

The Rate of Returns to Schooling: A Case Study of Urban China

Xu Zhang*

Abstract

This study seeks to study the rate of return to schooling in urban China by addressing two questions: First, have the returns to schooling in China been underestimated? Second, should Mincer earning function be directly applied? The study suggests a modified Mincer earning equation with quadratic term of years of schooling to reflect the differences in average rate of returns to additional year of schooling among various levels of education. The OLS estimates of returns to education are lower than those from the IV approach. The overall returns to education from OLS estimation based on the standard Mincer equation are in the range of 7%-8%. The IV approach reports a much higher rate of returns to education, 16% for college graduate, based upon the modified Mincer earning equation, while the figure from the OLS estimation is 10.24%.

Introduction

China has been experiencing profound economic reforms and a sustained average annual growth rate of over 6% for a few decades. High-speed economic growth result in more human capital investment in China and the rate of returns to schooling has been widely studied.

Researchers find the studies are interesting and important due to several reasons. First, returns to schooling indicate the productivity of education and show the incentive for private investment in one's human capital, which contributes to economic growth especially for developing countries. Second, different levels of returns to schooling among various groups of people (e.g. men and women, different ethnic groups, city and rural areas, etc) may shed some light on the assessment of efficiency of resource allocation pertaining to education. Third, estimation on economic returns to schooling has important policy implications. Psacharopoulos and Patrinos (2002) suggest government should design policies and incentives to promote private investment in human capital and ensure investment for low-income families based upon the evidence from research. Finally, examining returns to schooling in China can help to assess the structure of the Chinese labor market and the extent of economic transformation. Historically, Chinese culture highly values education. The study of this issue can signal private demand for education. If the returns to schooling are too low, it may hurt economic growth by not accumulating enough human capital.

* Farmingdale State College, State University of New York - History, Economics and Politics Department
2350 Broadhollow Road, Farmingdale, NY 11735 - xu.zhang@farmingdale.edu

Psacharopoulos (1994) surveys empirical studies of the returns to education across 98 countries and finds a world average of 10.1% , 9.6% in Asia and a range between 11.2% to 11.7% in low and middle income countries (those with a per capita income of less than \$2449). However, existing literature shows a much lower economic returns to education in China. Johnson and Chow (1997), using the data from 1988 Chinese Household Income Project (CHIP-88), estimates a return of 3.29%. Liu (1998) suggests the returns to education are in the range of 2.8% to 3.6% by using the same data set—CHIP 88. The CHIP 1995 (CHIP-95) data have an advantage over 1988 data as the former contain information on working hours which can be used to calculate hourly wage rate, while the latter collect data on annual earnings. Using CHIP-95, Li (2003) reports the overall return for an additional year of schooling is 5.4%. Chen and Hamori (2009) study the China Health Nutrition Survey (CHNS) of 2004 and 2006 and present the overall economic returns to schooling are in the range of 7% - 8%.

The conventional approach to estimate the returns to schooling is the standard Mincer earning equation (Mincer, 1974). Although hundreds of papers have examined this issue in many countries, using various data sets, and with different estimation techniques, few studies produce the “true” rate of return to education (Heckman, Lochner, & Petra, 2005).

The main problem discussed in the previous studies on the economic returns to education is the endogeneity of the schooling variable. Chosen years of schooling are not exogenous and tend to be correlated with unobservables in the error term of the earning function. The unobservable can be ability, which is positively correlated with years of schooling and with earnings, giving rise to “ability bias” (Card, 1999). Thus, commonly used estimation method – ordinary least squares (OLS) tends to overestimate the returns to education. Some studies employ various instrumental variables such as parental education (Heckman and Li, 2004; Li and Luo, 2004), parental income (Fleisher, Li, Li and Wang, 2005), and spouse’s education (Trostel et al., 2002). Card (1999) and Card and Lemieux (2001) state that IV estimates of the rate of return to education can be higher or lower than OLS estimates depending on the instrument variables chosen in the analysis. Heckman and Li (2004) use the data from the China Urban Household Income and Expenditure Survey in year 2000 to study the returns to higher education in China. They find the IV estimates of the economic returns to four-year college education are higher than the OLS estimates of the Mincer model. Fleisher, Li, Li and Wang (2005) use data from CHIP 1988, 1995, and 2002 to examine how selection and sorting affected the change of the economic returns to schooling for college graduates during China’s reforms between 1988 and 2002. They find the IV estimates of the returns to college education are sensitive to the use of a proxy for ability.

Panel data estimation can be another approach to control for unobserved individual fixed effects. Unfortunately, panel data are often not available, particularly for developing countries. Therefore, individual cross-sectional data are most commonly used in the study of returns to education in developing countries.

Most existing literatures directly follow the Mincer earning function to study the economic returns to education. The variable “years of schooling”, which is a proxy for investment in human capital, is the only

measure of education entering the regression analysis. Thus, by employing a Mincer earning function, previous studies simply assume the marginal returns to an additional year of schooling is constant and ignore the possibility that returns to schooling can vary at different level of education. In addition, previous studies estimate work experience, a proxy for general and specific job training in human capital theory (Mincer, 1974), on the basis of age and years of schooling but ignore the possibility that some group of people such as Intellectual Youth may spend time on totally unrelated jobs, for example, moving rocks, or growing plants. Intellectual Youth were sent to countryside for “re-education” during the Cultural Revolution when they were students in middle school and high school. Thus, estimates of job experience may be overestimated and therefore underestimate the returns to job experience, which influence the estimation of the returns to schooling.

Therefore, this study seeks to answer two questions: first, should the Mincer earning function be directly applied to estimate the returns to schooling in China? Second, whether the returns to schooling in China have been underestimated. The major findings are as follows. The study suggests a modified Mincer earning equation with quadratic term of years of schooling to reflect the differences in average rate of returns to additional year of schooling among various levels of education. The OLS estimates of returns to education are lower than those from the IV approach. The overall returns to education from OLS estimation based on the standard Mincer equation are in the range of 7%-8%. The IV approach reports a much higher rate of returns to education, 16% for college graduate, based upon the modified Mincer earning equation, while the figure from the OLS estimation is 10.24%. The household registration system plays a very important role in the determination of wage. Rural households tend to earn 30%-40% less than urban households. Workers working in state-owned enterprises tend to earn more than private sector.

The rest of the paper is organized as follows. The following section discusses the empirical model and data. Section 3 presents estimation results. Section 4 concludes.

The Empirical Framework and data

This study begins with the basic human capital earning function (Mincer, 1974):

$$\ln w_i = \alpha_0 + \beta_0 S_i + \beta_1 X_i + \beta_2 X_i^2 + \epsilon_i \quad (1)$$

where $\ln w_i$ denotes the natural logarithm of the hourly wage for individual i ; S_i denotes years of schooling of individual i ; X_i represents number of years of job experience (or age) for individual i ; ϵ_i is the error term, and β_0 shows the average rate of return to one additional year of schooling.

The ordinary least square (OLS) estimation procedure has been extensively used in previous studies. However, the OLS estimates may bias the rate of return to education due to the omission of the ability variables. Standard critique emphasizes an upward bias because of the positive correlation among the ability, earning and years of schooling. Controversies still exist regarding whether the omitted ability

variables bias the OLS estimates, and if they do, how much the potential bias would be. This paper will follow this standard approach and present the result from the OLS estimation first. In order to address the endogeneity problem, an instrument variable approach is also adopted in this study by using spouse's education as an instrument (Trostel et al., 2002). Spouse's education is correlated with one's schooling but not with one's wage rate and education is one of the positive assortative traits in marriage (Jepsen and Jepsen, 2002). The following two-equation framework is adopted to address the endogeneity problem of schooling:

$$\ln w_i = \alpha_0 + \beta_0 S_i + \beta_1 X_i + \beta_2 X_i^2 + \epsilon_i \quad (2)$$

$$S_i = \gamma Z_i' + u_i \quad (3)$$

where Z_i represents the vector of the instrument variable and the other exogenous variables.

This study uses urban sample of the third wave (in year 2002, the first and the second wave of data were collected in 1988 and 1995, respectively) of data from China Household Income Project (CHIP), a national cross-sectional study collectively designed by a team of Chinese and western economists and conducted by the Institute of Economics at the Chinese Academy of Social Science. The CHIP-2002 survey covers about 6835 households and 20632 individuals in urban areas of China.

CHIP-2002 contains very rich information about the earnings, which include regular salary/wage, income from secondary jobs, all kinds of transfer income and all these earning measures are in annual terms. Since this study focus on wages determination, I only look at regular salary/wage for annual earnings and exclude all the other sources of income. In separate survey questions, the respondents are interviewed to indicate the number of work-hours in an average day and the average number of work days per week. Thus, the hourly wage rate can be calculated by using annual earnings and the number of work-hours. Before CHIP-95, due to data limitation, annual earning, instead of hourly wage rate, was used as dependent variable in the analysis. Li (2003) states using annual earnings will underestimate earning differences between different educational groups as highly educated people tend to work fewer hours than less educated people.

Another advantage shared by both CHIP-2002 and CHIP-95 is respondents were interviewed to indicate at which year they found their first jobs, which makes it possible to calculate the actual job experience of the respondents, while the earlier studies estimate job experience through age and years of schooling. According to the human capital theory, job experience is a proxy for human capital investment on the job. At the beginning of one's career, human capital accumulation dominates depreciation. As years of experience increases, human capital depreciation will eventually dominate accumulation, thus, job experience enters the earning function in linear form as well as quadratic form. However, there is still a problem in estimating the rate of return to schooling when estimated actual job experience is employed in the regression as it may not reflect the possibility that one may be idle for a while to land a job. Thus, job experience may be overestimated while the returns to schooling may be underestimated if one has been unemployed for some time between the year he got his first job and year 2002.

There is an improvement in the measure of education as CHIP-2002 collects data on eight education levels based upon degrees as well as years of schooling through a separate interview question. If education levels are employed to estimate the returns to schooling, a particular number of years should be assigned to each education level. The existing literature takes the following transformation scheme for each education level: Graduate (Yanjiusheng, 19 years), college/university (Daxue, 16 years), Junior college (Dazhuan, 15 years), Technical secondary school (Zhongzhuan, 12 years), Senior middle school (Gaozhong, 12 years), Junior middle school (Chuzhong, 9 years), elementary school (Xiaoxue, 6 years) and Classes for eliminating illiteracy (2 years). Since some students may stay at the same grade for more than one year, using education level as measure of education may underestimate the years of schooling, and, therefore, overestimate the returns to schooling. In this study, years of schooling directly from the interview question will be used to measure the education background of the respondents.

A problem may arise in the standard Mincer earning equation as it assumes the average returns to schooling is constant and the measure of education is assumed to be linear with the earning variable, which neglects the possibility that contributions of various education levels to earning can be different. Thus, in this study I adopt a modified version of Mincer earning equation as follows:

$$\ln w_i = \alpha_0 + \beta_0 S_i + \beta_1 S_i^2 + \beta_2 X_i + \beta_3 X_i^2 + \epsilon_i \quad (4)$$

where years of schooling S_i and job experience X_i appear in both linear and quadratic terms.

Since this study focus on the contribution of schooling to wage, I restrict the sample to fulltime workers who were legally employed for the whole year at the time of interview. I also exclude those people younger than 16 years old as students are mandated to be in school before 16 years old by compulsory education law. Observations with inconsistent information on age and actual job experience and self-employed workers who report zero salary/wage are excluded as well. After exclusion, the sample contains about 8925 observations.

A few dummy variables are generated to capture the impact on earnings from the perspective of sex, ownership, household registration system and ways to find the jobs. D_female is set as 1 for females and 0 for males. D_SOE is set as 1 if the employers are state/provincial/local government enterprises. D_hukou is set as 1 for rural household, and as 0 for urban household registration. D_guanxi is set as 1 if the job is introduced by relatives or friends, and as 0 if otherwise. Intellectual Youth are a group of students who were sent to countryside for “re-education” by conducting heavy-labor work during the Cultural Revolution when the value of education was highly ignored. Thus, it is interesting to examine the wage determination for this special group of people and dummy variable D_IYouth is generated as 1 for being Intellectual Youth during Cultural Revolution, 0 for not. Table 1 reports the definition for each variable.

Results and Discussion

Table 2 presents the descriptive statistics of all the variables.

As discussed in previous section, the Mincer earning equation is estimated for hourly wage rate through OLS analysis. Two proxies are employed for the work experience: the age and the actual work experience. Household registration system (Hukou) is strictly implemented in China and becomes a constraint for a few issues such as job-hunting, school enrollment, healthcare benefits. A dummy variable for household registration system D_hukou is added to examine the impact of household registration system on wage rates. In addition, market condition reflected by the ownership of the enterprises can also influence the wage rate. Thus, a dummy variable for state owned enterprises D_SOE is incorporated in the regression. The wage difference between sex is another focus of study and a dummy variable D_female is generated to capture this effect. The results based on hourly wage by using OLS estimation procedure are given in Table 3.

The results from Table 3 show that the OLS estimates of return to education in 2002 are in the range of 7%-8%, which are below the world average. Households registered as rural earned about 40% less than urban households. Females earn about 13%-16% less than males while people work in state-owned enterprises earn around 32%-33% more than other enterprises. The impact of having experience as Intellectual Youth is insignificant in this regression and thus ignored. Only 879 observations reported their jobs were introduced by relatives/friends (through Guanxi) and the impact is insignificant and thus ignored as well.

Households who were not originally born in cities can obtain Hukou through the following ways: going to college, becoming officials, joining the army, land occupied by government, buying a house in the city and buying a Hukou. Therefore, the wage difference between rural and urban household may be overestimated as those people who obtain Hukou by going to college or becoming officials tend to earn more due to higher education background. In order to capture the pure effect from Hukou on the wage rate, OLS regression is conducted again in the sample excluding people who obtained their Hukou by going to college or becoming officials and the results are presented in Table 4. Clearly, the results from Table 4 show that household registered as rural earn less than urban household while the magnitude decreases by about 5%-6% when controlling for the way to obtain Hukou.

As explained in previous section, the Mincer earning equation ignores the possibility that the marginal returns to schooling can be different among various levels of education. To address this problem, a modified Mincer earning equation is proposed as equation (4) where a measure of education appears in both linear and quadratic forms. Table 5 reports the results of the OLS estimation on equation (4).

As Table 5 presents, the linear term of years of schooling becomes insignificant while the quadratic term of schooling is significant in the estimation on the modified Mincer earning equation. By using the model with actual experience, the average return to additional year of schooling is 0.0064 times number of years of schooling. Thus, it is predicted that the higher education one receives, the higher returns to

schooling. For example, for a college graduate, the average return to schooling is $0.0064 \times 16 = 10.24\%$; while for a junior middle school graduate, the average returns to schooling is $0.0064 \times 9 = 5.76\%$.

To address the endogeneity problem of schooling, this study follows the IV approach and matches the individual data with the household data. The data on head of household and spouse are pulled out. Spouse's education is used as a proxy for the years of schooling of head of household as education is a positive assortative trait. Spouse's education is related with years of schooling of head of household, but not connected with the wage rate. Table 6 presents the results from 2SLS estimation. Again, since quadratic term of schooling appears in the second stage of regression, the average return to schooling is about $0.01x$ years of schooling. For example, the average rate of return to schooling for a college graduate is $0.01 \times 16 = 16\%$, which is much higher than the OLS estimate 10.24%. Thus, the estimates of returns to schooling from the IV approach are higher than those from the OLS approach and the OLS estimates tend to underestimate the rate of returns to schooling.

CONCLUSION

This study uses the most recent data from the Chinese Household Income Project (2002) to examine the returns to education in urban China. The main goals are to investigate two questions: first, is standard Mincer earning equation outdated to estimate the returns to education in China? Second, has the rate of returns to education been underestimated from the OLS technique compared to the 2SLS approach?

The study suggests a modified Mincer earning equation with quadratic term of years of schooling to reflect the differences in average rate of returns to additional year of schooling among various levels of education. The OLS estimates of returns to education are lower than those from the IV approach. The overall returns to education from OLS estimation based on the standard Mincer equation are in the range of 7%-8%. The IV approach reports a much higher rate of returns to education, 16% for college graduate, based upon the modified Mincer earning equation, while the figure from the OLS estimation is 10.24%. Household registration system plays a very important role in the determination of wage. Rural households tend to earn 30%-40% less than urban households. Workers working in state-owned enterprises tend to earn more than private sector.

The study did not consider measurement errors, which can be an extension for future studies.

REFERENCES

- Card, D. (1999). "The causal effect of education on earnings." In O. Ashenfelter, & D. Card (Eds.) *Handbook of Labor Economics, Vol. 3b.* (pp. 1801- 1863)
- Card, D. & T. Lemieux. (2001). "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis." *The Quarterly Journal of Economics*, 116(2), 705 – 746.
- Chen, Guifu and Shigeyuki Hamori. (2009). "Economic Returns to Schooling in Urban China: OLS and the Instrumental Variables Approach." *China Economic Review* 20 (2009), 143 – 152.

- Fleisher, B.M.; Li, H.; Li, S. and X. Wang (2005). "Sorting, Selection, and Transformation of Return to College Education in China." Working Paper Nos. 05-07, Ohio State University.
- Heckman, J.J. and X. Li. (2004). "Selection Bias, Comparative Advantage and Heterogeneous Returns to Education: Evidence from China in 2000." *Pacific Economic Review*, 9(3), 155 – 171.
- Heckman, J.; Lochner, L. and T. Petra. (2005). "Earnings Functions, Rate of Return, and Treatment Effects: The Mincer Equation and Beyond." NBER Working Paper: 11544.
- Jepsen, Christopher and Lisa Jepsen. (2002). "An Empirical Analysis of the Matching Patterns of Same-Sex and Opposite-Sex Couples." *Demography*, 39(3), pp.435-453
- Johnson, E.N. and G.C. Chow. (1997). "Rates of Return to Schooling in China." *Pacific Economic Review*, 2(2), 101 – 113.
- Li, H. (2003). "Economic Transition and Returns to Education in China." *Economics of Education Review*, 22(3), 317 – 328.
- Li, H. and Y. Luo. (2004). "Reporting Errors, Ability Heterogeneity, and Returns to Schooling in China." *Pacific Economic Review*, 9(3), 191–207.
- Liu, Z. (1998). "Earnings, Education and Economic Reforms in Urban China." *Economic Development and Cultural Change*, 46(4), 697–725.
- Mincer, J. (1974). "Schooling, Experience and Earnings." New York NBER.
- Psacharopoulos, G. (1994). "Returns to Investment in Education: a Global Update." *World Development*, 22(9), 1325–1343.
- Psacharopoulos, G., and A.P. Harry. (2002). "Returns to Investment in Education: a Further Update." World Bank Policy Research Working Paper 2881.
- Trostel, P.; Walker, I. and P. Woolley. (2002). "Estimates of the Economic Return to Schooling for 28 Countries." *Labor Economics*, 9(1), 1- 16.

Table 1: Definition of Variables

Variable	Definition
Wage	Natural Log of Hourly Wage Rate
Age	Age in Years
Act_EXP	Actual Work Experience by Year 2002
Schooling	Years of Schooling
D_hukou	1 for Household Registered as Rural; 0 for Household Registered as Urban
D_female	1 for Females and 0 for Males
D_SOE	1 for Ownership of Present Work Unit as State/Provincial/Local State Enterprises; 0 for Private/Share-Holding Companies
D_guanxi	1 if Job is introduced by Relatives/Friends; 0 if Otherwise.
D_IYouth	1 for Intellectual Youth, 0 for not.

Table 2: Descriptive statistics of all variables

Variable	Number of Observations	Mean	Standard Deviation
Wage	8925	1.472	1.8096
Age	8996	40.7023	9.0738
Act_EXP	8968	20.5467	9.5908
Schooling	8996	11.5728	2.9653
D_hukou	8996	0.011	0.1043
D_female	8996	0.4369	0.496
D_SOE	7239	0.5959	0.4907
D_guanxi	9445	0.0931	0.2905
D_IYouth	11398	0.2092	0.4067

Source of Data: China Household Income Project 2002 (CHIP-02), obtained from Inter-University Consortium for Political and Social Research

Table 3: Rates of Return to Education (Fulltime Workers, OLS)

Dependent Variable	Wage	Wage
Schooling	0.0774** (23.58)	0.0808** (24.68)
Age	0.0948** (12.98)	
Age_Square	-0.0010** (-11.34)	
Act_EXP		0.0373** (10.34)
Act_EXP_Square		-0.0005** (-6.02)
D_hukou	-0.4274** (-4.99)	-0.3823** (-4.44)
D_female	-0.1569** (-8.36)	-0.1341** (-7.17)
D_SOE	0.3311** (17.29)	0.3196** (16.74)
Constant	-1.6796** (-11.03)	-0.1535** (-2.77)
Summary Statistics	$\bar{R}^2 = 0.1806$ F=243.23 N=6597	$\bar{R}^2 = 0.1851$ F=250.16 N=6584

Numbers in parentheses are t-statistics

**Denotes that variables are significant at the 5% level

Table 4: Rate of Returns to Education (Full-Time Workers, Excluding Individuals who Obtained Hukou by Going to College or Becoming Officials, OLS)

Dependent Variable	Wage	Wage
Schooling	0.0820** (23.84)	0.0842** (24.63)
Age	0.0944** (12.49)	
Age_Square	-0.0010** (-10.89)	
Act_EXP		0.0373** (10.06)
Act_EXP_square		-0.0005** (-5.84)
D_hukou	-0.3712** (-4.30)	-0.3227** (-3.72)
D_female	-0.1552** (-8.02)	-0.1317** (-6.82)
D_SOE	0.3420** (17.43)	0.3307** (16.90)
Constant	-1.7299** (-11.04)	-0.1975** (-3.47)
Summary Statistics	$\bar{R}^2 = 0.1858$ F=236.49 N=6193	$\bar{R}^2 = 0.1897$ F=242.03 N=6180

Numbers in parentheses are t-statistics

**Denotes that variables are significant at the 5% level

Table 5: Rate of Returns to Education (Full-Time Workers, OLS, Based on Equation (4) Modified Mincer Earning Equation)

Dependent Variable	Wage	Wage
Schooling	0.0189 (1.30)	0.0121 (0.84)
Schooling_Square	0.0027** (4.15)	0.0032** (4.90)
Age	0.0962** (13.12)	
Age_Aquare	-0.0010** (-11.52)	
Act_EXP		0.0387** (10.71)
Act_EXP_Square		-0.0006** (-6.31)
D_hukou	-0.4325** (-5.05)	-0.3860** (-4.50)
D_Female	-0.1530** (-8.15)	-0.1290** (-6.89)
D_SOE	0.3394** (17.22)	0.3170** (16.62)
Constant	-1.412** (-8.55)	-0.1790** (-2.05)
Summary Statistics	$\bar{R}^2 = 0.1826$ F=211.45 N=6597	$\bar{R}^2 = 0.1879$ F=218.61 N=6584

Numbers in parentheses are t-statistics

**Denotes that variables are significant at the 5% level

Table 6: Rate of Returns to Education (Head of Household and Spouse, 2SLS)

	1 st (Schooling)	Stage (Wage)	2 nd (Wage)	Stage	1 st (Schooling)	Stage (Wage)	2 nd (Wage)	Stage
Schooling_Hat_Square			0.0047** (10.17)				0.005** (10.80)	
IV: Spouse's Education	0.4650** (29.37)				0.4600** (29.27)			
Age	-0.3743** (-6.62)		0.1492** (8.35)					
Age_Square	0.0039** (6.06)		-0.0016** (-8.24)					
Act_EXP					-0.1715** (-6.37)		0.0469** (5.50)	
Act_EXP_Square					0.0027** (4.78)		-0.0007 (-4.22)	
D_hukou	-1.6187** (-3.22)		-0.3515** (-2.32)		-1.9175** (-3.81)		-0.2636* (-1.73)	
D_Female	-0.2631** (-2.52)		-0.0336 (-1.09)		-0.2743** (-2.63)		-0.0144 (-0.47)	
D_SOE	0.7130** (7.33.)		0.3139** (10.13)		0.7333** (7.55)		0.2947** (9.48)	
Constant	14.5236** (11.52)		-2.5472** (-6.13)		8.2242** (22.11)		0.0528 (0.40)	
Summary Statistics	$\bar{R}^2 = 0.3016$		$\bar{R}^2 = 0.1139$		$\bar{R}^2 = 0.3058$		$\bar{R}^2 = 0.1084$	
	F=214.84		F=64.17		F=218.64		F=60.58	
	N=2972		N=2949		N=2965		N=2942	

Numbers in parentheses are t-statistics

**Denotes that variables are significant at the 5% level

*Denotes that variables are significant at the 10% level