

Research on the Impact of Population Structure Changes on the Economic Development of Guangdong Province: Based on Principal Component Analysis

Han Gong*, Jiapeng Li and Zhanhui Huang

School of Economics, Guangzhou College of Commerce, Guangzhou 511363, China.

Abstract

The rapid economic development provides the basis for population growth, and population growth will cause changes in the population structure, which in turn will bring many opportunities and challenges to social and economic development. Taking Guangdong Province as an example, this article collects relevant data on the five major indicators of population age structure, population cultural structure, population urban-rural structure, population industrial structure, and economic development from 2000 to 2019, and uses principal component analysis to analyze the data. Through analysis, it is found that the proportion of the working-age population in the total population, the proportion of the population with high school (including secondary vocational) and above in the total population and the change in the proportion of employment in the tertiary industry has a positive impact on the change in Guangdong's per capita GDP; The number of years of education is the proportion of the population of junior high school and below in the total population, the proportion of the rural population in the total population, and the change in the proportion of the employed population in the secondary industry will have a negative impact on the changes in Guangdong's per capita GDP. Finally, in response to this result, this article also puts forward suggestions in four aspects: population age structure, population cultural structure, population urban and rural structure, and population industrial structure. I hope these suggestions can provide references for sustainable economic development.

Keywords

Population Structure; Economic Development; Principal Component Analysis.

1. Introduction

As the largest economic province in Guangdong, the improvement of its economic level has prompted a large number of people to flow into the local area. According to statistics, the permanent population of Guangdong Province at the end of 2019 was 115.21 million, an increase of 1.75 million over the same period. The average permanent population increased by 1.594 million in the past five years. The rapid population growth has not only brought tremendous power to the economic development of Guangdong Province, but also has an impact on the changes in the population structure of Guangdong Province. The changes in the population structure have an impact on economic development to a certain extent. Therefore, studying the population structure and the impact of its changes on economic development is particularly important. Only by clarifying the relationship between changes in the population structure of Guangdong Province and regional economic development can we better grasp the lifeline of economic development and improve the level of economic development.

2. Literature review

2.1. The impact of gender structure of population on economic development

Regarding the impact of the gender structure of the population on economic development, different scholars may have different views. The gender structure of the population will have a direct impact on the birth rate and the reproduction process of the population. Furthermore, according to the differences in the needs of different industries for different genders, changes in the gender structure of the population often induce changes in the industrial structure of the population, and ultimately affect economic development. In addition, the imbalance between males and females may cause social problems and inhibit economic development.

2.2. The impact of the age structure of population on economic development

Scholars such as Weiguo Wang (2004) cited economic viewpoints and pointed out that the proportion of the working-age population in the total population is in proportion to the labor supply, and an increase in labor supply will stimulate economic growth. [1] Through analysis, Jihong Zhang (2006) claimed that due to the implementation of the family planning policy in the 20th century, the number of newborn children has decreased, resulting in a decrease in the number of working-age populations, which will inevitably affect the supply of labor and thus affect economic development. [2] On the other hand, under the background of the implementation of the two-child policy, the social burden coefficient is gradually rising, which will lead to more raising consumption in the short term, and will provide a guarantee for the supply of labor in the long run.

2.3. The impact of the cultural structure of population on economic development

Yue Zhao (2011) discussed the economic benefits of improving the cultural quality of the population from two aspects. First of all, the cultural quality of the population determines the level of science and technology and labor productivity, and its improvement will promote economic development. Moreover, the improvement of the cultural quality of the population can convert knowledge reserves into productivity and drive the transformation of production methods, which is the key to changing the economic growth model. [3]

2.4. The impact of the urban and rural structure of population on economic development

Peng Xiao (2008) established the Granger causality test and found that the proportion of the urban population in the total population and the proportion of the rural population in the total population are both the reasons for the growth of per capita GDP. The increase in the urban population can promote the improvement of the level of urbanization and boost economic development. In addition, the transfer of rural population to cities to reduce the proportion of the rural population in the total population will not only provide surplus labor for cities, but also promote the process of urbanization and economic growth. [4]

2.5. The impact of the industrial structure of population on economic development

When it comes to the composition of the population industry structure in different regions, some scholars cling to the perspective that there are regional differences in China's population industry structure. Scholars such as Jiajun Qiao (1999) analyzed the regional differences in the structure of the population industry with the ratio of the employed population in the service industry to the employed population in agriculture and pointed out that the tertiary industry is more developed in areas with higher economic development levels, while the primary industry in areas with less economic development has the advantage. [5] Xiuhui Lin (1992) believes that

the development of tertiary industry in coastal areas will help promote economic development and accelerate the process of regional modernization. [6]

3. Empirical analysis

3.1. Variable selection and data processing

According to the above-mentioned scholars' viewpoint on the relationship between population structure and economic development, we consider building a principal component analysis model based on the 9 indicators in Table 1 and collect relevant data from the *China Statistical Yearbook (2000-2019)* and *Guangdong Statistical Yearbook (2000-2019)* to study the impact of changes in the population structure of Guangdong Province on economic development.

What needs to be explained here is that the standard deviation of the gender ratio in Guangdong from 2000 to 2019 was about 0.436 according to the gender ratio data in the *Guangdong Statistical Yearbook (2000-2019)*, which means that the gender ratio in Guangdong Province was relatively stable from 2000 to 2019, and the relationship with economic growth is not obvious. Therefore, the indicator of population gender structure is not considered to be included in the principal component analysis model.

Table 1. Related variables of the principal component analysis model

Variable	Type of indicator	Name of the indicator
Z_Y	Indicators for measuring economic development	Per Capita GDP (yuan)
Z_{X1}	Indicators for measuring the age structure of the population	The proportion of the population aged 15-64 in the total population (%)
Z_{X2}	Indicators for measuring the cultural structure of the population	The proportion of the population with a primary education level in the total population (%)
Z_{X3}		The proportion of the population with an education level of junior high school in the total population (%)
Z_{X4}		The proportion of the population with an education level of high school education (including secondary vocational school) in total population (%)
Z_{X5}		The proportion of the population with an education level of college and above in the total population (%)
Z_{X6}	Indicators for measuring the urban and rural structure of the population	The proportion of the rural population in total population (%)
Z_{X7}	Indicators for measuring the industrial structure of the population	The proportion of the employed population in the secondary industry (%)
Z_{X8}		The proportion of employment in tertiary industry (%)

3.2. Correlation test of variables

Before the principal component analysis, it is necessary to perform multicollinearity, KMO, and Bartlett sphericity tests on the variables to verify whether the principal component analysis method is feasible for analysis. The test results are shown in Table 2, Table 3, and Table 4.

Table 2. Results of multicollinearity diagnosis

Model	β	Std. Error	Standard coefficient	t-Statistic	Sig.	VIF
Constant	1.680e-015	0.020		0.000	1.000	
Z_{X1}	-0.086	0.102	-0.086	-0.845	0.416	25.105
Z_{X2}	0.318	0.310	0.318	1.025	0.327	230.039
Z_{X3}	-0.002	0.057	-0.002	-0.033	0.974	7.836
Z_{X4}	0.062	0.124	0.062	0.499	0.628	36.606
Z_{X5}	0.114	0.094	0.114	1.206	0.253	21.298
Z_{X6}	-0.055	0.081	-0.055	-0.683	0.509	15.709
Z_{X7}	0.480	0.149	0.480	3.222	0.008	53.127
Z_{X8}	1.047	0.208	1.047	5.039	0.000	103.356

Table 3. Correlation coefficient matrix

Variable	Z_{X1}	Z_{X2}	Z_{X3}	Z_{X4}	Z_{X5}	Z_{X6}	Z_{X7}	Z_{X8}
Z_{X1}	1.000	-0.893	0.592	0.821*	0.689	-0.901	0.723	0.543
Z_{X2}	-0.893	1.000	-0.286	-0.966	-0.901	0.853*	-0.585	-0.799
Z_{X3}	0.592	-0.286	1.000	0.158	-0.085	-0.378	0.544	-0.143
Z_{X4}	0.821*	-0.966	0.158	1.000	0.925*	-0.811	0.419	0.884*
Z_{X5}	0.689	-0.901	-0.085	0.925*	1.000	-0.749	0.339	0.893*
Z_{X6}	-0.901	0.853*	-0.378	-0.811	-0.749	1.000	-0.692	-0.582
Z_{X7}	0.723	-0.585	0.544	0.419	0.339	-0.692	1.000	0.004
Z_{X8}	0.543	-0.799	-0.143	0.884*	0.893*	-0.582	0.004	1.000

Table 4. KMO statistics and Bartlett sphere test value

	Value
Kaiser-Meyer-Olkin measure of sampling adequacy	0.650
Approximate chi-square	257.856
Bartlett's sphericity test Degrees of freedom	28
Sig.	0.000

Construct a multiple linear regression model of the independent variable Z_{Xi} ($i=1,2,3,4,5,6,7,8$) and dependent variable Z_Y . According to Table 2, except for the variable Z_{X3} , the Variance Inflation Factor of other variables are all bigger than 10, which indicates that the model has serious multicollinearity.

It can also be judged from Table 3 that the multiple linear regression model has multicollinearity because each independent variable has a certain degree of correlation, among which the independent variables Z_{X1} and Z_{X4} , Z_{X2} and Z_{X6} , Z_{X4} and Z_{X5} , Z_{X4} and Z_{X8} , Z_{X5} and Z_{X8} have a high degree of correlation exceeding 0.8, which can infer that these independent variables are highly correlated.

In addition, according to Table 4, the value of the KMO statistic to test the correlation between the variables is greater than 0.6, and the significance probability of the Bartlett sphere test is less than 0.05, which means that the null hypothesis (the variables are not independent of each other) should be rejected.

The above results show that there is a certain correlation between each variable. Thus, all variables can be used for principal component analysis.

3.3. Analysis of variance and extraction of principal components

Perform principal component analysis on the variable Z_{Xi} ($i=1,2,3,4,5,6,7,8$), and obtain the eigenvalues and Variance Contribution Ratio of each component as shown in Table 5.

Table 5. Interpreted total variance and principal component extraction results

Component	Eigenvalues	Initial eigenvalue		Extraction sums of squared loadings		
		Variance Contribution Ratio	Accumulated Variance Contribution Ratio	Eigenvalues	Variance Contribution Ratio	Accumulated Variance Contribution Ratio
1	5.472	68.406	68.406	5.472	68.406	68.406
2	1.796	22.445	90.851	1.796	22.445	90.851
3	0.479	5.983	96.834			
4	0.150	1.872	98.706			
5	0.043	0.537	99.243			
6	0.038	0.477	99.720			
7	0.020	0.247	99.967			
8	0.003	0.033	100.000			

Using the principal component analysis method, the original multiple independent variables can be replaced with several components to explain the influence of the independent variables on the dependent variables. Generally speaking, the Accumulated Variance Contribution Ratio of the selected components should reach 85% or more. [7] According to Table 5, the eigenvalues of the first principal component and the second principal component are bigger than 1, and the sum of the Variance Contribution Ratio of the first principal component and the second principal component is 90.851%, indicating that they can explain most of the information in the independent variable Z_{Xi} ($i=1,2,3,4,5,6,7,8$). Therefore, we consider extracting the first principal component and the second principal component to explain the influence of the independent variable Z_{Xi} ($i=1,2,3,4,5,6,7,8$) on the dependent variable Z_Y and conduct further analysis.

3.4. Factor loading matrix and eigenvector

Since we are using the principal component analysis method, we use the initial factor loading matrix for analysis. [8] The loading value of each factor in the first principal component and the second principal component shows in Table 6.

Under normal circumstances, the higher the loading value of the factor, the larger extent that the component can reflect the information of the factor.

Table 6. Initial loading value of each factor

Component	Z_{X1}	Z_{X2}	Z_{X3}	Z_{X4}	Z_{X5}	Z_{X6}	Z_{X7}	Z_{X8}
F1	0.932	-0.986	0.332	0.956	0.888	-0.921	0.620	0.765
F2	0.306	0.062	0.825	-0.234	-0.410	-0.165	0.639	-0.598

Based on Table 6, it can be seen that the factors Z_{X1} , Z_{X4} , Z_{X5} , and Z_{X8} in the first principal component have higher loading values, indicating that these factors are fully reflected in the first principal component, also the factors Z_{X3} and Z_{X7} in the second principal component have higher loading values, indicating that these factors are fully reflected in the second principal component.

Use Formula 1 [9] to calculate the feature vector of each factor in the first principal component and the second principal component to obtain the equation of the first principal component and the second principal component.

$$u_{ij} = \frac{\alpha_{ij}}{\sqrt{\lambda_i}} \quad (1)$$

What needs to be explained here is that in Formula 1, u_{ij} ($i=1,2$; $j=1,2,3,4,5,6,7,8$) represents the feature vector, α_{ij} ($i=1,2$; $j=1,2,3,4,5,6,7,8$) represents the initial loading value of each factor in the principal component, λ_i ($i=1,2$) represents the initial eigenvalue of the principal component. The calculation results are shown in Table 7.

Table 7. Feature Vector

Component	Z_{X1}	Z_{X2}	Z_{X3}	Z_{X4}	Z_{X5}	Z_{X6}	Z_{X7}	Z_{X8}
F1	0.398	-0.422	0.142	0.409	0.380	-0.394	0.265	0.327
F2	0.228	0.046	0.616	-0.175	-0.306	-0.123	0.477	-0.446

3.5. Fitting equation

By multiplying the feature vectors of each factor in the first principal component and the second principal component in Table 7 by the corresponding variable, we can obtain the expressions of the first principal component and the second principal component such as Formula 2 and As shown in Formula 3.

$$F1 = 0.398Z_{X1} - 0.422Z_{X2} + 0.142Z_{X3} + 0.409Z_{X4} + 0.380Z_{X5} - 0.394Z_{X6} + 0.265Z_{X7} + 0.327Z_{X8} \quad (2)$$

$$F2 = 0.228Z_{X1} - 0.046Z_{X2} + 0.616Z_{X3} - 0.175Z_{X4} - 0.306Z_{X5} - 0.123Z_{X6} + 0.477Z_{X7} - 0.446Z_{X8} \quad (3)$$

Using F1 and F2 as independent variables and variable Z_Y as dependent variables to establish a regression model. According to the results in Table 8, the comprehensive model of principal components is:

$$Z_Y = 0.391F1 - 0.284F2 + 9.604E - 017 \quad (4)$$

Table 8. Results of multicollinearity diagnosis

Model	β	Std. Error	Standard coefficient	t-Statistic
<i>Constant</i>	9.604E-017	0.032		0.000
F1	0.391	0.014	0.915	27.761
F2	-0.284	0.025	-0.380	-11.521

We can take formula 2 and formula 3 into formula 4, and get the fitting equation of the impact of population structure change on per capita GDP as shown in formula 5.

$$Z_Y = 0.091Z_{X1} - 0.178Z_{X2} - 0.119Z_{X3} + 0.210Z_{X4} + 0.235Z_{X5} - 0.1197Z_{X6} - 0.032Z_{X7} + 0.254Z_{X8} + 7.678e - 17 \quad (5)$$

$$R^2 = 0.982 \quad \bar{R}^2 = 0.979$$

It can be seen from the R-squared and the Adjusted R-squared that 98.2% of the total variation of the observed values of the dependent variable Z_Y is explained by the estimated sample regression model, which means that the model fully fits the sample data on the whole.

The coefficients of the independent variables in the fitting equation show that the proportion of the population aged 15-64 in the total population, The proportion of the population with an education level of high school education (including secondary vocational school) and above in the total population, and the change in the proportion of employment in the tertiary industry have a positive impact on the changes in Guangdong's per capita GDP, while the proportion of the population with an education level of junior high school and below in the total population, the proportion of the rural population in the total population, and the change in the proportion of the employed population in the secondary industry will have a negative impact on the changes in Guangdong's per capita GDP.

4. Conclusions and recommendations

The above-mentioned empirical analysis proves that changes in the population structure have a certain degree of influence on the development of the regional economy, which is consistent with the research conclusions of relevant scholars. Therefore, taking relevant measures to optimize the following four aspects of the population structure of Guangdong Province can better promote the economic development of Guangdong Province to a new level.

4.1. Optimize the age structure of the population and promote the development of population modernization

The empirical analysis results prove that the increase in the proportion of the working-age population has a positive impact on the changes in Guangdong's per capita GDP. In recent years, the working-age population in Guangdong has accounted for the majority of the proportion, but the population aging trend is becoming more and more obvious. In 2000, the proportion of the 65-year-old population in the province was 6.05%, and it has risen to 9.00% in 2019. At the

same time, the proportion of the 0-14-year-old population has also decreased. If this development trend continues, the proportion of the working-age population will gradually decrease in the future, and economic development will face challenges. Therefore, it is extremely important to promote the implementation of a comprehensive two-child policy. Although it will increase the social burden in the short term, it can guarantee labor supply in the long run, while slowing down the growth rate of population aging, and maintaining stable economic development to a certain extent.

4.2. Develop education and enhance the cultural quality of the population

Education is the foundation of prosperity, and the development of education can create new growth points for economic growth. The "Analysis of the Educational Status of the Employed Population in Guangdong" issued by the Guangdong Provincial Statistics Bureau in 2013 showed that the average schooling years of the employed population in the province in 2010 was 9.79 years, of which the average schooling years of the employed population in the Pearl River Delta region was 10.28 years, while the average schooling years of the employed population in the eastern, western and northern region is less than 9 years. It can be seen that most of the educational resources are concentrated in the Pearl River Delta, and the distribution of educational resources across the province is unbalanced. This is also the reason for the gap in the development level between the Pearl River Delta region of Guangdong Province and the eastern and western regions of Guangdong. The empirical analysis results prove that the development of high schools and higher education has a positive effect on economic development. Therefore, improving the education level of the eastern and western regions of Guangdong and coordinating the allocation of educational resources will improve the quality of the workers in the eastern, western and northern parts of Guangdong and promote regional economic development, which will ultimately provide an impetus for the economic development of the province.

4.3. Implement the rural revitalization strategy and improve the level of urbanization

Chairman Xi Jinping made an important deployment for the comprehensive implementation of the rural revitalization strategy at the 19th National Congress of the Communist Party of China. This is not only an important foundation for building a modern economic system, but also an inevitable requirement for the great rejuvenation of the Chinese nation. The empirical analysis results prove that the increase in the proportion of the rural population will hinder the growth of Guangdong's per capita GDP. And according to statistics, the urbanization level of Guangdong Province has shown a steady upward trend in the past 15 years, but the urbanization level of cities in the province has a large gap. In 2019, the urbanization rate of the Pearl River Delta in Guangdong Province reached 86.3%, while the average urbanization rate of the eastern, western and northern regions of Guangdong was 52.3%. The imbalance in the level of urbanization has led to a large influx of population into the Pearl River Delta region, further hindering the development of the eastern, western and northern regions of Guangdong. Therefore, it will help promote the level of urbanization in the region and the return of employed population, at the same time consolidate the foundation for the province's economic development by coordinating and improving the construction of urban public service facilities and transportation infrastructure in the eastern, western and northern region of Guangdong.

4.4. Improve the layout of the population industry structure and accelerate economic development

As a coastal province, Guangdong has close economic ties with inland and overseas, and vigorously developing the tertiary industry plays an important role in accelerating the modernization of the region. [10] The empirical analysis results also prove that the increase in

the proportion of employment in the tertiary industry will promote the growth of Guangdong's per capita GDP, however, the increase in the proportion of employment in the secondary industry will lead to a decrease in Guangdong's per capita GDP, since the development of the secondary industry will further aggravate the deterioration of the regional environment and hinder regional economic development. According to statistics, the population proportion of the tertiary industry in Guangdong Province in 2019 is 18.2%; 34.6%; 47.2%. The industrial structure distribution is similar to the inverted triangle-shaped population industrial structure model. The tertiary industry's contribution to the growth of regional GDP is close to 60%, which means that the tertiary industry has an extremely important position and role in promoting regional economic development in Guangdong Province. Therefore, optimizing the layout of the population industry, vigorously supporting the development of the tertiary industry in the province, and attracting more workers to engage in service industries such as finance, commerce, and technology will help provide strong momentum for economic growth.

References

- [1] Weiguo Wang, Yong Xu, Qiuying Li: A Quantitative Analysis of the Impact of China's Population Age Structure Change on Economic Development, Population and Development, Vol. 6(2004), p.1-8. (In Chinese)
- [2] Jihong Zhang: Analysis on the Influence of China's Population Structure on Social and Economic Development, Gansu Science and Technology, Vol. 1(2006), p.6-9. (In Chinese)
- [3] Yue Zhao: Discuss the Influence of Chinese Population Cultural Quality on Sustainable Economic Development (MS., He Bei University, China 2007) (In Chinese)
- [4] Peng Xiao: An Empirical Analysis of the Relationship between the Population Structure and Economic Growth in Henan Province, Chinese businessman, Vol. 15 (2008), p.219-219. (In Chinese)
- [5] Jiajun Qiao, Fengxian Lu: Problems and Countermeasures of China's Population Industry Structure, Journal of Henan University (Natural Edition), Vol. 29 (1999) No.4, p.41-45. (In Chinese)
- [6] Xiuhui Lin: Vigorously Develop the Tertiary Industry and Accelerate the Modernization of Coastal Areas, Journal of the Fujian Provincial Committee Party School of the Communist Party of China, Vol. 12(1992), p.37-37. (In Chinese)
- [7] Tiemei Gao: Econometric Analysis Methods and Modeling—EViews applications and examples (third edition) (Tsinghua University Press, China 2016) (In Chinese)
- [8] Yaoting Zhang, Kaitai Fang: Introduction to Multivariate Statistical Analysis (Science Press, China 1999) (In Chinese)
- [9] Chaofeng Li, Zhongbao Yang: The Eigenvector Calculation Problem in SPSS Principal Component Analysis, Statistics Education, Vol. 3 (2007), p.10-11. (In Chinese)
- [10] Xiuhui Lin: Vigorously Develop the Tertiary Industry and Accelerate the Modernization of Coastal Areas, Journal of the Fujian Provincial Committee Party School of the Communist Party of China, Vol. 12(1992), p.37-37. (In Chinese)