Analysis and Research on the Correlation Between Population and Economic Development Based on Inter Provincial Panel Data

Xuanhang Ma, Chun Luo*

School of Sciences, Shanghai Institute of Technology, Shanghai, 200235, China.

Abstract: This paper collects the population and economic development related index data of various provinces and cities, and makes a comparative analysis by using panel data regression and multiple linear regression of various provinces and cities. It is found that in recent years, the population of China has been increasing and the growth rate has slowed down, which has driven the economic development such as per capita GDP under certain conditions, but the development level of various provinces and cities is quite different, Different industrial structure adjustment policies should be adopted according to the industrial development of various provinces and cities in China.

Keywords: Population; Economic Development; Regression

Introduction

The interrelationship between population and economic development has always been a key topic of concern for various experts and scholars. In recent years, with the annual increase in gross domestic product, China's population has also increased year by year. However, in recent years, the birth rate of China's population has shown a fluctuating trend, And the stability of the three major industrial structures will also have a certain impact and impact on economic development.

1. Research objects and methods

This article takes various provinces and cities across the country as the research object to analyze the interrelationship between population size and economic development. By selecting a large number of indicators for population size and economic development, the correlation between population size and economic development is analyzed from a dual perspective of time series and cross-section, avoiding the single problem of one-dimensional data.

2. Analysis of the correlation between population size and economic development in various provinces and cities

2.1 Establishment of an Evaluation Index System for Population and Economic Development

Considering the stability of indicators and the accuracy of the evaluation system, the range method [1] is used to standardize the obtained data, eliminate the impact of dimensionality, and make the relevant data more representative. The formula is as follows:

$$X'_{ij} = \frac{X_{ij} - \min\{X_{1j}, X_{2j}, X_{3j}\cdots\cdots, X_{nj}\}}{\max\{X_{1j}, X_{2j}, X_{3j}\cdots\cdots X_{nj}\} - \min\{X_{1j}, X_{2j}, X_{3j}\cdots\cdots, X_{nj}\}}$$

The following standard data is still represented by X_{ij}.

This article takes the population size and related economic indicators of various provinces and cities across the country as the research object. The relevant data of population and economic indicators are queried through statistical yearbooks[2], and correlation analysis is conducted using Eviews software. The total population reflects the current population size of a province or city, and is a manifestation of population growth in a regional context. There are many indicators to measure the economic development of a region. Based on the understanding of the economic development status and the stability of economic coordination in this article, four indicators are selected: the output value of the primary industry, the output value of the secondary industry, the output value of the tertiary industry, and the per capita GDP.

2.2 Empirical Model Analysis

In this article, the explanatory variables selected are the population of each province and city, and the explanatory variables are (per capita GDP), (primary industry output value), (secondary industry output value), and (tertiary industry output value). By using the Eviews software, it can be seen that the skewness coefficient values of population, per capita GDP, primary industry output, secondary industry output, and tertiary industry output in various provinces and cities across the country are all greater than zero, showing a right leaning trend; The kurtosis coefficients of population and per capita GDP are all less than 3, showing a relatively "thin tail" shape, while the kurtosis coefficients of primary industry output, secondary industry output, and tertiary industry output are all greater than 3, showing a relatively "thick tail" shape.

Conduct panel regression analysis on the correlation between the population of various provinces and cities in China and the output value of the primary industry, the secondary industry, the tertiary industry, and per capita GDP.[3] The fixed effects model refers to a regression model where the intercept changes with each explanatory variable, but the slope coefficient remains unchanged. For different longitudinal time series, only the intercept term is different. The formula is as follows:

$$y_{it} = \lambda_i + \sum_{k=2}^{k} \beta_k x_{kit} + \mu_{it}$$

In order to judge whether the regression equation of panel data is a fixed effect model, Eviews software is used to analyze through F-test and Hausman test. The original assumption of F is $H_0 : \lambda_1 = \lambda_2 = \lambda_3 = \cdots = \lambda_{N-1} = 0$,the formula is as follows:

$$F = \frac{(SSE_\gamma - SSE_\mu)/(N-1)}{SSE_\mu/(NT-N-K)} \sim F(N-1, NT-N-K)$$

The results show that the regression coefficient of the panel regression model has passed the F-test and Hausman test, that is, the original hypothesis is rejected, indicating that the regression model of population and economic development belongs to the individual fixed effect model. The expression of the panel regression model is as follows:

$$Y = 0.43\,X_1 + 0.04\,X_2 + 0.64\,X_3 - 0.24\,X_4 + 0.09$$

2.3 Correlation analysis between population size and economic development based on multiple linear regression

The population changes in provinces and cities such as Shandong and Jiangsu have a significant impact on the per capita GDP and the output value of the tertiary industry, and exhibit a positive correlation, while there is basically no direct impact on the output value of the primary and secondary industries; The population changes in Tianjin, Shanghai, Zhejiang, and other places are positively proportional to the output value of the tertiary industry. Most provinces and cities have a negative correlation between the population and the output value of the secondary industry, and their relationship with per capita GDP is not significant. This indicates that these provinces and cities are relatively developed, and a small number of population changes are not directly related to the level of economic development; The population changes in Liaoning, Heilongjiang, and Jilin provinces and cities are positively correlated with the output value of the tertiary industry, while showing different negative correlations with other economic indicators; The population of most provinces and cities in the central region is negatively correlated with economic indicator variables, indicating that changes in

the population of these provinces and cities have a significant adverse impact on the level of economic development, and even excessive population growth can lead to a sharp decline in economic level; Looking at the economic development of provinces and cities in the northwest region, it can be found that the population size of these provinces and cities has little impact on economic development, and their estimated coefficients are all within 0.2.

2.4 Comparison of analysis results based on two types of regression analysis

In a regression model, multiple determinable coefficients are a monotonically increasing function of the number of explanatory variables. On this basis, degrees of freedom can be selected to modify the sum of residual squares and regression squares in multiple determinable coefficients, thereby introducing the modified determinable coefficients, which are calculated as follows:

$$\overline{R}^2 = 1 - \frac{\sum e_i^2 /(n-k)}{\sum (Y-\overline{Y})^2 /(n-1)}$$

$$\overline{R}^2 = 1-(1-R^2)\frac{\mathrm{n}-1}{n-k}$$

In actual analysis, the closer the determinability coefficient and the modified determinability coefficient are to 1, the better the model fitting effect. However, it can only indicate that the explanatory variables in the model have a greater joint impact on the dependent variable, and cannot directly indicate that the various explanatory variables in the model have a greater impact on the dependent variable. Therefore, further judgment is needed to analyze the model, and F-test[4] is used to determine whether the model is significant. There is also a certain variable relationship between the F-statistic and the determinable coefficient, as follows:,

$$F = \frac{n-k}{k-1}\cdot\frac{R^2}{1-R^2}$$

The results showed that the determinability coefficient and modified determinability coefficient of panel regression were both close to 1, indicating a good fitting effect. However, the results of the multiple linear regression test of each province and city are slightly worse than the panel data, which further reflects that the panel regression analysis reduces the error of one-dimensional data to a certain extent; Comparing and analyzing the results of the f test with the corresponding p-value, the panel regression results are less than the critical value, and the p-value is close to 0, indicating that the model has passed the test, indicating that the fitting effect of the model is significant, and the relationship between population size and economic development level is strong. However, the results of the multivariate linear regression test in various provinces and cities are somewhat unsatisfactory, with Chongqing, Gansu The regression relationship between the population size and economic development level of Hebei and other provinces and cities has passed the test, and the regression model of Yunnan, Zhejiang, Shanghai and other provinces and cities has not passed the F-test. In summary, the panel regression estimation model is more in line with practical needs.

In the panel data regression, the population and GDP per capita are significantly positively correlated, while in the provinces and cities analyzed separately, the population and GDP per capita in some provinces and cities are weakly negatively correlated. Secondly, to analyze the correlation between population size and the three major industries, it should be considered that China is an agricultural country with a large base of farmers, so a small change in population size has little impact on the output value of the primary industry. As for the impact of the tertiary industry, due to the different levels of economic development in different provinces and cities, there are also significant differences in demand for the tertiary industry. Therefore, for relatively developed provinces and cities, The increase in population has just promoted the development of the tertiary industry, and for provinces and cities with lower development levels, the excessive increase in population has actually led to the phenomenon of the tertiary industry slipping back.

3. Conclusion

The output value of the first, second, and third industries has a more significant driving effect on population growth, especially the growth of the output value of the first industry can promote the increase of population. From the correlation between population size and economic development level in each province and city, it can be seen that there are significant differences in the proportion of population size to economic development for different provinces and cities. Combined with the results of cluster analysis, it can be

seen that the economic development in the central region is good, while the economic growth rate in the western region is relatively slow.

References

[1] Lan YP, Zhang Y. Population growth, urbanization process, and China's economic development [J]. International Finance, 2020 (04): 31-37.

[2] Guo CW, Xun HF, Ma Y. Identification of magnetic source boundaries using the correlation coefficient method of standard deviation between horizontal and vertical derivatives [J]. Geology and Exploration, 2020,56 (02): 418-426.

[3] Wang C, Zhang Y, Li XG, Dong HL. Improved Particle Swarm Optimization Algorithm and Its Application in Cluster Analysis [J/OL]. Journal of System Simulation: 1-12.

[4] Chen LF. In depth discussion of range method in the evaluation of standard uncertainty A [J]. Journal of Metrology, 2019, 40 (02): 347-352.

About authors:

Ma Xuan Hang: (2000-), female, from Langfang, Hebei province, Han nationality, with a master's degree, the main research direction is financial mathematics and statistics.

Luo Chun: (1966-), male, born in Hefei, Anhui province, Han nationality, professor, doctor, master tutor, the main research direction is experiment design and entrepreneurship guidance.