Return to Education in Urban China

By

Haiyun (Jimmy), Du

An essay submitted to the Department of Economics

in partial fulfillment of the requirements for

the degree of Master of Arts

Queen's University

Kingston, Ontario, Canada

July 2014

Acknowledgement

This essay is dedicated to my parents, for without whom, none of this would've been possible. I would also like to thank my adviser, Professor Christopher Ferrall, whose guidance and patience has made this essay possible.

Table of content

Abstract

Building on a recent study by Xiu and Gunderson (2013), this paper studies the return to education in China while controlling for the credential effect associated with university attainment. This paper extends Xiu and Gunderson's study by accounting for endogeneity of the education regressors through the use of two instrumental variables (spousal years of education and spousal university attainment), and looking at the heterogeneity of return to education throughout the conditional distribution with quantile regressions. This paper finds that the IV estimates for additional years of schooling, and university attainment are much higher than their OLS counterparts; particularly, estimates for return to university attainment are reflective of the Chinese's societal attitude towards the importance of a university education. This paper also finds that, both additional years of schooling and university attainment offer greater return to income in the private, as opposed to the state owned sector, in contrast to Xiu and Gunderson's findings. The quantile regression results indicate that return to years of schooling decreases over the conditional distribution, a result that suggests education and talent are substitutes in China. Return to university attainment shows a weakly increasing trend over the conditional distribution in the entire sample, while two different patterns emerge when looking at the male and female subsamples separately. Females are shown to benefit from a greater increase to education over their male counterparts, and the returns to education have increased over time in China.

Introduction

The literature on return to education is divided along two schools of thought in regards to the purpose of education. On the one hand, we have human capital theory, as championed by Mincer (1971), and Becker (1962), which tells us that the role of education is to produce human capital. On the other hand, we have the screening hypothesis, which is derived from Spence's signaling model (1973), and tells us education serves only to signal to employers about the innate abilities of an individual, while adding no other value to the individual.

The traditional empirical approach to studying return to education is done by estimating the Mincer equation, which postulates that the log of one's income is a linear function of their years spent in schooling, and a quadratic polynomial in their years of labour market experience. As such, this approach attempts to study the return one gets from education by estimating the coefficient on the regressor for years spent in school. Although a good first approximation, this approach ignores the discrete increase in income one experiences by completing key stages of the education system, which can be thought of as the gain associated with the signaling effect of education, the so called sheepskin, or credential effect.

Labour economists have begun to study return to education while also controlling for possible sheepskin effects. This is done by augmenting the Mincer equation with additional terms that capture the completion of various stages of education system. Using this approach, one can separate out the human capital gains associated with additional years spent in school from the signaling effect gained by completing key stages of education deemed significant by society. Although many empirical studies on the

sheepskin effect have been done in the context of various countries, as are mentioned in the literature review section, almost none exists for the Chinese context.

Although many studies on the return to education have been done for China, almost all of them focus on estimating a Mincer type equation while only controlling for years of education or credentials obtained. (Liu and Zhang 2013) One exception is Xiu and Gunderson's 2013 "Credential Effects and the Return to Education in China".

Xiu and Gunderson (2013) estimates a basic Mincer equation while additionally controlling for the completion of key stages of the Chinese education system, as well as a female dummy, and a communist party membership dummy, For the completion of key stages, the authors control for several dummies associated with these key stages: junior high school, senior high school, technical college, 2 year college, and 4 year college. This basic model is then estimated using the standard OLS estimator, for the entire sample, as well as the male and female subsamples to compare. In addition, they also compare their results between the subsamples of workers in the state-owned sector, as well as the private sector.

Building on Xiu and Gunderson's study on the return to education in China, this paper makes several changes of its own. In particular, this paper attempts to control for endogeneity of education, and looks at the heterogeneity in the return to education across the distribution of the dependent variable.

Xiu and Gunderson's study controls for possible credential effects associated with the completion of a large number of key-stage education milestones. This paper instead focuses only on controlling for whether the individual was able to graduate from a 4 year university program, or a graduate program. There are several reasons for this. The first

reason is in regards to the problem of endogeneity. As with all other studies on the return to education, the question of endogeneity, particularly the problem of the upward bias observed in the estimates that results from not being able to control for the latent talent variable, is one that always get asked. Xiu and Gunderson addresses this issue by pointing to the infeasibility of finding the large number of instruments related to all 5 of its regressors of interest, and yet, are uncorrelated with the error term. Further, the authors justify not dealing with endogeneity by stating that, since the literature has shown that IV estimates for return to education tend to be larger than their OLS counterparts, and at the same time, since the use of survey data is susceptible to measurement errors, the estimates will tend to be lower than their OLS counterparts. Because of these two counteracting factors, the authors proceed with the use of simple OLS. By reducing the number of regressors of interest down to two, that of years of education and the completion of university, this paper can address the problem of endogeneity by instrumenting the two regressors with two widely used instruments in the literature, the spouse's years of education and the whether one's spouse went to university.

Another reason for focusing on the completion of university can be attributed to its importance in the Chinese society. In the Dynastic period of China's history, becoming a civil servant was the only way one could escape the subsistence living of a peasant (Merchants were seen to be the lowest class in the Chinese society, as they do not create anything on their own, rather, they simply sell the goods made by others (Barbieri-Low 2007)). In order to be selected for the state bureaucracy, candidates had to pass the Imperial Examination. Individuals spent their entire lives preparing for the examination, as success meant upward social mobility, and becoming one of the social elites

(Benjamin 2002). Because of the benefits associated with passing the Imperial Examination, education was heavily stressed in the Chinese society. This belief in the importance of education persists today (Mullins, 2005).

The second improvement this paper make to Xiu and Gunderson (2013) is by looking at the pattern of return to education throughout the entire conditional income distribution, rather than just at the mean. The paper accomplishes this by estimating quantile regressions. As well, IV quantile regressions are also estimated.

This paper is organized as follows. In the second section, I review the literature on return to education and the sheepskin effect. The third section briefly outlines the history of the Chinese urban labour market reforms. The fourth section outlines the methodology and estimation strategies used in this paper. The fifth section looks at justification for the use of the chosen IV. The sixth section describes the data, and the seventh section presents the summary statistics. The eighth section presents the regression results and discussions while the last section concludes.

Literature Review

Hungerford and Solon (1987) were the first ones to show, empirically, the existence of the so called "Sheepskin Effect". Using the Current Population Survey data on white males between the ages of 25 and 64, Hungerford and Solon estimated a Mincer equation while controlling for dummies that indicated whether one had finished years in the education system that corresponded to the average time it took to complete significant milestones (e.g. 12 years that represents high school completion, 16 years for a bachelor's

degree). The authors found significant estimates for all the coefficients on the credential dummies.

Following Hungerford and Solon's example, many other papers on the sheepskin effect followed by estimating the augmented Mincer equation, which includes both total years of schooling (S), and degrees obtained (D). Although the interpretation for the coefficients on S and D can be summarized as the economic gains associated with education, which operates through the two independent channels of human capital gains (as captured by S), and signaling one's ability to potential employers (captured by D), there hasn't been any real theoretical justification for why these two factors can independently vary, after all, earlier papers such as Hungerford and Solon simply derived their credential dummies from the years of schooling data. Flores-Lagunes et. al (2007) presented a human capital model to explain the independent variations in S and D, which gave theoretical justifications for the inclusion of both variables in the empirical model found in the literature. The model tries to explain why time spent in school (S) could have varied among individuals with the same degree obtained (D); which predicted that ability is negatively correlated with time spent in school for degree holders, while it is positively correlated among non-degree holders. In particular, the model leads to the specification of a wage equation in which the derivative with respect to S is negative among degree holders, and positive among non-degree holders.

Under this framework, researchers have found strong evidences for the existence of sheepskin effects in developed countries such as Japan (Bauer et al. 2005), USA, (Flores-Lagunes and Light 2010, and Jaeger and Page 1996), Canada (Ferrer and Riddell 2002), New Zealand (Gibson 2000), and Sweden (Antelius 2000). The studies done for

developing countries however, have found less convincing evidences. For example, evidence have been found for credential effect in Colombia (Mora and Muro 2008) and males in the Philippines (Schady 2003), yet, a lack of evidence was found in Libya (Arabsheibani and Manfor 2001), and Guatemala (Patrinos 1996).

Despite these large number of studies in other countries, there have been almost none on the sheepskin effect in China. Most of the studies done on return to Education in China either focused on the effect of completion of key-stages of education, and the comparison of wages for individuals in these groups alone, (e.g. Gustafsson and Li 2000, Heckman and Li 2004, Knight and Song 2003, Liu 1998, and Li 2003), or they focused on the effect of additional years of schooling by estimating the standard Mincer equation in studies such as Chen and Hamori (2009), Wang (2011), Fleisher et al. (2010), Johnson and Chow (1997), Liu (2007), Maurer-Fazio (1999), Yang (2005), and Zhang et al. (2005).

Additionally, some of these studies make use of instrumental variables to tackle the endogeneity problem. Heckman and Li's paper makes use of parent's education and year of birth, while Wang, and Chen and Hamori make use of spousal years of education. However, the use of parent's income as an instrument may be suspect, as Chen and Feng (2009) found that even after controlling for one's education, both father and mother's education had a direct and independent effect on one's income.

Chinese Background

In this section, I give a summary of the history of China's urban labor market reforms.

Before the economic reform that took place after the death of Mao in 1976, the typical urban worker would be assigned a job in one of the government agencies or state owned enterprise after completion of some education (Zhang et al., 2005). This assignment would be determined by the level of her education, field of study, and family background. Once assigned, it was very difficult and rare for someone to change their job, even within their work unit. In the work unit, the worker's wage was determined through a wage grade system, in which promotion to a higher grade depended on seniority, rather than productivity. The worker's wage would be kept low, but she would be compensated with various benefits that included free education for her children, healthcare, pension, and housing. Due to the assignment system, it was also very difficult for workers to be dismissed.

After the death of Mao, Deng Xiaoping initiated a series of economic and education reforms that radically departed from the system used under Mao. Starting in 1977, The National Higher Education Entrance Examination was used again to select students into the higher education system, whereas previously under Mao, students were selected based on political and family background. Shortly afterward in early 1980's, the wage grade system was replaced by a new, more flexible system, where the wage of workers also depended on the work unit profitability and their own productivity. In 1983, graduates, particularly those from higher education, were given the choice to choose to be either assigned a job by the government, or be allowed to seek their own employment in the private sector. In the same year, a labour contract system was also introduced to selected workers in the state owned sector. Under the contract system, workers signed a contract that specified the durations, responsibilities, and wages of their employment,

very similar to the ones used in western countries today. Once the contract expired, both the employee and employer could choose not to renew the contract. This contract system was gradually adopted to rest of the state owned sector, until about 40% of the urban workers was under some sort of contract in 1995 (Meng 2000).

Methodology and Estimation Strategy

The functional form of the paper's regression model follows that of the classic Mincer Equation, while controlling for higher education attainment and others.

$$\ln(HourIncome_i) = \beta_0 + \beta_1 Years_i + \beta_2 Higher_i + \beta_3\, exp_i + \beta_4\, exp_i^2 + \\ + \beta_5 Female_i + \beta_6 CCP_i + \sum \beta ProvinceDummy_i + \epsilon_i$$

The dependent variable is the natural log of the hourly income. The regressors of interest in this study are that of years of education $Years_i$, and the dummy for obtaining a post-secondary degree $Higher_i$. Following the Mincer's Equation, I also control for the individual's years of labour market work experience, and experience squared. In addition, I also control for the sex of the individual, the dummy for whether they are a member of the Chinese Communist Party (CCP), as well as dummies for the province of residence.

Due to the availability of data, the paper was also able to make use of hourly income as the dependent variable in this study. The 2002 income are deflated to 1995 levels using the national consumer price index. This dependent variable is constructed by dividing the total income by the numbers of hours worked in the survey year. The data for hours worked are self-reported by the individuals in the survey. In contrast, most of the studies on return to education in China have used monthly or annual incomes in the

past. Studies have found that the most educated individuals in China are often the ones that worked the least number of hours during the year. Because of this, studies that have used monthly or annual income tend to underestimate the return to education (Wang 2011, Li and Zax 2000, and Li 2003).

Controlling for the gender of the individual is also a standard procedure in the literature. As studies have shown repeatedly in the past, females often earn less than their male counterparts on average in the labour market (Hausmann et al. 2009).

As standard for return to education studies in China, I control for one's participation in the CCP. In contrast to having membership in a political party in democratic western countries, membership in the CCP is not open to anyone who simply wishes to join. Party membership is only given to individuals who meet the selection criteria, which often entail having greater socio-economic status and having connections within the party (Bian et al. 2001).

Lastly, I include the addition of provincial dummies in my model. Provinces in China vary significantly in terms of population, development status, standard of education and other aspects important to the study. In terms of population, provinces ranges from 5.6 million to over 100 million, more than the entire population of the Philippines.

Although entrance into higher education is dependent only on the National Higher Education Entrance Examination, the actual content of the exams, and the score needed to enter different universities is different across the provinces. In particular, the actual exams are the responsibility of the provincial government, which are usually written by a team of high school teachers and university professors commissioned by the province. As

such, the content of the exams may vary from province to province. As well, universities, especially the top ones, often have a quota on the number of students from the different provinces. This is done to ensure that students from disadvantaged, or less developed provinces are also able to have a chance of going to these university. Once the students receive their grades on the exam, they compare themselves to the score needed to enter the different universities, which are province specific (Hsieh, 2013). As well, ethnic minorities also require a lower score compared to their Han counterparts (the ethnic majority in China, making up almost 92% of the population). With some provinces having significantly more ethnic minorities than others, especially the ones in the south, west, and northeast, provincial heterogeneity in terms of university admission becomes considerable.

Another reason for provincial heterogeneity is the state of economic development amongst the provinces. When economic reforms were initiated in the late 1970's, a few selected coastal provinces were given the opportunity to engage in free trade with the rest of the world to test the policy. As a result of this, and reason of geographic proximity to the ocean in general, coastal provinces are much more developed than ones inland in general. In addition, Fleisher and Chen (1997) attribute the lower economic growth of China's coastal provinces to their inferior total factor productivity.

Standard OLS will first be estimated to give a set of baseline results used for comparison. The paper then proceeds with estimating the model with instrument variables, in particular, the endogenous regressors of years of education and higher education attainment are instrumented with spousal years of education and whether the spouse completed higher education. All of these estimations will be done for the entire

sample, as well as the male and female subsamples. Additionally, the procedure will be repeated using both the 1995 and the 2002 datasets to look at the changes in return to education over the different stages of China's economic development.

The paper then separately estimate the model for subsamples of workers who works in the state owned sector and the private sector. Both OLS and IV results for 1995 and 2002 are presented.

Next, the model will be estimated using quantile regression to examine the variation in return to education throughout the conditional distribution of income in China. Results for the 10^{th}, 25^{th}, 50^{th}, 75^{th}, 90^{th}, and 95^{th} percentiles are presented. Similarly, this will be done for the entire dataset, as well as male and female subsamples for both 1995 and 2002. And lastly, the paper looks at IV quantile regression; however, only higher education attainment is instrumented with spousal higher education attainment due to technical difficulties. The same percentiles and subsamples are used for IV quantile.

Instrumental validity

To deal with the possible endogeneity present in years of education and higher education attainment, I have chosen to use spouse's education and spouse's higher education attainment as instruments for the two suspect regressors. The use of spouse's education attainments as instruments for one's own is something that have been done previously in the literature; both in the Chinese context (Chen and Hamori 2009, Wang 2011), as well as the international context (Trostel et al. 2002, Arabsheibani and

Mussurov 2007). When attempting to use instrumental variables, one has to find suitable instruments that are both correlated with the endogenous regressors, while at the same time uncorrelated with the regression error. I discuss both of these aspects in turn in this section.

The idea that the education attainments of a couple being correlated can be found in Becker's work on positive martial assortative matching (Becker 1981). The theory suggests that individuals will tend to marry others whom they share common characteristics with, education being one obvious example of this. Furthermore, strong empirical evidence for positive martial assortative matching have also been found (Mare, 1991; Qian, 1998), with strong positive correlations found between one's own education level and their spouse's. Moreover, Chen and Hamori (2009), and Chong et al. (2009) have found strong evidence for assortative matching effect in China.

To confirm the results for my own study, tests for weak instruments are conducted. First stage regressions from the standard 2 stage least square procedure found that both of the instruments are significantly different from zero in explaining the variation in both of the endogenous regressors, the results can be found in the appendix. It is found that for every additional year the spouse had spent in school, one's own years of schooling increases by about half a year on average. This increase of half a year of additional schooling is also gained if the spouse had completed higher education. On the other hand, spousal attainment of higher education increases the probability of one's own higher education attainment by almost 30 percent, and each additional year of spousal education increases the probability of one's own higher education attainment by 1

percent. Moreover, the magnitude of these estimates are almost identical for both the 1995 and 2002 years, suggesting stability of the estimates over time.

The appendix reports two additional tests for weak instruments. In particular, the Cragg-Donald Wald F-statistics/Minimum eigenvalue statistics, as well as the Kleibergen-Paap rank Wald F statistic strongly reject the null for both years.

The second condition the instruments have to satisfy is instrument exclusion, which says that the instruments in question cannot directly influence the dependent variable. The way spousal education attainment can influence one's income is by increasing the individual's productivity, or it can be used as a signal in the market place. Wang (2010) used the CHIP 1995 and 2002 data sets to study the robustness of the casual effect of spousal college education on individual's income. His study concluded that the positive gain in income reflected positive assortative matching, rather than direct cross-productivity effects. Additionally, Wang (2011), also using the CHIP 1995 and 2002 datasets, found evidence for spousal years of education satisfying instrumental exclusion by testing the robustness of their results to the presence of a general dependence between potential wage and spousal education.

Additionally, even though my models are exactly identified, with 2 instruments and 2 endogenous regressors, I estimated two test models for each year, where only one of the 2 endogenous regressors is assumed to be actually endogenous, and instrumented that chosen endogenous regressor with both the instruments. The results are again found in the appendix. In this case, I was able to test for instrument exclusion by having an over identified model. The Hansen J statistics were all highly significantly for both of the models, and for each of the 1995 and 2002 samples. For the results, the first model in the

first column assumes years of education is endogenous, and the second column model assumes completion of higher education is endogenous.

Selection Bias

Another problem for studies on return to education concern the self-selection of women into the labour force. Because we only observe data for female workers who are in the labour market, if the factors that influence whether they enter the labour market in the first place also influence their income, then our estimate will be biased. Xiu and Gunderson utilized the Heckman two-step procedure and found no evidence for sample selection in the CHIP data sets. They comment that this result probably stems from the very high female participation rate during this period, which sat at 87 percent, slightly lower than the 96 percent for males. As well, female unemployment rate was also very low, and close to that of the male unemployment rate at 2.4 percent and 3 percent respectively. The high labour force participation of women compared to other countries can be attributed to the Chinese government's emphasis on gender equality in recent years (Wang 2011). Gustafsson and Li (2000) commented on how "women in China are quite similar to men in performing market work.", while Millimet and Wang (2006) found very little difference across the genders in full and part time jobs in China. Moreover, as noted by Wang (2011), "solving the selection-bias problem is not trivial; again, the availability of exogenous IVs is required. The instruments used in the literature are arguably problematic."

Another possible concern relates to the problem of selection into being married, as it is natural to assume that individuals whom are married are subject to the influence of unobservable that may also influence their incomes. As it will be examined more closely in the summary statistics section to follow, although the married individuals tend to be slightly older, and more educated on average than the unmarried ones, the difference between them are not large enough to be a source of major concern.

Data

This paper uses data from the Chinese Household Income Project (CHIP) 1995 and 2002. The CHIP data are cross sectional survey level data collected for the purpose of studying the determinants of employment income in China. CHIP contains data for households and individuals in both rural and urban China, the latter of which is the focus of this study. The urban data sets contain information on over 20,000 individuals across nearly 7,000 households in 11, and 12 representative Provinces for the years of 1995 and 2002 respectively, although after cleaning the data, the number of observations are far less than contained in the original data sets.

The data sets are widely used in the literature due to their high quality of data and documentations. They contains an extensive number of variables covering the work history of individuals, as well as their financial situation. They are also generally considered some of the best publicly available data sets in China, and are also a representative sample of urban Chinese households (Gustafsson et al. 2000)

One of the reason why the CHIP data sets are ideal for studies on the return to education is that they contain not only variables on years of schooling, but also degree

attainment, as well as actual, rather than potential, years of experience for the individuals. The variable for degree attainment is a categorical variable that asks for the highest level of education the individual was able to attain at the end of the survey years. For this study, I only focus on whether the individual was able to obtain at least a bachelor's degree from a traditional 4 year University; this includes students who were able to finish graduate degrees. I lump undergraduate and graduate students together because of the limitation in the data set. In particular, although the 2002 data set distinguishes between bachelor degree holders and graduate degree holders, the 1995 data set does not.

For this study, I focus on individuals who are at least 22 years of age and have at least 12 years of education (so the population in question is at least old enough to has the chance to graduate from university), and those that are no older than 65 (the mandatory retirement age, which could be extended via the Communist party's approval), and were not self-employed. As well, due to the study using the instrumental variables of spouse's years of education and higher education attainment, only those for whom these variables are present were used, namely, married couples.

Summary Statistics.

Mean and standard deviation for the variables used in the paper are presented in table 1A and 1B, which contain the summary statistics for the 2002 and 1995 data sets respectively. The sample means are shown for each variable, with the standard deviation found under them in the parenthesis.

Some noticeable features of the data include the increase in income earned over the 7 year period, with male workers income more than females. The average level of education attainment has also risen over the years, with people spending more years in school and a higher percentage of the population gaining a university degree. Membership in the Communist Party has also increased significantly, with the largest percentage increase occurring among the female workers. In addition to summary statistics on actual years experiences found in the data set, potential experience, a popular proxy to actual experience in the literature, is also constructed to compare. As we can see, potential experience, as defined by age minus years of schooling minus six, is higher than actual experience. This is to be expected, as the construction of potential experience assumes every year not spent in school after the age of six is used to gain experience, which is of course not true.

A summary table of the key variables broken down by provinces is also presented in table 1C. The provinces are numbered in the dataset, which is the way they are labeled in the summary table. Attempts had been made to find the corresponding province with their numerical labelling used in the dataset, but the information could not be obtained.

Lastly, table 1D summarizes the key variables, as well as age, for the unmarried individuals not used in the study. As we can see, on average, the unmarried individuals is a little bit younger (the average age for married individuals is around 40 in the data), earn a little less income, and have less education than the married ones. This is to be expected, as most of the unmarried individuals are still looking for their partner. Although the differences are noticeable, nothing too drastic or norm breaking is present in the data.

Table 1A: Summary Statistics - 2002

2002	Population	Male	Female
Actual Exp	20.81 (7.98)	21.59 (7.97)	19.95 (7.91)
Potential Exp	22.79 (17.91)	23.12 (17.89)	22.44 (17.9)
Income	14929.87 (9558.73)	16409.99 (10690.76)	13296.02 (7812.6)
TotalHour	2144.961 (442)	2174.669 (452.11)	2112.16 (428.35)
lnHourIncome	1.80 (0.62)	1.89 (0.61)	1.71 (0.62)
Years	13.83 (1.77)	14.01 (1.83)	13.64 (1.68)
Higher	0.22 (.42)	0.28 (.45)	0.16 (.37)
exp	20.81 (7.98)	21.59 (7.97)	19.95 (7.91)
expsqr	0.05 (0.34)	0.05 (0.04)	0. 05 (.03)
Female	0.48 (.5)		
CCP	0.46 (.5)	0.55 (.49)	0. 35 (.48)
SpouseEdu	13.83 (1.77)	13.64 (1.68)	14.01 (1.83)
SpouseHigher	0.22 (.42)	0.16 (.37)	0.28 (.45)
Obs	2,998	1,573	1,425

Table 1B: Summary Statistics - 1995

1995	Population	Male	Female
Actual Exp	20.75 (8.22)	21.59 (8.45)	19.64 (7.77)
Potential Exp	22.52 (16.25)	23.74 (16.51)	20.9 (15.48)
Income	7762.61 (4834.96)	8189.28 (5284.75)	7203.52 (4108.02)
TotalHour	1999.47 (330.4)	2015.97 (350.51)	1977.85 (300.78)
lnHourIncome	1.25 (.51)	1.3 (.49)	1.18 (.53)
Years	13.55 (1.68)	13.73 (1.72)	13.31 (1.59)
Higher	0.19 (.39)	0.23 (.42)	0.13 (0.34)
exp	20.75 (8.22)	21.59 (8.45)	19.64 (7.77)
expsqr	0.05 (0.04)	0.05 (0.04)	0.04 (0.03)
Female	0.43 (.5)		
CCP	0.39 (.49)	0.48 (0.5)	0.27 (0.44)
SpouseEdu	13.55 (1.68)	13.31 (1.59)	13.73 (1.72)
SpouseHigher	0.19 (.39)	0.13 (0.34)	0.23 (.42)
Obs	3,447	1,955	1,492

.

Table 1C – Summary Statistics by Provinces

Province			2002				1995	
	lnHourIncome	Years	Higher	Obs	lnHourIncome	Years	Higher	Obs
11	2.16 (.53)	13.57 (1.8)	.15 (.36)	270	1.54 (.41)	13.74 (1.77)	.19 (.39)	329
14	1.58 (.51)	14.13 (1.72)	.3 (.46)	288	.96 (.46)	13.63 (1.58)	.22 (.41)	337
21	1.72 (.58)	13.59 (1.75)	.21 (.41)	249	1.13 (.48)	13.25 (1.38)	.16 (.37)	381
32	1.89 (.62)	13.84 (1.77)	.22 (.42)	270	1.41 (.45)	13.51 (1.71)	.17 (.38)	343
34	1.7 (.52)	14 (2.04)	.22 (.42)	204	1.13 (.46)	13.57 (1.61)	.16 (.37)	237
41	1.54 (.54)	13.77 (1.52)	.19 (.4)	258	.95 (.48)	13.36 (.154)	.21 (.41)	255
42	1.67 (.57)	13.97 (1.79)	.24 (.43)	359	1.29 (.44)	13.5 (1.62)	1.8 (.39)	366
44	2.42 (.61)	13.53 (1.56)	.19 (.39)	248	1.96 (.51)	13.55 (1.81)	.2 (.4)	235
51	1.66 (.59)	13.59 (1.72)	.15 (.36)	211	1.2 (.46)	13.78 (2.02)	.21 (.4)	443
53	1.82 (.5)	14.19 (1.8)	.28 (.45)	327	1.17 (.33)	13.41 (1.53)	.16 (.37)	352
62	1.61 (.6)	13.67 (1.78)	.19 (.39)	193	1.01 (.41)	13.85 (1.67)	.24 (.43)	169
50	1.89 (.64)	13.88 (1.81)	.32 (.46)	131				

Table 1D: Comparison of Key Variables between married and unmarried

		2002			1995	
	Population	Male	Female	Population	Male	Female
lnHourIncome	1.72 (.67)	1.78 (.65)	1.63 (.68)	1.17 (.55)	1.21 (. 53)	1.15 (.57)
P113	13.73 (1.75)	13.84 (1.81)	13.6 (1.67)	13.42 (1.65)	13.53 (1.7)	13.32 (1.59)
Higher	.20 (.4)	.24 (.43)	.16 (.36)	.16 (.37)	.18 (.39)	.11 (.32)
Age	39.36 (9.32)	40.5 (9.58)	37.85 (8.74)	39.42 (8.93)	40.72 (9.3)	37.65 (8.65)
Obs	1628	983	645	1794	1102	692

Regression Results.

The OLS and IV estimation results for the year 2002 and 1995 are found in table 2A and 2B respectively. For both tables, column 1 gives the OLS results when estimated using the entire data set, column 2 and 3 give the male and female OLS results respectively, column 4 through 6 give the IV estimation results for the entire data set, male, and female subsamples in that order.

OLS results

We first look at the more recent 2002 OLS results, and then compare it with the 1995 OLS results.

2A: OLS and IV - 2002

	(1) OLS	(2) OLS Male	(3) OLS Female	(4) IV	(5) IV Male	(6) IV Female
Years	0.0518***	0.0453***	0.0584***	0.0923***	0.120***	0.0745**
	(0.00660)	(0.00903)	(0.00964)	(0.0265)	(0.0372)	(0.0363)
Higher	0.237***	0.205***	0.287***	0.496***	0.274*	0.695***
	(0.0273)	(0.0367)	(0.0414)	(0.112)	(0.143)	(0.168)
exp	0.0174***	0.0191**	0.0127	0.0244***	0.0310***	0.0152*
	(0.00557)	(0.00766)	(0.00802)	(0.00630)	(0.00883)	(0.00878)
expsqr	-1.908	-2.994*	0.104	-2.355*	-4.461**	0.452
	(1.291)	(1.745)	(1.890)	(1.417)	(1.907)	(2.040)
Female	-0.102***			-0.062***		
	(0.0197)			(0.0213)		
CCP	0.0807***	0.109***	0.0477	0.00754	0.0297	-0.0117
	(0.0204)	(0.0277)	(0.0304)	(0.0241)	(0.0349)	(0.0343)
_cons	1.136***	1.304***	0.897***	0.411	0.0953	0.582
	(0.117)	(0.159)	(0.168)	(0.380)	(0.540)	(0.519)
N	2998	1573	1425	2998	1573	1425

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Provincial dummies not included

2002 sample

From columns 1-3 of table 1B, we can see that an additional year of schooling is associated with about a 4 to nearly 6 percent increase in hourly wage. This result is consistent with the findings in other studies in the literature, which estimate a return of about 5-15 percent (Card 2001). We can also see that additional years of schooling has a greater impact on the wage of female workers, with the return being about 4.5 percent for males but nearly 6 percent for females, a difference of about 29 percent. Looking at the return to completion of higher education, we see a wage increase of about 24 percent for the population in general, 20.5 percent for males, and a much higher 27.8 percent for females. In particular, at these estimates, the return to university is roughly the same as the increase in wage gained from spending slightly more than 4 additional years in school, which is about the average time it takes to complete higher education, given that the study has grouped graduate students in the same group. Although this result is perhaps lower than what we would've guessed initially, given how much a university education is stressed in Chinese society, we have to keep in mind that this result indicates an additional increase from completing university, one that persists after we have controlled for the years spent in school. As we'll see later, once endogeneity has been addressed, the return to higher education is estimated to be much higher than under OLS.

Looking at the other coefficients, we see that additional years of experience result in a wage increase of 1.9 percent for males, and 1.27 percent for females, although the latter result is not significant. This result contrasts with the benefit of additional years of schooling, which benefits females more than males. One particularly interesting result that stands out is the insignificance of the squared years of experience variable. This

suggests that hourly wage is a linear, rather than quadratic function in years of experience in our case. This result differs from the one found in the literature, which points to the significance of the squared years of experience term.

The estimate on the female dummy indicates the existence of a highly significant income gender gap in China. The OLS estimates a negative difference of 10.2 percent in hourly wage for females. As well, Communist Party members seem to enjoy an average increase of 8 percent in wage compared to the general population, with males benefiting more from the membership.

1995 sample

We now turn our attention to the estimates found using the 1995 data set. At a first glance, we see that the return to both measures of education have increased by a significant margin over the 7 year period. For return to years of schooling, the 1995 estimates indicate a wage increase of about 3 percent for the general population, 1.87 percent for males, and 4.45 percent for females. Comparing these estimated returns to the 2002 ones, we see that the increase in return to additional years of schooling is largest for the male workers, whose estimated return has more than doubled over the period. There are also significant increases over the years associated with the return to completion of university, with all 1995 estimates being only around 15 to 17 percent. As is the case for 2002, females benefit significantly more than their male counterparts when it comes to additional education.

In contrast to 2002, female workers are the ones with a significant estimate for years of labour market experience, while the males estimate are insignificant in 1995. As was in 2002, the squared years of experience are also insignificant in 1995.

2B: OLS and IV - 1995

	(1) OLS	(2) OLS Male	(3) OLS Female	(4) IV	(5) IV Male	(6) IV Female
Years	0.0295***	0.0187***	0.0445***	0.0261*	0.0232	0.0486**
	(0.00471)	(0.00611)	(0.00746)	(0.0145)	(0.0206)	(0.0202)
Higher	0.151***	0.152***	0.167***	0.411***	0.240**	0.533***
	(0.0202)	(0.0246)	(0.0358)	(0.0726)	(0.0944)	(0.106)
exp	0.0136***	0.00672	0.0202***	0.0187***	0.00888	0.0286***
	(0.00442)	(0.00539)	(0.00760)	(0.00468)	(0.00554)	(0.00803)
expsqr	-0.796	0.487	-1.939	-2.031*	-0.0266	-3.959**
	(0.988)	(1.179)	(1.789)	(1.057)	(1.217)	(1.905)
Female	-0.0622***			-0.0423***		
	(0.0152)			(0.0160)		
CCP	0.0526***	0.0371*	0.0812***	0.0320*	0.0268	0.0510*
	(0.0154)	(0.0198)	(0.0251)	(0.0168)	(0.0205)	(0.0292)
_cons	0.830***	1.064***	0.472***	0.784***	0.969***	0.300
	(0.0837)	(0.110)	(0.130)	(0.202)	(0.290)	(0.283)
N	3447	1955	1492	3447	1955	1492

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Provincial dummies not included

The gender gap in income is also lower in 1995, with the estimates being only about 6 percent. Membership in the Communist Party also increase one's hourly wage. Although the benefit of party membership is not as high for the general population

compared to 2002, it seems however, that females are able to enjoy a much greater benefit.

IV results.

I discuss the results of this section while comparing it to the OLS results found above. Like the discussion on OLS results, I first present the estimates from the year 2002, and then look at the 1995 results to compare.

2002 Sample

A first look at the IV estimates reveal a higher return to both measures of education when compared to their OLS counterparts (in 2002), a result that is consistent with the literature. In particular, David Card (1999) found that IV estimate for years of schooling is generally 20-40 percent higher than OLS estimates in a survey of the literature.

Looking at the estimates for years of schooling, we see a wage increase of 9.2 percent for the population, 12 percent for males, and 7.5 percent for females. It should be noted that this is the only instance in the OLS/IV estimates where males benefit more from either forms of education than the females. Comparing these IV estimates to the OLS results from the same period, we find an increase of 77 percent for the population, 166 percent for males, and 29.3 percent for females of the IV over OLS; results that are much higher than the ones found by Card, especially for the males. As for the return to completion of university. IV estimates tell us that there are a wage increase of 49.6 percent, 27.4 percent and 69.5 percent when one completes university for the population, males, and females samples respectively; although the male estimate is not very significant, an interesting result. When comparing these estimates to the OLS ones, we

find a similarly high increase of 109 percent, 33.7 percent, and 142 percent for the three samples respectively. If we compare the magnitudes of the IV estimates for completion of university to the IV estimates for additional years of schooling in 2002 , we find that completion of university equates to the benefit of 5.37, 2.28, and 9.32 additional years of schooling (Obtained by dividing the completion of university estimates by the estimates for years of schooling). At least for the female workers, the estimate suggests that a university degree represents a significant financial benefit that is reflective of the importance the Chinese place on higher education.

For the benefit of labour market experience, the estimates in the IV model are also higher than their OLS counterparts. The population benefit for an additional year of labour market experience is 2.4 percent, up from 1.7 percent under OLS in the same period. Similarly, male workers benefit more from work experience compared to female workers.

As is expected, an income gap between the genders persist, with the IV estimate at negative 6.28 percent, up from the OLS estimate. Interestingly though, Communist Party membership, as well as the constant term, lose their significance under IV.

1995 sample

Contrary to our expectation, and the results for the 2002 sample, the IV estimates for years of education are much less significant, as well as being lower than their OLS counterparts. In particular, IV estimate for the general population is only significant at the 10 percent level, and the magnitude of the estimate is lower than the OLS one in 1995. At 2.61 percent, the IV estimate represents a decrease of 11 percent from the OLS results. At the same time, although IV estimates for the male and female workers increased from the

OSL, a loss of significance is incurred. The male estimate in particular is not even significant at the 10 percent level.

As for the estimates for completion of university, the results are more consistent with the 2002 findings. The IV estimates indicate a wage increase of 41.1 percent, 24 percent, and 53.3 percent for the population, males, and females if they complete university. Comparing the results to the 1995 OLS estimates, the IV estimates represent an increase of 172 percent, 58 percent, and 219 percent for the three samples. When compared to return to additional years of schooling estimated under IV in 1995, we find that the benefit of a university education equates to an incredible 15.75, 10.34, and 10.97 years of schooling. These results are even more incredible than the one found for the 2002 year sample.

Likewise, labour market experience also has greater estimates under IV, with female workers experiencing a near 3 percent increase in wage for each years spent working.

Discussion of persistent trends found under OLS and IV

Although the OLS estimates are subject to omitted variable bias, some qualitative aspects of the resulting patterns persist even after we control for endogeneity. In particular, the fact that females benefit more from additional education is a phenomena found in nearly all regression results, with the only exception of years of education under IV in 2002, something that is consistent with the finding in other studies. One popular explanation given for this phenomena is that education, in addition to increasing human capital and providing a signal, serves to "reduce the gap in male and female income attributable to factors such as discrimination, tastes, and circumstances (Dougherty

2005). The line of reasoning goes like this. Societal influences such as discrimination, and traditional views on the type of jobs suitable for women negatively affects the wage of females in comparison to the males. As a women becomes more educated, she will become less influenced by these factors; either by her increased willingness to find employment outside of the lower paying careers typically associated with women, or by fully taking advantage of her higher income potentials, which is enabled by her ability to pay for childcare as a result of her better job offers.

Another persistent pattern observed in our results is the significantly increased return to incomes when comparing between 1995 and 2002. This result is somewhat expected, as China has experienced tremendous economic growth within this period, and the higher return to education over time can be seen as the result of a more developed labour market.

Private vs State Owned Sector

In this section, we follow Xiu and Gunderson's analysis and look at the return to education in the state owned versus private sector. The state owned sector in China consists of the traditional government functions, as well as commercial firms owned by the Chinese government. These state owned firms allows the Chinese government to interfere directly in the market, and is a prominent feature of the Chinese economy, a product of the so called Socialism with Chinese Characteristics.

Due to their contrasting nature, and goal of firms in the two sectors, it is natural to expect a certain degree of heterogeneity to be present in the return to education among them. In particular, private sector firms seek to maximize profit, and employs the most

productive workers. Market forces and competition induce private firms to pay competitive wages to employees, whom are regularly assessed on their productivity even after being employed. On the other hand, entities in the state owned sector care less about profit maximization, rather, they focus on achieving goals deemed worthwhile by the central government. As well, workers in the state owned sector are less likely to be laid off on productivity related reasons. As a result, candidates that aspire to work in the state owned sector are more likely to be judged on educational credentials. Pons and Blanco (2005) point to the fact that education credentials are often a requirement for government jobs. Pay raises often depend on seniority and education level rather than actual productivity. On the other hand, they also suggest the idea that, if completion of key stages of education are supposed to signal productivity, something valued by the private sector, and less so by the public ones, then education credentials should have a greater effect in the private sector. Because of these contrasting ideas on just how education affects income in the different sectors, this analysis is done to give evidence and more weight to the arguments.

In Xiu and Gunderson's analysis, they found that years of education has a greater effect on incomes in the private sector, while education credentials have a greater effect in the public sector. The paper now present its own findings, based on both OLS and IV estimation of the regression model. The regression results are summarized in table 3A and 3B, which contain the estimation from the year 2002 and 1995 respectively. I will discuss the results from the year 2002 first, then compare it with the ones from 1995.

2002 sample

3A: OLS and IV Sectors - 2002

	(1) OLS Public	(2) OLS Private	(3) IV Public	(4) IV Private
Years	0.0551***	0.0537***	0.0570***	0.0729***
	(0.00364)	(0.00634)	(0.00980)	(0.0177)
Higher	0.205***	0.292***	0.536***	0.792**
	(0.0269)	(0.0771)	(0.0957)	(0.390)
exp	0.0163***	0.0110	0.0210***	0.0183*
	(0.00526)	(0.00876)	(0.00549)	(0.00991)
exp_sqr	-0.000140	-0.0000379	-0.000167	-0.0000906
	(0.000120)	(0.000203)	(0.000123)	(0.000227)
Female	-0.0934***	-0.169***	-0.0680***	-0.141***
	(0.0175)	(0.0310)	(0.0189)	(0.0340)
CCP	0.0933***	0.0961***	0.0361	0.0288
	(0.0183)	(0.0356)	(0.0221)	(0.0427)
_cons	1.123***	1.170***	0.969***	0.772***
	(0.0804)	(0.138)	(0.137)	(0.221)
N	2388	543	2388	543

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Provincial dummies not included

In contrasts to Xiu and Gunderson's findings, OLS estimates find that years of education have a greater impact on incomes in the public sector, while completion of university has a greater impact in the private sector; the reverse of their findings. In

particular, at a wage increase of 29.2 percent for completion of university in the private sector, the benefit is significantly higher than the 20.5 percent experienced in the public sector. Another unexpected result from the OLS estimation is the insignificance of years of experience in the private sector.

In contrast again to the previous OLS results, we find that after instrumenting for both years of education and university attainment, it is found that both these regressors have a greater effect on incomes in the private sector, and by a large margin.

It should also be noted that the insignificance of the years of experience squared term under both type of estimations, suggesting that incomes is only a linear function in experience in our case. The gender gap in incomes is found again under both OLS and IV, with the gap estimated being higher under OLS, and more pronounced in the Private sector, as is expected. Communist party membership is found to be insignificant under IV, contrasting with the highly significant result under OLS.

1995 sample

When we look at the results from 1995, yet another different picture emerges. Looking at the OLS results first, we see that, in contrast to the 2002 results, years of education is found to be more beneficial in the private sector, while higher education credential increases incomes more in the public sector. This result is consistent with the one found by Xiu and Gunderson, but the opposite of the OLS result for 2002. University credential is found to be insignificant in the private sector, another result that is difficult to explain. It is also evident that the return to education has increased from 1995 to 2002, a sign of the greater presence of market forces in the Chinese economy.

We now look at the IV results. Similar to the pattern for return to education found for the 2002 data set, we see that both years of education and university credential benefit workers more in the private sector. And once again, the magnitude of the benefits are lower than the 2002 counterparts, with return to years of schooling almost doubling the 1995 ones in 2002.

3B: OLS and IV Sector - 1995

	(1) OLS Public	(2) OLS Private	(3) IV Public	(4) IV Private
Years	0.0206***	0.0398***	0.0181**	0.0321***
	(0.00308)	(0.00327)	(0.00712)	(0.0101)
Higher	0.139***	0.0589	0.366***	0.641**
	(0.0215)	(0.0411)	(0.0760)	(0.261)
exp	0.0170***	0.0380***	0.0202***	0.0424***
	(0.00457)	(0.00562)	(0.00405)	(0.00537)
exp_sqr	-0.000113	-0.000588***	-0.000182**	-0.000703***
	(0.0000983)	(0.000129)	(0.0000913)	(0.000129)
Female	-0.0684***	-0.127***	-0.0561***	-0.115***
	(0.0143)	(0.0166)	(0.0147)	(0.0174)
CCP	0.0407***	0.124***	0.0239	0.106***
	(0.0145)	(0.0201)	(0.0166)	(0.0234)
_cons	0.942***	0.466***	0.903***	0.486***
	(0.0725)	(0.0794)	(0.101)	(0.120)
N	2210	1237	2210	1237

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Provincial dummies included

In contrast to 2002, we find evidence for significance of years of schooling in all our estimates for 1995, a reassuring result. The puzzling aspect of the findings, however, is that one would've guessed experience should've been more likely to be more insignificant in 1995, rather than 2002, given that the Chinese economy was more developed in the latter period. Years of experience squared is once again found insignificant in 1995. We can observe that the pattern of gender gap in incomes between the sectors is the same in 1995 as in 2002, however, the magnitude of the gap is noticeably higher in 2002, suggesting the gap widening. This result is interesting, as despite the increased return to education for females over time, and the higher percentage of female workers with university education in 2002, the income gap still managed to grow by a fair margin.

Discussion of sectorial results

In general, although OLS results found contradicting evidence for which sector harbored a greater return to education, the IV estimates both found that private sector valued it more. This suggest that, despite the formal requirement of having a certain level of education attainment to work in the public sector, the private market has developed to such an extent that both the human capital gain, as well as the signaling aspect of education translate into higher increase in incomes in the private sector of the Chinese economy.

Quantile Regression.

In this section, I depart from the analysis conducted by Xiu and Gunderson, and move on to explore an aspect of return to education not covered by their paper. Throughout my analysis so far, the paper has assumed that return to education is a constant that applied to all individuals who passed through the system. Of course, this assumption is unrealistic, as it is quite obvious that the benefit of education varies from individual to individual. Although there are many different factors that contribute to this heterogeneity, and many different manifestation of it, I focus on studying one single aspect of this issue. In particular, I will look at how return to education varies across the distribution of income. Traditional OLS models model the dependent variable at the mean of its conditional distribution, ignoring the variation that may occur at the other parts. To study this variation, I will estimate the paper's model using the quantile regression, which models the conditional distribution of the dependent variable at various percentiles.

In the sections to follow, I will estimate the model that has been used throughout the paper at six different points on the conditional distribution. Like the analysis above, I will present the results by estimating the model using the population, male, and female subsamples for both the 1995, and 2002 years. I have chosen to estimate the return to education at 6 different percentiles, these being the 10^{th}, 25^{th}, 50^{th}, 75^{th}, 90^{th} and 95^{th} percentiles.

Tables 4A, 4B, and 4C contains the estimates for the three samples in 2002, while 4D, 4E, 4F contains the estimates for the three samples in 1995. In addition, I have included line graphs of the estimates for years of education and university attainment in

graphs 4a to 4e, where the line graph with small letter corresponds to the estimates taken from the table labeled using the same letter capitalized. In each of the line graphs, the higher graph represents the return to higher education, and the line on bottom represents the return to years of schooling.

2002 sample

From the results for the population, we can see a clear downward trend in the magnitude of the estimates for years of education, with increase in wage starting at 6.1 percent at the 10^{th} percentile and gradually decreasing until it reaches 2.8 percent at the top of the distribution. Although the return to years of schooling for the two sexes are both clearly downward sloping, the magnitude of the change over the conditional distribution is larger for the female workers, with the lowest income females receiving almost 3.6 percent less increase in wage compared to the highest income ones; though, it should be pointed out that the estimates for the 95^{th} percentile are not significant for either sex. As for the credential effect captured by the completion of university, we can make out a clear increasing trend in return as we move further up in the distribution. When we separate between male and female however, two distinct pictures seem to emerge. As for the pattern of return to higher education, the estimates for males actually decrease from the 10^{th} to the 25^{th} percentile before steadily increase over the rest of the distribution. For females, a contrasting pattern is seen. The return to higher education instead increases from the 10^{th} to the 25^{th} percentile, decrease until reaching a local minimum at around the 75^{th} percentile before increasing over the right tail of the distribution. In the paper's estimates, the highest return to university for females is actually attained at the 25^{th} percentile.

4A: Quantile Population - 2002

	(1) 10th	(2) 25th	(3) 50th	(4) 75th	(5) 90th	(6) 95th
Years	0.0609***	0.0566***	0.0534***	0.0444***	0.0376***	0.0279**
	(0.0154)	(0.00894)	(0.00645)	(0.00637)	(0.0102)	(0.0126)
Higher	0.200***	0.197***	0.219***	0.224***	0.275***	0.289***
	(0.0599)	(0.0352)	(0.0269)	(0.0270)	(0.0429)	(0.0529)
exp	0.0131	0.0165**	0.0128**	0.0205***	0.0247***	0.0299**
	(0.0129)	(0.00732)	(0.00530)	(0.00515)	(0.00895)	(0.0118)
expsqr	-1.434	-1.778	-0.473	-1.942	-3.490*	-5.169*
	(2.998)	(1.698)	(1.229)	(1.190)	(2.020)	(2.643)
Female	-0.152***	-0.123***	-0.0873***	-0.0925***	-0.0700**	-0.0615*
	(0.0436)	(0.0256)	(0.0191)	(0.0189)	(0.0297)	(0.0364)
CCP	0.112**	0.108***	0.0610***	0.0296	0.0389	0.107***
	(0.0452)	(0.0266)	(0.0200)	(0.0196)	(0.0311)	(0.0383)
_cons	0.570**	0.784***	1.145***	1.460***	1.772***	2.136***
	(0.275)	(0.158)	(0.114)	(0.112)	(0.187)	(0.231)
N	2998	2998	2998	2998	2998	2998

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Provincial dummies not included

4B: Quantile Male - 2002

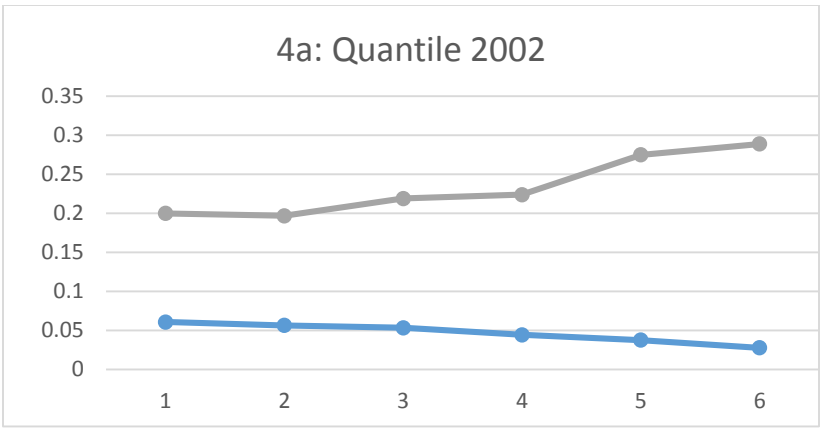
	(1) 10th	(2) 25th	(3) 50th	(4) 75th	(5) 90th	(6) 95th
Years	0.0543***	0.0493***	0.0497***	0.0376***	0.0258*	0.0240
	(0.0173)	(0.0127)	(0.00692)	(0.0109)	(0.0138)	(0.0177)
Higher	0.200***	0.169***	0.185***	0.241***	0.272***	0.286***
	(0.0670)	(0.0484)	(0.0277)	(0.0446)	(0.0565)	(0.0701)
exp	0.0136	0.0160	0.0157***	0.0164*	0.0257**	0.0343**
	(0.0156)	(0.0107)	(0.00585)	(0.00887)	(0.0126)	(0.0157)
expsqr	-2.503	-2.460	-1.866	-1.464	-3.875	-6.724**
	(3.529)	(2.393)	(1.314)	(1.982)	(2.749)	(3.355)
CCP	0.138***	0.112***	0.102***	0.0560*	0.107***	0.158***
	(0.0502)	(0.0375)	(0.0214)	(0.0330)	(0.0408)	(0.0517)
_cons	0.738**	0.954***	1.294***	1.701***	1.967***	2.247***
	(0.320)	(0.230)	(0.125)	(0.193)	(0.253)	(0.311)
N	1573	1573	1573	1573	1573	1573

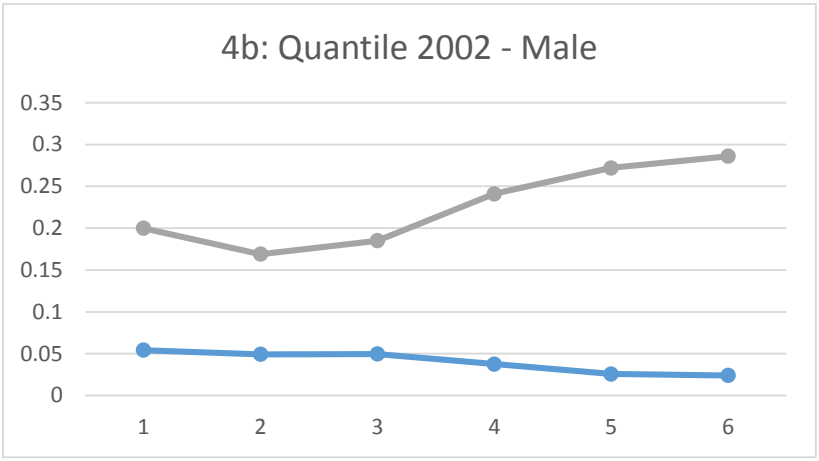
Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Provincial dummies not included

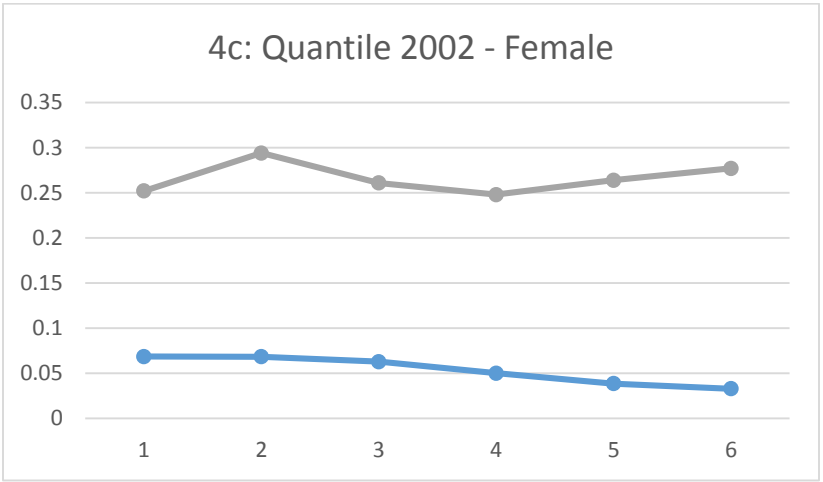
4C: Quantile Female - 2002

	(1) 10th	(2) 25th	(3) 50th	(4) 75th	(5) 90th	(6) 95th
Years	0.0686***	0.0685***	0.0629***	0.0502***	0.0387***	0.0330
	(0.0205)	(0.0128)	(0.0112)	(0.00972)	(0.0149)	(0.0227)
Higher	0.252***	0.294***	0.261***	0.248***	0.264***	0.277***
	(0.0798)	(0.0538)	(0.0496)	(0.0443)	(0.0699)	(0.0998)
exp	0.00917	0.00467	0.0106	0.0240***	0.0247*	-0.00316
	(0.0160)	(0.0102)	(0.00908)	(0.00805)	(0.0130)	(0.0214)
expsqr	1.956	2.366	0.865	-2.345	-3.043	3.346
	(3.831)	(2.432)	(2.192)	(1.957)	(3.095)	(5.074)
CCP	0.0407	0.0547	0.00327	-0.00760	-0.00199	0.0941
	(0.0579)	(0.0378)	(0.0352)	(0.0318)	(0.0489)	(0.0700)
_cons	0.122	0.559**	0.827***	1.208***	1.590***	2.095***
	(0.351)	(0.220)	(0.191)	(0.167)	(0.266)	(0.420)
N	1425	1425	1425	1425	1425	1425

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Provincial dummies not included







Looking at the other explanatory variables, we see an increasing pattern to the return to additional years of experience across the three samples, although most of the estimates are not significant for the lower percentiles. And for females in particular, nearly all the estimates show no significance at the 10 percent level. Years of experience

squared is still insignificant, consistent with previous results above. The gender gap in incomes decreases sharply as we move towards higher income workers, with the lowest income female workers earning 15 percent less than male ones.

1995 sample.

4D: Quantile Population - 1995

	(1) 10th	(2) 25th	(3) 50th	(4) 75th	(5) 90th	(6) 95th
Years	0.0391***	0.0264***	0.0247***	0.0206***	0.0173*	0.0202
	(0.00829)	(0.00579)	(0.00460)	(0.00581)	(0.00885)	(0.0128)
Higher	0.160***	0.153***	0.137***	0.157***	0.180***	0.142***
	(0.0340)	(0.0240)	(0.0196)	(0.0248)	(0.0364)	(0.0544)
exp	0.0108	0.00768	0.0133***	0.0145***	0.0211***	0.00941
	(0.00777)	(0.00500)	(0.00394)	(0.00495)	(0.00752)	(0.0107)
expsqr	-0.202	0.785	-0.602	-1.082	-2.541	-0.228
	(1.795)	(1.159)	(0.909)	(1.136)	(1.750)	(2.475)
Female	-0.0806***	-0.0697***	-0.0416***	-0.0455**	-0.0546**	-0.0532
	(0.0249)	(0.0171)	(0.0140)	(0.0178)	(0.0272)	(0.0404)
CCP	0.100***	0.0386**	0.0279*	0.0368**	0.00665	0.0321
	(0.0262)	(0.0182)	(0.0148)	(0.0187)	(0.0280)	(0.0414)
_cons	0.242	0.686***	0.916***	1.195***	1.383***	1.567***
	(0.147)	(0.0989)	(0.0784)	(0.0996)	(0.152)	(0.215)
N	3447	3447	3447	3447	3447	3447

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Provincial dummies not included

4E - Quantile Male - 1995

	(1) 10th	(2) 25th	(3) 50th	(4) 75th	(5) 90th	(6) 95th
Years	0.0218**	0.0171***	0.0114	0.0158**	0.0197*	0.0138
	(0.0109)	(0.00579)	(0.00707)	(0.00699)	(0.0118)	(0.0158)
Higher	0.134***	0.155***	0.171***	0.177***	0.132***	0.134**
	(0.0461)	(0.0232)	(0.0287)	(0.0278)	(0.0470)	(0.0623)
exp	-0.00687	0.00526	0.0146**	0.0102*	0.0154	0.0156
	(0.0104)	(0.00490)	(0.00611)	(0.00598)	(0.0110)	(0.0142)
expsqr	3.138	1.120	-1.049	0.0221	-0.933	-1.262
	(2.317)	(1.104)	(1.364)	(1.318)	(2.443)	(3.155)
CCP	0.0782**	0.0345*	0.0296	0.0122	-0.00631	-0.0470
	(0.0379)	(0.0181)	(0.0222)	(0.0215)	(0.0372)	(0.0478)
_cons	0.569***	0.830***	1.105***	1.327***	1.455***	1.711***
	(0.196)	(0.0994)	(0.123)	(0.121)	(0.219)	(0.289)
N	1955	1955	1955	1955	1955	1955

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Provincial dummies not included

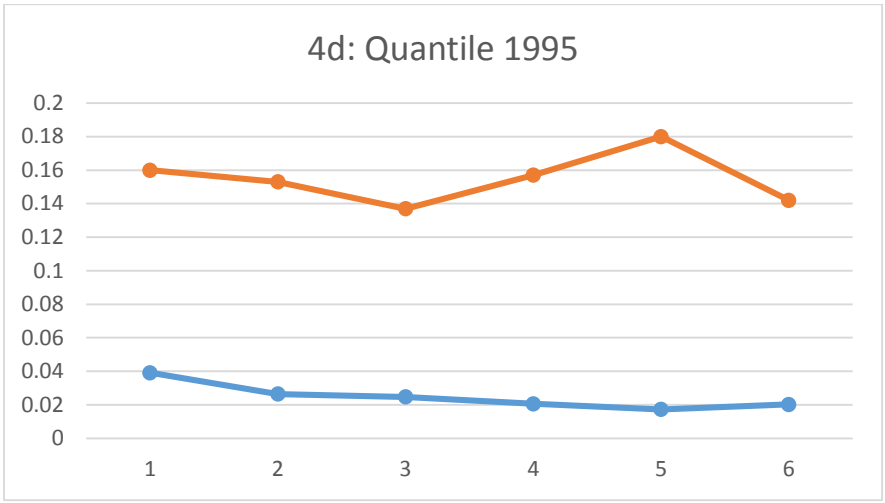
4F - Quantile Female - 1995

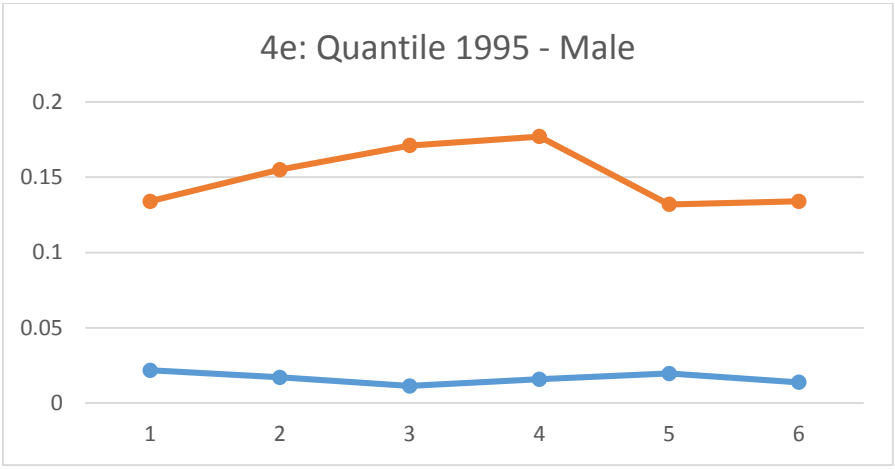
	(1) 10th	(2) 25th	(3) 50th	(4) 75th	(5) 90th	(6) 95th
P113	0.0535***	0.0531***	0.0386***	0.0444***	0.0128	0.0124
	(0.0135)	(0.0143)	(0.00759)	(0.0109)	(0.0150)	(0.0252)
Higher	0.202***	0.170***	0.133***	0.126**	0.197***	0.249**
	(0.0523)	(0.0640)	(0.0360)	(0.0519)	(0.0673)	(0.115)
exp	0.0125	0.0102	0.0160**	0.0204**	0.0189	0.00921
	(0.0114)	(0.0126)	(0.00700)	(0.00991)	(0.0131)	(0.0197)
expsqr	0.348	0.744	-1.198	-2.551	-2.575	0.00745
	(2.753)	(3.045)	(1.716)	(2.456)	(3.279)	(4.975)
CCP	0.141***	0.0414	0.0567**	0.0692*	0.0557	0.0783
	(0.0380)	(0.0466)	(0.0260)	(0.0376)	(0.0488)	(0.0817)
_cons	-0.0961	0.229	0.633***	0.741***	1.344***	1.512***
	(0.227)	(0.238)	(0.127)	(0.183)	(0.247)	(0.379)
N	1492	1492	1492	1492	1492	1492

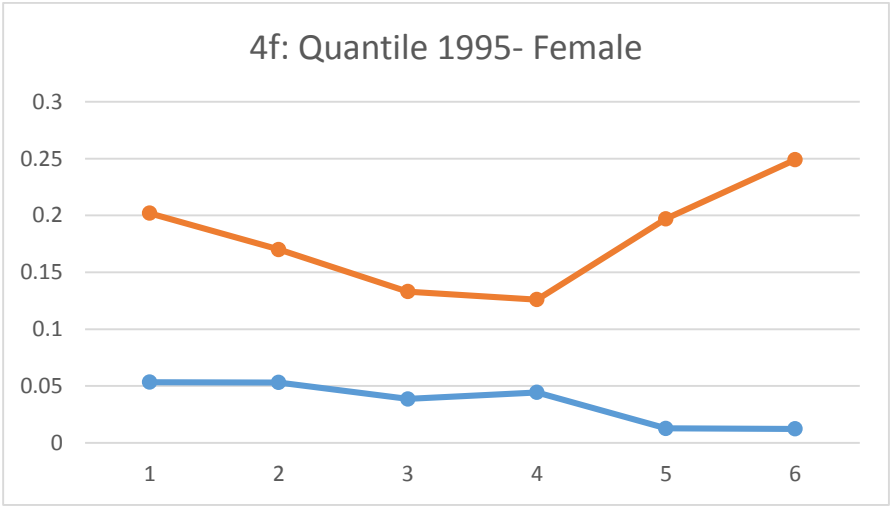
Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Provincial dummies not included

Return to years of schooling still shows a downward pattern when we look at the 1995 sample, although the slope of decrease is much flatter, resulting in a smaller difference in return to schooling between the highest and lowest earners in the population. Contrasting sharply with the case for 2002, return to higher education shows a weak downward sloping behavior, with the changes from percentile to percentile much

more radical. Return to education shows a much higher degree of heterogeneity in 1995 than it does in 2002 in the population.







When we look at the return to higher education for the two sexes, a different pattern of estimates can be seen for 1995 as well. As opposed to exhibiting a decrease in return at the lower percentiles before steadily increasing over the higher percentiles for males, the estimates instead increase over the 10^{th} to 75^{th} percentiles before dropping sharply at the right lower tail. In particular, the estimates for the 90^{th} and 95^{th} percentiles are around the same magnitude as the estimate for male workers in the 10^{th} income percentile, with the highest return estimated to be obtained for workers at the 75^{th} percentile. For females, return to university shows a clear convex pattern, with the estimates decreasing sharply over the 10^{th} to 75^{th} percentiles before rising sharply over the 90^{th} to 95^{th} percentiles.

Returns to years of experience are for the most part not significant, something that warrants further investigation. The gender gap in income again drops sharply as we move away from the lowest earners, this time however, rises back up again for the very higher earners.

Discussion and comparison

Previous studies on return to years of schooling using quantile regressions have found an upward trend amongst developed countries (Martins and Pereira 2004, Mwabu and Schultz 1996), and a downward trend in developing countries, including China (Wang 2011, Patrinos et al. 2006). In the literature, the percentiles of the conditional income distribution are interpreted as levels of individual ability (Chernozhukov et al. 2007, Arias et al. 2001). Under this interpretation, the observed downward trend in return to years of schooling can be seen as evidence for the substitutability between education

and talent, or that more able students benefit less from education on the margin, a result consistent with Ashenfelter and Rouse (1998).

To explain the downward trend observed, one can point to the distinct exam-oriented education system used in China. Because the education system trains its students on exam taking and passing tests (Li et al.2005), very few real world skills useful in the job market are taught. As a result, the more able student who succeed in the job market may have not benefited from the formal education process in the first place, which give evidence to ability and education being substitutes in China.

Another possible explanation is related to the mobility provided by education. In China, citizens are subjected to a strict household registration system known as hukou. This registration system limits where the individual may live and work. Under this system, it is possible that less able students can take advantages of the greater mobility provided by getting more education, as greater schooling may allow them to find employment in a better place (Wang 2011). Indeed, Knight and Yueh (2004) have found that education and mobility are positively correlated.

As for the shapes of return to completion of higher education, a plausible and encompassing explanation looks at the importance of education credentials in relation to the type of jobs the workers at the different percentile tend to work in. If we look at the population in general, an increasing pattern can be seen. Instead of interpreting the percentiles as levels of ability, we instead take them at face value, as the percentiles at which the workers earns wage at. In this sense, the low wage earners are likely to be graduates employed in menial jobs that doesn't necessarily require a university degree, and the higher earners are more likely to be professionals in careers that are more

technical in nature, which in most cases require higher education. This argument also applies to the observed pattern of return to university for males, although the drop at the 90^{th} and 95^{th} percentile in 1995 poses an interesting question.

For the female workers, the steady drop in return to university at left of the distribution can be perhaps attributed to the type of jobs chosen by women as they move up in the income distribution, particularly, most would often choose jobs traditionally associated with females, such as teacher, nurse, or secretary that does not necessarily need a 4 year university education, but rather, technical skills that are gained on the job or in technical schools. At the top, however, the jobs with the highest income potential are the same for women as they are for men, which are business and government positions, both of which are less affected by discrimination, as can be seen by the narrowed income gender gap estimated.

IV Quantile

A natural extension to the quantile regressions in the context of this paper is by incorporating the instrument variables. Although attempts have been made to use both of the spousal education IV, due to the limitations of the user written program ivqreg used in this paper, only one instrument can be used. In particular, the following results were obtained by estimating an IV quantile model with the regressor for university attainment instrumented by spousal university attainment. It is clear that the author could've chosen either to instrument years of schooling or university attainment. This paper has decided to go with the latter to focus on what the author believes is the more important measure of

education in China. As well, Wang (2011) used the same method to study return to years of schooling, although he did not control for university attainment in his model.

Due to the obvious problem of endogeneity that is present when we estimate a model with 2 endogenous regressors, but controlling only for one, the following results should be taken with a grain of salt and be seen as an attempted first approximation.

Like the presentation for Quantile regressions, results are found in tables 5A to 5F, and the visualized estimates to years of schooling and university attainment found in graphs 5a to 5f.

2002 Sample

One of the first thing that should be obvious to the reader is the highly questionable, and insignificance of the estimates for return to years of schooling. Although the estimates for the population and the male workers are reasonable at the 10^{th} percentile, most of the other are not even significant at the 10^{th} percent level, and worse yet, some are even significant and negative. This result is clearly the product of choosing to instrument only university attainment, and as such, estimates to years of schooling in this section of the study should be ignored.

Looking at the focus of this section, return to university attainment are estimated to be increasing at an astonishing rate over the conditional distribution, With the exception of returns slightly decreasing at the lower percentiles, as well as a sharp drop seen between the 90^{th} and 95^{th} percentiles for female workers, the exponential shaped pattern is otherwise seen consistently in all the subsamples. As can be seen in estimates, returns increase steadily over the 10^{th} to 75^{th} percentiles before exploding for very higher earners in the right tail of the conditional distribution. At around the 95^{th} percentile,

graduating from a university entails an average increase in wage in the range of 180 percent for all workers, and for the female workers income at the 90th percentile, an increase of 280 percent is estimated.

5A: IV Quantile 2002

	(1) 10th	(2) 25th	(3) 50th	(4) 75th	(5) 90th	(6) 95th
Higher	0.512***	0.701***	0.831***	0.855***	1.731***	1.832***
	(0.0459)	(0.0374)	(0.0353)	(0.0386)	(0.0653)	(0.0847)
Years	0.0215**	-0.00862	-0.0224***	-0.00943	-0.00339	-0.00229
	(0.0106)	(0.00865)	(0.00816)	(0.00893)	(0.0151)	(0.0196)
exp	0.0143	0.00879	0.0149**	0.0188**	0.0223	0.0188
	(0.00973)	(0.00793)	(0.00748)	(0.00819)	(0.0138)	(0.0179)
expsqr	-1.366	0.740	-0.568	-1.184	-2.066	-2.372
	(2.252)	(1.835)	(1.731)	(1.896)	(3.205)	(4.156)
Female	-0.123***	-0.0777***	-0.0571**	-0.0481	-0.0706	-0.0673
	(0.0355)	(0.0289)	(0.0273)	(0.0299)	(0.0505)	(0.0654)
CCP	0.109***	0.0655**	0.0582**	-0.00332	-0.0101	0.0296
	(0.0366)	(0.0298)	(0.0281)	(0.0308)	(0.0521)	(0.0676)
_cons	0.970***	1.583***	1.989***	2.118***	2.301***	2.564***
	(0.200)	(0.163)	(0.154)	(0.168)	(0.284)	(0.368)
N	2998	2998	2998	2998	2998	2998

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Provincial dummies not included

5B: IV Quantile Male - 2002

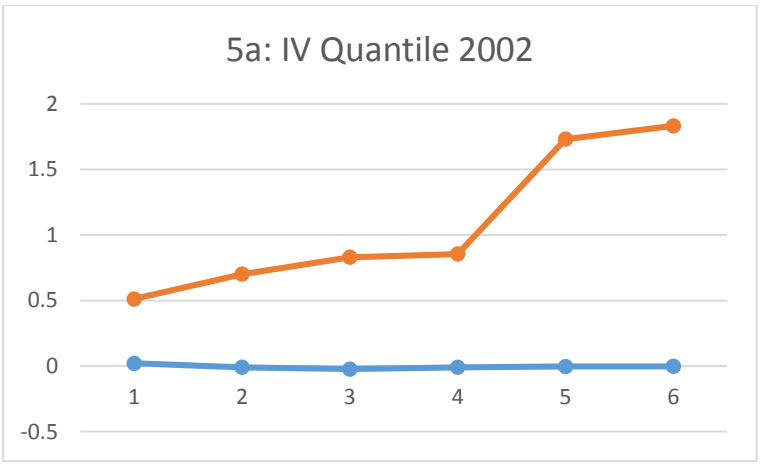
	(1) 10th	(2) 25th	(3) 50th	(4) 75th	(5) 90th	(6) 95th
Higher	0.261***	0.264***	0.619***	0.619***	0.752***	1.742***
	(0.0682)	(0.0542)	(0.0518)	(0.0565)	(0.0712)	(0.133)
Years	0.0459***	0.0332***	-0.0163	-0.000173	0.00173	-0.00443
	(0.0136)	(0.0108)	(0.0103)	(0.0113)	(0.0142)	(0.0266)
exp	0.0125	0.0145	0.0136	0.0120	0.0283**	0.0287
	(0.0132)	(0.0105)	(0.0100)	(0.0110)	(0.0138)	(0.0258)
expsqr	-2.224	-1.966	-1.289	0.237	-3.670	-5.000
	(2.961)	(2.353)	(2.249)	(2.455)	(3.091)	(5.788)
CCP	0.117**	0.105***	0.0915**	0.0252	0.0502	0.0825
	(0.0478)	(0.0380)	(0.0363)	(0.0396)	(0.0499)	(0.0935)
_cons	0.862***	1.123***	2.037***	2.173***	2.185***	2.613***
	(0.266)	(0.211)	(0.202)	(0.220)	(0.277)	(0.520)
N	1573	1573	1573	1573	1573	1573

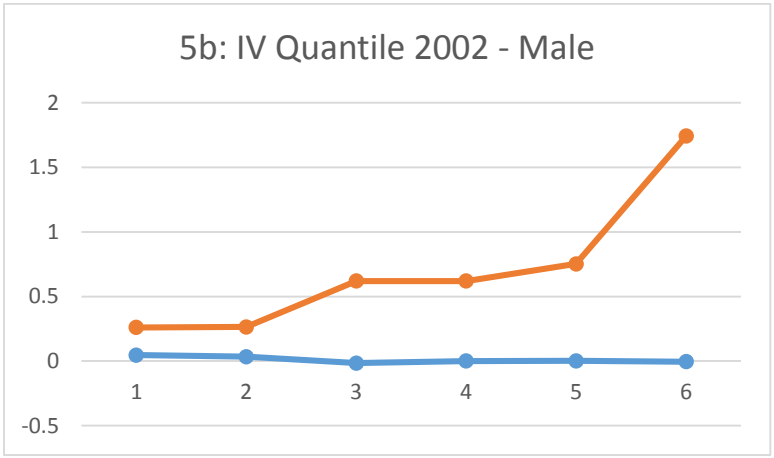
Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Provincial dummies not included

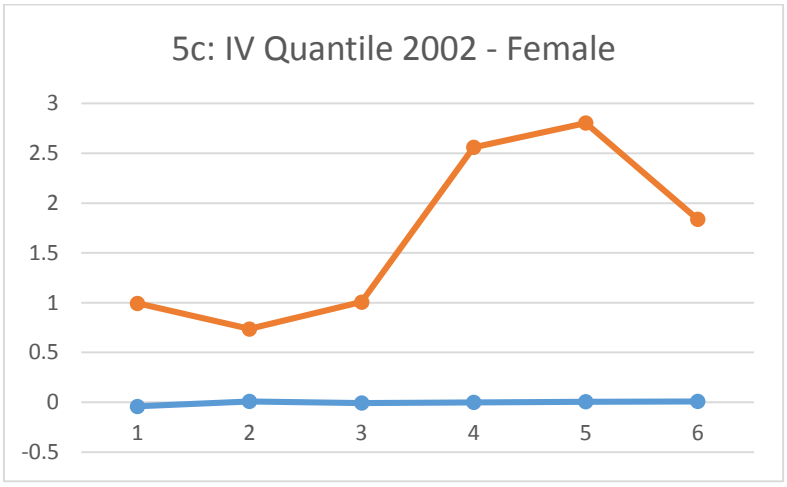
5C: IV Quantile Female - 2002

	(1) 10th	(2) 25th	(3) 50th	(4) 75th	(5) 90th	(6) 95th
Higher	0.994***	0.737***	1.006***	2.560***	2.804***	1.838***
	(0.0668)	(0.0502)	(0.0483)	(0.0856)	(0.117)	(0.106)
Years	-0.0417**	0.00930	-0.00709	-0.000886	0.00582	0.00896
	(0.0178)	(0.0134)	(0.0129)	(0.0229)	(0.0313)	(0.0284)
exp	-0.0158	0.00211	0.0110	0.0238	0.0279	0.00174
	(0.0156)	(0.0117)	(0.0113)	(0.0199)	(0.0273)	(0.0248)
expsqr	6.935*	2.795	1.770	-1.796	-3.657	1.822
	(3.756)	(2.820)	(2.717)	(4.815)	(6.588)	(5.975)
CCP	0.0462	0.0527	0.0114	-0.0654	-0.0378	-0.0149
	(0.0602)	(0.0452)	(0.0435)	(0.0771)	(0.106)	(0.0957)
_cons	1.737***	1.317***	1.671***	1.846***	1.967***	2.390***
	(0.313)	(0.235)	(0.226)	(0.401)	(0.549)	(0.497)
N	1425	1425	1425	1425	1425	1425

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Provincial dummies not included


5a: IV Quantile 2002
2
1.5
1
0.5
0
-0.5
1
2
3
4
5
6


5b: IV Quantile 2002 - Male
2
1.5
1
0.5
0
-0.5
1
2
3
4
5
6


5c: IV Quantile 2002 - Female
3
2.5
2
1.5
1
0.5
0
-0.5
1
2
3
4
5
6

1995 sample

5D: IV Quantile 1995

	(1) 10th	(2) 25th	(3) 50th	(4) 75th	(5) 90th	(6) 95th
Higher	0.384***	0.388***	0.394***	0.580***	1.716***	2.265***
	(0.0376)	(0.0298)	(0.0272)	(0.0308)	(0.0618)	(0.0967)
Years	0.0117	0.00257	-0.00116	-0.0124*	-0.0197	-0.0209
	(0.00795)	(0.00632)	(0.00575)	(0.00652)	(0.0131)	(0.0205)
exp	0.0127*	0.00932	0.0138**	0.0236***	0.0300**	0.0264
	(0.00749)	(0.00595)	(0.00542)	(0.00614)	(0.0123)	(0.0193)
expsqr	-0.516	0.444	-0.855	-3.184**	-4.294	-3.668
	(1.730)	(1.374)	(1.250)	(1.417)	(2.844)	(4.455)
Female	-0.0729***	-0.0561***	-0.0345*	-0.0255	-0.0330	-0.0678
	(0.0270)	(0.0215)	(0.0195)	(0.0221)	(0.0444)	(0.0696)
CCP	0.0751***	0.0592***	0.0316	0.0190	-0.00458	-0.0167
	(0.0279)	(0.0222)	(0.0202)	(0.0229)	(0.0459)	(0.0719)
_cons	0.509***	0.919***	1.222***	1.517***	1.724***	1.925***
	(0.143)	(0.114)	(0.104)	(0.117)	(0.235)	(0.369)
N	3447	3447	3447	3447	3447	3447

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Provincial dummies not included

5E: IV Quantile Male - 1995

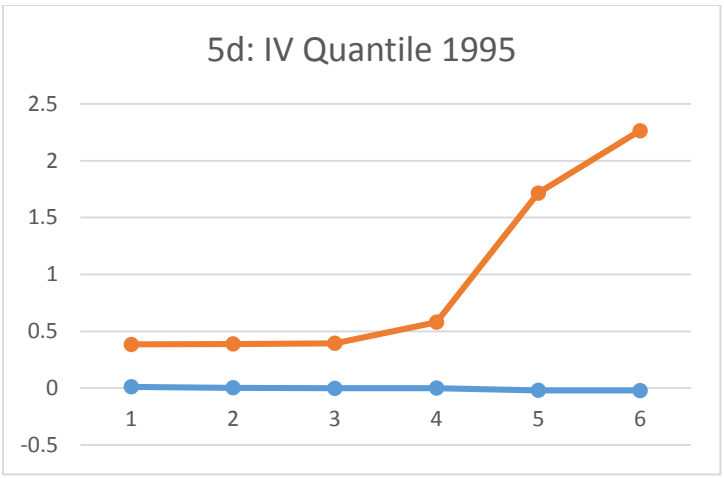
	(1) 10th	(2) 25th	(3) 50th	(4) 75th	(5) 90th	(6) 95th
Higher	0.208***	0.212***	0.270***	0.301***	0.670***	1.462***
	(0.0559)	(0.0451)	(0.0414)	(0.0453)	(0.0627)	(0.115)
Years	0.0189**	0.0105	-0.000707	-0.00367	-0.0274***	-0.0248
	(0.00948)	(0.00764)	(0.00702)	(0.00768)	(0.0106)	(0.0195)
exp	-0.00444	0.00565	0.0111*	0.00990	0.0178*	0.0249
	(0.00910)	(0.00733)	(0.00673)	(0.00737)	(0.0102)	(0.0187)
expsqr	2.626	1.150	-0.411	-0.00703	-1.604	-2.847
	(2.027)	(1.633)	(1.500)	(1.642)	(2.271)	(4.163)
CCP	0.0698**	0.0271	0.0208	0.0122	-0.0157	-0.0565
	(0.0329)	(0.0265)	(0.0243)	(0.0266)	(0.0369)	(0.0676)
_cons	0.616***	0.916***	1.287***	1.596***	2.015***	2.075***
	(0.174)	(0.140)	(0.129)	(0.141)	(0.195)	(0.357)
N	1955	1955	1955	1955	1955	1955

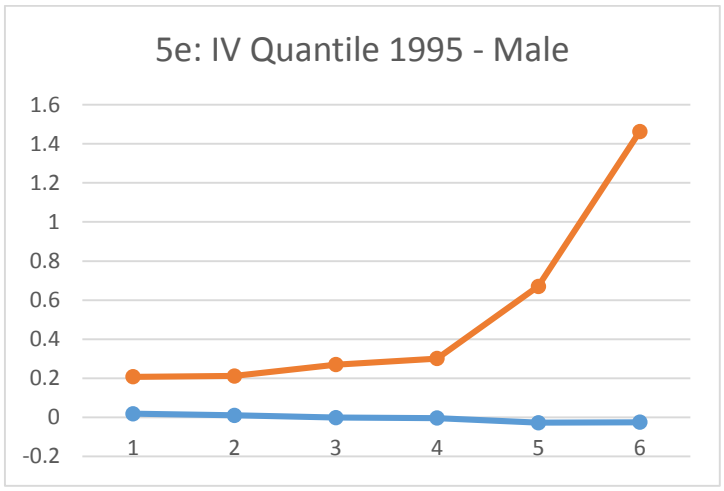
Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Provincial dummies not included

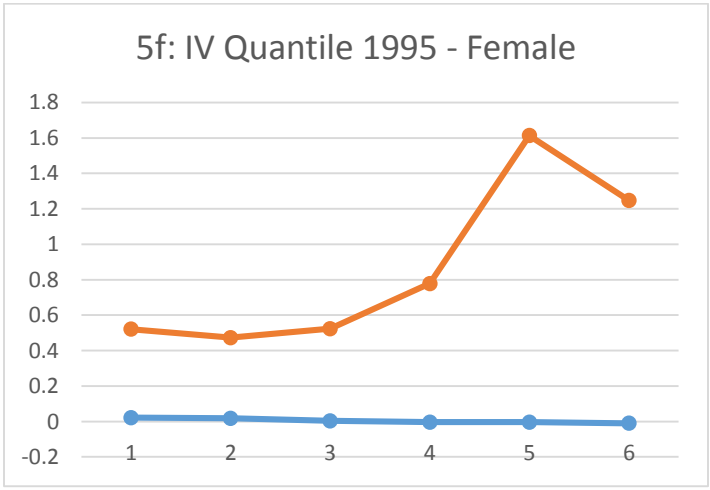
5F: IV Quantile Female - 1995

	(1) 10th	(2) 25th	(3) 50th	(4) 75th	(5) 90th	(6) 95th
Higher	0.521***	0.473***	0.524***	0.778***	1.613***	1.248***
	(0.0517)	(0.0401)	(0.0372)	(0.0431)	(0.0701)	(0.0784)
Years	0.0219	0.0180*	0.00364	-0.00396	-0.00410	-0.0102
	(0.0135)	(0.0105)	(0.00970)	(0.0113)	(0.0183)	(0.0205)
exp	0.0154	0.0127	0.0166*	0.0299***	0.0348**	0.0237
	(0.0131)	(0.0101)	(0.00938)	(0.0109)	(0.0177)	(0.0198)
expsqr	-0.317	-0.148	-1.170	-4.859*	-5.920	-2.867
	(3.209)	(2.491)	(2.306)	(2.676)	(4.352)	(4.867)
CCP	0.120**	0.0730*	0.0306	0.0363	0.0463	0.0555
	(0.0485)	(0.0376)	(0.0348)	(0.0404)	(0.0658)	(0.0735)
_cons	0.255	0.605***	1.029***	1.299***	1.372***	1.616***
	(0.232)	(0.180)	(0.167)	(0.193)	(0.314)	(0.352)
N	1492	1492	1492	1492	1492	1492

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Provincial dummies not included







Similar to the 2002 sample, the estimates for return to years of schooling are not worth paying too much attention to. The same pattern for return to university attainment

is again seen in all three subsamples for the 1995 year, with the estimates for the most parts lower than their 2002 counterparts, as is expected. The only exception is the 226 percent return to university attainment found in the general population at the 95th percentile, which is actually higher than the 95th percentile estimate in 2002. This result can be perhaps be explained by the fact that China was very early into the Economic reform in 1995, with only some selected cities and individuals given the chance to trade with foreign partners. Couple with the fact that university graduates were fewer back then, the very successful individuals were able to command a substantially higher income than their peers.

Although the IV quantile results in this section probably suffer heavily from endogeneity of the years of schooling regressor, the incredible rate of return to university estimates should nevertheless be taken as an indication of the importance of a university degree in China's labour market.

Conclusion

By restricting the number of credential regressors, this paper was able to control for endogeneity of measures of education while narrowing its focus to the return to higher education attainment, an achievement revered throughout Chinese history. With the use of instrumental variables, this study was able to find higher estimates for return to education than using OLS alone, a result that is consistent with the literature. In particular, the paper finds the IV estimates for return to higher education attainment, after controlling for years of education, are orders of magnitudes higher than the return to an additional year of schooling, a result that confirms the importance of university education

in Chinese society. As well, return to education are estimated to have a greater impact on incomes earned for female than male workers, a result that suggests giving females more access to education as a mean to counteract the observed income gender gap.

Also by controlling for endogeneity, this paper was able to find evidence for greater returns to both years of schooling and university attainment in the private sector, a result that contrasts the findings of Xiu and Gunderson.

The quantile regression results found a decreasing pattern for return to years of schooling, something that is consistent with other quantile findings for developing countries. Return to university showed a weakly increasing pattern in the population, while 2 contrasting patterns for the males and females subsamples.

Lastly, the IV quantile estimates shows an exponential growth pattern that is associated with the return to university as we move to the right of the conditional income distribution. The richest were found to benefit greatly from university attainment.

Due to the choice of instruments used in this paper, the samples were restricted to married couples in the data sets. Even though the dataset is fairly recent, the average individual is well into their mid age in the study. Combined with the fact that the ban National Higher Education Entrance Examination was only lifted some 30 years ago, the results in this study cannot be applied to the young people of today. With the rapid pace of development occurring in China, it is no doubt that the magnitude and pattern of return to education in China will be vastly different today between someone who has long left the classroom, and a recent graduate. In order to examine casual effect education has on income that are relevant for policy makers, young workers, and students today, new identification strategies will have to be used than the one found in this paper.

Bibliography

Antelius J. (2000) Sheepskin Effects in the Returns to Education: Evidence on Swedish Data' Working Paper Series 158, Trade Union Institute for Economic Research.

Arabsheibani G. R. and Manfor L. (2001) 'Non-linearities in Returns to Education in Libya', *Education Economics* 9(2): 139–144.

Arabsheibani, G.R. and A. Mussurov. 2007. "Returns to Schooing in Kazakhstan." *Economics of Transition* 15:341–364.

Arias, O., K. F. Hallock, and W. Sosa-Escudero. 2001. "Individual heterogeneity in the returns to schooling: Instrumental variables quantile regression using twins data." *Empirical Economics* 26:7–40.

Ashenfelter, O. and C. Rouse. 1998. "Income, Schooling, and Ability: Evidence from a New Sample Identical Twins." *Quarterly Journal of Economics* 113:253–284.

Barbieri-Low, Anthony J. 2007. *Artisans in Early Imperial China*. Seattle & London: University of Washington Press.

Bauer T. K., Dross P. J. and Haisken-DeNew J. P. (2005) 'Sheepskin Effects in Japan', *International Journal of Manpower* 26(4): 320–335.

Becker, Gary S. (1962) "Investment in Human Capital: A Theoretical Analysis" *Journal of Political Economy*, 70(5) Part 2: *Investment in Human Beings*, 9-49.

Becker, G.S. 1981. *A Treatise on the Family*. Cambridge: Harvard University Press

Bian, Yanjie, Xiaoling Shu, and John R. Logan. "Communist Party membership and regime dynamics in China." *Social Forces* 79, no. 3 (2001): 805-841.

Card, D. 1999. Handbook of Labor Economics, volume 3, chapter The causal effect of education on earnings. Amsterdam: North Holland.

Card D. (2001) 'Estimating the Return to Schooling: Progress on Some Persistent Econometric
Problems', *Econometrica* 69(5): 1127–1160.

Chen G. and Hamori S. (2009) 'Economic Returns to Schooling in Urban China: OLS and the
Instrumental Variables Approach', *China Economic Review* 20(2): 143–152

Chen, Y. and S. Feng. 2009. "Parental Education and Wages: Evidence from China." *IZA Discussion Paper No. 4218.*

Chernozhukov, V., C. Hansen, and M. Jansson. 2007. “Inference approaches for instrumental variable quantile regression.” *Economics Letters* 95:272–277.

Chong, H., H. Li, P.W. Liu, and J. Zhang. 2009. “Why Does Spousal Education Matter for Earnings? Assortative Mating and Cross-productivity.” *Journal of Labor Economics* 27:633–652.

Dougherty, Christopher. "Why are the returns to schooling higher for women than for men?" *Journal of Human Resources* 40, no. 4 (2005): 969-988.

Elman, Benjamin. 2002. *A Cultural History of Civil Examinations in Late Imperial China*. London: Univ. of California Press.

Ferrer A. M. and Riddell W. C. (2002) ‘The Role of Credentials in the Canadian Labour Market’ *Canadian Journal of Economics* 35(4): 879–905.

Fleisher, Belton M., and Jian Chen. 1997. “The Coast-Noncoast Income Gap, Productivity, and Regional Economic Policy in China.” *Journal of Comparative Economics* 25 (2): 220-236.

Fleisher B., Li H. and Zhao M. Q. (2010) ‘Human Capital, Economic Growth, and Regional Inequality in China’, *Journal of Development Economics* 92(2): 215–231.

Flores-Lagunes, Alfonso, and Audrey Light. 2007. "Interpreting Sheepskin Effects in the Returns to Education." *Education Research Section*. Princeton: Princeton University.
Flores-Lagunes A. and Light A. (2010) ‘Interpreting Degree Effects in the Returns to Education’,
Journal of Human Resources 45(2): 439–467.

Gibson J. (2000) ‘Sheepskin Effects and the Returns to Education in New Zealand: Do They Differ by Ethnic Groups’, *New Zealand Economic Papers* 34(2): 201–220.

Gustafsson B. and Li S. (2000) ‘Economic Transformation and the Gender Earnings Gap in Urban China’, *Journal of Population Economics* 13(2): 305–329.

Hausmann, Ricardo, Saadia Zahidi, Laura Tyson, Ricardo Hausmann, Klaus Schwab, and Laura D'Andrea Tyson. "The global gender gap report 2009." World Economic Forum, 2009.

Heckman J. and Li X. (2004) ‘Selection Bias, Comparative Advantage and Heterogeneous Returns to Education: Evidence from China in 2000’, *Pacific Economic Review* 9(3): 155–171.

Hsieh, Pei-Tseng Jenny. 2013. *Education in East Asia*. Bloomsbury Academic.

Hungerford, Thomas, and Gary Solon. 1987. "Sheepskin Effects in the Returns to Education." *Review of Economics and Statistics* 69: 175-77.

Jaeger D. A. and Page M. E. (1996) 'Degrees Matter: New Evidence on Sheepskin Effects in the Returns to Education', *The Review of Economics and Statistics* 78(4): 733–740.
Johnson E. N. and Chow G. C. (1997) 'Rates of Return to Schooling in China', *Pacific Economic Review* 2(2): 101–113.

Knight, J. and L. Yueh. 2004. "Job Mobility of Residents and Migrants in Urban China." *Journal of Comparative Economics* 32:637–660.

Knight J. and Song L. (2003) 'Increasing Urban Wage Inequality in China', *Economics of Transition* 11(4): 597–619.

Li H. (2003) 'Economic Transition and Returns to Education in China', *Economics of Education Review* 22(3): 317–328.

Li, H., P.W. Liu, N. Ma, and J. Zhang. 2005. "Does Education Pay in Urban China? Estimating Returns to Education Using Twins." *CUHK Discussion Papers # 00013* .

Li H. and Zax J. (2000) 'Economic Transition and Labor Supply in Urban China' in Wang Y. and Chen A. (eds.) *China's Labor Market and Problems of Employment,* Chengdu, China: Southwest University of Economics and Finance Press: 217–233 (in Chinese).

Liu, Elaine, and Shu Zhang. 2013. "A Meta-Analysis of the Estimates of Return to Schooling in China." Unpulished.

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(3): 542–564.

Mare, R. 1991. "Five Decades of Educational Assortative Mating." American Sociological Review 56:15–32.

Martins, P.S. and P.T. Pereira. 2004. "Does Education Reduce Wage Inequality? Quantile Regression Evidence From 16 Countries." *Labour Economics* 11:355–371.

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(1): 17–40.

Meng, X. 2000. *Labour market reform in china.* Cambridge Univ. Press.

Michael Spence (1973). "Job Market Signaling". *Quarterly Journal of Economics* (The Quarterly Journal of Economics, Vol. 87, No. 3) **87** (3): 355–374.

Millimet, D.L. and L. Wang. 2006. "A Distributional Analysis of the Gender Earnings Gap in Urban China." *B.E. Journal of Economic Analysis & Policy* 5: Article 5.

Mincer, Jacob (1974) *Schooling, Experience and Earnings*, New York: National Bureau of Economic Research.

Mora J. J. and Muro J. (2008) 'Sheepskin Effects by Cohorts in Colombia', *International Journal of Manpower* 29(2): 111–121.

Mullins, Richard. 2005. "Test Fever: what the gaokao means to today's students, teachers and parents." *China Today* 54 (6): 34-36.

Mwabu, G. and T.P. Schultz. 1996. "Education Returns Across Quantiles of the Wage Function: Alternative Explanations for Returns to Education by Race in South Africa." *American Economic Review* 86:335–339.

Patrinos H. A. 1996. 'Non-linearities in the Returns to Education: Sheepskin Effects or Threshold Levels of Human Capital', *Applied Economics Letters* 3(3): 171–173.

Patrinos, H.A., C. Ridao-Cano, and C. Sakellariou. 2006. "Estimating the Returns to Education: Accounting for Heteroneity in Ability." *World Bank Policy Research Working Paper 4040* .

Pons E. and Blanco J. M. (2005) 'Sheepskin Effects in the Spanish Labour Market: A Public–Private Sector Analysis', *Education Economics* 13(3): 331–347.

Qian, Z. 1998. "Changes in Assortative Mating: The Impact of Age and Education." *Demography* 35:279–292.

Schady N. R. (2003) 'Convexity and Sheepskin Effects in the Human Capital Earnings Function: Recent Evidence for Filipino Men', *Oxford Bulletin of Economics and Statistics* 65(2): 171–196.

Trostel, P., I. Walker, and P. Woolley. 2002. "Estimates of the economic return to schooling for 28 countries." *Labour Economics* 9:1–16.

Wang, L. 2010. "Does Spousal Education Matter for Earnings?" *Unpublished*.

Wang, Li. 2013. “How does Education Affect the Earnings Distribution in Urban China” *Oxford Bulletin of Economics and Statistics* 75: 435–454

Xiu, Lin, and Morley Gunderson. 2013. “Credential Effects and the Return to Education in China.” *Labour* 27 (2): 225-248.

Yang D. T. (2005) ‘Determinants of Schooling Returns during Transition: Evidence from Chinese Cities’, *Journal of Comparative Economics* 33(2): 244–264.

Zhang J., Zhao Y., Park A. and Song X. (2005) ‘Economic Returns to Schooling in Urban China, 1988 to 2001’, *Journal of Comparative Economics* 33(4): 730–752.

Appendix

First stage result - 1995

	(1) Years	(2) Higher
SpouseEdu	0.471***	0.0105***
	(0.0103)	(0.00109)
SpouseHigher	0.433***	0.296***
	(0.107)	(0.0113)
exp	-0.0972***	-0.00964***
	(0.0156)	(0.00165)
exp_sqr	0.00162***	0.000209***
	(0.000361)	(0.0000382)
Female	-1.001***	-0.0784***
	(0.0552)	(0.00584)
CCP	1.218***	0.0721***
	(0.0623)	(0.00659)
_cons	7.595***	0.0671***
	(0.234)	(0.0247)
Cragg-Donald Wald F statistic	**328.389**	
(Kleibergen-Paap rk Wald F statistic	**140.825**	
N	8238	8238

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

First stage result - 2002

	(1) Years	(2) Higher
SpouseEdu	0.474***	0.0135***
	(0.0175)	(0.00150)
SpouseHigher	0.429***	0.310***
	(0.116)	(0.0202)
exp	-0.0361*	-0.00490**
	(0.0211)	(0.00240)
exp_sqr	-0.000397	-0.0000172
	(0.000495)	(0.0000555)
Female	-0.530***	-0.0994***
	(0.0651)	(0.00800)
CCP	1.376***	0.120***
	(0.0727)	(0.00982)
_cons	7.366***	0.0493
	(0.349)	(0.0336)
Cragg-Donald Wald F statistic	**227.514**	
Kleibergen-Paap rk Wald F statistic	**130.660**	
N	5340	5340

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Overidentification tests - 2002

	(1) Years Instrumented	(2) Higher Instrumented
Years	0.0956***	0.0348***
	(0.00692)	(0.00550)
Higher	0.0983***	0.757***
	(0.0346)	(0.0948)
exp	0.0206***	0.0207***
	(0.00457)	(0.00464)
exp_sqr	-0.000163	-0.000162
	(0.000104)	(0.000106)
Female	-0.122***	-0.0939***
	(0.0155)	(0.0167)
CCP	0.0586***	0.0691***
	(0.0182)	(0.0180)
cons	0.564***	1.219***
	(0.107)	(0.0789)
Hansen J statistic	**34.076**	**28.490**
N	5340	5340

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Overidentification tests - 1995

	(1) Years Instrumented	(2) Higher Instrumented
Years	0.0536***	0.0229***
	(0.00472)	(0.00379)
Higher	0.0690***	0.563***
	(0.0257)	(0.0812)
exp	0.0260***	0.0279***
	(0.00370)	(0.00379)
exp_sqr	-0.000280***	-0.000347***
	(0.0000814)	(0.0000845)
Female	-0.101***	-0.0953***
	(0.0114)	(0.0116)
CCP	0.0801***	0.0856***
	(0.0131)	(0.0127)
_cons	0.419***	0.720***
	(0.0779)	(0.0596)
Hansen J statistic	**35.688**	**9.007**
N	8238	8238

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$