

An Empirical Study on the Impact of Educational Gap on Income Gap

Yu Wang

College of Science, North China University of Technology
Beijing City, China
wangyu_1013@126.com

Hongmei Li

College of Science, North China University of Technology
Beijing City, China
lhmei6280@163.com

Abstract—Both in theory and in experience, education is considered as an effective means of promoting economic growth, increasing household income and reducing income inequality. It is not clear whether the reform of education system in recent years has redefined the relationship between education gap and income gap. Based on the data of six provinces and cities in different parts of China in 1996-2015, we use panel data regression to analyze and prove that the extent to which the income gap is affected by education has risen and education inequality has greatly affected income inequality. Besides, although the narrowing of the education gap has not significantly improved income inequality, there is no denying that education expansion is conducive to control the income distribution gap.

Keywords—education gap; income gap; Gini coefficient; years of education

I. INTRODUCTION

Both the popularizing of elementary education and the expanding of high education are important goals in many countries and regions. Since the reform and opening-up, the Chinese government has adopted a series of policies and measures in the fields of education, which has greatly expanded the range of public education. Although education inequality still exists, the education gap is reduced. Meanwhile, the income gap in China is also changing dramatically. During the past 20 years, the income gap has an increasing trend. In this paper, we discuss the education factors which have influence on income gap, in order to provide feasible suggestions to narrow the income gap in China.

II. DATA SOURCES AND MEASUREMENT METHODS

In this paper the data are from Statistical Yearbook of China, China Population Statistics Yearbook, and the local statistical yearbook of relevant provinces and cities. Based on the data of people over 15 years old, we divide the people into 5 groups, which are the illiteracy or literacy less, primary school, junior high school, high school, college and above. According to China's geographical division, we collect data in six representative regions, which are Beijing, Liaoning, Jiangsu, Hubei, Sichuan and Xinjiang.

In this paper, we choose three variables, which are the average years of education, the Gini coefficient on education, and Gini coefficient on income. The measurement methods are as follows:

A. Measurement of Average Years of Education

Let μ be the average years of education, we get

$$\mu = \sum_{i=1}^n p_i y_i, \quad i=1, 2, \dots, 5$$

Where p_i is the population proportion in different group, y_i is the corresponding years of education. In this paper, $y_1 = 0, y_2 = 6, y_3 = 9, y_4 = 12, y_5 = 16$.

B. Measurement of the Gini Coefficient on Education

There are direct and indirect methods to measure the Gini coefficient on education. We use the latter. In this paper, we employ the Thomasetal's (2001) calculation formula of the improved Gini coefficient based on the education distribution feature which is as follows [4]:

$$G^E = \frac{1}{\mu} \sum_{i=2}^n \sum_{j=i-1}^1 p_i |y_i - y_j| p_j \quad (1)$$

Where G^E is the Gini coefficient based on the education distribution feature. y_i and y_j are different education years in different groups, p_i and p_j are the corresponding proportions, μ is the average years of education, n is the number of groups [1].

Based on the data, the Lorentz Curve of Education is shown in Figure 1. The vertical axis represents the cumulative percentage of education years and the horizontal axis represents the cumulative percentage of population. The dotted line of 45 degree represents the education equal line, while the curve represents the Lorentz Curve of Education. Therefore [2][6]:

$$G^E = \frac{\text{The area of A}}{\text{The area of A} + \text{The area of B}} \quad (2)$$

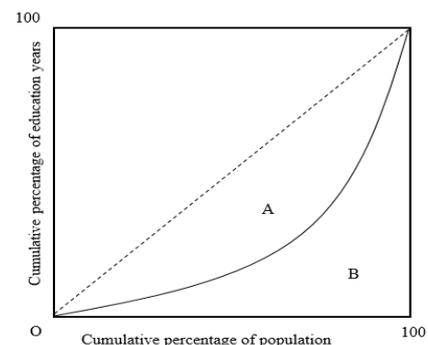


Fig. 1 Lorentz Curve of Education

C. Measurement of the Gini Coefficient on Income

Similarly, we can calculate the Gini coefficient on income. The corresponding Lorenz Curve of Income is shown in Figure 2.

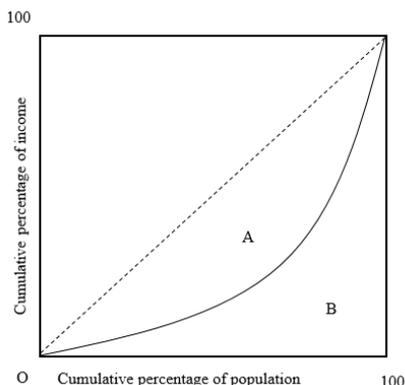


Fig. 2 Lorenz Curve of Income

The formula of the Gini coefficient on income is:

$$G = \frac{\text{The area of A}}{\text{The area of A + The area of B}} \tag{3}$$

III. THE EMPIRICAL ANALYSIS

A. Panel Data Regression

In this paper, we develop the panel data regression model for depth analysis. The formula is:

$$y_{it} = \alpha + X_{it}\beta + u_{it} \quad i = 1, \dots, N; t = 1, \dots, T \tag{4}$$

where, i is the individual, t is the time, N of the cross section is the number of individuals, and T is the dimension of the time series, α is a scalar, X_{it} is the observation of individual i during t period, Y_{it} is the explained variable, μ_{it} is the single factor error [3].

The formula of μ_{it} is:

$$u_{it} = \mu_i + v_{it} \tag{5}$$

Where μ_i is special effects of invisible individuals, v_{it} is residual random disturbance.

B. Data Empirical Analysis

In the model based on the data in this paper, period $T=20$, number of individuals $N=120$. It has not only the cross sectional data but also the time series data, which meets the basic requirements of the micro panel. We use panel econometric analysis method, through the likelihood ratio test to choose the individual fixed effects model. In the model, the income gap G is explained variable, and the education gap G^E and the average years of education E are explanatory variables.

1) Data Analysis

a) Variable calculation

Based on the collected educational data, calculate the percentage of accumulated years of education and the percentage of the population. And then draw the education Lorenz curve, according to the equation (2) to calculate the education Gini coefficient. The following six provinces and cities in 2015 data, for example, the provinces and cities in 2015 to calculate the Gini coefficient of education (Table 1).

TABLE I THE PERCENTAGE OF THE ACCUMULATED POPULATION AND YEARS OF EDUCATION IN 2015

| | the percentage of the accumulated population | the percentage of accumulated years of education |
|-----------------|--|--|
| Beijing | 0.00% | 0.00% |
| | 7.90% | 2.99% |
| | 34.50% | 22.29% |
| | 55.56% | 42.67% |
| | 100.00% | 100.00% |
| Liaoning | 0.00% | 0.00% |
| | 16.89% | 8.85% |
| | 65.05% | 51.54% |
| | 81.84% | 71.38% |
| | 100.00% | 100.00% |
| Jiangsu | 0.00% | 0.00% |
| | 22.71% | 10.56% |
| | 62.48% | 46.96% |
| | 82.16% | 70.97% |
| | 100.00% | 100.00% |
| Hubei | 0.00% | 0.00% |
| | 23.24% | 10.69% |
| | 63.47% | 48.02% |
| | 83.54% | 72.85% |
| | 100.00% | 100.00% |
| Sichuan | 0.00% | 0.00% |
| | 34.59% | 18.05% |
| | 72.47% | 56.83% |
| | 87.72% | 77.65% |
| | 100.00% | 100.00% |
| Xinjiang | 0.00% | 0.00% |
| | 25.11% | 12.87% |
| | 67.42% | 52.43% |
| | 83.30% | 72.24% |
| | 100.00% | 100.00% |

According to the table 1, we get the education of the six provinces and cities Lorenz curve. The figure 3 shows the Lorenz curves of the six provinces and cities.

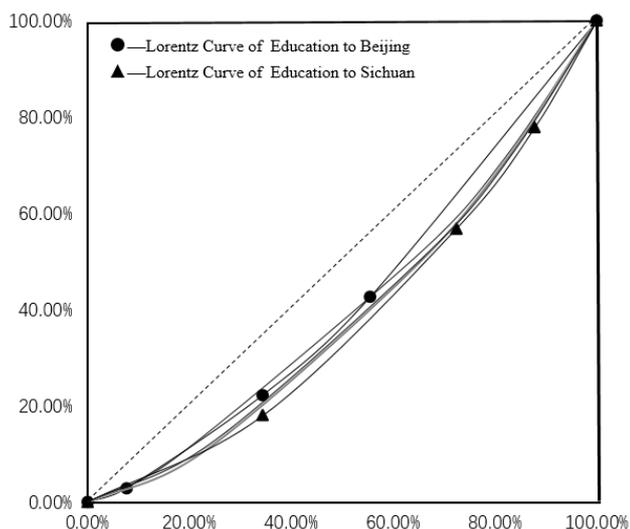


Fig. 3 Lorenz Curve of Education to six provinces and cities in 2015

The six curves have many overlapping parts. The curve of the dot mark is the Beijing Lorenz curve; the triangular mark curve is the educational Lorenz curve of Sichuan Province. We can see that the level of education in Beijing is relatively average, different from Sichuan Province.

Based on the equation (2) we can calculate the Gini coefficient of education of six provinces and cities in 2015, and for the convenience of the study, the income Gini coefficient is used by the statistical data of the local governments. The Gini coefficient of education and income is as follows:

TABLE II THE GINI COEFFICIENT OF EDUCATION AND INCOME

| provinces and cities | the Gini coefficient of education | the Gini coefficient of income |
|----------------------|-----------------------------------|--------------------------------|
| Beijing | 0.165051885 | 0.2936 |
| Liaoning | 0.183726137 | 0.367 |
| Jiangsu | 0.214184053 | 0.3775 |
| Hubei | 0.214789538 | 0.3904 |
| Sichuan | 0.237532625 | 0.3852 |
| Xinjiang | 0.205223501 | 0.4348 |

Similarly, the Gini coefficient of Chinese education and the Gini coefficient of income are calculated and plotted as shown in figure 4.

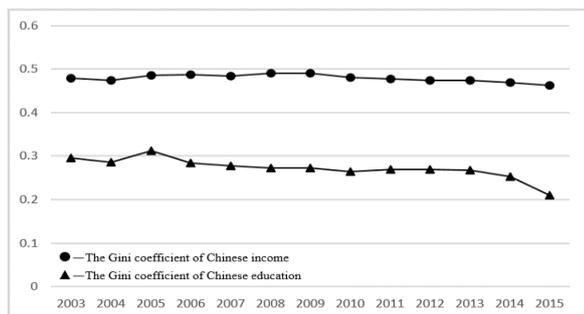


Fig. 4 The Gini coefficient of Chinese education and income

In Figure 4, the dotted mark represents the Gini coefficient of income, and the triangular markings represent the education Gini coefficient. It can be seen that, in China, the income Gini coefficient and education Gini coefficient there is a dependency.

Next, we use the panel analysis method to analyze the above data of six provinces and cities.

b) Likelihood Ratio Test

TABLE III DATA OF RESULTS OF LIKELIHOOD RATIO TEST

| Effects Test | Statistic | d.f. | Prob. |
|---------------------------------|------------|---------|--------|
| Cross-section F | 102.087772 | (5,93) | 0.0000 |
| Cross-section Chi-square | 224.405428 | 5 | 0.0000 |
| Period F | 3.937635 | (19,93) | 0.0000 |
| Period Chi-square | 70.831573 | 19 | 0.0000 |
| Cross-Section/Period F | 51.846100 | (24,93) | 0.0000 |
| Cross-Section/Period Chi-square | 319.897587 | 24 | 0.0000 |

In the above table, the value of Cross-section F is very large, and the p value is less than 0.05, which indicating that the model conforms to the fixed effect model.

c) The Results of Measurement

TABLE IV DATA OF RESULTS OF MEASUREMENT REGRESSION IN 1996-2015

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|----------------|-------------|--------------------|-------------|--------|
| C | 0.253549 | 0.044323 | 5.720469 | 0.0000 |
| E | 0.015361 | 0.003604 | 4.262050 | 0.0000 |
| G ^E | -0.146549 | 0.081980 | 1.787622 | 0.0265 |
| R-squared | 0.913020 | Adjusted R-squared | | 0.9075 |
| F-statistic | 167.9497 | Prob(F-statistic) | | 0.0000 |

In table 4, we can get the equation:

$$G = 0.254 + 0.0154E - 0.147G^E \tag{6}$$

In the equation, it can be seen that the average length of education and income gap are changing in the same direction, while the education gap and the income gap are in the opposite.

As can be seen in table 4, firstly, the coefficient of E and G^E are significant when the significance level is 0.05. The adjusted R² > 0.90, which proves that the education gap and the average years of education are important factors affecting income gap and the two factors can account for more than 90% of the income gap.

Secondly, in table 4, we can see that the coefficient of G^E is -0.147, which means that in the case of other factors unchanged, the Gini coefficient on income will change 0.147 unit in the opposite direction, when the Gini coefficient change 1 unit. In the same way, the coefficient of E is 0.0154, which means that in the case of other factors unchanged, the Gini coefficient on income will change 0.0154 unit in the same direction, when the average years of education change 1 unit.

Thirdly, from the tend of the change, the average years of education is yearly increasing, and the education inequality is decreasing. However, the income gap is widening year by year. The supply and demand relationship in the labor market is also the factor that affects the relationship between education and income gap. When education years increase, education degree and educational level are improved. People's knowledge literacy and ability with education have the advantage in the labor market, thus the income gap is widened [5].

d) Unit Root Test

TABLE V DATA OF RESULTS OF UNIT ROOT TEST

| Method | Statistic | Prob.** | Cross-sections | Obs |
|--|-----------|---------|----------------|-----|
| Null: Unit root (assumes common unit root process) | | | | |
| Levin, Lin & Chu t* | -4.87147 | 0.0000 | 6 | 101 |
| Breitung t-stat | -2.30127 | 0.0107 | 6 | 95 |
| Null: Unit root (assumes individual unit root process) | | | | |
| Im, Pesaran and Shin W-stat | -6.71009 | 0.0000 | 6 | 101 |
| ADF - Fisher Chi-square | 58.0835 | 0.0000 | 6 | 101 |
| PP - Fisher Chi-square | 87.3809 | 0.0000 | 6 | 108 |

In table 5, it is the result of unit root test of the first order difference of sequence. The LLC test, LM test, ADF test, and PP test are significant at the significance level of 0.05. It indicates that the data is fluctuating in time series, but the first-order difference time series is stable. Thus the data can be used for analysis, and the above panel analysis is meaningful.

IV. CONCLUSION

From the analysis mentioned above, we can draw the following conclusions:

- (1) Education is an important variable in the factors that affect the income gap, which has strong explanatory power.
- (2) The narrowing of the education gap has not significantly reduced the income gap. The mechanism of action is: the higher education degree a person has, the greater personal ability he has, the greater contribution to society he makes and the more income the society assigns to him.
- (3) The average length of education will affect the reverse change of education gap. That is, the increase of the average years of education has narrowed the gap of education, but the impact on the income gap is not obvious.

ACKNOWLEDGMENT

This research was financially supported by NCUT research project of education, teaching reform and curriculum construction (No. XN009-068), and high level talents of Beijing university cross training "real training program" project.

REFERENCES

- [1] Li Shou, Lai De Sheng, etc. China's income distribution research report [M]. Social Science Literature Publishing House, 2013 (1): 181-189.
- [2] Hu Jing Jing. Research on the Income Gap of Chinese Urban and Rural Residents since the Reform and Opening - up [M]. People's Publishing House, 2013 (1): 162-165.
- [3] Bai Zhong Lin, Zhang Xiao Tong. Economic Analysis of Panel Data Metrology [M]. Machinery Industry Press, 2010 (1): 10-14.
- [4] Du Peng. Empirical Analysis of China's Educational Development on Income Gap [J]. Nan Economic Research, 2005 (4): 47-52.
- [5] Bai Xue Mei. Education and Income Inequality: A Study of Chinese Experiences [J]. Managing the World, 2004 (6): 53-58.
- [6] Chen Zong Sheng. A proposal for Estimating Total Gini Coefficient [J]. Economic Research, 2002(5):84-87.