PAFTAD

PACIFIC TRADE AND DEVELOPMENT WORKING PAPER SERIES

Paper No. 35-05

Economic Restructuring, Heterogeneous Returns to Schooling and the Evolution of Wage Inequality in Urban China

Rong Zhu

AUTHOR AFFILIATION: National Institute of Labour Studies, Flinders University, Adelaide, Australia.

CONTACT: rongzhu0101@gmail.com

ACKNOWLEDGEMENTS: _I would like to thank Elisabetta Magnani, Denzil Fiebig, Bill Schworm, Denise Doiron, Gigi Foster, Shiko Maruyama, seminar participants at the Melbourne Institute of Applied Economic and Social Research and participants at the 35th Pacific Trade and Development (PAFTAD) Conference for helpful comments. The usual disclaimer applies.

PAFTAD International Secretariat
Crawford School of Public Policy
College of Asia and the Pacific
The Australian National University
Canberra ACT 0200 Australia

Economic Restructuring, Heterogeneous Returns to Schooling and the Evolution of Wage Inequality in Urban China*

Rong Zhu†

Abstract

This paper investigates the evolution of individual-level differences in private returns to education in urban China during 1995 and 2002, a period featuring radical labor and ownership restructuring. While stochastic dominance tests show strong evidence of the trend of rising returns to education, the dispersion in schooling returns is found to diminish dramatically within and across different population subgroups after the economic restructuring. The convergence of schooling returns is interpreted as evidence of a more functioning and increasingly integrated urban labor market for wage earners in China. I also find that the change in the dispersion in schooling coefficients is not responsible for the increase in wage inequality from 1995 to 2002.

JEL classification: I20; J24

Keywords: Returns to schooling, individual heterogeneity, economic transition

*I would like to thank Elisabetta Magnani, Denzil Fiebig, Bill Schworm, Denise Doiron, Gigi Foster, Shiko Maruyama, seminar participants at the Melbourne Institute of Applied Economic and Social Research and participants at the 35th Pacific Trade and Development (PAFTAD) Conference for helpful comments. The usual disclaimer applies.

†National Institute of Labour Studies, Flinders University, Adelaide, Australia. E-mail address: rongzhu0101@gmail.com.

1 Introduction

The economic returns to education in China have been well documented in the existing literature. OLS and IV methods have been widely used to assess how an additional year of education contributes to the increase in average earnings.[1] The literature shows that the rates of return to schooling in urban China were generally below 4% in the 1980s (Byron and Manaloto, 1990; Meng and Kidd, 1997; Liu, 1998), and they were much lower than the world average (10.1%) and Asian average (9.6%) reported in Psachaporoulos (1994). More recent studies find a trend of rising schooling returns among urban workers during China's economic transition (Maurer-Fazio, 1999; Li, 2003; Knight and Song, 2003; Yang, 2005; Zhang et al., 2005). For example, using annual survey data of urban households collected by China's National Bureau of Statistics, Zhang et al. (2005) find that the average rate rose from 4.0% in 1988 to 6.7% in 1995 before reaching 10.2% in 2001.

The fact that differential rates of private return to education may exist in China's urban labour market was acknowledged in previous literature. To capture the potential nonlinearity and heterogeneity, researchers either interact the schooling variable with other controls in Mincer type equations or use an alternative specification with dummy variables indicating discrete levels of schooling. Although widely used, these specifications typically assume that people within a particular group have the same rate of return to education and heterogeneity among individuals within the subgroup is usually not considered. For example, using data from the China Household Income Project of 1988 and 1995, Yang (2005) estimates city–specific schooling coefficients for sampled cities. He finds that the mean of these city–level returns increased from 3.1% in 1988 to 5.1% in 1995, and the dispersion in these estimates rose significantly during the seven years. However, Zhang et al. (2005) find that the returns to education in urban China have demonstrated an apparent trend of decreasing dispersion at the province level. While Yang (2005) and Zhang et al. (2005) provide important and interesting findings, policy makers may want to know how the dispersion in returns measured at the individual level changes during

[1] See Zhang et al. (2005) for a comprehensive survey of the studies on private returns to schooling in China. Liu (2007) is the unique study I know of that examines the external returns to education in Chinese cities.

China's economic transition. As discussed in Harmon et al. (2003), schooling returns are highly likely to vary at the individual level. Card (1995, 1999) shows how the schooling premia could differ even among individuals sharing the same levels of education with a simple static model. In the presence of heterogeneous schooling returns, OLS regression estimates an average effect for the whole population and IV regression could only recover the local average treatment effect of education on earnings for *compliers*, a subgroup of population whose education levels would otherwise be induced to change with the values of instrumental variables (Imbens and Angrist, 1994; Card, 1999, 2001). These average estimates apparently conceal the potential individual dispersion in returns to education, which might contain useful information for policy prescriptions.

There are several reasons why the heterogeneity in schooling returns is worthy of special attention. First, there are good reasons to believe that rate of return to schooling varies across/within groups. Some workers may have attended higher quality schools. Even they share the same amount of schooling, the schooling coefficient could differ; Second, the heterogeneity in schooling returns may inform us of education's effects on wage inequality. This is especially important for a transitional economy such as China. The existing literature on returns to education in China has shown that the average rate of return has increased significantly over time, which has been accompanied by the increase in wage inequality. From the average return, it is natural to conclude that the rising schooling return explains the rising wage inequality. What is missing from looking at the average return is that how the variability of schooling coefficients has changed. With the average return increasing over time, if there is a substantial decrease in the dispersion in schooling returns, the schooling returns may help narrow down the wage inequality rather than drive it up. As a result, looking at the mean rate without checking the heterogeneity in schooling coefficients may lead to misleading conclusions.

To take into account the possible heterogeneity in schooling returns, some studies turn to advanced econometric techniques. Using data from the 2002 China Urban Household Investment and Expenditure Survey, Heckman and Li (2004) estimate the marginal treatment effect of college education on earnings allowing for heterogeneous returns, and they find evidence of differential impacts of college education on earnings.[2] Unfortunately, the marginal treatment effect estimation, which falls into the program evaluation literature, could only measure the treatment effect of a binary variable (for example, college education or not) and the effect is only estimated for a subgroup of people who are at the margin between enrolling in colleges or not (Carneiro et al., 2001). If the evaluation of the heterogeneity for the whole population and for all schooling levels is the central focus, this method does not offer satisfactory solutions.

I know of two studies regarding the heterogeneity itself as being of interest and aiming at uncovering the magnitude of dispersion in schooling coefficients. Harmon et al. (2003) extend the Mincer specification to include the dispersion in the rate of return to schooling by treating the return as a random coefficient that follows a normal distribution. Using the UK labour Force Survey from 1993 to 2000, they find that the dispersion in returns to education in the UK was quite substantial and there was no significant change in the dispersion during the 1990s. With data from the US National Longitudinal Survey of Youth, Koop and Tobias (2004) estimate various Bayesian hierarchical models to investigate the unobserved heterogeneity in returns to schooling. Their empirical results reveal evidence of heterogeneity in schooling returns, which are more likely to follow a continuous distribution rather than a discrete one. The two studies, while insightfully assessing the possible dispersion in schooling returns, both use data sets from developed countries and rely on the assumptions of function forms or specifications of prior distributions over parameters.

This paper seeks to fill the gap in the existing literature by examining the evolution of individual–level differences in the marginal returns to education in urban China between 1995 and 2002. Although there is a growing literature cautioning the potential heterogeneity in individuals' rates of return to education in China (Heckman and Li, 2004; Yang, 2005; Wang et al., 2009), no study has systematically examined it. To circumvent the possible limitations from restrictive assumptions, the nonparametric local linear regression for mixed continuous and categorical data types developed by Li and Racine (2004) is used to estimate observation–specific returns to education. The schooling coefficients obtained

[2] The same semiparametric methodology has also been applied in Wang et al. (2009), which examines the impact of higher education on wage inequality in China.

are realistically allowed to vary at the individual level and are also robust since the method does not impose any distribution or functional form assumptions. With individual–specific schooling coefficients, I make distributional comparisons of returns to schooling among different population subgroups and measure the magnitude of heterogeneity within and across subgroups using inequality indexes.

The radical labour restructuring in China between 1995 and 2002 offers a good opportunity to study the evolution of rates and dispersion in schooling returns under different levels of labour market competitiveness. Before 1995, although some gradualist policies had been introduced to reform the urban labour market since China's economic reform was initiated in the late 1980s, state–owned enterprises under Chinas socialist system were still the principal employers in the urban labour market in the mid–1990s, providing about 60 percent of the total employment and 75 percent of the formal employment. Public sector's commitment to safeguarding the welfare of urban workers persisted well into the mid–1990s. Although there were efforts to increase labour market flexibility, the system that public sectors provided lifetime employment and benefits such as housing, health care and pensions to urban workers, remained quite intact before the mid 1990s, and involuntary dismissal of employees in state sectors was very rare, if not impossible. The wages were more determined by the seniority and ranking in the state firms than by the individual productivity. The rigid labour system and inefficient governance under state ownership resulted in increasingly redundant labour and substantial financial losses.

Aiming at reversing the trend of financial insolvency in state sectors within a three–year period (1997–2000), a radical labour retrenchment program was launched in 1997. Guaranteed lifetime employment for urban workers in the state sector was abandoned and replaced with massive layoffs and forced early retirements. Many state–owned enterprises were privatized and former workers in these units had to find jobs through market channels. As in early phases of other transitional economies in Eastern and Central Europe and the former Soviet Union, the radical labour restructuring led to widespread labour dislocation and the domination of job destruction over job creation (Dong and Xu, 2009). To alleviate the pain of massive labour adjustment, the Chinese Government adopted many policies to support those newly laid–off workers. In addition, the employment expansion of private sectors, along with the nationwide robust economic development, greatly compensated the reduction of employment in state and collective enterprises (Giles et al., 2006; Dong and Xu, 2009). As a result, China experienced a relatively smooth transition in employment shifts within and between sectors. From 1995 to 2001, the number of workers employed in the state–owned sector fell from 113 million to 67 million, a relative decline of 40%. In the meantime, employment in the urban collective sector went down from 31.5 million to 12.9 million (Giles et al., 2006). The state and collective units, which jointly accounted for over 75 percent of all urban employment in 1995, fell to only 37 percent in 2002 (Dong and Xu, 2008). For a more detailed description of the labour restructuring process, I refer to Giles et al. (2006), Dong and Xu (2008) and Dong and Xu (2009).

Using two waves of urban surveys from the China Household Income Project of 1995 and 2002, this paper is the first to systematically investigate the evolution of individual heterogeneity in returns to schooling in urban areas during China's economic restructuring, although I look at the correlations between education and wages only. The year 1995 represents the time when state and collective units remained the principal employers of urban workers, and the year 2002 marks the time when the public sector greatly reduced its impact on urban employment and when the urban labour system became more market–oriented. The nonparametric regression results show that the rates of return to schooling in urban China increased significantly over time and women were rewarded more for education than men. Substantial individual differences in returns to education are uncovered both across and within groups defined by gender, schooling level, occupation, ownership and region in both years. The stochastic dominance tests show strong evidence that the schooling coefficient distribution in 2002 for each gender group first–order stochastically dominates the coefficient distribution in 1995, indicating a trend of rising returns to education during China's economic transition (Maurer-Fazio, 1999; Yang, 2005). I also find that the dispersion in schooling returns diminishes considerably within and across different subgroups. I interpret the narrowing heterogeneity in schooling returns as evidence of a more functioning and increasingly unified urban labour market after the economic restructuring, at least for wage earners. Increased labour market competitiveness and facilitated labour mobility not only serve as equilibrium forces to reduce disparities in returns to education among different regions (Heckman, 2005; Zhang et al., 2005), but also equalize the price of human capital among individual wage earners. While it is commonly believed that individual returns to education tend to converge as the competitiveness and integration of the labour market increase (Heckman, 2005), this paper is the first empirical study that investigates this argument systematically. I also find empirical evidence that the change in the dispersion in schooling coefficients was not a contributor to the increase in the Gini coefficients of hourly wages in urban China from 1995 to 2002.

The remainder of this chapter is organized as follows. Section 2 describes a theoretical model that motivates this study. Section 3 is a description of the data. In Section 4, I describe the empirical approach taken. Section 5 presents the heterogeneous returns to education, makes comparisons of schooling coefficient distributions among different population subgroups, compares the magnitude of heterogeneity within and across subgroups and evaluates the impact of schooling coefficient heterogeneity on wage inequality in urban China. Section 6 contains concluding remarks.

2 Theoretical Motivation

I use a simple static model formulated by Card (1995) and restated in Card (1999, 2001) to illustrate how the returns to education could vary at the individual level. The Becker–type model of optimal schooling assumes that each individual faces a market opportunity locus that gives the levels of earnings and costs associated with schooling choices.

Let $W(S)$ denote the average wage an individual will receive if he or she acquires the schooling level S. The individual is assumed to choose S to maximize the utility function given by:

$$U(S, W) = log(W(S)) - C(S) \tag{1}$$

where $log(W(S))$ represents the overall economic benefits from acquiring education level S and C is an increasing convex function in S which measures the costs of schooling. An optimal schooling choice meets the following first–order condition:

$$\frac{W'(S)}{W(S)} = C'(S) \tag{2}$$

The left–hand side measures the percentage change of wage resulting from one more year of education, and the right–hand side denotes the marginal cost involved. When the marginal benefit is a decreasing function of S and the cost is increasing in S, the two components could be simply specified as:

$$\frac{W'(S)}{W(S)} = w_i - k_1 S \tag{3}$$

$$C'(S) = c_i + k_2 S \tag{4}$$

where k_1 and k_2 are two non–negative constants. w_i is a random variable corresponding to the factors that may affect one's return to schooling. c_i corresponds to the tastes for schooling, access to funds or other known or unknown factors affecting one's costs of schooling. This specification implies that the optimal schooling choice is equal to $S_i^*=(w_i-c_i)/(k_1+k_2)$, which is linear in individual–specific heterogeneity terms. Individual i's marginal return to schooling could be correspondingly obtained as:

$$\beta_i^* = \frac{W'(S^*)}{W(S^*)} = w_i - k_1 S^* = w_i \frac{k_2}{k_1+k_2} + c_i \frac{k_1}{k_1+k_2} \tag{5}$$

Since w_i and c_i are two random variables, the equilibrium of this static model entails a distribution of marginal returns to education across the population. It is likely that the background factors (w_i and c_i) might lead to dispersion in schooling coefficients. Even among people who share the same level of education, the private returns could differ because of the randomness of w_i. The prediction of this simple model motivates me to empirically investigate the differences in economic returns to schooling at the individual level.

3 Data and Descriptive Statistics

For this study, I use the two cross sectional data sets from the China Household Income Project of 1995 and 2002 (CHIP1995 and CHIP2002).[3] The two comparable urban surveys were organized respectively in 1996 and 2003 by the Institute of Economics at the Chinese Academy of Social Sciences for the reference periods of 1995 and 2002. To ensure data representativeness and survey comparability, the two samples were both drawn by the National Bureau of Statistics. CHIP1995 was surveyed at 69 cities in 11 provinces in China encompassing Beijing, Guangdong, Jiangsu and Liaoning from eastern areas, Anhui, Henan, Hubei and Shanxi from the central part, as well as Gansu, Sichuan and Yunnan from western China. The 2002 survey covered the same regions.[4] CHIP1995 covered 6,928 households and 21,688 individuals, and 6,835 households and 20,632 individuals were surveyed in CHIP2002. These data sets are commonly regarded to be representative of individuals with urban household registration status (urban *Hukou*) in China and they have been widely used in existing studies (Li, 2003; Liu, 2007).

In this paper, I focus on urban workers aged between 16 and 60 with positive earnings and, following the existing literature (Li, 2003; Yang, 2005), I exclude those people who are self–employed individuals, retirees, students or homemakers.[5] Owners of private or individual enterprises are also excluded because of the difficulty to separate their wages from profit income. Observations with missing values on any variables used in this study are also dropped.[6] My final sample consists of 10,466 and 9,492 individuals for 1995 and 2002 respectively.

Table 1 presents the means and standard deviations of earnings and selected individual characteristics separately by year and gender. The annual income is the sum of reported

[3]The two waves of CHIP data can be accessed at the Inter–University Consortium for Political and Social Research (ICPSR: http://www.icpsr.umich.edu/icpsrweb/ICPSR/).

[4]Chongqing, which was a sub–provincial city in Sichuan Province surveyed in CHIP1995, became a province–level municipality in 1997. As a result, there were 12 provinces surveyed in CHIP2002.

[5]I focus on urban residents only. Rural–urban migrants were most likely not surveyed in CHIP1995 (Gustafsson and Li, 2000). There is a separate data set for rural–to–urban migrants in CHIP2002 but I exclude it from this study to ensure comparability of estimation results between the two waves.

[6]I dropped the observations with missing values on variables displayed in Table 1. Individuals with missing occupation, ownership, industry and province information are also deleted. See Section 5 for more details.

Table 1: Summary Statistics by Year and Gender

	1995		2002	
	Male	Female	Male	Female
Annual income	7246.63	6062.05	12180.44	9971.39
	(4547.71)	(4207.03)	(8482.79)	(7073.28)
Hourly wage	3.43	2.90	5.86	5.06
	(2.59)	(2.19)	(5.17)	(4.86)
Log(Hourly wage)	1.04	0.85	1.54	1.35
	(0.62)	(0.68)	(0.68)	(0.73)
Age	39.63	37.08	41.72	38.78
	(9.89)	(8.71)	(9.22)	(8.45)
Schooling	11.05	10.43	11.57	11.46
	(2.95)	(2.79)	(3.05)	(2.86)
Experience	22.58	20.65	24.15	21.32
	(10.47)	(9.55)	(10.37)	(9.55)
CCP–member (%)	33.64	15.31	36.38	20.63
	(0.47)	(0.36)	(0.48)	(0.41)
Minority (%)	4.33	4.33	3.86	3.92
	(0.20)	(0.20)	(0.19)	(0.19)
Observations	5,520	4,946	5,261	4,231

Note: *CCP–member* denotes Chinese Communist Party membership and *Minority* represents ethnic minority. Income and wages are in 1995 *yuan* with earnings in 2002 deflated by the national consumer price index. Standard deviations are reported in parentheses.

earnings in all forms received from the current job in each year (including regular wages, bonuses, subsidies and all other income from the work unit). Hourly wages are calculated using annual income and working hours. Following Yang (2005), to enhance earnings comparability, I deflate the earnings in 2002 with the national consumer price index so that all income and wages are measured in 1995 *yuan*. *Experience*, the potential labour market experience, is calculated as age minus schooling minus six.[7]

Earnings had increased substantially for both genders over the seven–year period. The average hourly wage for male workers increased from 3.43 *yuan* in 1995 to 5.86 *yuan* in 2002, and for females it rose from 2.90 to 5.06 *yuan* in the mean time. On average, male workers earned more income and wages than females in both 1995 and 2002. Completed years of schooling had increased by 0.5 years for men and one year for women during the seven years. The gender gap in educational attainments narrowed to only 0.1 years in 2002. For both 1995 and 2002, male workers were slightly older and with more labour market experience. The proportions of Communist Party members (*CCP–member*) rose by 2.7% for males and by 5.3% for females between 1995 and 2002. Moreover, ethnic minorities (*Minority*) made up 4.3% of both male and female workers in 1995, but the two proportions reduced to less than 4% for both genders in 2002.

In Table 2, I report the average years of completed schooling for selected groups. For almost every group, higher education levels were observed in 2002 than in 1995 and men had higher education attainments than women. Years of working experience have been categorized into three subgroups: 0–15, 16–30 and over 30. Newer labour market entrants generally spent more time on education than older ones. In terms of job types, technicians and managerial staff were the two occupations requiring most education, while employees at private firms were among those with the least levels of schooling. Differences in average years of schooling are also observed among employer ownership categories. State–owned units employed the most educated workers in 1995, but it was foreign firms and joint ventures that hired the most educated employees in 2002.[8] In addition, I group those

[7] A few negative values of *Experience* are recorded as zero.

[8] The deepened economic transition in China witnessed a massive downsizing in public sectors. In my sample, the proportion of workers in state–owned units reduced from 82% in 1995 to 32% in 2002. The share of employment in urban collectives went down from 15% to 7%.

Table 2: Average Years of Education for Selected Groups

		1995		2002	
		Male	Female	Male	Female
Experience	0-15	12.22	11.78	13.60	13.40
	16–30	10.93	10.25	11.80	11.18
	Over 30	9.92	8.39	9.79	9.36
Occupation	Technician	12.94	12.31	13.40	13.12
	Office worker	11.14	10.73	12.15	12.09
	Skilled worker	9.67	9.54	9.98	10.36
	Managerial staff	12.27	11.96	13.06	13.03
	Other jobs	9.28	8.97	9.93	10.02
Ownership	State–owned	11.28	10.80	11.10	11.25
	Private firms	9.42	9.09	10.34	10.17
	Foreign firms	10.82	10.55	12.25	12.23
	Urban collective	9.52	9.12	10.10	10.07
	Other types	9.25	9.06	12.07	12.03
Region	Coastal	11.10	10.49	11.61	11.51
	Inland	11.02	10.40	11.55	11.42

Note: Joint ventures are included in foreign firms.

provinces surveyed into two regions according to their geographic locations and economic resemblances in China: the more developed coastal regions and the less developed inland provinces.[9] From 1995 to 2002, I also see the increase in the attained years of schooling for both the richer coastal areas and the poorer inland regions in China. Cross–gender differences in average education levels, although still existent, were found to be negligible in these regions in 2002.

4 Empirical Methodology

The empirical approach takes three steps. First, I obtain an observation–specific schooling coefficient for each individual by employing nonparametric kernel regression. Then I compare the coefficient distributions for different groups through stochastic dominance tests and assess the heterogeneity within and between them using generalized entropy inequality measures. Finally, I evaluate the impact of differential schooling returns on wage inequality in China.

[9]The coastal provinces are Beijing, Guangdong, Jiangsu and Liaoning. All other provinces are located in the inland regions in China.

4.1 Local Linear Kernel Estimation

I use the local linear kernel method developed by Racine and Li (2004) and Li and Racine (2004) to do the estimation, which has an advantage over conventional kernel methods in its capability to smooth both categorical and continuous variables.The nonparametric method is more flexible and robust than parametric methods since functional form and distribution assumptions are avoided and interactions that may exist among all variables are also allowed. Furthermore, this method could generate an observation–specific coefficient estimate for each covariate, which is crucial for assessing the individual heterogeneity in returns to schooling.

The model is specified as $W_i{=}F\,(X_i, Y_i){+}\varepsilon_i$. W_i is the dependent variable, which is the logarithmic hourly wage in my case, and F is the unknown functional form. X_i represents the vector of continuous variables (*Schooling* and *Experience*) and I use s to denote its dimension. Y_i is the vector of unordered discrete variables (*Male*, *Communist–Party membership*, *Minority* and *Province*) and l denotes its dimension. For a typical individual j, I find $\gamma\,(X_j){=}(\alpha_j, \beta\,(X_j))'$ to minimize the following objective function:

$$\sum_{i=1}^{N} \left(W_i - \alpha_j - (X_i - X_j)'\,\beta\,(X_j)\right)^2 K(\widehat{h}, \widehat{\lambda}) \tag{6}$$

where α_j predicts W_j and $\beta\,(X_j)$ is the vector of the partial derivative of $F\,(X_j, Y_j)$ with respect to X for individual j. $K(\widehat{h}, \widehat{\lambda})$ is the multivariate product kernel for mixed categorical and continuous variables, which is equal to $\prod_{q=1}^{s} \frac{1}{\widehat{h}_q} g(\frac{x_{qi}-x_{qj}}{\widehat{h}_q}) \prod_{q=1}^{l} m(y_{qi}, y_{qj}, \widehat{\lambda}_q)$. g is the second–order Gaussian kernel and h_q is the bandwidth for the qth component of X. m is the kernel function for an unordered discrete variable, which equals to 1 if $y_{qi}{=}y_{qj}$ and λ_q $(0{\leq}\lambda_q{\leq}1)$ otherwise.[10] If the bandwidths selected for discrete variables are all equal to 0, the product kernel $\prod_{q=1}^{l} m(y_{qi}, y_{qj}, \widehat{\lambda}_q)$ for discrete variables becomes an indicator function. Consequently, the conventional frequency–based kernel method is a special case of the nonparametric approach I use.

[10] I do not have ordered discrete predictors in this analysis. If there is an ordered categorical variable Z_i, Li and Racine (2004) suggest using the kernel function $n(z_i, z_j, \mu)$ which equals to 1 if $z_i{=}z_j$ and $\mu^{|z_i - z_j|}$ otherwise. μ $(0{\leq}\mu{\leq}1)$ is the smoothing parameter for the ordered discrete variable.

After running OLS regression of W_i on $(1, X_i{-}X_j)$ with weight $K(\widehat{h},\widehat{\lambda})^{\frac{1}{2}}$, the coefficient vector on continuous variables for individual j could be consistently estimated as:

$$\begin{aligned}\widehat{\gamma}(X_j) = \left(\widehat{\alpha}_j, \widehat{\beta}\,(X_j)\right)' = &\left[\sum_{i=1}^{N} K(\widehat{h},\widehat{\lambda}) \begin{pmatrix} 1 & (X_i - X_j)' \\ X_i - X_j & (X_i - X_j)(X_i - X_j)' \end{pmatrix}\right]^{-1} \\ &\sum_{i=1}^{N} K(\widehat{h},\widehat{\lambda}) \begin{pmatrix} 1 \\ X_i - X_j \end{pmatrix} W_i\end{aligned} \tag{7}$$

The basic idea of this nonparametric approach is that when estimating the coefficient for individual j, more weights are assigned to individuals who share similar labour market characteristics to individual j, and less weights are assigned to individuals with less similar personal characteristics.

The selection of smoothing parameters (h, λ) is crucial in kernel estimation, which involves a tradeoff between bias and variance. In this analysis, I employ the least squares cross–validation procedure, which can select the asymptotically optimal bandwidths in the sense of minimizing mean square errors (Li and Racine, 2007). The objective function is given by $CV(h,\lambda){=}\frac{1}{N}\sum_{i=1}^{N}(W_i{-}\widehat{F}_{-i}(X_i, Y_i))^2$, where $\widehat{F}_{-i}(X_i, Y_i)$ is the leave–one–out estimator of $F(X_i, Y_i)$ and (h, λ) is the vector of smoothing parameters selected to minimize $CV(h, \lambda)$. I refer to Li and Racine (2004) for more details about the kernel method and Hayfield and Racine (2008) for how to implement this approach.

This robust and flexible approach generates a unique coefficient estimate of *Schooling* for each individual and the estimates are realistically allowed to vary among individuals. However, these nonparametric estimates could also be inconsistently estimated without controlling for unobserved ability. Similar to previous studies using the CHIP data, the data does not have sufficient information for me to carefully address the unobserved ability bias (Li, 2003; Yang, 2005). Previous studies tend to use the IV methods to correct the unobserved ability bias. However, whether omitted ability seriously affects the returns to education is still an open question (Griliches, 1977; Angrist and Krueger, 1991; Card, 1999, 2001). Ashenfelter et al. (1999) suggest that much of the difference between IV and

OLS estimates is the result of publication bias rather than ability bias, and Manski and Pepper (2000) also suggest IV estimates may be biased upwards. Harmon et al. (2003) also conclude that there is no advantage to use IV methods to correct the ability bias. Moreover, the bias from ignoring individual abilities is somehow compensated by the bias in the opposite direction from disregarding measurement errors. In a recent survey, Card (1999) summarizes the literature and suggests that the omitted variable bias is actually small. The nonparametric estimates herein are comparable with the existing studies on returns to education in China since they also tend not to control for the bias from omitted ability (Liu, 1998; Li, 2003; Yang, 2005; Zhang et al., 2005). My examination of individual heterogeneity in the schooling coefficients without controlling for the unobserved ability is also in line with Harmon et al. (2003) and Koop and Tobias (2004). As discussed in both Harmon et al. (2003) and Koop and Tobias (2004), when the dispersion in schooling returns is the focal interest, the result need not be affected by the omitted ability bias.

4.2 Methodology to Assess the Heterogeneity

4.2.1 Stochastic Dominance Tests

I use the stochastic dominance approach developed by Abadie (2002) to make distributional comparisons of the returns to schooling between different population subgroups and know who benefits most from additional education.

Assuming that β_i is the coefficient on education for individual i, I want to compare the estimate distributions between two groups: group A ($\{\beta_i^A\}_{i=1}^{N_A}$) and group B ($\{\beta_i^B\}_{i=1}^{N_B}$). Let $F_A(\beta)$ and $F_B(\beta)$ represent the cumulative distribution functions of $\{\beta_i^A\}_{i=1}^{N_A}$ and $\{\beta_i^B\}_{i=1}^{N_B}$ respectively.

The two null hypotheses I want to test are:

(1). Equality of distributions:

$F_A(\beta)=F_B(\beta) \quad \forall\beta \in \mathscr{B}$, where $\mathscr{B}$ denotes the union support for β^A and β^B;

(2). First–order stochastic dominance:

F_A dominates F_B if $F_A(\beta)\leq F_B(\beta) \quad \forall\beta \in \mathscr{B}$, with strict inequality for some β.

When testing the two null hypotheses, β_i is replaced with its nonparametric estimate $\widehat{\beta}_i$. F_A and F_B are also replaced with their corresponding empirical distribution functions $\widehat{F}_A$ and $\widehat{F}_B$.[11] As in McFadden (1989) and Abadie (2002), the two test statistics are defined as:

$$T_{ED} = (\frac{N_A N_B}{N_A + N_B})^{\frac{1}{2}} \sup_{\beta \in \mathscr{B}} |\widehat{F}_A(\beta) - \widehat{F}_B(\beta)| \tag{8}$$

$$T_{FSD} = (\frac{N_A N_B}{N_A + N_B})^{\frac{1}{2}} \sup_{\beta \in \mathscr{B}} \left(\widehat{F}_A(\beta) - \widehat{F}_B(\beta)\right) \tag{9}$$

where T_{ED} is the two–sample Kolmogorov–Smirnov statistic to test the hypothesis of equal distributions between group A and group B, and T_{FSD} is the generalized Kolmogorov–Smirnov statistic to test the null hypothesis of first–order stochastic dominance of F_A over F_B.

However, the asymptotic distributions of T_{ED} and T_{FSD} under the null are generally unknown since they depend on the underlying distribution of the data. Abadie (2002) suggests approximating the distributions of test statistics by resampling from the pooled samples and recomputing the test statistics. A four–step bootstrap strategy is thus developed to make the inference about hypotheses possible: (i) Let T be the generic notation for T_{ED} and T_{FSD}. Calculate the statistic $\widehat{T}$ for the original coefficient samples of $\{\widehat{\beta}_i^A\}_{i=1}^{N_A}$ and $\{\widehat{\beta}_i^B\}_{i=1}^{N_B}$; (ii) Resample (N_A+N_B) observations with replacement from the pooled sample of $(\{\widehat{\beta}_i^A\}_{i=1}^{N_A}; \{\widehat{\beta}_i^B\}_{i=1}^{N_B})$, and divide the observations into two groups with sample sizes N_A and N_B. Use the two generated samples to obtain $\widehat{T}_r$; (iii) Repeat step (ii) R times (R=300 in my implementation); (iv) Obtain the p–values of the tests by calculating the relative frequency of $(\widehat{T}_r > \widehat{T})$, which is equal to $\frac{1}{R}\sum_{r=1}^{R} I(\widehat{T}_r > \widehat{T})$. Reject the null hypothesis if the p–value obtained is smaller than some significance level α, $0<\alpha<0.5$.[12]

[11]The empirical cumulative distribution function for group j is defined as $\widehat{F}_j(\widehat{\beta}) = \frac{1}{N}\sum_{i=1}^{N} I(\widehat{\beta}_i^j \leq \widehat{\beta}^j)$, where I is the indicator function.

[12]Ideally, I should compute the nonparametric estimates within each bootstrap replication to take into consideration the uncertainty of estimates. However, this is highly computationally difficult since I need to reestimate the bandwidths for each bootstrap replication. As a result, the bootstrapped p–values I obtain might differ slightly from their *true* values. Eren and Henderson (2008) point out that when a large p–value is obtained, it is unlikely that accounting for such uncertainty would alter the inference.

4.2.2 Generalized Entropy Measures and Heterogeneity Decompositions

I assess the magnitude of heterogeneity in schooling coefficients with the generalized entropy measures. As indexes of inequality with nice properties such as mean independence, scale independence, symmetry and Pigou–Dalton transfer principle, these measures could also be additively decomposed among groups into a "within–group" inequality component (the weighted average of inequality within each subgroup) and a "between–group" component (the inequality across those subgroups when each person's coefficient was equal to the subgroup's mean coefficient), which is the property that the popular Gini coefficient and the coefficient of variation do not offer (Cowell, 2000).

I employ two commonly used measures, the Theil T (denoted with T) and the Theil L (denoted with L).[13] Assume β_i is the return to education for individual i, $\overline{\beta}$ is the average education return and let N denote the number of observations, then the Theil T is defined as $T=\frac{1}{N}\sum_{i=1}^{N}\frac{\beta_i}{\overline{\beta}}ln(\frac{\beta_i}{\overline{\beta}})$ and the Theil L index is defined as $L=\frac{1}{N}\sum_{i=1}^{N}ln(\frac{\overline{\beta}}{\beta_i})$.[14] Suppose the population could be divided into G groups, s^g is the estimates share of group g and $\overline{\beta}^g$ is the average schooling coefficient for group g, then the two inequality measures could be decomposed as:

$$T = T_{within} + T_{between} = \sum_{g=1}^{G} s^g T^g + \sum_{g=1}^{G} s^g ln\left(\frac{\overline{\beta}^g}{\overline{\beta}}\right) \tag{10}$$

$$L = L_{within} + L_{between} = \frac{1}{G}\sum_{g=1}^{G} L^g + \frac{1}{G}\sum_{g=1}^{G} ln\left(\frac{\overline{\beta}^g}{\overline{\beta}}\right) \tag{11}$$

T_{within}, the "within–group" inequality ($\sum_{g=1}^{G} s^g T^g$) measured using the Theil T index, is a weighted average of inequality within subgroups. $T_{between}$ denotes the "between–group" inequality ($\sum_{g=1}^{G} s^g ln(\frac{\overline{\beta}^g}{\overline{\beta}})$) measured across different subgroups. Similarly, I use L_{within} and $L_{between}$ to denote the "within–group" inequality ($\frac{1}{G}\sum_{g=1}^{G} L^g$) and the "between–group" inequality ($\frac{1}{G}\sum_{g=1}^{G} ln(\frac{\overline{\beta}^g}{\overline{\beta}})$) when Theil L is used as the inequality index.

[13] The Theil L is sometimes referred to as the mean log deviation measure.

[14] I replace each β_i with its nonparametric estimate $\widehat{\beta}_i$ when doing the calculation.

5 Results

5.1 Nonparametric Regression Results

5.1.1 Estimation Results Using Basic Control Variables

In this section, I present the nonparametric coefficient estimates of *Schooling* obtained using basic explanatory variables only (*Experience*, *Male*, *Communist–Party membership*, *Minority* and *Province*). A schooling coefficient is obtained for each individual and I follow the approach employed by Henderson et al. (2006) to describe the estimate distribution. I present the nonparametric mean estimate, the estimates corresponding to the 10th, 25th, 50th, 75th and 90th percentiles of the coefficient distribution in Table 3 (labeled $Q10$, $Q25$, $Q50$, $Q75$ and $Q90$). As a comparison, I also run an OLS regression of log hourly wages on the aforementioned control variables and squared *Experience*. I obtain a coefficient estimate of *Schooling* for each sample using the Mincer type specification.

Table 3: Estimates of Returns to Education Using Basic Controls

	1995			2002		
	All	Male	Female	All	Male	Female
OLS	0.0624	0.0541	0.0724	0.0899	0.0800	0.1036
	(0.0021)	(0.0028)	(0.0033)	(0.0027)	(0.0034)	(0.0044)
NP mean	0.0699	0.0624	0.0782	0.1014	0.0932	0.1115
	(0.0020)	(0.0024)	(0.0027)	(0.0028)	(0.0032)	(0.0035)
$Q10$	0.0378	0.0360	0.0508	0.0724	0.0687	0.0866
	(0.0030)	(0.0037)	(0.0041)	(0.0046)	(0.0055)	(0.0059)
$Q25$	0.0514	0.0445	0.0634	0.0876	0.0821	0.0996
	(0.0024)	(0.0030)	(0.0036)	(0.0035)	(0.0043)	(0.0044)
$Q50$	0.0672	0.0584	0.0742	0.1021	0.0931	0.1134
	(0.0026)	(0.0030)	(0.0034)	(0.0031)	(0.0037)	(0.0041)
$Q75$	0.0834	0.0772	0.0933	0.1182	0.1068	0.1278
	(0.0028)	(0.0035)	(0.0038)	(0.0035)	(0.0039)	(0.0059)
$Q90$	0.1062	0.0935	0.1145	0.1312	0.1187	0.1371
	(0.0038)	(0.0041)	(0.0053)	(0.0046)	(0.0050)	(0.0059)
Observations	10,466	5,520	4,946	9,492	5,261	4,231

Note: For the OLS estimates, Huber–White standard errors to correct heteroskedasticity of unknown form are reported in parentheses. Standard errors for the nonparametric estimates are obtained via bootstrapping with 300 replications. All the estimates are significant at the 1% level.

The OLS estimates of returns to education displayed in Table 3 confirm the findings in previous literature that the average rates of schooling return in urban China have increased over time (Maurer-Fazio, 1999; Yang, 2005; Zhang et al., 2005) and women have higher average returns to education than men (Li, 2003; Zhang et al., 2005).[15] The return to an additional year of schooling is estimated to have increased from 6.2% in 1995 to 9% in 2002. However, the local linear kernel regression indicates that the OLS estimate understates the effect of schooling on wages for urban workers. Using the nonparametric approach, the marginal rate of return to an additional year of schooling is estimated to rise from 7% to over 10% during the seven–year period, a relative increase of 42.9%.

I use the consistent model specification test developed by Hsiao et al. (2007) to test the specification of my linear models. The linear specification is rejected at conventional confidence levels for each sample. The simple linear specification I use may omit important interactions and nonlinear relationships among variables. Thus, it is rejected by the specification test for each sample. While misspecified, the linear specification still provides a very good approximation to the nonparametric mean coefficients.

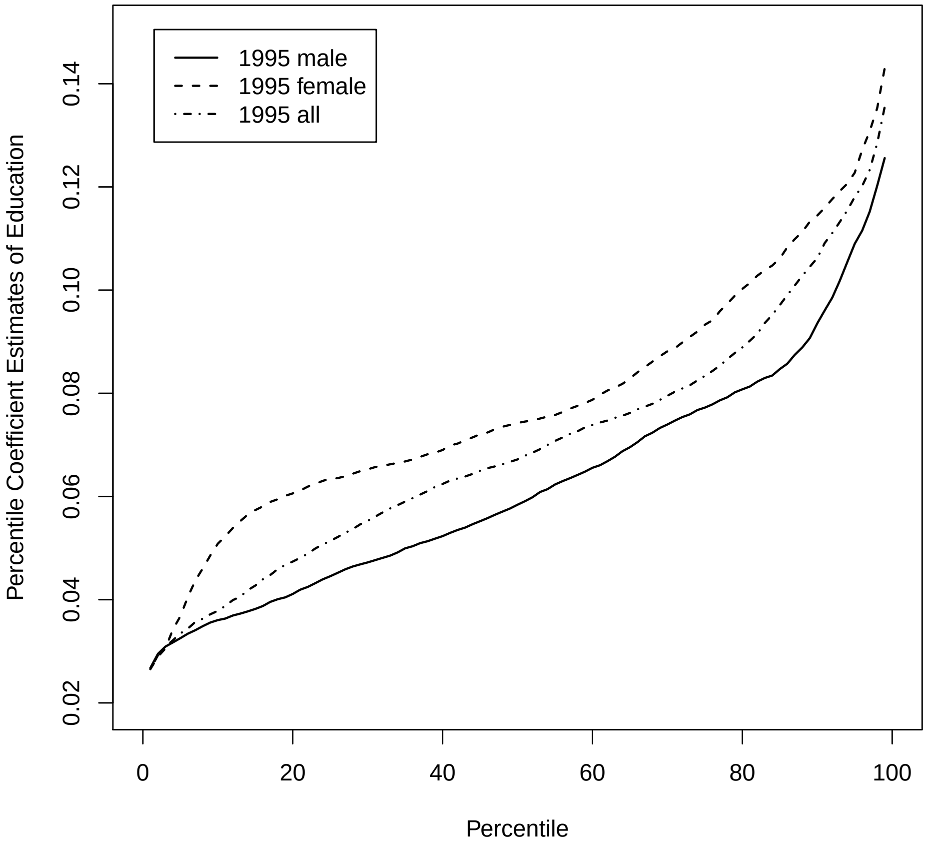
Complementary to the formal statistical test, I also compare the parametric and nonparametric models in light of their ability to fit the sample. The within–sample fitting measures I use are R^2, mean squared error (MSE) and mean absolute error (MAE). Supposing that the dependent variable is W_i and the predicted dependent variable is $\widehat{W_i}$, then R^2 is the squared correlation coefficient between W_i and $\widehat{W_i}$, the mean squared error is defined as $MSE=\frac{1}{N}\sum_{i=1}^{N}[W_i-\widehat{W_i}]^2$, and the mean absolute error is defined as $MAE=\frac{1}{N}\sum_{i=1}^{N}|W_i-\widehat{W_i}|$. The three measures all show that the nonparametric approach fits the samples better. For example, when switching from OLS estimation to kernel regression, the R^2 increases from 0.3163 to 0.3656 for the 1995 sample and for the 2002 sample it rises from 0.2800 to 0.3206. The MSE and MAE reduce from 0.2958 and 0.3948 to 0.2700 and 0.3725 for the 1995 sample, and the two measures decrease from 0.3612 and 0.4569 to 0.3430 and 0.4408 for the 2002 sample. This improvement in sample fitting shows the

[15] The trend of increased returns to education has also been found in other transition economies such as Poland (Rutkowski, 1996; Keane and Prasad, 2006), Russia and Ukraine (Gorodnichenko and Sabirianova, 2005) and Hungary (Jolliffea and Campos, 2005).

flexibility of the nonparametric method since it could relax functional form assumptions and allow for any nonlinearities and interactions in and among all variables.

From the nonparametric estimate distributions, I observe substantial increases in the average and median schooling returns over time. For instance, an additional year of education led to 7.8% growth in wages for female workers in 1995 and it went up to over 11.2% in 2002. The median rate of return for them rose from 7.4% to 11.3% in the mean time. In addition, less than 25% of female workers had schooling returns over 10% in 1995. However, about 90% of women had returns exceeding 10% in 2002. I refer to Yang (2005) and Zhang et al. (2005) for two nice discussions explaining the rising returns to education in urban China. Furthermore, a comparison of the values of $Q10$, $Q25$, $Q50$, $Q75$ and $Q90$ in Table 3 reveals that women enjoyed higher returns to education than men in both years. The median schooling coefficients were respectively 1.6% and 2.0% higher for female workers in 1995 and 2002.

By comparing the values of $Q10$, $Q25$, $Q50$, $Q75$ and $Q90$, I find considerable individual differences in education returns within each gender group. For example, among the male workers surveyed in CHIP1995, the schooling coefficient at the 90th percentile of their estimate distribution (9.45%) is more than twice as large as it is at the 10% percentile (3.60%). To better describe the dispersion in estimates, for every year, I plot the schooling coefficient corresponding to every percentile of each gender group's estimate distribution in Figure 1. Much dispersion is found in the effects of schooling on wages for each group. In addition, from 1995 to 2002, the percentile ratio $Q90/Q10$ of schooling returns reduced from 2.81 to 1.82 between 1995 and 2002. It decreased from 2.60 to 1.63 for male workers and from 2.25 to 1.58 for female workers. This provides preliminary evidence that the within–group heterogeneity in schooling returns diminished from 1995 to 2002 for both genders in urban China, although the rates of returns to schooling had increased substantially. Using inequality measures with the most desirable decomposability properties, I will carefully investigate the differences and changes in the schooling coefficient heterogeneity for selected subgroups from 1995 to 2002 in Section 5.



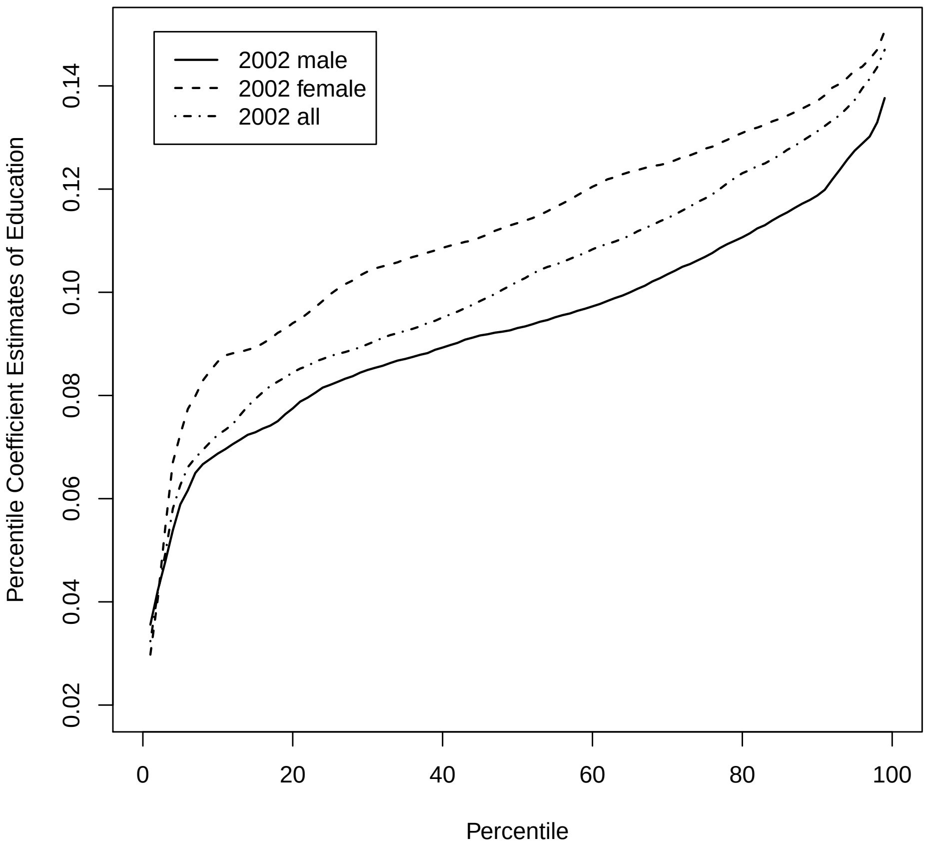


Figure 1: Heterogeneity in Returns to Education for 1995 and 2002

5.1.2 Schooling Coefficients for Selected Groups

I summarize the first, second and third quartile rates of schooling return ($Q25$, $Q50$ and $Q75$) for selected groups in Table 4. The schooling coefficients for subgroups defined by schooling level, job experience, occupation, ownership and region were generally higher in 2002 than in 1995. Moreover, individual differences in schooling coefficients are found within each subgroup.

Among all schooling levels, college graduates had the least marginal returns to education in 1995. I find increasing marginal returns to education levels before college education in 1995. For a typical senior high school graduate whose return equals to the median value of 0.0684, if he/she furthers his/her education by four more years to a college degree, his/her hourly wage could be increased by around 29.4%. For year 2002, the increasing marginal returns to education are found among all education levels. For individuals receiving primary education or less, the median value of marginal return was 9.56%. The value is respectively 10.01% and 10.45% for junior and senior high school graduates. The median value is the largest for college graduates, reaching 10.47%. This finding confirms the empirical regularity summarized in Deng and Li (2009) that returns to education increase with education levels in China. This finding implies that by increasing the education levels of low income individuals in China, the wage inequality could be reduced.

Consistent with the findings in Liu (1998), Maurer-Fazio (1999) and Li (2003), returns to education were higher for new workers. In 1995, the median value of schooling coefficients for workers with 15 years of labour market experience or less was 3.5 percentage points higher than that for those with over 30 years of post–schooling working experience. For the 2002 sample, the value was 2.5%. Possible explanations for this difference among experience groups include a vintage effect, rising quality of education, or greater mobility among younger workers because of fewer employer–specific investments (Li, 2003; Zhang et al., 2005). Among the occupation categories, office workers and skilled workers shared very similar median returns to education in 1995 and 2002. Technicians had less returns than the two professions in 1995 but more in 2002. It was the managerial staff who were rewarded least from an additional year of schooling in both years.

Table 4: Nonparametric Estimates of Returns to Education for Selected Groups

		1995			2002		
		$Q25$	$Q50$	$Q75$	$Q25$	$Q50$	$Q75$
School level	Primary or less	0.0515	0.0656	0.0822	0.0806	0.0956	0.1130
	Junior high	0.0530	0.0674	0.0824	0.0875	0.1001	0.1169
	Senior high	0.0522	0.0684	0.0858	0.0882	0.1045	0.1225
	College and above	0.0433	0.0649	0.0841	0.0888	0.1047	0.1286
Experience	0-15	0.0753	0.0882	0.1082	0.0998	0.1126	0.1281
	16-30	0.0490	0.0637	0.0749	0.0897	0.1021	0.1186
	Over 30	0.0404	0.0515	0.0653	0.0751	0.0885	0.1051
Occupation	Technician	0.0503	0.0664	0.0816	0.0887	0.1039	0.1198
	Office worker	0.0554	0.0712	0.0870	0.0871	0.1031	0.1211
	Skilled workers	0.0536	0.0710	0.0836	0.0871	0.0978	0.1144
	Managerial staff	0.0376	0.0484	0.0650	0.0744	0.0906	0.1056
	Other jobs	0.0601	0.0741	0.0898	0.0917	0.1076	0.1232
Ownership	State–owned	0.0500	0.0666	0.0826	0.0867	0.0988	0.1161
	Private firms	0.0640	0.0814	0.0970	0.0932	0.1073	0.1247
	Foreign firms	0.0642	0.0847	0.1110	0.0936	0.1088	0.1263
	Urban collective	0.0577	0.0686	0.0838	0.0905	0.1077	0.1196
	Other types	0.0649	0.0782	0.0978	0.0867	0.1011	0.1176
Region	Coastal	0.0471	0.0621	0.0791	0.0871	0.1013	0.1148
	Inland	0.0552	0.0722	0.0851	0.0879	0.1023	0.1198

Note: All the estimates are significant at the 1% level.

The returns to education are also found to be higher in private sectors as opposed to public ones. Private sectors generally operate at a higher degree of market mechanism. Thus, the reward for education is higher. For example, workers from foreign firms and joint ventures enjoyed the highest wage effects of education, and their median rate of schooling returns in 1995 was 1.81 percentage points (=0.0847–0.0666) higher than that in state–owned sectors. However, there is a trend of convergence in this gap. It diminished by only 1 percentage point (=0.1088–0.0988) in 2002, although there were substantial increases in returns to education for each ownership category over time. The narrowed differentials are likely to be attributable to the convergence in the wage–setting behavior of the public and private sectors during the public sector reform. In addition, in line with the findings of Li (2003) and Heckman (2005), but in contrast with Liu (1998), the schooling coefficients were higher for the poorer inland provinces. However, these differentials also equalized during the seven–year period and the differences became much less apparent in 2002.

5.2 Evaluation of the Heterogeneity in Schooling Returns

5.2.1 Stochastic Dominance Analysis

The statistics used to describe the schooling coefficient distributions ($NP\ mean$, $Q10$, $Q25$, $Q50$, $Q75$ and $Q90$) literally reveal two main findings in the presence of heterogeneous effects of education: (i) The schooling coefficients were larger in 2002 than in 1995 for males, females and the pooled sample (males and females); (ii) Women were rewarded more for an additional year of schooling than men in both 1995 and 2002. In this section, using a statistical test developed by Abadie (2002), I first test the null hypothesis of equal schooling coefficient distributions between gender and year subgroups. Once it has been determined that the coefficient distributions are different from one another, I use the first–order stochastic dominance tests to check whether there are any orderings or rankings of schooling coefficient *distributions* between different subgroups.

Table 5: Stochastic Dominance Tests, p–values

	Equality of distributions	First–order stochastic dominace
$\widehat{\mathscr{B}}^A_{2002}/\widehat{\mathscr{B}}^A_{1995}$	0.0000	0.9841
$\widehat{\mathscr{B}}^F_{2002}/\widehat{\mathscr{B}}^F_{1995}$	0.0000	0.9733
$\widehat{\mathscr{B}}^M_{2002}/\widehat{\mathscr{B}}^M_{1995}$	0.0000	0.9967
$\widehat{\mathscr{B}}^F_{1995}/\widehat{\mathscr{B}}^M_{1995}$	0.0000	0.8367
$\widehat{\mathscr{B}}^F_{2002}/\widehat{\mathscr{B}}^M_{2002}$	0.0000	0.7067

Note: p–values are obtained via bootstrapping with 300 replications.

The test results are displayed in Table 5. I use $\widehat{\mathscr{B}}^g_y$ to denote the coefficient distribution for year y and group g, where y=1995, 2002 and g=F (female), M (male) and A (all). When testing stochastic dominance, $\widehat{\mathscr{B}}^F_{1995}/\widehat{\mathscr{B}}^M_{1995}$, for example, means the null hypothesis is that the distribution of schooling coefficients for women first–order stochastically dominates the coefficient distribution for men in the 1995 sample. The null hypothesis is rejected if the p–value obtained is smaller than some significance level α ($0<\alpha<0.5$).

I can easily reject equality of distributions at conventional test levels for each case in Table 5. Significant differences or changes in schooling coefficients are found for every two

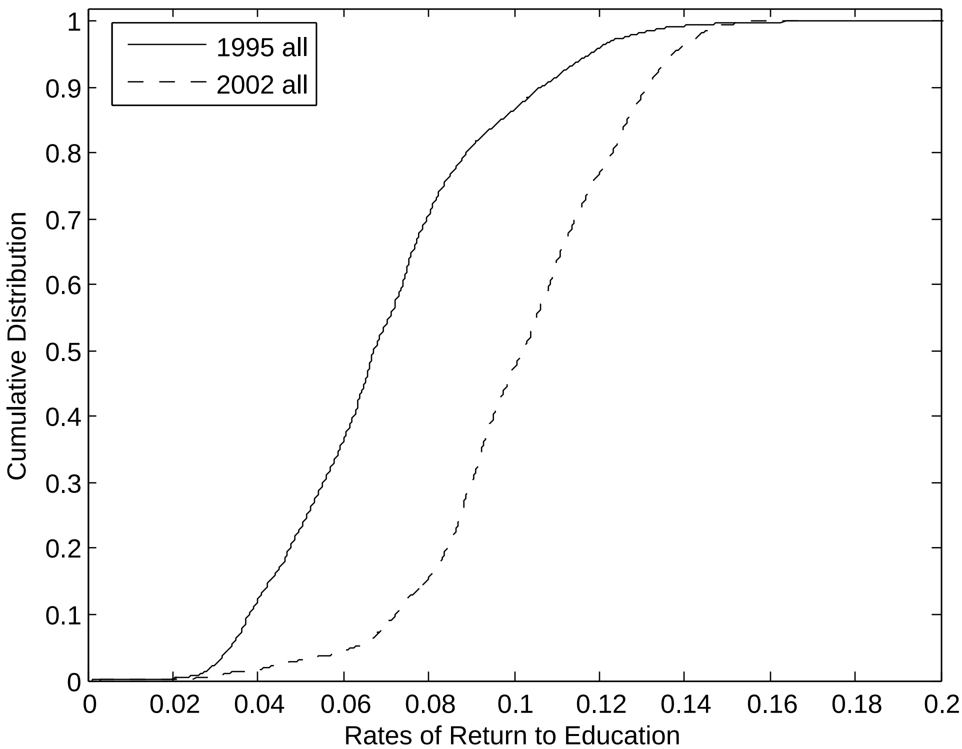


Figure 2: Cumulative Distributions of Schooling Coefficients for 1995 and 2002

groups that I have considered. In terms of rankings, I find strong evidence of first–order stochastic dominance of one group over the other for each of the five cases. These ordering results for *distributions* are much stronger than what I can conclude from just literally comparing the few statistical *values* of schooling returns between different subgroups. For example, I find that people enjoyed higher returns to education in 2002 than in 1995 at the 10th, 25th, 50th, 75th and 90th percentiles as well as at the mean of the schooling coefficient distribution. The stochastic dominance test provides further evidence that for *every* point $\widehat{\beta}^A$ in the union support of $\widehat{\mathscr{B}}^A_{1995}$ and $\widehat{\mathscr{B}}^A_{2002}$, the proportion of people with returns exceeding $\widehat{\beta}^A$ in 2002 is always at least as large as that in 1995, with strict inequality holding for some points. To make this point more clear, I graphically illustrate this ordering relation in Figure 2, in which the cumulative distribution of schooling coefficients for 1995 almost never lies below that for 2002. In addition, for any social welfare functions satisfying certain basic requirements such as Pareto dominance, anonymity, and the population invariance principle, social welfare resulting from returns to education is greater for the first–order dominant estimate distribution in 2002 than for the dominated distribution in 1995, no matter what the exact welfare function form is. The same conclusion also holds for each gender group. These results confirm the rising trend in returns to education when the labour market experiences increasing openness and competitiveness (Heckman, 2005). I also find the coefficient distribution for females first–order stochastically dominates the distribution for males in each year, indicating women indeed benefited more from education than men.

5.2.2 Heterogeneity Measurement and Decomposition

Apart from the evidence of orderings and rankings between the schooling coefficient distributions for different groups, I want to measure the magnitude of the coefficient heterogeneity within and between them. The evolution of between–group heterogeneity shows that how the group divisions affect the schooling return differences. The within–group heterogeneity, which focuses on people with similar characteristics, is a better measure of the impact of institutional changes' on schooling return dispersion. In Table 6, I report the heterogeneity measured using the Theil T and Theil L indexes for selected groups in each year.

The heterogeneity measured using Theil T declined dramatically from 0.0651 in 1995 to 0.0291 in 2002, and using Theil L it decreased from 0.0690 to 0.0338, showing considerable reduction in the dispersion in schooling coefficients over time.[16] The narrowed heterogeneity for these groups is consistent with the findings of Zhang et al. (2005) that from 1988 to 2001 the different OLS estimates of return to education across different groups tend to converge in more competitive labour markets due to factor price equalization. The conclusion still holds even if the rates of return are not assumed to be the same within each particular group and the schooling premia is realistically allowed to vary at the individual level. This finding is also in line with the prediction of Heckman (2005) that more facilitated labour

[16]Note that the Theil indexes require that the support of the schooling coefficients to be positive real numbers. In my estimation, there are respectively seven and five negative estimated coefficients for the 1995 and 2002 samples. I have dropped them when computing the Theil indexes. I believe that dropping the few values will not lead to different conclusions. Yang (2005) uses the standard deviation and the Gini coefficient of city–level estimates to assess the dispersion in returns to education. For my estimates, using the two measures leads to the same conclusion. The Gini coefficient decreased from 0.2026 to 0.1294 between 1995 and 2002, and the standard deviation of the schooling coefficients diminished from 0.0255 to 0.0237. Additionally, the coefficient of variation went down from 0.3654 in 1995 to 0.2336 in 2002.

mobility in China will lead to the equalization of rates of return to human capital.

Table 6: Heterogeneity Measured Using Theil Indexes

		Theil T		Theil L	
		1995	2002	1995	2002
		0.0651	0.0291	0.0690	0.0338
Gender	Male	0.0668	0.0260	0.0688	0.0298
	Female	0.0515	0.0242	0.0556	0.0298
School level	Primary and below	0.0639	0.0359	0.0706	0.0457
	Junior high	0.0576	0.0212	0.0602	0.0221
	Senior high	0.0666	0.0269	0.0699	0.0300
	College and above	0.0870	0.0415	0.0911	0.0520
Occupation	Technician	0.0673	0.0303	0.0718	0.0361
	Office worker	0.0610	0.0338	0.0645	0.0401
	Skilled worker	0.0589	0.0221	0.0612	0.0243
	Managerial staff	0.0647	0.0385	0.0667	0.0444
	Other jobs	0.0524	0.0219	0.0553	0.0247
Ownership	State–owned	0.0673	0.0265	0.0714	0.0300
	Private firms	0.0570	0.0209	0.0577	0.0234
	Foreign firms	0.0578	0.0196	0.0636	0.0206
	Urban collective	0.0601	0.0213	0.0607	0.0257
	Other types	0.0508	0.0335	0.0525	0.0393
Region	Coastal	0.0784	0.0245	0.0818	0.0273
	Inland	0.0563	0.0320	0.0596	0.0379

The two measures (Theil T and Theil L) are all lower for females than males in both years, although I find the strong evidence that women's schooling coefficient distribution first–order stochastically dominates that of men's. For each gender group, the heterogeneity in schooling estimates also decreased over time. When calculating the ratio of men's Theil T index relative to that of women's, the ratio fell from 1.2971(=0.0668/0.0515) in 1995 to 1.2374(=0.0260/0.0242) in 2002, suggesting a narrowed difference in the extent of heterogeneity in returns to education between males and females. This conclusion still holds if the ratios are calculated using the alternative Theil L index.

I also report the within–group dispersion in returns to schooling for each category of schooling level, occupation, ownership and region. I observe a trend of decreasing heterogeneity in schooling coefficients for every category. In addition, I compare the extent of dispersion in returns to education across different subgroups. Among all the schooling levels, the extent of heterogeneity was the largest for college graduates in both 1995 and

2002, which helped drive up the overall dispersion in schooling coefficients. Among all the job categories, the occupation with the maximum heterogeneity in schooling returns switched from technicians and professionals in 1995 to managerial staff in 2002. In addition, the magnitude of heterogeneity seems inversely related to each sector's openness to competition in both years. State–owned enterprises, which were found to have the lowest median returns to education, had the most dispersion among all ownerships in both years. Workers in foreign firms and joint ventures who benefited most from more education had the minimum extent of heterogeneity in education returns in 2002. I also find reduced dispersion in schooling coefficients in the state sector over the seven years, which might reflect the fact that the state sector reformed itself to be more market–oriented in the presence of persistent competition from private sectors. Furthermore, more heterogeneity in schooling coefficients was found for the more developed eastern and coastal provinces in 1995, but the heterogeneity for these provinces became smaller than that for China's inland regions in 2002.

The above indexes only measure the within–group heterogeneity in schooling coefficients. To investigate how the heterogeneity between/across different subgroups had changed over time, I decompose each heterogeneity measure into a weighted average of inequality within different subgroups (within–group inequality) and the inequality between/across different subgroups (between–group inequality). The Theil T and Theil L decompositions show consistent results and they are displayed in Table 7.

As I observe a sharp decrease in heterogeneity for every subgroup defined by gender, schooling level, occupation, ownership and region in Table 6, the within–group heterogeneity in schooling returns, which is a weighted average of the inequality within each subgroup, also diminished substantially from 1995 to 2002. The between–group inequality, which contributes much less to the overall heterogeneity than the within–group inequality, also shows a declining trend over time for most groups. For example, the between–gender heterogeneity in schooling returns declined from 0.0064 to 0.0040 when the Theil T is used as the measure of overall dispersion. Moreover, the between–group heterogeneity for the categories respectively defined by occupation, ownership and region in 1995 are all at least

Table 7: Decompositions of Heterogeneity Using Theil Indexes

		Theil T		Theil L	
		1995	2002	1995	2002
		0.0651	0.0291	0.0690	0.0338
Gender	Within–group	0.0587	0.0251	0.0625	0.0298
	Between–group	0.0064	0.0040	0.0065	0.0040
School level	Within–group	0.0648	0.0289	0.06867	0.0336
	Between–group	0.0003	0.0002	0.0003	0.0002
Occupation	Within–group	0.0601	0.0280	0.0637	0.0326
	Between–group	0.0050	0.0011	0.0053	0.0012
Ownership	Within–group	0.0643	0.0287	0.0682	0.0334
	Between–group	0.0008	0.0004	0.0008	0.0004
Region	Within–group	0.0640	0.0290	0.0679	0.0337
	Between–group	0.0011	0.0001	0.0011	0.0001

twice as large as the between–group heterogeneity in 2002.

I interpret the narrowing heterogeneity in schooling returns from 1995 to 2002 as evidence of a more functioning and increasingly integrated urban labour market in China. The radical labour restructuring program between 1997 and 2000, which abandoned lifetime employment for state sector workers, forced massive laid–off workers from state ownership to find jobs in private and foreign sectors through market channels, and greatly reduced the importance of public sectors in urban employment, has intensified the role of market forces in the allocation of human resources in the urban labour market (Dong and Xu, 2009). The previous rigid labour system and inefficient governance had created severe labour redundancy and skill mismatch (Dong and Xu, 2009), which might result in smaller wage differentials but quite dispersed returns to human capital even among people sharing similar levels of education. After the restructuring, urban workers are experiencing increasingly more incentives and freedom to move between different jobs with higher returns across different sectors. The increased labour market integration and more facilitated labour mobility not only serve as equilibrium forces to reduce disparities in returns to education among different regions as found in Zhang et al. (2005), but also result in more equalized prices of human capital among individual wage earners.

5.3 Evolution of the Effects of Differential Schooling Returns on Urban Wage Inequality

China's economic restructuring was accompanied by a dramatic increase in wage inequality (Whalley and Xing, 2010). In my sample, the Gini coefficient of hourly wages had risen from 0.3328 in 1995 to 0.3737 in 2002 (see Table 8). As I find the evidence of convergence in returns to years to schooling, it is natural to predict that the rapid increase in wage inequality was not attributable to the change in the dispersion in schooling coefficients. In this section, I quantify the effects of differential schooling premia on wage inequality in urban China and check whether the empirical results are consistent with my prediction.

To investigate the impact of the schooling coefficients' on wage inequality, I resort to the method developed by Firpo et al. (2009), which builds upon the concept of the influence function. The influence function is widely used in robust statistics to represent the influence of an individual observation on a distributional statistic such as quantile, Gini coefficient or other measures of interest. By adding the influence function back to the statistic of interest (in my case, the Gini coefficient), the recentered influence function (RIF) is obtained. Firpo et al. (2009) show that by running an OLS regression of the recentered influence function of Gini coefficient on covariates (schooling coefficient in my case), the marginal effects of schooling coefficients ($\widehat{\gamma}$) on the Gini coefficient of wages could be recovered. The regression results are displayed in Table 8. More details about this method are shown in the Appendix.

In the regressions using full samples, the coefficients of returns to schooling are positive, showing that schooling coefficients have positively contributed to wage inequality in both years. As expected, the differential schooling coefficients play a diminishing impact on wage inequality from 1995 to 2002, indicating that the dispersion in schooling coefficients was not responsible for the *increase* in overall earnings inequality after the economic restructuring. As the schooling coefficients I have obtained are marginal returns to education given individuals' education levels, I also do the estimations using subsamples of people with similar education levels and check whether the dispersion in schooling returns contributed to the rising within–group wage inequality from 1995 to 2002. As shown in Table

Table 8: Schooling Premia and Wage Inequality

		All	Primary or less	Junior High	Senior High	College and above
1995	*Gini*	0.3328	0.3506	0.3338	0.3156	0.3089
	$\widehat{\gamma}$	1.5329	0.9813	1.7121	1.4484	1.4914
		(0.1368)	(0.2778)	(0.2345)	(0.2236)	(0.4360)
	R^2	0.0119	0.0059	0.0134	0.0129	0.0096
	N	10,466	2,101	3,938	3,221	1,206
2002	*Gini*	0.3737	0.3552	0.3685	0.3687	0.3301
	$\widehat{\gamma}$	0.6880	0.9217	0.9507	0.2605	0.8768
		(0.1838)	(0.3527)	(0.3598)	(0.3532)	(0.3112)
	R^2	0.0015	0.0055	0.0022	0.0002	0.0045
	N	9,492	1,227	3,180	3,312	1,773

Note: Standard errors are reported in parentheses.

8, the dispersion in schooling returns contributed less in magnitude to the wage inequality in 2002 than in 1995 among the people sharing similar levels of education. As a result, I conclude that the change in the dispersion in schooling coefficients was not responsible for the increase in the wage inequality from 1995 to 2002.

6 Conclusion

This paper is the first study that examines the individual–level heterogeneity in private returns to education in urban China. Using local linear kernel estimation, I obtain an observation–specific schooling coefficient for each individual. Substantial individual differences in returns to education are found for each group defined by gender, schooling level, occupation, ownership and region in both 1995 and 2002. The nonparametric regression results also show that the rates of schooling returns in urban China have increased significantly after the massive downsizing of public sectors and the employment shifts within and between different ownership sectors.

To better describe the dispersion in returns to education for different groups, I first compare schooling coefficient distributions through stochastic dominance tests. The test results show strong evidence that the schooling coefficient distribution in 2002 for each gender group dominates the corresponding distribution in 1995, and the coefficient distribution for females first–order stochastically dominates those for males in each year. Then

I examine the magnitude of heterogeneity within each group using two generalized entropy measures: Theil T and Theil L indexes. The heterogeneity in schooling returns in urban China diminished within each gender, schooling level, occupation, ownership and region group from 1995 to 2002, although their rates of education returns have increased substantially over time. In addition, although men were found to have lower returns to education than women, the magnitude of heterogeneity in schooling returns was more pronounced among male workers in both years. Finally, for the groups respectively defined by gender, schooling level, occupation, ownership and region status, I decompose each heterogeneity measure into a within–group inequality and a between–group inequality. Both of them had declined considerably after the economic restructuring process. I interpret the increased homogeneity in schooling returns from 1995 to 2002 as evidence of a more functioning and an increasingly competitive urban labour market for wage earners in China. I also find empirical evidence that the dispersion in schooling coefficients was not responsible for the increase in wage inequality from 1995 to 2002.

Appendix A Firpo et al. (2009)'s Method Based on Recentered Influence Function

The empirical approach builds upon the concept of influence function, which is widely used in robust statistics to represent the influence of an individual observation on a distributional statistic such as quantile, Gini coefficient or other measures of interest. By adding the influence function back to the statistic of interest, the recentered influence function (RIF) can be obtained. Firpo et al. (2007) and Firpo et al. (2009) show that by running an OLS regression of the recentered influence function on covariates, the marginal effects of controls on that statistic can be recovered.

Following the notation in Firpo et al. (2007), let Y denote the wage variable and μ represent its mean, then the Gini coefficient of wages is defined as $\beta^{Gini}(F_Y) = 1-2\mu^{-1}R(F_Y)$. $R(F_Y)$ equals to $\int_0^1 GL(p, F_Y)dp$ with $p(y)=F_Y(y)$ and $GL(p, F_Y)$ is the generalized Lorenz ordinate of F_Y given by $GL(p, F_Y)=\int_{-\infty}^{F^{-1}(p)} zdF_Y(z)$. The recentered influence function

(RIF) of Gini is obtained as:

$$RIF(y, \beta^{Gini}) = 1 + \lambda(F_Y)y + \delta(y, F_Y) \tag{12}$$

where $\lambda(F_Y)$ equals $2\mu^{-2}R(F_Y)$ and $\delta(y, F_Y)$ is equal to $-2\mu^{-2}[y(1-p) + GL(p, F_Y)]$. A linear specification is typically assumed for $RIF(y, \beta^{Gini})$:

$$E[RIF(y, \beta^{Gini})|X] = X'\gamma \tag{13}$$

Then the $\widehat{\gamma}$ estimated as $(X'X)^{-1}X'RIF(y, \beta^{Gini})$ in the RIF-OLS regression can be used to estimate the marginal effect of X on the Gini coefficient of Y. Since the $RIF(y, \beta^{Gini})$ is never observed in practice, following Firpo et al. (2009), I replace all unknown components with their sample estimators in my empirical application.

References

Abadie, A. (2002). "Bootstrap tests for distributional treatment effects in instrumental variable models". *Journal of American Statistical Association 97*, 284–292.

Angrist, D. J. and A. B. Krueger (1991). "Does compulsory schooling attendance affect schooling and earnings?'. *Quarterly Journal of Economics 106*, 979–1014.

Ashenfelter, O., C. Harmon, and H. Oosterbeek (1999). "A review of estimates of the schooling/earnings relationship, with tests for publication bias". *Labour Economics 6*, 453–470.

Byron, R. P. and E. Manaloto (1990). "Returns to education in China". *Economic Development and Cultural Change 38*, 783–796.

Card, D. (1995). "Earnings, schooling and ability revisited". *Research in Labor Economics 14*, 23–48.

Card, D. (1999). "The causal effect of education on earnings". In O. Ashenfelter and

D. Card (Eds.), *Handbook of Labour Economics*, Volume 3, pp. 1801–1863. North Holland, Amsterdam.

Card, D. (2001). "Estimating the return to schooling: Progress on some persistent econometric problems". *Econometrica 69*, 1127–1160.

Carneiro, P., J. J. Heckman, and E. Vtylacil (2001). "Estimating the return to education when it varies among individuals". Working paper. University of Chicago.

Cowell, F. (2000). "Measurement of inequality". In A. Atkinson and F. Bourguignon (Eds.), *Handbook of Income Distribution*, Volume 1, pp. 87–166. North Holland, Amsterdam.

Deng, Q. and S. Li (2009). "What lies behind rising earnings inequality in urban China? Regression-based decompositions". *CESifo Economic Studies 55*, 598–623.

Dong, X. and C. L. Xu (2008). "The impact of China's millennium labour restructuring program on firm performance and employee earnings". *Economics of Transition 16*, 223–245.

Dong, X. and C. L. Xu (2009). "Labor restructuring in China: Toward a functioning labor market". *Journal of Comparative Economics 37*, 287–305.

Eren, O. and J. D. Henderson (2008). "The impact of homework on student achievement". *Econometrics Journal 11*, 326–348.

Firpo, S., M. N. Fortin, and T. Lemieux (2007). "Decomposing wage distributions using recentered influence function regressions". Working paper, Department of Economics, University of British Columbia.

Firpo, S., M. N. Fortin, and T. Lemieux (2009). "Unconditional quantile regressions". *Econometrica 77*, 953–973.

Giles, J., A. Park, and F. Cai (2006). "How has economic restructuring affected China's urban workers?". *China Quarterly 185*, 61–95.

Gorodnichenko, Y. and K. Sabirianova (2005). "Returns to schooling in Russia and Ukraine: A semiparametric approach to cross-country comparative analysis". *Journal of Comparative Economics 33*, 324–350.

Griliches, Z. (1977). "Estimating the returns to schooling: Some persistent econometric problems". *Econometrica 45*, 1–22.

Gustafsson, B. and S. Li (2000). "Economic transformation and the gender earnings gap in urban China". *Journal of Population Economics 13*, 305–329.

Harmon, C., V. Hogan, and I. Walker (2003). "Dispersion in the economic return to schooling". *Labour Economics 10*, 205–214.

Hayfield, T. and J. S. Racine (2008). "Nonparametric econometrics: The np package". *Journal of Statistical Software 27*, 1–32.

Heckman, J. (2005). "China's human capital investment". *China Economic Review 16*, 50–70.

Heckman, J. and X. Li (2004). "Selection bias, comparative advantage and heterogeneous returns to education: Evidence from China in 2000". *Pacific Economic Review 9*, 155–171.

Henderson, J. D., A. Olbrecht, and S. W. Polachek (2006). "Do former athletes earn more at work? A nonparametric assessment". *Journal of Human Resources 41*, 558–577.

Hsiao, C., Q. Li, and J. Racine (2007). "A consistent model specification test with mixed categorical and continuous data". *Journal of Econometrics 140*, 802–826.

Imbens, G. and J. Angrist (1994). "Identification and estimation of local average treatment effects". *Econometrica 62*, 467–475.

Jolliffea, D. and N. F. Campos (2005). "Does market liberalisation reduce gender discrimination? Econometric evidence from Hungary, 1986-1998". *Labour Economics 12*, 1–22.

Keane, P. M. and E. S. Prasad (2006). "Changes in the structure of earnings during the Polish transition". *Journal of Development Economics 80*, 389–427.

Knight, J. and L. Song (2003). "Increasing wage inequality in China: Extent, elements and evaluation". *Economics of Transition 4*, 317–328.

Koop, G. and J. L. Tobias (2004). "Learning about heterogeneity in returns to schooling". *Journal of Applied Econometrics 19*, 827–849.

Li, H. (2003). "Economic transition and returns to education in China". *Economics of Education Review 22*, 317–328.

Li, Q. and J. Racine (2004). "Cross-validated local linear nonparametric regression". *Statistica Sinica 14*, 485–512.

Li, Q. and J. Racine (2007). *"Nonparametric Econometrics: Theory and Practice"*. Princeton University Press.

Liu, Z. (1998). "Earnings, education and economic reforms in urban China". *Economic Development and Cultural Change 46*, 697–725.

Liu, Z. (2007). "The external returns to education: Evidence from Chinese cities". *Journal of Urban Economics 61*, 642–564.

Manski, C. F. and J. V. Pepper (2000). "Monotone instrumental variables: With an application to the returns to schooling". *Econometrica 68*, 997–1013.

Maurer-Fazio, M. (1999). "Earnings and education in China's transition to a market economy: Survey evidence from 1989 and 1992". *China Economic Review 10*, 17–40.

McFadden, D. (1989). "Testing for stochastic dominance". In T. Fomby and T. Seo (Eds.), *"Studies in the Economics of Uncertainty in Honor of Josef Hadar"*. Springer-Verlag.

Meng, X. and M. P. Kidd (1997). "Labor market reform and the changing structure of wage determination in China's state sector duing the 1980s". *Journal of Comparative Economics 25*, 403–421.

Psachaporoulos, G. (1994). “Returns to investment in education: A global update”. *World Development 22*, 1325–1344.

Racine, J. and Q. Li (2004). “Nonparametric estimation of regression functions with both categorical and continuous data”. *Journal of Econometrics 119*, 99–130.

Rutkowski, J. (1996). “High skills pay off: The changing wage structure during economic transition in Poland’. *Economics of Transition 4*, 89–111.

Wang, X., B. M. Fleisher, H. Li, and S. Li (2009). “Access to higher education and inequality: The Chinese experiment”. IZA Discussion Paper 2823.

Whalley, J. and C. Xing (2010). “The regional distribution of skill premia in urban China”. NBER Working Paper 16575.

Yang, D. T. (2005). “Determinants of schooling returns during transition: Evidence from Chinese cities”. *Journal of Comparative Economics 33*, 244–264.

Zhang, J., Y. Zhao, A. Park, and X. Song (2005). “Economic returns to schooling in urban China, 1988-2001”. *Journal of Comparative Economics 33*, 730–752.