Effects of Urbanization on the Resources Allocation of Basic Education: A Regression Modeling Analysis Based on Panel Data

Xiaoqing Xu[1,2,a*], Hong Gao[3,b]

[1]*School of Humanities and Law, Northeastern University, Shenyang, China*
[2] *School of Management, Shenyang JianZhu Unversity, Shenyang, China*
[3]*China CITIC Bank Corporation Limited, Anyang, China*
[*a]*xxqwzx@126.com*, [b]*707542388@qq.com*

ABSTRACT

With the development of urbanization, the allocation level of basic education resources is constantly improving. This paper studies the impact of urbanization on the allocation of basic education resources based on the panel data of 14 prefecture-level cities in Liaoning Province from 2009 to 2016. Co-integration test shows that there is a long-term stable equilibrium relationship between the allocation level of basic education resources and urbanization. Using generalized least squares method (FGLS) and panel modified standard deviation (PCSE) to adjust covariance, the paper takes the proportion of urban population as an explanatory variable and the income gap between urban and rural areas as a control variable, the regression analysis shows that urbanization has a positive impact on education expenditure, while the control variable, the income gap between urban and rural areas, has a negative impact on education expenditure. The results are of great significance to realize the rational allocation of basic education resources and the balanced development of education.

Keywords: *Urbanization; Basic Education; Regression Modeling Analysis; FGLS; PCSE*

1. INTRODUCTION

With the development of urbanization, the allocation level of basic education resources is constantly improving. Urbanization needs to provide indispensable support for the development of basic education, and it also has a great impact on the development of basic education. Existing studies have generally discussed the influencing factors of educational development level or unbalanced educational development. In a review of empirical studies on the influencing factors of education inequality, it is found that more empirical studies use education Gini coefficient to measure relative education inequality, and increasing education expenditure can reduce education gap [9].

On the relationship between urbanization process and education development, some researchers have also conducted a series of studies. A study shows that basic education resources can significantly promote urbanization, However, there are regional differences in the relationship between them [10]. Another study shows that there is a long-term equilibrium relationship among financial expenditure on education, economic growth and urbanization. At the same time, there is a two-way causal relationship between education investment and new urbanization, and education investment has a positive impact on new urbanization [8].

Most of the existing achievements focus on the relationship between education and economic growth and the imbalance of education development, but lack of urbanization as the macro-transformation background affecting the development of basic education, and lack of panel data to analyze the impact of urbanization on the allocation of basic education resources from the perspective of time and space.

Based on this, the paper focuses on the following issues: Does the urbanization process have an impact on the allocation of basic education resources? If there is an impact, how big is the impact? Is this influence different from city to city?

Using co-integration test and regression analysis, this paper studies the influence degree of urbanization on the allocation of basic education resources, which has

A. El-Hashash et al. (Eds.): IEIT 2022, ACSR 100, pp. 675–680, 2023.
https://doi.org/10.2991/978-94-6463-058-9_107

practical significance for realizing the rational allocation of basic education resources and the balanced development of education under the background of urbanization.

2. METHOD

2.1. Theoretical Framework

The main phenomenon of China's urbanization process at the current stage is the continuous concentration of rural population and industries in cities and towns, and the continuous diffusion of urban life concepts and ways to rural areas, which brings about the growth of residents' income and changes in consumption behavior. In addition, it emphasizes quality rather than speed, and the salient features are the focus on basic public service and the quality of human life, which inevitably require the allocation of basic education resources in the direction of more equalization.

The core of new urbanization is "urbanization of people". According to Schultz's "Human Capital Theory" [5], human capital is the sum of the expenditure on education and vocational training of workers and the opportunity cost of their education. The core of human capital is to improve the quality of population, therefore, investment in education is the main part of human capital investment.

The logical relationship is shown in Figure 1.

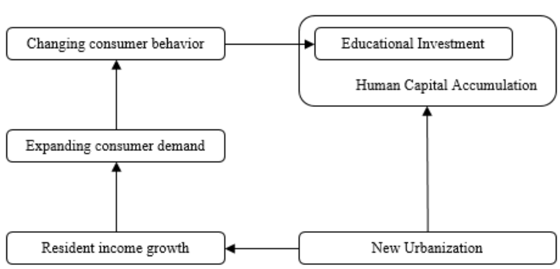

Figure 1: The logical relationship between education development and new urbanization

In summary, urbanization can have an impact on the allocation of educational resources through the growth of residents' income, expansion of consumption demand, and changes in consumption behavior. Investment in education is a major part of human investment in new urbanization. To study the impact of urbanization on educational investment, the following function is established:

$$\mathrm{H} = f(x) = f(\text{ urbanization, income }) \quad (1)$$

2.2. Variables

Drawing on previous research results and Grange's idea of causality, the paper explores whether there is a long-term equilibrium relationship between urbanization and basic education resource allocation level at the level of Liaoning province, and constructs a regression model about them.

Considering the availability and intuitiveness, the explanatory variable in this paper is chosen as the proportion of education expenditure to GDP (EPR), and the larger the value, the higher the degree of investment in the allocation.

The main explanatory variable is the urbanization rate (UR). For the measurement of urbanization rate in this paper, the population share indicator method (urban population/total population), which is based on the resident population, is used to measure the urbanization rate of each prefecture-level city in Liaoning Province by the level of population urbanization, and this data is more easily available.

In addition, the choice of control variables needs to be considered in the empirical analysis. To reflect the influence, the control variable in the paper is chosen as the urban-rural income gap, expressed as the ratio of per capita disposable income of urban permanent residents to per capita disposable income of rural permanent residents (INCOME).

2.3. Sources of Data

The paper takes 14 prefecture-level cities in Liaoning Province (Shenyang, Dalian, Anshan, Fushun, Benxi, Dandong, Jinzhou, Yingkou, Fuxin, Liaoyang, Panjin, Tieling, Chaoyang and Huludao) as research objects, and selects short panel data from 2009 to 2016 for analysis. The data come from Liaoning Statistical Yearbook and Liaoning Demographic Annual Report.

2.4. Descriptive Statistical Analysis

The description of the main variables and descriptive statistical properties of the regression model are shown in Table 1.

Table 1: Panel data describing statistical properties

	EPR(%)	UR(%)	INCOME
Average	13.123	62.067	2.125
Mean	13.1252	63.425	2.082
Max	17.296	80.550	2.736
Min	7.330	32.200	1.696
SD	2.217	12.874	0.261

As shown in Table 1, the level of urbanization varies widely among different regions and periods. According to the standard deviation coefficient, the differences in EPR and INCOME between different regions and periods decrease in order relative to UR.

Figure 2 shows trend of education expenditure in GDP are not exactly the same for each prefecture; the overall trends for Shenyang, Fuxin, and Chaoyang are all rising and then falling and then rising again, except for the other prefectures, which all experience a downward trend before bottoming out.

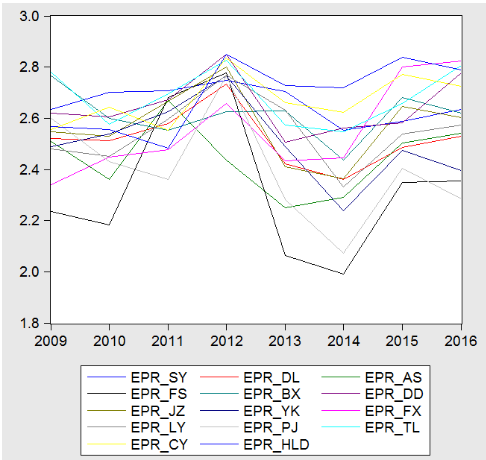

Figure 2: Time trend of 14 prefecture-level cities

3. RESULTS

3.1. Stationary Test and Co-integration Test

This paper first tests the stability of each variable, and then carries out the co-integration -test after the precondition of the integration test. Co-integration test is a method to examine the long-term equilibrium relationship between variables. If a group of non-stationary variables has a stationary linear combination, that is, this linear combination counteracts the random trend of this group of variables, then this group of variables is co-integration. In order to investigate whether there is a long-term relationship between the allocation level of basic education resources and urbanization level, it is necessary to carry out co-integration test on two variables: the proportion of education expenditure to GDP (EPR) and urbanization rate (UR).

3.1.1. Stationary Test

Firstly, the unit root test is carried out to investigate the stability of the panel data. According to the properties of panel data, the unit root test can be divided into two types: one is that all sections have the same root, and the unit root test methods include LLC test [6], Breitung test [1] and Hadri test [4]; The other is that each section has different roots. The unit root test methods include IPS test (Im, Pesaran and Shin, 2003), Fisher-ADF test [7] and Fisher-PP test [2]. In order to increase the robustness of the test results, this paper adopts two methods to carry out unit root test, namely, the same root unit root test LLC test and the different root unit root test Fisher-ADF test. The original assumptions of the two methods are that there is a unit root. By observing the trend of the original data sequence diagram, the intercept term should be considered in determining the unit root test mode.

The LLC test was performed on the original panel data for the variables EPR and UR, and the test results are presented in Table 2: the P-value corresponding to the test statistics for both was 0.0000 at the 5% significance level, which is less than 0.05, means that the original hypothesis is rejected and there is no unit root. The ADF test results are shown in Table 3: the P-values are less than 0.05 at the 5% significance level. thus, it can be seen that the original panel data both pass the smoothness test, indicating that the variables EPR and UR are both zero-order single integer panel data, which satisfy the prerequisites, and the co-integration can be conducted. The test can be conducted for EPR and UR.

Table 2: LLC unit root test

Variables	Statistic	Prob.**	Results
EPR	-7.79414	0.0000	Stable
UR	-26.9642	0.0000	Stable

Table 3: ADF unit root test

Variables	ADF	Prob.**	Results
EPR	-3.10086	0.0010	Stable
UR	-9.07040	0.0000	Stable

Note: "**" indicates rejection of the original hypothesis and acceptance of the alternative at the 5% significance.

3.1.2. Co-integration Test

On the basis of the previous panel data passing unit root tests, Co-integration tests are conducted on the variables EPR and UR to test whether there is a long-run equilibrium relationship between the data. The methods are divided into two categories: one is the panel Co-integration test based on the unit root test of residual data, which is based on the Engle-Granger two-step test, and the specific methods are mainly the Kao test and the Pedroni test; the other is based on Johansen Co-integration test. In this paper, the Kao test and Pedroni test are used, which have the advantage of allowing for max individual differences and also different short-run dynamics, and two original hypothesis is that there is no Co-integration between them. As is shown in Table 4 and Table 5.

Table 4: Kao test

	t-Statistic	Prob
ADF	-1.893000	0..292

Table 5: Pedroni test

	Statistic	Prob.	
Panel v-Statistic	-1.822942	0.9658	Accept
Panel rho-Statistic	0.675420	0.7503	Accept

Panel PP-Statistic	-5.590636	0.0000	Reject
Panel ADF-Statistic	-7.756801	0.0000	Reject
Group rho-Statistic	2.305713	0.9894	Accept
Group PP-Statistic	-7.079695	0.0000	Reject
Group ADF-Statistic	-7.596387	0.0000	Reject

Note: Prob. indicates the p-value of the statistic

The data in Table 4 shows that the statistic P value of Kao test is 0.0292, and if it passes the test of 10% significance level, it rejects the original hypothesis, indicating that there is a co-integration relationship between variables.

It can be seen from Table 5 that Panel v-Statistic, Panel rho-Statistic and Group rho-Statistic did not pass the test of 10% significance level, and other statistics all passed the test of 10% significance level.

In Pedroni test, Panel ADF-Statistic and Group ADF-Statistic have better small sample properties than other statistics, so this model mainly refers to Panel ADF-Statistical and Group ADF-Statistical to judge whether there is co-integration relationship. The data in Table 5 shows that Panel ADF-Statistic and Group ADF-Statistic are both 0.0000, which have passed the test of 10% significance level. Therefore, it can be judged that there is a long-term equilibrium relationship between the allocation level of basic education resources and urbanization in 14 prefecture-level cities in Liaoning Province.

3.2. Multiple Regression Analysis

3.2.1. Regression Modeling

The general form of panel data models based on panel data and used to analyze the relationship between the effects of variables:

$$y_{it} = \alpha_{it} + \beta_{it}x_{it} + \mu_{it} \quad (2)$$
$$i = 1,2,3 \ldots N; \quad t=1,2,3\ldots,T$$

Note:

y_{it}- The explained variable;

x_{it}- The explanatory variable;

α_{it} β_{it}- The parameters;

i - Different individuals;

t - Different times;

μ_{it}- The random disturbance term;

The coefficients in the model change with time and individual, so they can reflect the influence of time factors and individual differences neglected in the model. In practical application, it is divided into variable coefficient model, variable intercept model and mixed estimation model.

As the number of cross-sections is larger than periods, and the units of the cross-section are all units of the total, it is more reasonable to use the fixed-effects model. On this basis, F-test is used to determine the choice of the variable coefficient model, variable intercept model, and mixed estimation model. In the case that the parameters do not vary over time, the two hypotheses are tested as follows:

$$H_1: \beta_1 = \beta_2 = \Lambda = \beta_N \quad (3)$$

$$H_2: \alpha_1 = \alpha_2 = \Lambda = \alpha_N \quad \beta_1 = \beta_2 = \Lambda = \beta_N \quad (4)$$

If hypothesis$H_2$is accepted, it is a mixed estimation model and no further tests are performed. If H_2 is rejected, H_1 is tested. if $H_1$is accepted, it is a variable intercept model; if hypothesis H_1 is rejected, it is a variable coefficient model.

The sums of squares of residuals estimated with the three models were calculated as follows, in order:

S_1=1.591179 S_2=1.995689 S_3=3.119565

The F is constructed and modeled as follows:

$$F_2 = \frac{S_3 - S_1/[(N-1)(k+1)]}{S_1/\big(NT - N(k+1)\big)}$$
$$\sim \mathrm{F}[(N-1)(k+1), N(T-k-1)] \quad (5)$$

$$F_1 = \frac{S_2 - S_1/[(N-1)k]}{S_1/\big(NT - N(k+1)\big)}$$
$$\sim \mathrm{F}[(N-1)k, N(T-k-1)] \quad (6)$$

Due to N=14, T=8, K=3, therefore F_2 =1.72404, F_1=0.68445, When the significance level α=0.05,

$F_{0.05}$ (39,70)=1.570558, $F_{0.05}$ (26,70)=1.65434

Thus, F_2>$F_{0.05}$ (39,70),

Reject H_2 , because $F_1 < F_{0.05}$ (26,70),Then H_1 is accepted and the fixed effects variable intercept regression model should be chosen.

Referring to the production function, the explanatory variable of the regression model is EPR, the explanatory variable is UR, and the control variable is INCOME.

Finally, the level of basic education resource allocation is necessarily influenced by other factors besides those mentioned above, but since it is impossible to list them all, the omitted variables are grouped into the residual term.

In order to eliminate effects of heteroscedasticity and autocorrelation in the panel data, the variables were taken as natural logarithms, respectively, the improved regression model was as follows:

$$LnEPR_{it} = C_i + \beta_{it}LnUR_{it} + \gamma_{it}LnINCOME_{it} + \varepsilon_{it} \quad (7)$$

Note:

EPR_{it}-The share of education expenditure in GDP in period t of city i,;

UR_{it}-The urbanization rate of population in period t of city i;

$INCOME_{it}$-The urban-rural income gap in period t of city i;

C_i- The intercept term;

β_{it} ,γ_{it}- The parameters to be estimated;

ε_{it}- The error term.

3.2.2. Regression Result

In this paper, we first estimate the model using OLS, which considers individual fixed effects and time fixed effects, but the model does not take into account between-group heteroscedasticity and contemporaneous correlation, so the results are not fully significant and need to be further tested. The estimation results are shown in column (1) of Table 6.

In order to effectively deal with the complex panel error structure, this paper also adopts FGLS method for estimation, which adds a weight to the explanatory variables so that the variances are the same. It can not only eliminates heteroskedasticity, but also performs t-tests and F-tests under OLS on the estimators. The specific results are shown in column (2) of Table 6.

In the case that the number of time T is smaller than N, the standard deviation of FGLS does not fully reflect their variances, and in order to remedy the shortcomings, this paper further adopts the panel corrected standard error (PCSE) to deal with the complex panel error structure. The PCSE method corrects the standard deviation of the residual terms by substituting them into the diagonal matrix on the basis of retaining the OLS estimated parameters. The specific results are shown in column (3) of Table 6.

From column (1) of Table 6, it can be seen that the coefficient estimates of UR are not significant at the 5% level, and others are below the critical value and significant. From columns (2) and (3) of Table 6, it can be seen that results of both FGLS and PCSE are consistent: the Durbin-Watson stat obtained using FGLS and PCSE are 2.037514 and 2.154529, indicating that the residuals of the equation are not auto correlated; the R-squared is close to 0.5, indicating the equation has a good fit; and the Prob is

less than 0.05, which indicates that it has good significance. In addition, the parameter estimates are also significant, and the coefficient estimates of the variables UR and INCOME are significant at the 5%. Analyzing the signs of the regression coefficients and their magnitudes, the UR is +0.330994, indicating that it has a positive effect, and the slope term indicates that every 1% increase drives the increase as a percentage of GDP by 0.33%; while the control variable urban-rural income gap has a negative effect on the expenditure. The slope term indicates that each 1% increase in urban-rural income gap will drive education investment to decrease 0.42% of GDP.

Table 6: Results of panel data regression analysis for 14 prefecture-level cities in Liaoning Province

	(1)		(2)		(3)	
Model form	OLS		FGLS		PCSE	
Coefficients and P	Coefficient	Prob.	Coefficient	Prob.	Coefficient	Prob.
C	1.720614	0.0406	1.511974	0.0204	1.511974	0.0217
UR	0.300663	0.1365	0.330994	0.0371	0.330994	0.0394
INCOME	-0.539339	0.0045	-0.424636	0.0000	-0.424636	0.0000
Durbin-Watson stat	2.053091		2.037514		2.154529	
R-squared	0.431667		0.495811		0.495811	
Prob(F-statistic)	0.000001		0.000000		0.000000	

This paper analyzes the different situations of 14 cities in Liaoning. As is shown in Table 7.In the fixed-effect variable intercept model, the mean is 1.511974, and the estimation result of each city C_i, which is the deviation of the "i" city intercept value from the average intercept value. According to Table 7, it can be seen that there are some differences in the proportion of education investment in GDP among the 14 prefecture-level cities, among which Shenyang, Dandong, Jinzhou, Fuxin, Liaoyang, Tieling, Chaoyang, and Huludao deviate positively from the average respectively, indicating that these regions have a greater influence on the support, among which Huludao has the greatest support. On the contrary, the level of others is negatively deviated, which indicates that these regions have less influence.

Table 7: Estimation results of C_i for the difference in the level of allocation of basic education resources by region

Region	C_i	Region	C_i
Shenyang	0.047388	Yingkou	-0.067040
Dalian	-0.143203	Fuxin	0.005988
Anshan	-0.168166	Liaoyang	0.002681
Fushun	-0.281791	Panjin	-0.177927
Benxi	-0.011793	Tieling	0.138002
Dandong	0.027389	Chaoyang	0.225317
Jinzhou	0.082613	Huludao	0.320542

3.2.3. Robustness Analysis

Since FGLS and PCSE use different methods to deal, the paper uses both methods for estimation, and the results are the same except for the slightly different Prob values, which indicates that the regression model is robust and the parameter is accurate.

4. CONCLUSION

By the Co-integration test, the paper found that there is a long-term stable equilibrium relationship between the urbanization rate and the allocation level of basic education resources in Liaoning Province.

Through regression analysis, it can be seen that urbanization has promoted the level of basic resource allocation across regions in Liaoning Province while the urban-rural income gap has hindered local government investment in education.

Based on the above conclusions, under the process of urbanization, the government should be implementing some specific policies to increase investment in education in terms of human, material and financial resources, reduce the income gap between urban and rural areas and regional differences, so that more basic education resources can be effectively allocated and utilized.

ACKNOWLEDGEMENTS

This research was supported by Science Research Funding Project of the Education Department of Liaoning Province(lnqn201920).

REFERENCES

[1] Breitung, J., R Brüggemann. (2000). Uncovered interest parity: what can we learn from panel data?. Sfb Discussion Papers, 323-337.

[2] Choi, I. (2001). Unit root tests for panel data. Journal of International Money and Finance, 20(2), 249-272.

[3] Doan T.(2013). IPSHIN: RATS procedure to implement Im, Pesaran and Shin panel unit root test. Boston College Department of Economics.

[4] Hadri, K., Guermat, C., Whittaker, J.(2003). Estimation of technical inefficiency effects using panel data and doubly heteroscedastic stochastic production frontiers. Empirical Economics, 28(1), 203-222.

[5] Klein, P. G., Chair, R. D. P. . (2006). T.w. schultz and the human-capital approach to entrepreneurship. Review of Agricultural Economics, 28(3), 344-350.

[6] Levin A , Lin C . Unit Root Test in Panel Data: new results. 1993.

[7] Maddala, G. S., Wu, S. (1999). A comparative study of unit root tests with panel data and a new simple test. Oxford Bulletin of Economics and Statistics, 61(S1), 631-652.

[8] Sun Z., Chen R.(2014). Research on education investment, economic growth and new urbanization construction[J]. Reform and Strategy, (11):112-115.

[9] Xiong Y., Liu Z., Zhou C.(2014). Initial endowment, resource allocation, educational expansion and educational equity[J]. Tsinghua University Education Research, (6):96-103.

[10] Zhou H.(2016). An empirical study on the impact of regional basic education resource allocation on new urbanization in China[J]. Northwest Normal University Journal (Social Science Edition), (2):93-98.