A new Copula-CoVaR approach incorporating the PSO-SVM for identifying systemically important financial institutions

Tingting Zhang & Zhenpeng Tang

Submit your article to this journal ⬀

Article views: 135

View related articles ⬀

View Crossmark data ⬀

Routledge
Taylor & Francis Group

A new Copula-CoVaR approach incorporating the PSO-SVM for identifying systemically important financial institutions

Tingting Zhang and Zhenpeng Tang

School of Economics and Management, Fujian Agriculture and Forestry University, Fuzhou, China

ABSTRACT

The effective identification of systemically important financial institutions (SIFIs) is key to preventing and resolving systemic financial risks; thus, it is of great research significance for emerging countries to supervise SIFIs and manage systemic financial risks. Since traditional research on identifying SIFIs does not consider emerging machine learning models, it is difficult to properly fit the characteristics of actual financial institutions' asset distribution. This paper proposes a new method for measuring SIFIs, integrating the PSO-SVM model into the Copula-CoVaR model. This new PSO-SVM-Copula-CoVaR model is meant to evaluate China's SIFIs based on the publicly traded price data of Chinese listed financial institutions. The empirical results show that, compared with the traditional parameter method (GARCH model) and the nonparametric method (kernel density estimation), the marginal distribution estimation method using the PSO-SVM method can better fit the distribution of an institution's financial asset return sequence. That is, the model proposed in this paper helps regulatory authorities improve the list of SIFIs more reasonably and implement effective regulatory measures.

1. Introduction

The global financial crisis of 2008 demonstrated that, when SIFIs get into trouble or go bankrupt, the whole financial system suffers great damage due to the spillover effect, which may eventually lead to a serious economic crisis (Fan et al., 2018; Johnson & Mamun, 2012; Zhang et al., 2021). Thus the effective identification of SIFIs is key for regulators hoping to deal with the systemic risks of the financial system (Chen et al., 2021).

As of now, this paper is one of the few studies that attempting to introduce machine learning models into the identification of SIFIs. The main identification methods of SIFIs relevant in this task are the indicator-based approach and the

---

CONTACT Zhenpeng Tang ✉ zhenptaf@126.com

market-based approach (Li et al., 2021; Xu, 2011). The indicator-based approach involves constructing a corresponding index system to evaluate the importance of financial institutions based on the core concepts of SIFIs.

The market-based approach measures a single financial institution's risk contribution to the whole system based on financial market data, determining its systemic importance; compared with the indicator-based approach's index method, it can effectively overcome the problem of data lag and improve the accuracy of measuring institutional spillover risk. Among this series of market model methods, the CoVaR method, which considers the spillover effects of financial institutions, has received widespread attention (Bernal et al., 2014).

However, CoVaR is essentially a linear correlation function between returns, and there has yet to be discussion of the specific dependency structure, which has some drawbacks when it comes to describing the tail risk dependency (Benoit et al., 2013; Kleinow et al., 2014; López-Espinosa et al., 2012). Considering the complex structure of interdependence between financial institutions, scholars have introduced different copula functions to capture the linear and non-linear tail-related characteristics of such institutions and the entire financial system (Bernardi et al., 2017; Jaeger-Ambrozewicz, 2013; Liu, 2015)

The main difference between the indicator-based approach and the market-based approach lies in their perspective for understanding the meaning of systemically important financial institutions and their related data. Although it is intuitive and simple to measure SIFIs with an indicator-based approach (Lu & Hu, 2014), such an approach needs more comprehensive data (Brämer & Gischer, 2013); additionally, the frequency of statistical data involved is usually low, so it is difficult to capture the changing characteristics of high-frequency data. As a result, the dynamic spillover effect of systemic risk is not considered. Compared with the indicator-based approach, the market-based approach is more forward-looking because it can adopt higher frequency data.

In summary, the market-based approach based on publicly traded data is more suitable for the identification of systematic financial institutions, and the CoVaR model combined with the copula function has the advantage of portraying the dependence of tail risk (Sun et al., 2020; Wu et al., 2021).

Despite this approach's suitability, the existing research still has several shortcomings. Firstly, with the development and complexity of the financial market, the traditional marginal distribution model based on parametric and nonparametric methods has difficulty fitting the distribution characteristics of actual financial assets. Secondly, while the rapidly internationalized and liberalized Chinese stock market is becoming increasingly important for global financial markets (Yao et al., 2018; Zhang, 2017), few studies have explored China's financial markets. Finally, with the asset structure of China's financial market changing, it is difficult to properly conduct financial stability supervision by focusing only on the banking market.

Compared with traditional statistical or simulation methods, machine learning can optimize loss or return functions by using historical data, which can better quantify the complex and changeable characteristics of financial risks. Among the available frameworks, support vector machine (SVM) is a theoretical framework and general

method for establishing machine learning with limited samples (Dutta, 2022; Luo et al., 2020; Vapnik, 1995, 1998). Not only does SVM have a strict theoretical basis, it can better solve practical problems (e.g., small sample, nonlinearity, high dimension and local minimum point) (Kim, 2003; Lee, 2007). This paper introduces SVM to model the edge distribution.

Therefore, this paper endeavors to make up for the deficiencies in the existing research and to make a more accurate and reasonable evaluation of China's SIFIs, based on the publicly traded price data of 27 listed financial institutions in China's capital market from January 2011 to March 2021. To do so, this paper constructs a new PSO-SVM-Copula-CoVaR model for research. Firstly, the PSO-SVM model, which is optimized by the particle swarm optimization algorithm (PSO) (Eberhart & Kennedy, 1995), is used to model the edge distribution of stock price data, and the estimation results are compared with the traditional edge distribution modeling estimation methods (such as the GARCH model and kernel density estimation). Secondly, the Copula-CoVaR function is used to describe the multivariate distribution and tail-dependent risks of financial variables completing the risk assessment of systemically important financial institutions. The results show that the risk estimation model incorporating the PSO-SVM algorithm can effectively improve the accuracy of SIFI recognition.

This paper contributes by filling the gaps in the existing literature on the identification of systemically important financial institutions, introducing machine learning algorithm into the traditional Copula-CoVaR model, and exploring the application of machine learning algorithms in the field of financial risk management. In addition, most existing studies are focused on the banking sector, and few have explored the identification of systemically important financial institutions in China's financial market. The research in this paper provides a unique reference for financial market risk research in emerging countries.

The remaining parts of this paper are arranged as follows: Section 2 is a review of the current literature. Section 3 briefly introduces the sub-models that constitute the PSO-SVM-Copula-CoVaR model and the specific steps for constructing the PSO-SVM-Copula-CoVaR model. Section 4 conducts an empirical analysis based on the actual data of China's financial market. Section 5 is the conclusion of this paper.

2. Literature review

This part reviews the existing literature on the identification of SIFIs by focusing on three aspects: indicator-based approach, market-based model approach, and Copula-CoVaR method.

Research on the indicator-based approach has shown that the index system for identifying global systemically important financial institutions (G-SIFIs) used by the Financial Stability Board (FSB) considers cross-border transaction activities, business scale, relevance, substitutability, and complexity. Lo (2009) analyzed the systemic risks of institutions based on factors such as business concentration, financial leverage, relevance, risk sensitivity, and closeness between institutions. Billio et al. (2012) used the Granger causality test to establish a financial network based on the stock price

data of major financial institutions around the world; they then measured the systemic importance of financial institutions based on indicators such as transmittance and acceptance. Guo (2013) ranked China's systemically important banks based on different evaluation index systems used in China and internationally. Thomson (2015) pointed out that assessing the importance of financial institutions should include factors such as scale, concentration, relevance, contagion, and environment.

The research into the market-based approach showed that studies commonly used market model methods including the Shapley value (Tarashev et al., 2009), systematic impact index (SII), vulnerability index (VI) (Zhou, 2010), conditional value at risk (CoVaR) (Adrian & Brunnermeier, 2009), marginal expected shortfall (MES) (Acharya et al., 2017), component expected shortfall (CES) (Banulescu & Dumitrescu, 2015), and SRISK (Banulescu-Radu et al., 2021; Brownlees & Engle, 2017).

Furthermore, in the research based on the Copula-CoVaR model, Karimalis and Nomikos (2018) and Hakwa et al. (2012) proposed a model to measure the contribution of marginal system risk based on the copula theory and the CoVaR method; this model not only characterizes the linear correlation between financial institutions and the entire financial system but also captures nonlinear tail correlation characteristics. Reboredo and Ugolini (2015) and Boako and Alagidede (2018) applied the Copula-CoVaR method to the European debt market and African stock market, respectively. Compared with other methods of estimating CoVaR (e.g., quantile regression, tail risk network, and multivariate GARCH model), the Copula-CoVaR method can estimate the dynamic and asymmetric tail correlation between data more flexibly (Bernardi et al., 2017; Mainik & Schaanning, 2014; Patton, 2006; Reboredo & Ugolini, 2015).

Furthermore, in the research based on the Copula-CoVaR model, Karimalis and Nomikos (2018) and Hakwa et al. (2012) proposed a model to measure the contribution of marginal system risk based on the copula theory and the CoVaR method; this model not only characterizes the linear correlation between financial institutions and the entire financial system but also captures nonlinear tail correlation characteristics. Reboredo and Ugolini (2015) and Boako and Alagidede (2018) applied the Copula-CoVaR method to the European debt market and African stock market, respectively. Compared with other methods of estimating CoVaR (e.g., quantile regression, tail risk network, and multivariate GARCH model), the Copula-CoVaR method can estimate the dynamic and asymmetric tail correlation between data more flexibly (Bernardi et al., 2017; Mainik & Schaanning, 2014; Patton, 2006; Reboredo & Ugolini, 2015).

3. Methodology

This section briefly describes the PSO-SVM model and Copula-CoVaR model that constitute the PSO-SVM-Copula-CoVaR model; it also, introduces the main steps for constructing the PSO-SVM-Copula-CoVaR model.

3.1. The PSO-SVM model

Determining edge distribution estimation is equivalent to solving the following integral equation:

$$\int_0^1 \theta(x-t)p(t)dt = F(x) \tag{1}$$

This article uses independent and identically distributed data x_1, x_2, $\cdots$, x_l to construct the following empirical distribution function::

$$F_l(x) = \frac{1}{l}\sum_{i=1}^{l}\theta(x-x_i) \tag{2}$$

At the same time, the boundary conditions (0,0), (1,1) are added to solve the problem of edge distribution by using a support vector machine.

The objectives for this paper are as follows: firstly, define the corresponding regression problem in image space. Secondly, construct the kernel function $K(x_i, x_j)$ by using the support vector machine method. Thirdly, construct the cross kernel function $k(x_i, t)$. Fourthly, according to the kernel function $K(x_i, x_j)$, use the support vector method to solve the regression problem-that is, find out the support vector x_i, $i = 1, 2, \ldots, N$ and the corresponding coefficients $\beta_i = (\alpha_i^* - \alpha_i)$, $i = 1, 2, \ldots, N$, in which $F(x)$ is the desired edge distribution.

$$F(x) = \sum_{k=1}^{N}\beta_k^0 K(x_i, x) \tag{3}$$

3.2. The Copula-CoVaR model

Firstly, VaR_α^i is the value in the α quantile.

$$\Pr\left(x_t^i \leq VaR_\alpha^i\right) = \alpha \tag{4}$$

$CoVaR_\beta^{j|i}$ is defined as the VaR of institution j (or the financial system) under the condition that $x_t^i \leq VaR_{\alpha, t}^i$ when certain events occur in institution i. $CoVaR_\beta^{j|i}$ is the value of the β quantile under the conditional probability distribution.

$$\Pr\left(x_t^j \leq CoVaR_{\beta, t}^{j|i}|x_t^i \leq VaR_\alpha^i\right) = \beta \tag{5}$$

In Formula (5), x_t^j and x_t^i are the returns sequence of sequence j and i, respectively. The following is for solving $CoVaR_\alpha^{j|i}$ by analyzing the copula function.

$$\Pr\left(x_t^j \leq CoVaR_{\beta, t}^{j|i}|x_t^i \leq VaR_{\alpha, t}^i\right) = \alpha\beta \tag{6}$$

According to Sklar's theorem:

$$C(F_j(CoVaR_{\beta, t}^{j|i}), F_i(VaR_{\alpha, t}^i)) = \alpha\beta \tag{7}$$

Based on Formula (7), $u_j = F_j(CoVaR_{\beta,t}^{j|i})$ can be obtained, and then:

$$CoVaR_{\beta,t}^{j|i} = F_i^{-1}(u_j) \tag{8}$$

Here, $u_j = F_j(x_t^j)$, $u_i = F_i(x_t^i)$, F_j and F_i are the marginal distributions of x_t^j and x_t^i, respectively.

3.3. The PSO-SVM-Copula-CoVaR model

Based on the sub-model introduced above, this paper constructs a new PSO-SVM-Copula-CoVaR model to identify systemically important financial institutions in China. Figure 1 is a flowchart of the PSO-SVM-Copula-CoVaR model, which involves the following four main steps:

Step 1: Model the marginal distribution of the stock price data of financial institutions. The PSO-SVM marginal distribution estimation method of is used to model the marginal distribution of the stock price data of each listed financial institution, and the copula data of each financial institution is obtained through probability integral conversion.

Step 2: The union sequence of financial institution assets is constructed, and the union sequence $z^i = F(x_1^i, x_2^i) = C(F_1(x_1^i), F_2(x_2^i))$ is obtained by using the definition of the copula function, where z^i represents the overall distribution characteristics of institution i.

Step 3: Estimate the risk spillover value of institution i. Based on Step 2, the Copula parameters are estimated by the maximum likelihood method, and CoVaR is calculated with the Copula-CoVaR model.

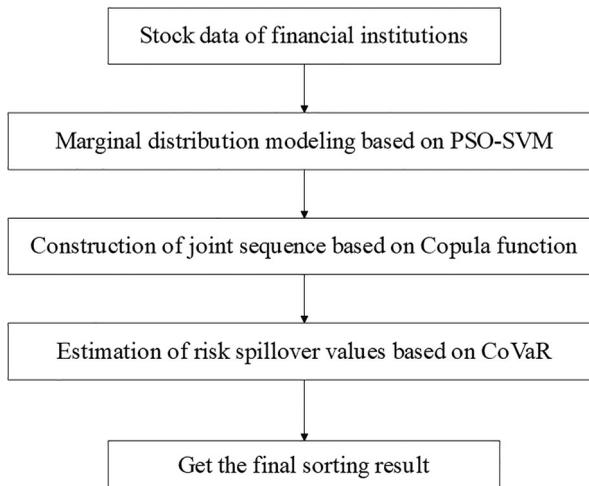Step 4: Sort the estimated CoVaR values of each financial institution to get the final result.

Figure 1. The flow chart of PSO-SVM-Copula-CoVaR model.
Source: Authors.

4 Empirical results

4.1. Data selection and basic analysis

In an efficient market, the roles, statuses, and interrelationships of institutions in the financial system can be fully reflected in the volatility and interrelationship patterns of returns in the stock market. At the same time, the daily return rates of stocks haves the beneficial characteristics of easy access and large sample size, which can better reflect the financial system's real situation.

China's financial institutions can be divided into four categories (i.e., banking, insurance, securities, and trust) according to their industries, among which banking and insurance are the two most important pillars. With the development of interest rate marketization, the traditional asset-liability business of commercial banks has been overtaken by increasingly diversified investment and financing channels; this also makes the business dealings among financial institutions in the system closer. Therefore, this paper uses 28 financial institutions as samples, choosing the daily stock return rate of each listed financial institution as its research object. The total sample is comprised of 16 banks and 12 insurance/securities/trust institutions.

In addition, considering the balance of the sample number (which covers more listed companies) and sample size (which accounts for a longer time interval), the sample time period is from January 4, 2011 to June 30, 2022, covering a total of 2,191 trading days. In this sample, the stock price rate of return $r_t = lnP_t - lnP_{t-1}$, $t = 1, 2, \ldots, T$. where r_t is the rate of return, and P_t is the closing price at time t. The analysis data comes from Wind Information; R and Matlab software are used to process the data. Table 1 shows the basic information of the selected financial institutions, and Table A2 in Appendix A shows the results of the descriptive statistical analysis of the data.

The results in Table A2 in Appendix A show that the maximum value of the daily stock return of financial institutions is ±0.1, which is determined by the ups and downs of the Chinese stock market. The standard deviation results show that the volatility of non-bank financial institutions, especially securities institutions, is significantly greater than that of banking institutions; it also demonstrates that, the volatility of small-and medium-sized joint-stock banks is greater than that of state-owned commercial banks. In the data's distribution pattern, the skewness of the yield data of all financial institutions is greater than 0, and the kurtosis values of most financial institutions are greater than 3. These distributions are representative of the typical 'peak and thick tail' distribution characteristics of financial assets, which is confirmed again by the value of the Jarque-Bera statistics.

4.2. Construction and comparison of different marginal distribution models

In the calculation of the Copula-CoVaR model, it is particularly important to establish a suitable marginal distribution model for financial assets. At present, the modeling of marginal distribution is mainly achieved through the parametric method and non-parametric method. The parametric method mainly refers to the introduction of parameter distributions (e.g., normal distribution and I distribution) based on the

Table 1. Basic information of financial institutions.

Number	Code	Full name of the institution	Abbreviation
1	000001.SZ	PingAn Bank	PAB
2	002142.SZ	Bank of Ningbo	BONB
3	600000.SH	Shanghai Pudong development Bank	SPDB
4	600015.SH	HuaXia Bank	HXB
5	600016.SH	China Minsheng banking Corp., Ltd	CMBC
6	600036.SH	China merchants bank	CMB
7	601009.SH	Bank of Nanjing	BONJ
8	601166.SH	Industrial Bank	IB
9	601169.SH	Bank of Beijing	BOBJ
10	601288.SH	Agricultural Bank of China	ABC
11	601328.SH	Bank of communications	BOCOM
12	601398.SH	Industrial and commercial Bank of China Ltd.	ICBOC
13	601939.SH	China construction bank	CCB
14	601988.SH	Bank of China	BOC
15	601998.SH	China CITIC bank	CNCB
16	601818.SH	China Everbright Bank	CEBB
17	000686.SZ	Northeast securities	NES
18	000728.SZ	Guoyuan securities	GYS
19	000783.SZ	Changjiang securities	CJS
20	600030.SH	CITIC securities	CITICS
21	600109.SH	Sinolink securities	SLS
22	600837.SH	Haitong securities	HTS
23	601318.SH	Ping An insurance (group) Co. of China, Ltd.	PAICOC
24	601601.SH	China Pacifific insurance (group) Co., Ltd.	CPIC
25	601628.SH	China life insurance (group) co., Ltd.	CLIC
26	600643.SH	Shanghai Aj Group Co., Ltd	SAGC
27	000563.SZ	Shanxi International Trust & Investment Corp., Ltd.	SITI
28	601099.SH	The Pacific Securities Co., ltd	PSC

Source: Authors.

GARCH model. This method, used to describe the residual distribution of yield series, reflects the series's characteristics of peak, thick tail, aggregation, and asymmetry. The non-parametric method refers to directly using the kernel density function or the empirical distribution function to describe the marginal distribution of the return sequence. In the interest of effective econometric modeling, the marginal distribution that best reflects the distribution characteristics of financial assets should be selected for research. If the fitting result of the selected marginal distribution model is quite different from the actual financial asset distribution characteristics, it will result in greater errors during the selection of the subsequent copula model, the estimation of parameters, and the calculation of CoVaR. Thus, this paper seeks the marginal distribution closest to the distribution characteristics of financial assets using three methods: the GJR-GARCH model, kernel density estimation, and the PSO-SVM algorithm.

4.2.1. The GJR-GARCH model

The development of the GARCH model in structural form and distribution is meant to better describe the characteristics of asset return series. The empirical data shows the advantages of the asymmetric GARCH model in the GARCH model (Abad et al., 2014); the GJR-GARCH model is found to be more effective than other asymmetric models in describing the behavior of asymmetric fluctuations in the financial market. Therefore, this paper adopts the GJR-GARCH model as the representative model of the GARCH model in the parametric method.

The GJR-GARCH (1,1,1) model proposed by Glosten et al. is as follows:

$$\varepsilon_t \equiv z_t h_t, \varepsilon_t | I_{t-1} \sim N(0, h_t^2) \tag{9}$$

$$h_t^2 = w + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1}^2 + \gamma S_{t-1} \varepsilon_{t-1}^2 \tag{10}$$

In these formulas, z_t is an independent and identically distributed random variable, and I_{t-1} represents the set of all available information in the $t-1$ period. If $\varepsilon_t < 0$, then $S_{t-1} = 1$, otherwise $S_{t-1} = 0$. The parameters w, γ, α, and β are all non-random real numbers, and γ is a parameter that reflects the asymmetry of fluctuations. In order to ensure a positive conditional variance, the parameters must adhere to the following constraints: $w > 0$, $\alpha \geq 0$, $\beta \geq 0$, and $\alpha + \beta \geq 0$. Additionally, $\alpha + \frac{\gamma}{2} + \beta < 1$ is a necessary and sufficient condition for the wide-sense stationarity of the GJR-GARCH model.

4.2.2. The kernel estimation method
In the research of kernel density estimation method, the minimum square error of candidate distribution and empirical distribution is taken as the reference for selecting the optimal kernel function and window width. Firstly, under the condition of the same window width, the optimal kernel function is selected as Gaussian kernel function, rather than a box kernel function, epanechnikov kernel function, or and triangle kernel function. Secondly, the optimal window width is found to be 0.0001 in the neighborhood (0,5) with the default window width of 0.003 as the center.

4.2.3. The PSO-SVM estimation method
Although the non-parametric kernel density estimation method overcomes the shortcomings of the historical simulation method, problems still persist, including its heavy dependence on sample data selection and slow response to emergencies. The support vector machine is a novel small sample learning method with a solid theoretical foundation. The SVM's goal is to obtain the optimal solution in accordance with the existing information rather than just the optimal value when the number of samples tends to infinity. The algorithm converts the actual problem into a high-dimensional feature space through nonlinear transformation, which can ensure that the algorithm has good generalizability. A small number of support vectors determine the final result, which basically does not involve probability measures and the law of large numbers. These vectors, can grasp key samples and eliminate a large number of redundant samples.

Consider the study of the marginal distribution of the SVM estimation method. In the parameter-setting scenario with penalty coefficient $C = 1$ and window width $\sigma = 16$, the error rate of three-fold cross-validation is taken as the reference to select the optimal kernel function and window width; the radial basis kernel is also selected as the estimation function of marginal distribution in this scenario. Secondly, the PSO algorithm is used to find the optimal parameter combination, and the search intervals of the penalty coefficient and window width are [0.5,3] and [10,20], respectively.

Finally, the optimal parameter combination is found as penalty coefficient $C = 1.99556$ and window width $\sigma = 16.99585$.

4.3. Comparison of the results of different marginal distribution estimation methods

The purpose of using the above three methods to estimate the marginal distribution of the return rate series is to find the marginal distribution that best reflects the distribution characteristics of financial assets. The other purpose of the marginal distribution estimation is to estimate the parameters of the copula function. Therefore, this paper adopts the maximum likelihood method to estimate the six types of copula functions (Gaussian copula, t-copula, Clayton copula, Gumbel copula, Frank copula, and Joe copula), as well as investigating the optimal copula function form by using the information criterion. Relevant copula functions are described in the Appendix. It is estimated that the minimum mean square error considers the optimal joint distribution under different methods. The specific estimation results are shown in Table 3.

As can be seen from Table 2, the accuracy of the model's analysis results are affected by the limitations of the econometric model itself, such as the need to assume the conditional distribution of the residuals in advance, and its inability to adapt well to the characteristics of the financial price data. However, the edge distribution estimation method based on PSO-SVM has an absolute advantage in both the performance of marginal distribution and the results of copula function. The

Table 2. Mean square error of different marginal distribution methods.

Institutions	GJR-GARCH	Kendisty	PSO-SVM
PAB	0.0037	0.0014	**0.0001**
BONB	0.0018	0.0013	**0.0002**
SPDB	0.0061	0.0022	**0.0002**
HXB	0.0096	0.0033	**0.0005**
CMBC	0.0094	0.0056	**0.001**
CMB	0.004	0.0014	**0.0002**
BONJ	0.0023	0.0024	**0.0001**
IB	0.0033	0.0015	**0.0001**
BOBJ	0.0094	0.0045	**0.0001**
ABC	0.0145	0.0272	**0.0004**
BOCOM	0.0114	0.0072	**0.0002**
ICBOC	0.0068	0.0108	**0.0001**
CCB	0.0053	0.0059	**0.0005**
BOC	0.0188	0.0191	**0.0003**
CNCB	0.0106	0.005	**0.0007**
CEEB	0.0193	0.0102	**0.0002**
NES	0.0085	0.0016	**0.0004**
GYS	0.006	0.0014	**0**
CJS	0.0077	0.0019	**0.0001**
CITICS	0.003	0.0012	**0.0001**
SLS	0.0037	0.0013	**0**
HTS	0.0093	0.0016	**0.0001**
PAICOC	0.0038	0.0012	**0**
CPIC	0.0107	0.0032	**0.0007**
CLIC	0.0073	0.0011	**0**
SAGC	0.0064	0.0015	**0.0001**
SITI	0.0078	0.0025	**0.0005**
PSC	0.0026	0.0011	**0.0002**

Notes: Bold value is the minimum value.
Source: Authors.

marginal distribution estimation of PSO-SVM does not just avoid the loss of data volume caused by excluding some samples (like the traditional nonparametric method), but it also obtains a distribution that is closer to the return series of financial assets than the parametric method's distribution. In addition, because of the existence of support vectors, the method is not limited by the sample data. In view of these findings, this paper uses an integrated system composed of a machine learning method (PSO-SVM) and copula function to further study the identification of systemically important financial institutions at the micro level.

4.4. Result analysis based on the Copula-CoVaR model

Table 3 shows CoVaR values calculated through different edge distribution estimation methods based on the Shanghai Securities Composite Index. Table 3 shows the CoVaR values of the top five financial institutions calculated under different methods. According to the weak-form efficient market hypothesis, market prices fully reflect all available information, thus capturing more micro-level information while taking into account the ever-changing market risks.

The table makes it evident that, due to the special characteristics of China's financial market, commercial banks were born out of the 'universal' national banking system in the early days of the founding of the People's Republic of China. Even though

Table 3. CoVaR results of different marginal distribution estimation methods (calculation results based on Shanghai Securities Composite Index).

Institutions	GJR-GARCH	Kendisty	PSO-SVM
PAB	−0.0272	−0.0782	−0.0782
BONB	−0.0333	−0.0305	−0.0303
SPDB	−0.017	−0.0234	−0.0232
HXB	−0.0142	−0.0227	−0.0227
CMBC	−0.0121 (#4)	−0.0267	−0.048
CMB	−0.03	−0.0758	−0.0758
BONJ	−0.0291	−0.026	−0.0254
IB	−0.0232	−0.1018	−0.1018
BOBJ	−0.0109 (#2)	−0.0207	−0.0207 (#5)
ABC	−0.0112 (#3)	−0.0153 (#2)	−0.0153 (#2)
BOCOM	−0.0136	−0.0178 (#4)	−0.0179 (#4)
ICBOC	−0.0143	−0.017 (#3)	−0.017 (#3)
CCB	−0.0131 (#5)	−0.0199	−0.0474
BOC	−0.0103 (#1)	−0.0148 (#1)	−0.0149 (#1)
CNCB	−0.0173	−0.0234	−0.023
CEEB	−0.0168	−0.0214	−0.0214
NES	−0.0353	−0.037	−0.0369
GYS	−0.0308	−0.0357	−0.0357
CJS	−0.0308	−0.0819	−0.0819
CITICS	−0.0368	−0.0328	−0.0328
SLS	−0.0412	−0.0424	−0.0424
HTS	−0.0281	−0.0319	−0.0321
PAICOC	−0.027	−0.082	−0.082
CPIC	−0.0436	−0.0875	−0.0875
CLIC	−0.0484	−0.0812	−0.0812
SAGC	−0.0277	−0.039	−0.039
SITI	−0.0353	−0.0892	−0.0892
PSC	−0.0338	−0.0335	−0.033

Notes: #1–#5 are financial institutions ranked 1 to 5, respectively, in the calculation results.
Source: Authors.

the financial system has diversified, large commercial banks still occupy the dominant position in China's financial system, making them systemically important. The results of different estimation methods consistently show that the banking industry's listed companies have an advantage in systemic importance, indicating that the financial system is still bank-led; some joint-stock commercial banks and city commercial banks also have systemic importance that cannot be ignored and should also be considered in the important matters of financial supervision.

Furthermore, in order to verify the robustness of the model, Table 4 shows the value of CoVaR calculated based on the Shenzhen Securities Component Index. A comparison of Tables 3 and 4 reveals that the ranking of SIFIs is consistent, which proves the robustness of the model constructed in this paper.

Our results prove that, more than other types of financial institutions, banks are still the main contributors to financial systemic risks. Since the risk contagion between financial institutions is experiencing evolving and dynamic change, the estimation may deviate from the actual figure if the size of financial institutions is considered the most important determinant in identifying SIFIs. In the practice of financial supervision, we should not only focus on the traditional large-scale state-owned commercial banks with systematic importance; we should also strengthen the supervision of financial institutions with high vulnerability, potential damage and high debt ratios, even if the asset scale is not large enough.

Table 4. CoVaR results of different marginal distribution estimation methods (calculation results based on Shenzhen Securities Component Index).

Institutions	GJR-GARCH	Kendisty	PSO-SVM
PAB	−0.0272	−0.0781	−0.0781
BONB	−0.0332	−0.0302	−0.0299
SPDB	−0.0169	−0.0231	−0.0228
HXB	−0.0142	−0.0225	−0.0224
CMBC	−0.012 (#4)	−0.0592	−0.0592
CMB	−0.0299	−0.0271	−0.0268
BONJ	−0.0291	−0.0258	−0.0251
IB	−0.0231	−0.0269	−0.027
BOBJ	−0.0108 (#2)	−0.0205	−0.0204 (#5)
ABC	−0.011 (#3)	−0.0149 (#2)	−0.0149 (#2)
BOCOM	−0.0134	−0.0176 (#4)	−0.0175 (#4)
ICBOC	−0.0141	−0.0165 (#3)	−0.0165 (#3)
CCB	−0.013 (#5)	−0.0195 (#5)	−0.0474
BOC	−0.0103 (#1)	−0.0146 (#1)	−0.0146 (#1)
CNCB	−0.0172	−0.0232	−0.0227
CEEB	−0.0167	−0.0212	−0.021
NES	−0.0352	−0.037	−0.0369
GYS	−0.0307	−0.0357	−0.0358
CJS	−0.0307	−0.0346	−0.0347
CITICS	−0.0367	−0.0328	−0.0327
SLS	−0.0411	−0.0423	−0.0423
HTS	−0.028	−0.0812	−0.0812
PAICOC	−0.0269	−0.0819	−0.0819
CPIC	−0.0435	−0.0875	−0.0875
CLIC	−0.0482	−0.0812	−0.0812
SAGC	−0.0277	−0.0389	−0.0389
SITI	−0.0352	−0.0794	−0.0794
PSC	−0.0337	−0.0333	−0.0327

Notes: #1–#5 are financial institutions ranked 1 to 5, respectively, in the calculation results.
Source: Authors.

5. Conclusions

Since the global financial crisis in 2008, accurately identifying SIFIs has become an increasingly prominent concern, especially considering that traditional methods of identifying SIFIs ignore the application of machine learning models. Traditional measurement models have requirements for stable data and sample size, and their recognition capabilities are limited.

The empirical results show that, on the one hand, the marginal distribution estimation method based on PSO-SVM can better fit the distribution of financial asset return sequences (both in terms of the performance of marginal distribution and the results of the copula function); in short, the method can effectively identify China's systemically important financial institutions. Findings indicate that the ranking of SIFIs is roughly the same as the ranking in reality, meaning that the PSO-SVM-Copula-CoVaR model constructed in this paper can more effectively identify China's SIFIs and provide a new method for the regulatory authorities to identify SIFIs and improve regulatory efficiency.

However, due to the varying levels of information disclosure from financial institutions, this study is limited by the fact that the data used is single stock price data. Once the financial market information disclosure system is improved, future studies can include other types of financial data in the research of the model, which would better realize the dynamic identification of SIFIs.

Disclosure statement

The authors declare that they have no conflict of interest.

Funding

ORCID

Tingting Zhang http://orcid.org/0000-0001-9478-9564

References

Abad, P., Benito, S., & López, C. (2014). A comprehensive review of value at risk methodologies. *The Spanish Review of Financial Economics*, *12*(1), 15–32. https://doi.org/10.1016/j.srfe.2013.06.001

Acharya, V. V., Pedersen, L. H., Philippon, T., & Richardson, M. (2017). Measuring systemic risk. *Review of Financial Studies*, *30*(1), 2–47. https://doi.org/10.1093/rfs/hhw088

Adrian, T., & Brunnermeier, M. K. (2009). CoVaR. *Journal of Labor Economics*, *55*(6), 301–348.

Banulescu, G. D., & Dumitrescu, E.-I. (2015). Which are the SIFIs? A component expected shortfall approach to systemic risk. *Journal of Banking & Finance*, *50*, 575–588. https://doi.org/10.1016/j.jbankfin.2014.01.037

Banulescu-Radu, D., Hurlin, C., Leymarie, J., & Scaillet, O. (2021). Backtesting marginal expected shortfall and related systemic risk measures. *Management Science*, *67*(9), 5730–5754. https://doi.org/10.1287/mnsc.2020.3751

Benoit, S., Colletaz, G., & Hurlin, C. (2013). A theoretical and empirical comparison of systemic risk *m*easures. HEC Paris Research Paper.

Bernal, O., Gnabo, J. Y., & Guilmin, G. (2014). Assessing the contribution of banks, insurance and other financial services to systemic risk. *Journal of Banking & Finance*, *47*, 270–287. https://doi.org/10.1016/j.jbankfin.2014.05.030

Bernardi, M., Durante, F., & Jaworski, P. (2017). CoVaR of families of copulas. *Statistics & Probability Letters*, *120*, 8–17. https://doi.org/10.1016/j.spl.2016.09.005

Billio, M., Getmansky, M., Lo, A. W., & Pelizzon, L. (2012). Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *Journal of Financial Economics*, *104*(3), 535–559. https://doi.org/10.1016/j.jfineco.2011.12.010

Boako, G., & Alagidede, P. (2018). Systemic risks spillovers and interdependence among stock markets: International evidence with CovaR-copulas. *South African Journal of Economics*, *86*(1), 82–112. https://doi.org/10.1111/saje.12182

Brämer, P., & Gischer, H. (2013). An assessment methodology for domestic systemically important banks in Australia. *Australian Economic Review*, *46*(2), 140–159. https://doi.org/10.1111/j.1467-8462.2013.12008.x

Brownlees, C. T., & Engle, R. F. (2017). Srisk: A conditional capital shortfall measure of systemic risk. *Review of Financial Studies*, *30*(1), 48–79. https://doi.org/10.1093/rfs/hhw060

Chen, H., Li, Z., & Liu, K. (2021). Research on identification of systemically important financial institutions from the perspective of macro-prudential based on the index model. In *Proceedings—2nd International Conference on E-Commerce and Internet Technology, ECIT* (pp. 325–330).

Dutta, H. (2022). A consensus algorithm for linear support vector machines. *Management Science*, *68*(5), 3703–3725. https://doi.org/10.1287/mnsc.2021.4042

Eberhart, R., & Kennedy, J. (1995). New optimizer using particle swarm theory. In *Proceedings of the International Symposium on Micro Machine and Human Science* (pp. 39–43).

Fan, H., Amalia, A., & Gao, Q. (2018). The assessment of systemic risk in the Kenyan banking sector. *Complexity*, *2018*, 1–15. https://doi.org/10.1155/2018/8767836

Guo, W. (2013). Assessment of the systematic importance of Chinese listed banks: Indicator approach based positive analysis. *Modern Economic Science*, *35*(2), 28–35.

Hakwa, B., Jäger-Ambrożewicz, M., & Rüdiger, B. (2012). *Measuring and analysing marginal systemic risk contribution using CoVaR: A copula approach*. Working Paper.

Jaeger-Ambrozewicz, M. (2013). A critical note on delta-CoVaR. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2210045

Johnson, M., & Mamun, A. (2012). The failure of Lehman brothers and its impact on other financial institutions. *Applied Financial Economics*, *22*(5), 375–385. https://doi.org/10.1080/09603107.2011.613762

Karimalis, E. N., & Nomikos, N. K. (2018). Measuring systemic risk in the European banking sector: A copula CoVaR approach. *The European Journal of Finance*, *24*(11), 944–975. https://doi.org/10.1080/1351847X.2017.1366350

Kim, K. J. (2003). Financial time series forecasting using support vector machines. *Neurocomputing*, *55*(1–2), 307–319. https://doi.org/10.1016/S0925-2312(03)00372-2

Kleinow, J., Horsch, A., & Garcia, M. (2014). Determinants of systemically important banks in Latin America. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2441868

Lee, Y. C. (2007). Application of support vector machines to corporate credit rating prediction. *Expert Systems with Applications*, *33*(1), 67–74. https://doi.org/10.1016/j.eswa.2006.04.018

Li, H., Wu, Y. J., Zhang, S., & Zou, J. (2021). Temporary rules of retail product sales time series based on the matrix profile. *Journal of Retailing and Consumer Services*, *60*, 102431. https://doi.org/10.1016/j.jretconser.2020.102431

Liu, G. W. (2015). Simulating risk contributions of credit portfolios. *Operations Research*, *63*(1), 104–121. https://doi.org/10.1287/opre.2015.1351

Lo, A. W. (2009). *Hedge funds, systemic risk, and the financial crisis of 2007–2008: Written testimony for the house oversight committee hearing on hedge funds*. SSRN Working Paper Series.

López-Espinosa, G., Moreno, A., Rubia, A., & Valderrama, L. (2012). Short-term wholesale funding and systemic risk: A global CoVaR approach. *Journal of Banking & Finance*, *36*(12), 3150–3162. https://doi.org/10.1016/j.jbankfin.2012.04.020

Lu, J., & Hu, X. (2014). Novel three-bank model for measuring the systemic importance of commercial banks. *Economic Modelling*, *43*, 238–246. https://doi.org/10.1016/j.econmod.2014.08.007

Luo, J., Yan, X., & Tian, Y. (2020). Unsupervised quadratic surface support vector machine with application to credit risk assessment. *European Journal of Operational Research*, *280*(3), 1008–1017. https://doi.org/10.1016/j.ejor.2019.08.010

Mainik, G., & Schaanning, E. (2014). On dependence consistency of CoVaR and some other systemic risk measures. *Statistics & Risk Modeling*, *31*(1), 49–77. https://doi.org/10.1515/strm-2013-1164

Patton, A. J. (2006). Modelling asymmetric exchange rate dependence. *International Economic Review*, *47*(2), 527–556. https://doi.org/10.1111/j.1468-2354.2006.00387.x

Reboredo, J. C., & Ugolini, A. (2015). Systemic risk in European sovereign debt markets: A CoVaR-copula approach. *Journal of International Money and Finance*, *51*(C), 214–244. https://doi.org/10.1016/j.jimonfin.2014.12.002

Sun, X., Liu, C., Wang, J., & Li, J. (2020). Assessing the extreme risk spillovers of international commodities on maritime markets: A GARCH-Copula-CoVaR approach. *International Review of Financial Analysis*, *68*, 101453. https://doi.org/10.1016/j.irfa.2020.101453

Tarashev, N., Borio, C., & Tsatsaronis, K. (2009). The systemic importance of financial Institutions. *BIS Quarterly Review*.

Thomson, J. B. (2015). On systemically important financial institutions and progressive systemic mitigation. *DePaul Business and Commercial Law Journal*, *8*(2), 135.

Vapnik, V. (1995). *The nature of statistical learning theory* (pp. 988–999). Springer.

Vapnik, V. (1998). *Statistical learning theory*. John Wiley and Sons.

Wu, F., Zhang, Z., Zhang, D., & Ji, Q. (2021). Identifying systemically important financial institutions in China: new evidence from a dynamic copula-CoVaR approach. *Annals of Operations Research*. https://doi.org/10.1007/s10479-021-04176-z

Xu, C. (2011). Review on the identification methods of systemically important financial institutions. *International Financial Research*, *11*, 57–64.

Yao, S. J., He, H. B., Chen, S., & Ou, J. H. (2018). Financial liberalization and cross-border market integration: Evidence from China's stock market. *International Review of Economics & Finance*, *58*, 220–245. https://doi.org/10.1016/j.iref.2018.03.023

Zhang, D. (2017). Oil shocks and stock markets revisited: Measuring connectedness from a global perspective. *Energy Economics*, *62*, 323–333. https://doi.org/10.1016/j.eneco.2017.01.009

Zhang, Z., Zhang, D., Wu, F., & Ji, Q. (2021). Systemic risk in the Chinese financial system: A copula-based network approach. *International Journal of Finance & Economics*, *26*(2), 2044–2063. https://doi.org/10.1002/ijfe.1892

Zhou, C. (2010). Are banks too big to fail? Measuring systemic importance of financial institutions. *International Journal of Central Banking*, *6*, 205–250.

Appendix A

The binary distribution joint distribution of copula function consists of two parts: distribution function of variables and copula function representing variables. The form of the joint distribution function of the corresponding strength parameters φ_0 and $\Delta\varphi$ is as follows:

$$F(\varphi_0, \Delta\varphi) = C\big(F_1(\varphi_0), F_2(\Delta\varphi); \theta\big) = C(u_1, u_2; \theta)$$

In this formula, $u_1 = F_1(\varphi_0)$ and $u_2 = F_2(\Delta\varphi)$ are the distribution functions of strength parameters, and θ is the parameter of copula function. Table A1 shows the relevant information of the six functions involved in this study.

See Table A2.

Table A1. Six different copula functions.

Copula	Definition of $C(u_1, u_2; \theta)$	θ limits
Gaussian	$\Phi_\theta\big(\Phi^{-1}(u_1), \Phi^{-1}(u_2); \theta\big)$	$[-1, 1]$
T	$T_2\big(T_v^{-1}(u_1), T_v^{-1}(u_2); \theta, v\big)$	$[-1, 1]$
Clayton	$\big(u_1^{-\theta} + u_2^{-\theta} - 1\big)^{-1/\theta}$	$(0, \infty)$
Gumbel	$\exp\left\{-\left[(-ln u_1)^\theta + (-ln u_2)^\theta\right]^{1/\theta}\right\}$	$[1, \infty)$
Frank	$-\dfrac{1}{\theta}\ln\left[1 + \dfrac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)}{e^{-\theta} - 1}\right]$	$(-\infty, \infty) \setminus \{0\}$
Joe	$1 - \left((1 - u_1)^\theta + (1 - u_2)^\theta - (1 - u_1)(1 - u_2)^\theta\right)^{1/\theta}$	$[1, \infty)$

Source: Authors.

Table A2. Descriptive statistics of financial institutions' yield series.

Institutions	Mean value	Maximum	Minimum	Standard deviation	Skewness	Kurtosis	The value of JB
PAB	−0.0004	0.1004	−0.1002	0.0217	0.4291	3.2719	1048.0822
BONB	0	0.1004	−0.0998	0.021	0.3585	2.7689	749.5641
SPDB	0	0.1003	−0.1002	0.0161	0.5181	6.9072	4464.9874
HXB	0	0.1005	−0.1003	0.0159	0.5266	7.2739	4944.0742
CMBC	0	0.1001	−0.0999	0.0156	0.7018	10.0187	9365.2931
CMB	−0.0005	0.1003	−0.0991	0.0184	0.3989	3.2757	1041.2027
BONJ	0	0.1003	−0.0996	0.018	0.3873	4.4822	1894.4847
IB	−0.0006	0.1005	−0.1002	0.0187	0.4973	4.8305	2226.9294
BOBJ	0	0.0996	−0.0997	0.0145	0.5356	8.5778	6838.5964
ABC	0	0.1012	−0.099	0.0119	0.6253	12.5867	14638.5466
BOCOM	0	0.101	−0.0999	0.0136	0.9544	12.7164	15128.866
ICBOC	0	0.0853	−0.099	0.012	0.2323	8.8938	7258.5516
CCB	0	0.1004	−0.0994	0.0142	0.5048	8.1101	6112.7821
BOC	0	0.1014	−0.1004	0.0118	0.8192	15.4914	22201.6703
CNCB	0	0.1009	−0.1003	0.0173	0.8425	8.1371	6319.3392
CEEB	0	0.1014	−0.0992	0.0156	0.929	7.4127	5344.8072
NES	0	0.1005	−0.1001	0.0248	0.2152	3.7053	1274.4205
GYS	0	0.1005	−0.1001	0.0241	0.2447	3.7559	1313.9505
CJS	0	0.1007	−0.1008	0.0241	0.3832	3.9736	1499.8072
CITICS	−0.0008	0.1004	−0.1001	0.0235	0.4675	4.2214	1711.8527
SLS	−0.0005	0.1005	−0.1003	0.0282	0.2921	2.8126	756.0823
HTS	−0.0007	0.1005	−0.1001	0.023	0.4589	4.5373	1962.1611
PAICOC	−0.0002	0.0964	−0.1	0.0191	0.2638	3.0363	870.1473
CPIC	0	0.102	−0.1001	0.0255	0.4426	3.8837	1453.0792
CLIC	−0.0005	0.1004	−0.1001	0.0223	0.6184	4.2513	1794.9185
SAGC	0	0.1005	−0.1002	0.0254	0.1512	3.3439	1032.7094
SITI	0	0.1017	−0.1007	0.0267	0.2726	3.3897	1079.7279
PSC	0	0.1001	−0.1001	0.0218	0.2471	2.2402	482.4474

Source: Authors.