Hindawi
Mathematical Problems in Engineering
Volume 2022, Article ID 4487654, 13 pages
https://doi.org/10.1155/2022/4487654

Research Article

Statistical Analysis of the Effect of Population Quality on Residents' Income Distribution under the Data Fusion Algorithm

Lihua Wu,[1,2] Xiaomei Guo ,[3] and Xiao Zang[4]

[1]*School of Economics and Management, Tongji University, Shanghai 200092, China*
[2]*Business School, Jinggangshan University, Ji'an, Jiangxi 343009, China*
[3]*Humanities and Law School, Chengdu University of Technology, Chengdu, Sichuan 610000, China*
[4]*Law School, Chengdu University, Chengdu, Sichuan 610000, China*

Correspondence should be addressed to Xiaomei Guo; 9920090025@jgsu.edu.cn

Received 22 March 2022; Revised 10 May 2022; Accepted 27 May 2022; Published 8 July 2022

Academic Editor: Xiantao Jiang

After years of development, today's society is a prosperous era of the knowledge economy. Different from the previous period of development by consuming natural resources, people have realized that talents with technology and knowledge have become an important factor in promoting the development of social production. That is to say, the improvement of population quality has not only an impact on its own development but also an important impact on the comprehensive strength and competitiveness of its location. Population quality is the stipulation of the population in terms of quality, also known as population quality. It includes ideological quality, cultural quality, physical quality, and so on. Data fusion is a technology for the comprehensive analysis and processing of information from multiple sources. Therefore, this paper comprehensively considers the influence of ideological and moral quality and population structure. Under the algorithm of fusion data, the impact of population quality on residents' income is explored. This paper uses the analytic hierarchy process to weight each indicator and calculate the comprehensive score. This paper ranks the quality of population and economic development of each province, city, and autonomous region in China according to their high scores. The results of the experiments in this paper indicate a relevant analysis with the income distribution of the nation's residents, and the correlation is −0.9493. There exists a negative relation between the overall quality of Chinese residents and the income gap.

1. Introduction

Economic development is the foundation for promoting social development. In the era of social development full of competition and challenges, the prosperity and civilization of a country is strongly correlated with quality of its citizens as well as quantity of talents. Therefore, based on the exploration of economic development, many scholars have begun to find that economic growth in the current era has long been different from the past. Only relying on land and labor to increase, people realize the effect of population quality on economic development. The improvement of population quality can promote the progress of science and technology and the improvement of labor productivity, thereby effectively enhancing the competitiveness of products and realizing the improvement of economic benefits. On the other hand, the improvement of population quality can accelerate the speed of knowledge update to a certain extent, so as to convert more knowledge into productivity. In today's information technology era, what drives future development is the quality of workers and the quantity and quality of intellectuals. In China's economic development, it was clearly proposed to shift the economic construction to the track of science as well as technology and improve the quality of workers to build an affluent society. It is also emphasized that population quality is an important content of development. Through these explanations, it can clearly show the importance of the improvement of population quality in development.

The disparity of wealth between China's regions is serious, the quality of the population is unevenly distributed, and the income gap is gradually widening. There are many factors that affect the income of residents. But in the era of the knowledge economy, the level of population quality plays a decisive role. The research on the effect of population quality on income distribution in this paper can effectively coordinate the population size, structure, quality, and distribution of a country. On the other hand, it promotes the mutual coordination between social resources and environmental resources of the state and the requirement of taking the road of sustainable development. Therefore, it is particularly important to systematically study the status quo of China's population quality and income distribution, as well as the imbalance between regions, to explore the relationship between the two and propose corresponding policy suggestions for narrowing the income gap.

The innovation of this paper is based on two points: First, this paper is based on the basic theory of population quality and human resources. It is combined with the results of past investigations to further construct metrics. Second, in the analysis of the impact of population quality on residents' income, this paper uses the component analysis method of multivariate statistics to analyze and compare the economic development of population quality.

2. Related Work

The improvement of population quality has the ability to turn the disadvantage of the population base into an advantage for the current national development. Therefore, the study of population quality has attracted the attention of many scholars. In the developing countries, populations are mostly clustered in rural areas. Hossein et al. focused on a representative breakdown of active traveling aggregates (RACs) in rural Iran. The results show that descriptors in rural RACs are strongly influenced from the proximity of villages to the nearest cities [1]. The China Statistical Yearbook shows that both the area of grain planting and the total agricultural output are expanding, at the same time, the difference in income is narrowing in urban and rural areas. But at the same time, farmers' income is increasingly dependent on nonagricultural employment, indicating that subsidies are less important to farmers' income realization. By analyzing the relationship between farmers' living conditions and subsidies, Bai discusses farmers' attitudes and evaluations from the perspective of subsidy recipients [2]. Sabetova and Egorova examine the problem of unequal income of the population in a Russian federal state and identifies elements that influence it. Population income disparities reflect differences in how individuals and their groups participate in social processes, as well as inherent and inevitable differences in the government's distribution of the nation's wealth among citizens. He concludes that wild fluctuations in the dynamics of the effective income of the population are undesirable and that the efforts of national and regional authorities aimed at improving income policy are justified [3]. The aim of Cymbranowicz is to show the relationship between work and deprivation within a background of the phenomenon of the developing poor. It could be said that there is poverty as well as deprivation within their context. Thinking employment as a contributing element to poverty eradication and full participation in employment is the best way to combat it and fight exclusion from society [4]. Fiedler and Lividini presents a review of the Family Consumption and Expenditure Survival Survey (HCES) - a regular multi-purpose survey conducted in about 120 countries. With its household-based subnational representation, the HCES could go past the constraints of traditional data sources that have long been used to inform food policy [5]. Maurer examines legislation that applies to New York residents for income tax purposes. The individual considers that he has established permanent residence, and they may be unhappy to find that they are still a resident of New York State for income tax purposes [6]. Whether in terms of prices, wages or wealth, crossing institutional or regulatory borders is often accompanied by discontinuities. Mathä et al. find a large wealth gap between Luxembourg residence families and those commuting across borders [7]. Stewardship and repair for polluted settings involves the participation affected local residents. Recently, Mcintyre et al. aimed to tackle this gap by using cross-sectional survey data from 2,009 adult residents of communities near 13 contaminated sites in Australia [8]. González Canché provides a way to measure the wealth level of a group of nonresident students enrolled at an institution. He also creates a measure of the heterogeneity of the nonresident student population by considering the wealth levels of the states in which these students reside [9]. Residential programs need to guide residents in handovers. Murphy et al. designed and implemented a novel handover education workshop aimed at improving residents' confidence and performance. Surveys before and after the workshop showed that self-confidence, comfort, and knowledge are all improved ($P < 0.001$) [10]. The research results fully demonstrate the importance that scholars around the world attach to the study of population quality. However, there are many factors that affect the research direction of population quality, and the data are huge and complex, and only some of them are explored in the research process.

2.1. The Impact of Population Quality on Income Distribution of Residents

2.1.1. The Externality of Population Quality and the Theory of Resident Income Distribution. The investment of human capital to improve the quality of the population can not only bring direct economic income but also bring certain benefits to others. Therefore, the spillover of human capital shows that population quality has the nature of public goods, which can indirectly affect the income of residents [11]. The following is a look at the spillover of population quality from the perspectives of households, enterprises, and society.

First of all, within a resident family, the parents' good educational background will improve the educational opportunities and level of the next generation, and the intergenerational transmission and mutual influence of the

population quality level will increase the income gap. In addition, if a member of the family suffers from a serious disease, the family will bear high medical expenses, which will directly affect the investment in human capital and the final population quality of other members, and thus affect the income level.

Secondly, within the enterprise, the spillover of population quality is produced when the incentive mechanism of groups and individuals achieves the interaction and mutual learning of group members. Due to the different advantages and characteristics of team members, they can share their own unique skills, knowledge, and information with each other, and they will reach a state of mutual influence in the joint work, and then improve together. Generally speaking, the more knowledge-intensive the enterprise's human capital, the stronger the spillover. The spillover of human capital will correspondingly improve the production efficiency of the enterprise, thereby increasing the income level of the employees of the enterprise. In the same way, there will also be spillovers of population quality in enterprises with poor development trends and weak competitiveness. Due to the poor production efficiency of enterprises, excellent employees will be lost. The production efficiency of enterprises is further reduced, the benefits are even worse, and the income of employees is further reduced. Therefore, the spillover caused by the mobility caused by the difference in the quality of the labor force will exacerbate the continuous widening of the personal income gap of residents.

Finally, from a social point of view, a country with a higher population quality level also has higher productivity of its social members and a stronger ability to adapt to job demands. It can achieve a more reasonable allocation of human resources, the unemployment rate will be reduced, and members of society can do their best. It can not only increase the stability and civilization of the society but also reduce the burden of relief and subsidies for the unemployed. It indirectly promotes social and economic growth. Conversely, if the unemployment rate increases and the social order deteriorates, the state needs to increase unemployment benefits accordingly. The funds used to improve the quality of the population will decrease, labor productivity will decrease, and the income level of the entire society will decline. Therefore, the spillover of population quality will cause income gaps among members in different regions and countries.

About the basic concept of population quality, British scholars believe that it refers to the health status, technology, and ability level of the country as a unit. The earliest and most influential definition of population quality in China is that it is people's understanding and the ability and conditions to transform the world, but this cognition is controversial [12]. Until 2006, a widely recognized research description was put forward, which refers to the social functions and influences displayed by the structure and composition of the population under certain historical conditions.

According to the current academic research on the content of population quality, it generally includes "two-element method," "three-element method," and "four-element method." The traditional population quality often refers to the physical quality of the population [13]. The quality of the population is usually called morality, intelligence, and physique. Ideological quality is the state of consciousness that governs people's behavior, cultural quality is people's ability to understand and transform the world, and physical quality is the natural condition and basis for population quality. Population group is the unity of quality and quantity, and the two are interrelated and restrict each other.

2.1.2. Mathematical Derivation of Theoretical Expectations. This paper proposes to study poverty from the perspective of income distribution changes, because income distribution changes not only include income growth and income distribution but also include changes caused by heterogeneity. If we simply consider the impact of income changes on poverty from the perspective of income growth and income distribution, it is likely to overestimate or underestimate the growth and dispersion effects of income impact on poverty. Therefore, it is more reasonable and practical to study poverty from the perspective of income distribution changes by introducing the factor of heterogeneity [14].

Initially, most scholars only studied poverty decomposition from the perspective of income growth and income distribution. Some scholars proposed a poverty decomposition method with path dependence. The main formula is as follows:

$$\xi_{A-1} - \xi_A = \left(\xi_{(A+1)}, A - \xi_A\right) + \left(\xi_{(A\neq 1)} - \xi_{(A=1)}, A\right). \quad (1)$$

The poverty line used in this decomposition method is selected based on the base period, and income is decomposed into growth effects and redistribution effects. After that, some scholars believe that the poverty line should be selected according to the reporting period, that is, the current poverty level needs to be considered. Therefore, the proposed poverty decomposition formula is given by

$$\xi_{(B-1)} - \xi_B = \left(\xi_{(B\neq 1)} - \xi_{B(B\neq 1)}\right) + \left(\xi_{B(B\neq 1)} - \xi_B\right). \quad (2)$$

The path-dependent poverty decomposition method is intuitive and easy to understand. However, the income growth effect and the redistributive effect are not equal to the total effect of income on poverty. In addition, when different poverty lines are selected for decomposition, the results obtained by this method are inconsistent and cannot be compared and analyzed, and the decomposition method must be improved.

Therefore, based on the poverty decomposition method with path dependence, a residual term is introduced to explain the unexplained part of poverty in the decomposition method except for the income growth effect and redistribution effect:

$$\xi_{(C-1)} - \xi_C = \left(\xi_{(C+1)}, C - \xi_C\right) + \left(\xi_{C(C\neq 1)} - \xi_C\right) + E. \quad (3)$$

The residual term of this method is composed of the interaction term of average income and income distribution.

Although path dependence is eliminated, the existence of the residual term is controversial. Based on this, some scholars use the Shapley value decomposition method based on strict game theory, combine the influence effects of income changes and distribution changes in the base period and the reporting period, and take the mean value to decompose the poverty index. It distributes benefits based on the Shapley value, which avoids egalitarianism in distribution. It is more reasonable and fair, and it also reflects the process of mutual games among the members. The decomposition formula is as follows:

$$\begin{aligned}\xi_{(x-1)} - \xi_x &= \left(\xi_{(x\neq 1)}, X - \xi_x\right) + \left(\xi_{(x\neq 1)} - \xi_{x(x\neq 1)}, X\right),\\ \xi_{(x-1)} - \xi_x &= \left(\xi_{(x\neq 1)} - \xi_x, (x\neq 1)\right) + \left(\xi_x, (x\neq 1) - \xi_x\right).\end{aligned} \tag{4}$$

Take the average from the formula to get:

$$\begin{aligned}\xi_{(x-1)} - \xi_x = {} & \frac{\left(\xi_{(x\neq 1)}, X - \xi_x\right) + \left(\xi_{(x\neq 1)} - \xi_{(x\neq 1)}, X\right)}{2}\\ & + \frac{\left(\xi_{(x\neq 1)} - \xi_x, (x\neq 1)\right) + \left(\xi_x, (x\neq 1) - \xi_x\right)}{2}.\end{aligned} \tag{5}$$

This method solves both the residual term problem and the symmetry problem, so it is widely used in poverty decomposition research [15].

2.1.3. Establishment of an Indicator System for Evaluating Various Regions. In the description, this paper expounds on the basic theory of population quality in detail and makes a further framework for the evaluation index system of population quality in the description. Based on this framework, this paper makes a detailed analysis of its factors, which is the important content of this section.

The labor force is one of the power sources of population development, and the quality of the workforce is profoundly associated both with quality and with the body. It also has a close relationship with the quality of the material produced which can further affect social and economic development. Therefore, the importance of physical fitness can be seen. From a regional perspective, if people's physical fitness is good, they can provide more excellent human resources and provide advantages for the region's economic prosperity and scientific and technological research and development. Otherwise, it will cause social problems related to the population and create obstacles to development. The physical quality includes the natural and social attributes of a region, that is, the average life expectancy, crude mortality, and neonatal mortality.

Figure 1(a) presents statistics on life expectancy and crude death rates in 2000 and 2010. From the point of view of the probability of crude death, there is little difference between regions. The highest crude death rate in 2000 was about 7.8% in Yn, and the lowest was 5.2% in Hn. After ten years of development, the highest and lowest death rates in 2010 were 6.9% in Tb and 4.2% in Gd, respectively. However, life expectancy per capita is stable, with a life expectancy exceeding 80 in developed cities, such as Sh. The Gd region, which is expected to be the shortest, is also above 66. Figure 1(b) lists infant and child mortality rates in 2000 and 2010. Compared with the crude mortality rate, there are serious regional differences in the infant mortality rate. In 2000, the infant mortality rate in Bj was only 3.8‰, but the mortality rate in Yn was as high as 70.3‰. With the development of the social economy, the infant mortality rate in Bj dropped to 2.1‰ in 2010, while the high mortality rate in Yn remained at 57.8‰. With the improvement of living standards, and the improvement of medical conditions, the human quality of each region has been significantly improved.

Scientific and cultural quality needs to be accumulated over a long period. As the main manifestation of human transformation of the world and cognition, it is also the core of population quality and a huge driving force for social and economic development. It can guarantee the innovation of production technology and promote social development by supplying excellent scientific research talents to society. The criteria for measuring the cultural quality of the population include the probability of illiteracy, the ratio of the number of people with a college degree or above, and the average number of years of education. The illiteracy rate is the proportion of the population who cannot read and write beyond school age.

Figure 2 shows the statistical map of population quality. Figure 2(b) shows the illiteracy rate in the region. It is easy to see that after ten years of development, the quality of the residents has improved significantly. However, from the horizontal comparison, there are serious problems in the quality of population culture in each region, and the quality of culture shows a serious bifurcation. In terms of illiteracy rate, the illiteracy rate in Tb was as high as 47.3%, while the illiteracy rate in Bj was only 4.9% in 2000. There is a huge difference between the two. By 2010, illiteracy rates had declined in all regions, and essentially everyone could read and write, with most provinces having illiteracy rates below 5%. The lack of highly educated people is further illustrated by the fact that the percentage of the population with higher education is much higher in Bj than in any other status, with most provinces having a rate of less than 4%. After a decade of development, there has been a significant improvement, but the polarization is still serious, with the majority of the population below 10%. In 2000, the average years of education in Bj reached 9.6 years. At this time, the average age of Tb was only 3 years. Even after a decade of development, when Bj improved to 11.6 years, Tb was still less than 6 years old and only at the elementary school level.

The quality of labor skills of the population is the main content of the quality of the population, and it is the supplement of the cultural quality of the population. The cultural quality of the population focuses on the ability of theoretical knowledge. And it focuses on people's practical ability, emphasizing the application of scientific and cultural knowledge to practice to realize its value, and the combination of theory and practice provides an important guarantee for the country to enhance its competitiveness and stamina. The commonly used indicators include working-

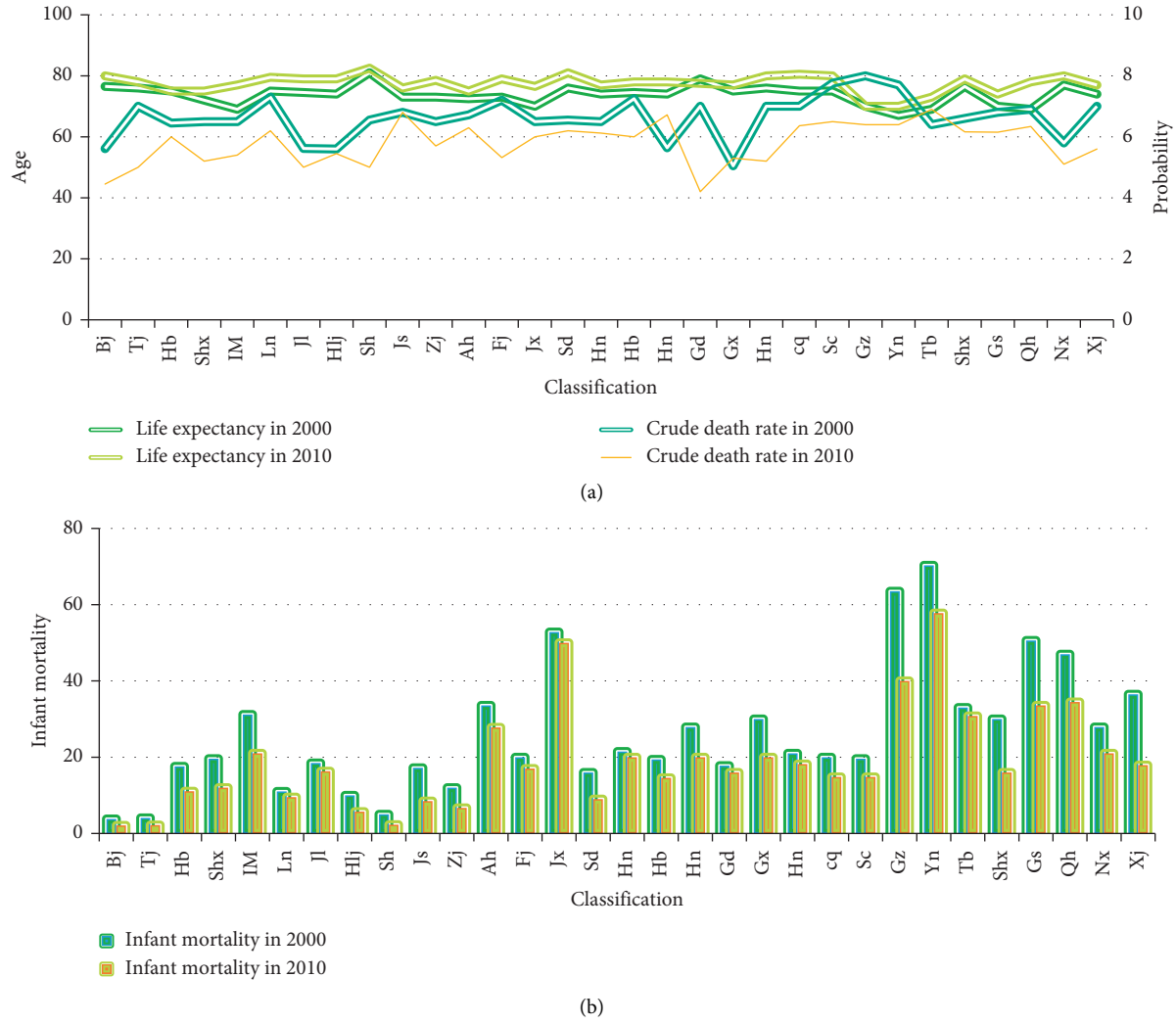

Figure 1: Comparison of changes in the physical fitness of the population in different regions. (a) Life expectancy and crude mortality. (b) Infant mortality.

age population, number of patent applications per 10,000 people, etc. It fully considers the practicality of labor skills, and two indicators were chosen to measure the R&D full-time equity and patentee grants per 100,000 people. Full-time equivalent of R&D personnel: It is an indicator commonly used to compare human input in science and technology. It refers to the sum of the workload of personnel engaged in R&D activities throughout the year whose cumulative working time accounts for 90% or more of the total working time and the workload of part-time personnel converted from actual working hours.

As can be seen from Figure 3, the overall growth of labor skills and quality of employees in various regions of China is fast, but the phenomenon of polarization is becoming more and more serious. For example, in terms of the full-time equivalent of R&D personnel, the full-time equivalent of R&D personnel in most Chinese provinces in 2010 was more than double that in 2000. Among them, Gz Province, the fastest growing province, reached 18 times that of 2000 in 2010, but because the base is too small, it still appears to be insufficiently invested compared with developed regions. The East is far ahead of the Midwest. In terms of the number of patents granted per 10,000 people, it was generally low in 2000, and most provinces grew slowly in the past ten years. A few provinces with fast growth are concentrated in the eastern coast, and the eastern provinces are much higher than the central and western provinces. Among them, the number of patents granted in Zj is 50 times that of Tb.

The ideological and moral quality mainly includes an outlook on life, moral outlook, ideological quality, etc., which can directly affect the actual process of social and economic development. When the mainstream ideological and moral qualities of a society are in line with the requirements of social and economic development, it will promote the development of this society. Otherwise, it will hinder the development of this society. Since some index values cannot be obtained, the number of traffic accidents per 100,000 population and the crime rate are finally selected as measurement indicators.

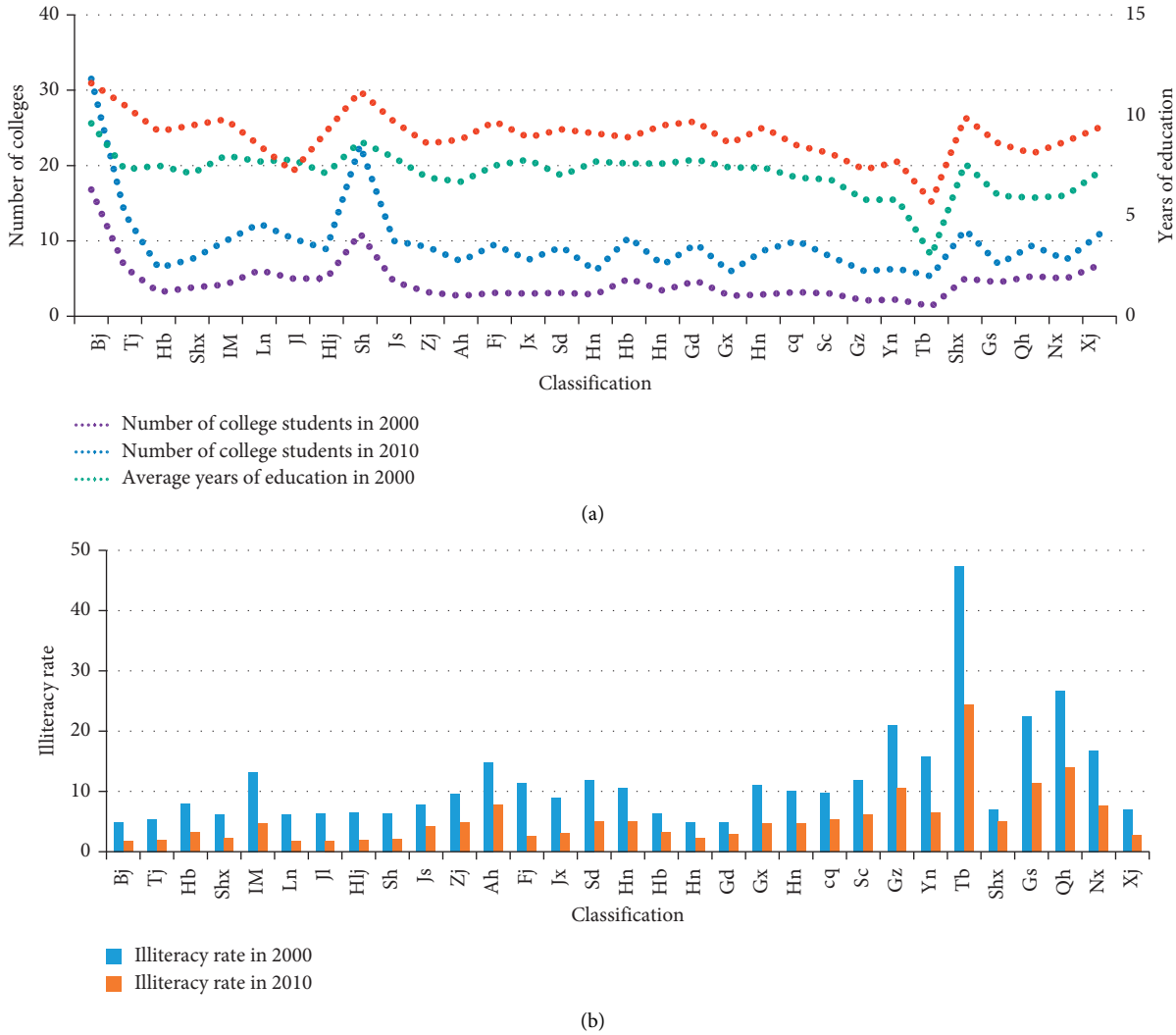

Figure 2: Basic situation of population cultural quality. (a) Percentage and percent of university students with a mean number of years of schooling. (b) Regional illiteracy rate.

Figure 4(a) is the statistical graph of the number of traffic accidents in the area, and Figure 4(b) is the crime rate statistics of the area. In terms of crime rates, crime rates in the vast majority of provinces increased over the decade. The few provinces with a drop in crime, except Bj, are concentrated in the central and western regions. It shows that the rectification of social security is effective. In addition, it also shows that while the economy in eastern China has developed rapidly, there has also been an increase in crime incidents. In terms of traffic accident fatality rates, China shows low rates in the central region and high levels in the eastern and western regions. The reasons for this are as follows: On the one hand, the eastern region is economically developed, the traffic conditions are complex, and traffic accidents occur frequently, resulting in high mortality. On the other hand, although there are fewer traffic accidents in the western region, the death rate is also higher due to the backward medical equipment.

Population composition can be considered from two aspects: the geographical distribution of the population and the natural composition of the population. Human survival depends to a large extent on the geographical distribution of the population, and the geographical distribution of the population reflects the level of social and economic development in each region. The natural composition of the population refers to the gender and age composition of the population. It is not only related to the reproduction of the population itself, but also directly affects the production activities of the national economy and social and family activities. The population sex ratio, the total population dependency ratio, and the number of nonrural households as a share of the total population were selected to measure the demographic status.

China's population structure has the following characteristics: First, the sex ratio of the population is unreasonable, and all provinces have more men than women. But what is gratifying is that the situation in most provinces has improved in the past ten years, and the sex ratio of the population has steadily declined. Only Tj and Tb have seen a significant increase in education. Second, the total dependency ratio of the population has dropped sharply, indicating that population aging is common in all provinces in

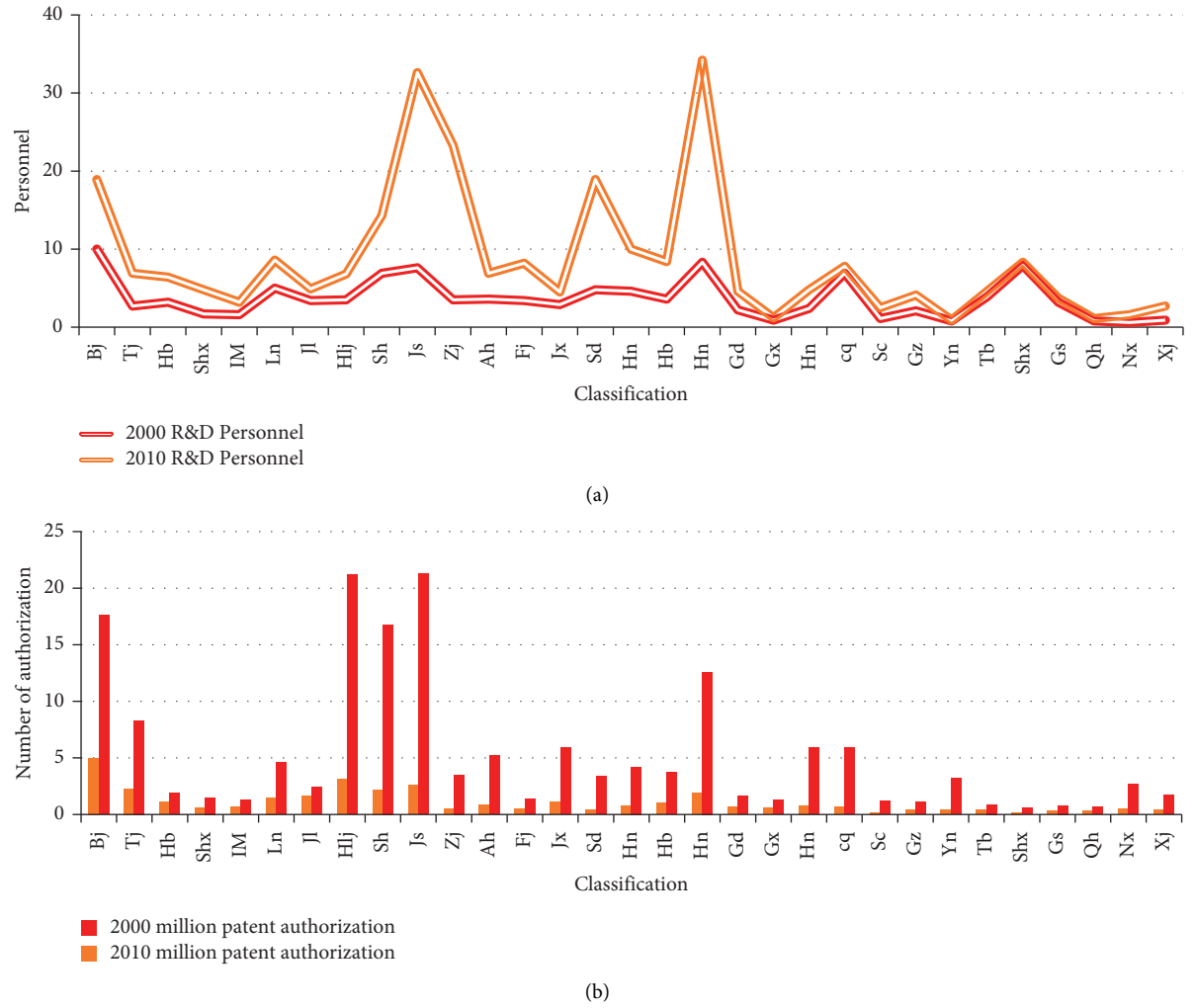

Figure 3: The basic situation of the labor skills quality of the population. (a) Full-time equivalent of R&D personnel. (b) Number of patent authorizations per 100,000 people.

China, and the trend is becoming more and more obvious. Third, China's urbanization process is steadily advancing, but the rural population is still the main population. The proportion of nonagricultural hukou in all provinces in China has increased over the past decade. Among them, Bj is close to 80%, Sh is close to 90%, but Gz, Yn, Tb, and other places in the west are still less than 20%, and most other provinces are only below 40%.

The sex ratio of the population refers to the ratio of the number of men and women in the population. The imbalance of sex ratio is related to people's traditional ideas on the one hand and may increase social instability on the other hand, leading to an increase in crime. The population dependency ratio refers to the ratio of the nonworking-age population to the working-age population. If the index value is too large, it will cause a heavy burden on society and a relative shortage of labor. The indicator of the proportion of nonagricultural hukou also reflects the overall quality of the population to a certain extent. The general quality of urban residents is generally higher than that of rural areas, so the urbanization process is accelerated.

In Table 1, since the 13 indicators selected in this paper are inconsistent in dimension, to avoid the effect of different dimensions on the results, this paper first standardizes the data. In terms of this article, we choose to perform a linear transformation on the original data.

$$\text{Normalized value} = \frac{(\text{original value} - \text{minimum value})}{(\text{maximum value} - \text{minimum value})}. \tag{6}$$

However, from the indicators selected in this paper, we can find that indicators such as illiteracy rate are negative indicators, the smaller the value, the better will be the demographic quality. In this paper, we adopt a fuzzy membership function to standardize the indicators to eliminate the positive and negative points. That is, the number (membership) between 0 and 1 is used to indicate the degree to which an object belongs to a concept. When the value of the membership degree is in the interval [0-1], the maximum and minimum values are expressed as $A_{\max}$ and $A_{\min}$, and the expressions are as follows.

Figure 4: The basic situation of the labor skill quality of the population. (a) Crime rate. (b) Population traffic accident number.

Table 1: Comprehensive evaluation table of population quality.

Overall layer	System layer	Variable	Feature layer
Comprehensive quality of the population	Healthy	$X1$	Average life expectancy
		$X2$	Infant mortality
		$X3$	Crude death rate
	Scientific culture	$X4$	Illiteracy rate
		$X5$	Average years of education
		$X6$	The proportion of college stud
	Labor skills	$X7$	The R&D staff was full at that time
		$X8$	Number of patents granted per 100,000 population
	Ideological and moral	$X9$	Traffic accidents per 100,000 population crime rate
		$X10$	
	Population structure	$X11$	Population dependency ratio
		$X12$	Population sex ratio
		$X13$	Proportion of nonagricultural hukou

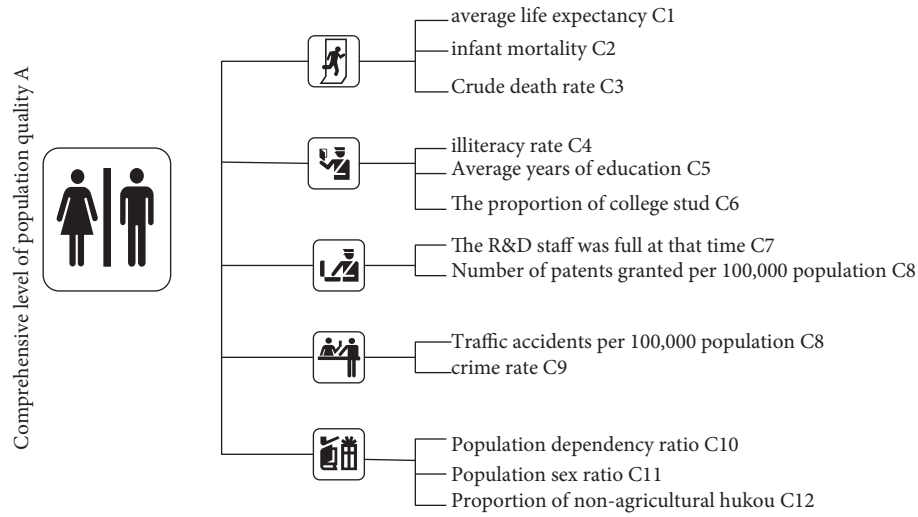

Figure 5: Progressive hierarchy of population quality.

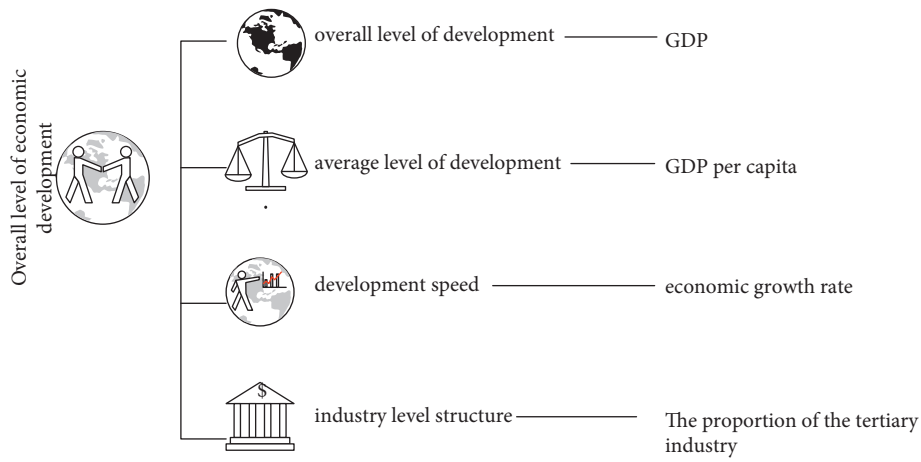

Figure 6: Progressive hierarchy of economic development levels.

In the case of positive indicators,

$$\overline{A}_i = \frac{A_i - A_{\min}}{A_{\max} - A_{\min}}. \tag{7}$$

In the case of a negative indicator,

$$\overline{A}_i = \frac{A_{\max} - A_i}{A_{\max} - A_{\min}}. \tag{8}$$

2.2. Impact of Population Quality on Residents' Income

2.2.1. Determination of the Weighted Components of the Index System. The weighted principal component TOPSIS model mainly combines the principal component analysis method and the TOPSIS method, as shown in Figures 5 and 6 for the progressive hierarchical structure of population quality and the progressive hierarchical structure of economic development level.

First, the principal component decision matrix was obtained by principal component analysis, and then the TOPSIS value function model was used to convert the low-dimensional system into a one-dimensional system, and the comprehensive evaluation value was calculated. The decision matrix is obtained by the vector norm method, denoted as $J = \{J_{in}\}$. Then, assign the matrix-vector weight $M = (m_1, m_2, \ldots, m_i)$ to form a weighted canonical matrix $K = \{K_{in}\}$, where

$$K_{in} = M_{in} \times J_{in}. \tag{9}$$

Confirmation for positive understanding A^+ and negative understanding A^- is expressed as

$$\begin{aligned} A^+ &= \max(A_{in}), \\ A^- &= \min(A_{in}). \end{aligned} \tag{10}$$

After that, by calculating the distances L_i^+ and L_i^- from the data samples to the positive and negative understanding

Table 2: The contribution of variance and cumulative variance of each component.

Element	Eigenvalues	Standard deviation	Variance contribution rate, %	Cumulative variance contribution rate, %
1	2.75	2.75	53.8	53.8
2	132	132	12.5	66.3
3	1.23	1.23	9	75.3
4	0.97	0.97	6.8	82.1
5	0.93	0.93	6.1	88.2
6	0.76	0.76	4.1	92.3
7	0.58	0.58	2.4	94.7
8	0.46	0.46	1.5	96.2
9	0.41	0.41	1.2	97.5
10	0.38	0.38	1	98.5
11	0.33	0.33	0.77	99.2
12	0.24	0.24	0.4	99.6
13	0.1	0.1	0.9	100

Table 3: Spatial measured impact of population quality on resident income.

Variable	Regression coefficients	Standard deviation	t	P
GDP per capita	0.23	0.11	2.14	0.033
Constant term	−0.11	0.06	−1.95	0.051
Average life expectancy	0.18	0.095	1.86	0.062
Average years of education	0.22	0.108	1.9998	0.046
Patents granted per 100,000 population	0.223	0.075	2.98	0.003
Crude death rate	0.057	0.061	0.95	0.344
Proportion of nonagricultural households				
R^2	0.939	Log likelihood	42.62	
Akaike info criterion	−71.25	Schwarz criterion	−61.21	

and then calculating the ideal proximity L_i, that is, the comprehensive evaluation index, the process is as follows:

$$L_i^+ = \sqrt{\sum_{i=1}^{m} \left(A_{in} - A_i^+\right)^2},$$

$$L_i^+ = \sqrt{\sum_{i=1}^{m} \left(A_{in} - A_i^-\right)^2}, \tag{11}$$

$$L_i = \frac{L_i^-}{L_i^+ + L_i^-}, \quad i = (1, 2, \ldots, n).$$

Here, L_i is in the interval of [0-1], the value is nearer to 1, the more effective the solution is.

The cumulative variance contribution of the first four principals reached 82.1% using R language programming for analysis of principal variance. It means that the first four principal components can contain 82.1% of the information in the original index. To a certain extent, it can replace the original index for dimensionality reduction analysis, so the first four principal components are selected. The percentage contribution of various main constituents is listed in Table 2.

2.2.2. Spatial Quantitative Analysis of Resident Income and Population Quality. The average life expectancy at the level of physical health in Table 3 passed the significance test with a significance level of 0.1, it shows that average life expectancy significantly affects the income of the population. If the average life expectancy (normalized) changes by 1 unit, the GDP per capita of the population (normalized) will change by 0.18 units. The average educational expenditure per 10,000 people, which represents the quality of science and culture, passed the significance test of 0.05. It shows that the quality of science and culture has a great effect on income of the population. When the average education expenditure per 10,000 people (after normalization) changes by 1 unit, the per capita GDP (after normalization) will change by 0.22 units. The number of patents granted per 100,000 population representing labor skill quality and the proportion of nonagricultural hukou representing the population structure both passed the significance test with a significance level of 0.01, indicating that these two population quality indicators have a significant impact on per capita income. For each additional unit (normalized) per capita GDP will increase by 0.223 and 0.4 units (normalized), respectively. The P value of the crude divorce rate, which represents ideological and cultural quality, is 0.34, which fails the significance test, but its regression coefficient is positive, indicating that ideological and cultural quality has a certain positive impact on residents' income. This variable is not significant, the reason may be that when the indicator system is established, the crude divorce rate cannot fully represent the ideological and cultural quality of residents due to the availability of data.

The fitted model in Table 4, $R^2 = 0.95$, indicates that the model fits well, and the spatial autocorrelation coefficient $\lambda = 0.8975$. It passed the significance test with a significance level of 0.01, indicating that there is indeed a strong correlation between regions. In terms of the regression coefficients of the specific explanatory variables, except for the

Table 4: Spatial error estimation results of population quality on resident income.

Variable	Regression coefficients	Standard deviation	t	P
Constant term	−1.3	0.042	−3.13	0.002
Average life expectancy	0.302	0.075	4.004	0.0001
Average years of education	0.28	0.091	3.13	0.002
Authorizations per 100,000 people	0.22	0.05	4.39	0
Crude death rate	0.07	0.075	2.98	0.003
Proportion of nonagricultural households	0.057	0.06	1.27	0.203
λ	−0.9	0.24	−3.71	0.0002
R^2	0.95	Log likelihood	44.35	
Akaike info criterion	−76.7	Schwarz criterion	−68.1	

Table 5: Estimated results of the spatial lag of population quality on residents' income.

Variable	Regression coefficients	Standard deviation	t	P
GDP per capita	0.036	0.032	1.15	0.25
Constant term	0.069	0.018	14.24	0
Average life expectancy	0.18	0.0321	2.18	0.029
Average years of education	−1.09	0.02	−5.44	0
Patents granted per 100,000 population	0.133	0.039	3.45	0.0006
Crude death rate	0.057	0.061	0.95	0.344
Proportion of nonagricultural households	−0.28	0.038	−7.36	0
D_1	0.88	0.042	20.93	0
D_2	−0.56	0.023	−24.83	0
R^2	0.993	Log likelihood	93.6	
Akaike info criterion	−169.2	Schwarz criterion	−156.3	

Table 6: Estimated results of the spatial error in terms of population quality and residents' income.

Variable	Regression coefficients	Standard deviation	t	P
GDP per capita	0.25	0.013	20.37	0
Constant term	0.21	0.031	6.69	0
Average life expectancy	−0.12	0.012	−10.89	0
Average years of education	0.028	0.03	0.92	0.36
Patents granted per 100,000 population	0.063	0.015	4.21	0
Crude death rate	−0.27	0.023	−11.9	0
λ	0.94	0.041	23.04	0
D_1	−0.58	0.015	−38.32	0
D_2	−0.56	0.023	−24.83	0
R^2	0.997	Log likelihood	99.8	
Akaike info criterion	−183.5	Schwarz criterion	−172.1	

crude divorce rate, all the other variables have a significant effect on GDP per capita, and they all pass the test of significance at the 0.01 level. Moreover, the regression coefficients are all positive, indicating that each variable has a positive impact on GDP per capita.

2.2.3. Impact of Population Quality on Residents' Income by Division of Economic Regions. Using GeoDa software, taking GDP per capita as the dependent variable, and dummy variables for regional factors were added according to the explanatory variables selected section, and the fitting outcomes are shown in Tables 5 and 6, where

$$D_1 = \begin{cases} 1, \text{East area}, \\ 0, \text{other}, \end{cases} \quad D_2 = \begin{cases} 1, \text{Central Region}, \\ 0, \text{other}. \end{cases} \tag{12}$$

Specifically, the fitting effect of the spatial econometric model before and after adding the dummy variables of geographical factors was compared and analyzed. After adding dummy variables, the goodness-of-fit R^2 of both the spatial lag and the error model is improved accordingly. In addition, the Log$L$ value, AIC value and the absolute value of Schwartz's discriminant coefficient are all higher than the regression model fitted before adding dummy variables. Combined with the significant results of the two geographical dummy variables, it is further demonstrated that when fitting the measurement model of the impact of population quality on residents' income, the geographical factors that divide the east, middle and west by economic regions should be fully considered. Secondly, observing the fitting results of the two spatial econometric models after adding dummy variables. In the spatial lag model, the spatial

autoregression coefficient $P = 0.036$, which did not pass the significance test, which is different from the actual situation. In the spatial error model, the spatial autocorrelation coefficient is $\lambda = 0.94$, which is significant. The significance test with a significance level of 0.01 indicates that there is a significant correlation in the distribution of income between regions. Judging from the size of R^2, the goodness of fit in the error model is $R^2 = 0.997 > R^2 = 0.993$ in the lag model. In addition, the absolute values of Log$L$ value, AIC value, and Schwartz criterion discriminant coefficient are all larger than those of the lag model, which further indicates that the fitting effect of the spatial error model is better.

Based on the conclusions, the spatial error model including dummy variables of geographical factors should be selected in the empirical analysis. From the significance of the regression coefficients of each explanatory variable, in the spatial error model, except for the number of patents granted per 100,000 population, other variables are significant. It shows that each decomposition index of population quality has a significant impact on residents' income. Although the patent granted price is not significant per 100,000 persons representing labor skills, the regression coefficient is 0.028, indicating a positive effect on the income of the population.

From the perspective of regional differences between the east and the west, there is a certain degree of difference in the income of residents in the economic regions of the east, the middle, and the west. The spatial error model obtained is as follows:

$$\begin{aligned}\mathrm{RJ_{GDP}} = {} & 0.25 + 0.21\mathrm{PJ_{YQSM}} - 0.12\mathrm{MWRP_{JJYJF}} + 0.028\mathrm{MSWR_{KZLSQS}} \\ & + 0.063\mathrm{CLHL} - 0.27\mathrm{FNHK_{SZBZ}} + 0.83D_1 - 0.58D_2.\end{aligned} \tag{13}$$

If the spatial errors of the three regions are considered separately, the regression model should be expressed as follows.

East is given as

$$\begin{aligned}\mathrm{RJ_{GDP}} = {} & 0.25 + 0.21\mathrm{PJY_{QSM}} - 0.12\mathrm{MWRP_{JJYF}} + 0.028\mathrm{MSWR_{KZLSQS}} \\ & + 0.063\mathrm{CLHL} - 0.27\mathrm{FNHK_{SZBZ}} + 0.83D_1 - 0.58D_2.\end{aligned} \tag{14}$$

Middle is given as

$$\begin{aligned}\mathrm{RJ_{GDP}} = {} & 0.25 + 0.21\mathrm{PJY_{QSM}} - 0.12\mathrm{MWRP_{JJYF}} + 0.028\mathrm{MSWR_{KZLSQS}} \\ & + 0.063\mathrm{CLHL} - 0.27\mathrm{FNHK_{SZBZ}} - 0.58.\end{aligned} \tag{15}$$

West is given as

$$\begin{aligned}\mathrm{RJ_{GDP}} = {} & 0.25 + 0.21\mathrm{PJY_{QSM}} - 0.12\mathrm{MWRP_{JJYF}} + 0.028\mathrm{MSWR_{KZLSQS}} \\ & + 0.063\mathrm{CLHL} - 0.27\mathrm{FNHK_{SZBZ}}.\end{aligned} \tag{16}$$

From the regression estimation results, it can be seen that the per capita income is higher in the central and western regions in the eastern region, while the western region is higher than the central region under the same conditions.

3. Conclusion

This paper studies and analyzes the effect of population quality on residents' income distribution. This paper first uses the weighted principal component TOPSIS model weighting method to calculate the comprehensive quality index of the population in various regions of China, and then inductively divides the income stages of residents in each region. This paper then uses time-series data to study the relationship between the overall quality of the population and residents' income from the perspective of the whole country and quantitatively studies the effect of population quality on residents' income. It is found that the population quality is high in the east and low in the west, which is consistent with the level of economic development. This paper uses the idea of a comprehensive evaluation to establish the comprehensive index of population quality calculated by the index system and classify it, and divide the 31 provinces into four categories. 3 provinces belong to the high population quality group, 5 high population quality groups, 15 provinces belong to the medium population quality group, and 8 provinces belong to the low population quality group. There are three main reasons for this status quo: First, China is a multi-ethnic country, and the different production and way of life among ethnic groups make it very different in the level of economic development. At the same time, this can also cause differences in population quality. Secondly, the economic development status of different regions is the decisive factor that leads to the difference in the distribution of population quality. Finally, the construction of health care is the most intuitive cause of the development of population quality. Due to the differences in the level of economic development of each region, there are bound to be significant differences in investment in education and health, which leads to differences in the distribution of population quality.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This research study was sponsored by the National Social Science Foundation of China in 2016, "Research on Poverty Alleviation Mechanism of Asset Income in Luoxiaoshan Area from the Perspective of Factor Allocation Efficiency," under 16XJL007. The authors thank the project for supporting this article.

References

[1] D. Hossein, M. Vaziri, and H. Timmermans, "Classification of rural residents on the basis of activity travel characteristics:

case study of villages in Iran," *Transportation Research Record*, vol. 2451, no. 1, pp. 20–27, 2014.

[2] S. Bai, "Labor-income distribution structure of farm households and their evaluation on grain direct subsidy in a wheat production village of shandong province, China," *Studies in Regional Science*, vol. 48, no. 2, pp. 187–202, 2018.

[3] T. V. Sabetova and G. N. Egorova, "Analysis of the factors forming the observed inequality in the income of population in Russia," *Proceedings of the Voronezh State University of Engineering Technologies*, vol. 80, no. 2, pp. 449–457, 2018.

[4] K. B. Cymbranowicz, "Analysis and assessment of the 'Working Poor' Phenomenon among young people in the European union," *Acta Universitatis Lodziensis Folia oeconomica*, vol. 3, no. 348, pp. 91–112, 2020.

[5] J. L. Fiedler and K. Lividini, "Monitoring population diet quality and nutrition status with household consumption and expenditure surveys:suggestions for a Bangladesh baseline," *Food Security*, vol. 9, no. 1, pp. 63–88, 2017.

[6] J. E. Maurer, *University of Miami Business Law Review*, vol. 26, no. 3, p. 4, 2018.

[7] T. Y. Mathä, A. Porpiglia, and M. Ziegelmeyer, "Wealth differences across borders and the effect of real estate price dynamics:evidence from two household surveys," *Journal of Income Distribution*, vol. 26, no. 3, pp. 1–35, 2018.

[8] E. Mcintyre, J. Prior, I. L. Connon, J. Adams, and B. Madden, "Sociodemographic predictors of residents worry about contaminated sites," *Science of the Total Environment*, vol. 643, pp. 1623–1630, 2018.

[9] M. S. González Canché, "The heterogeneous non-resident student body:measuring the effect of out-of-state Students' Home-state wealth on tuition and fee price variations," *Research in Higher Education*, vol. 58, no. 2, pp. 141–183, 2017.

[10] H. J. Murphy, A. C. Karpinski, A. Messer et al., "Resident workshop standardizes patient handoff and improves quality, confidence, and knowledge," *Southern Medical Journal*, vol. 110, no. 9, pp. 571–577, 2017.

[11] C. L. Tsai, "Statistical analysis of the item-count technique using Stata," *STATA Journal*, vol. 19, no. 2, pp. 390–434, 2019.

[12] K. V. Anderson and S. R. Daniewicz, "Statistical analysis of the influence of defects on fatigue life using a gumbel distribution," *International Journal of Fatigue*, vol. 112, pp. 78–83, 2018.

[13] K. Yang, M. Liu, Y. Wang et al., "Carbon dots derived from fungus for sensing hyaluronic acid and hyaluronidase," *Sensors and Actuators B: Chemical*, vol. 251, no. 8, pp. 503–508, 2017.

[14] D. Lam, M. Wei, and D. Wunsch, "Clustering data of mixed categorical and numerical type with unsupervised feature learning," *IEEE Access*, vol. 3, no. 2, pp. 1605–1613, 2015.

[15] M. Al Harrach, V. Carriou, S. Boudaoud, J. Laforet, and F. Marin, "Analysis of the sEMG/force relationship using HD-sEMG technique and data fusion:A simulation study," *Computers in Biology and Medicine*, vol. 83, pp. 34–47, 2017.