Regression Analysis of the Impact of Guangdong Province’s Tertiary Industry on Export Trade Based on the Diamond Model

Minjing Peng\textsuperscript{1a}, Min Liang\textsuperscript{2b}

\textsuperscript{1}School of Economics & Management, Wuyi University WYU SEM Jiangmen, China
\textsuperscript{2}School of Economics & Management, Wuyi University WYU SEM Jiangmen, China
\textsuperscript{a}peng.mj@wyu.edu.cn
\textsuperscript{b}Corresponding author:260303350@qq.com

Abstract. With the advent of the era of big data, data analysis can objectively explain many events and behaviors with scientific results. China is actively promoting the development of the service industry and opening up to the outside world. Guangdong is the largest foreign trade province. It is of great significance to analyze the impact of the tertiary industry on the export trade by means of data analysis. Based on Michael Porter diamond model, the data of six variables from 2010 to 2019 are selected under the collection of a large number of data, such as the number of employment in the tertiary industry, research and development expenditure. Unit root test and co-integration test are used to analyze. At last, using the method of stepwise regression, the regression equation model of Guangdong Province’s export trade volume under the influence of the tertiary industry is established. Through the model, it is concluded that the research expenditures and development level of the tertiary industry can significantly promote the export trade of goods in Guangdong Province, and then improve the competitiveness of export trade. In addition, this paper also stands on the perspective of the tertiary industry, for the development of commodity export trade in Guangdong Province to put forward relevant suggestions.

1. Introduction

In the context of big data, data has become the most active factor of production after land, labor, capital, and technology. Data flow has become a basic and strategic resource for promoting high-quality economic development and improving people’s lives. Therefore, our country will continue to focus on data analysis, statistical operations and other capabilities, and display the optimization in a scientific and reasonable way.

Guangdong is at a critical juncture of transformation and upgrading. Tertiary industries such as information transmission, software and information technology have played an important role in resisting the impact of COVID-19 and promoting trade and export. Guangdong is an extroverted economy. In recent years, the export dependence of Guangdong’s economy has reached more than 90%, and export has driven the growth of the provincial economy. Even if affected by the epidemic and the trade war between China and the US in 2020, Guangdong will still create a trade surplus of US $3,709.580 billion, and its total export volume will be much higher than that of other provinces and cities. Therefore, based on the diamond model, from the perspective of the tertiary industry, scientific
data analysis methods are used to analyze Guangdong's commodity export trade, which provides a reference for Guangdong's deployment of commodity trade strategies.

Referring to the production factors, demand conditions, enterprise strategy, enterprise organization, competitive factors and related industrial factors in the diamond model, this paper sets the number of employment in the tertiary industry, research expenditure, per capita regional GDP, foreign investment, development level of the tertiary industry, and added value of producer services as variables. Through unit root test, the stationary test of variables is carried out, and the competitiveness of the tertiary industry to commodity export trade in Guangdong Province is explored and analyzed.

2. Research methods and variable setting

Under the policy of opening to the outside world, China's overall trade volume increased from 2015 to 2020. However, in recent years, under the influence of trade war and national novel coronavirus, the growth trend of domestic trade slowed down. Meanwhile, as an important seaport province in eastern China, Guangdong Province was affected in import and export trade. Therefore, it is necessary to analyze the export trade of Guangdong Province by means of data analysis. The impact of the tertiary industry on the export trade of Guangdong Province can be seen directly and scientifically through the results of data analysis.

Therefore, the variables of the research analysis are set under the element index based on the diamond model. The co-integration regression equation is obtained by unit root test and co-integration test. In order to reduce the collinear relationship between multiple variables, the stepwise regression mathematical analysis method is adopted to construct the equation model, and the impact of the tertiary industry on the export trade in Guangdong Province is explored through the data results. Using the data results to explore the impact of the tertiary industry on export trade in Guangdong Province.

2.1. Data Processing Method

2.1.1. Unit root test

The purpose of unit root test is to eliminate the false regression and false correlation problems caused by unstable data to the model, so the unit root stationarity test of data variables is a necessary operation. The stationarity of time series is important for data modeling. The instability of time series data may lead to the failure of traditional t-test, or false correlation between independent variables, resulting in model distortion. Therefore, unit root test is needed to determine whether the data has stationarity by checking whether the data has unit root. The following procedures are established for the data:

\[ y_t = \alpha y_{t-1} + u_t, u_t \sim IID(0, \sigma^2) \]  
\[ y_t = \alpha y_{t-1} + u_t + \mu, u_t \sim IID(0, \sigma^2) \]  
\[ y_t = \alpha y_{t-1} + u_t + \beta t, u_t \sim IID(0, \sigma^2) \]

In formula (1) - (3), \( u_t \sim IID(0, \sigma^2) \) , \( y_0 = 0 \). Where \( \mu \) is the displacement term and \( \beta t \) is the time trend term. For the unit root test of time series \( y_t \) in the above model, the null hypothesis and the alternative hypothesis are respectively:

\[ H_0: \alpha = 1 \]
\[ H_1: \alpha < 1 \]

According to the test principle, if the value of degree of freedom(DF) is greater than or equal to the critical value, the null hypothesis is accepted, indicating that it is non-stationary. If the value of degree of freedom is less than the critical value, the null hypothesis is rejected, indicating that it is stable.

2.1.2. Cointegration test

In the case of unstable time series data, a reasonable model can be constructed through co-integration test. The idea of cointegration test is to put the existing first-order integrated variables together for analysis, and make linear combination through variables, so as to eliminate their random trend and get
their long-term linkage trend. At present, EG-ADF detection has a wide range of applications. The main steps of EG-ADF detection are as follows:

\[ y_t = \gamma_0 + \gamma_1 x_t + \epsilon_t \]
\[ e_t = y_t - (\gamma_0 + \gamma_1 x_t) \]
\[ \Delta e_t = \eta e_{t-1} + \sum_{i=1}^{m} \lambda_i \Delta e_{t-i} + \epsilon_t \]
\[ \Delta e_t = \alpha + \beta t + \eta e_{t-1} + \sum_{i=1}^{m} \lambda_i \Delta e_{t-i} + \epsilon_t \]
\[ \Delta e_t = \alpha + \beta t + \eta e_{t-1} + \sum_{i=1}^{m} \lambda_i \Delta e_{t-i} + \epsilon_t \]

The least square method is used to estimate the model as shown in formula (4), and then the corresponding residual sequence is calculated to obtain formula (5). Next, equation (6-8) is used to detect the stationarity of the residual sequence. If the null hypothesis is rejected after testing, the residual sequence is stationary, which means that there is a co-integration relationship between \( y \) and \( x \), and formula (4) is a co-integration regression equation. If the null hypothesis is accepted, then the residual sequence is non-stationary and there is no co-integration relationship between the dependent variable and the independent variable.

2.2. Data Selection and Variable Setting

2.2.1. Introduction to diamond model

The diamond model was put forward by Michael Porter, an American strategic management expert. Porter's diamond model theory is a method to understand the global competitive position of a country or a region, which is suitable for analyzing the competitiveness of various regions or industries\[4\]. Porter diamond model theory proposes that the domestic economic environment has a great impact on competitive advantage, mainly including the following basic factors: production factors, demand conditions, enterprise strategy, enterprise institutions and competitive factors, related industry factors, opportunities, and government behavior\[5\]. The relationship of several elements in Porter’s diamond model is shown in Figure 1, in which the four elements of production factor, demand condition, enterprise strategy, enterprise organization and competition factor, and relevant industry factor are the decisive factors.

![Figure 1: Diamond model elements](image)

2.2.2. Variable setting

The model was constructed from 2010 to 2019, and 10 years of data were used for analysis. The data mainly came from statistical resources such as National Statistical Yearbook and Statistical Yearbook of Guangdong Province.

2.2.2.1. Explained variable

The export volume of goods (\( Y \)): The competitiveness of the export trade of goods is shown by the export volume of goods. The greater the export volume of goods indicates, to some extent, the stronger
the export trade capacity of the region or country, thus the stronger the competitiveness of the export trade. Therefore, the export volume of goods is taken as the explanatory variable.

2.2.2.2. Explanatory variables
Based on the elements of diamond model, the number of employment in the tertiary industry \( X_1 \) and the research expenditure \( X_2 \) are taken as the production factors. The demand condition factor is the per capita GDP \( X_3 \); The enterprise strategy, structure and competition factors are foreign investment \( X_4 \); The relevant industrial factors are the development level of the tertiary industry \( X_5 \) and the added value of producer services \( X_6 \). Given the desirability of the data, government action and opportunity are not included as explanatory variables for the time being.

3. Process of empirical analysis

3.1. Model Construction
The explanatory variables are set according to each element of the diamond model. These six explanatory variables all have an important impact on the export volume of goods, namely the explained variables. Therefore, the model is set as follows:

\[
\ln Y = C + \alpha_1 \ln X_1 + \alpha_2 \ln X_2 + \alpha_3 \ln X_3 + \alpha_4 \ln X_4 + \alpha_5 \ln X_5 + \alpha_6 \ln X_6 + \eta
\] (9)

According to the theoretical knowledge of diamond model and econometrics model, the model set in this paper is shown in model formula(9). In model formula(9), "C" represents the constant term, "\( \alpha_n \)" represents the parameter to be estimated, and "\( \eta \)" represents the random error term.

3.2. the Unit Root Test of the Model
The unit root test was carried out on all variables. As shown in Table I, variable \( \ln Y, \ln X_1, \ln X_2, \ln X_3, \ln X_6 \) has the stationarity of original data, and variable B is stable after first-order difference, while variable \( \ln X_4 \) is unstable after first-order difference and second-order difference.

Therefore, the variables selected successively are \( \ln X_1, \ln X_2, d.\ln X_3, \ln X_5, \ln X_6 \).

Table 1  Unit root test of model variables

| variable | The ADF statistics | ADF statistic threshold | Test results |
|----------|--------------------|-------------------------|--------------|
| \ln Y    | -3.580             | -3.750 -3.000 -2.630    | stability    |
| \ln X_1  | -4.090             | -4.380 -3.600 -3.240    | stability    |
| \ln X_2  | -6.440             | -4.380 -3.600 -3.240    | stability    |
| \ln X_3  | -3.234             | -4.380 -3.600 -3.240    | instability  |
| d.\ln X_3| -3.741             | -4.380 -3.600 -3.240    | stability    |
| d.\ln X_4| -1.662             | -4.380 -3.600 -3.240    | instability  |
| d.\ln X_5| -1.443             | -4.380 -3.600 -3.240    | instability  |
| d.\ln X_6| -2.664             | -4.380 -3.600 -3.240    | instability  |
| \ln X_4  | -2.664             | -4.380 -3.600 -3.240    | instability  |
| \ln X_5  | -2.664             | -4.380 -3.600 -3.240    | instability  |
| \ln X_6  | -5.478             | -4.380 -3.600 -3.240    | stability    |

3.3. Cointegration Test
In this paper, the co-integration test adopts the traditional and classical EG-ADF test method, and the ordinary least square method is used to estimate the residual sequence, and the ADF test is carried out on the residual sequence. Therefore, the time trend of the residual sequence, and it can be found that the residual sequence has no fixed time trend.

In addition, the ADF test results of residual sequence are shown in Table II. \( Z (t) \) value is -2.568, which rejects the unit root null hypothesis, so the residual sequence is stable. In conclusion, there is a stable co-integration relationship between explanatory variables and explained variables, and regression analysis can be conducted on the model.
Table 2  Residual sequence ADF test results

| statistic threshold | Z(t) value |
|---------------------|------------|
| Test Statistic      | -2.568     |
| 1% Critical Value   | -2.660     |
| 5% Critical Value   | -1.950     |
| 10% Critical Value  | -1.600     |

3.4. Regression Analysis Treatment

According to the test of variance inflation factor, the VIF value is greater than 10, and the model has serious collinearity. In order to reduce the influence of multiple linearity, linear regression is carried out by means of stepwise regression. Variables are selected according to the criteria set. As shown in Table III, input variables $\ln X_2$ and $\ln X_5$ are obtained.

Table 3  Stepwise regression process variable selection

| model | variable           | R    | R2     | Adjusted R2 |
|-------|--------------------|------|--------|-------------|
| 1     | (constant), $\ln X_2$ | 0.951| 0.905  | 0.893       |
| 2     | (constant), $\ln X_2$, $\ln X_5$ | 0.983| 0.965  | 0.955       |

The results in Table IV shows that $\ln X_2$ and $\ln X_5$ is selected as the regression variables, the final model regression results are shown in Table IV, P value of the variable coefficients of $\ln X_2$ and $\ln X_5$ are significant, therefore the regression equation of the In model formula (10) are as follows:

$$ \ln Y = 12.537 + 0.939 \times \ln X_2 + 0.854 \times \ln X_5 $$

(10)

Table 4  The final regression results of the model

| Parameters | Coefficient | Standard Deviation | T test | P values |
|------------|-------------|--------------------|--------|----------|
| Constant   | 12.537      | 1.079              | 11.613 | 0.000    |
| $\ln X_2$  | 0.939       | 0.200              | 4.688  | 0.002    |
| $\ln X_5$  | 0.854       | 0.244              | -3.505 | 0.010    |

3.5. Results and Discussion

3.5.1. The impact of research and development spending on export trade

It can be seen from Table IV that the research and development expenditure of Guangdong Province has a significant positive relationship with the export trade of goods. Every 1% increase in research expenditure will promote the increase of 0.93% in the export trade of goods. With the progress of scientific and technological research and development, the computer, software and other industries in the tertiary industry dedicated scientific research and innovation is also promoting the export of high-tech products. In recent years, the export of high-tech products in Guangdong has maintained a steady growth, with an average annual export of more than 210 billion US dollars, accounting for more than 30 percent of the province's total export.

3.5.2. The influence of tertiary industry development level on export trade

According to the model, the development level of the tertiary industry also has a significant positive relationship with the export trade of goods. The contribution rate of the tertiary industry to the export will increase by 0.854 when the level of the tertiary industry increases by one index. In the process of developing the tertiary industry, the asset-light and asset-heavy mode of the tertiary industry has the characteristics of short cycle and quick results, which will bring a competition mechanism to small and medium-sized enterprises. Through competition, the service quality can be improved, the cost can be reduced and the quality of export trade can be improved.

3.5.3. The influence of other variables

According to the results of the unit root test, the variables selected by the regression are $\ln X_1$, $\ln X_2$, $\ln X_3$, $\ln X_5$, $\ln X_6$ respectively. Therefore, a better regression model is obtained by combining the
The final regression model results are shown in Table V, with F value of 15.24 and P value of 0.0242, indicating that the model is relatively significant on the whole.

Table 5 Final regression model results

|       | F(5.3) | P       | R-squared | Adj R-squared | Root MSE |
|-------|--------|---------|-----------|---------------|----------|
| value | 15.24  | 0.0242  | 0.9621    | 0.8990        | 0.02387  |

In addition, the coefficient of determination of the model is 0.9621, and the coefficient of determination of the modified model is 0.8990, indicating that the interpretation of the model is relatively high and the goodness of fit is good. Therefore, the regression equation (11) of the model is as follows:

\[
\ln Y = -1.05 \ln X_1 + 1.86 \ln X_2 + 1.19 \ln X_3 + 2.33 \ln X_5 + 1.16 \ln X_6 + 17.84
\]  

(11)

Through stepwise regression to build the model, the tertiary industry employment, GDP per capita, foreign investment, the added value of producer services coefficient of p values have no significance, which are no significant effects of export trade in Guangdong Province. According to the unit root test and co-integration test, there is a stable co-integration relationship. Since the sample time span is 10 years, there is no significant relationship between these variables in a short period of time in recent years, but in terms of long-term effect, there will be a mutually stable relationship. Other variables have a trend effect, but the long-term effect is not obvious. Research expenditure and the development level of the tertiary industry will significantly affect Guangdong's export trade.

4. Conclusions

Using the data of Guangdong Province from 2010 to 2019, this paper, based on the diamond model, establishes the number of employment in the tertiary industry, research expenditure, per capita regional GDP, foreign investment, the development level of the tertiary industry, and the added value of producer services as the indicators of the tertiary industry to analyze the export trade measurement model of Guangdong Province.

Through the research and results of the impact of the tertiary industry on the export trade in Guangdong Province, we can know that research expenditure and development can obviously stimulate the export trade. In the future development, it is suggested to cultivate science and technology service industry to promote service export trade, and improve the development of new information technology such as cloud computing and big data, which can also promote the service-oriented transformation development of manufacturing industry. At the same time, Guangdong Province should focus on the Guangdong-Hong Kong-Macao Greater Bay Area, deepen urban cooperation, policy linkage, participation and penetration of goods trade and other dimensions, promote the development of the tertiary industry in the Greater Bay Area, and further lay a solid foundation for the growth of foreign trade.

Acknowledgment

This research was supported by The Natural Science Foundation of Guangdong Province (Grant No. 2015A030313642), and Science and Technology Planning Project of Guangdong Province (Grant No. 2014B040404072).

References

[1] Jin Chaohui, and Zhu Mennan. "The Impact of RMB Real Exchange Rate Changes on Export Trade." International Trade Issues. 05(2021):143-160.
[2] Wang collinearity's Export Trade: A Case Study of China's High-tech Export." Price Monthly. 05(2021):48-53.
[3] Gan Jingyun. "The Influence of Human Capital on Complexity of Service Export" Modern Marketing(Next Issue). 04(2021):48-49.
[4] Rui Mingjie. "New diamond model of industrial competitiveness ". Social Sciences. 04(2006):68-73.
[5] Zhuang, H., & Huang, J., & Chen, J. (2009). An empirical analysis of China's service trade competitiveness based on the "diamond model". Finance & Trade Economics (03), 83-89.