurate estimation values[2]. While domestic-related research needs to catch up, Lina Dai used the Monte Carlo simulation method. It estimated the loss distribution using the nonparametric and kernel density methods to measure commercial banks' operational risk loss, respectively[3]. Junming Xie and Binghui Hu integrated the loss distribution method, POT model, and confidence model to analyze and measure the operational risk of four central state-owned banks[4].

Copula's theory has dramatically progressed in applying risk management in recent years. Most scholars use Copula functions to construct the correlation structure of different operational risk units. Embrechts and Puccetti use Copula functions to construct the dependence structure and calculate the upper definitive bound of VaR[5]. Shi Yongfen et al.

MSEA 2024, May 24-26, Jinan, People's Republic of China

DOI 10.4108/eai.24-5-2024.2350131

compared the differences in the effectiveness of different Copula functions in fitting the dependence structure of operational risk units[6]. In summary, Copula is a new method that can reflect nonlinearity, asymmetry, and tail correlation between variables. However, most of the studies assume that the various types of operational risk are entirely dependent on each other, and to calculate the operational risk, the VaR values of the various types of operational risk are added up, which is inconsistent with the actual situation.

For a long time, the operational risk data of some banks have been missing, and the accuracy of measuring operational risk is rarely compared between a single model and a multivariate model based on the same perspective, which results in the current situation that the operational risk of the banking industry is difficult to control effectively. Given this, the main contribution of this paper is to fully utilize the large-sample data in Codd's operational risk database, compare the measurement values (VaR) of commercial banks' operational risk by a single EVT (POT model under the extreme value theory) and a multivariate EVT-pair-Copula model from the perspective of the combination of multivariate models and theories, and effectively overcome the operational risk measurement accuracy due to the irrational criteria of model use. Risk measurement inaccuracy and overestimation or reduction of risk control capitalization capacity, thus providing a basis for establishing and implementing operational risk prevention and control strategies in the Chinese banking industry.

2 Model Construction

2.1 POT modelling to construct marginal distributions

Extreme value theory is a measure of the extent of risky losses in extreme market conditions, assuming a series of asset return series $Z_1, Z_2, \ldots, Z_n$, whose distribution parameters can be described by $F_\mu(y)$:

$$F_\mu(y)=P(Z-\mu\leq y|Z>\mu)=\frac{F(\mu+y)-F(\mu)}{1-F(\mu)}=\frac{F(Z)-F(\mu)}{1-F(\mu)}, y\geq 0 \tag{1}$$

According to the PBH theorem[7], if the threshold μ is large enough, the excess distribution function is approximated as a GPD distribution, i.e., a generalized Pareto distribution:

$$G_{\varepsilon,\beta}(y)=\begin{cases} 1-(1+\frac{\varepsilon y}{\beta})^{-\frac{1}{\varepsilon}}, \varepsilon \neq 0 \\ 1-e^{-\frac{y}{\beta}}, \varepsilon = 0 \end{cases} \tag{2}$$

The standard empirical cumulative method fits the middle portion of the standardized residuals Z_t. In contrast, the standard GPD method is used to model the middle portion of the sequence where the upper and lower bounds are confirmed. Finally, the marginal distribution of the standard residual test sequence can be derived as:

$$F(Z)=\begin{cases} \frac{N_{\mu L}}{N}(1-\varepsilon\frac{Z-\mu_L}{\beta(\mu)})^{-\frac{1}{\varepsilon}} & Z<\mu_L \\ \phi(Z) & \mu_L\leq Z\leq\mu_\mu \\ 1-\frac{N_{\mu\mu}}{N}(1+\varepsilon\frac{Z-\mu_\mu}{\beta(\mu)})^{-\frac{1}{\varepsilon}} & Z>\mu_\mu \end{cases} \tag{3}$$

Where the lower and upper thresholds of the standardized residual series are μ_L and μ_μ, respectively, $\phi(Z)$ is the empirical distribution function of the centre portion of the data; N is the overall sampling size, which N_μ is the data observed when the threshold is exceeded; β is the scale parameter and ε is the shape parameter.

2.2 Threshold determination

2.2.1 MEF chart method

$$e(\mu)=\frac{\beta}{1-\varepsilon}+\frac{\varepsilon}{1-\varepsilon}\mu \tag{4}$$

where $0<\varepsilon<1,\sigma+\varepsilon\mu>0$, and e(μ) approximates a linear function when $x\geq\mu$. As a result, the image will tend to be linearized at a particular observation, which is the threshold μ.

2.2.2 Hill diagram method

$$H_{k,n}=\frac{1}{k}\sum_{i=1}^{k} ln\frac{X_i}{X_k}=\frac{1}{\varepsilon_k} \tag{5}$$

The principle of choice is to use as a threshold μ the operational risk loss value X_k corresponding to the horizontal coordinate k of the start of the stabilization region in the Hill.

2.2.3 Hkkp validation method

$$\gamma(k)=\beta_0+\beta_1 k+\varepsilon,(k=1,\ldots,n),\varepsilon\sim(0,1) \tag{6}$$

Where k is the number of loss values exceeding the threshold μ, $\gamma(k)$ is an estimate of the excess number for k, and β_0 is the actual value of the shape parameter[8], which is obtained by regression using the OLS method and comparing the μ that minimizes the variance of the shape parameter is done as the threshold value.

2.3 Risk metrics

In a general market volatility environment, the maximum loss an asset or a portfolio can sustain within a specific confidence interval is the value-at-risk-VaR. Where the confidence interval is p, and the VaR is

$$VaR = \mu + \frac{\beta}{\varepsilon}((\frac{n}{N_\mu}(1-p))^{-\varepsilon}-1) \tag{7}$$

2.4 Characterizing the structure of interdependence: the EVT-pair-Copula model

$$F(z_1,z_2)=C(F_1(z_1),F_2(z_2)) \tag{8}$$

The joint density function can be further obtained through the Copula function $C:[0,1]^2\rightarrow[0,1]$ of Eq. (8):

$$f(z_1,z_2)=c(F_1(z_1),F_2(z_2))f1(z_1)f2(z_2) \tag{9}$$

where the density function of the Copula function is $c(u1,u2)=\frac{(u_1,u_2)}{\partial u_1\partial u_2}$, and $F_n(z_n)$ is obtained according to Eq. (3). $f_n(z_n)$ is the density function of the marginal distribution function $F_n(z_n)$.

In summary, the parameters of the EVT-pair-Copula model are estimated as follows: first, using the EVT model to extract the standardized residuals to the operational units; second, transforming the data of the standardized residuals into a uniform distribution obeying (0,1); and third: calculating the Pearson correlation coefficients between any two operational risk units on this basis, and determining, based on the two-by-two dependent frequency histograms, the corresponding Copula type.

3 Empirical analysis

3.1 Data description

In this paper, we refer to the research samples of the Cord database. Xie Peirong[9], as well as the operational risk loss events of China's commercial banks collected from the public media as samples, and the operational risk events collected from 2012 to 2021, a total of 275 cases, which are summarized and organized into Internal fraud (Internal), External fraud (External), Irregular execution and delivery (Irregular), Collusion between inside and outside (Collusion), Financial corruption (Financial), and Information technology vulnerabilities (Technology). Given the limited data available, analyzing each commercial bank individually was impossible, so the loss data were considered a bank as a whole.

Table 1. Statistics on the six operational risk units (in billions of yuan)

Risk Module	skewness	peak	JB	average	max	min	standard	frequency
External	10.73	121.24	92072.7	3.40	312.69	0.00005	26.64	153
Internal	6.81	47.77	1003.6	5.45	190.00	0.00300	26.36	52
Irregular	5.24	28.61	1085.1	385.52	755.00	0.00800	131.98	32
Collusion	4.21	18.40	218.2	8.51	120.00	0.00050	25.64	21
Financial	2.48	5.16	13.4	2.75	20.00	0.00337	5.88	15
Technology	2.87	8.50	10.5	0.03	0.20	0.00003	0.06	12

From the statistical characterization of losses for the six operational risk units in Table 1, it was found that there are significant loss events in each unit of operational risk, with the most significant loss amounting to ￥75.5 billion for execution and delivery violations. However, the number of high losses occurs less frequently, which aligns with the characteristics of high losses and low frequency. In addition, compared with the normal distribution, all six operational risk units are not symmetrically distributed, and there is a severe right skew phenomenon. Eqs. (4), (5), and (6) are used to determine the optimal thresholds for the six operational units, as shown in Table 2. Then, the value with the smallest $\chi 2$ is screened by the $\chi 2$ test as the final threshold, and the $\chi 2$ test given in the table is the test result corresponding to the finalized threshold.

Table 2. Table of threshold selection results

	Internal	External	Irregular	Collusion	Financial	Technology
MEF	0.34	7.22	10	2.15	0.642	0.0058
Hill	0.006	0.0055	1.98	0.90	0.642	0.0064
Hkkp	0.49	0.3	10	0.90	0.642	0.0058
μ	0.49	0.3	10	2.15	0.642	0.0058
$\chi 2$	0.1708	0.0160	0.0273	0.5245	0.6227	0.2726

Through Table 2, it can be demonstrated that the $\chi 2$ of the thresholds finally estimated by the Hkkp validation method is overall very small compared to the more subjective image observation methods such as MEF plots and Hill plots.

3.2 Parameter estimation of the POT model

Maximum likelihood estimation of the morphology parameter ε and the scale parameter of the POT model for the six operational units is performed by substituting them into Eq. (7). As can be seen from Table 3, the VaR values of the six operational risk units are ￥119.65 billion, ￥71.77 billion, ￥305.95 billion, ￥846.86 billion, ￥127.93 billion, and ￥270 million, respectively, at 99% confidence level, which sums up to ￥1344.5 billion for the operational risk units other than Financial.

Table 3. Parameter estimation results of the POT model

	Internal	External	Irregular	Collusion	Financial	Technology
μ	0.49	0.3	10	2.15	0.642	0.0058
ε	0.9668	1.0064	1.5004	1.1716	1.0714	0.7662
β	4.6988	1.8921	4.6542	0.0148	0.9193	0.0177
VaR	1196.5	717.7	3059.5	8468.6	12793	2.7

Using the estimated morphology parameter ε and scale parameter, we do the CDF and empirical CDF plots to check the fitting effect of the data fitting of the six operational risk units, respectively. We find that the data fitting of the operational risk units of Internal, External, Irregular, and Collusion is better. In contrast, the fitting effect of the remaining two operational units is slightly worse. The most miniature ideal is Financial, probably because the country maintains a zero-tolerance attitude toward finances, resulting in an alarming decrease in its occurrence, followed by Technology, along with the development of new technologies such as the Internet, there are numerous means of telecommunication fraud, and hackers and foreign spies have caused severe monetary losses to commercial banks in an unnoticeable way.

3.3 EVT-Pair-Copula Structural Modelling

It was found that the number of Financial occurrences decreased significantly from year to year, so the study was conducted only for the tail data of the five operating units for which suprathreshold data had been identified, and based on this, three functions, namely, normal Copula, t Copula, Clayton Copula in Archimedean Copula, were adopted to measure them.

Bootstrap repeated sampling was used to obtain the loss gaps for the five operational risk units, and Monte Carlo simulation was used to obtain the loss sample data for each operational risk unit, which was transformed into new data obeying U(0,1) and to estimate the spearman rank correlation coefficients between the losses of these five operational risks. Among them, External has the highest correlation with Irregular. So, in this paper, Irregular is used as the node of Collusion, and Internal and External are used as the nodes of Technology and Collusion. Finally, we get the frequency histogram of the two variables.

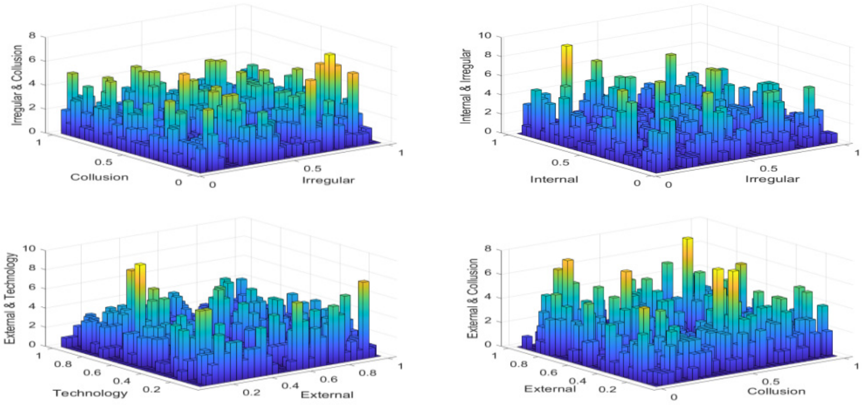

Fig. 1. Frequency histogram of two two variables

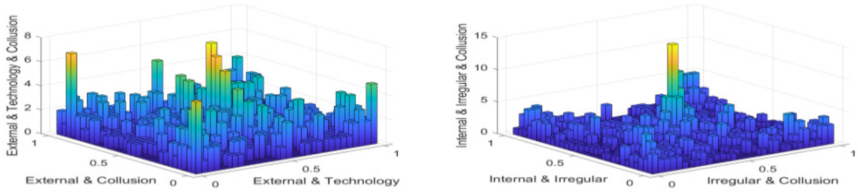

Fig. 2. Frequency histograms for (External & Technology & Collusion) and (Internal & Irregular & Collusion)

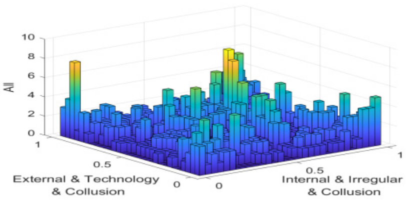

Fig. 3. Frequency histogram of the five central operational risk units

From Fig. 1, it is initially observed that Irregular and Collusion are fitted with normal Copula sensitive to tail data, Irregular and Internal are also fitted with normal Copula sensitive to tail data, External and Technology are fitted with normal Copula, External and Collusion were fitted with normal Copula and Matlab software was used to estimate the parameters of the respective Copula functions and generate the corresponding fitted values. Next, the two normal Copula are obtained. Then new Copula functions are constructed to obtain the right graph of Fig. 2, which is fitted appropriately with t Copula, i.e., t Copula for External, Technology, and Collusion; new Copula functions are constructed with the other two normal Copula to obtain the left graph of Fig. 2, which is also fitted appropriately with t Copula fitted appropriately, i.e., the t Copula of Irregular, Collusion, and Internal, the respective function parameters are estimated accordingly, and the fitted values are generated. Finally, the final

Copula function with the two constructed Copulas to obtain Figure 3 was more sensitive to the tail data, so Clayton Copula was fitted.

Parameter estimation is done using the Clayton Copula function, and finally, 2000 samples for each type of operational risk are simulated by the Monte Carlo method. With 100,000 simulations, the confidence level of 99.9% VaR is 1296.098 billion yuan. Compared with the total loss of 1344.5 billion yuan that is summed up by the first five operational risk units in the POT model using only extreme value theory, the total loss is reduced by 3.36%, so the application of EVT-Pair-Copula model to measure the correlation between the operational risk units and to construct the correlation structure estimation is more accurate. It can reduce the total loss of operational risk and save the cost of regulatory capital for operational risk for China's commercial banks.

4 Conclusion

In this paper, a total of 275 operational risk loss events of commercial banks from 2012 to 2021 are selected as samples, and a single EVT model and a multivariate EVT-pair-Copula model are chosen to measure and cross-sectionally compare the operational risk of commercial banks under different models. The results show that: first, the VaR calculated under the POT model based on EVT theory is 1344.5 billion yuan; second, the VaR calculated under the EVT-pair-Copula model using Monte Carlo simulation is 1296.098 billion yuan; third, the total VaR of the operational risk unit calculated by the EVT-Pair-Copula model is higher than that of the traditional EVT-based POT model the VaR of each operational risk unit summed up by about 3.36%. Applying the EVT-Pair-Copula model to construct the correlation structure of the operational risk unit can significantly reduce the operational risk capital requirement, achieve the risk diversification effect of the asset portfolio, and enhance the bank's profitability. Fourth, with the continuous development of information digitization, the loss events of IT vulnerability are increasing in frequency and amount, so it is crucial to increase the new type of risk prevention and management.

References

[1] Mc Neil A J.(1999) Extreme Value Theory for Risk Managers [M].Internal Modeling and CAD II, Published by RISK Books,1999:1-23.https://www.researchgate.net/publication/2470539.

[2] Dutta K, Perry J.(2006)A Tale of Tails: An Empirical Analysis of Loss Distribution Models for Estimating Operational Risk Capital. Federal Reserve Bank of Boston.10.2139/ssrn.918880

[3] Dai, L N.(2020) Measurement of Operational Risk in Commercial Banks - Based on Internal and External Loss Data and Expert Opinion. Statistical Theory and Practice, 59-66. DOI:10.13999/j.cnki.tjllysj.2020.02.009.

[4] Xie J M, Hu B H.(2023)Research on Operational Risk Measurement of Commercial Banks under the Perspective of Data Integration.Creditworthiness,72-77+86.

[5] Embrechts P, Puccetti G.(2006)Bounds for functions of dependent risks.Finance and Stochastics,341-352.

[6] Wu F, Liu M M.(2023)A Study of Risk Spillover Effects in Financial Markets Based on Dynamic Vine Copula Modeling. Operations Research and Management,179-185. DOI:CNKI:SUN:JIRO.0.2018-01-008.

[7] KÜhn R, Neu P.(2003)The Functional correlation approach to operational risk in banking organizations.Physica A: Statistical Applications,650-666. DOI:10.1016/S0378-4371(02)01822-8.
[8] Pickands J.(1975)Statistical inference using extreme order statistics.Annals of Statistics,119-131.
[9] Xie P R.(2021) A Study on Operational Risk Measurement in Chinese Commercial Banks: Based on Optimized Extreme Value POT Model.Hangzhou: Zhejiang Institute of Science and Technology.DOI:CNKI: CDMD:2.1014.366375.