The Educational level of the Population, Economic Development, and Urbanization: Based on Experimental and Mathematical Statistics Analysis

Fengyu Lyu[1, a]
[a]18918687821@sina. cn

[1]College of Science, Donghua University, Shanghai, China

Abstract—The educational level of the population has a significant influence on urbanization and economic development. This article uses China's inter-provincial panel data from 2005 to 2019 and analyzes the positive effect of population education on the urbanization rate based on linear regression. Through the intermediary effect model, it is found that economic development is the mechanism of this positive driving effect. This paper also manually collates and calculates the data of annual per capita education years in cities, towns, and townships. Based on the heterogeneity analysis of city, town, and township types, it is found that education level in townships has the strongest driving effect on urbanization compared with cities and towns. At the same time, in terms of different regions, the central region has the greatest promotion effect, followed by the northeast and western regions, and the eastern region has the least promotion effect.

Keywords-urbanization; the educational level of the population; economic development; mediating effect model; heterogeneity analysis.

1 INTRODUCTION

Urbanization is a natural historical process in which non-agricultural industries gather in urban areas and rural populations are concentrated in cities along with the development of industrialization. It is an objective trend in the development of human society and an important symbol of national modernization. The improvement of the educational level of the population will help the economy transform from a country with a large population to a country with strong human resources, and will also affect the progress of urbanization. As a result, whether the enhancement of the educational level of the population is beneficial for urbanization is a crucial question to be discussed and analyzed. At the same time, the research of the relationship among the educational level of the population, the macroeconomic development, and the rate of urbanization is meaningful to conduct since usually, the regions with a relatively larger degree of urbanization have higher economic growth. This dissertation empirically analyzes the interaction mechanism among the educational level of the population, economic development, and the urbanization rate by using the mediation effect model, and it also provides developing countries Chinese cases and experiences for the sake of promoting the development of urbanization.

MSIEID 2022, December 09-11, Chongqing, People's Republic of China

DOI 10.4108/eai.9-12-2022.2327693

2 LITERATURE REVIEW

In recent years, with the progress of the census, various population issues, especially the educational level of the population, have attracted people's attention. Based on the fifth and sixth national censuses, Zhang Shuang took the average years of education per capita and the proportion of the population with various educational levels as the core indicators and analyzed the overall educational level changes of labor resources in a time series from the variations in the quantity and quality of the labor force [1]. Li Yanyan collected educational data of China's population from 2007 to 2017 and used cluster analysis, GIS technology, and concentration index to analyze the discrepancies in educational status. According to his research, in the past decade, the educational level of the northeast and east, and middle regions was generally better than that of the western region. Almost all regions matched their educational standard with regional economic development except for the mismatch between the economic development and the educational level of some western regions such as Xinjiang [2]. Du Rui and Li Xia used multivariate statistics, discrete analysis, regression analysis, and established a scatter plot for the quantitative study [3]. Some scholars have studied the relationship between the economic development of individual regions and the level of education. Li Jie together with his colleagues used SPSS to construct a statistical model to analyze the situation in Suzhou City [4]. Xiong Cuilin and Li Xudong constructed a gray correlation model of the interaction of multiple factors in the economy to analyze the coordination of the educational level of the population and the economic development in Guizhou [5]. Wu Ling and Tang Wen also used the same method to analyze the situation in Liangshan, but the results showed that the coordination between regional education and economic growth is not high [6]. Some scholars found out the reasons behind the issue. Zhong Shuiying and others used the theory of economic growth, demographic age structure, and educational level variables, together with those fitted data from 2000 to 2013 to construct a dynamic panel model. They found that the improvement of educational level offset the dissipation of the demographic dividend, and the economic growth engine varied from depending on the cheap labor force to the enhancement of productivity of all factors of production [7]. Wang Cengceng and others used principal component and cluster analysis, combined with the analysis of the differences in educational levels in various regions in China to reach out a conclusion that economic transformation and upgrading have increased demand for talents, so education is the key to boost economy [8].

Some scholars have revealed the reasons for the education gap between urban and rural areas through research. Huang Weihai and Liu Menglu found that there was no Kuznets inverted U-shaped curve in the education gap between urban and rural populations, because of the improvement of the economic level did not reduce the gap, and the growth rate of the average education level of the population in towns and counties was always lower than that in the cities. The investment in middle and high school education compared to elementary schools had a more tremendous effect on the education gap between urban and rural areas. Besides, population migration has caused the heterogeneity of the urban and rural population structure, which has also led to the expansion of the gap [9]. Some scholars believe that the higher the level of education, the easier it is for people to migrate to cities. From the perspective of social identity, Zeng Peng and Xiang Li constructed a multi-group structural equation model based on survey data to analyze the differences in education levels and influencing factors in the urbanization of the central and western regions. It is concluded that the differences in personal

characteristics have a significant impact on the urbanization willingness of the population transferred from agricultural regions. Furthermore, the research showed that social identity is significantly positively correlated with this willingness, among which the correlation between elementary school and below is the most significant, and the correlation between college and above is the least [10]. Wang Huiying used the 2SLS method to take the university expansion policy as an instrumental variable of educational level. The research showed that the new generation of the floating population with high educational levels has a better economic integration status, but their participation in social organizations, political activities, and interpersonal skills needed to be developed [11].

Based on the above-mentioned existing literature, the possible contributions of this article are mainly four points. Firstly, most papers study the relationship between the educational level of the population and economic development or whether the educational level of the population plays a positive role in promoting urbanization. Few articles have studied the interactive mechanism among the educational level of the population, economic development, and the rate of urbanization. Secondly, by taking into account the urban-rural gap in the educational level of the population, this dissertation classifies the educational level according to city, county, and town. It also considers the different effects of educational level on urbanization in different regions and then develops the heterogeneity analysis of the relationship between urban and rural areas. Thirdly, to demonstrate the relationship between the three in detail, this dissertation uses the latest and the most comprehensive data from the National Bureau of Statistics of 31 provinces, cities, and autonomous regions for 15 years from 2005 to 2019. At the same time, due to the lack of data from the Bureau of Statistics, and to obtain data on the number of years of education in cities, counties, and towns, this dissertation adopts the “China Population and Employment Statistics Yearbook” published in 2020 and downloads 45 tables of urban and rural areas from 2005 to 2019. Fourthly, given that China is a developing country with a dual economic structure of urban and rural areas, the conclusions drawn in this article based on rich data samples and suggestions on regional development are pertinent and can provide relevant cases and experiences for other developing countries.

3 PREPARE YOUR PAPER BEFORE STYLING

3.1 Variables Selection

3.1.1 Explained variable

The explained variable in this article is the urbanization rate. Based on the limitations of existing data, this article uses the ratio of urban population to year-end permanent population (*urb*) to measure.

3.1.2 Explained variable

The core explanatory variable of this dissertation is the educational level of the population, and the average number of years of education per capita (*edu*) will be selected in the model to measure. In addition, this paper also selects the educational level of the rural, town, and urban populations in each province as the explanatory variables. The model will use the average years of education per person in the countryside (*vedu*), the average years of education per person in

towns (*tedu*), and the average years of education per person in the cities (*cedu*) to conduct heterogeneity analysis.

3.1.3 Mediating variable

To reveal the internal mechanism of the population's educational level affecting the urbanization rate, this paper selects economic development as a mediating variable to test the mediating effect and uses per capita gross regional product (*pgdp*) data to measure it.

3.2 Data Sources

The empirical part of this dissertation selects 31 provinces, cities, and autonomous regions in China as the targets of empirical analysis. The analysis collects sample data from the National Bureau of Statistics and the "China Population and Employment Statistics Yearbook" from 2005 to 2019. The missing data use linearity. The missing data is complemented by linear interpolation and cubic spline interpolation. In the empirical stage, to eliminate the influence of heteroscedasticity and volatility, we did logarithm treatment on the per capita GDP data (*lpgdp*).

3.3 Model Setting

The research model in this paper is a one-variable linear regression, and the details are as follows:

$$urb = \alpha_0 + \alpha_1 edu + \varepsilon_1 \tag{1}$$

In the equation, *urb* is the explained variable, which is the rate of urbanization, *edu* is the core explanatory variable, which is per capita years of education, ε_1 is the random error term.

3.4 Descriptive Statistics of Variables

The definitions and descriptions of all variables in this article are demonstrated in Table 1. In addition, this paper conducts the descriptive and statistical analysis of all variables, and the results are shown in Table 2. It can be seen from the statistical characteristics of the variables that China's inter-provincial average urbanization rated from 2005 to 2019 was 0. 532, with a standard deviation of 0. 147. The average regional GDP per capita is RMB 40, 085/person, and the standard deviation is 26, 545, indicating that the development status of various regions in China is quite disparate, and it is necessary to carry out regional heterogeneity analysis.

Table 1 The definition and demonstration of variables

Types of Variables	Names of Variables	Signs	Definitions
Explained variable	Rate of urbanization	*urb*	Urban population / year-end permanent population
Mediating variable	Economic development	*pgdp*	Reginal gross domestic production (GDP) per capita
Explanatory variables	Educational level per capita	*edu*	(Population in elementary schools×6+Population in junior high schools×9+Population in high schools×12+Population in

			college and above×16)/Number of populations aged 6 and over

Table 2 Descriptive statistics

Variable	Unit	Obs	Mean	Std. Dev.	Min	Max
urb	/	465	0. 532	0. 147	0. 207	0. 942
pgdp	RMB/person	465	40085	26545	5218	164563
edu	year	465	8. 726	1. 214	3. 738	12. 68
vedu	year	465	7. 396	0. 884	3. 24	9. 801
tedu	year	465	8. 864	0. 951	4. 211	12. 35
cedu	year	465	10. 42	0. 958	4. 61	13. 16

4 EMPIRICAL ANALYSIS

4.1 Correlation Analysis

Figure 1 depicts the relationship between the number of years of education per capita and the urbanization rate. From the preliminary view of the trend on the graph, there is a good positive correlation between the number of years of education per capita and the urbanization rate, but the relationship between the variables and the mechanism of interaction is needed to be further confirmed by regression analysis in this dissertation.

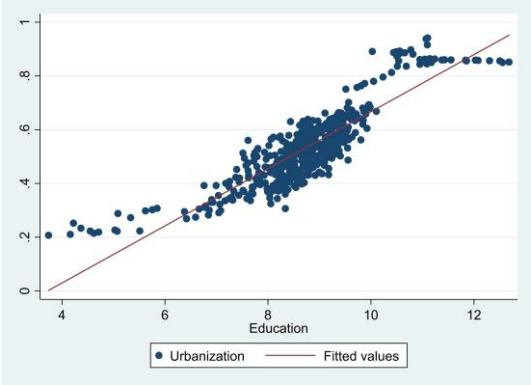

Figure 1. Correlation between urbanization rate and number of the years of education per capita

4.2 Benchmark Regression

Table 3, Model (1), reports the main regression results of the relationship between the number of years of education per capita and the urbanization rate. Specifically, the estimated coefficient of the core explanatory variables of the model is 0. 09603, which is extremely positive at the 1% significance level: R^2=0. 7798, The goodness of fit of the model is prominent. The regression results preliminarily show that the average number of years of education per capita has a positive effect on urbanization in general. That is, as the educational level per capita increases, the urbanization rate will gradually increase. The main reasons for this phenomenon

are as follows: Firstly, the increase in the standard of education in towns and cities has increased people's willingness to migrate from rural areas to towns and cities since they can enjoy better education, thereby gaining more work and promotion opportunities and better living standard. Secondly, the improvement of rural education has increased the quality of people, making them more capable of migrating and living in cities for a long time, which helps them be competent of high-efficiency and fast-paced life of the city.

Table 3 Results of regression analysis

	(1)	**(2)**	**(3)**
	urb	**lpgdp**	**urb**
edu	0. 0960***	0. 8539***	0. 0307***
	-36. 059	-45. 056	-5. 5727
lpgdp			0. 0750***
			-13. 0734
cons	-0. 3063***	2. 9466***	-0. 5166***
	(-11. 7444)	-16. 6286	(-19. 1185)
N	465	465	465
R^2	0. 7798	0. 8674	0. 8961

4.3 Mechanism Test

The above estimation results indicate that the increase in the number of years of education per capita will significantly increase the rate of urbanization. Thus, what kind of mechanism is used to achieve the effect mentioned above? This paper constructs the following mediating effect model to test how the mechanism of economic development interacts with the other two factors. In this empirical analysis, economic development is measured by the logarithmic GDP per capita.

Among them, lpgdp is the mediating variable, which is the GDP per capita after logarithm. According to the test method of the mediating effect, the first step is to regress the equation (1) to test whether the positive pushing effect of *edu* on *urb* exists. If the coefficient α_1 is positive, it indicates that it does have a positive pushing effect. The second step is to estimate equation (2), the relationship between *edu* and the mediating variable. The estimated coefficient of the expected coefficient β_1 is positive. The third step is to estimate equation (3). When the values of α_1 and β_1 are positive, and the values of γ_1 and γ_2 are both positive, if α_1 is greater than γ_1, it indicates the existence of a positive mediation effect.

$$lpgdp = \beta_0 + \beta_1 edu + \varepsilon_2 \tag{2}$$

$$urb = \gamma_0 + \gamma_1 edu + \gamma_2 lpgdp + \varepsilon_3 \tag{3}$$

In this dissertation, the mediating effect model is regressed. The coefficient β_1 of the core explanatory variable, the average number of years of education per capita in the model (2), Table 3, is significantly positive. It can be verified that the number of years of education per capita which can promote the rise of urbanization is also able to improve economic development. It is mainly because of the improvement of educational level, which can improve the quality of the labor force. Furthermore, it increases the efficiency of the labor force and the total output, the GDP per capita of a region, and the development of the regional economy.

The regression results of model (3) in Table 3 show that the number of years of education per capita, the γ_1 coefficient is 0. 03069, and the GDP per capita, the coefficient γ_2 is 0. 07505. The degree of the fitting is R^2=0. 8961, which is significant. α_1, γ_1, γ_2 are all greater than 0 and γ_1 is less than α_1, which indicates that the GDP per capita and the number of average years of education both play a positive role in promoting the urbanization rate. After using the regional GDP per capita as an explanatory variable, the direct and positive promotion effect of the number of average years of education per capita on the urbanization rate is reduced (γ_1 is less than α_1), indicating the existence of a mediating effect. That is, the positive promotion effect of the number of average years of education per capita on regional GDP per capita is reflected in the urbanization rate through the latter. The reason why the increase in the educational level per capita can promote urbanization has been mentioned above. Besides, the reason why economic development can promote urbanization is that the secondary and tertiary industries account for a large proportion of the economy. The improvement of economic development depends on the secondary and tertiary industries to great extent. Moreover, cities are dominated by secondary and tertiary industries. Therefore, improving the economy largely depends on the improvement of the urban economy. Cities and towns need more labor to meet their development so the increased employment opportunities will promote urbanization of the population. At the same time, the development of the urban economy will attract more people to yearn for a higher living standard in cities and towns. On the other hand, the development of the rural economy will increase the capital of rural people per capita and provide economic conditions for moving to cities and towns.

Finally, according to $\gamma_0 = -0.5166 < 0$ and $\alpha_0 = -0.3063 < 0$, indicating that if the effects of the GDP per capita and the number of years of education per capita are excluded, reverse urbanization may occur. Without educational conditions and economic foundation, cities and towns will lose their attractiveness. This is because that access to such resources is the main source of opportunity and ideal life, which is the purpose and main motivation for people to migrate from suburbs to cities and towns. In addition, the difference between rural and urban areas is not tremendous. It further shows the importance of China's economic development and people's educational level to urbanization.

4.4 Mechanism Test

4.4.1 Heterogeneity analysis of cities, counties, and towns

Table 4, Model (1), Model (2), and Model (3) respectively report the results of regression analysis of the impact of the number of average years of education per capita on the urbanization rate in suburbs, towns, and cities. The estimated coefficients of explanatory variables are 0. 1142, 0. 0469, and 0. 0554 respectively, which are all significantly positive at the significance level of 1%. The results of the regression analysis show that, for China's urbanization rate, the increase in the number of years of education per capita in urban and rural areas is conducive to the improvement of the urbanization rate. However, the improvement of the years of education per capita in rural areas has a stronger promotion effect on the urbanization rate. Thus, the enhancement of educational level per capita in rural areas is more helpful for urbanization. This is mainly because that after those who live in rural areas for years get higher education, they are more competent to the high-quality life in cities and towns.

Therefore, they are capable of meeting higher requirements for sophisticated labor skills. Comparatively, the effect of the improvement of urban educational level on urbanization is not obvious, though the improvement of urban education can increase the gap between urban and rural education. which can attract more people to migrate city to have a better living standard. Instead, the people in the town, have already lived in towns and almost have no intention to leave the town. It is because once they leave the town, the quality of life will be greatly decreased. Therefore, they will not choose to migrate, which has no great impact on the urbanization rate.

Table 4 Results of regression analysis

	(1)	**(2)**	**(3)**
	urb	**urb**	**urb**
vedu	0. 1142***		
	-25. 6606		
tedu		0. 0469***	
		-11. 238	
cedu			0. 0554***
			-15. 9072
cons	-0. 3130***	0. 1158***	-0. 0458
	(-8. 3001)	-2. 7368	(-1. 1137)
N	465	465	465
R^2	0. 5933	0. 3447	0. 4219

4.4.2 Regional heterogeneity analysis

According to the division method of The National Bureau of Statistics in 2011, all provinces in China are divided into four regions: eastern, central, western, and northeast, which can help to reveal the regional characteristics of the impact of years of education per capita on urbanization rate. In Table 5, models (1), (2), (3), and (4) respectively report the results of the regression analysis of the impact of years of education per capita on urbanization rate in eastern, central, western, and northeastern regions. According to the analysis, the estimated coefficients of explanatory variables are 0. 0714, 0. 1266, 0. 0971, and 0. 1014 respectively. The results of the regression analysis are significantly positive at the significance level of 1% indicating that, for China's urbanization rate, the number of years of education per capita has the most significant effect on the urbanization rate in central China, followed by northeast China. The urbanization rate in eastern and western China is less affected by the number of years of education per capita.

The central region used to be dominated by agriculture with many rural areas and little urban development. Based on the development of rural education which has had a strong impact on urbanization, combining with the resources of the central parts of a good education and overall planning including the “Rise of Central China” and “Belt and Road Initiative” strategy, the rise of this region in educational attainment is the strongest driver of urbanization. In 2003, the CPC Central Committee and The State Council issued Several Opinions on the Implementation of the Revitalization Strategy of The Northeast and Other Old Industrial Bases, which implemented the revitalization strategy of northeast China. Later on, it came to an important period for the exploration of new-type urbanization in Northeast China. At this stage, the improvement of the educational level of the population in northeast China plays a relatively crucial role in promoting urbanization. In western China, the terrain and geological conditions

are complex, and the infrastructure construction is relatively backward. As a result, the indigenous population is small. Although there are rural areas, it is difficult to develop into cities. Therefore, although the educational level of rural areas has a certain impact on urbanization, population migration and urban construction are still a big problem. For eastern regions, whose economic development has been in the final stage of industrialization, are the most developed place in China mainly occupied by cities and towns. As the development of urban education had a weak impact on urbanization in the past, the educational level of the population has the least promoting effect on urbanization in the eastern and western regions.

Table 5 Results of regression analysis

	(1)	**(2)**	**(3)**	**(4)**
	urb	**urb**	**urb**	**urb**
edu	0. 0714***	0. 1266***	0. 0971***	0. 1014***
	-13. 8288	-20. 7421	-27. 8714	-12. 3524
cons	-0. 0248	-0. 6254***	-0. 3344***	-0. 3376***
	(-0. 4681)	(-11. 3159)	(-10. 5974)	(-4. 3340)
N	150	90	180	45
R^2	0. 8424	0. 84	0. 8229	0. 782

5 CONCLUSIONS AND POLICY IMPLICATIONS

According to the above empirical results, this dissertation holds that the educational level of the population plays an important role in promoting urbanization and economic development. The development of urbanization depends on the promoting effect of the educational level of the population and the economic development caused by the increase in educational level. Compared with the educational level in cities and towns, improving the educational level in rural areas has the most obvious promotion effect on urbanization. Besides, compared with the eastern and western regions, improving the educational level of the central and northeastern regions will promote urbanization more strongly. The eastern regions have the least positive effect on urbanization while the central regions have the most significant effect.

The implications of the above conclusions lie in the following four aspects. Firstly, improve the educational level of the population, it is suggested that the government can increase the educational expenditure, hire high-quality teachers with high academic qualifications, conduct disciplinary training for existing teachers, carry out international academic exchanges, and give classified guidance to students. This method can ensure that every student can become a useful talent as far as possible. In addition, enterprises can train employees to improve the quality of the labor force. Secondly, promoting the interaction between urbanization and economic development by investing and distributing advanced equipment can improve labor efficiency and increase regional output and income. Thirdly, for the sake of ensuring high-quality education and institutions in cities, teachers can be encouraged to conduct overseas academic exchanges to improve the academic level of institutions. Moreover, it is suggested that increasing the construction of educational resources such as senior middle schools in towns can attract more rural students to study in cities. The government can increase and expand rural compulsory education resources, strengthen the construction and improvement of school

atmosphere and environment to attract prominent teachers with higher salaries and a favorable academic atmosphere. Fourthly, with the help of good educational resources, the central region should strengthen the connection and agglomeration between large cities and other small and medium-sized cities and small towns, which can give full play to the radiating and leading role of the central cities in the urban agglomerations in the central region. Therefore, it can further strengthen the population-gathering function of small cities and towns absorbing rural surplus labor. Based on a strong industrial foundation, the northeast regions need to form a new type of urbanization advantage from the technological reform of new industries to promote urban migration. Instead, for the western regions with the full implementation of the western development strategy, the investment must be strengthened and the construction of infrastructures such as roads, medical care, culture and sports, and public services must be strengthened as well. For the developed urban areas in eastern regions, it is suggested to strengthen the development of tertiary industry, develop foreign trade and commodity exchanges, increase imports and exports, and further develop the urban economy to attract more population.

REFERENCES

[1] Zhang, S. (2014) An empirical study on the educational level of the Chinese labor force based on the time series analysis of the census data platform, Northwest population, 35(03): 73-78.

[2] Li, Y. Y. (2019) Analysis of the spatial-temporal development of educational status in China under the new situation. Shanghai Education and Research, 10: 5-10.

[3] Du, R. , Li, X. (2016) The relationship between population education level and GDP. Global Market Information Herald, 04: 84.

[4] Li, J. , Fei, S. L. , Su, T. (2016) Statistical analysis of population education level and economic development. Journal of Jilin Institute of Chemical Technology, 33(07): 95-98.

[5] Xiong, C. L. , Li, S. Y. (2012) Grey correlation analysis of education level and economic development in Guizhou. Journal of Guizhou normal university (Natural Science Edition), 30(01): 119-122.

[6] Wu, L. , Tang, W. (2016) Grey correlation analysis of education level and economic growth in Liangshan Prefecture. Urban Geography, (18): 226.

[7] Zhong, S. Y. , Zhao, Y. , Ren, J. R. (2016) Research on the substitution effect of “Education Dividend” on “Demographic Dividend”. Chinese Population Science, 02: 26-34+126.

[8] Wang, C. C. , Guo, Y. , Yin, Z. M. (2019) A study on the differences in the educational level of the population in different regions of my country under the background of economic transition and upgrading. Journal of Liaoning University of Technology (Social Science Edition), 21(01): 30-33.

[9] Huang, W. H. , Liu, M. L. (2016) The evolution and influence mechanism of the population education gap between towns and towns. Educational Economics Review, 1(03): 56-76.

[10] Zeng, P. , Xiang, L. (2017) Mathematical statistics and management. Mathematical statistics and management, 36(02): 273-286.

[11] Wang, H. Y. (2020) The impact of education on the social integration of the new generation of floating population. Market Weekly, 33(09): 124-125+174.