EDUCATION RETURN IN URBAN CHINA:

EVIDENCES FROM CHNS DATASET

Lili Kang*

Department of Management, Birmingham Business School

Fei Peng

CREW, Birmingham Business School

*Correspondance author

Email: lxk699@bham.ac.uk

Preliminary, not for quotation

EDUCATION RETURN IN URBAN CHINA:

EVIDENCES FROM CHNS DATASET

ABSTRACT

This essay investigate the returns to schooling in urban China using pooled CHNS (China Health and Nutrition Survey) dataset of the 1990s (including 1991, 1993 and 1997) and the 2000s (including 2000, 2004 and 2006). Based on the standard Mincerian human capital earning function, we adopt the OLS and IV models to see the difference of controlling endogeneity of schooling or not.

Education is measured in two ways: years of schooling (input) and highest educational qualification (output, 5 categories of educational levels). We find that OLS estimate of the returns to schooling in the 2000s is higher than the 1990s. The returns to "college and above" and "professional school" increase across time, while the returns to "upper and lower middle school" decrease from 1990s to 2000s. In addition, if we do not control for endogeneity bias of schooling, the OLS estimates of the returns to schooling will be underestimated, especially in the 2000s.

Introduction

For at least a decade following the economic reforms initiated in the early 1980s, scholars have found unusually low returns to education in the urban China. Returns to education provide the information about the incentives for human capital accumulation, the efficiency of resource allocation, and the distributional consequences of differences in human capital. With the CHNS (China Health and Nutrition Survey) data, our paper adopt the ordinary least square (OLS) and instrumental variable (IV) methodology to presents a significant change in education returns over the 1990s and 2000s in urban China.

In this paper, we answer three important questions: Do the returns to education in urban China differ between the 1990s and the 2000s? Do the OLS estimates differ if considering the interactive effects of education and ownership? Do the OLS estimates, i.e., estimates which ignore the endogeneity bias, differ significantly from the estimates that consider this bias?

The rest of this paper is organised as follows. In the next section, we briefly discuss the evolution of China's labour market institutions and ownership structure which affect wage determination. In Section III, we provide a literature review on the returns to education in China. In section IV, we present the OLS and IV empirical specifications. Data descriptive and estimates of the returns to schooling are presented in Section V. Section VI concludes.

Evolution of labour market

From the foundation of PRC to the late 1970s, state-owned sectors dominate in the urban China, and the Bureau of Labour and Personnel centrally determined and controlled the wages of all workers through a grade system to reduce labour costs for rapid industrialization. Low wages were made possible by state-subsidized food prices and state provision of non-wage benefits to workers and their families, such as housing, child care, medical insurance, and pensions. At the same time, the government effectively eliminated most of the direct private costs of education by waiving all tuitions and fees for college students and by providing living stipends to students from poor families.

This heavy planning led to poor effect incentive which depressed productivity and innovation. In the early 1980s, Deng Xiaoping reformed the economy beginning with the rural "Household Responsibility System". However, the welfare guarantees to urban workers

hindered the urban reforms until the middle and late 1990s. In October 1984, the Communist Party passed the "Resolution on Economic Institutional Reform," which changed the fixed wage quota system to a floating total wage system, by allowing profitable firms to pay higher salaries and bonus to more productivity workers (Dai, 1994).

In 1986, the State Council issued "Temporary Regulations on the Use of Labour Contracts in State- Run Enterprises," formally introduced labour contracts to end the system of permanent employment (Meng, 2000). Zhang et.al. (2005) show that by 1997, one hundred million employees had signed labour contracts with their employers. Most workers who quit state-owned enterprises voluntarily moved to the non-state sector. Since the early 1990s, non-state enterprises, including foreign, private, and mixed ownership enterprises, have emerged as prominent players in the labour market. By competing aggressively with the public sector, these firms provided an impetus for state-sector restructuring.

In addition, the institutional barriers to off-farm labour participation have been attenuated. The SSBa (1990–2000) have documented a series of policies that loosened restrictions on labour mobility out of agriculture. During 1986 to 1995, the percentage of rural labour force employed in township and village enterprises (TVEs) increased from 12.8% to 22.2% (SSBb, 1996).Maurer-Fazio (1999) show the rising significance of education as a determinant of off-farm earnings, a result that implies individuals are being rewarded more for their human capital, which is a sign of well-functioning markets.

Literature review

Many researchers used OLS methodology to examine the rates of returns to education in China. However, plenty of literatures (such as, Heckman and Li, 2004; Li and Luo, 2004) pointed out the endogenous bias of education and used the IV methodology to cope with this problem. Normally, the IV estimators are higher than the OLS estimators of the conventional Mincerian model.

Chen and Hamori (2009) examines economic returns to schooling in urban China using OLS and IV methodologies for women and men with the CHNS 2004 and 2006 pooled data, and their instruments for schooling is spouse education. Heckman and Li (2004) use the 2000 data from the China Urban Household Income and Expenditure Survey (CUHIES) to identify

the returns to higher education for young people in the urban areas of the six provinces. The instruments employed are parental education and year of birth, and the IV estimator of average return to four-year college attendance was 43% (on average, 11% annually).

Li and Luo (2004) use the CHIP 1995 data and apply various IV estimations to estimate returns to schooling for young workers in urban China. Their result is robust using either parental education or sibling variables as instruments. Yang (2005) uses CHIP 1988 and 1995 data to study the changes over time in returns to education for a large number of Chinese cities. On average, he reports that the estimated rates of return at the city level increased from 3.1% in 1988 to 5.1% in 1995.

Fleisher, Li, Li and Wang (2005) adopt three methods, i.e., OLS, IV, and semi-parametric (SPIV) to estimate how selection and sorting influenced the evolution of the private returns to schooling for college graduates with CHIP 1988, 1995 and 2002 dataset. All three methods show a substantial increase in returns to schooling between 1995 and 2002. The IV and SPIV estimates of the returns to college education also turn out to be sensitive to the use of a proxy for ability. The IV instruments are parental schooling and parental income.

Empirical Specifications

The Mincerian human capital earning function is employed to estimate the returns to education. The dependent variable is the natural logarithm of hourly wage. Two measures of education are used in this paper. One is the average years of schooling. Another measure is the educational levels, which we classify into 5 categories: college and above; professional, technical or vocational school; upper middle school; lower middle school; primary school and illiteracy or semi-illiteracy.

The square of experience is consistent with a declining return to experience and with the shape of age-earnings profiles observed in many dataset. In addition to human capital, wages are also affected by demographic factors, such as gender, marital status, and the market conditions. For example, ownership dummies are added to explore the rent of State Owned Enterprise (SOE) or public sector, year dummies to reflect the deepening of economic reform, as well as provincial dummy variables may reflect the provincial inequality in China (Fleisher et al 2009).

$$\ln W_i = \beta_0 + \beta_1 E_i + \beta_2 X_i + \varepsilon_i$$

where W_i is hourly wage rate of individual i, E_i represent education (years of schooling, or educational levels), X_i is a vector of control variables including gender dummy (male as 1), labour market experience (age – years of schooling – 6) and its square, marital status dummy (ever married as 1), ownership dummies, year dummies and provincial dummies; ε is an error term $\varepsilon \sim N(0, \sigma^2)$.

Plenty of literature mentioned the biases that may be caused principally by the endogenous nature of schooling. (see Dearden 1999a, b, and surveys of this literature in Card 1999 and Blundell, Dearden and Sianesi 2005). In our CHNS sample, endogeneity can arise from measurement error in schooling, since the schooling information is provided in levels rather than in years. Second, endogeneity can arise because of omitted ability. That is, the return coefficient β_1 is biased (upwards) because chosen schooling levels are (positively) correlated with omitted ability, while ability is (positively) correlated with the wage rate. Third, urban populations in China place more importance on academic background, and the academic background has a strong influence on individual employment, wages, and promotion. Adult education is therefore popular and positively correlated with the wage rate (Chen and Hamori, 2009). Moreover, Card (1999) argue that OLS estimates of β_1 are biased downwards because individuals with high discount rates choose low levels of schooling, that is, schooling with higher marginal rates of return.

Mcintosh (2006) argues that the most common methodology adopted to correct for such biases has been an instrumental variables approach, isolating exogenous variation in education received. For example, Harmon and Walker (1999) discuss a number of potential variables with which to instrument education choices in the UK. We adopt an IV approach and using the "piped / tap water in house or courtyard" and "estimated market value of house / apartment"[1] variables as instruments. Piped water represents the public investment and the market value of house represents the household wealth and environment for education.

The following two-equation model describes the natural logarithm of hourly wage and years of schooling are normally applied to cope with the endogeneity of schooling:

$$\ln W_i = \alpha X_i' + \beta Y S_i + \mu_i$$

---

[1] Details of tap water and electricity for lighting please see appendix B.

$$YS_i = \delta Z_i' + v_i$$

Where Z_i denotes the vector of observed instrumental variables with the properties suggested above and the other exogenous variables, same as the variables in the above OLS regression. Our data contain two potential instruments for years of schooling YS_i: tapped water and house value.

Data description and empirical results

The dataset used in this paper is the China Health and Nutrition Survey (CHNS). It is conducted by the China's National Institute of Nutrition and Food Safety, the Chinese Centre for Disease Control and Prevention, and the University of North Carolina at Chapel Hill. The survey employs a multistage random-cluster sampling process to draw households from nine administrative divisions (Guangxi, Guizhou, Heilongjiang, Henan, Hubei, Hunan, Jiangsu, Liaoning, and Shandong)[2]. Most households have been followed up across all five waves, but some deletions and additions have occurred based on community participation. In 1991, the CHNS surveyed 15,917 individuals from 3,795 households. Sample sizes are relatively similar across waves. Beginning in 1993 and continuing through subsequent waves, the survey added new households in the sample areas that were formed by individuals included in the previous waves. From 1997 onward, the survey added new households and communities to replace those that were no longer participating.

We use six waves of CHNS data (1991, 1993, 1997, 2000, 2004 and 2006) to compare returns of education in 1990s and the new century. The six sample years represent distinct phases of economic reform in China. Specifically, the year 1991 represents the early stage of urban reform that started in 1982, the year 1993 represents the middle stage of urban economic transitions after the reform re-started in 1992, and the year 1997 and 2000 are affected by the 1997-98 Asian financial crises, and by 2004 and 2006 economic transition had entered a mature stage.

The sample used in this study is selected based on individuals aged 18-65 who work in sectors besides "Agriculture, forestry, animal husbandry, fishery and water conservancy" sector. It only includes workers with positive annual gross income. Owners of private or

---

2 For details of surveyed administrative divisions, please see appendix A.

individual enterprises for the primary and secondary job have been excluded, because it is difficult to separate their wages from the profit income. Observations with missing values on education, experience, etc. have been dropped.

Nominal annual earnings include regular wages, bonuses, all kinds of subsidies and in-kind wages from the work unit. Nominal annual earnings are converted into real annual earnings by deflating by provincial urban CPI (Consumer Price Index) with year 1995 as 100. As presented in Figure 1, annual earnings increase from 2,668.16 Yuan in 1991, to 4,463.60 Yuan in 1997, triple to 10,180.81 Yuan in 2004, then jump to 14,977.48 Yuan in 2006.

 DeBrauw and Rozelle (2004) mentioned that previous studies may have mis-measured wages by using a wage measure that endogenizes part of the individual's decision regarding the amount of labour to allocate to off-farm work in rural China. If so, wages for the educated could be systematically understated relative to the less educated. Using data that allow for a more appropriate measure of the wage—the hourly wage instead of a daily or monthly wage—might help provide better estimates. So, we use the calculated real average hourly wage rate with the real annual earning and annual total working hours[3]. With the long working hours, the average hourly wage rate is very low in China, especially in 1991 and 1993, only around 2 Yuan per hour. It jumps to above 6.01 Yuan per hour in 1997, and remains steady in 2000 and 2004, then jumps again in 2006 to 9.28 Yuan per hour. We use the estimated market value of house or apartment to reflect the household wealth. In 1991, the average house value is 33 thousands Yuan, increase to 55 thousands Yuan in 1993, then stay around 75 thousands Yuan in 1997 and 2000, followed by two dramatic increase to 166 thousands Yuan in 2004 and 229 thousands Yuan in 2006, which is consistent with the rapid housing price increase in urban China.

---

3 For details of annual earning and hourly wage rate, please see appendix B.

Figure 1 Trend of main variables over 1991-2006

Estimated House maket value
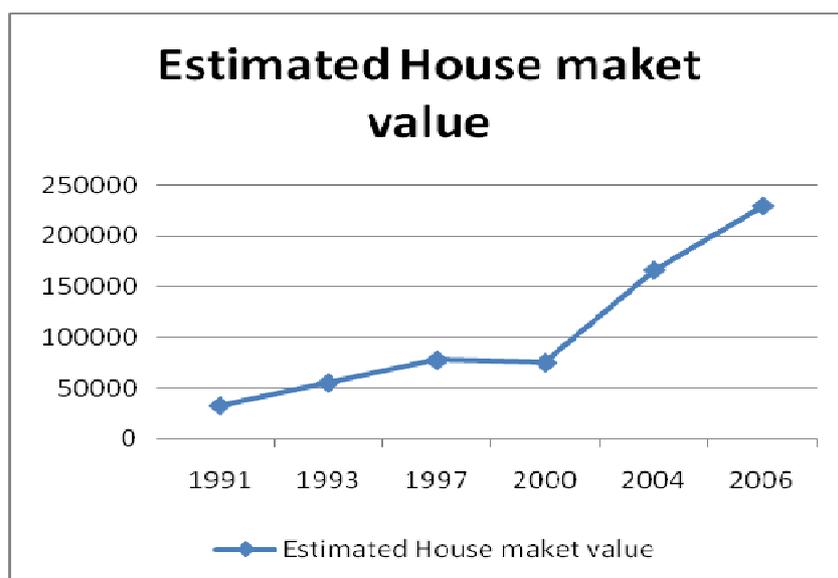
Note: Annual earning, hourly wage rate and the house market value are real values, adjusted by the provincial urban CPI (consumer price index).

Table 1 shows the data description of main variables, comparing the 1990s and the 2000s. The average years of schooling increases from 8.35 years to 9.6 years. Among the 5 categories of education levels, the "lower middle school" level dominates the whole sample period, above 30 percent, followed by the "upper middle school" level, about 20 percent. Persons with higher education increase from 10 percent to 20 percent, and persons with vocational degrees increase from 9 percent to 13 percent. The "Primary School and below" is the baseline educational level, decrease from 23 percent to 13 percent.

The average estimated working experience is higher in the 2000 at 24.56 years, compared to 22.31 years in the 1990. About 52-54 percent people are once married in the surveys. The ownership variable as three categories: SOE (state-owned enterprises) and public sector, collective enterprises and private or foreign enterprises[4]. The SOE and public sector dominate across the two periods, around 51-52 percent. The percent of the private or foreign enterprises increase from 25 percent to 32 percent, while the baseline – the number of collective enterprises decreased from 23 percent to 16 percent. Nearly 89-90 percent of households have piped/tap water in house or courtyard across the surveyed period.

---

[4] Details of enterprise categories please see appendix B.

Table 1 Data description

Description	1990s			2000s		
	Obs.	Mean	Std. Dev.	Obs.	Mean	Std. Dev.
Hourly Wage Rate (Yuan)	4225	3.16	38.03	4008	7.63	30.47
Years of Schooling (Year)	6512	8.35	4.03	5011	9.60	3.89
College and Above (%)	5500	0.10	0.30	4469	0.20	0.40
Professional and Technical School (%)	5500	0.09	0.29	4469	0.13	0.34
Upper middle School (%)	5500	0.21	0.40	4469	0.19	0.39
Lower middle School (%)	5500	0.38	0.49	4469	0.34	0.48
Primary School and Below (%)	5500	0.23	0.42	4469	0.13	0.34
Estimated Experience(Year)	6512	22.31	12.98	5011	24.56	12.30
Gender (1=Male)	6537	0.52	0.50	5045	0.54	0.50
Marital Status (1=Ever Married)	6493	0.82	0.38	4996	0.87	0.34
SOE or Public Sector (%)	5368	0.51	0.50	3636	0.52	0.50
Collective Enterprises (%)	5368	0.23	0.42	3636	0.16	0.36
Private or Foreign Enterprises (%)	5368	0.25	0.43	3636	0.32	0.47
Water(1=having tap water)	6484	0.89	0.32	5037	0.90	0.30
Estimated House Market Value (Yuan)	2250	53,222	83,941	1246	117,187	162,178

Note:
1. 1990s include the pooled 1991, 1993 and 1997 data; 2000s include the pooled 2000, 2004 and 2006 data.
2. Hourly wage and estimated house market value are real value, adjusted by the provincial urban CPI (consumer price index).
3. Provincial dummy variables are not reported.

Table 2 show the baseline OLS estimators for 1990 and 2000s time periods, with the logarithm of hourly wage rate as the dependent variable. The returns to one more year of schooling increase from 5.5 percent in 1990s to 8.5 percent in 2000s. According to the educational levels, we find that all educational levels have significantly higher returns than the baseline "primary school and below" level, and higher level, higher return. From 1990s to 2000s, the returns to "college and above" and "professional school" increase 13 percent and 2 percent, respectively. On the contrary, the returns to "upper middle school" and "lower middle school" decrease 4 – 5 percent.

Comparing the 1990s and 2000s, the contribution of one more year of experience decreases from about 3 percent to 1.6 percent. Male workers always earn more than female workers, especially in the 2000s, nearly 20 percent higher. In the 1990s, workers in the private or foreign enterprises earn the most, about 36 percent higher than collective enterprises, decrease to 27 percent higher in the 2000s; while in the 2000s, workers in the SOE or public sectors earn the highest wage rate, about 30 percent higher than the collective enterprises.

Table 2 Baseline OLS estimates

Dependent: Ln (hourly wage)	1990s	2000s	1990s	2000s
Years of schooling	0.055***	0.085***		
	0.003	0.005		
College and above			0.649***	0.771***
			0.042	0.068
Professional school			0.448***	0.469***
			0.04	0.068
Upper middle			0.310***	0.272***
			0.036	0.067
Lower middle			0.181***	0.129**
			0.033	0.064
Experience	0.027***	0.016***	0.032***	0.016***
	0.004	0.005	0.004	0.006
Experience square	-0.000***	0	-0.000***	0
	0	0	0	0
Male	0.112***	0.192***	0.121***	0.201***
	0.019	0.027	0.02	0.027
Ever Married	0.046	-0.057	0.027	-0.017
	0.034	0.051	0.036	0.052
SOE or Public Sector	0.068***	0.322***	0.04	0.298***
	0.024	0.038	0.025	0.039
Private or Foreign Enterprises	0.361***	0.274***	0.359***	0.270***
	0.053	0.043	0.056	0.043
R-squared	0.168	0.205	0.181	0.219
N	4006	2915	3687	2817

Note:
1. 1990s include the pooled 1991, 1993 and 1997 data; 2000s include the pooled 2000, 2004 and 2006 data.
2. Hourly wage is real value, adjusted by the provincial urban CPI.
3. The italic values are heteroskedastciticy- robust standard error.
4. The significant levels are * for 10%; ** for 5% and *** for 1%.
5. Provincial dummy variables and year dummies are not reported.

When we consider the interactive effect of years of schooling and ownership, the table 3 tells us that in the 1990s, one year of schooling can increase 8.5 percent hourly wage rate in the private or foreign enterprises, 5.5 percent in the SOE or public sector, and 5 percent in the collective enterprises. In the 2000s, in the SOE or public sector, workers can benefit 9 percent higher wage rate from one years of schooling, remain 8.5 in private firms and 5.8 percent in the collective enterprises. In sum, higher educated workers prefer private or foreign enterprises in 1990s, and SOE or public sector in 2000s. The effects of experience and gender are similar as in the baseline OLS regressions.

Table 3 OLS estimates with interactive dummies

Dependent variable: Ln (hourly wage)	1990s	2000s
Schooling	0.050***	0.058***
	0.004	*0.006*
Schooling*SOE or public sector	0.005*	0.032***
	0.002	*0.004*
Schooling*Private or Foreign Enterprises	0.035***	0.027***
	0.006	*0.004*
Experience	0.028***	0.017***
	0.004	*0.005*
Experience square	-0.000***	0
	0	*0*
Male	0.113***	0.191***
	0.019	*0.027*
Ever Married	0.045	-0.061
	0.034	*0.051*
R-squared	0.166	0.207
N	4006	2915

Note:
1. 1990s include the pooled 1991, 1993 and 1997 data; 2000s include the pooled 2000, 2004 and 2006 data.
2. Hourly wage is real value, adjusted by the provincial urban CPI.
3. The italic values are heteroskedastciticy- robust standard error.
4. The significant levels are * for 10%; ** for 5% and *** for 1%.
5. Provincial dummy variables and year dummies are not reported.

Next, we analyze the returns to schooling with the instrumental variables estimators. Greene (2008) mentions that, many standard estimators including OLS (Ordinary least squares) and IV (instrumental-variable) could be thought as special cases of GMM (Generalized method of moments) estimators. The latter estimator has a clear advantage over IV estimator. If heteroscedasticity is present in the model GMM is more efficient whereas heteroscedasticity is not present, the GMM is no worse asymptotically than IV (Baum et al. 2003). In this paper we use cross-sectional data, and for that reason we could expect that model error is heteroscedastic. To alleviate possible heteroscedasticity problem apart from 2SLS (two-stage least squares) IV estimates, we employ GMM based Instrumental Variable estimates.

As shown in table 4, the GMM rate of returns to schooling is 25 and 38 percent in 1990s and 2000s respectively, higher than the OLS estimators (6 percent and 9 percent). The result of the endogeneity test rejects the null hypothesis that the OLS estimates are consistent, with GMM C statistic Chi2 are 29.76 and 32.37. The first-stage F-statistics (29.99 and 20.36) confirm the joint significance of the two instrumentals to explain endogenous "years of

schooling" volatility. Furthermore, the Hansen's J Chi2 statistics (22.03 and 11) reject the two instruments are over-identified.

Different from the OLS estimators, one year of experience has higher impact in the 2000s (8.6 percent) than in the 1990s (3.7 percent). The gender gap disappears and ever married workers are in inferior situation compared to single person, showing about 6.5 percent lower wage rate. Across the surveyed period, workers in the private or foreign enterprises earn the highest wage rate, especially in the 1990s, at 61 percent higher than workers in the collective enterprises, decrease to 24.6 percent in the 2000s. SOE or public sectors are not attractive for workers.

Table 4 Instrumental GMM models

Dependent variable: Ln (hourly wage)	1990s	2000s
Years of schooling	**0.258***	**0.386***
	0.044	0.062
Experience	**0.037***	**0.086***
	0.01	0.02
Experience square	**0**	**-0.001****
	0	0
Male	**-0.001**	**0.125**
	0.055	0.079
Ever Married	**-0.12**	**-0.649***
	0.088	0.184
SOE or Public Sector	**-0.248***	**-0.242**
	0.093	0.176
Private or Foreign Enterprises	**0.611***	**0.246***
	0.118	0.134
Exodgeneity test		
GMM C statistic Chi2	29.7574	32.3665
p-value	0.0000	0.0000
Instruments validity test		
First stage F-statistic	29.99	20.36
p-value	0.0000	0.0000
Over-identification test		
Hansen's J chi2	22.0289	10.9963
p-value	0.0000	0.0009
N	**1315**	**792**

Note:
1. 1990s include the pooled 1991, 1993 and 1997 data; 2000s include the pooled 2000, 2004 and 2006 data.
2. Hourly wage is real value, adjusted by the provincial urban CPI.
3. The italic values are heteroskedastciticy- robust standard error.
4. The significant levels are * for 10%; ** for 5% and *** for 1%.
5. Provincial dummy variables and year dummies are not reported.

Conclusions

We examined the returns to schooling in urban China using pooled CHNS dataset of the 1990s (including 1991, 1993 and 1997) and the 2000s (including 2000, 2004 and 2006). In this paper, we answer three important questions:

(1) Do the returns to schooling in urban China differ between the 1990s and the 2000s?
(2) Do the OLS estimates differ if considering the interactive effects of education and ownership?
(3) Do the OLS estimates, i.e., estimates which ignore the endogeneity bias, differ significantly from the estimates that consider this bias?

First, we find that OLS estimate of the returns to schooling in the 2000s is higher than the 1990s. The returns to "college and above" and "professional school" increase across time, while the returns to "upper and lower middle school" decrease from 1990s to 2000s. Second, whether considering the schooling interactive dummy with ownership or not, the OLS estimates are similar. Finally, we find that OLS without control for endogeneity bias may underestimate the true rates of returns of schooling, especially in the 2000s.

Reference

Appleton S, Song L, Xia Q (2005). 'Has China crossed the river? The evolution of wage structure in urban China during reform and retrenchment'. Journal of Comparative Economics, 33(4): 644–663

Baum K., Schaffer M., Stillman S., (2003). 'Instrumental variables and GMM: Estimation and testing', STATA Journal.

Bishop J, Luo F J, Wang F (2005). 'Economic Transition, Gender Bias, and Distribution of Earnings in China'. Economics of Transition, (13): 239–259

Blundell, R., Dearden, L. and Sianesi, B. (2005). 'Evaluating the impact of education on earnings in the UK: models, methods and results from the NCDS', Journal of the Royal Statistical Society Series A, Vol. 168, pp. 473–512.248 Bulletin _ Blackwell Publishing Ltd 2006

Byron, Raymond P., Manaloto, Evelyn Q., (1990). 'Returns to education in China'. Economic Development and Cultural Change 38, 783–796.

Card, D. (1999). 'The causal effect of education on earnings', in Ashenfelter O. and Card D. (eds), Handbook of Labor Economics, Vol. 3A, Elsevier Science North-Holland, Amsterdam, New York and Oxford, pp. 1801–1863.

Chen, Baizhu, Feng, Yi, (2000). 'Determinants of economic growth in China: Private enterprises, education, and openness'. China Economic Review 11 (1), 1–15.

Chen, Guifu, Shigeyuki Hamori, (2009). 'Economic returns to schooling in urban China: OLS and the instrumental variable approach'. China Economic Review 20, 143-152

Dai, Yuanchen, (1994). 'Zhongguo lao dong li shi chang pei yu yu gong zi gai ge'. Beijing: Zhongguo lao dong chu ban she. (Chinese Labour Market Training and Wage Reform. Beijing: Chinese Labor Press.)

Dearden, L. (1999a). 'The effects of families and ability on men's education and earnings in Britain', Labour Economics, Vol. 6, pp. 551–567.

(1999b). 'Qualifications and Earnings in Britain: How Reliable are Conventional OLS Estimates of the Returns to Education', Working Paper No. 99/7, Institute for Fiscal Studies.

de Brauw A., Huang J, Rozelle S, Zhang L, Zhang Y (2002). 'The evolution of China's rural labour markets during the reforms'. Journal of Comparative Economics, 30(2): 329–353

de Brauw, Alan D., Rozelle, Scott, (2008). 'Reconciling the returns to education in off-farm wage employment in rural China'. Review of Development Economics, 12(1): 57-71

Demurger, Sylvie, (2001). 'Infrastructure development and economic growth: An explanation for regional disparities in China'. Journal of Comparative Economics 29, 95–117.

Fleisher, Belton M., Chen, Jian, (1997). 'The coast–noncoast income gap, productivity, and regional economic policy in China'. Journal of Comparative Economics 25, 220–236.

Fleisher, Belton M., Wang, Xiaojun, (2004). 'Skill differentials, return to schooling, and market segmentation in a transition economy: The case of mainland China'. Journal of Development Economics 73 (1), 315–328.

Fleisher, Belton M., Wang, Xiaojun (2005). 'Returns to schooling in China under planning and reform'. Journal of Comparative Economics, 33(2): 265-277

Greene, William H., (2008). 'Econometric Analysis', 6th ed., Pearson/Prentice Hall,

Heckman, James J., (2003). 'China's investment in human capital'. Economic Development and Cultural Change 51(4), 795–804.

Heckman, James, (2005) . 'China's Human Capital Investment', China Economic Review, 15 50–70.

Heckman, J. J., & Li, X. (2004). Selection bias, comparative advantage and heterogeneous returns to education: evidence from China in 2000. Pacific Economic Review, 9(3), 155−171.

Knight J, Song L N (2003). 'Increasing urban wage inequality in China: Extent, elements and evaluation'. Economics of Transition, 11(4): 597–619

Li, Haizheng, (2009), 'Higher Education in China', NBER working paper, http://www.nber.org/~confer/2008/augm08/li.pdf

Li, H., & Luo, Y. (2004). Reporting errors, ability heterogeneity, and returns to schooling in China. Pacific Economic Review, 9(3), 191−207.

Maurer-Fazio, Margaret, (1999). 'Earnings and Education in China's Transition to a Market Economy—Survey Evidence from 1989 and 1992'. China Economic Review. 10, 1:7–40, Spring

McIntosh, S. (2006), 'Further Analysis of the Returns to Academic and Vocational Qualifications', Oxford Bulletin of Economics and Statistics, 68(2): 225-51.

Meng, Xin, Kidd, Michael P., (1997). 'Labour market reform and the changing structure of wage determination in China's state sector during the 1980s'. Journal of Comparative Economics 25 (3), 403–421.

Murray, Michael P., (2006), 'Avoiding invalid instruments and coping with weak instruments', The Journal of Economic Perspectives, 20 (4): 111-132.

Nyberg, Albert and Scott Rozelle, (1999). 'Accelerating China's Rural Transformation',Washington, DC: World Bank

Psacharopoulos, George, (1994): 'Returns to Investment in Education: A Global Update', World Development 22: 1325–43.

SSBa, State Statistical Bureau, Zhongguo Tongji Nianjian (China Statistical Yearbook). Beijing: State Statistical Bureau, 1990–2000.

SSBb, State Statistical Bureau of China, (1996). 'China Regional Economy: A Profile of Years of Reforms and Opening-Up'. China Statistical Publisher, Beijing.

Yang, Dennis T., (1994). 'Knowledge spillovers and labor assignments of the farm household'. Ph.D. dissertation. University of Chicago.

Yang, D. T. (2005). Determinants of schooling returns during transition: evidence from Chinese cities. Journal of Comparative Economics, 33(2), 244−264.

Zhao, Yaohui, (1997). 'Labour migration and returns to rural education in China'. American Journal of Agricultural Economics 79, 1278–1287

Zhao Yaohui (2002). 'Earnings differentials between State and Non-State Enterprises in urban China'. Pacific Economic Review, 7(1): 181–197

Zhang Junsen, Zhao Yaohui, Park A, Song Xiaoqing (2005). 'Economic returns to schooling in urban China, 1988 to 2001'. Journal of Comparative Economics, 33: 730–752

Appendix A: Surveyed provinces in CHNS dataset

The provinces sampled are broadly representative of China's rich urban regional variation. They include Liaoning, Heilongjiang, Jiangsu, Shandong, Henan, Hubei, Hunan, Guangxi, and Guizhou. Two of these are dynamic high-growth provinces in China's east coastal region (Jiangsu and Shandong), two are located in the northeast (Liaoning and Heilongjiang), and one is heavily industrialized (Liaoning), three are located in the middle region (Henan, Hubei, and Hunan), and two are provinces in the southwest with concentrated populations of ethnic minorities (Guangxi and Guizhou).

Administrative Divisions of the People's Republic of China (PRC)

Table: Surveyed administrative divisions (highlighted in the map)

Administrative Division	code	CHNS dataset
Liaoning	21	1989, 1991, 1993, 2000, 2004, 2006
Heilongjiang	23	1997, 2000, 2004, 2006
Jiangsu	32	1989, 1991, 1993, 1997, 2000, 2004, 2006
Shandong	37	1989, 1991, 1993, 1997, 2000, 2004, 2006
Henan	41	1989, 1991, 1993, 1997, 2000, 2004, 2006
Hubei	42	1989, 1991, 1993, 1997, 2000, 2004, 2006
Hunan	43	1989, 1991, 1993, 1997, 2000, 2004, 2006
Guangxi	45	1989, 1991, 1993, 1997, 2000, 2004, 2006
Guizhou	52	1989, 1991, 1993, 1997, 2000, 2004, 2006

Appendix B: Constructing variables

Hourly wage rate

The annual earnings generally includes regular wages, other income from the work unit such as hardship allowance, private enterprise proprietor's pre-tax net income, individual enterprises proprietor's pre-tax net income, income of employees of individual enterprise, income of re-employed retired member, other employee income, second-job income, property income such as interest, dividends, net profits from stock/bond trading, property rentals, transfer income, and income from household sideline production.

1. CHNS-1991 urban and rural survey

   Annual earning = C8*12 + (I9+I11+I12+I13+I13A+I14)*12 + I19

   1) C8: monthly wage
   2) I9-I14: monthly subsidy
   3) I19: annual bonus

   Hourly wage rate = Annual earnings / (daily working hours * weekly working days * 50)

2. CHNS-1993 urban and rural survey

   Annual earning = C8*12 + (I9+I11+I12+I13+I13A+I14)*12 + I19

   4) C8: monthly wage
   5) I9-I14: monthly subsidy
   6) I19: annual bonus

   Hourly wage rate = Annual earnings / (daily working hours * weekly working days * 50)

3. CHNS-1997 urban and rural survey

Annual earning = C8*12 + (I9+I11+I12+I13+I13A+I14)*12 + I19

1) C8: monthly wage

2) I9-I14: monthly subsidy

3) I19: annual bonus

Hourly wage rate = Annual earnings / (daily working hours * weekly working days * 4 * annual working months)

4. CHNS-2000 urban and rural survey

Annual earning = C8*12 + I14A*12 + I19

1) C8: monthly wage

2) I14: monthly subsidy

3) I19: annual bonus

Hourly wage rate = Annual earnings / (daily working hours * weekly working days * 4 * annual working months)

5. CHNS-2004 urban and rural survey

Annual earning = (C8*12 + I14A*12 + I19) + (C8A*12 + I14B*12 + I19A)

1) C8: monthly wage for primary job

2) C8A: monthly wage for secondary job

3) I14A: monthly subsidy for primary job

4) I14B: monthly subsidy for secondary job

5) I19: annual bonus for primary job

6) I19A: annual bonus for secondary job

Hourly wage rate = Annual earnings / $\sum_{i=1,2}$ (daily working hours * weekly working days * 4 * annual working months)

i=1 for the primary job; i=2 for the secondary job

6. CHNS-2006 urban and rural survey

Annual earning = (C8*12 + I14A*12 + I19) + (C8A*12 + I14B*12 + I19A) + I101 + I103

1) C8: monthly wage for primary job

2) C8A: monthly wage for secondary job

3) I14A: monthly subsidy for primary job

4) I14B: monthly subsidy for secondary job

5) I19: annual bonus for primary job

6) I19A: annual bonus for secondary job

7) I101: annual other cash income

8) I103: annual other non-cash income

Hourly wage rate = Annual earnings / $\sum_{i=1,2}$ (daily working hours * weekly working days * 4 *

annual working months)

i=1 for the primary job; i=2 for the secondary job

Ownership

1. SOE and public sector

   CHNS1991, CHNS1993, CHNS1997: state enterprise or institute

   CHNS2000: government units, state enterprise or institute

   CHNS2004 and CHNS2006: government department; state service / institute; state-owned enterprise

2. Collective enterprises

   CHNS1991, CHNS1993, CHNS1997, CHNS2000, CHNS2004 and CHNS2006: Small collective enterprise (such as township enterprise); large collective enterprise (such as county, city or provincially owned enterprise)

3. Private or Foreign enterprises

   CHNS1991: Joint venture; individual or private enterprises

   CHNS1993: Individual; three source invested enterprise and household business

   CHNS1997, CHNS2000, CHNS2004 and CHNS2006: private, individual enterprise; three-capital enterprise (owned by foreigners, overseas Chinese and joint venture)

Tap water

CHNS1991, CHNS1993, CHNS1997, CHNS2000, CHNS2004, CHNS2006 (1=having tap water, if choosing 1or 2 for question L1)

L1: How does your household obtain drinking water?

1- Piped/tap water in house

2- Piped/tap water in courtyard

Estimated Market Value of House

CHNS1991, CHNS1993, CHNS1997, CHNS2000, CHNS2004, CHNS2006

L18. How much is this house/apartment worth? (Yuan)