

Analysis of Interprovincial Trade Relationships Based on Improved Gravity Modeling

Yixin Yang ^{1, a}, Jin Chen ^{2, b, *}

¹School of Management Science and Engineering, Anhui University of Finance and Economics, Bengbu, Anhui 233030, China

²School of Accountancy, Anhui University of Finance and Economics, Bengbu, Anhui, 233030, China

^a1438122537@qq.com, ^b1730607338@qq.com

Abstract

This paper analyzes the inter-provincial trade flows estimated by the improved gravity model using the complex network method. The article collects the distance and annual GDP value of each province from Baidu map and China Statistical Yearbook, and uses the gravity model to estimate the inter-provincial trade flow. Considering that the inter-provincial trade in the current year may be related to the GDP value of the previous year, the original model is improved by introducing the Gi-1 index and the results are normalized by multiplying them by 100 as the weights of the network constructed, and a fully connected network is established with the name of each province as a node connected network, and do the analysis of the network's node strength, weighted tight centrality, weight entropy of the three indicators, while using spectral clustering to further explore the trade relationship between the provinces.

Keywords

Complex network, gravity model, inter-provincial trade.

1. Introduction

Interprovincial trade in goods and services is of great significance in promoting the economic development of inland provinces. The study of inter-provincial trade helps to understand the industrial division of labor and collaboration between provinces and to explore the competitive advantages and disadvantages of each province's industry.

At present, the analysis of inter-provincial trade in China can be roughly divided into three categories. One category is the analysis of the structure and attributes of the trade exchange situation itself. Shi Bingzhan [1] estimates the potential of China's interprovincial trade; Shang Yong[2] explores the characteristics of China's interprovincial trade using interregional input-output tables. The second category examines the impact of interprovincial trade on other areas. Zhang Shaojun [3] and others analyzed the contribution of interprovincial trade to China's economic growth; Liu Yi'ang [4] and others explored the phenomenon of segmentation of China's interprovincial product markets. The third category analyzes the impact of other factors on the development of interprovincial trade. Zhou Xianghong [5] analyzed the impact of informatization gap on interprovincial trade; Xu Xianxiang [6] and others studied the association between railroad freight transport and interprovincial trade patterns.

Meanwhile, since there is no clear publicly available data on interprovincial trade, most scholars use input-output tables or gravity models, which are estimated through a series of calculations. In this paper, an improved gravity model is proposed, which is used to estimate the interprovincial trade in 2022 and conduct related analysis.

2. Improved Gravity Model

Gravity model as a tool for estimating interprovincial trade flows, especially applicable when the specific trade data is not known, the traditional gravity model using the GDP value of each of the two provinces and the distance between them to estimate the value of trade, it is usually believed that the better the economic development of the province in that year, the closer the distance to other provinces, the lower the trade cost, the easier it is to develop trade, and vice versa, the trade volume is smaller. On this basis, this paper argues that interprovincial trade flows are not only related to the economic situation of the current year, but the economic volume of the previous year also concerns the willingness of the province to carry out trade in the current year. The specific formula is as follows.

$$F_{ij} = K \frac{\sqrt{G_i * G_{i-1}} * \sqrt{G_j * G_{j-1}}}{D_{ij}^\theta} \quad (1)$$

Where, F_{ij} denotes the trade flow between two provinces i and j , G_i and G_{i-1} denote the GDP values of province i in the current year and the previous year, D_{ij} is the distance between the two provinces, k is the gravitational constant, usually taken as 1, and θ is the distance friction coefficient, usually taken as 2.

The calculated trade flows for selected provinces in 2022 are obtained as shown in Table 1.

Table 1. Trade flows in selected provinces in 2022

	Beijing	Tianjin	Hebei	Shanxi	Inner Mongolia
Beijing		34947.03	20080.48	4125.23	3791.52
Tianjin	34947.03		6637.06	1457.04	924.76
Hebei	20080.48	6637.06		19720.15	2582.44
Shanxi	4125.23	1457.04	19720.15		2613.78
Inner Mongolia	3791.52	924.76	2582.44	2613.78	

Considering the length of the article, this paper only lists the flow situation between five provinces, which can be seen that the economic communication between Beijing and Tianjin provinces is closer, while the trade relationship between Beijing and Inner Mongolia and Shanxi is poorer, but it also indirectly reflects the superiority of the economic development of Beijing and Tianjin.

3. Network Construction and Analysis

After obtaining the inter-provincial trade flow value, this paper takes each province and municipality directly under the central government as a node, and takes the flow value as the weight to construct a fully connected trade network. Considering the readability of the image, the flow value is normalized and multiplied by 100 to constrain it between 0 and 100 as the actual weight. Figure 1 illustrates the interprovincial trade flows in 2022.

From Figure 1, it can be seen that Anhui and Jiangsu have the closest trade flows, reflecting the great similarity between these two provinces in terms of product economy; Zhejiang also possesses close economic ties with Shanghai and Jiangsu, and the trade relations between Beijing and Tianjin, and Shandong and Hebei rank at the top of the list, except for the provinces of Hainan, Xinjiang, Yunnan, Tibet, Heilongjiang, Jilin, Qinghai, and Liaoning, which are all provinces in the 2022 there are more economic exchanges. Meanwhile, as a fully-connected network, each province has direct economic relations with each other, which reflects the frequent trade links among provinces in the country.

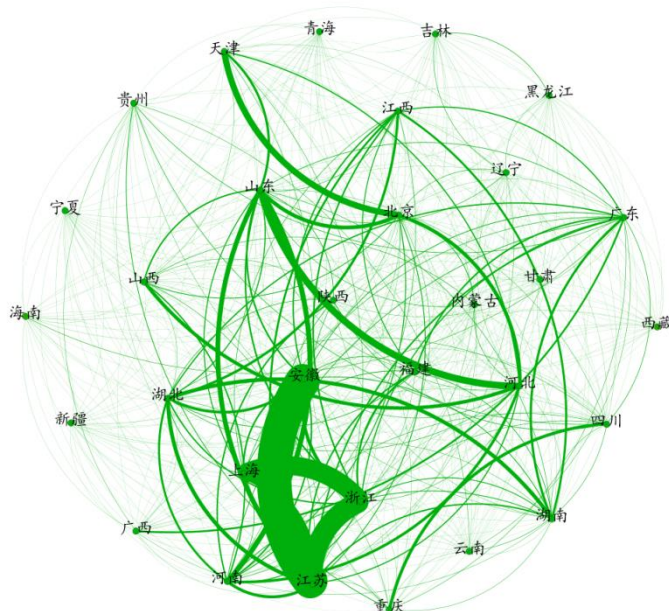


Figure 1. Inter-provincial trade flow network diagram

3.1. Node strength analysis

Under the premise of fully connected network, trying to study the importance of nodes, it is no longer possible to measure it with degree value. Considering that the weight of each edge is different, the weighted degree [7] is used to reflect the importance of different nodes by comparing the size of the sum of the weights of each node's connectivity, which is given by the following formula.

$$k_i^w = \sum_{j \in N(i)} w_{ij} \tag{2}$$

Where, $N(i)$ denotes the set of nodes that are directly connected to node i , and w_{ij} is the connection weight between nodes i and j .

After the calculation, the information of the nodes with the top and bottom three node weighting degrees is displayed in Table 2.

Table 2. Weighting degree of some nodes

Node Label	Weighting Degree	Node Label	Weighting Degree
Jiangsu	275.56	Qinghai	1.14
Zhejiang	158.84	Xinjiang	0.93
Anhui	146.71	Tibet	0.09

Jiangsu has the largest weighted degree of 275.56, which indicates that Jiangsu Province has the largest trade flows with other provinces, showing Jiangsu's convenient geographic environment and strong economic capital in terms of economic circulation; while Tibet has the smallest weighted degree of 0.09, which may be caused by the remoteness of Tibet's location and its vast geographic area, which makes the trade cost higher.

3.2. Weighted close centrality

Closeness centrality is an important index to assess the centrality of nodes in weighted networks, which reflects the average distance of a node to all other nodes in the network, however, in a fully connected network, the distance of any node to any other node is 1, which makes the interpretability of the final result extremely poor, therefore, this paper adopts the weighted closeness centrality [7] to do the assessment of node situation, which takes into account the weight of the edges. In weighted networks, lower weights usually mean tighter connections, which is calculated as follows.

$$Cc(i) = \frac{1}{\sum_{j \neq i} d(i,j)} \quad (3)$$

Where, $Cc(i)$ denotes the weighted tight centrality of node i , and $d(i,j)$ is the shortest path length from node i to node j considering the weights of each edge.

For the calculation results of this index, this paper still shows the results of the ranking of the first and last provinces for introduction.

Table 3. Weighted tight centrality of some nodes

Node labels	Weighted tight centrality	Node labels	Weighted tight centrality
Tibet	316.24	Guangdong	77.91
Jilin	306.85	Sichuan	83.43
Heilongjiang	300.85	Jiangsu	89.45

Comparing this result with the node strength, Tibet, as the region with the largest weighted tight centrality, possesses the lowest node weighting, considering that the actual meaning of the network weight is the size of the trade flow, and the larger its weight, the more tightly connected it implies the region is, which is opposite to the usual interpretation of the weighted tight centrality, i.e., the smaller the weighted tight centrality of the node is computed, the more it indicates that the region is located in the center of the domestic trade network.

According to the data in the table, Guangdong is at the center of the whole trade network, followed by Sichuan, indicating that these provinces have the closest trade relations with other provinces.

3.3. Weight Entropy

Weight entropy is a metric used to measure the complexity and diversity of the distribution of edge weights in the network, which can help us understand the uniformity of network connections, the higher the weight entropy, the more uniform the distribution of weights in the network, and vice versa, the more uneven the distribution of weights in the surface network, indicating that the trade transactions of some provinces may be far more important than other inter-provincial trade. The formula for weight entropy is as follows.

$$H = - \sum_{i=1}^n p_i \log(p_i) \quad (4)$$

Where, p_i is the value after weight normalization.

The weight entropy of the whole network is calculated to be 4.23, which indicates that the weight distribution of the whole network is extremely uneven. Combined with the previous analysis, it can be seen that the degree of connection between places like Jiangsu and Hebei and other provinces is much larger than that of Tibet and Xinjiang, which indicates that the economic development of these provinces is more prominent, and the trade exchanges between them may occupy a more important position in the whole trade network, which is also more conducive to the domestic economy's take off.

3.4. Spectral clustering

Spectral clustering [8] is a clustering algorithm based on graph theory, which is mainly used for data segmentation and cluster identification, and it can utilize the feature vectors in the similarity matrix of the data to perform clustering. Using this method, similar features between provinces can be found and classified into different categories, and provinces located in the same category will have stronger trade between them than provinces located in different categories.

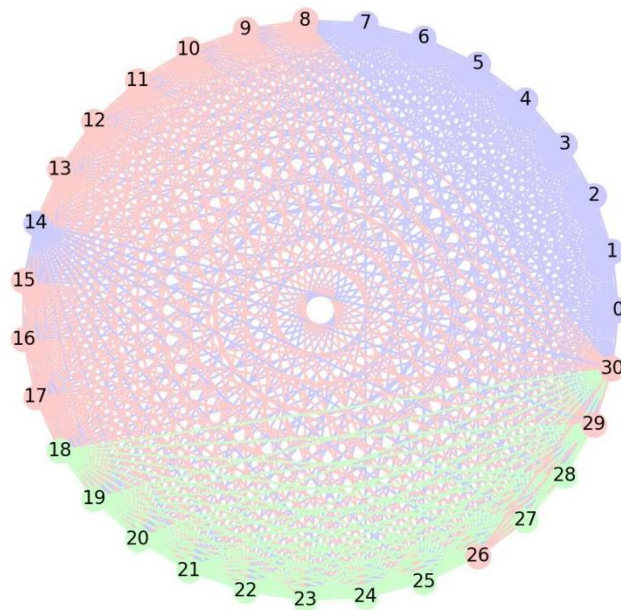


Figure 2. Network clustering results

According to the clustering results, it is obvious to know that Beijing, Tianjin, Hebei, Shanxi, Inner Mongolia, Liaoning, Jilin, Heilongjiang, Shandong are categorized as the first category, Shanghai, Jiangsu, Zhejiang, Anhui, Fujian, Jiangxi, Henan, Hubei, Hunan, Shaanxi, Ningxia, Xinjiang belong to the second category, and Guangdong, Guangxi, Hainan, Chongqing, Sichuan, Guizhou, Yunnan, Xizang, Gansu, Qinghai are the third category. Provinces belonging to the same category are bound to have more frequent trade exchanges, which, combined with the previous analysis, can also testify to the correctness of the classification.

Acknowledgments

This study was supported by Undergraduate Research Innovation Fund Program of Anhui University of Finance and Economics (XSKY23031ZD).

References

- [1] B.Z. Shi, R.E. Zhang. Estimation of China's interprovincial trade potential - A perspective based on the comparison of domestic and international trade [J]. *International Trade Issues*,2021(12):49-65.
- [2] Shang Yong. Analysis of China's interprovincial trade based on interregional input-output tables[J]. *Statistics and Decision Making*,2017(21):134-137.
- [3] Zhang Shaojun,Li Shantong. Research on the contribution of inter-provincial trade to China's economic growth[J]. *Research on Quantitative Economics and Technical Economics*,2017,34(02)
- [4] Liu Yi'ang,Lai Desheng. Research on the factors of product market segmentation in China based on gravity model--panel data from inter-provincial railroad freight trade[J]. *Economic Jingwei*, 2016,33(01).
- [5] ZHOU Xianghong,CUI Zhaocai. Research on the mechanism of informationization gap affecting interprovincial trade imbalance--an empirical test based on railroad freight data from 2003 to 2012[J]. *Journal of Public Management*,2020,17(01).
- [6] XU Xianxiang, LI Shun. China's interprovincial trade patterns:a study based on railroad freight[J]. *World Economy*,2012,35(09).
- [7] HUANG Hui, LI Ruiji. Research on key node identification of aviation manufacturing supply chain based on complex network[J]. *Advances in Aeronautical Engineering*,2023,14(06):167-177.

- [8] HAO Lili, ZHANG Pan, LI Hongmin, et al. Evaluation of China's green innovation efficiency and influence mechanism based on spectral clustering[J]. Journal of Northeast Normal University (Natural Science Edition),2024,56(01):46-54.