Wage Income and Returns to Education in Rural China: Based on Quantile Regression Method*

Ying WANG[1], Zhiwang LV[1], Zhi QU[2]

[1] China Agricultural University, Beijing

Email:wang-shengbo@163.com, lvzhiw@cau.edu.cn

[2] Chinese Academy of Social Sciences, Beijing

Email: quzhi4125@163.com

Abstract:

Employing quantile regression method, this paper estimated the Mincer equation of the wage income of Chinese rural residents. The results showed that, as income quantiles increased, returns to rural education declined slowly at first and then made a small rise. Through comparing wage income returns to education between male and female, this paper found that the returns to education of female were significantly higher than those of male at both 10% and 25% income quantiles. In addition, by dividing rural non-farm workers into five groups based on the economic nature of their units, the essay found that the diversity of units' economic nature and the differences in rates of education returns among people were the causes of income inequality among rural non- agricultural workers.

Key Words: Wage income, Returns to education, Non-agricultural employment, Quantile regression

*Ying WANG, Master degree candidate, College of Economics and Management, China Agricultural University, Beijing. Zhiwang LV, associate professor, College of Economics and Management, China Agricultural University. Zhi QU, PhD candidate, Graduate School of Chinese Academy of Social Sciences, Beijing. The project is supported by the National Social Science Fund of China (Grant No.13CJY019)

1. INTRODUCTION

In the process of economic development and economic restructuring, labor transfer among different industries is inevitable. Since Chinese reform and opening-up in 1978, the proportion of agriculture in economic structure has declined constantly, while the proportion of non-agricultural has increased, which produces a large number of rural surplus labor transferring from agriculture to non-agricultural. On one hand, township enterprises provide non-agricultural employment opportunities for rural surplus labor. On the other hand, as the demand for labor in urban areas increases, more and more farmers migrate to city and work in off-farm industries. Large numbers of non-agricultural employment opportunities make rural surplus labor easier to improve their income, however, substantial previous studies show that the probability of non-agricultural employment and the off-farm income level among different migrant works vary greatly because of different quality of human capital, especially the difference of education level. Using CHIP data, Knight and Song (2001) found that the probability of finding non-agricultural work for rural residents with high school education is 10% and 20% higher than that of illiteracies in 1988 and 1995 respectively. Moreover, after rural surplus labor obtain non-agricultural employment opportunities, education can still play a significant positive role on non-agricultural income. de Brauw (2002) confirm that education has good influence on both migrant workers and local non-agricultural employment, and the function of education increases continuously as the time passes. Deng Quheng (2009) pointed out that education can improve non-agricultural employment opportunities for rural residents, while under the premise of having access to off-farm employment, education can also improve non-farm income levels.

Since the U.S. economist Mincer proposed Mincer equation in 1958 which combines personal income with education level and work experience, the equation has become the most common method for national scholars to study wage income and returns to education. Using the logarithm of annual income as the dependent variable, Li Shi and Li Wenbin (1994) indicated that the rate of returns to education of rural non-farm workers is 0.02 based on the CHIP data of 1988. In Deng Quheng’s (2007) paper, research data also comes from CHIP， but the dependent variable is the logarithm of hourly wage and the rate of returns to education of migrant workers in 2002 is 0.0585. Although the estimated results of Mincer equation vary because of different estimation methods, and various dependent variable selection and research data selection, what most scholars believe is the fact that the rate of returns to education is relatively low at the beginning of the reform and opening up, and with reform's deepening and market improvement, human capital plays a more important role in income distribution, and thus the rate of returns to education increases.

In previous studies, most scholars use ordinary least squares (OLS) method to estimate Mincer equation, thus the rate of returns to education reflects, with other conditions unchanged, how average income changes with different education levels. However, due to the prevalence of income inequality, income distribution presents a high degree of skewness. Consequently, using average income to measure the changes is not appropriate and the estimate results are often biased. In order to avoid these problems, this paper uses quantile regression method to estimate the rate of returns to education. Contrasted with OLS method, quantile regression estimates the equation under arbitrary income quantiles comprehensively and particularly, and thus the results can reflect how income is impacted by the level of education in a certain income quantile rather than how average income is influenced by different education levels (Xing Chunbing, 2008). Furthermore, if income distribution reveals a sharp kurtosis, a thick tail or significant heteroscedasticity, the OLS estimate result would no longer be the best linear unbiased estimator (BLUE) and the robustness of estimator parameters would be poor, while using quantile regression method will be a good way to solve these defects. What is more, the heterogeneity of each sample makes it difficult to explain capacity variance when analyzing the returns to education (Li Xuesong and James Heckman, 2004), while quantile regression method assumes that the ability of people in higher income quantile is stronger than that of people in lower income quantile. Hence, using this method can not only evade the difficulty of finding suitable instrumental variable for "ability" variable but also avoid the estimation error brought by the heterogeneity of sample.

2. RESEARCH DATA AND METHODOLOGY

2.1 Research Data

The data in this paper comes from China General Social Survey (CGSS) in 2010 which is conducted by National Survey Research Center (NSRC) at Renmin University of China. This survey is the first national, comprehensive, and continuous large-scale social survey project in China covering residents' income, medical and health care, education, unemployment insurance and some other aspects. The respondents include 11785 urban and rural residents coming from 31 provinces, autonomous regions and municipalities. The samples in this paper are consisted of 1223 rural residents who have non-agricultural work experience more than 1 year and are currently engaged in non-agricultural work. Besides, the non-agricultural employment income (i.e. the wage income) of the samples is greater than zero. Descriptive statistical analyses of research data are as follows.

The survey (CGSS) does not ask respondents the years of education, so the education years in this paper are reckoned through education background: for respondents who have no education, their education years are 0; for respondents who go to primary school, their education years are 6; if respondents go to junior high school, then the education years are 9; if respondents go to senior high school, then the education years are 12; junior college is equivalent to 15 years of education; undergraduate college corresponds to 16 years of education; the years of education for masters and PhDs are 19.

Table 1 Annual Income Distribution Statistics of Non-agricultural Workers in 2009

	Less than 10 thousand yuan	10 thousand to 20 thousand yuan	20 thousand to 30 thousand yuan	30 thousand to 50 thousand yuan	50 thousand to 100 thousand yuan	More than 100 thousand yuan
Proportion	29.19%	34.83%	17.83%	11.12%	5.15%	1.88%

Data sources: analyzing according to 2010 CGSS data

Table 2 Statistics of the Education Levels of Non-agricultural Workers

	Illiteracy	Primary School	Junior High School	Senior High School	Junior College	Undergraduate College	Graduate College
Proportion	4.17%	23.79%	46.85%	19.87%	4.01%	1.23%	0.08%

Data sources: analyzing according to 2010 CGSS data

Table 3 Unit's Economic Nature Distribution Statistics of Non-agricultural Workers

	State-owned or State-controlled	Collectively-owned or Collectively-controlled	Privately-owned or Privately-controlled	Foreign Capital Investment	Others
Proportion	8.34%	5.56%	67.05%	1.47%	17.58%

Data sources: analyzing according to 2010 CGSS data

Table 4 Descriptive Statistical Analysis of Sample

	Total（100%）		**Male（62.55%）**		**Female（37.45%）**	
Variable	Mean	Std. error	Mean	Std. error	Mean	Std. error
Annual Wage Income	21036.31	45035.61	24718.97	55379.71	14885.14	15347.35
Education Years	8.84	3.07	9.00	2.84	8.57	3.40
Non-agricultural Employment Years	11.81	9.07	13.30	9.50	9.32	7.68

Data sources: analyzing according to 2010 CGSS data

The statistical data in table 1 and table 2 indicate that annual non-agricultural income of migrant workers in 2009 present a highly right-skewed distribution, while the years of education show an approximate normal distribution. In all samples, the average annual non-agricultural income of migrant workers is 21036.31 yuan, the average years of education is 8.84 years, and the average years of non-agricultural employment is 11.81 years. Table 4 compares female non-agricultural workers with male non-agricultural workers. The average years of education for male and female workers are close, but average income of male worker is significantly higher than that of female worker, which indicates gender differences play an important role in income determination. Thus, reducing gender discrimination will make positive effect on abating income inequality.

In addition, Table 5 displays the relationship between wage income and education levels. As income levels increase, the average years of education and average wage income increase. There exists a positive correlation between education and wage income. However, in terms of the gap between different income levels, average years of education changes little between adjacent income levels while average wage income changes a lot. Table 5 shows that the higher the income level is, the bigger the gap of average wage income between adjacent income levels exists.

Table 5 Average Education Years and Wage Income of Different Income Levels

	Less than 10 thousand yuan	**10 thousand to 20 thousand yuan**	**20 thousand to 30 thousand yuan**	**30 thousand to 50 thousand yuan**	**50 thousand to 100 thousand yuan**	**More than 100 thousand yuan**
Average Education Years	7.90	8.85	9.36	9.71	10.02	10.17
Average Wage Income	5665.94	12788.67	21615.78	33560.66	58709.52	229630.4

Data sources: analyzing according to 2010 CGSS data

2.2 Research Methodology

This paper uses extended Mincer equation to estimate the rate of returns to education:

$$LnY = \beta_0 + \beta_1 Edu + \beta_2 Exp + \beta_3 Exp^2 + \sum_i \lambda_i x_i + \varepsilon$$

Thereinto, LnY is the logarithm of migrant workers' annual wage income, Edu and Exp represent the knowledge labors gained from education and the experience labors gained from the work respectively. In order to quantify knowledge and experience, years of education and years of non-agricultural employment are chosen as proxy variables in this paper. The coefficients β_1 and β_2 denote the growth rate of personal wage income when Edu or Exp adds one unit, and$\beta_1$represents the rate of returns to education. What is more, in a period of time when labor start to work, working experience accumulates as wage income increases. There exists a positive correlation between working experience and wage income during this period. When wage income of migrant workers reaches a certain level, working experience and wage income may present a negative correlation because of technological progress and the decline of working effort and enthusiasm. Taking this nonlinear relationship into account, squared term of working years should be introduced into the equation, and its coefficient$\beta_3$is often negative. In addition, for the sake of analyzing the impact of other factors on wage income, some other control variables such as gender and units' economic nature can be introduced into the equation.

As mentioned above, the annual wage income of migrant workers shows a highly right-skewed distribution in 2009, thus if the Mincer equation was estimated by OLS method, the results would lack robustness and credibility. Therefore, this paper uses quantile regression method to estimate the Mincer equation. Compared with OLS method which can only describe the impact of independent variable on partial changes of dependent variable, quantile regression is more accurate in describing the influence of independent variable on distribution shape changes of dependent variable. Especially when the dependent variable appears skewed distribution, quantile regression is able to capture characteristics of the tails of distribution and to analyze the impact of independent variable on various distribution of dependent variable in different income quantiles comprehensively. Based on the distribution of wage income in table 1, this paper use quantile regression method to estimate the Mincer equation at 10%(the lowest quantile), 25%(lower quantile), 50%(median quantile), 75%(higher quantile), and 90%(the highest quantile) income quantiles respectively with the help of Stata12.0.

3. QUANTILE REGRESSION ESTIMATION OF RETURNS TO EDUCATION

Numerous previous studies show that there is a positive correlation between education and wage income (Liu Longsheng, 2008), and the statistic result in table 5 confirms this regulation as well. In order to study the difference of returns to education under various income levels and the trend of returns to education in income distribution, this paper uses quantile regression method to estimate Mincer equation. The results are presented in table 6.

The estimated results in Table 6 show that the rates of returns to education and working experience vary in different income levels. The rate of returns to working experience increases as income level rises while the rate of returns to education declines slowly at first and then makes a small rise as income quantiles increases. The rate of returns to education of migrant workers at the lowest quantile (10%) is higher than that of migrant workers at the highest quantile (90%), and the rate of returns to education of migrant workers at 75% income quantile is the lowest. Specifically, wage income of migrant workers at 10% and 25% income quantiles will increase by 6.90% and 6.52% respectively with each additional year of education, while wage income of migrant workers at 90% and 75% income quantiles will increase by 6.31% and 5.93% respectively with each additional year of education. Therefore, if all migrant workers get an additional year of education, the growth rate of migrant workers at 10% income quantile would be 0.59% and 0.97% higher than that of migrant workers at 90% and 75% income quantiles respectively, which means that improving the education level of migrant workers at

10% and 25% income quantiles will be conducive to reduce income inequality among migrant workers.

Table 6 Quantile Regression Result of Mincer Equation

		Q=10%	Q=25%	Q=50%	Q=75%	Q=90%
Education		0.0690***	0.0652***	0.0652***	0.0593***	0.0631***
		(0.0156)	(0.0121)	(0.0093)	(0.0103)	(0.0120)
Working Experience		0.0296**	0.0311**	0.0405***	0.0467**	0.0645***
		(0.0146)	(0.0119)	(0.0088)	(0.0086)	(0.0134)
Squared terms of Working Experience		-0.0007	-0.0008**	-0.0009***	-0.0011	-0.0015**
		(0.0005)	(0.0003)	(0.0003)	(0.0002)	(0.0004)
Gender		0.2548***	0.3154***	0.2749***	0.2516***	0.2265**
		(0.0860)	(0.0676)	(0.0556)	(0.0623)	(0.0938)
Units; economic nature	State-owned/controlled	0.4948***	0.1868	0.0429	0.0160	-0.2564**
		(0.1582)	(0.1217)	(0.1301)	(0.1304)	(0.1502)
	Collectively-owned/controlled	-0.1284	-0.0342	-0.0267	-0.0769	-0.2807
		(0.2268)	(0.1954)	(0.1512)	(0.1648)	(0.1859)
	Privately-owned/controlled	0.2548*	0.2703***	0.1052	0.1419	-0.0605
		(0.1336)	(0.1049)	(0.0864)	(0.0997)	(0.1217)
	Foreign Capital Investment	0.7778***	0.6481***	0.4322***	0.2589	0.3964
		(0.2044)	(0.1956)	(0.1645)	(0.3072)	(0.5216)
Constant Term		7.3525***	7.8402***	8.4132***	8.8946***	9.3949***
		(0.2209)	(0.1608)	(0.1344)	(0.1374)	(0.1657)
Pseudo R^2		0.0646	0.0680	0.0773	0.0666	0.0770

Annotation: ***, **, and * indicate that estimated parameters are significant at the significance level of 1%, 5%, and 10% respectively, the number in ( ) is the standard error of estimated parameter resulted from bootstrap method after repeated sampling for 500 times.

Coefficients of dummy variable of gender in Mincer equation are positive and the coefficients decrease gradually as income level rises, which indicates that there is a gender discrimination among migrant workers and that the discrimination against female is more apparent in low-income level. Specifically, in the case of the same education level and working experience, the annual wage income of male migrant workers at 25% income quantile is 31.54% higher than that of female migrant workers and the proportion is 22.65% at 90% income quantile. The result indicates that the income inequality between male and female migrant workers is bigger in low-income level than that in high-income level. Thus, reducing gender discrimination against women and making more female migrant workers gone into high-income level will make contributions to reducing income inequality.

In table 7, the further analysis of returns to education focusing on male and female migrant workers reveals that the rate of returns to education of male migrant workers is different from that of female migrant workers and that the variation trends of returns to education are not the same in income distributions of two gender groups. In male migrant workers, the rate of returns to education presents a W-shaped fluctuation as income level increases, while the rate of returns to education of female migrant workers increases at first and then decreases. At both 10% and 25% income quantiles, the rates of returns to education of female migrant workers are significantly higher than that of male migrant workers. Therefore, improving education background of female migrant workers, especially the background of female migrant workers in low-income level, can effectively reduce income inequality between male and female migrant workers.

Table 7 Quantile Regression Result of Gender Differences of Returns to Education

	Q=10%	Q=25%	Q=50%	Q=75%	Q=90%
Returns to Education of Male	0.0735*** (0.0236)	0..0639*** (0.0189)	0.0746*** (0.0124)	0.0620*** (0.0140)	0.0635*** (0.0168)
Returns to Education of Female	0.0787*** (0.0306)	0.0923*** (0.0153)	0.0586*** (0.0128)	0.0527*** (0.0161)	0.0481*** (0.0173)

Annotation: ***, **, and * indicate that estimated parameters are significant at the significance level of 1%, 5%, and 10% respectively, the number in ( ) is the standard error of estimated parameter resulted from bootstrap method after repeated sampling for 500 times.

Quantile regression result of Mincer equation based on units' economic nature indicates that the difference of economic nature of non-agriculture units is one of the reasons that create income inequality among different migrant workers. At every income quantile, migrant workers who work in foreign invested units can have evidently higher wage income than those who work in other kinds of units, and the difference is more apparent in low-income level. Besides, the wage income for migrant workers in privately-owned or privately-controlled units is also relatively high. Results in table 6 evince that migrant workers have disadvantages in wage income if he works in state-owned/controlled or collectively-owned/controlled units and the disadvantages are more serious in high-income level.

In Table 8, most quantile regression results of returns to education based on units' economic nature are significant at 1%, 5% or 10% level. In these estimated parameters which are significant, the rate of returns to education of migrant workers who work in foreign invested units is high, which can explain why migrant workers in foreign invested units have higher wage income from the perspective of education. The rate of returns to education of migrant workers in collectively-owned or collectively-controlled units is also relatively high, and as income quantiles increases the rate increases slightly after a sharp drop. Meanwhile, the rate is higher than that of migrant workers in privately-owned/controlled, state-owned/controlled or others units in each income level. Accordingly, improving education background of migrant workers in collectively-owned/controlled units will not only be helpful to reduce income inequality in this kind of unit but also be conducive to reduce income inequality between collectively-owned/controlled units and other kinds of units.

In addition, the economic nature of "others" contains neighborhood committees, village committees, commonweal organizations, etc. In table 6, most coefficients of dummy variable of economic nature types are positive except collectively-owned/controlled units, which indicate that the wage income of migrant workers in “others” units is higher than that of migrant workers in collectively-owned/controlled units and generally lower than that of migrant workers in the other three types of units.

Table 8 Quantile Regression of Returns to Education Based on Units’ Economic Nature

	Q=10%	Q=25%	Q=50%	Q=75%	Q=90%
State-owned/controlled	0.0617*	0.0502**	0.0615***	0.0753***	0.0957***
	（0.0352）	（0.0254）	（0.0221）	（0.0246）	（0.0222）
Collectively-owned/controlled	0.1165***	0.1104***	0.0805**	0.0939***	0.1106*
	（0.0445）	（0.0431）	（0.0366）	（0.0335）	（0.0617）
Privately-owned/controlled	0.0610***	0.0690***	0.0729***	0.0770***	0.0797***
	（0.0197）	（0.0173）	（0.0104）	（0.0169）	（0.0131）
Foreign Capital Investment	0.1793***	0.1351*	0.1026	-0.0629	0.0522
	（0.0676）	（0.0813）	（0.0911）	（0.1059）	（0.0997）
Others	0.0810***	0.0882**	0.0402	0.0078	0.0173
	（0.0292）	（0.0395）	（0.0266）	（0.0240）	（0.0462）

Annotation: ***, **, and * indicate that estimated parameters are significant at the significance level of 1%, 5%, and 10% respectively, the number in ( ) is the standard error of estimated parameter resulted from bootstrap method after repeated sampling for 500 times.

4. CONCLUSION

Based on the above statement, this paper obtains the following conclusions:

(1) There is a positive correlation between education and wage income, and the rate of returns to rural education declines slowly at first and then makes a small riseas income quantiles increases. Thus improving the education background of migrant workers at 10% and 25% income quantiles will be conducive to reduce income inequality among migrant workers.

(2) There exists gender discrimination against female migrant workers, and this discrimination is more serious in low-income level. Besides, the rate of returns to education of female migrant workers is positive at each income quantile, and the rates of returns to education of female migrant workers are significantly higher than that of male migrant workers at both 10% and 25% income quantiles. Therefore, improving education background of female migrant workers, especially the background of female migrant workers in low-income level, can effectively reduce income inequality between male and female migrant workers.

(3) The difference of economic nature of non-agriculture units is one of the reasons that create income inequality among different migrant workers. Non-agriculture employees in different units have various rates of returns to education. Accordingly, increasing education background

of migrant workers, improving economic market structure, and narrowing policy differences on units of disparate economic nature will be propitious to reduce income inequality.

REFERENCES

[1] de Brauw,et al., 2002,The evolution of China's rural labor markets during the reforms, *Journal of Comparative Economics*, 30 (2),329-353.

[2] John B. Knight, and Lina Song,2001, Economics growth, economic reform, and rising inequality in China,In *China's Retreat fromEquality*, eds., Carl Riskin,et al., 84-122, New York: M.E. Sharp.

[3] Deng Quheng, 2007, The influence of reform on taxes and fees for education expenses of rural residents: a research based on PSM and DID method, working paper.(*in Chinese*)

[4] Deng Quheng, ed.,2009, *Education, Income Growth and Income Inequality: Empirical Analysis on Rural China*, Shanghai: Truth and Wisdom Press. (*in Chinese*)

[5] Liu Longsheng, 2008, Influences of education and experience on Chinese residents, *Journal of Quantitative & Technical Economics*, 25 (4): 75-84.(*in Chinese*)

[6] Li Shi, and Li Wenbin, 1994, Estimation of returns to personal education investment in China, In *Research on Chinese Residents' Income Distribution*, eds., Zhao Renwei, and Keith Griffin,289-336, Beijing: China Social Sciences Press.(*in Chinese*)

[7] Li Xuesong and James J. Heckman, 2004, Heterogeneity, selection bias and the return to education: anempirical analysis based on Chinese micro-data, *Economic Research Journal*, 50(4), 91-99.(*in Chinese*)

[8] Xing Chunbing, 2008, Quantile regression, return to education and income distribution, *Statistical Research*,25(5), 43-49.(*in Chinese*)