

Research on Influencing Factors of Export Complexity of Chinese High-tech Industry Based on Big Data Analysis

WANG JUNSHENG¹, Ziqi ZHONG^{1*}, Wang Haoran¹

¹Department of International Business, Graduate School of Pan-Pacific International Studies, Kyung Hee University, Yongin-Si, Gyeonggi-Do, South Korea

Abstract. Export trade can measure the economic level of a country, but it can only reflect the amount of exports, but not the quality of exported products and the technical content of exported products. Therefore, domestic and foreign scholars have begun to study the complexity of export technology. The development of big data technology makes it possible to analyze the export complexity using big data analysis technology. With the rapid development of high-tech industries represented by high-end manufacturing, there is more and more research on the export of high-tech industries. Based on the existing research results, this article first introduces the current export profile of high-tech products and explains the concept of export complexity. Then, the flow of big data analysis was sorted out. Finally, this paper theoretically analyzes the influence of industrial agglomeration on industrial export complexity, and uses big data analysis and regression verification. The results show that industrial agglomeration has a significant role in promoting the export complexity of China's high-tech products.

1 Introduction

Since the reform and opening up, China's foreign trade has made tremendous development. In recent years, the main body of China's foreign trade is slowly shifting from traditional labor-intensive industries to capital-intensive industries, and the proportion of traditional labor-intensive industries in total exports has been declining year by year [1]. In this large realistic background, most researchers have turned their research focus to the study of the technical complexity of export commodities.

At present, China's high-tech industrial product exports mainly include the following three aspects.

1.1 Independent Innovation Ability is Poor

China's high-tech products is lack of core technology and awareness of property rights protection, the specific performance of low-end products exceeds supply, and high-end products rely on imports [2].

1.2 The Added Value of Product Exports is Low

Most of China's high-tech product exports are processing trade, which is at the low end of the industrial chain. The profit depends on cheap labor, not advanced technology.

1.3 China's High-tech Products Have a High Degree of Dependence on Foreign Technology

Because of the lack of independent innovation awareness, most of them do not care about technological progress by imitating or purchasing advanced foreign technology [3].

Therefore, in order to improve the status of products in the high-tech industry, we must strengthen the awareness of independent innovation, enhance core competitiveness, improve production technology, increase the technical complexity of export products, and optimize the export structure [4].

2 Overview of export complexity

The concept of export complexity was first proposed by Hausmann et al. They believed that the complexity of export technology reflects the export commodity structure of a country or region, and can reflect the technical content and international division of labor of the country or region's exports to a certain extent. The higher the export complexity of a commodity, the higher the technical content contained in the commodity, and the more competitive the region's foreign trade; Conversely, the lower the export complexity, the lower the quality of the commodity, the lower the technological content, and the greater the dependence on natural resource factors, the region's foreign trade is not competitive [5]. At present, the research on export complexity mainly focuses on the following three aspects:

*Corresponding author's e-mail: chichung@khu.ac.kr

- Study the relationship between export complexity and economic growth.
- Using export complexity indicators to measure the export complexity of a country or industry.
- Analysis of factors affecting export complexity.

In these three aspects of research, more attention is paid to what factors play a role in the dynamic evolution of a country's export complexity. Regarding the research on the influencing factors of export complexity, the existing literature mainly analyzes from the perspectives of foreign direct investment, economic development, human capital, R & D investment, and system quality. There are few literatures from the perspective of industrial agglomeration. This article starts from the perspective of industrial agglomeration and attempts to explore the impact mechanism of industrial agglomeration on export complexity based on big data analysis methods [6].

3 Big data analysis

For big data analysis, it can be regarded as a process of layer-by-layer expansion and analysis of data. In the information age, a large amount of data is generated every day. How to mine valuable information from these large amounts of data is the basic task of big data analysis [7]. This article validates the viewpoint of this article based on the regression analysis method in data mining. In the context of the information age, big data analysis also came into being. The following is a summary of the basic methods of big data analysis methods.

Based on the bellow five points, the big data of high-tech industry exports are initially sorted out: Effective data extraction and noise removal are the first steps in data analysis. Because of the inconsistent time span of the data, the data structure, the inconsistency of the previous and the previous years, and the inconsistent range of the information records of the previous and the other years, the directly collected data cannot be directly used for analysis. To this end, the original data must be filtered, and the data with high correlation and complete information retention must be retained for analysis [8].

3.1 Visual Analysis

The main requirement of data analysis for big data analysis is data visualization. Visual analysis can more intuitively analyze and process data, allowing users to find results from the data itself.

3.2 Data Mining Algorithm

The main body corresponding to data mining is the machine. Using data clustering, data segmentation, and outlier analysis to analyze the data, not only the amount of data processed is huge, but also the processing speed is faster.

3.3 Predictive Analysis Capabilities

Data mining enables data analysts to obtain valuable information about the data, while predictive analysis is the

data analysts' prediction and judgment on the development trend of data based on the results of visual analysis and data mining.

3.4 Semantic Engine

Due to the complexity of big data, data analysis is faced with new challenges. Therefore, advanced data analysis tools are needed to extract, analyze, and judge data. Based on this, a semantic engine is generated, which can intelligently rely on documents to extract relevant information.

3.5 Data Storage

The data warehouse is a relational database established according to a specific model for the convenience of multi-dimensional analysis and multi-angle display of data. In the design of business intelligence systems, the construction of a data warehouse is the key, the premise of realizing big data analysis, the key to integrating business system data, and supporting the normal operation of the system with data extraction, conversion, and loading, while enabling data Query and system access to build a basic platform for multi-combination data analysis and data mining.

4 Analysis of factors affecting export complexity

With the development of China's economy and technological innovation, the geographical agglomeration of various industries is becoming more and more obvious, especially high-tech industrial clusters. On the one hand, it can promote regional economic growth and give play to regional advantages. It also makes high-tech industrial clusters further promote the transformation and upgrading of China's foreign trade structure. The development of high-tech industries plays a very important role in a country's economic development. The following analyzes the impact of industrial clustering on high-tech industries and verifies the analysis through big data analysis [9].

Industrial clusters can reduce the cost of enterprises entering the market through the sharing of resources such as infrastructure, information resources, and labor. When high-tech products participate in international trade, due to the professionalism and complexity of the customers and sales channels they face, this product has great uncertainty in trade exports. Industrial clusters can improve the technological innovation capabilities of enterprises in the cluster through knowledge spillover effects and talent concentration effects, so that high-tech enterprises can quickly respond to and adjust to various uncertainties in the international market. Therefore, industrial clusters can enhance the ability of enterprises to resist external risks and improve the ability of enterprises to export products, thereby increasing the complexity of product exports. The following is a detailed analysis of the effects of knowledge spillover effects and talent concentration effects on export complexity [10].

4.1 The Influence of Knowledge Spillover Effect in Industrial Cluster on Export Complexity

The development of high-tech industries mainly depends on technological innovation. However, technological innovation requires a lot of manpower, material and financial resources, and the risk of research and development is very large, which is a long-term process. Therefore, independent research and development of enterprises must have strong capital and high-tech talent resources. However, most of China's high-tech enterprises are small and medium-sized enterprises. They do not have sufficient funds and lack independent research and development capabilities. Therefore, for these small and medium-sized high-tech enterprises, they need to receive outside help. The spillover of knowledge in the cluster is an important way for SMEs to quickly absorb relevant technological innovation knowledge.

The improvement of the overall technological innovation level within the industrial cluster depends on the improvement of the technological level of small and medium-sized enterprises that are relatively backward in R & D technology. Whether it is a large high-tech enterprise or a small and medium-sized high-tech enterprise, knowledge spillover can increase the overall knowledge stock in the cluster, stimulate cluster innovation activities, reduce enterprise innovation costs and risks, and thus promote the overall innovation level within the industrial cluster. It helps to optimize the export structure of enterprises in the cluster and increase the complexity of export of high-tech products.

4.2 The Influence of Talent Accumulation Effect on Export Complexity in Industrial Clusters

Talent agglomeration refers to the concentration of talents in a specific space or professional field, and close contact with each other to promote the sharing of information and knowledge, so that the regional industrial economy can obtain sustained and strong competitive advantages. A large number of talents are gathered in the industrial cluster, and the competition between high-tech industries is fierce. Therefore, enterprises with technological advantages can obtain market share, and in order to not be eliminated by the market, the technological laggards must also increase research and development efforts to catch up with the technologically advanced ones. This kind of competition within the cluster keeps the atmosphere of continuous innovation in the cluster and promotes the continuous development of technology. There is competition within the industrial cluster and mutual cooperation among cluster members. They have their own strengths and there are differences in skills. Formal and informal exchanges between technicians and close cooperation may produce new innovations. Ji Xiaoli and others analyzed the innovation path of talent gathering, and believed that talent gathering provided the convenience of sharing hidden knowledge, and hidden knowledge determines innovation; talent gathering strengthens innovation incentives, reduces innovation risks, and helps improve innovation efficiency. Industrial agglomeration and talent agglomeration complement each

other, promote each other, and jointly promote the development of innovation capabilities of enterprises in the cluster, thereby helping to improve the technical content and quality of enterprises' export products.

4.3 Big data analysis and verification

This paper selects the data from 2002 to 2019 in 9 provinces and 2 cities in the Yangtze River Economic Belt for empirical testing. The relevant data comes from the "China High-tech Industry Statistical Yearbook", the relevant provincial and municipal statistical yearbooks and the national research network database.

In order to further verify the impact of high-tech industrial clusters on export complexity, the following general regression model was established to test:

$$\ln Extec_{kt} = \alpha_0 + \alpha_1 LE_{kt} + \gamma X_{kt} + \varepsilon_{kt} \quad (1)$$

Among them: *Extec* is the explanatory variable, that is, the export complexity of high-tech products. *LE* stands for the concentration of high-tech industries. If the coefficient α_1 is positive, it means that the higher the concentration of high-tech industries, the higher the export complexity. *X* represents various control variables. In this paper, several variables of GDP per capita, patented technology, foreign direct investment and human capital are added. *K* represents the province, and *t* represents the year, and $\ln(*)$ represents the logarithm, and ε represents the random interference term that follows the normal distribution.

With reference to the research of Rodrik and Hausmann, the formula (2) measures the export complexity of products:

$$PRODY_i = \sum_j \left\{ \frac{\frac{x_{ji}}{X_j}}{\left(\sum_j \frac{x_{ji}}{X_j} \right)} Y_j \right\} \quad (2)$$

Among them: *PRODY_i* refers to the export complexity of product *i*. *X_{ji}* represents the export value of product *i* in a country or region *j*, and *X_j* represents the export value of a country or region *j*. *Y_j* represents a country or region *j* GDP per capita. The greater the value of *PRODY_i*, the higher the export technical level of commodity *i*, and the more obvious the comparative advantage of exports.

After calculating the export complexity of a product using equation (2), the export complexity of a country or region can be calculated by the following formula:

$$EXPY_{jt} = \sum \left\{ \frac{x_{jit}}{X_{jt} PRODY_{IT}} \right\} \quad (3)$$

Where: *EXPY_{jt}* is the export complexity of country *j* in year *t*, and *x_{jit}* is the export value of product *i* in country *j* in year *t*, *X_{jt}* is the total export value of country *j* in year *t*, and *PRODY_i*, it is product *i* in year *t* Export complexity.

Perform regression analysis based on data, regression model and export complexity calculation formula:

4.3.1 Correlation coefficient analysis. Before regression, the correlation analysis of the relevant variables in the model is carried out. The results are shown in the following table. It can be seen from the table that the correlation coefficient between the explained variable,

that is, the export complexity of high-tech products and each explanatory variable is positive; Although the correlation coefficient between the export complexity and the industrial cluster is not very large, from the perspective of the coefficient sign, the expectation of a positive correlation between the two is consistent. The correlation analysis is shown in Table 1.

Table1. Correlation analysis

	<i>Ln Extec</i>	<i>LE</i>	<i>ln pcGDP</i>	<i>ln FDI</i>	<i>ln human</i>	<i>ln patent</i>
<i>Ln Extec</i>	1.0000					
<i>LE</i>	0.1658	1.0000				
<i>ln pcGDP</i>	0.7364	0.0873	1.0000			
<i>ln FDI</i>	0.7131	-0.1178	0.8129	1.0000		
<i>ln human</i>	0.7742	0.3077	0.7811	0.7857	1.0000	
<i>ln patent</i>	0.7776	0.2697	0.8992	0.7398	0.8701	1.0000

4.3.2 General linear regression analysis between high-tech industry agglomeration and export complexity. The effect of high-tech industrial agglomeration on export complexity has been described above. To further verify the effect of industrial agglomeration on export complexity, use stata11. 0 software performs the following OLS linear regression and conducts empirical tests. The results are shown in Table 2.

Table2. General linear regression of the effect of industrial agglomeration on export complexity

Explanatory variables	Explained variable (<i>Ln Extec</i>)				
	Model 1	Model 2	Model 3	Model 4	Model 5
<i>LE</i>	0.257**	0.158*	0.288***	0.0465	0.0159
	2.14	1.95	3.56	0.49	0.17
<i>ln pcGDP</i>		1.196***	0.595***	0.437***	0.0671
		13.78	4.08	3.03	0.31
<i>ln FDI</i>			0.411***	0.153	0.198*
			4.96	1.52	1.96
<i>ln human</i>				0.463***	0.298**
				4.17	2.31
<i>ln patent</i>					0.265**
					2.42
<i>_cons</i>	8.237***	-3.664***	0.0367	-0.865	2.326
	55.67	-4.21	0.04	-0.82	1.36
<i>N</i>	166	166	166	166	166
<i>Adj. R-sq</i>	0.23	0.546	0.605	0.642	0.652

Note: ***, **, and * indicate significance at 1%, 5%, and 10% significance levels, respectively.

As can be seen from the regression results in Table II, in the model (1), the sign of the coefficient between the agglomeration of high-tech industries and export complexity in the Yangtze River Economic Belt is positive, and it is significantly positive at the 5% significance level of. This shows that the previous assumptions are correct, indicating that the higher the concentration of high-tech industries in the Yangtze River Economic Belt, the higher the export complexity. The high-tech industry clusters are conducive to the increase in product export complexity. It

can be seen from the models (2) and (3) that after adding the variables of GDP per capita and foreign direct investment, the coefficient between the industrial agglomeration and the export complexity is still significantly positive. Models (4) and (5) After adding the human capital and patented technology variables, although the coefficient between industrial agglomeration and export complexity is not significantly positive, it has not changed the sign of the coefficient.

5 Conclusion

It can be seen from the above analysis that industrial agglomeration has a significant promotion effect on the export complexity of China's high-tech products, which shows that industrial agglomeration has made a tremendous contribution to optimizing the structure of China's export commodities and improving the technical content of export commodities. In addition, among many control variables, the level of per capita technology, patented technology, foreign direct investment, and the degree of infrastructure improvement are also conducive to increasing the complexity of high-tech product exports. The degree of industrial agglomeration has different effects on the technical complexity of export commodities in different regions and different levels of economic development. On the one hand, the knowledge spillover effect and the talent accumulation effect in the industrial cluster promote the technological innovation of the enterprise by increasing the knowledge stock in the high-tech industrial cluster, and thus contribute to the increase of the export complexity of the enterprise; The exchange of knowledge and information and close cooperation have further strengthened the innovation incentives, thereby promoting the improvement of the technological content of enterprises' export products.

References

1. B. Sheng, X.F. Qian, J.L. Huang, etc, Ten years of transformation after China's entry into the WTO: Review and prospect of China's foreign trade development. *International Economic Review*, 2011, vol.5, pp.84.
2. Y. Bian, L. Ma, Y.Y. Wang, Influencing factors and countermeasures for the development of high-tech industrial clusters. *Technology and Management*, 2009, vol.11, pp.33.
3. W. Fang, M. Yang, The effect of knowledge spillovers of high-tech industrial clusters on enterprise technology catch-up, *Technological progress and countermeasures*, 2018, vol.23, pp.66.
4. Y. Xuan, S.Y. Xuan, Empirical research on industrial agglomeration, technological innovation approach and export of high-tech enterprises, *International Trade Issues*, 2012, vol.5, pp.25-56.
5. R.D. Yang, Research on the Influencing Factors of China's Industrial Products Export Growth: Empirical Analysis Based on Panel Data by Industry from 1994 to 2005, *World Economy*, 2008, vol.8, pp.32-42.

6. D.Z. Cheng, "An International Comparative Analysis of the Complexity of China's Service Exports-Also Explaining the" Service Trade Balance Paradox ", " Economic Research Working Paper, 2013, pp.4-6.
7. D. Fernandez-Arias, M. Lopez-Martln and T. Montero-Romero. Financial Soundness Prediction Using a Multi-Classification Model: Evidence from Current Financial Crisis in OECD Banks, Computational Economics, 2017, vol.4, pp.1-23.
8. M. Islami, J.R. Kurz-Kim. A Single Composite Financial Stress Indicator and Its Real Impact in the Euro Area, International Journal of Finance and Economics, 2014, vol.19, pp.204-211.
9. J.F. Gao, Representation of China's foreign trade export technology structure and its in-depth analysis, Reform, 2015, vol.5, pp.90-95.
10. Y.M. Huang, W.J. Zhang, Measurement of China's export complexity and analysis of influencing factors, World Economic Research, 2017, vol.12, pp.59-64.