

Returns to Education and Informal Labour Markets in Urban China: Evidence from CULS3

QU Xiaobo

Institute of Population and Labor Economics, CASS

Albert Park

Department of Economics, the Hong Kong University of Science and Technology

This paper, based on the latest CULS3 data^①, provides empirical evidence and demonstrates features of returns to education by comparing local vs. migrant labors, formal vs. informal employment, formal vs. informal jobs and formal vs. informal sectors in China urban area. In particular we still focus on the difference of returns to regular education and vocational education among different group. Besides, we also mainly pay attention to the influence on individual's returns to education if other family members are engaged in formal employment and if characteristics of other family members are controlled, including household size, share laborers, share male labourer, mean education of other labor, highest education of other labor, mean age of other labor, and if certain city variables are controlled, such as GDP per capita, GDP growth rate, government spending per capita, etc.

The plan of this paper is as follows. In the first section, we present a brief introduction and provide a brief review literature and informality of labor market in urban china on this subject. In the second section, we give some statistical description of relevant variable used in this paper. The next section discusses and analyzes the estimation results by using Heckman, switching regression, quantile regression and OLS methods. The final section makes a preliminary conclusion based on the empirical analysis.

1. Introduction

The urban economy in the most developing counties can be classified into the formal or modern sector and the informal or traditional sector and is also a universal phenomenon (Kazuhiro, 2007). This has become a stylized fact that the informal sector is often the most important employment opportunity in developing countries (Isabel & Andrey, 2011). Extensively researched

^① CULS3 data is third round China urban labor survey.

topics for urban informality mainly focus on two aspects; one is features of informal sector and distribution; the other is theoretical research on the causes and consequences of informality. For example, According to ILO (1972, 2002), Kazuhiro (2007), etc., the informal sector is characterized by small firm sizes, unskilled jobs, low wages, and the loose enforcement of laws and regulations, and it incorporates sectors such as petty trading, domestic services, repair services, and basic manufacturing. According to Jamal and Weeks (1988), rural–urban income differentials are substantial in many sub-Saharan African nations but the income differentials for unskilled workers are very small. Ranis and Stewart (1999) find that about 70% of the labor forces in the informal sector in Manila are engaged in trading and only about 13% are in manufacturing. Similar numbers are found for many Latin American economies as well (Marcouiller et al., 1997). For theoretical research on the causes and consequences of informality, it can be seen from Maloney (2004), Fields (2005), Loayza and Rigolini (2006), Amaral and Quintin (2006), Bennett and Estrin (2007), Bennett (2008) and World Bank (2007). However, this literature is dominated by the a priori assumption that the informal sector is homogenous. In fact, there some evident differences in the different type of informal employment. Results of informal sector employment is to associate with earnings of work, especially the proliferation of informal self-employment in developing countries is not economically inefficient (Gustavo Yamada, 1996).

By nearly any measure, China is the world’s largest transition economy and the world’s largest developing country. So that improved understanding of informality in China may provide key lessons and insights for other countries. Beginning in the late 1990s, China’s urban labor market experienced a rapid diversification of ownership types. From 1995 to 2005, the percentage of China’s urban labour force that was undocumented (not officially registered as employees of formal sector work units or as self-employed workers) increased from 9 percent to 36 percent (Park & Cai, 2008). Despite the great importance of research on informal issues, there is a glaring lack of research and data sources for studying China’s informal labor market. The few studies of informal employment in China, mostly completed by members of this project’s research team, are primarily descriptive due to data limitations, and focus on defining and documenting the extent of informal employment (Cai and Park, 2007; Wu and Cai, 2006). China passed a landmark Labor Law in 1994, formalizing for the first time the system of labor contracting and employment (Cai, Du, and Wang, 2009).

At the same time, the last decade has witnessed an explosion of rural-urban migration in China. Data from the most recent NBS national rural household survey finds that rural migrants numbered 153 million in 2010, with up to ninety percent migrating to urban areas, where they form a large constituency within the growing informal sector. Like the rest of the world, China has

been negatively impacted by the current global economic crisis. The informal sector, especially through self-employment opportunities, provides a cushion, or alternative employment option, for dislocated workers from both the formal and informal sectors. This paper provides empirical evidence and features between earnings of education and different type of Informal Labour employment, to be helpful understanding informal sector and employment results in China.

2. Statistical Description and Features

Table 1-1 and table 1-2 are statistics description of relevant variables. In this draft paper we focus on analysis of returns to education with different labor group in China urban labor market, we select samples women aged 16 to 55 and men aged 16 to 60 for CULS3 data. It should be pointed out that the Mean-value of all variables in the table1 is weighted^①.

We can see from the statistics of table 1-1 that the average age of migrants is younger than average age of local urban labor and in pooled samples, it is 38.5. The average age of local urban labor almost equals average age of labor in pooled samples. Distribution of gender is basically similar in the three groups, with proportion of male accounts for about 51%. Distribution of labors' marital status is also similar in these three groups. However, levels of schooling shows considerable differences. Firstly, it is apparent from the table1-1 that proportion of middle school and below in the migrants is the highest, subsequently in pooled samples. The proportion of middle school and below in the local labor is only 28.1%. Secondly, the proportion of regular high school and vocational high school in local labor is 31.3% and 9.64% respectively, which is higher than migrants group. Moreover, the proportion of vocational higher education in the local labor is also higher than migrants group. As for the level of regular higher education, the proportion of college and above in the local labor is 13.9%, the proportion of migrants is 8.8%.

Year of schooling in local labor is 12.2, which is higher than migrants group and all samples. Therefore, both level of education and years of schooling show a significant predominance in the local labor than migrants group in terms of human capital. Proportion of children under 16 for migrant family group is 32.2%, which is higher than local labor and pooled samples. Seemingly, Proportion of old people over 60 for the migrant family is 23.93%, higher than local labor and

^① The weights for urban resident sample is defined simply as the city local resident population (form urban statistical yearbook) divided by the city sample size (number of households).

The migrant weights correct for the probability that a neighborhood is selected (based on the number of local resident households, from sampling data or from the neighborhood form if sampling data is missing), the probability that a migrant is chosen in a given neighborhood (based on the size of the migrant population in the neighborhood according to the neighborhood form), and the ratio of the city migrant population to the city migrant household sample size.

pooled sample. This shows that Family structure presents a visible difference between in local labor and migrants. But interestingly, we can see from table1 that monthly wage of migrants group is higher than that of pooled sample and local labor, which are 2424.0, 2394.1 and 2453.0 yuan, respectively. The reason might be that Weekly working hours in migrants group is higher than that of the pooled sample and local labor, which is 54.79, 49.11 and 44.60 hours, respectively.. Accordingly, hourly wage in local labor is higher than hourly wage in migrants group. Finally, due to difference of the size of population between migrants and urban, it is can be seen from table1 that sample profiles of CULS3 data have evident difference in both labor groups.

Table1-1: Statistical Description of Relevant Variable for Local and Migrants

Variable	All	Local	Migrants
Age	38.5	39.5	33.8
Sex (male=1)	51.3%	51.4%	50.8%
Marriage (married=1)	76.0%	76.9%	76.1%
Middle school and below	32.4%	28.1%	53.1%
Regular high school	29.5%	31.3%	21.3%
Vocational high school	9.11%	9.64%	6.64%
Vocational higher education	15.9%	17.1%	10.2%
Regular college and above	13.0%	13.9%	8.80%
Years of Schooling	11.9	12.2	10.6
Work experience	22.0	23.0	17.4
Proportion of children under 16 in the family	26.4%	25.2%	32.2%
Proportion of old people up 60 in the family	15.0%	13.4%	22.8%
Monthly wage (yuan/month)	2424.0	2394.1	2453.0
Weekly working hours (hour)	47.3	45.0	55.8
Hourly wage (yuan/hour)	11.9	12.8	10.9
Shanghai	15.2%	16.9%	13.8%
Wuhan	18.6%	17.0%	19.8%
Shenyang	15.4%	15.1%	15.6%
Fuzhou	16.3%	15.3%	17.0%
Xian	17.0%	17.7%	16.5%
Guangzhou	17.7%	18.1%	17.3%
Obs	11,041	5432	5640

Note: in the Levels of schooling, Middle school and below=1, Regular high school =2, Vocational high school =3, Vocational higher education =4, college and above=5. Mean-value of all variables is weighted.

It is can be seen from table 1-2 that the average age of migrants is younger than average age of formal employment and in formal sector, which is 39.4 and 38.2 respectively. The average age of informal employment has a slightly difference in the informal sector. Distribution of gender is basically similar in both formal employment and formal sector groups, with proportion of male accounts for about 57%. Distribution of labors' marital status is also similar in both formal employment and in formal sector. However, levels of schooling shows considerable differences. Firstly, it is apparent from the table1-1 that proportion of middle school and below in the informal sector is the highest, subsequently in the informal employment. The proportion of middle school and below in the formal employment is only 19.3%. Secondly, the proportion of regular high school and vocational high school in the formal employment is 27.9% and 10.3% respectively, which is higher than informal employment group and informal sector group. Moreover, the proportion of vocational higher education in the formal employment group is also higher than formal sector group. As for the level of regular higher education, the proportion of college and above in the formal employment is 20.6%, the proportion of informal employment is 5.5%.

Table 1-2: Statistical Description of Relevant Variable for Formal vs. Informal Employment and Formal vs. Informal Sector

Variable	Employment		Sector	
	Formal	Informal	Formal	Informal
Age	39.4	35.6	38.2	38.6
Sex (male=1)	58.7%	52.2%	57.4	53.8
Marriage (married=1)	80.9%	79.6%	78.9%	87.7%
Middle school and below	19.3%	50.8%	22.8%	55.8%
Regular high school	27.9%	25.2%	27.1%	27.1%
Vocational high school	10.3%	8.5%	10.2%	7.5%
Vocational higher education	22.0%	10.1%	21.0%	6.6%
Regular college and above	20.6%	5.5%	18.9%	3.0%
Years of Schooling	13.0	10.6	12.7	10.2
Work experience	20.4	18.7	19.4	22.1
Proportion of children under 16 in the family	27.2%	31.2%	27.6%	31.8%
Proportion of old people up 60 in the family	14.9%	18.4%	15.8%	16.8%
Monthly wage (yuan/month)	2654.8	2569.6	2622.9	2658.3
Weekly working hours (hour)	44.2	60.1	46.4	65.7

Hourly wage (yuan/hour)	14.4	9.5	14.3%	11.1
Shanghai	16.7%	14.5%	17.9%	10.3%
Wuhan	18.4%	20.1%	16.0%	26.5%
Shenyang	13.4%	16.1%	13.3%	18.2%
Fuzhou	17.5%	15.0%	17.8%	12.9%
Xian	13.3%	20.0%	16.1%	18.4%
Guangzhou	20.7%	14.3%	19.0%	13.7%
Obs	5278	5757	7570	3471

Note: in the Levels of schooling, Middle school and below=1, Regular high school =2, Vocational high school =3, Vocational higher education =4, college and above=5. Mean-value of all variables is weighted.

It is seen from table 1-2, year of schooling in formal employment is 13.2, which is higher than informal employment group and all samples; year of schooling in formal sector group is 12.7, which is higher than informal sector group. Therefore, both level of education and years of schooling show a significant predominance in the formal employment and formal sector than in the informal employment and informal sector group in terms of human capital. Proportion of children under 16 for informal employment group is 31.2%, which is higher than formal employment and formal sector samples. Seemingly, Proportion of old people over 60 for the informal employment group is 18.4%, higher than formal employment and formal sector sample. This shows that Family structure presents a visible difference between in formal employment and informal employment. But interestingly, we can see from table1-2 that monthly wage of informal sector group is higher than that of formal sector and formal employment group, which are 2658.3, 2622.9 and 2654.8 yuan, respectively. The reason might be that Weekly working hours in informal sector group is higher than that of the formal employment and formal sector, which are 65.7, 44.2 and 46.4 hours, respectively. Accordingly, hourly wage in formal sector group is higher than hourly wage in informal sector group. Finally, due to difference of the size between formal employment and informal employment, it is can be seen from table1-2 that the size of informal sector in Shanghai and Guangzhou are the smallest, compared to other four cities. The size of informal sector in Wuhan, Shenyang, Fuzhou and Xi'an is comparatively high.

Table1-3: Variance of log -Hourly Wage in Different Educational Levels for Migrants and Local

	Local		Migrants	
	Mean	standard deviation	Mean	standard deviation

Middle school and below	1.99	0.640	1.95	0.696
Regular high school	2.18	0.634	2.18	0.751
Vocational high school	2.27	0.549	2.17	0.697
Vocational higher education	2.48	0.589	2.46	0.671
Regular college and above	2.81	0.586	2.90	0.683

Note: Mean-value of all variables is weighted.

Table1-4: Mean of Formal and Informal Workers Hourly Wage for Different Educational Levels

	Formal workers	Informal workers (include self-work)	Informal workers (exclude self-work)
Middle school and below	11.61	7.14	6.80
Regular high school	12.91	8.19	8.02
Vocational high school	12.03	8.89	8.82
Vocational higher education	15.86	10.30	9.76
Regular college and above	21.72	16.90	16.46

Note: Mean-value of all variables is weighted.

Table 1-3 is the distribution of log-hourly wage of different educational levels in local labor and migrants. Table 1-4 shows us the distribution of hourly wage of different educational levels in formal and informal sector, the informal sector falls into two categories: self-employed included and self-employed excluded.

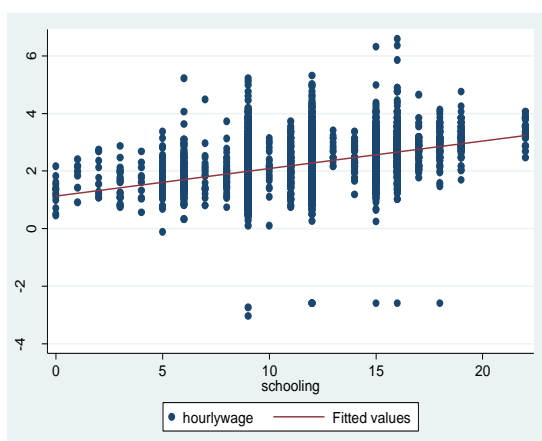


Figure 1-1: Log Hourly wage plot in formal sector

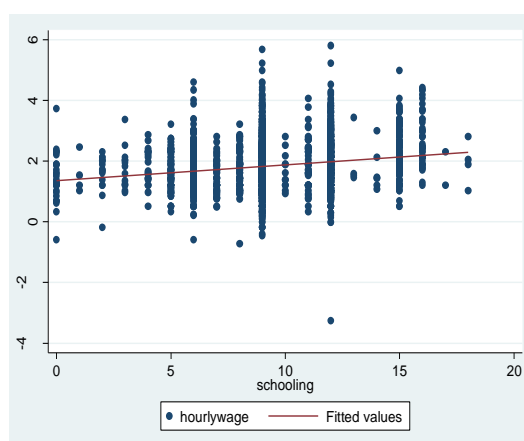


Figure 1-2: Log Hourly wage plot in informal

Figure 1-1 and figure 1-2 are plot and fitted value between log-hourly wage and years of schooling in the formal and informal sector. The figure 1-1 is a plot about formal sector, and the

figure 1-2 is a plot about informal sector. We can see clearly that the line of fitted value in informal sector tend more horizontal than formal sector. That is to say, the slop of fitted value line in formal sector is evident bigger than informal sector.

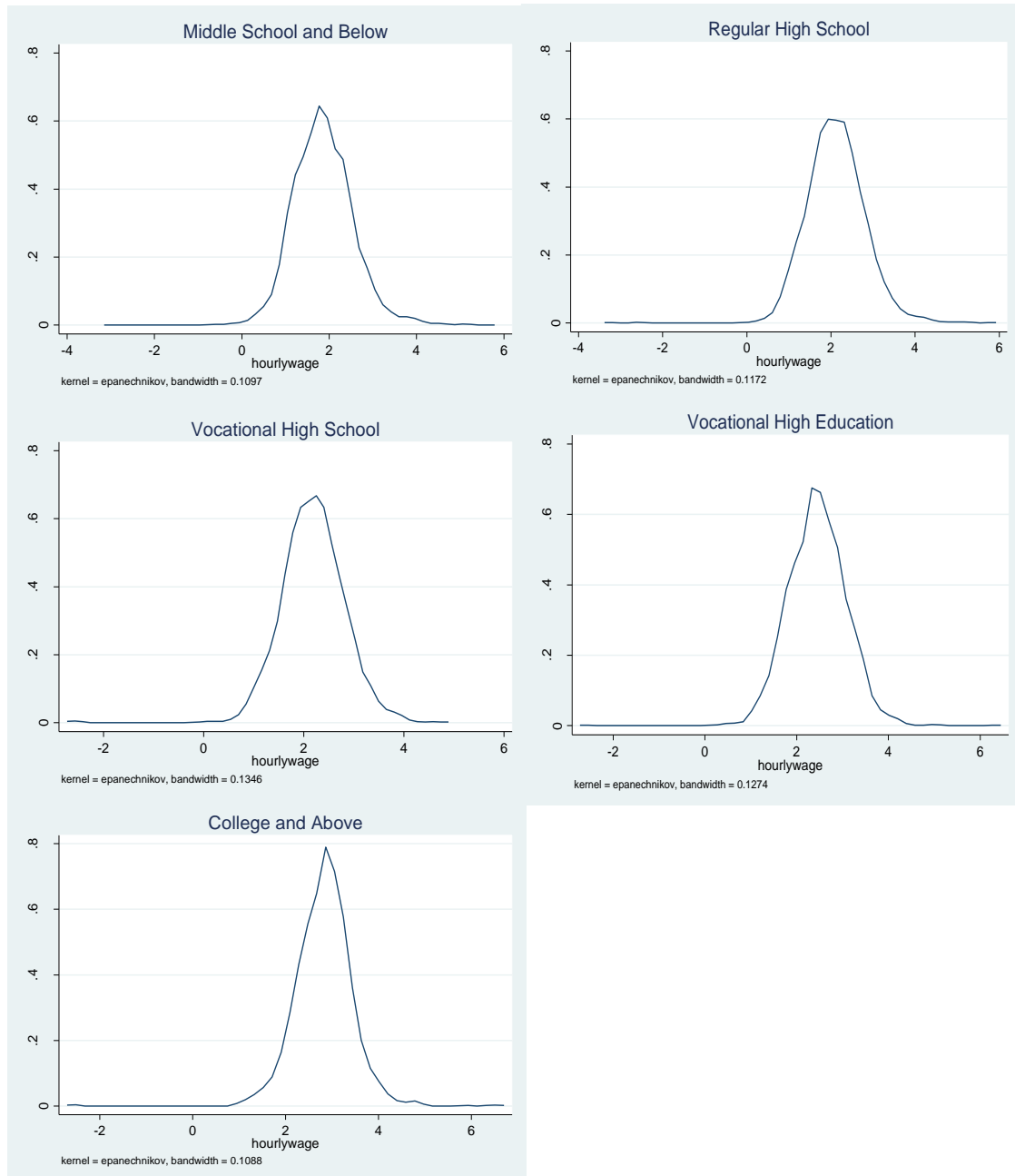


Figure 2: Density distribution of Log hourly wage in different educational level

Figure 2 is density distribution of Log hourly wage in different educational level for pooled sample women aged 16 to 55 and men aged 16 to 60. It is can be seen that with the raise of

educational level, density distribution of log-hourly wage tend likely more concentration. Moreover, density proportions of log-hourly wage also tend higher with the raise of educational level.

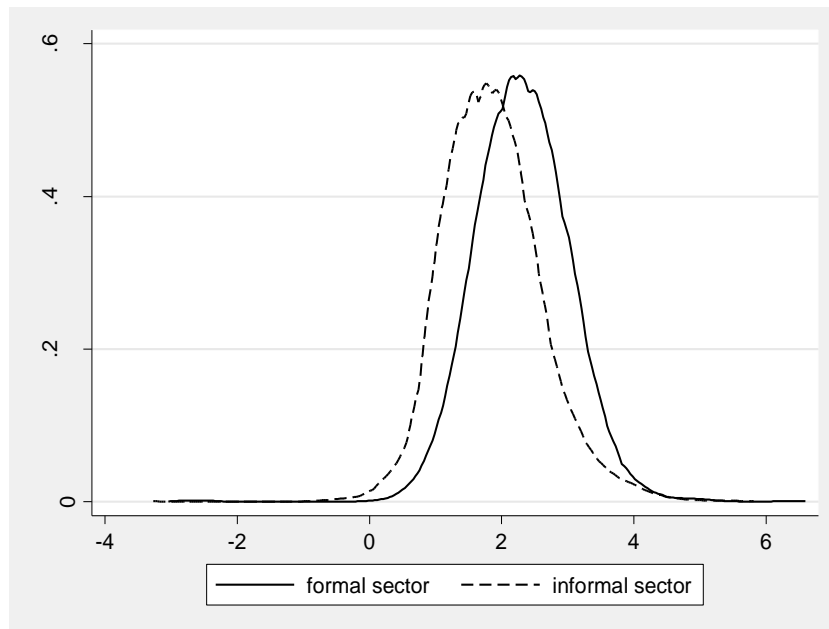


Figure 3: density distributions of log-hourly wages in the formal and informal sector

Figure 3 is the density distribution of log-monthly wages in the formal and informal sector. The horizontal axis stands for log-monthly wages, the vertical axis stands for density distribution in the formal and informal sector. We can see from the figure3 that there is a high degree of overlap between formal and informal sector in terms of log-monthly wages. At the same time, we also see that log-hourly wage of informal sector are more dispersed than that of formal sector form figure3.

3. Returns to Education and Employment Outcomes Using Different Methods

The benchmark model for the development of empirical estimation of the returns to education is the key relationship derived by Mincer (1974). This paper will adopt expanded Mincer wage function to estimate returns to education for China urban labor. Among all specification, we controlled Gender, age, the number of kids under 6 in the family, the number of old people over 60 in the family and city dummy variables, etc. The methods used in this paper

include Heckman, OLS method, quantile regressions and switching regressions, in these specifications of regress that contains controls for work experience and other individual characteristics, such as gender, age, schooling and city dummy variable influence both wage function and select function. In the Heckman regression method, the number of children under six in the family and the number of old people up sixty in the family only influence select function.

3.1 Returns to education on different categories of workers, employment statuses

In the all specification, dependent variable is logarithm of hourly wage. At the same time, we controlled Gender, age, the number of kids under 6 in the family, the number of old people over 60 in the family and city dummy variables, etc. In this section, all regressions use Heckman and OLS method, in these controlling variables, Gender, age, schooling and city dummy variable influence both wage function and select function. In the Heckman method, the number of children under six in the family and the number of old people over sixty in the family only influence select function. In the wage function, observational sample is women aged 16 to 55 and men aged 16 to 60.

Table 2-1: Estimating Results of Pure Mincer by OLS and Heckman f from Pooled Sample (I)

	OLS		Heckman	
	Coef.	P> z	Coef.	P> z
<i>Wage function</i>				
Schooling	0.1019	0.000	0.1062	0.000
experience	0.0176	0.000	0.0260	0.000
experience2	-0.0003	0.000	-0.0006	0.000
_cons	0.9299	0.000	0.7615	0.000
<i>Select function</i>				
child16			-0.0140	0.619
old60			0.0029	0.934
schooling			0.0461	0.000
experience			0.0742	0.000
experience2			-0.0023	0.000
_cons			-0.0218	0.836
lambda			0.2233***	0.0638(Std.)
Obs	10846		10846	

Note: P value is significant probability. “***, **, *” significant at 1%, 5%

Table 2-2: Estimating Results of Pure Mincer by OLS and Heckman from Pooled Sample (II)

	OLS		Heckman	
	Coef.	P> z	Coef.	P> z

<i>Wage function</i>				
Regular high school	0.2259	0.000	0.2179	0.000
Vocational high school	0.2891	0.000	0.2959	0.000
Vocational higher education	0.5817	0.000	0.6008	0.000
College and above	0.9306	0.000	0.9640	0.000
experience	0.0231	0.000	0.0294	0.000
experience2	-0.0005	0.000	-0.0006	0.000
_cons			1.6909	0.000
<i>Select function</i>				
child16			-0.0190	0.509
old60			-0.0030	0.932
Regular high school			-0.1387	0.001
Vocational high school			0.1017	0.116
Vocational higher education			0.3365	0.000
College and above			0.6596	0.000
experience			0.0899	0.000
experience2			-0.0026	0.000
_cons			0.2682	0.000
lambda			0.1480***	0.0472(Std.)
Obs	10846		10846	

Note: P value is significant probability. “***, **, *” significant at 1%, 5%, Sample is women Aged 16 to 55 and Men aged 16 to 60.

Table 2-1 we used OLS and Heckman method to get the estimating Results of pure Mincer equation if adopting years of schooling to measure returns of education. We found that while returns to education in the pooled sample by OLS are 10.19%, it is 10.62% by heckman method.

Table 2-2 we used OLS and Heckman method to get the estimating Results of pure Mincer equation if adopting level of schooling to measure returns of education. We found that returns to education of vocational high school is higher than return to education of regular high school in pooled sample, which are 0.2891 and 0.2959 by OLS and Heckman regression respectively, higher than coefficients of regular high school from table 2-2. It can be seen from Table 3-1 that regression coefficients of college and above are the highest in level of education whether OLS regression or Heckman regression.

Table 2-3: Estimating Results of Adding Dummy Variable into Equation by OLS and Heckman from Pooled Sample (I)

	OLS		Heckman	
	Coef.	P> z	Coef.	P> z
<i>Wage function</i>				
Schooling	0.1048	0.000	0.1076	0.000

experience	0.0259	0.000	0.0334	0.000
experience2	-0.0005	0.000	-0.0008	0.000
Sex	0.2157	0.000	0.2686	0.000
City dummy		Yes		Yes
_cons	0.9373	0.000	0.7852	0.000
<i>Select function</i>				
Schooling			0.0368	0.000
experience			0.0913	0.000
experience2			-0.0028	0.000
Sex			0.7541	0.000
child16			-0.0401	0.170
old60			-0.0074	0.841
City dummy				Yes
_cons			-0.2277	0.061
lambda			0.1856***	0.0642(Std.)
Obs		10844		10844

Note: P value is significant probability. “***, **, *” significant at 1%, 5%, Shanghai takes as reference city. Sample is Aged 16 to 55 and Men aged 16 to 60.

Table 2-4: Estimating Results of Adding Dummy Variable into Equation by OLS and Heckman from Pooled Sample (II)

	OLS		Heckman	
	Coef.	P> z	Coef.	P> z
<i>Wage function</i>				
Schooling	0.1067	0.000	0.1114	0.000
Experience	0.0262	0.000	0.0348	0.000
experience2	-0.0005	0.000	-0.0008	0.000
Sex	0.2145	0.000	0.2746	0.000
Migrants	-0.0353	0.032	-0.0724	0.001
City dummy		Yes		Yes
_cons	0.9013	0.000	0.6973	0.000
<i>Select function</i>				
Schooling			0.0572	0.000
Experience			0.0935	0.000
experience2			-0.0028	0.000
Sex			0.7528	0.000
Migrants			0.5224	0.000
child16			-0.0444	0.125
old60			-0.0251	0.494
City dummy				Yes
_cons			-0.6241	0.000
lambda			0.2156***	0.0600(Std.)
Obs		10844		10844

Note: P value is significant probability. “***, **, *” significant at 1%, 5%, Shanghai take as reference city. Sample is women Aged 16 to 55 and Men aged 16 to 60.

Table 2-3 is that we take dummy variables of sex and city into consideration if adopting years of schooling to measure returns of education. Both OLS and Heckman estimation shows that, Returns to education in both OLS and Heckman regression results increased, specifically, returns to education of OLS regression result is 10.48%, and 10.76% in the Heckman regression.

Table 2-4 is that except for dummy variables of sex and city into consideration, we also take dummy variables of migrants into consideration adopting years of schooling to measure returns of education. Both OLS and Heckman estimation shows that, Both OLS and Heckman estimation shows that, Returns to education in both OLS and Heckman regression results increased slightly, specifically, returns to education of OLS regression result is 10.67%, and 11.14% in the Heckman regression.

Table 3-1: Estimating Results Both Local Workers and Migrants Workers Exclude

	Self-employed (I)			
	Local		Migrant	
	Coef.	P> z	Coef.	P> z
<i>Wage function</i>				
Schooling	0.1141	0.000	0.0939	0.000
Experience	0.0333	0.000	0.0364	0.000
experience2	-0.0007	0.000	-0.0009	0.000
Sex	0.2581	0.000	0.2774	0.000
City dummy		Yes		Yes
_cons	0.6891	0.000	0.9829	0.000
<i>Select function</i>				
Schooling	0.0804	0.000	0.0510	0.000
Experience	0.0930	0.000	0.0402	0.000
experience2	-0.0028	0.000	-0.0014	0.000
Sex	0.7849	0.000	0.8464	0.000
child16	0.0079	0.831	-0.2384	0.000
old60	-0.0398	0.396	0.0165	0.701
City dummy		Yes		Yes
_cons	-0.9109	0.000	0.2282	0.144
lambda	0.1897**	0.0933(Std.)	0.1112***	0.0390(Std.)
Obs		4812		3587

Note: P value is significant probability. “***, **, *” significant at 1%, 5%, Shanghai takes as reference city. Sample is women Aged 16 to 55 and Men aged 16 to 60.

If adopting years of schooling to measure returns of education, it can be seen from table 3-1 that returns to education in local workers and migrant workers is 11.4% and 9.39% respectively, which exclude self-employed in local workers and migrant workers. It suggests that both local workers and migrant workers show a significant growth in terms of returns to education. Moreover, male has dramatically promoted to the hourly wage in both estimating regression, which shows hourly wage of male is higher than female. Table 3-1 shows coefficient of work experience is significantly positive and coefficient of square work experience is significant negative. This explains that effect of work experience on dependent variable (logarithm of hourly wage) presents reverse “U”, which is consistent with the general theory of human capital. From table 3-1 we can see that all variables including, gender, work experience and years of schooling etc have a significant effect on employment option except proportion of children under 16 and proportion of old people over 60 in the family .

Table 3-2: Estimating Results Both Local Workers and Migrants Workers Exclude

	Self-employed (II)			
	Local		Migrant	
	Coef.	P> z	Coef.	P> z
<i>Wage function</i>				
Regular high school	0.2012	0.000	0.2092	0.000
Vocational high school	0.3525	0.000	0.2553	0.000
Vocational higher education	0.6026	0.000	0.5709	0.000
College and above	0.8894	0.000	0.9383	0.000
experience	0.0343	0.000	0.0418	0.000
experience2	-0.0007	0.000	-0.0011	0.000
Sex	0.1976	0.000	0.2813	0.000
City dummy		Yes		Yes
_cons	1.8094	0.000	1.7524	0.000
<i>Select function</i>				
Regular high school	0.0336	0.450	0.1562	0.017
Vocational high school	0.3510	0.000	0.1742	0.089
Vocational higher education	0.6223	0.000	0.3607	0.000
College and above	0.9298	0.000	0.5245	0.000
experience	0.1071	0.000	0.0429	0.000
experience2	-0.0030	0.000	0.0015	0.000
Sex	0.7295	0.000	0.8629	0.000
child16	0.0189	0.555	0.2393	0.000
old60	-0.0654	0.121	0.0153	0.723
City dummy		Yes		Yes
_cons	-0.3048	0.001	0.6237	0.000

lambda	-0.1289*	0.0684(Std.)	0.0737*	0.0370(Std.)
Obs		4812		3587

Note: P value is significant probability. “***, **, *” significant at 1%, 5%, Shanghai takes as reference city. Sample is women Aged 16 to 55 and Men aged 16 to 60.

Table 3-2 is regression results of Heckman method by adopting level of schooling to measure returns of education, which exclude self-employed in local workers and migrant workers. Firstly, it shows returns to education increases with level of education enhancement both in the local workers and in the migrant workers. Moreover, returns to education of Vocational high school are higher than return to education of regular high school in both local workers and migrant workers. After exclude self-employed from migrants’ sample, regression coefficient of vocational high education in migrant employed workers is also higher than regular high school. The probable reason is migrants of vocational high education are mainly engaged in manufacturing, construction, hotels and catering services industries as opposed to local labor, few migrants of vocational high education is self-employed. Secondly, it can be seen from table 3-2 that regression coefficients of college and above are the highest in level of education whether in local workers and in migrant workers. Interestingly, regression coefficient of college and above in migrant group is higher than in local workers. Regression coefficients of college and above are 0.8894 and 0.9383 in both group samples, respectively. It is clearly that returns to education for college and above in the migrants is higher than local labor. Therefore, table 3-2 shows that migration contributes to improving returns to education for labor.

Table 4-1: Estimating Results of Returns to Education of Self-employed

Years of Schooling		Level of Education	
<i>Wage function</i>		<i>Wage function</i>	
Schooling	0.0713***	Regular high school	0.2648***
experience	0.0234**	Vocational high school	0.1806***
experience2	-0.0006***	Vocational higher education	0.5010***
Sex	0.2524***	College and above	0.8274***
City dummy	Yes	experience	0.0253***
_cons	1.1475***	experience2	-0.0007***
<i>Select function</i>		<i>Select function</i>	
Schooling	0.0114	Sex	0.2456***
experience	0.0354	City dummy	Yes
experience2	-0.0010	_cons	1.8787***
Sex	0.2761	<i>Select function</i>	
child16	-0.0269	Regular high school	0.4660
old60	0.2322	Vocational high school	-0.2029
City dummy	Yes	Vocational higher education	0.2235
_cons	0.8992	College and above	0.0429
		experience	0.0528

lambda	0.6164***	experience2	-0.0012
Obs	2487	Sex	0.2614
		City dummy	Yes
		_cons	0.6364
		lambda	-0.4195***
		Obs	2487

Note: “***, **, *” significant at 1%, 5%, and 10%. Sample is women Aged 16 to 55 and Men aged 16 to 60.

Table 4-1 is separate regression results with self-employed, which are returns to education of Heckman method by adopting years of schooling and educated level for self-employed workers. If adopting years of schooling to measure returns of education, it can be seen from table 4-1 that returns to education in self-employed workers is 7.13%, which is lower than local labor and migrants. If adopting educational level to measure returns to education of self-employed workers, it is found from table 4-1 returns to education increases with level of education enhancement, except for returns to education of vocational high school. Regression coefficients of vocational higher education and college and above are 0.5010 and 0.8274, which are lower than local labor and migrants as well (table 3-2).

Table 4-2: Estimating Results of Formal vs. Informal Employment (I)

	Formal employment		Informal employment	
	Coef.	P> z	Coef.	P> z
<i>Wage function</i>				
Schooling	0.0928	0.000	0.0732	0.000
experience	0.0193	0.000	0.0319	0.000
experience2	-0.0004	0.000	-0.0009	0.000
Sex	0.2011	0.000	0.2110	0.000
City dummy		Yes		Yes
_cons	1.1973	0.000	1.1650	0.000
<i>Select function</i>				
Schooling	0.0325	0.072	0.0033	0.910
experience	-0.0001	0.995	-0.0431	0.148
experience2	0.0001	0.789	0.0011	0.155
Sex	0.0264	0.804	0.1616	0.423
child16	0.0898	0.289	-0.1055	0.433
old60	-0.1996	0.109	-0.1553	0.267
City dummy		Yes		Yes
_cons	2.1605	0.000	2.4841	0.000
lambda	-0.1189*	0.0837(Std.)	-0.2178*	0.1086(Std.)
Obs		5772		2744

Note: P value is significant probability. “***, **, *” significant at 1%, 5%, and 10%. Shanghai take as reference

city. Definition of formal and informal employment is just based on social insurance in this table.

Table 4-2 is regression results of Heckman method by adopting years of schooling to measure returns of education between formal employment and informal employment, which also exclude self-employed. It should be pointed out that definition of formal and informal employment is just based on social insurance in the table 1-4. Firstly, from table 4-2 we can see returns to education increase with improving years of schooling whether in the formal employment or informal employment. It also can be seen from table 4-2 that returns to education in the formal employment and informal employment exclude self-employed are 9.28% and 7.32% respectively. Secondly, it is evident that returns to education of formal employment are significant higher than that of informal employment.

Table 4-3: Estimating Results of Formal vs. Informal Employment (II)

	Formal employment		Informal employment	
	Coef.	P> z	Coef.	P> z
<i>Wage function</i>				
Regular high school	0.1496	0.000	0.1484	0.000
Vocational high school	0.2220	0.000	0.2566	0.000
Vocational higher education	0.4810	0.000	0.4625	0.000
College and above	0.7564	0.000	0.8492	0.000
experience	0.0248	0.000	0.0390	0.000
experience2	-0.0005	0.000	-0.0011	0.000
Sex	0.2014	0.000	0.2271	0.000
City dummy		Yes		Yes
_cons	2.0286	0.000	1.7260	0.000
<i>Select function</i>				
Regular high school	-0.0564	0.707	-0.0829	0.712
Vocational high school	0.1767	0.385	0.5605	0.030
Vocational higher education	0.0872	0.601	0.0794	0.743
College and above	0.3588	0.053	0.3627	0.397
experience	0.0060	0.727	-0.0352	0.215
experience2	0.0000	0.979	0.0010	0.189
Sex	0.0230	0.826	0.1820	0.356
child16	0.0839	0.317	-0.0975	0.451
old60	-0.1981	0.111	-0.1631	0.231
City dummy		Yes		Yes
_cons	2.4533	0.000	2.3689	0.000
lambda	-0.1136*	0.0794(Std.)	-0.2854**	0.1277(Std.)
Obs		5772		2744

Note: P value is significant probability. “***, **, *” significant at 1%, 5%, and 10%. Shanghai take as reference city. Sample is women Aged 16 to 55 and Men aged 16 to 60. Definition of formal and informal employment is

just based on social insurance in this table.

Table 4-3 is regression results of Heckman method by adopting educational level to measure returns of education both formal employment and informal employment exclude self-employed. It can be seen from table 4-3 that returns to educational increase with improving education level whether in the formal employment or informal employment. Besides, returns of education of vocational high school is higher than regular high school both formal employment and informal employment exclusive self-employed. As well as, it is evident that returns to education of formal employment and informal employment are significant higher than that of self-employed.

Table 4-4: Estimating Results of Formal vs. Informal Jobs (I)

	Formal jobs		Informal jobs	
	Coef.	P> z	Coef.	P> z
<i>Wage function</i>				
Schooling	0.0949	0.000	0.0666	0.000
experience	0.0194	0.000	0.0324	0.000
experience2	-0.0004	0.000	-0.0008	0.000
Sex	0.2031	0.000	0.1833	0.000
City dummy		Yes		Yes
_cons	1.1357	0.000	-0.0223	0.609
<i>Select function</i>				
Schooling	0.0027	0.879	0.0011	0.139
experience	-0.0001	0.883	0.0420	0.811
experience2	0.0159	0.879	-0.0218	0.896
Sex	0.0569	0.473	0.1234	0.458
child16	-0.2783	0.009	-0.0223	0.609
old60	0.0391	0.017	-0.0379	0.176
City dummy		Yes		Yes
_cons	2.2611	0.000	2.3499	0.000
lambda	0.0970*	0.0673(Std.)	0.1351*	0.0931(Std.)
Obs		5956		2528

Note: P value is significant probability. “***, **, *” significant at 1%, 5%, and 10%. Shanghai take as reference city. Sample is women Aged 16 to 55 and Men aged 16 to 60.

Table 4-4 is regression results of Heckman method by adopting years of schooling to measure returns of education between formal jobs and informal jobs, which also exclude self-employed. Firstly, from table 4-4 we can see returns to education increase with improving years of schooling both formal jobs and informal jobs. It can be seen from table 4-4 that returns to education in the

formal jobs and informal jobs exclusive self-employed are 9.49% and 6.66% respectively. It is evident that returns to education of formal jobs are significant higher than that of informal job. Moreover, returns to education of formal jobs are also higher than self-employed, but returns to education of informal jobs are lower than that of self-employed.

Table 4-5: Estimating Results of Formal vs. Informal Jobs (II)

	Formal jobs		Informal jobs	
	Coef.	P> z	Coef.	P> z
<i>Wage function</i>				
Regular high school	0.1939	0.000	0.0780	0.014
Vocational high school	0.2370	0.000	0.2684	0.000
Vocational higher education	0.5067	0.000	0.4357	0.000
College and above	0.7824	0.000	0.7480	0.000
experience	0.0234	0.000	0.0383	0.000
experience2	-0.0005	0.000	-0.0009	0.000
Sex	0.2035	0.000	0.1951	0.000
City dummy		Yes		Yes
_cons	1.9790	0.000	1.6566	0.000
<i>Select function</i>				
Regular high school	-0.0177	0.907	-0.0478	0.815
Vocational high school	0.2846	0.156	0.1493	0.650
Vocational higher education	0.1664	0.319	-0.0503	0.890
College and above	0.3828	0.022	-0.0190	0.964
experience	0.0077	0.655	-0.0354	0.164
experience2	-0.0001	0.718	0.0011	0.114
Sex	0.0147	0.887	0.0352	0.838
child16	0.0558	0.479	-0.0165	0.921
old60	-0.2766	0.009	0.1348	0.420
City dummy		Yes		Yes
_cons	2.5902	0.000	2.0662	0.000
lambda	-0.1037*	0.0664 (Std.)	0.1393*	0.0941(Std.)
Obs		5956		2565

Note: P value is significant probability. “***, **, *” significant at 1%, 5%, and 10%.Shanghai take as reference city. Sample is women Aged 16 to 55 and Men aged 16 to 60.

Table 4-5 is regression results of Heckman method by adopting educational level to measure return of education both formal jobs and informal jobs exclude self-employed. It can be seen from table 5-5 that returns to education increase with improving educational level both in the formal jobs and informal jobs. It is worth pointing out that regression result of levels of schooling in the formal jobs is visibly higher than that in the informal jobs at the same level of education according

to table 4-5. Besides, returns of education of vocational high school are higher than regular high school whether in the formal jobs and informal jobs exclusive self-employed. As well as, it is evident that regression coefficients of regular high school, vocational higher education and college and above in the formal jobs are significant higher than informal jobs, except for vocational high school, which is lower than informal jobs.

Table 4-6: Estimating Results of Formal vs. Informal Sectors (I)

	Formal Sectors		Informal Sectors	
	Coef.	P> z	Coef.	P> z
<i>Wage function</i>				
Schooling	0.1060	0.000	0.0669	0.000
experience	0.0272	0.000	0.0231	0.001
experience2	-0.0005	0.000	-0.0006	0.000
Sex	0.2052	0.000	0.2607	0.000
City dummy		Yes		Yes
_cons	0.9159	0.000	1.2481	0.000
<i>Select function</i>				
Schooling	0.0398	0.015	-0.0073	0.803
experience	-0.0056	0.716	0.0146	0.700
experience2	0.0003	0.384	-0.0005	0.508
Sex	0.0099	0.917	0.2535	0.123
child16	0.0129	0.864	-0.0421	0.671
old60	-0.1245	0.277	-0.1244	0.332
City dummy		Yes		Yes
_cons	1.9135	0.000	1.6402	0.011
lambda	0.2144*	0.1357(Std.)	0.4728***	0.1093(Std.)
Obs		7541		3409

Note: P value is significant probability. “***, **, *” significant at 1%, 5%, and 10%. Shanghai take as reference city. Sample is women Aged 16 to 55 and Men aged 16 to 60.

Table 4-6 is regression results of Heckman method by adopting years of schooling to measure returns of education between formal sector and informal sector. Firstly, from table 4-6 we can see returns to education increase with improving years of schooling both formal jobs and informal jobs. It can be seen from table 4-6 that returns to education in the formal sector and informal sector are 10.6% and 6.69% respectively. It is evident that returns to education of formal sector are significant higher than that of informal sector. Moreover, returns to education of formal jobs are also higher than self-employed, but returns to education of informal sector are lower than that of self-employed.

Table 4-7: Estimating Results of Formal vs. Informal Sectors (II)

	Formal Sectors		Informal Sectors	
	Coef.	P> z	Coef.	P> z
<i>Wage function</i>				
Regular high school	0.2213	0.000	0.2040	0.000
Vocational high school	0.3106	0.000	0.1924	0.001
Vocational higher education	0.5957	0.000	0.4284	0.000
College and above	0.9014	0.000	0.8021	0.000
experience	0.0319	0.000	0.0271	0.000
experience2	-0.0006	0.000	-0.0007	0.000
Sex	0.2070	0.000	0.2752	0.000
City dummy		Yes		Yes
_cons	1.8355	0.000	1.7681	0.000
<i>Select function</i>				
Regular high school	-0.0236	0.858	0.0628	0.711
Vocational high school	0.2700	0.126	-0.4853	0.061
Vocational higher education	0.2089	0.177	-0.2855	0.358
College and above	0.3832	0.012	0.0557	0.894
experience	-0.0005	0.972	-0.0008	0.980
experience2	0.0002	0.517	-0.0003	0.689
Sex	0.0076	0.935	0.2328	0.129
child16	0.0119	0.875	-0.0427	0.662
old60	-0.1205	0.293	-0.1207	0.320
City dummy		Yes		Yes
_cons	2.2404	0.000	1.8161	0.000
lambda	0.1120*	0.0536(Std.)	0.4829***	0.0971(Std.)
Obs		7541		3409

Note: P value is significant probability. “***, **, *” significant at 1%, 5%, and 10%. Shanghai take as reference city. Sample is women Aged 16 to 55 and Men aged 16 to 60.

Table 4-7 is regression results of Heckman method by adopting educational level to measure returns of education between formal sector and informal sector. It is worth pointing out that regression results of educational level in the formal sector is significantly higher than that in the informal sector at the same level of education according to table 4-7. Moreover, it is clearly seen that return to education of vocational high school is higher than return to education of regular high school in the formal sector, but which is lower than return to education of regular high school in the informal sector. Returns to education of vocational higher education is higher than return to education of vocational high school whether in the formal sector or in the informal sector. Regression coefficient of college and above in the formal sector and informal sector are 0.9014

and 0.8021 respectively, which are obviously higher than regression coefficients of vocational higher education.

3.2 Returns to education using switching regression and occupation dummy

The table 5-1 is main finding of returns of educational years. It shows the returns to education are much lower at 7 percent in the informal sector compared to 11.1 percent in the formal sector, while the average was 10.1 percent for all workers (see Table 2-1). After controlling for sex, city effect, returns to schooling goes down in both the sectors, in the formal sector from 10.5 to 11.1 percent and in informal sector from 6.7 to 4.2 percent respectively, while it was 10.8 for all workers (see Table 2-2). The addition of occupation dummy, professional reduces the returns to years of education to 10.5 percent as well as it lowers the returns of different levels of education in the formal sector and informal sector (Table 6 etc.).

Table 5-1: Returns to years of education in the Formal and Informal Sector: With and without occupation (Estimated from OLS and switching regression)

	<u>OLS</u>		<u>ML (switching regression)</u>	
	without occupation	with occupation	without occupation	with occupation
Formal sector				
Years of schooling	0.105***	0.099***	0.111***	0.105***
experience	0.027***	0.027***	0.030***	0.030***
Exp. squared	-0.001***	-0.001***	-0.001***	-0.001***
occupation dummy		0.146***		0.158***
Informal sector				
Years of schooling	0.067***	0.065***	0.042***	0.036***
experience	0.022***	0.018***	0.031***	0.036***
Exp. squared	-0.001***	-0.001***	-0.001***	-0.001***
occupation dummy		0.131***		0.135***
Sector choice equation, Pr(formal sector)				
Years of schooling			0.187***	0.196***
experience			0.042***	0.041***
Exp. squared			-0.001***	-0.001***
occupation dummy				0.247***
Observations			10843	8825

* Significant at 10%; ** significant at 5%; *** significant at 1%

Note: other explanatory variables Sex, cities are included in the regression.

Table 5-2 is returns to levels of education in the formal and informal Sector by estimated

switching regression. As we can see, in the formal sector inclusion of the occupation dummy lowers the returns of different levels of education in the formal sector (Table 6-1). Returns to Vocational higher education fall from 55.6% to 49.0% while the drop is even higher at g College and above level from 85.1 percent to 77.5%. In the informal sector introduction of the occupation dummy reduces the returns to Vocational higher education by about 13 percentage points and that of College by only 11 percentage points. Returns to Vocational higher education in formal sector is higher than in formal sector both with occupation dummy and without occupation dummy form table 10, 55.6% and 32.6% respectively. Similarly, Returns to College and above education level in formal sector is higher than in formal sector both with occupation dummy and without occupation dummy form table 9, 85.1% and 61.8% respectively.

Table 5-2: Returns to levels of education in the Formal and Informal Sector: With and without occupation (Estimated from switching regression)

Wage function	Estimated switching regression			
	Without occupation		With occupation	
	Formal Jobs	Informal Jobs	Formal Jobs	Informal Jobs
Reference category : Middle school and below				
Regular high school	0.203***	0.127***	0.200***	0.093*
Vocational high school	0.295***	0.108***	0.269***	0.078*
Vocational higher education	0.556***	0.326***	0.490***	0.191***
College and above	0.851***	0.618***	0.775***	0.504***
experience	0.036***	0.031***	0.035***	0.036***
experience2	-0.0007***	-0.0008***	-.0007***	-0.0009***
occupation (professional=1)			.173***	0.126***
Selection function				
Reference category : Middle school and below				
Regular high school			0.617***	
Vocational high school			0.983***	
Vocational higher education			1.435***	
College and above			1.880***	
experience			0.030***	
experience2			-0.0008***	
occupation (professional=1)			0.263***	
Identifying		0.506***	0.564***	
Observations		10843	8825	

* Significant at 10%; ** significant at 5%; *** significant at 1%

Note: other explanatory variables Sex, cities are included in the regression.

Table 5-3 is returns to levels of education in the formal and informal jobs by estimated

switching regression. As we can see, in the formal sector inclusion of the occupation dummy lowers the returns of different levels of education in the formal jobs (Table 6-3). Returns to Vocational higher education fall from 54.5% to 50.6% while the drop is even higher at g College and above level from 81.3 percent to 76.4%. In the informal jobs introduction of the occupation dummy reduces the returns to Vocational higher education by about 13 percentage points and that of College by only 5 percentage points. Returns to Vocational higher education in formal jobs is higher than in formal sector both with occupation dummy and without occupation dummy form table 12, 54.5% and 38.3% respectively. Similarly, Returns to College and above education level in formal jobs is higher than in formal jobs both with occupation dummy and without occupation dummy form table 9, 81.3% and 58.0% respectively.

Table 5-3: Returns to levels of education in the Formal and Informal Jobs: With and without occupation (Estimated from switching regression)

Wage function	Estimated switching regression			
	Without occupation		With occupation	
	Formal Jobs	Informal Jobs	Formal Jobs	Informal Jobs
Reference category : Middle school and below				
Regular high school	0.207***	0.142***	0.204***	0.098***
Vocational high school	0.278***	0.206***	0.261***	0.150***
Vocational higher education	0.545***	0.383***	0.506***	0.254***
College and above	0.813***	0.580***	0.764***	0.523***
experience	0.025***	0.030***	0.025***	0.032***
experience2	-0.0005***	-0.0008***	-0.0005***	-0.0008***
occupation (professional=1)			0.155***	0.195***
Observations	5978	5063	5978	5063
Selection function				
Reference category : Middle school and below				
Regular high school	0.534***		0.612***	
Vocational high school	0.858***		0.898***	
Vocational higher education	1.280***		1.297***	
College and above	1.803***		1.815***	
experience	0.031***		0.033***	
experience2	-0.0004***		-0.0004***	
occupation (professional=1)			0.069***	
Identifying	0.530***		0.589***	
Observations	10843		8825	

* Significant at 10%; ** significant at 5%; *** significant at 1%

Note: other explanatory variables Sex, cities are included in the regression.

3.3 Returns to education within the formal and informal sectors: quantile regression

We also estimate quantile regressions for both the formal and informal sector to understand returns to human capital investments across segments within the formal and informal sector. As known to all, Quantile regression helps to analyze returns to education across the earnings distribution. The returns to education along the distribution of earnings using quantile regressions are increasing and significant in both the sectors, except at the 50th quantile at the top of the distribution in the formal sector. We found using quantile regressions that the returns to education are significantly similar across the earnings distribution not only in the informal sector, but also in the informal sector. The inter quantile regression further shows that in the formal sector the difference in the returns are significant only in the secondary education segments of the distribution, while in the 90th quantile is not significantly different from zero (Table 6). However, the inter quantile regression shows that in the in formal sector the difference of each educational level in the returns are not significant.

Table 6: Returns to education by quantile regression estimates between Formal and Informal Sector

Variables	(1) q10	(2) q25	(3) q50	(4) q75	(5) q90
Formal sector					
Years of schooling	0.105***	0.107***	0.115***	0.106***	0.096***
Inter quantile		0.0018	0.0097**	0.0008	0.009
Regular high school	0.197***	0.170***	0.182***	0.176***	0.104
inter quantile		0.027	-0.014	-0.096	-0.427
Vocational high school	0.296***	0.269***	0.246***	0.209***	0.111**
inter quantile		0.027	0.050*	0.019	0.022
Vocational higher education	0.545***	0.516***	0.538***	0.497***	0.351***
inter quantile		0.035	-0.006	0.102***	0.105
College and above	0.826***	0.812***	0.812***	0.703***	0.583***
inter quantile		0.013	0.069*	0.017	0.041***
Observations	5930	5930	5930	5930	5930
Informal sector					
Years of schooling	0.039***	0.041***	0.042***	0.047***	0.048***
Inter quantile		0.0005	0.002	0.007	0.007
Regular high school	0.150***	0.156***	0.152***	0.126***	0.138***
inter quantile		0.152	0.179	0.186	0.036
Vocational high school	0.199***	0.241***	0.249***	0.194***	0.259***
inter quantile		0.032	-0.003	0.044	0.082
Vocational higher education	0.434***	0.442***	0.412***	0.450***	0.489***
inter quantile		0.015	0.014	0.028	0.031

College and above	0.716***	0.737***	0.774***	0.790***	0.973***
inter quantile		0.033	-0.108	-0.005	0.111
Observations	4913	4913	4913	4913	4913

* Significant at 10%; ** significant at 5%; *** significant at 1%

Note: other explanatory variables Sex, cities are included in the regression.

3.4 Influence of certain city variables on individual's returns to education

Table 7-1, table 7-2 and table 7-3 are OLS regression results that we also pay attention to the influence on individual's returns to education if certain city variables are controlled. Table 7-1 is regression results of individual's educational level and some city features estimated city fixed effects and the interaction terms, such as GDP per capita, GDP growth rate, government spending per capita, etc. we can see from table 7-1 that individual's returns to education has evident growth if certain city variables are controlled, Returns to education of regular high school is 1.0244 in the table 7-1, returns to education of vocational high school, vocational higher education and college and above are 1.1941, 1.8506 and 2.0618 respectively. It is clearly seen from table 7-1 that city share of informal workers, GDP growth rate and other city variable, etc have the significant influence on with individual returns of education.

Table 7-1: Estimating Results of Returns to Education controlled city variables by OLS(I)

	Coefficients	Std. Err
Regular high school(edug2)	1.0244***	0.3830
Vocational high school(edug3)	1.1941**	0.5259
Vocational higher education (edug4)	1.8506***	0.4252
College and above(edug5)	2.0618***	0.4690
experience	0.0336***	0.0021
experience2	-0.0008***	0.0001
Sex (male=1)	0.2356***	0.0115
City2(Wuhan)	-0.5059***	0.0309
City3(Shenyang)	-0.4156***	0.0318
City4(Fuzhou)	-0.2250***	0.0338
City5(Xian)	-0.6018***	0.0309
City6(Guangzhou)	0.1238***	0.0365
edug2*city_infshare	0.2064	0.4224
edug2*gdp_pc	0.0077	0.0110
edug2*gdp_rate	-5.6319**	2.6373
edug2*gov_payper	0.0000*	0.0000
edug3*city_infshare	-0.5204	0.5100
edug3*gdp_pc	-0.0280*	0.0144

edug3*gdp_rate	-3.6301	3.6650
edug3*gov_payper	0.0000	0.0000
edug4*city_infshare	1.6774***	0.4443
edug4*gdp_pc	-0.0237**	0.0117
edug4*gdp_rate	-9.9526***	2.9351
edug4*gov_payper	0.0000	0.0000
edug5*city_infshare	2.7497***	0.5011
edug5*gdp_pc	-0.0216*	0.0116
edug5*gdp_rate	-11.0549***	3.2382
edug5*gov_payper	0.0000	0.0000
_cons	1.7331***	0.0312
Obs		10844

Note: “***, **, *” significant at 1%, 5%, and 10%. Dependent variable is log-hourly wage. Sample is women Aged 16 to 55 and Men aged 16 to 60.

Table 7-2: Estimating Results of Returns to Education controlled city variables by OLS(II)

	Coefficients	Std. Err
Regular high school(edug2)	0.6911***	0.0604
Vocational high school(edug3)	0.6614***	0.0788
Vocational higher education (edug4)	0.9502***	0.0561
College and above(edug5)	1.1405***	0.0564
infemp* city_infshare*edug2	1.4190***	0.4847
infemp* city_infshare*edug3	0.0189	0.7111
infemp* city_infshare*edug4	1.3241**	0.6758
infemp* city_infshare*edug5	-1.7734**	0.8704
infemp* city_infshare	-2.3153***	0.1046
infemp*edug2	-0.2124**	0.0852
infemp*edug3	0.0466	0.1379
infemp*edug4	-0.2651**	0.1165
infemp*edug5	0.3134**	0.1474
city_infshare *edug2	-3.3100***	0.3291
city_infshare *edug3	-2.8091***	0.3865
city_infshare *edug4	-3.2404***	0.3013
city_infshare *edug5	-2.6361***	0.3146
_cons	2.1505***	0.0175
Obs		10845

Note: “***, **, *” significant at 1%, 5%, and 10%. Dependent variable is log-hourly wage. Sample is women Aged 16 to 55 and Men aged 16 to 60.

Table 7-2 is regression results of individual’s educational level and city share of informal workers. Returns to education of regular high school is 0.6911 in the table 7-2, returns to

education of vocational high school, vocational higher education and college and above are 0.6614, 0.9502 and 1.1405 respectively. Especially, the cross-term of schooling years and city share of informal workers has also the significant influence on with individual returns of education. Therefore, there is a relationship between the sizes of informal in a city and returns to education for individuals. Returns to education are to interact city share of informal workers with individual educational level.

Table 7-3: Estimating Results of Returns to Education controlled city variables by OLS(III)

	Coefficients	Std. Err
Regular high school(edug2)	1.0629***	0.3837
Vocational high school(edug3)	1.3777***	0.5278
Vocational higher education (edug4)	1.9526***	0.4262
College and above(edug5)	2.2769***	0.4686
experience	0.0327***	0.0021
experience2	-0.0007***	0.0001
Sex (male=1)	0.2362***	0.0115
city_infshare	-1.8649***	0.2403
gdp_pc	0.0965***	0.0074
gdp_rate	-9.0585***	1.4990
gov_payper	0.0000***	0.0000
edug2*city_infshare	0.1701	0.4229
edug2*gdp_pc	0.0093	0.0110
edug2* gdp_rate	-5.9255**	2.6394
edug2*gov_payper	0.0000**	0.0000
edug3*city_infshare	-0.5611	0.5118
edug3*gdp_pc	-0.0337**	0.0145
edug3*gdp_rate	-4.5184	3.6729
edug3*gov_payper	0.0000	0.0000
edug4*city_infshare	1.4991***	0.4461
edug4*gdp_pc	-0.0260**	0.0118
edug4*gdp_rate	-10.3115***	2.9384
edug4*gov_payper	0.0000	0.0000
edug5*city_infshare	2.5906***	0.5010
edug5*gdp_pc	-0.0299***	0.0115
edug5*gdp_rate	-11.9114***	3.2362
edug5*gov_payper	0.0000	0.0000
_cons	2.7176***	0.2255
Obs		10844

Note: “***, **, *” significant at 1%, 5%, and 10%. Dependent variable is log-hourly wage. Sample is women Aged 16 to 55 and Men aged 16 to 60.

Table 7-3 is regression results of individual's educational level and some city features keeping just the interaction terms, such as interaction terms of GDP per capita and educational level, GDP growth rate, government spending per capita and city informal share, etc. it is see from table 7-3 that individual's returns to education of regular high school is 1.0629 in the table 7-3, returns to education of vocational high school, vocational higher education and college and above are 1.377, 1.9526 and 2.2769 respectively, which are evidently higher than that of regression results in the table 7-1. Table 7-4 is regression results of individual's educational level and some city features keeping just the interaction terms.

Table7-4: Estimating Results of Returns to Education controlled city variables by OLS(IV)

	Formal workers		Informal workers	
	Coefficients	Std. Err	Coefficients	Std. Err
Regular high school(edug2)	-0.0008	0.5967	1.1091**	0.4983
Vocational high school(edug3)	0.4230	0.7393	1.1020	0.7732
Vocational higher education (edug4)	0.8345	0.5833	2.3311***	0.7118
College and above(edug5)	1.0440*	0.5968	3.5498***	1.0586
experience	0.0244***	0.0028	0.0313***	0.0031
experience2	-0.0006***	0.0001	-0.0008***	0.0001
Sex (male=1)	0.2315***	0.0156	0.2206***	0.0164
city_infshare	-1.8860***	0.5116	-1.3884***	0.2684
gdp_pc	0.0859***	0.0120	0.0808***	0.0088
gdp_rate	-14.3855***	2.7630	-6.6626***	1.7423
gov_payper	-0.0001***	0.0000	0.0000***	0.0000
edug2*city_infshare	1.0777	0.6971	-0.9918*	0.5361
edug2*gdp_pc	0.0166	0.0158	0.0037	0.0149
edug2*gdp_rate	-0.8997	4.1442	-5.3751	3.3602
edug2*gov_payper	0.0000	0.0000	0.0000	0.0000
edug3*city_infshare	0.1964	0.7339	-3.0648***	0.7722
edug3*gdp_pc	-0.0327*	0.0191	-0.0197	0.0244
edug3*gdp_rate	-0.0856	5.1012	-0.8326	5.4442
edug3*gov_payper	0.0000	0.0000	0.0000	0.0000
edug4*city_infshare	1.5315**	0.6652	0.2354	0.7744
edug4*gdp_pc	-0.0318**	0.0157	0.0015	0.0255
edug4*gdp_rate	-4.0255	3.9833	-12.932**	5.0844
edug4*gov_payper	0.0000	0.0000	0.0000*	0.0000
edug5*city_infshare	3.0158***	0.6869	-2.6080**	1.0441
edug5*gdp_pc	-0.0297**	0.0152	0.0557**	0.0328
edug5*gdp_rate	-6.1618	4.0973	-15.632**	7.4130
edug5*gov_payper	0.0000	0.0000	-0.0001***	0.0000
_cons	3.9296***	0.4089	2.3025***	0.2633
Obs	5930		4913	

Note: “***, **, *” significant at 1%, 5%, and 10%. Dependent variable is log-hourly wage. Sample is women Aged 16 to 55 and Men aged 16 to 60.

4. Summary Conclusions

This paper examines returns to education of different labor group by adopting Heckman, quantile regression, switching regression model and OLS method in all regressions based on CULS3 data (China Urban Labor Survey). Specification of regress include controls for work experience and other individual characteristics, such as gender, age, years of schooling or educational level and city dummy variable influence both wage function and select function. According to the empirical analysis, some important conclusions are summarized as the following:

Firstly, According to providing estimation in this paper, there is a general trend that returns to education increases with improving level of education or education in China urban by migration (by local labor-migrants), employment (by formal-informal), jobs (by formal-informal) and sector (by formal-informal). Specifically, returns to education of in the pooled sample, local labor and migrants exclude self-employed workers are respective 10.76%, 11.4% and 9.39%. It suggests that both local labor and migrants shows a significant growth in terms of returns to education. Moreover, migration contributes to improving returns to education for labor. While returns to education in self-employed workers are 7.13%, returns to education in the formal employment and informal employment exclude self-employed workers are 9.28% and 7.32%, respectively. Returns to education in the formal jobs and informal jobs exclusive self-employed workers are 9.49% and 6.66%, respectively. Returns to education in the formal sector and informal sector are 10.6% and 6.69%, respectively.

Secondly, returns to vocational education have obviously increased in different labor group, especially in the level vocational high school and vocational higher education. With the rise of educational level, returns to education in whether both local workers and migrant workers or both formal employment and informal employment tend to increase. In addition, returns to vocational high school are higher than that of regular high school. But at similar educational level, returns to educations of informal sector are still below than formal sector.

Thirdly, it is clearly from quantile regression results that the returns to education are significantly similar across the earnings distribution not only in the formal sector, but also in the informal sector. Inter quantile regression shows that in the in formal sector the difference of each educational level in the returns are not significant. Returns to education present considerable gap between formal sector and informal sector if using to switching regression model, compared to no

control whether or not any other family member is a formal worker. Individual's returns to education have evident growth if certain city variables are controlled, compared to pure Mincer equation. Individual's returns to education have evident growth if certain city variables are controlled, compared to pure Mincer equation.

Finally, Earnings of informal sector in urban China share similar features with some other developing countries, for example, Monthly wages reflect earning opportunities in the informal sector better than hourly wages. Not all employment of informal sector or informal employment is inferior to employment of formal sector with regard to wage. Despite the evident difference in mean earning, densities of informal and formal sector log-earning overlap to large extent. Wage dispersion is greater in the informal sector compared to formal sector.

References

- Kazuhiro Yuki, 2007, Urbanization, informal sector, and development, *Journal of Development Economics*, 84: 76-103.
- Isabel Günther and Andrey Launov, 2011, Informal employment in developing countries: Opportunity or last resort? *Journal of Development Economics*, 2011.01.05
- Jamal, V., Weeks, J., 1988. The vanishing rural–urban gap in sub-Saharan Africa, *International Labor Review*, 127 (3), 271-292.
- Ranis, G., Stewart, F., 1999. V-goods and the role of the urban informal sector in development. *Economic Development and Cultural Change*, 47 (2), 259-288.
- Marcouiller, D., de Castilla, V.R., Woodruff, C., 1997. Formal measures of the informal-sector wage gap in Mexico, El Salvador, and Peru. *Economic Development and Cultural Change* 45 (2), 367-392.
- Maloney, W.F., 2004. Informality revisited. *World Development* 32 (7), 1159–1178.
- Fields, G.S., 2005. A guide to multisector labor market models. Paper prepared for the World Bank Labor Market Conference, http://www.ilr.cornell.edu/directory/gsf2/Multisector_mar05.pdf.
- Loayza, N., Rigolini, J., 2006. Informality trends and cycles. World Bank Policy Research Working Paper No.4078.
- Amaral, P., Quintin, E., 2006. A competitive model of the informal sector. *Journal of Monetary Economics*, 53 (7), 1541–1553.
- Bennett, J., Estrin, S., 2007. Informality as a stepping stone: entrepreneurial entry in a developing economy. IZA Discussion Paper No.2950.
- Bennett, J., 2008. Formality, informality, and social welfare. IZA Discussion Paper No.3550.

- World Bank, 2007. *Informality: Exit and Exclusion*. World Bank, Washington D.C.
- Cai, Fang, Yang Du and Meiyang Wang (2009), "Migration and Labor Mobility in China", Report for UDNP.
- Cai, Fang, Yang Du and Meiyang Wang (2009), "Employment and Inequality Outcomes in China", Report for OECD.
- Cai, Fang, Yang Du and Meiyang Wang (2006), "Employment in China's Fast Growing Region: What Can We Learn from It to Make Employment Expansion Keeping Up with Economic Growth?" Report for the World Bank.
- Albert Park and Fang Cai, 2007, *the Informalization of the Chinese Labor Market*, unpublished paper.
- Cai, Fang, and Yaowu Wu. 2006. *China's urban informal employment: scale and characteristics*, mimeo.