risks

Article

Exploring Systemic Risk Dynamics in the Chinese Stock Market: A Network Analysis with Risk Transmission Index

Xiaowei Zeng [1], Yifan Hu [2,*], Chengjun Pan [2] and Yanxi Hou [2,*]

1 School of Management, Fudan University, Shanghai 200433, China; xwzeng20@fudan.edu.cn
2 School of Data Science, Fudan University, Shanghai 200433, China
* Correspondence: huyf23@m.fudan.edu.cn (Y.H.); yxhou@fudan.edu.cn (Y.H.)

Abstract: Systemic risk refers to the potential for a disruption in one part of a financial system to trigger a cascade of adverse effects, impacting the functioning of the system. Despite the progress on novel systemic risk measures, research on dynamics of systemic risk network structure and its community effect is still in its initial state. In this study, we utilize price data from 107 representative Chinese stocks spanning the period from 2017 to 2022. A systemic risk network is derived from the Risk Transmission Index based on TENET and the QR–Lasso model. By utilizing DBSCAN, HITS and community detection algorithms on the network, we aim to propose a more suitable definition of systemically important companies, explore the interrelationships between companies, and discuss its plausible reasons for dynamics structural changes. The empirical findings demonstrate a substantial involvement of insurance companies in both contributing to and receiving systemic risk within the analyzed context. We identify prominent risk output and input centers, and emphasize the profound impact of the COVID-19 pandemic on the dynamics of systemic risk.

Keywords: Chinese stock market; systemic risk; CoVaR; network analysis; community detection

check for updates

Citation: Zeng, Xiaowei, Yifan Hu, Chengjun Pan, and Yanxi Hou. 2024. Exploring Systemic Risk Dynamics in the Chinese Stock Market: A Network Analysis with Risk Transmission Index. *Risks* 12: 56. https://doi.org/10.3390/risks12030056

Academic Editor: Aivaz Kamer Ainur

Received: 17 February 2024
Revised: 12 March 2024
Accepted: 16 March 2024
Published: 20 March 2024

1. Introduction

Systemic risk, characterized by rapid transmission, close interconnections, and broad-ranging implications, is the inherent risk that exists in complex financial systems. Once triggered, it can jeopardize the stability of the entire financial system, leading to catastrophic consequences for the economy and society. For example, in recent years, one of the major shocks to the global financial system has been the outbreak of COVID-19. This pandemic has severely impacted economic activities worldwide, triggering a series of chain reactions. For businesses, it has led to supply chain disruptions and production halts, posing a significant threat to the global economic system. For investors, the substantial volatility in stock markets, bonds, and currency markets has raised their concerns about economic prospects, causing a decrease in market liquidity. Financial institutions have also faced challenges, as the interconnectivity among institutions has enabled the risk of one institution to rapidly spread throughout the entire financial system.

Systemic risk has gained widespread attention since the 2008 financial crisis, and many economists and authoritative institutions have provided specific definitions of systemic risk. Kaufman and Scott (2003) summarize the definition of systemic risk into three categories: the first category views it as a macro shock, causing significant adverse impacts on most entities in the entire system simultaneously. The other two categories perceive systemic risk as an evolving process. Specifically, the second category defines systemic risk as a causal relationship, that is, a chain reaction resulting from the close direct connections between financial institutions and markets (knock-on reaction), while the third category emphasizes the similarity of third-party risk exposures among financial institutions, attributing systemic risk to the similarity among institutions instead. Our approach aligns more with the concepts from the second and third definitions, focusing on the exploration of the connections of risk exposures among stocks to investigate systemic risk in the Chinese stock market.

Worldwide scholars also show research interest and propose various measures of systemic financial risks. Kupiec (1995) proposes one of the most commonly used measures, Value at Risk (VaR). To address its inability to identify extreme loss, Artzner et al. (1999) propose expected shortfall (ES). Based on these measures, Tobias and Brunnermeier (2016) propose Conditional Value at Risk (CoVaR) and Conditional Expected Shortfall (CoES) to quantify the conditional systemic risk, namely, the risk the whole financial system faces conditional on institution stress. By contrast, Acharya et al. (2017) propose marginal expected shortfall (MES) and systemic expected shortfall (SES) to examine a financial institution's stress conditional on systemic stress. Brownlees and Engle (2017) introduce SRISK to measure the expected capital shortfall of a financial institution conditional on a severe market decline.

When it comes to Chinese financial markets, understanding and measuring systemic risk is an important task, as the Chinese financial markets have matured significantly over the past few decades, and systemic risk is one crucial concern in its economic system. In January 2022, the People's Bank of China issued the "Macroprudential Policy Guidelines (Trial)", which, in its third chapter, provides instructive guidance on monitoring, identifying, and assessing systemic financial risks. Therefore, an abundant amount of research focuses on measuring systemic financial risk. Numerous methods are developed and applied to Chinese financial markets, such as the logit model in (Lin and Zheng 2016), the STV model in (Zhang et al. 2003), and the DCC-GARCH model in (Fang et al. 2012). On the other hand, researchers are making efforts to design innovative indicators to measure systemic risk. For example, Lai (2011) construct the China Financial Stress Index, which primarily consists of term spreads, stock market volatility, and many common and useful variables. Li et al. (2016) develop a systemic risk-monitoring indicator based on the risk interdependence structure.

Although these studies have made innovative progress on the quantitative modeling methods of systemic risk, it is hard to model systemic risk under the network structure, given consideration of the fact that there is a large number of financial institutions, each contributing to potential systemic risk more or less. Some studies address the network structure of financial risk including (Diebold and Yılmaz 2014; Hautsch et al. 2015; Härdle et al. 2016).

In terms of systemic risk index, Diebold and Yılmaz (2014) introduce various connectedness metrics derived from components of variance decomposition, asserting that the variance decomposition matrix is equivalent to a network adjacency matrix. Hautsch et al. (2015) construct a tail risk network with VaR as the weight of the edge, introducing the Lasso method to quantile regression for the identification of the relevant tail risk transmitters for each company. Härdle et al. (2016) propose a so-called TENET model, which uses a non-linear and variable selection-enhanced quantile regression method to model network risk driven by tail events. This approach not only captures the network effects of liquidity but also highlights the importance of tail event-driven inter-connectivity. The model, using a single index model, investigates non-linear factors in tail interdependence, providing meaningful explanations and a good model fit. In our research, we follow the work of Härdle et al. (2016) and develop a Risk Transmission Index based on the TENET mode. Such an index takes not only the spillover effects across financial institutions but also the market capitalization into account. Also, Härdle et al. (2016) use in-degree and out-degree measures to identify systemic important financial institutions (SIFIs) based on their network indices. We further suggest applying the HITS algorithm to gain another perspective of the network structure. In particular, the hub and authority values calculated by the HITS algorithm can assist researchers in identifying input and output risk centers.

When it comes to the network structure, Diebold and Yılmaz (2014) examine the network structures of 16 stocks, a relatively small sample size that limits the identification of the overall network structure. Härdle et al. (2016) analyze 100 stocks and categorize the market into four groups: depositories, insurers, broker–dealers, and others. Hautsch et al. (2015) focus on the 57 largest financial companies. While these studies aim to identify the systemic important companies, they place less emphasis on the network structure. The division of systemic risk community is mainly based on depositories, insurers, broker–dealers, and others. The market in China primarily consists of secondary industries, while most studies concentrate on systemic risks within the financial system. Therefore, we aim to apply research methods from social network analysis to derive an unsupervised clustering on the system risk network, and conduct a comprehensive study on the dynamics of the network structure in recent years. Specifically, we define and utilize the Risk Transmit Index as the edge weights of the network. We then employ the community detection techniques on the systemic risk network to identify five risk communities and analyze their industry composition.

Taking into account the dynamics of the network, previous studies (Diebold and Yılmaz 2014; Härdle et al. 2016) select specific time points, often during the period of the 2007–2008 financial crisis, to visualize the network structure and analyze the fluctuation of their systemic risk indices over time. In our study, we analyze systemic risks during the period of 2017–2022, which allows us to investigate whether the COVID-19 pandemic has influenced systemic risk. We find that clustering algorithms can help identify the period when COVID-19 had an impact on the stock market. Based on the clustering results, we divide the time into three periods: pre-pandemic, pandemic, and post-pandemic eras. We will further analyze the network structure and the systemic importance of companies during these three periods.

More specifically, this paper intends to address the following two questions for Chinese stock markets:

1. How can we quantify the inter-flow of systemic risk in the network of Chinese stock markets? We construct a systemic risk network for individual stocks, observe how systemic risk propagates within the stock market, explore the major risk output and input centers, and analyze the primary constituents of systemic risk in the Chinese stock market.
2. How does the major systemic risk contributors influence the systemic risk across communities and over time? We employ clustering algorithm on the risk dynamics over time to obtain small time periods, and then find out how the major risk centers behave regarding both detected communities and time periods in the system risk network we construct for the Chinese stock market.

To address the above two questions, we collect the data for 107 representative stocks in the Chinese stock market from 2017 to 2022, along with 30 macroeconomic factors representing China's economic development trends. We develop our approach based on the TENET model. We use the parameters from the single index model to construct a systemic risk network for Chinese markets, define the network-based risk indicator using the CoVaR measure, and then employ some network algorithms to detect the community and time-variant effects. Our analysis demonstrates that classical network metrics can be used for systemic risk analysis and they possess strong interpretability.

This paper proceeds as follows. We first introduce the models employed in this study in Section 2 in detail. The procedures of data preprocessing, empirical results, and analysis are described in Section 3. We conclude the findings, limitations, and further research recommendations on the systemic risk in the Chinese stock market in Section 4.

2. Materials and Methods

2.1. CoVaR Based on a Single Index Model

2.1.1. Estimation of VaR and CoVaR

In financial risk management, Value at Risk (VaR) is a widely applied risk measure that assesses and quantifies the potential loss in the value of a portfolio or investment over a specified time horizon. To take into account the risk transmission relationship among companies when extreme events take place, we can extend the measurement of VaR by examining the probability that another company experiences an extreme event when one occurs, which is the concept of Conditional Value at Risk (CoVaR). In the following paragraphs, we introduce the estimators for VaR and CoVaR.

Let X_t^i represent the loss variable for company i at time t. Given the macro state variables M_{t-1}, the VaR of X_t^i at a risk level τ, $\mathrm{VaR}_\tau(X_t^i)$, is defined as

$$\mathrm{VaR}_\tau(X_t^i) = \underset{\theta}{\operatorname{argmin}}\, \mathbb{E}\Big(\rho_\tau(X_t^i - \theta) \mid M_{t-1}\Big), \tag{1}$$

where $\Big\{\rho_\tau(u) = (\tau - I_{(u<0)})u\Big\}$ is the so-called check function. Thus, some statistical methodologies, like the quantile regression model, can be used to estimate the Value at Risk for regression data. An alternative definition of $\mathrm{VaR}_\tau(Y)$ is the τ quantile of return yields, $\Big\{\Pr\Big(X_t^i \le \mathrm{VaR}^i_{\tau,t} \mid M_{t-1}\Big) = \tau\Big\}$, which expresses VaR as a monetary amount, denoting the maximum expected loss at a specified level.

We assume that the loss follows a linear relationship with the 30 macroeconomic indicators. Since there are too many predictors in the model, it may lead to some model efficiency problems, such as a decrease in the explanatory power and the generalization ability of the model. The financial system is strongly impacted only by some of the macro state variables at certain times. So we introduce an L1-regularization (Lasso) in estimation and solve the following equation:

$$\widehat{\gamma}^i_\tau = \arg\min_{\gamma^i}\Bigg\{\frac{1}{n}\sum_{i=1}^{n}\Big[\rho_\tau\Big(X_t^i - (\gamma^i)^\top M_{t-1}\Big) + \lambda||\gamma^i||_1\Big]\Bigg\} \tag{2}$$

$$\widehat{\mathrm{VaR}}^i_{\tau,t} = \widehat{\alpha}^i + (\widehat{\gamma}^i_\tau)^\top M_{t-1}. \tag{3}$$

We find that the estimation can effectively capture the trends of VaR, and most results can pass the backtesting, which will be explained in the later section.

Given M_{t-1}, the macro state variables, and B^j_{t-1}, the company characteristics computed based on its balance sheet CoVaR of a paired variable (X_t^i, X_t^j) at risk level τ is denoted by $\mathrm{CoVaR}^{j|i}_{\tau,t}$ such that

$$\Pr\Big(X_t^j \le \mathrm{CoVaR}^{j|i}_{\tau,t} \mid X_t^i = \mathrm{VaR}^i_{\tau,t}, M_{t-1}, B^j_{t-1}\Big) = \tau. \tag{4}$$

Here, we apply TENET model in (Härdle et al. 2016) and calculate CoVaR for any paired log returns $\{(X_t^i, X_t^j)\}$ of stock i and j given the macro state variables $\{M_t\}$. The TENET model is motivated by the need to mitigate both model inefficiencies arising from an excessive number of predictors, and difficulties aroused by non-linear dependencies between pairs of financial assets. Härdle et al. (2016) apply a single index model with a SCAD penalty to address the problem. Suppose X_t^j represents the loss variable of financial institution j, following

$$X_t^j = g\Big(\beta^\top_{j|R_j} R_t^j\Big) + \epsilon_t^j, \tag{5}$$

where $\Big\{R_t^j := \{X_t^{-j}, M_{t-1}, B^j_{t-1}\}\Big\}$ represents the information set; $\Big\{X_t^{-j} := X_t^1, \ldots, X_t^{j-1}, X_t^{j+1}, \ldots, X_t^k\Big\}$ represents the explanatory variables, including the loss variable of all financial institutions other than institution j; and k denotes the number of financial institutions. $\widetilde{R}_{j,t}$ is defined as $\Big\{\widehat{\mathrm{VaR}}^{-j}_{\tau,t},\ \ M_{t-1}, B_{j,t-1}\Big\}$, where $\widehat{\mathrm{VaR}}^{-j}_{\tau,t}$ is estimated in (3) except for the j-th institution. $\widehat{\beta}_{j|\widetilde{R}_j}$ is defined as $\Big\{\widehat{\beta}_{j|-j}, \widehat{\beta}_{j|M}, \widehat{\beta}_{j|B_j}\Big\}^\top$. Then, the estimators are defined as

$$\widehat{\beta}^{j|\widetilde{R}_j} := \Big\{\widehat{\beta}^{j|-j}, \widehat{\beta}^{j|M}, \widehat{\beta}^{j|B_j}\Big\}, \tag{6}$$

which are obtained from the following equation:

$$\begin{aligned} L_n(\beta, g(\cdot)) &:= \frac{1}{n}\sum_{j=1}^{n}\sum_{t=1}^{n}\rho_\tau\left\{X_t - g\left(X_j^\top\beta\right) - g'\left(X_j^\top\beta\right)X_{tj}^\top\beta\right\}\omega_{ij}(\beta) + \sum_{l=1}^{p}P_\lambda(\beta_l) \\ \left(\widehat{\beta}, \widehat{g}(\cdot)\right) &:= \arg\min_{\beta, g(\cdot)} L_n(\beta, g(\cdot)), \end{aligned} \tag{7}$$

where $\omega_{ij}(\beta)$ represents a Gaussian kernel function, and P_λ serves as a penalty function. In this model, β not only reflects the influence of all financial institutions except j but also encapsulates the impact of macroeconomic factors and the firm itself. Its estimate is obtained through the single index model. The greater the absolute value of β, the more significant the influence that the corresponding variable will have on the tail risk of firm j. The penalty Smoothly Clipped Absolute Deviation (SCAD) is used, where $P_\lambda(x)$ is the SCAD penalty function, defined as

$$P_\lambda(x) = \begin{cases} \lambda|x| & \text{if } |x| \le \lambda \\ \frac{2a\lambda|x| - x^2 - \lambda^2}{2(a-1)} & \text{if } \lambda \le |x| \le a\lambda\,, \\ \frac{1}{2}(a+1)\lambda^2 & \text{if } |x| > a\lambda \end{cases} \tag{8}$$

where x is the coefficient being penalized, λ is a regularization parameter (often determined by cross-validation), and a is a constant (typically chosen as 3.7 in the original SCAD proposal) (Fan and Li 2001). Then, we estimate the CoVaR for company X_t^j by

$$\widehat{\text{CoVaR}}_{\tau,t}^{j|i} = \widehat{g}\left(\widehat{\beta}_{j|\widetilde{R}_j}^\top \widetilde{R}_{j,t}\right). \tag{9}$$

The above approach can force some coefficients to shrink to zero smoothly, realizing the target of variable selection. The key advantage of SCAD over the L1 penalty is that it tends to shrink coefficients more gently for larger values, leading to less bias for large coefficients. This can be beneficial in situations where some predictors have a significant impact on the response variable. Therefore, compared to other variable selection methods based on criteria like AIC and BIC, the SCAD penalty is more computationally efficient and cost-effective (Fan and Li 2001).

For the computational aspect, we use an iterative approach to treat $\widehat{\beta}_\tau$ and $\widehat{g}(\cdot)$ as a finite-dimensional optimization problem. The convergence of the solution is demonstrated in (Zhang et al. 2003). The penalty function γ_λ employs the SCAD method, and the selection of λ is determined through an improved BIC model (Duffie and Pan 1997), with validation performed using the GCV criterion. Additionally, the value of λ can reflect the average individual risk, which is quite explanatory.

2.1.2. Risk Fluid in the Network

Furthermore, the analysis of systemic risk necessitates the incorporation of not only a company's market capitalization but also its positioning within the financial network. The magnitude of a company's market capitalization serves as an indicator of its pronounced importance within the financial system. The assessment of a company's significance in the financial system extends beyond its size and encompasses its capacity to exert influence on other entities or be subject to external influences. In order to comprehensively address both dimensions, we formulate a distinct measure, the Risk Transmission Index (RTI), which contributes to the establishment of a risk network within the Chinese stock market.

Definition 1 (Risk Transmission Index)**.** *This metric quantifies the risk impact of both company i on company j and company j on company i at time t, which is defined as*

$$RTI_t^{i|j} := MC_t^i \times \left|\widehat{D}_t^{i|j}\right| \times MC_t^j, \tag{10}$$

where MC_t^i represents the market capitalization of company i at time t, and $\widehat{D}^{j|\widetilde{R}_j}$ measures the gradient that quantifies the marginal effect of covariates at $R_{j,t} = \widehat{R}_{j,t}$, which is defined as

$$\begin{aligned} \widehat{D}^{j|\widetilde{R}_j} &:= \left.\frac{\partial \widehat{g}\left(\widehat{\beta}_{j|R_j}^{\top} R_{j,t}\right)}{\partial R_{j,t}}\right|_{R_{j,t}=\widetilde{R}_{j,t}} \\ &= \widehat{g}'\left(\widehat{\beta}_{j|\widetilde{R}_j}^{\top} \widetilde{R}_{j,t}\right) \\ &= \left\{\widehat{D}^{j|-j}, \widehat{D}^{j|M}, \widehat{D}^{j|B_j}\right\}^{\top}. \end{aligned} \tag{11}$$

Specifically, $\widehat{D}^{j|-j}$ can assess the spillover effects among various financial institutions.

Now, we can construct the edges of the systemic network by the RTI between all pairs of institutes. The network is defined as a collection composed of two subsets, $\{G = (S, E)\}$, where S represents k nodes comprising k financial institutions, and E denotes the edges between different nodes. Since this paper primarily focuses on studying the mutual risk impacts among institutions, the network analysis uses the RTI at time t as the edge weight $A_t^{i|j}$. The construction of a weighted adjacency matrix of dimensions $k \times k$ at time t is articulated as follows:

$$A_t = \left(\mathrm{RTI}_t^{i|j}\right)_{k\times k} = \begin{pmatrix} 0 & \mathrm{RTI}_t^{1|2} & \mathrm{RTI}_t^{1|3} & \dots & \mathrm{RTI}_t^{1|k} \\ \mathrm{RTI}_t^{2|1} & 0 & \mathrm{RTI}_t^{2|3} & \dots & \mathrm{RTI}_t^{2|k} \\ \mathrm{RTI}_t^{3|1} & \mathrm{RTI}_t^{3|2} & 0 & \dots & \mathrm{RTI}_t^{3|k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathrm{RTI}_t^{k|1} & \mathrm{RTI}_t^{k|2} & \mathrm{RTI}_t^{k|3} & \dots & 0 \end{pmatrix}. \tag{12}$$

2.2. Systemic Risk Network Analysis

Our study employ advanced network analysis techniques to gain a comprehensive understanding of the complicated structure of the systemic risk network. We particularly focus on scrutinizing the dynamic alterations in systemic risk within the Chinese stock market, discerning variations under diverse economic conditions. The logistics of our network analysis is depicted in Figure 1.

The initial phase of our approach involves two parallel aspects of procedures—identifying input and output risk centers under the context of systemic risk network, and finding the community constituents of systemic risk. To achieve the former goal of identifying pivotal centers, we implement the Hyperlink-Induced Topic Search (HITS) algorithm. This algorithm, recognized for its efficacy in revealing influential nodes within a network, aids in pinpointing key elements in the systemic risk network. For the latter goal of detecting communities, we incorporate a community detection algorithm to unveil the underlying structures and interrelationships within the network, contributing to a more nuanced comprehension of the systemic risk landscape in the Chinese stock market. Furthermore, both HITS and community detection allow us to extend our analysis to a time-variant aspect.

Using $\sum_{i,j} \mathrm{RTI}_t^{i|j}$ as the systemic measure at the time t, we then plot the systemic risk dynamics over time from 2017 to 2022 and observe an obvious time pattern, which motivates us to employ the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm on the risk dynamics. DBSCAN is useful in an in-depth examination of the temporal patterns inherent in the systemic risk, due to its capacity to discern complex and evolving patterns, particularly in our non-linear contexts.

After dividing the whole time periods into a certain number of periods, we start to discuss the relationship between the risk enters and the communities' risk dynamics. We focus on how the most important input and output risk centers interact with one another in the communities as well as how they influence the systemic risk regarding different time periods. Through these multifaceted analytical techniques, we seek to unravel the intricate dynamics and contributing factors influencing systemic risk in this financial domain.

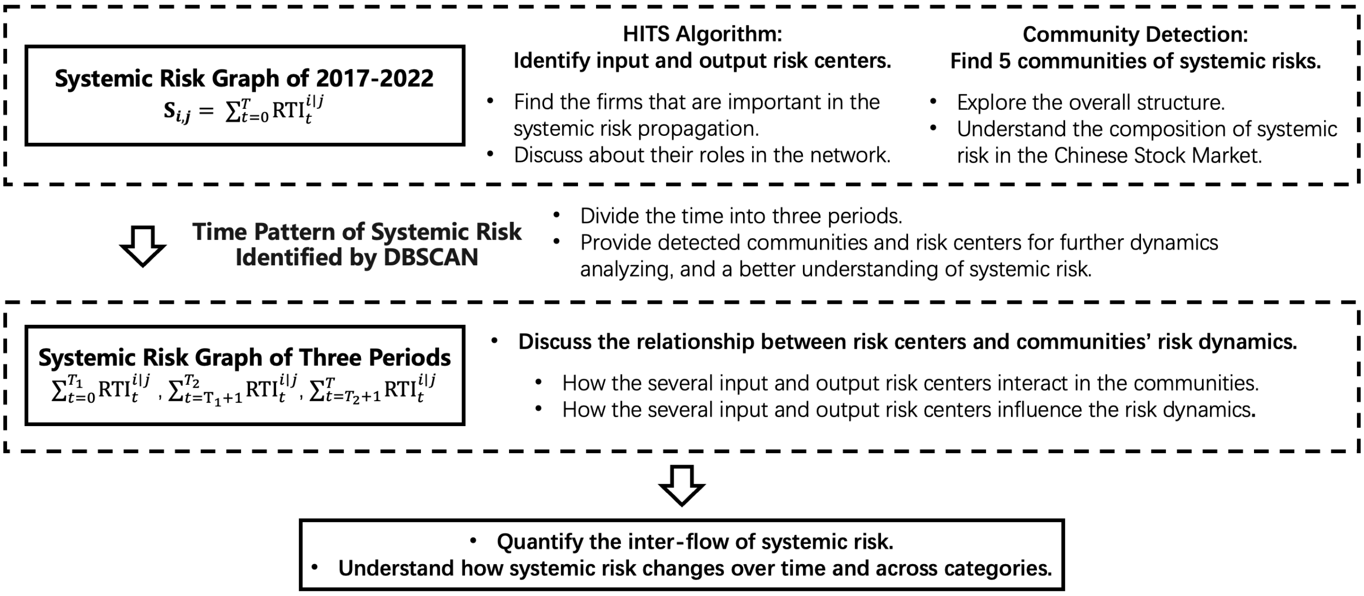


Figure 1. Diagram of the model composition for system risk network analysis.

2.2.1. Identifying Input and Output Risk Centers by HITS Algorithm

One mission of our paper is to understand how risk is transmitted in the network. One way is to discover where the systemic risk happens and where the risk will be transmitted, and then define these two kinds of nodes as output and input risk centers.

Despite the fact that the RTI we define shows the extent to which financial risk is transmitted into or out of one stock, it is insufficient to extract the input and output centers only based on the RTI. The risk associated with a particular stock's output may not solely be influenced by its direct incoming connections. Instead, it could be significantly affected by the risk transmitted from other stocks that are indirectly connected.

A similar phenomenon was first discovered by Kleinberg (1999) in the context of web networks that authoritative web pages do not stand out based on conventional network metrics such as in-degree and out-degree, and they might not use self-identifying terms on their pages. These issues pose challenges in detecting authoritative web pages. To filter authoritative pages from the huge information set on the Internet, he proposes the Hyperlink-Induced Topic Search (HITS) algorithm. HITS is a network analysis algorithm that assesses the network structure by identifying authority nodes and hub nodes, with hub nodes being pages linked to many related authorities. To be specific, in our research context, the graph S is defined by $\sum_{t=1}^{T} R_t^{i|j}$ as the in-degree (risk input) of company j from company i and out-degree (risk output) of company i to company j. Then, we calculate the hub and authority values according to Algorithm 1.

Therefore, the results of HITS can capture the nature of systemic risk and the centrality of stocks in the network. In this context, authority nodes represent the input centers of systemic risk, while hub nodes represent the output centers of systemic risk in the stock market. These input and output centers may have practical usages in macro-control and may help researchers to grasp how risk is transmitted in the network.

Algorithm 1 HITS

1: Sort the in-degrees of all the stocks decreasingly and construct a focused subgraph $S_\sigma \subset S$ with the top t of them.
2: Set a specific value d.
3: **while** not converge **do**
  for each stock $s \in S_\sigma$ **do**
    Add all the stocks to which stock s points to S_σ.
    if the in-degree of s is smaller than d **then**
      Add all the stocks that point to stock s to S_σ.
    else
      Add d stocks that point to stock s randomly to S_σ.
    Update the authority weight $a^{(s)}$ and the hub weight $h^{(s)}$.

$$a^{(s)} \leftarrow \sum_{p:(p,s)\in E} h^{(p)}$$

$$h^{(s)} \leftarrow \sum_{p:(s,p)\in E} a^{(p)}$$

    Normalize the weights so that $\sum_{s\in S_\sigma}(a^{(s)})^2 = 1, \sum_{s\in S_\sigma}(h^{(s)})^2 = 1$.
4: The nodes with relatively large weights are labeled as the authority nodes and hub nodes.

2.2.2. Identify Constituents of Systemic Risk by Community Detection Algorithm

In the examination of systemic risk within the stock market, our objective is to categorize stocks into distinct communities based on their risk transmission indices. Subsequently, an analysis of the industry composition and correlated features within these systemic risk communities will be conducted. It is hypothesized that within the same community, risks tend to be frequently transmitted, while transmission between different communities is comparatively infrequent.

The delineation of these communities serves the purpose of facilitating a comprehensive understanding of the interrelationships among various stocks. This, in turn, enables researchers and governmental entities to implement more nuanced and effective policies in the realm of macroeconomic regulation. To achieve this, we employ the modularity optimization algorithm as presented in the work by Blondel et al. (2008) and leverage community detection algorithms to the systemic risk network. The modularity of a network is defined as

$$Q = \frac{1}{2m}\sum_{i,j}\left[w_{ij} - \frac{k_i k_j}{2m}\right]\mathbf{1}\{\text{node } i, j \text{ are in the same community}\}, \tag{13}$$

where w_{ij} represents the edge weight from node i to node j, $k_i = \sum_j w_{ij}$, $m = \frac{1}{2}\sum_{i,j} w_{ij}$, $\mathbf{1}$ is the indicator function. If we think of a community as a node, then w_{ij} is the sum of edge weights from the nodes in community i to the nodes in community j.

This algorithm can partition a network into communities characterized by densely interconnected nodes, with correspondingly sparse connections between nodes belonging to distinct communities (Algorithm 2). In contrast to alternative community detection algorithms, our chosen algorithm strikes a judicious balance between computational efficiency and accuracy, thus expeditiously identifying highly modular partitions in extensive networks. Furthermore, it elucidates a comprehensive hierarchical community structure, contributing to a more nuanced comprehension of systemic risk dynamics within the stock market.

Algorithm 2 Community detection.

1: **Define**
- b: the number of iterations, $b = 1, \ldots, B$.
- B: the number of iterations after global convergence.
- n_b: the current number of communities in the bth iteration.
- $c_i^{(b)}$: the ith node in the bth iteration.
- $C_i^{(b)}$: the ith community in the bth iteration.
- $\mathcal{C}^{(b)}$: the set of communities in the bth iteration.
- $\Delta Q_{im}^{(b)}$: gain in modularity when placing $c_i^{(b)}$ into $C_m^{(b)}$.

2: Assign a different community $C_i^{(0)}$ for each stock $s_i \in S, i = 1, \ldots, k$.
3: Let $\mathcal{C}^{(0)} = \{C_1^{(0)}, \ldots, C_k^{(0)}\}$.
4: **while** $\mathcal{C}^{(b+1)} \neq \mathcal{C}^{(b)}$ **do**
  Treat the current communities $C_i^{(b)} \in \mathcal{C}^{(b)}$ as nodes $c_i^{(b)}$, $i = 1, \ldots, n_b$, where
  while $\exists\, i$ s.t. $C_i^{(b)}$ changes **do**
    for each node $c_i^{(b)}$ **do**
      Calculate $\Delta Q_{im}^{(b)}$'s, $m \in \{m : c_j^{(b)} \in C_m^{(b)}$ and $w_{c_i c_j}^{(b)} \neq 0\}$.
      if $\Delta Q_{im}^{(b)} \leq 0, \forall m$ **then**
        Keep $c_i^{(b)} \in C_i^{(b)}$.
      else
        Remove $c_i^{(b)}$ from $C_i^{(b)}$ and let $c_i^{(b)} \in C_{m^*}^{(b)}$, $m^* = \arg\max_m \{\Delta Q_{im}^{(b)}\}$.
  Let $\mathcal{C}^{(b+1)} = \{C_i^{(b)} : C_i^{(b)} \neq \emptyset\}$.
5: **Return** $\mathcal{C}^{(B)}$.

2.2.3. Identify Time Pattern of Systemic Risk by DBSCAN Algorithm

In order to comprehend the dynamic nature of systemic risk, an evaluation is conducted for each time instance, denoted as t, wherein an aggregate measure for the overall systemic risk is derived by summing the values of $A_t^{i|j}$, expressed as $\sum_{i \neq j} \mathrm{RTI}_t^{i|j}$. To delineate temporal patterns and inherent structures within the evolving systemic risk landscape, we employ the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm as introduced by Ester et al. (1996). This algorithm, widely recognized in the fields of data mining and machine learning, proves advantageous due to its capacity for density-based clustering (Algorithm 3).

Algorithm 3 DBSCAN

1: **Define**
- ε: the maximum distance that a point can be from another point to be considered as a neighbor.
- MinPts: the minimum number of data points required to form a dense region.

2: **for** each data point in the dataset **do**
  Count the number of points within the ε-neighborhood of that point.
  if the count is greater than or equal to MinPts **do**
    Mark the point as a core point, which is the central point in dense regions.
3: **for** each core point **do**
  Create a cluster and include all reachable points within the ε-neighborhood.
4: Assign border points to the cluster, where border points represent the data points that are within ε distance of a core point.
5: Identify noise points that are not core points or border points. These noise points are treated as outliers in this study.

Given the intricate and non-linear nature of the temporal evolution of systemic risk, DBSCAN is particularly suitable for our analysis. It excels in grouping closely packed data points within high-density regions, while designating points situated in low-density areas as noise. As a result of applying DBSCAN, the time series is effectively clustered into distinct temporal periods, allowing for an in-depth analysis of the underlying factors contributing to these discernible patterns.

To summarize, in order to capture the non-linear relationship and some potential interaction effects within the data, we adopt the TENET model in (Härdle et al. 2016). Combining this approach with the QR–Lasso model which possesses the variable selection tool, we employ the single index model to explore systemic risk in the Chinese stock market. Subsequently, we define a risk transmission index to measure the mutual risk impact among companies. We utilize this index as edge weights to construct a systemic financial risk network for the Chinese stock markets.

In addition, apart from using the fundamental network metrics like degree centrality to explore the network characteristics, we employ some advanced network analysis methods to delve into the systemic risk pattern. On one hand, we employ the Hyperlink-Induced Topic Search (HITS) algorithm in (Kleinberg 1999) to analyze risk output and input centers in the Chinese stock market and to study the propagation patterns within the systemic risk network. On the other hand, we employ the community detection algorithm in (Blondel et al. 2008) to decompose the network structure of the systemic risk for the Chinese stock markets and conclude some practical interpretations. Then, based on the visualization of systemic risk measurements obtained from the single index model, we apply the DBSCAN clustering algorithm to group all moments from 2017 to 2022 into three time segments, indicating three different risk periods in practice. Finally, we analyze the interaction between the input and output risk centers across communities and explore how these risk contributors influence the systemic risk in different time periods by taking the actual economic context of China into account.

3. Results

3.1. Sample and Data

3.1.1. Data Description

We obtain data from the RESSET database. Based on the Shanghai Composite Index, we select 107 representative stocks in the Chinese stock market spanning from 2012 to 2022, 30 macroeconomic factors representing economic development trends in China, 6 factors from the three-factor model, and 5 indicators from the balance sheet.

The Shanghai 180 Index (SZ180) is a stock index introduced by the Shanghai Stock Exchange, comprising the top 180 stocks listed on the Shanghai Stock Exchange's main board. These stocks represent a substantial market capitalization, good liquidity, and strong industry representation as depicted in Figure 2a,b. They span various sectors, including manufacturing, finance, construction, etc. In this study, we exclude 73 companies from this initial list, primarily consisting of newer financial institutions with limited data and companies listed on the Science and Technology Innovation Board (STAR Market). Therefore, the selected stocks in our study account for approximately 26.21% of the total market capitalization, which does not reach 50% as depicted in Figure 2c.

Nevertheless, the stocks we select still possess good representativeness for the following reasons:

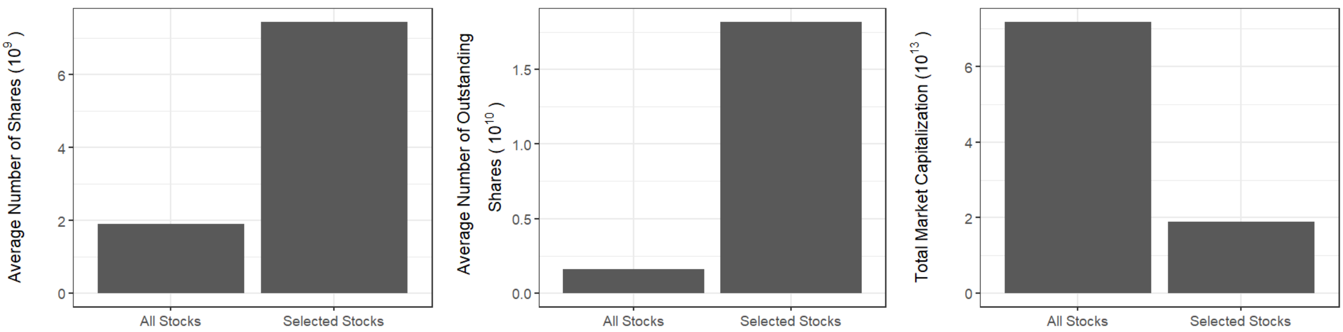
1. The "Pareto Principle" or the "80/20 Rule" is evident in the Chinese stock market. This principle, first proposed by the renowned Italian economist Vilfredo Pareto in 1897, states that 20% of the population holds 80% of the wealth. In the context of the stock market, this translates to the fact that 20% of the stocks tend to be profitable in the long term, while the remaining 80% often incur losses (Wu et al. 2010). This suggests that using a smaller subset of stocks to represent the overall market is reasonable. By calculating the long-term returns for each stock based on the annual average closing

prices in 2012 and 2022, we observe that approximately 73% of the stocks in the market have zero or negative returns, which aligns with the "80/20 Rule".

2. The industry distribution of the selected stocks is similar to that of the broader market. As depicted in Figure 3a,b, the industry distribution of the 107 stocks chosen in this study closely mirrors the industry distribution of the overall market, effectively reflecting the operational characteristics of the securities market.
3. The selected stocks in our study exhibit significant market capitalization and liquidity. As shown in Figure 2a,b, using the average number of shares per company as a measure of market size, the selected stocks in this study have an average number of shares that is 3.91 times higher than the market average. Additionally, using the latest available data on the number of outstanding shares as a measure of liquidity, the average number of outstanding shares for the selected stocks in this study is 11.42 times higher than the market average.

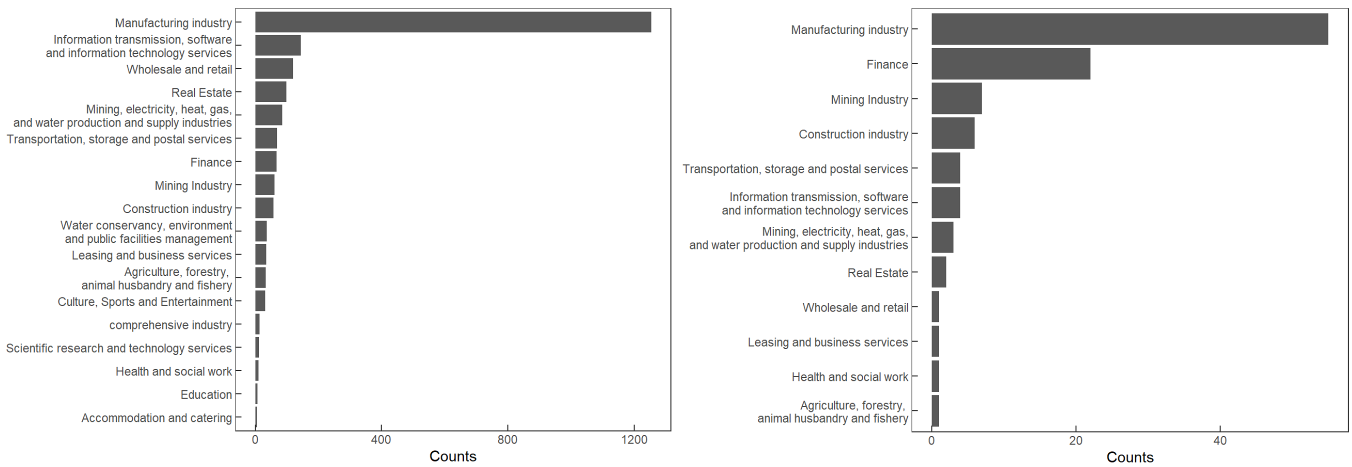
We do not choose the Shanghai 50, CSI 300, or STAR 50 for the following reasons:

- Shanghai 50: The selection criteria for Shanghai 50 are essentially the same as those for Shanghai 180, but it comprises only 50 stocks, which provides a less comprehensive representation of the stock market due to its smaller sample size.
- CSI 300: Within the CSI 300, 30% of the stocks come from the financial industry, which does not align well with the industry distribution of the overall market.
- STAR 50: The STAR 50 index consists of the 50 largest STAR Market-listed companies, which are relatively newer and may not provide sufficient data for analysis due to their recent establishment.



(**a**) (**b**) (**c**)

Figure 2. Comparison of three criteria between the whole market and selected stocks, including (**a**) market size, (**b**) stock liquidity, and (**c**) market capitalization.



(**a**) (**b**)

Figure 3. Comparison of industry distribution between (**a**) broader market, and (**b**) selected stocks.

3.1.2. Macroeconomic Indicators

We select a total of 30 macroeconomic indicators with distinctive characteristics related to the Chinese economy, covering daily, monthly, and quarterly time periods. These macro-state variables will be applied in the CoVaR and the TENET model. We categorize these indicators based on their reflection of economic conditions as illustrated in Table 1.

Table 1. Macroeconomic indicators.

Macroeconomic Indicators	Frequency	Category	Stationarity Transformation
Export Value (Current value)	Monthly	Foreign Trade Indicators	Log 1st Difference
Import Value (Current value)	Monthly		Log 1st Difference
Real Estate Development Composite Prosperity Index (Current value)	Monthly	Real Estate Related Indicators	1st Difference
Fixed Asset Investment Completed Value (Accumulated YoY)	Monthly		1st Difference
Real Estate Development Investment Completed Value (Accumulated value)	Monthly		12th Difference
Total Retail Sales of Consumer Goods (Current YoY)	Monthly	Consumer Related Indicators	1st Difference
Consumer Price Index (CPI) (Current YoY)	Monthly		None
Consumer Confidence Index (Current value)	Monthly		1st Difference
Per Capita Disposable Income of Urban Residents (Accumulated value)	Quarterly		4th Difference
Per Capita Consumption Expenditure of Urban Residents (Accumulated value)	Quarterly		4th Difference
Value Added of Wholesale and Retail Trade (Current YoY)	Quarterly		4th Difference
Electricity Generation Output (Accumulated value)	Monthly	Energy Logistics Related Indicators	12th Difference
Total Freight Volume (Accumulated value)	Monthly		12th Difference and 1st Difference
Railway Freight Volume (Accumulated YoY)	Monthly		1st Difference
Purchasing Price Indices of Raw Material (PPIRM) (Current value)	Monthly	Commodity Related Indicators	1st Difference
Retail Price Index (RPI) (Current YoY)	Monthly		1st Difference
Producer Price Index (PPI) (Current YoY)	Monthly		1st Difference
Corporate Goods Price Index (CGPI) (Current YoY)	Monthly		1st Difference
China Commodity Price Index (Current value)	Monthly		1st Difference
Money & Quasi-money(M2)	Monthly		1st Difference
Interbank Repo Benchmark Interest Rate (Current value)	Daily	Financial Market Indicators	1st Difference
Weighted Average Overnight Interbank Borrowing Rate (Current value)	Monthly		1st Difference
China Government Bond Yield (10-year)	Monthly		1st Difference
Total Outstanding Loans of Financial Institutions (Domestic and Foreign Currency) (Current YoY)	Monthly		1st Difference
Total Social Financing Scale (Stock) (Current YoY)	Monthly		1st Difference
Keqiang Index (Accumulated value)	Monthly	Macroeconomic Overall Indicators	1st Difference
Gross Domestic Product (GDP) (Current value)	Quarterly		Log 1st Difference
Value Added of the Primary Industry (Current value)	Quarterly		4th Difference
Value Added of the Secondary Industry (Current value)	Quarterly		4th Difference
Value Added of the Tertiary Industry (Current value)	Quarterly		4th Difference

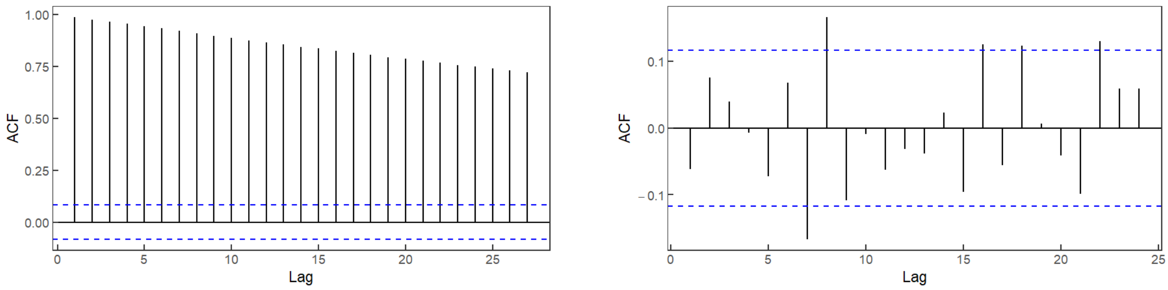
Below, we provide a brief introduction to these seven categories of macroeconomic indicators:

1. Foreign Trade Indicators reflect the foreign trade status of a country or region, the level of foreign trade activities, and international competitiveness.
2. Real Estate-Related Indicators reflect the activities and investment conditions in the real estate market.
3. Consumer-Related Indicators reflect the consumption behavior and capacity of residents, indicating the strength of economic consumption activities and changes in consumer confidence.
4. Energy Logistics-Related Indicators reflect the activity level and demand situation in the financial market's energy and logistics sectors.
5. Commodity-Related Indicators reflect the supply and demand relationships, cost pressures, and market price fluctuations of commodities. They hold significant reference value for economic analysis and decision-making.
6. Financial Market Indicators reflect the operational status of the financial market and interest rate levels.
7. Macroeconomic Overall Indicators reflect the overall economic scale and growth conditions of a country or region, indicating the overall economic development and the relative contributions of various industries.

3.1.3. Data Preprocessing

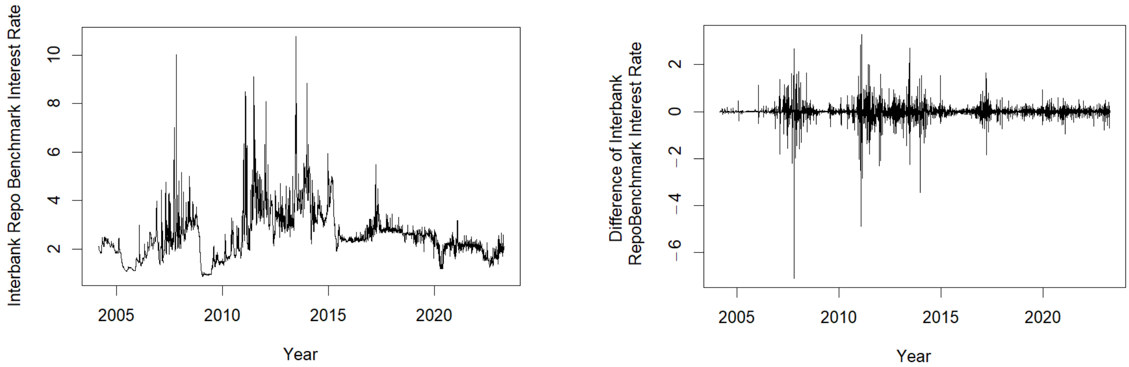
We examine the stationarity of the data and applied stationarity transformations based on autocorrelation function (ACF) plots, partial autocorrelation function (PACF) plots, as well as relevant literature and empirical knowledge. In this study, we use the augmented Dickey–Fuller (ADF) test to test the stationarity of the time series data, with the null hypothesis being that the series is non-stationary and the alternative hypothesis being that the series is stationary. The ADF tests are conducted on the following data at a significance level of 0.05.

1. Unprocessed Stock Price Data. Before any transformation, the ADF tests show that only five stocks' price series out of all the stocks are stationary. Subsequently, we take the logarithm of the price data and then differentiate it. The ADF tests are performed again, and the results indicate that, after the stationarity transformation, all stocks reject the non-stationary null hypothesis, thus rendering the price series stationary. Consequently, we utilize the rolling forecast to obtain biweekly stock price data with the rolling window being 5 years. We illustrate the transformation progress with the ACF plot of Shanghai Pudong Development Bank (C600000, Shanghai, China) before and after stationarity transformation in Figure 4.
2. Macroeconomic Indicators. Among all the daily and monthly data, only the monthly CPI data pass the ADF test. By observing ACF and PACF plots and relying on empirical knowledge, we determine the appropriate transformation methods for each indicator. After applying these transformations, all data become stationary.
   Regarding quarterly data, due to the large time intervals and relatively limited data spanning only 10 years with 45 data points, the ADF test alone cannot effectively confirm the stationarity of the data. Additionally, there is significant volatility in the quarterly data during the period from 2020 to 2022, which we believe can influence the overall stationarity of the data. Therefore, we employ empirical rules and refer to ACF and PACF plots as well as the reduction in p-values from the ADF tests to decide on the data transformation methods, which partially improve data stationarity. The specific methods used for stationarity transformation of the macroeconomic data are detailed in Table 1. Moreover, we illustrate such effect of 1st difference for Interbank Repo Benchmark Interest Rate in Figure 5.
   Finally, we utilize cubic spline interpolation to adjust all indicators to a biweekly time cycle.
3. **Three-factor Model and Balance Sheet.** The ADF tests show that all the data are stationary at a significance level of 0.05, and no further stationarity processing is required.



(**a**) (**b**)

Figure 4. ACF plots for stock C600000 : (**a**) before stationarity transformation; (**b**) after log 1st difference. If the calculated auto-correlation value exceeds the dash line, it indicates the presence of auto-correlation in the data



(**a**) (**b**)

Figure 5. The effect of 1st difference for Interbank Repo Benchmark Interest Rate. (**a**) Interbank Repo Benchmark Interest Rate; (**b**) the 1st difference for Interbank Repo Benchmark Interest Rate.

It is worth noting that only the preprocessed data from 2017 to 2022 are used for subsequent modeling.

3.2. Systemic Risk Network Analysis of the Chinese Stock Market

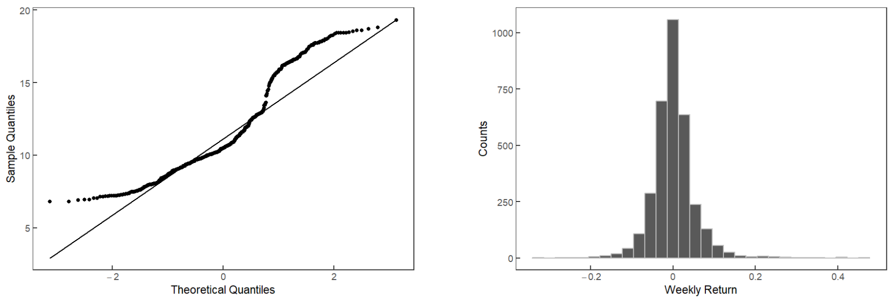
3.2.1. Feasibility Testing of QR–Lasso Model

In this study, we utilize the aforementioned 107 representative stocks in the Chinese stock market to analyze how the individual stock's Value at Risk (VaR) for biweekly return rates at a 0.1 risk level is influenced by macroeconomic factors during the period from 2017 to 2022.

We assume that different macroeconomic factors have varying degrees of influence on stock market risk under different economic conditions. In this model, at time t, we employ macroeconomic data from the period from $t - 14$ to t to regress against contemporaneous return rates, constructing quantile regression models with Lasso regularization for each t.

QR–Lasso is feasible and should be used in this context for the following reasons: First, the stock market return rates exhibit a peaked and heavy-tailed distribution. To assess the normality of the stock market return rate data, the Shapiro–Wilk test is selected, which is considered suitable for small sample sizes ($N = 563 < 5000$). The null hypothesis H_0 assumes that the data follow a normal distribution, while the alternative hypothesis H_1 assumes that the data do not originate from a normal distribution.

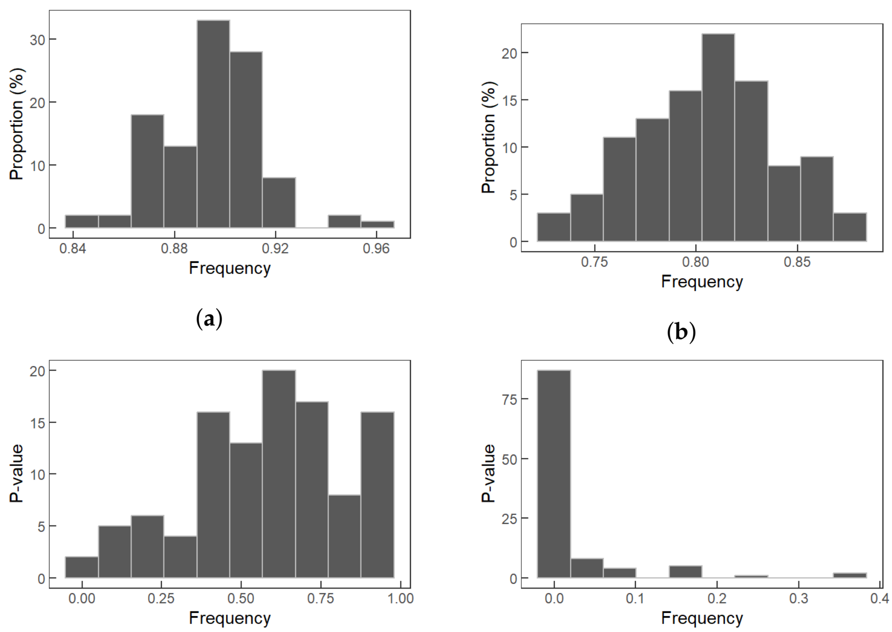
The results indicate that at a significance level of 0.05, all stocks reject the assumption of normality. This rejection reflects the riskiness of stocks and underscores the necessity of this study. For example, a QQ plot for stock C600000 is generated (Figure 6a), showing that the tails of the sequence are consistently above the normal quantiles. The data points do not form a straight line, providing a more intuitive illustration of the heavy-tailed nature of the stock market return rate distribution. Furthermore, histograms of weekly return rates for stocks from different industries are plotted. Figure 6b is an example of the financial industry. It is observed that all industries exhibit a pronounced peaked and heavy-tailed distribution, emphasizing the existence of risk in the stock market.



(**a**) (**b**)

Figure 6. Heavy tail phenomenon in stock market, (**a**) QQ plot for stock C600000; (**b**) distribution of weekly return in the financial industry.

Secondly, we apply the backtesting method to determine whether the VaR predictions align with actual portfolio losses. The ideal scenario is the observation that from 2017 to 2022, 95% of the return rates are higher than the model's VaR estimates. Additionally, Kupiec (1995) proposes a hypothesis testing method to assess whether the model conforms to its assumptions. Based on the results in Figure 7, the model proposed in this study can generally pass the backtesting, indicating that it can effectively reflect the risk characteristics of the Chinese stock market.



(**c**) (**d**)

Figure 7. Backtesting results of QR–Lasso model and quantile regression model, including (**a**) proportion results of QR–Lasso model; (**b**) proportion results of quantile regression model; (**c**) *p*-value results of QR–Lasso model; (**d**) *p*-value results of quantile regression model.

Third, the QR–Lasso model yields superior results. Comparing the QR–Lasso model with the traditional quantile regression model as shown in Figure 7 and Table 2, it is evident that the traditional model achieves approximately 85% of the return rates, exceeding the estimates. The hypothesis testing results reveal that the *p*-values are concentrated around 0, suggesting that the linear model does not align with the risk characteristics of the Chinese stock market, while the QR–Lasso model performs well.

Table 2. Proportion of exceeding estimated VaR and passing backtesting for the QR–Lasso model and quantile regression model.

Model	25% Quantile of Proportion Exceeding Estimated VaR	75% Quantile of Proportion Exceeding Estimated VaR	Proportion Passing Backtesting (0.05)
QR–Lasso Model	0.88	0.91	98.1%
Quantile Regression Model	0.78	0.82	14.0%

3.2.2. Feasibility Testing of Systemic Risk Network

It is well known that degree, in-degree, out-degree, closeness centrality, betweenness centrality, and similar metrics are commonly used to measure network structures (Table 3). The distribution of degrees can be employed to examine whether the network follows a power-law distribution, validating the feasibility of modeling. In-degree and out-degree respectively represent the total amount of systemic risk received and transmitted by individual stocks. Closeness centrality reflects whether a stock has a dominant role in the CoVaR of other stocks. Betweenness centrality measures the probability that a node lies on the shortest path between any two other points, which can be interpreted as a stock's ability to propagate risk in a systemic risk network.

Table 3. Network metrics and their interpretations.

Indicator	Definition	Interpretation
Degree	$\sum_{j\neq i} MC_{i,t} \times \left|\widehat{D}^t_{i\|j}\right| \times MC_{j,t} + \sum_{j\neq i} MC_{j,t} \times \left|\widehat{D}^t_{i\|j}\right| \times MC_{i,t}$	Judge if the network follows a power-law distribution
In-degree	$\sum_{j\neq i} MC_{i,t} \times \left|\widehat{D}^t_{i\|j}\right| \times MC_{j,t}$	Total systemic risk transmitted by individual stocks
Out-degree	$\sum_{j\neq i} MC_{j,t} \times \left|\widehat{D}^t_{i\|j}\right| \times MC_{i,t}$	Total systemic risk received by individual stocks
Closeness Centrality	$C_B(i) = 1/\sum_y d(i,j)$, where $d(i,j)$ is the shortest path from i to j	Indicates whether a stock has a dominant role in other stocks' CoVaR
Betweenness Centrality	$g(v) = \sum_{s\neq v\neq t} \sigma_{st}(i)/\sigma_{st}$, where $\sigma_{st}(i)$ calculates if i lies on the shortest path between s and v	Measures a stock's ability to propagate risk within the system

We test the fundamental assumptions of the network model based on the five metrics mentioned above. Figure 8 shows that the degree distribution broadly follows a power-law distribution, which aligns with the assumption of the Matthew effect—that is, in the stock market, most of the risk is concentrated in a few stocks.

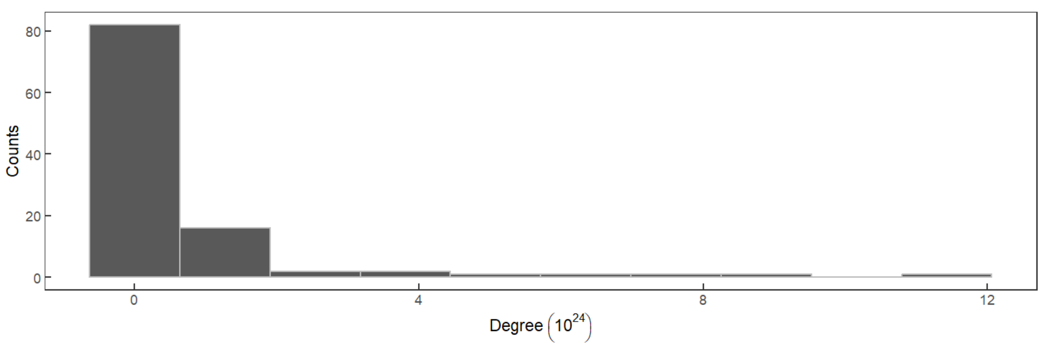
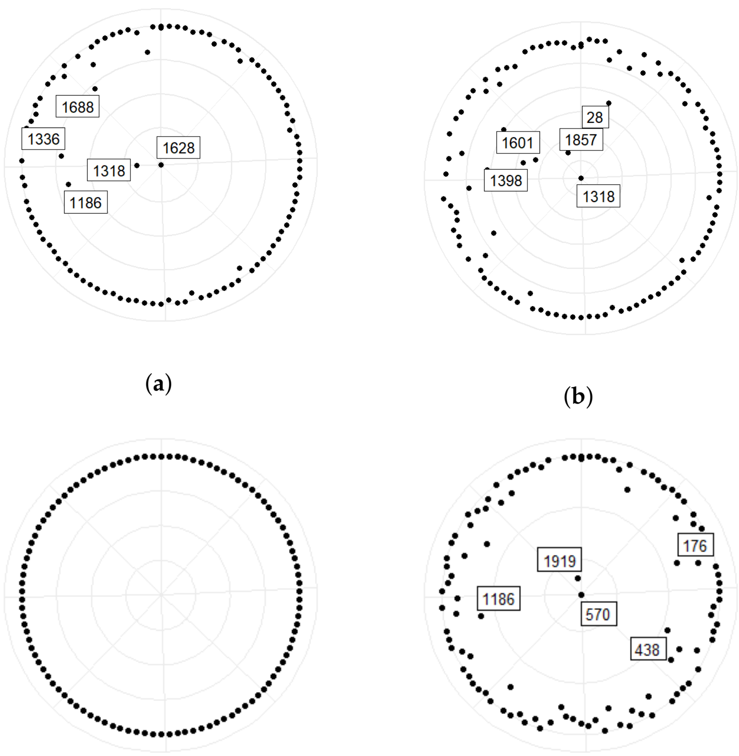


Figure 8. Distribution of degree.

Then, we visualize four important network indices in the graph, where we denote each node by its stock code modded by 600,000. Notably, financial and insurance companies like Ping An Insurance (PingAn, 1318, Shenzhen, China) and China Life Insurance Company Ltd. (CLIC, 1628, Beijing, China) are significant risk contributors (Figure 9a), while insurance companies like PingAn and Industrial companies like PetroChina Company Ltd. (CNPC, 1857, Beijing, China) are major receivers of systemic risk in the Chinese network (Figure 9b). From the analysis of closeness centrality (Figure 9c), it is evident that the closeness centrality for all nodes in the stock market is extremely small ($<1 \times 10^{-20}$). This suggests that no single stock significantly influences another stock. However, when examining betweenness centrality (Figure 9d), companies like Hang Seng Electronics (00570, Hangzhou, China) and Tongwei Company Ltd. (438, Chengdu, China). play a prominent role as risk transmitters within the stock market.

In summary, our model aligns with various network analysis metrics and can effectively apply the paradigm of network analysis for further examination.



(**a**) (**b**) (**c**) (**d**)

Figure 9. Four network metrics, including (**a**) out-degree, (**b**) in-degree, (**c**) closeness centrality and (**d**) betweenness centrality.

3.2.3. Primary Output and Input Centers of Systemic Risk

Merely analyzing degree-related metrics does not provide enough insights into the output and input centers of the stock market's systemic risk network. This is because the risk output by a stock may not necessarily originate from the stock itself but could be transmitted from another stock. The previous analysis of closeness centrality and betweenness centrality also indicates that there are no stocks with strong influences on other stocks, but there are stocks that play significant roles in risk transmission within the network. This suggests the importance of considering risk transmission within the network structure. In the graph, we denote each node by its stock code modded by 600,000.

As shown in Figure 10, risk output centers are PingAn, China Pacific Insurance Company Ltd. (CPIC, 1601, Shanghai, China), and Xinhua Insurance (NCI, 1336, Beijing, China), with the hub as 0.81, 0.55, and 0.21, respectively. Regarding the risk input centers, the two biggest risk input centers are CLIC and NCI, with authority as 0.99 and 0.08. Insurance industries in China seem to produce risk as well as receiving risk at the same time.

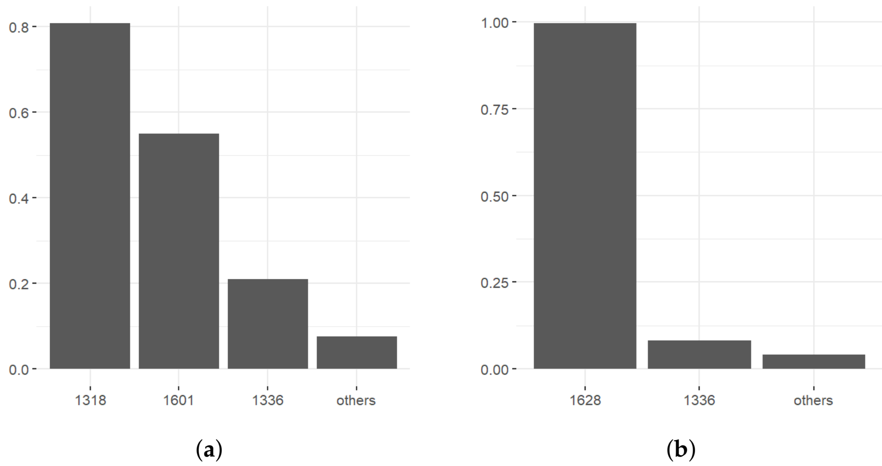


Figure 10. Stocks with top hub values and top authority values: (**a**) risk output centers; (**b**) risk input centers. The biggest three risk output centers are PingAn, CPIC and NCI. The biggest two risk input centers are CLIC and NCI.

3.2.4. Industry Composition of Systemic Risk

In network research, it is common to categorize nodes to create tightly connected groups within each group while keeping connections sparse between groups, and to conduct research on each group of nodes. A similar approach can be applied to the study of systemic risk in the Chinese stock market network. In this study, a community detection algorithm is employed to partition the systemic risk in the Chinese stock market into five major categories, using modularity as the evaluation metric for the community detection algorithm. A higher modularity indicates tighter connections within groups and sparser connections between groups. In this paper, the modularity obtained from the community detection algorithm is 0.477, falling within the normal range $[0.3, 0.7]$, indicating that the algorithm has produced a favorable community partition.

We analyze the characteristics of systemic risk in each community through risk output centers, risk acceptance centers, and industry structures. Regarding the industry composition as shown in Figure 11, we have the following:

1. Communities 1, 2, and 3 exhibit a diverse composition, encompassing emerging industries such as fin-tech, biopharmaceuticals, optical fiber, and emerging manufacturing sectors. Community 2 comprises automotive manufacturing entities, exemplified by corporations such as SAIC Motor Group (104, Shanghai, China) and Fuyao Glass Industry Group Company Ltd (660, Fuzhou, China) . Conversely, Community 3 is primarily characterized by healthcare enterprises, notably encompassing pharmaceutical companies such as Hengrui Pharmaceutical (276, Lianyungang, China) and Huahai Pharmaceutical (521, Taizhou, China) .
2. Community 4 is primarily led by the insurance industry, including PingAn, CPIC and NCI.
3. Community 5 is primarily dominated by real estate, transportation, construction, and manufacturing sectors, with some closely associated entities in the banking industry. This sector includes enterprises such as the China Railway Construction Corporation 1186, Beijing, China) , China Shipbuilding Industry Company Limited, and Industrial and Commercial Bank of China (ICBC, 1398, Beijing, China).

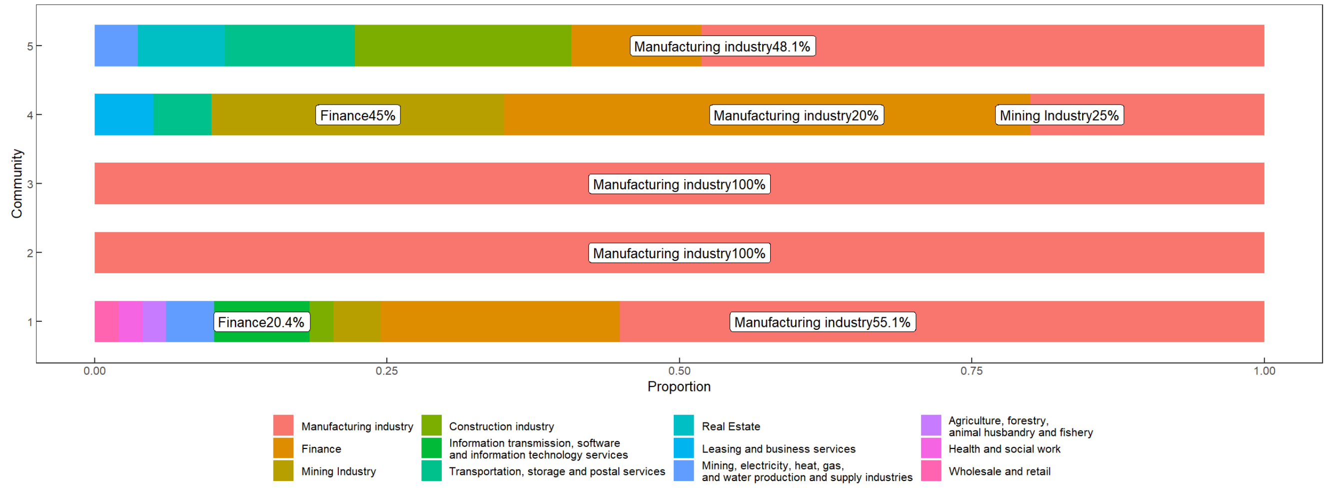


Figure 11. Industry composition of each community.

3.2.5. Systemic Risk over Time

To further study the characteristics of systemic risk in the Chinese stock market under different temporal and economic backgrounds, we cluster the changing risk of the Chinese stock market into three categories based on systemic risk over time. In this analysis, we use the RTI as a measure of systemic risk at each point in time and apply the DBSCAN clustering method.

The temporal clustering results are shown in Figure 12. The systemic risk in the Chinese stock market remains relatively low before December 2019 and after June 2021. During the first wave of the COVID-19 pandemic, disruptions in the supply chain and production halts in some industries due to pandemic control measures lead to heightened market tension, resulting in a sharp increase in systemic risk within the stock market.

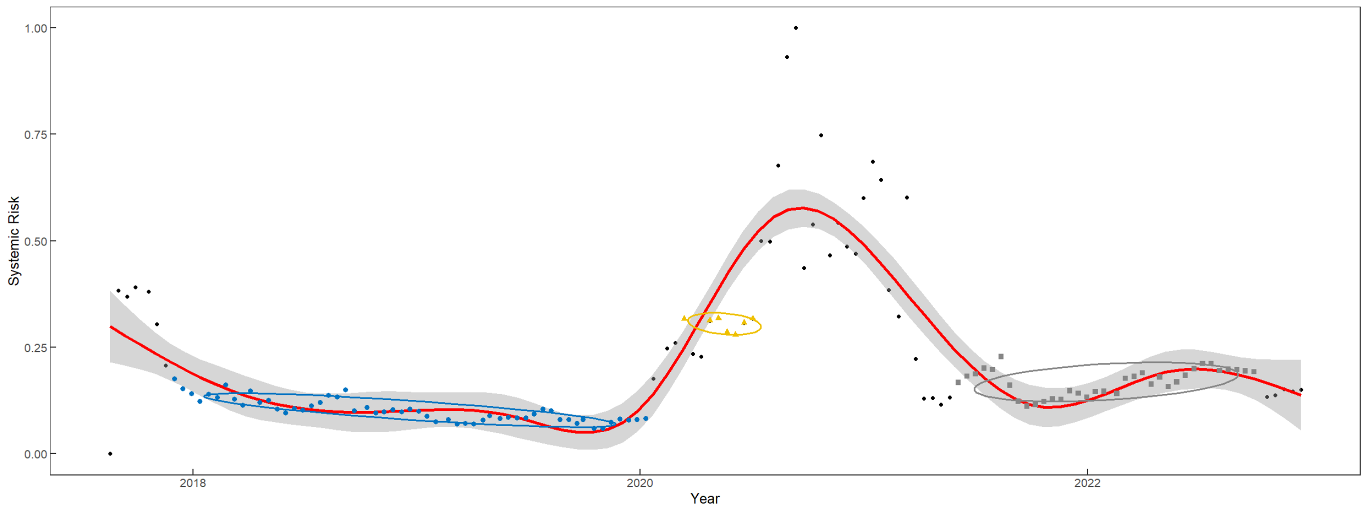


Figure 12. Clustering results of systemic risk dynamics in the Chinese stock market over time. Color blue indicates the time before the beginning of pandemic. Color grey refers to the period after the first strike of pandemic on the stock market. Black and Yellow points in the middle indicates the period when the pandemic is undergoing.

During the COVID-19 pandemic, the reduced labor mobility lead to the labor shortages. So the supply-side industrial chains are blocked. Because of the industrial clustering phenomenon, the local epidemic situation could affect the operation of the entire industry chain nationwide. The shrinkage of offline businesses, along with residents' fear towards the epidemic, lead to a decrease in consumer willingness. Furthermore, the development of the COVID-19 pandemic influences the volatility and trading activity of the domestic stock market. Empirical research (Lin 2023) has found a positive relation between the the daily growth rates of COVID-19 patients and the volatility of the domestic stock market. These impacts are particularly significant at the beginning of the epidemic outbreak. However, effective control measures implemented to manage the pandemic contribute to a return to normal systemic risk levels in 2021.

3.2.6. Dynamics of Systemic Risk Structures in Chinese Stock Market

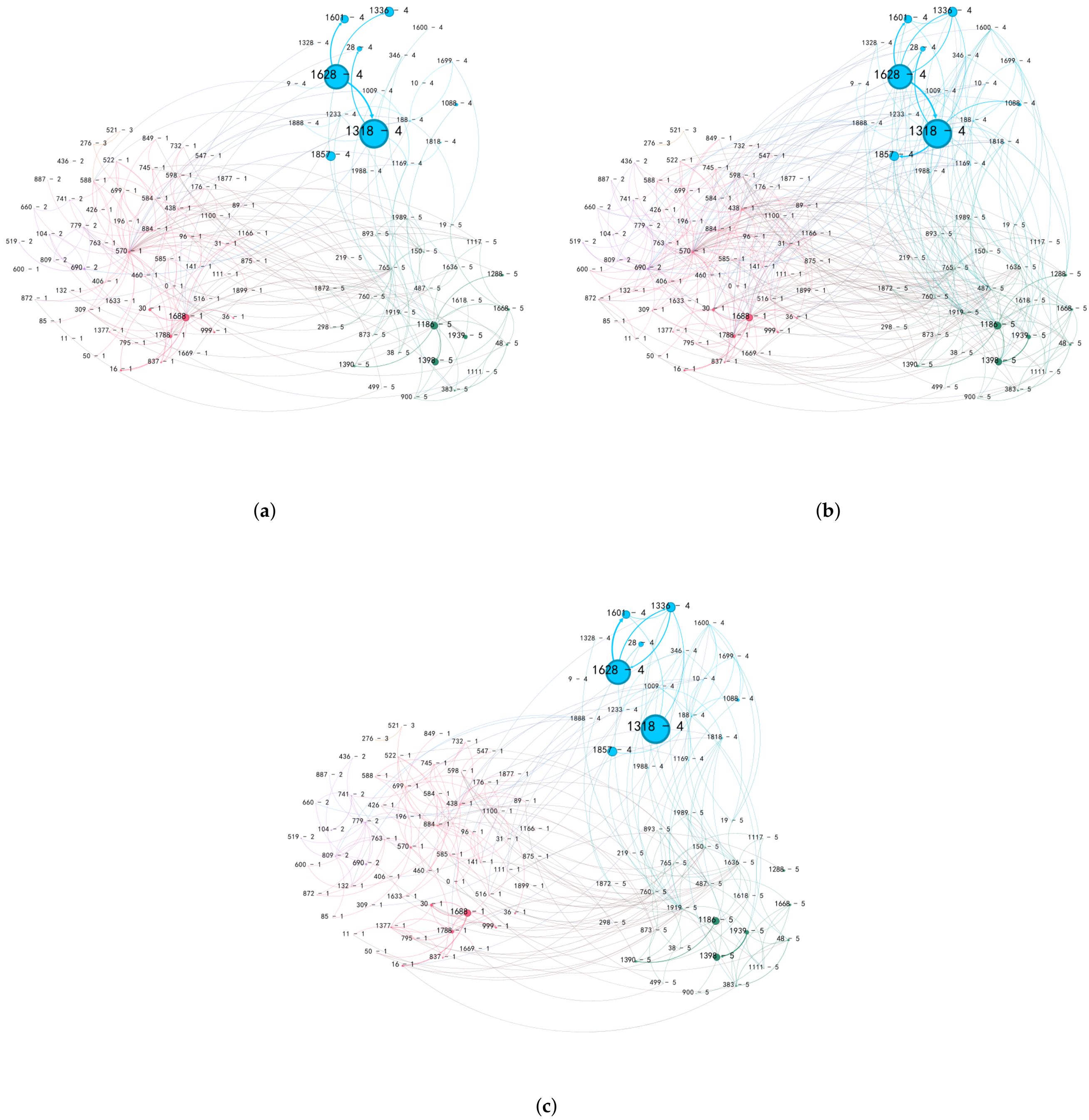
Through the analysis of DBSCAN, we divide the period from 2018 to 2022 into three time intervals: before December 2019, from December 2019 to June 2021, and after June 2021. Subsequently, we analyze the systemic risk network graphs for these three time periods as shown in Figures 13 and 14a,b.

Before December 2019, the systemic risk in the network is primarily transmitted between the insurance companies, including PingAn, CPIC, and NCI. This may be due to the significant clustering effect in the Chinese insurance industry. The five largest insurance companies PingAn, People's Insurance Company of China Ltd. (PICC, 01319, Beijing, China) , CLIC, CPIC, and NCI, have maintained over 50 percent of the market share over the long term. By examining their equity information, we find that the Central Huijin Investment Company Ltd. holds substantial shares in PingAn, CLIC, and CNI, which may lead to similarities in their stock price changes. These two aspects explain why the risk transmission effect among these companies is particularly pronounced in both risk input and output.

From December 2019 to June 2021, the network becomes dense, and during this period, the Chinese stock market is impacted by the initial wave of the COVID-19 pandemic, leading to significant economic uncertainty and an increase in systemic risk in the Chinese stock market. Still, most systemic risk is transmitted from PingAn to CNPC. This may be due to the strategic cooperation agreement signed at the end of 2017 between PingAn Bank[1] (Shenzhen, China) and CNPC. PingAn Bank, which is under the control of PingAn, agreed to provide a RMB 20 billion credit line to CNPC and its subsidiaries to support their various credit operations at PingAn Bank or its branches. At the same time, CNPC and PingAn Bank agreed to strengthen cooperation in various financial service domains. After June 2021, as the economy begins to recover, systemic risk in the Chinese stock market decreases to a lower level. Also, PingAn stops transmitting risk to CNPC . From the perspective of the community, evidence from Figure 13 underscores that Community 1 exhibits heightened internal connections. Within Community 2, the insurance industry exhibits complex interconnections and assumes a predominant role in bearing systemic risks. Notably, Industrial and Commercial Bank of China and China Railway Construction emerge as entities with the highest degrees within Community 3.

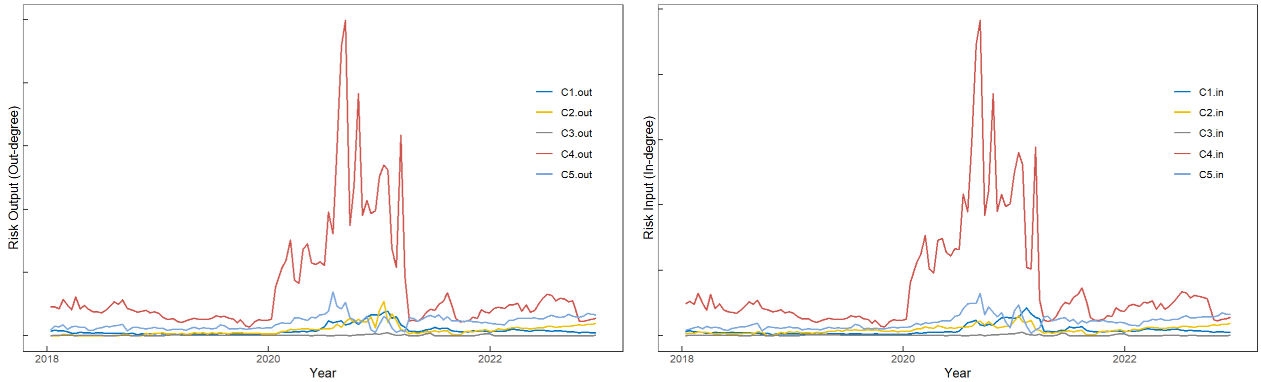
Through an examination of Figure 14a,b, we discern that Community 4 emerges as the principal component for both systemic risk inputs and outputs in the network, with other communities contributing to systemic risk at a lower level. Moreover, the systemic risk within Community 4 attains its peak during the pandemic period, suggesting a pronounced impact of the epidemic on the insurance industry in the short term.

By contrast, the systemic risk within the manufacturing industry exhibits fewer fluctuations during the epidemic period. From Figure 14a,b, it is observed that the systemic risk within Communities 1, 2, 3, and 5 experiences subtle increases from 2020 to 2022.



(**a**) (**b**) (**c**)

Figure 13. Systemic risk network of different time periods, (**a**) before December 2019, (**b**) from December 2019 to June 2021, and (**c**) after June 2021.



(**a**) (**b**)

Figure 14. (**a**) Risk output and (**b**) Risk input for each community.

4. Discussion

4.1. Conclusions

This paper defines the Risk Transmission Index and uses it to construct the systemic risk network in the Chinese stock market. Moreover, this paper further applies HITS, community detection, and DBSCAN clustering algorithms to analyze the systemic risk in the dimensions of time and community composition, resulting in the following conclusions:

1. By analyzing the four conventional network metrics—in-degree, out-degree, closeness centrality, and betweenness centrality—insurance companies serve both as the main contributors and major receivers of the systemic risk in the Chinese stock. No single stock significantly influences the others, but companies like Hang Seng Electronics play a pivotal role in risk propagation.
2. By analyzing the main risk output and input centers obtained from the HITS algorithm, we find that the biggest risk output centers are PingAn, CPIC and NCI; the biggest risk input centers are CLIC and NCI.
3. By examining the temporal evolution of systemic risk in the Chinese stock market, we conclude that a pre-2020 period is characterized by relatively low systemic risk. However, the onset of the COVID-19 pandemic's initial wave instigates significant shifts. Stringent pandemic control measures precipitate disruptions in the supply chain, production halts in select industries, and a pervasive sense of market tension, thereby engendering a noteworthy upsurge in systemic risk within the Chinese stock market. After the implementation of effective COVID-19 control measures, systemic risk reverts to normal levels in early 2021.
4. Moreover, an exploration of community characteristics derived from a community detection algorithm underscores that the sources of risk in the Chinese stock market from 2018 to 2022 predominantly manifest within sectors such as the secondary industry, emerging industries, and insurance. Each of these categories exhibits a pronounced internal correlation. The classification approach employed herein primarily hinges on the interplay of risk among companies, differing from conventional categorizations such as insurance, financial services, and others. Within this framework, the principal sources of risk emanate from the insurance sector. It is plausible that events such as pandemics can induce systemic risk in the insurance industry, and as these entities engage in non-traditional business activities, such endeavors may further contribute to systemic risks within the network.

From the previous analysis, we can lead to some constructive investment recommendations applicable to the Chinese stock market. Since the systemic risk in the network is primarily transmitted between the insurance companies, we suggest that investors should avoid concentrating all funds in the same industry or company; instead, diversifying investments across different industries, types of companies, and regions can be an effective method to mitigate risks. Additionally, due to the substantial transmission of systemic risk from PingAn to CNPC, it is also advised that investors not include them in the same investment portfolio.

4.2. Limitation and Further Work

The Risk Transmission Index defined in this paper is calculated at each moment t. In our systemic risk network within a specific time period T, we use $\sum_t A_t, t \in T$ as the adjacency matrix. However, there may be better measures of systemic risk transmission that allow us to compute the adjacency matrix on a time-period basis.

We use the HITS algorithm to identify input and output centers because it is superior to the PageRank algorithm, which can only calculate importance without differentiating between input and output. Future research may propose more reasonable network analysis algorithms for application in the background of systemic risk.

Furthermore, we have only conducted an association study on the relationships among individual stocks, industries, communities, time periods, and systemic risks. However, this cannot lead to causal conclusions. Therefore, further research is needed to analyze the causal effects involved.

Author Contributions: Conceptualization, Y.H. (Yifan Hu) and Y.H. (Yanxi Hou); methodology, X.Z., Y.H. (Yifan Hu) and Y.H. (Yanxi Hou); software, X.Z. and Y.H. (Yifan Hu); validation, X.Z., Y.H. (Yifan Hu) and C.P.; formal analysis, X.Z.; resources, C.P.; data curation, C.P.; writing—original draft preparation, X.Z., Y.H. (Yifan Hu), C.P. and Y.H. (Yanxi Hou); writing—review and editing, X.Z., Y.H. (Yifan Hu), C.P. and Y.H. (Yanxi Hou); visualization, X.Z. and C.P.; project administration, Y.H. (Yanxi Hou); funding acquisition, Y.H. (Yanxi Hou). All authors have read and agreed to the published version of the manuscript.

Funding: Hou's work was supported by the National Natural Science Foundation of China Grants 72171055 and 72161160340.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are available upon request.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ACF	Autocorrelation Function
ADF	Augmented Dickey–Fuller
CLIC	China Life Insurance Company Ltd.
CNPC	PetroChina Company Ltd.
CPIC	China Pacific Insurance Company Ltd.
CoES	Conditional Expected Shortfall
CoVaR	Conditional Value at Risk
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
ES	Expected Shortfall
HITS	Hyperlink-Induced Topic Search
MES	Marginal Expected Shortfall
NCI	Xinhua Insurance
RTI	Risk Transmission Index
PACF	Partial Autocorrelation Function
PingAn	Ping An Insurance
SCAD	Smoothly Clipped Absolute Deviation
SES	Systemic Expected Shortfall
SIFI	Systemic Important Financial Institution
SZ180	Shanghai 180 Index
VaR	Value at Risk

Note

[1] Ping An Bank is listed on the Shenzhen Stock Exchange, so we didn't include it in our research.

References

Acharya, Viral V., Lasse H. Pedersen, Thomas Philippon, and Matthew Richardson. 2017. Measuring systemic risk. *The Review of Financial Studies* 30: 2–47. [CrossRef]

Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber, and David Heath. 1999. Coherent measures of risk. *Mathematical Finance* 9: 203–28. [CrossRef]

Blondel, Vincent D., Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. 2008. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment* 10: 10008. [CrossRef]

Bo, Lin, and Zheng Li. 2016. Research on the Measurement and Monitoring of Systemic Risk in Finance. *Nankai Journal: Philosophy, Social Sciences Edition* 4: 150–60.

Brownlees, Christian, and Robert F. Engle. 2017. SRISK: A conditional capital shortfall measure of systemic risk. *The Review of Financial Studies* 1: 48–79. [CrossRef]

Diebold, Francis X., and Kamil Yılma. 2014. On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics* 182: 119–34. [CrossRef]

Duffie, Darrell, and Jun Pan. 1997. An overview of value at risk. *Journal of Derivatives* 4: 7–49. [CrossRef]

Ester, Martin, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. In *kdd*. München: Institute for Computer Science, University of Munich, Oregon: AAAI Press, vol. 96, pp. 226–31.

Fan, Jianqing, and Runze Li. 2001. Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association* 96: 1348–60. [CrossRef]

Fang, Yi, Shengmin Zhao, and Daoping Wang. 2012. Measurement of Systemic Risk in China's Financial Institutions: A Study Based on the DGC-GARCH Model. *Financial Regulation Research* 11: 26–42.

Hautsch, Nikolaus, Julia Schaumburg, and Melanie Schienle. 2015. Financial network systemic risk contributions. *Review of Finance* 19: 685–738. [CrossRef]

Härdle, Wolfgang Karl, Weining Wang, and Lining Yu. 2016. Tenet: Tail-event driven network risk. *Journal of Econometrics* 192: 499–513. [CrossRef]

Kaufman, George G., and Kenneth E. Scot. 2003. What is systemic risk, and do bank regulators retard or contribute to it? *The Independent Review* 7: 371–91.

Kleinberg, Jon M. 1999. Authoritative sources in a hyperlinked environment. *Journal of the ACM (JACM)* 46: 604–32. [CrossRef]

Kupiec, Paul H. 1995. *Techniques for Verifying the Accuracy of Risk Measurement Models*. Washington, DC: Division of Research and Statistics, Division of Monetary Affairs, Federal Reserve Board, vol. 95, pp. 73–84.

Lai, Juan. 2011. Research on Systemic Risk in China's Finance and Its Prevention. Doctoral dissertation, Jiangxi University of Finance and Economics, Nanchang, China.

Li, Zhihui, Yuan Li, and Zheng Li. 2016. Research on Monitoring Systemic Risk in China's Banking Industry: Implementation and Optimization Based on SCCA Technology. *Financial Research* 429: 92–106.

Lin, Yuanlong. 2023. The Impact of the COVID-19 Pandemic on the Chinese Stock Market.Master's thesis, Jilin University, Jilin, China.

Tobias, Adrian, and Markus K. Brunnermeier. 2016. CoVaR. *The American Economic Review* 106: 1705.

Wu, Qingwei, Tao Hong, and Jun Ma. 2010. A Review of the Long Tail Theory. *Journal of Zhoukou Normal University* 27: 124–29.

Zhang, Yuanping, and Gang Sun. 2003. Theoretical Analysis and Empirical Study on Financial Crisis Early Warning System. *International Financial Research* 10: 32–38.