IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14139

“Study Hard and Make Progress Every Day”: Updates on Returns to Education in China

Jie Chen
Francesco Pastore

FEBRUARY 2021

I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14139

"Study Hard and Make Progress Every Day": Updates on Returns to Education in China

Jie Chen
UNSW Sydney

Francesco Pastore
University of Campania "Luigi Vanvitelli" and IZA

FEBRUARY 2021

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

"Study Hard and Make Progress Every Day": Updates on Returns to Education in China

In this paper, we apply Generalized Propensity Score Matching (GPSM) method, which deals with a continuous treatment variable, to estimate the returns to education in China from 2010 to 2017. Results are compared with OLS estimates from the classical Mincerian equation, as well as estimates from two instrumental variable methods (i.e., 2SLS and Lewbel). We use the Chinese General Social Survey data, including a subset newly released in 2020. We find that returns to education in China experienced a slight decrease in 2010-2015, but reverted back in 2017. With the more flexible GPSM method, we also find that returns to university education remain higher than returns to secondary or compulsory education. The GPSM estimates are also closer to OLS estimates, compared to both instrumental variable methods.

JEL Classification: I26, J30

Keywords: returns to education, endogeneity, continuous treatment, sample selection, GPSM, IV, Lewbel, China

Corresponding author:
Francesco Pastore
University of Campania "Luigi Vanvitelli"
Corso Gran Priorato di Malta
I-81043, Capua (Caserta)
Italy

E-mail: francesco.pastore@unicampania.it

1 Introduction

Returns to education is a topic of longstanding research interest (Card, 1999). In 2020, university graduates in China totaled 8.74 million, setting a new record. Employability of young adults is a key issue facing the Chinese government, especially given the economic slowdown due to COVID-19. Understanding the most recent changes to returns of education can be of great interest for policy makers at all levels. In fact, it allows us to understand the extent to which the demand and supply of skills proceed at the same pace. We use the most recent wave of the Chinese General Social Survey released in 2020 and collected in 2017.

In 2015, China's president Xi brought up a term called "supply-side reform", emphasizing on the importance of improving the quality of production and optimizing its efficiency. In the same year, China's premier Li Keqiang, announced the "Made in China 2025" strategy in his Government Work Report. This strategy marks a new era of the Chinese manufacturing industry. Previously, "Made in China" tends to be associated with cheap and low-quality goods. With the supply-side reform and "Made in China 2025", the Chinese government aims to upgrade its manufacturing infrastructure, especially technology-driven industries, so that products from China can be more technology-intensive and hence more competitive in the global market. The introduction of these two visions has seen an increasing demand of high-quality skilled workers. Consequently, "shortage of skilled workers" coexists with "oversupply of university graduates".[1]. Given this background, understanding how the returns to education have changed over the recent years can benefit the policy makers as well as candidates for different education institutions.

Our paper provides several additions to the existing literature. The first consists of using the most recent data available. Moreover, we adopt a new set of instrumental variables (i.e., enrolment ratio of upper secondary entrants relative to lower secondary graduates, whether a subject's parent passed away when (s)he was 14 years old). The enrolment ratios predict the probability that an individual continues his or her education, but are unlikely to directly relate to one's earnings in the job market. One restriction of the enrolment ratios, though, is that they can only be applied locally to part of the dataset. Thus, we add the third instrumental variable, as it has been conveniently used in the existing literature to instrument one's education (Case et al., 2004; Asadullah and Xiao, 2020). Third, in addition to the more conventional OLS and 2SLS estimates, we adopt Lewbel method - a heteroskedasticity-based estimating method for linear regressions with an endogenous regressor - to control for endogeneity bias as an alternative and a robustness check. Fourth, we use Generalized Propensity Score Matching (GPSM) given our continuous treatment (i.e., years of education) and compare its estimates with estimates from the other methods, aiming to gauge its applicability in the economics of education. Lastly, we control for a larger number of covariates than in previous studies (e.g., parents' years of education, English skills, party membership, union membership, and health capital variables).

We find that returns to education in China remain relatively stable in 2010-2017, with a slight decrease in the years 2010-2015 but a strong reversion in 2017. We also find that

[1] http://opinion.people.com.cn/n/2014/0514/c1003-25017019.html

university graduates keep receiving higher returns than those who do not receive any higher education. In other words, it is still very convenient to invest in education in China. No matter how the supply is increasing, apparently the demand for skills is increasing at a quicker pace, although maybe less than expected. The huge increase in supply recently experienced in China, as well as in other countries, might contribute or is already contributing to a technological boost of the country's production system, although the effects are still not there. With the Chinese government's increasing support on technology-intensive industries, university graduates (especially those major in STEM areas) are expected to keep receiving high returns on education.

The paper is structured as follows. In Section 2, we review the existing literature on returns to education in China and motivate shortly the paper by accounting for the recent evolution of the market for skills in China. The ensuing two sections present data, methodology, and findings. We conclude in Section 5.

2 Literature review

Only a few studies measure the returns to education in China, one of the leading economies in the world. As can be seen from the literature summary in Table 1, studies on this topic almost paused in the recent decade. This is strange, since the evolution of the market for human capital in China is key to understanding the direction of the dramatic and continuous process of structural change the Chinese economy is undergoing (Goldin and Katz, 2007; Mason, 1996).

The earliest study exploring returns to education in China was by Jamison and Van der Gaag (1987). The authors used survey data from a one-off project in Gansu province and found that the return rate was around 5% in 1980s. Several other studies emerged in the 1990s and early 2000s and used Mincer's equation to estimate the return rate (Byron and Manaloto, 1990; Johnson and Chow, 1997; Meng and Kidd, 1997). By and large, they found the return rate to be below 5% . Over time, the returns to education seem to have increased, a phenomenon that has been documented by the existing literature (Ren and Miller, 2012). For example, Yang (2005) compared data from 1988 to those from 2001 and found the returns in the latter year to be higher than those in the earlier year. Same patterns appear in Wang (2013) and Zhang et al. (2005). The increasing returns to education is intuitive, given that China's economy has been developing quickly over the past 40 years. Interestingly, Asadullah and Xiao (2020) found the returns to education decreased between year 2010 and 2015. The authors attributed this decline to the upsurge of educated workers due to the higher education expansion since 1999.

Table 1: A brief summary of literature on returns to education in China

Paper	Returns to education(%)	Methodology	Data	Location
Jamison and Van der Gaag (1987)	4.5 (urban males); 5.5 (urban females)	Mincer's equation	A one-off survey among 2154 individuals (1985)	Hui county, Gansu Province
Byron and Manaloto (1990)	around 1.4 - 4	Mincer's equation	A one-off survey among 800 adults (1986)	Nanjing (city)
Johnson and Chow (1997)	4.02 - rural; 3.29 - urban	Mincer's equation	1988 Chinese Household Income Project	National (18938 obs)
Meng and Kidd (1997)	2.5 - 1981; 2.7 - 1987	Mincer's equation	A one-off survey by the Institute of Quantitative Economics of the Chinese Academy of Social Sciences (1989)	National (50900 obs)
Tao Yang (1997)	2.3	Mincer's equation	A one-off survey among 204 farmers (1990)	Sichuan Province
Yang (2005)	3.1 (1988); 5.1 (1995)	Mincer's equation	Chinese Household Income Project (CHIP)	National (26101 obs)
Zhang et al. (2005)	4.0-10.2 from 1988 to 2001	Mincer's equation	14 consecutive annual surveys carried out by the Urban Survey Organization (USO) of the National Bureau of Statistics	Six provinces , annual obs range from 5404 to 7853
Wang (2013)	3.6 (1995); 6.6 (2002)	Mincer's equation	CHIP 1995, 2002	National
	4.4 (1995); 8.8 (2002)	IV method (Spousal education)		
Asadullah and Xiao (2020)	6.7-7.5 (2010); 6.2-6.9 (2015)	Mincer's equation	China General Social Survey (CGSS) 2010, 2015	National (6278 obs)
	21.4 (2010); 16.4 (2015)	IV method (parents' education, whether a subject lost parent at the age of 14)		
	11.6 (2010); 8.8 (2015)	Lewbel's method		

Notes: Extensive literature on returns to education in China can be found in a recent meta-analysis by Churchill and Mishra (2018).

The most recent contribution and the most similar to ours on returns to education in China is Asadullah and Xiao (2020). They use the same data as ours, but only with reference to the 2010 and 2015 waves. In addition to classic OLS and 2SLS estimates, they also provide Lewbel estimates of endogeneity-corrected returns to education. They find a general, uncorrected rate of return of 9.7% for 2010 (9.2% for 2015). After correcting for endogeneity bias, they find 21.4% for 2010 (16.4% for 2015) with 2SLS and 11.6% for 2010 (8.8% for 2015) with Lewbel. Asadullah and Xiao (2020) conclude that the returns to education have declined between 2011 and 2015 and attribute the decline to 'an expanded supply of educated workers and diminishing returns to human capital'.

So far, very few have applied the GPSM method in estimating the returns to education (Uysal, 2015) and no previous study did it for the case of China. Using a unique dataset from Britain, Uysal (2015) analyzes the returns to education for females and males separately. The author uses a doubly-robust procedure, combining the GPS method with a weighting method. The author finds that, compared to no qualification, higher education brings significant wage premiums for both genders, although lower level of education does not seem to bring significant premiums for females. To the best of our knowledge, we are the first to compare the GPSM estimates with the Lewbel estimates.

3 Data description

Data are pooled from six CGSS annual surveys: 2010, 2011, 2012, 2013, 2015, and 2017. We first drop observations without wage income, then we merge all six datasets together. Next, we summarize the hourly wage and drop potential outliers (those with more than three standard deviations away from the mean). Lastly, we follow Asadullah and Xiao (2020) and keep only working age individuals (i.e. women aged 18-55 and men aged 18-60). We are left with 24,832 observations in total. Among these remaining 24,832 observations, around 24.4% are estimated to have completed higher education or above (i.e., vocational college, university, or postgraduate). A summary of the main variables of interest is presented in Table 2.

Table 2: Summary statistics of control variables for waged workers, 2010-2017

Variable	Mean	SD	Max.	Min.	N
Panel A: baseline control variables					
Hourly wage	16.200	26.765	400.641	0.003	24832
Ln (hourly wage)	2.109	1.189	5.993	-5.966	24832
Years of education	10.373	4.032	19	0	24821
Years of experience	26.881	11.212	54	0	24821
Age	43.253	9.341	60	18	24832
Female (yes=1)	0.418	0.493	1	0	24832
Ethnicity (minority=1)	0.087	0.282	1	0	24832
Panel B: Additional control variables					
hukou (agricultural=1)	0.593	0.491	1	0	24832
Marital status					
Single	0.117	0.322	1	0	24832
De facto	0.009	0.096	1	0	24832
Married	0.818	0.386	1	0	24832
Re-married	0.015	0.123	1	0	24832
Separated	0.007	0.082	1	0	24832
Divorced	0.022	0.148	1	0	24832
Widowed	0.011	0.103	1	0	24832
Sectors of industry					
Agricultural job (base group)	0.306	0.461	1	0	18678
State owned enterprise	0.211	0.408	1	0	18678
Collectively owned enterprise	0.053	0.224	1	0	18678
Privately owned enterprise	0.405	0.491	1	0	18678
Hong Kong, Macau or Taiwan funded enterprise	0.005	0.072	1	0	18678
Foreign funded enterprise	0.020	0.139	1	0	18678
Location					
Eastern (base group)	0.384	0.486	1	0	24832
Central	0.223	0.417	1	0	24832
Western	0.263	0.440	1	0	24832
Northeast	0.130	0.336	1	0	24832
Father's years of education	4.288	4.446	19	0	24298
Mother's years of education	6.060	4.559	19	0	24069
Good English skill (at/above the average level=1)	0.144	0.351	1	0	22605
Union membership (yes=1)	0.152	0.359	1	0	24628
Party membership (yes=1)	0.113	0.317	1	0	24761
Health capital					
Height (in *cm*)	166.229	7.779	196	64	24798
Health capital					
Physical health - normal (base group)	0.193	0.394	1	0	22597
Physical health - poor	0.086	0.280	1	0	22597
Physical health - good	0.722	0.448	1	0	22597
BMI					
Normal (18.5 $\leqslant$BMI$<$ 25, base group)	0.706	0.456	1	0	24781
Underweight (BMI$<$ 18.5)	0.070	0.256	1	0	24781
Overweight (25 $\leqslant$BMI$<$ 30)	0.199	0.399	1	0	24781
Obese (BMI$\geqslant$ 30))	0.026	0.158	1	0	24781
Instruments					
Eratio$_{vus}$	0.236	0.090	0.490	0.044	6471
Eratio$_{aus}$	0.343	0.129	0.678	0.137	6471
Eratio$_{auni}$	0.555	0.395	2.204	0.198	2391
Eratio$_{vuni}$	0.452	0.142	0.972	0.152	2391
Lost parent at age 14	0.077	0.266	1	0	24501

Notes: the data include waged workers from the Chinese General Social Survey in 2010, 2011, 2012, 2013, 2015, and 2017. 'Years of experience' equals age minus the sum of years of education and 6 (the legal school entry age in China).

The *Years of education* are calculated based on a mapping of the Chinese education system onto the International Standard Classification of Education.[2] The exact mapping is as follows: compulsory education is equivalent to 9 years of education; upper secondary education (both vocational and academic) is equivalent to 12 years of education; vocational college education is equivalent to 15 years of education; academic university education is equivalent to 16 years of education; master's education is equivalent to 19 years of education; and doctoral education is equivalent to 22 years of education. In CGSS, master's degree and doctoral degree are not differentiated and coded jointly as "postgraduate degree and above". Given that doctoral graduates are only around 10% of master's graduates in China, we decide to map "postgraduate degree and above" to master's degree (i.e., 19 years of education) for convenience. Existing literature conventionally assign 16 years to 'college and above' level of education Li (2003). However, the conventional specification could very likely lead to overestimation of returns to education. Our mapping can alleviate the overestimation problem.

For the *Years of experience*, we use the so-called potential work experience (i.e., age - years of education - 6). We don't know when people actually started to work to compute the actual work experience. Noting this fact is very important especially for women, since women are more likely to leave the job market temporarily for parenthood and take part-time jobs (Meara et al., 2020). *Sectors of industry* represents the sector in which one works. *Location* corresponds to the province in which one completed the survey.[3] As in Asadullah and Xiao (2020), our health capital variables include height, one's perceived physical health status, and one's Body Mass Index.

4 Methodology

As a starting point, we first estimate a Mincerian education production function as follows.

$$Y_i = \alpha + \beta edu_i + \gamma X_i + \epsilon_i$$

where Y is the natural logarithm of hourly working income. edu_i is the years of education corresponding to different education qualifications in China. X_i is a list of control variables which differ between the baseline and the full specifications. In the baseline, X includes years of experience, experience-squared, gender, ethnicity, and yearly fixed effects where applicable. In the full specification, we additionally control for marital status, *hukou* status[4], sectors of industry, provincial fixed effects.

Additionally, a list of not-so-conventional covariates are added in response to findings

[2] The mapping is prepared by the OECD, accessible via this link.

[3] In total, we have data from 31 provincial-level divisions: Anhui, Fujian, Guangdong, Guizhou, Hainan, Hebei, Henan, Hubei, Hunan, Gansu, Jiangxi, Jiangsu, Qinghai, Shaanxi, Shandong, Shanxi, Sichuan, Yunnan, Zhejiang, Heilongjiang, Jilin, and Liaoning, Tibet, Inner Mongolian, Ningxia, Guangxi, Xinjiang, Beijing, Chongqing, Shanghai, and Tianjin.

[4] *hukou* is a household registration system used in mainland China. Workers with rural *hukou* tend to suffer wage penalty compared to workers with urban *hukou* (Asadullah and Xiao, 2020). The *hukou* status can also affect one's access to government-subsidized public services and social welfare programs (Song, 2014).

from more recent literature. For example, according to the seminal work by Card (1999), when parents' years of education are used as instruments, as is usually done in the literature, the IV estimates can be much larger than the OLS estimates. Meanwhile, parents' education has been found to be a significant factor affecting children's education outcomes (Card, 1999; Altonji and Dunn, 1995). Thus, we control for parents' years of education in our full specification. We also control for the interaction between marital status and gender, as existing literature has demonstrated marital status affects men and women's earnings differently (Antonovics and Town, 2004; Juhn and McCue, 2017). Additionally, existing studies have identified significant wage premium related to party membership and union membership in China (Gunderson et al., 2016; Ma and Iwasaki, 2021). Lastly, health capital variables have also been documented as a potential channel leading to differential work payment (Schultz, 2002; Baum and Ford, 2004; Asadullah and Xiao, 2020).

As for the 2SLS method, we construct a novel set of instruments, inspired by Dai and Martins (2020) and Asadullah and Xiao (2020). First, we calculate two enrolment ratios based on the Educational Statistics Yearbook of China from 1987 to 2015: $Eratio_{vus}$, which equals the number of students enrolled in vocational upper secondary schools in year t divided by the number of lower secondary graduates in year t; $Eratio_{aus}$, which equals the number of students enrolled in academic upper secondary schools in year t divided by the number of lower secondary graduates in year t. The construction of those two ratios follows Dai and Martins (2020). However, our calculated ratios are based on more thorough data from 1987 to 2015, compared to the 1987-2007 period in Dai and Martins (2020). Meanwhile, our instruments serve a different purpose from those in Dai and Martins (2020). Dai and Martins (2020) use those ratios to instrument students' choice between academic- and vocational-track upper secondary education, while we use these ratios to instrument students' years of education.

These ratios are *relevant*, because they can predict the probability of a student enrolling into the vocational or academic upper secondary school in a specific province in a given year. Since only academic upper secondary graduates can attend the national college entrance exam, these ratios can reflect the probability of a student pursuing tertiary education. Specifically, the changes in $Eratio_{aus}$ should be in line with the changes in university recruitment. Figure 1 confirms this is indeed the case. Meanwhile, these ratios satisfy the *exclusion* restriction because there are strong reasons to believe that those quotas cannot affect the working income three or more years later. On the one hand, even if a student enters the job market immediately after finishing upper secondary education, the labor supply is a mixture of university graduates, secondary graduates, and others. On the other hand, it is common for young people to migrate and work in a different city or province. In neither case is the enrolment quota likely to directly affect the working income. One major restriction of these enrolment ratio instruments is that they only apply locally to a subsample, namely individuals who graduated from lower secondary schools between 1987 -2015 and continued education after the graduation. We thus include a third instrumental variable - a dummy indicating whether an individual's parent passed away when the individual was 14 years old. Several existing papers have instrumented schooling years with the timing of parental death (Case et al., 2004; Schultz, 2002; Asadullah and Xiao, 2020).

In addition to the conventional IV analyses, we also adopt Lewbel's method which is capable of generating heteroskedasticity-based instruments when no external instruments are available. The Lewbel's method can both generate legitimate estimators for an endogenous regressor on its own and serve as a robustness check to conventional IV analysis (Lewbel, 2012). On the one hand, 'years of education' is a classical example of endogenous variable (Card, 1999), making the Lewbel method an effective complement to conventional 2SLS or OLS. On the other hand, the Lewbel method has been widely used since its introduction.[5] However, both instrumental methods suffer a major drawback - they do not fully utilize the entire sample. As can be seen from Table 5, the 2SLS and Lewbel methods utilize less than half of the available sample, at best.

To fully utilise the sample and to flexibly capture any differential marginal effects of education, we apply the GPSM method introduced by Hirano and Imbens (2004) and developed into a Stata package by Bia and Mattei (2008). This method is similar to conventional propensity score matching in that it attempts to reduce estimation bias by matching on observables. Its main advantage, though, is its capability of estimating the treatment effect of a continuous treatment variable in response to different doses of a treatment. In this paper, we apply a more recent version of the GPS package modified by Guardabascio and Ventura (2014), which allows more flexibility in estimating the GPS when the treatment variable is not normally distributed.

The GPS method is implemented in multiple steps. The first step is to verify whether the conditional distribution of the treatment variable (i.e., years of education) is normally distributed. If not, one needs to specify an alternative distribution. Given the right-skewed nature of years of education, gamma distribution is utilized in our case. Second, the generalized propensity score (GPS) is estimated with generalized linear models. Third, the balancing property is evaluated among different levels of treatments. The logic of the balance test is such that, for the ith treatment level out of n levels, all other $n-1$ treatment levels can serve as its 'control group'. Specifically, the balancing test takes four steps. One first divides the GPS within each treatment level into m quantiles. Then mean differences of each covariate are calculated between individuals in a quantile of a treatment level and those in the same quantile but in all other treatment levels. Next, weighted average of the m differences for each covariate within each treatment is calculated. Lastly, t-tests are conducted on those weighted-averaged covariates between each treatment level and all other treatment levels. Mean difference test results on the 2017 data pre- and post-matching is reported in Table 7 in the Appendix.

Ideally, all the mean difference tests should be insignificant after conditioning on GPS. Yet, these 'perfect matching' situations are often hard to meet in practice. Similar to Bia et al. (2009), the balancing property of our covariates, although not fully achieved, is greatly improved conditional on the GPS. Specifically, the majority of our covariates are not significantly different after balancing; even the unbalanced covariates, their actual difference is reduced by aroung 90% post-matching - a dramatic improvement. After balancing, the

[5] Lewbel (2012), the paper introducing the methodology, has been cited more than 900 times as of January 2021.

conditional expectation of the outcome (i.e., ln hourly wages) is estimated given the treatment and the GPS. Lastly, the average potential outcome (i.e., the returns to education in response to different treatment levels), as well as the marginal treatment effect, is reported along with bootstrapped standard errors.

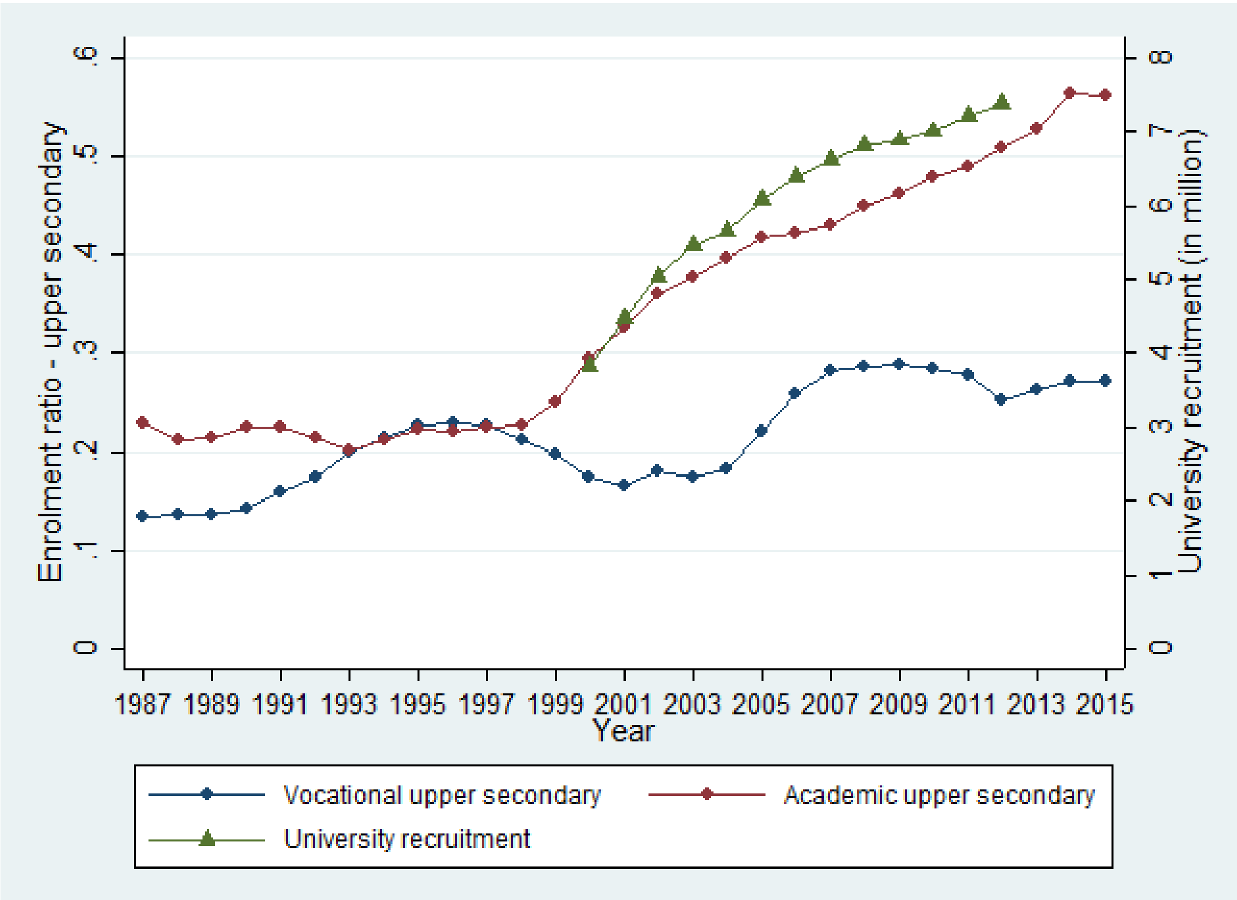


Figure 1: Upper secondary enrolment ratio vs. University recruitment

Note: data related to secondary (both lower and upper) education are manually extracted by the authors from the Educational Statistics Yearbook of China from 1987 to 2015. The vocational (academic) upper secondary enrolment ratio equal vocational (academic) upper secondary entrants of a year divided by lower secondary graduates of the same year. All ratios are compared with Dai and Martins (2020); if a discrepancy arises, we carefully double check our data entry and modify accordingly. The university recruitment data are extracted by the authors from the website of the Ministry of Education of China. The university recruitment data are only available for years 2003-2015, corresponding to upper secondary entrants in years 2000-2012.

5 Findings

Table 3: OLS regressions for returns to education

Dependent variable: ln hourly wages	Specifications (1)	(2)	(3)	(4)
Years of education	0.152***	0.134***	0.080***	0.066***
	(0.002)	(0.002)	(0.003)	(0.003)
Female (yes=1)		-0.262***	-0.285***	-0.064
		(0.012)	(0.013)	(0.044)
Married			0.184***	0.234***
			(0.024)	(0.030)
Female× Married				-0.153***
				(0.044)
Province FE	No	Yes	Yes	Yes
N	24821	24821	18669	15892
Adjusted R^2	0.267	0.349	0.465	0.476

Notes: the data include waged workers from the Chinese General Social Survey in 2010, 2011, 2012, 2013, 2015, and 2017. Specification (1) only controls for years of education. Specification (2) additionally controls for years of experience, experience-squared, gender, ethnicity, and yearly fixed effects. Specification (3) controls for extra and conventionally controlled factors including marital status (with multiple categories), *hukou*, sectors of industry, and location fixed effects. Specification (4) adds pertinent covariates that are sometimes uncontrolled in the literature (i.e. interactions of gender and marital status, parental education, union membership, party membership, whether an individual report having 'average or above' level of English skills, self-perceived health condition, height, and Body Mass Index categories). All specifications control for year fixed effects and report robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 3 presents estimates based on four different specifications of the Mincerian earnings function. An extra year of education relates to 13.4% increase in hourly wages in the baseline specification (column (2)) and 6.6% increase in the full specification (column (4)). The baseline specification shows that female workers suffer a 26.2% wage gap compared to male workers. Interestingly, the economically and statistically significant gender wage difference disappears in the full specification, indicating a potential overestimation of the wage gap in the other specifications due to omitted variable bias. A closer investigation reveals that the gender wage gap is largely attributable to 'marriage discrimination'[6] towards women. In other words, married female workers face a 15% wage disadvantage compared to unmarried female workers. This observation is in line with Polachek (2019) in that gender wage gap is much higher among married individuals than among unmarried ones. Yearly estimations based on the same four specifications are also provided in Table 6 in the Appendix. In terms of the returns to education, yearly results are rather similar to the overall estimates. Basically, estimates in the baseline specification (column (2)) are much larger than those in the full specification (column (4)). We notice that, in all four specifications, the returns to education are the lowest in 2013 and the highest in year 2011. Compared to the findings in Asadullah and Xiao (2020), our results reveal richer yearly dynamics. To further understand the extent to which our controlled covariates alter the estimated returns to education, we present in Table 4 the baseline and full specifications in greater detail. An extra year of education relates to 0.13% increase in hourly wages in the baseline specification and 0.06% increase in the full specification. As expected, experience increases hourly wages at a decreasing rate. Gender and ethnicity coefficients are insignificant in the full specification. In fact, the coefficient of ethnicity even changed from negative to positive. However, as previously mentioned, gender differences manifest itself when gender is interacted with marital status. Married male workers enjoy a large premium of more than 23%, whereas married female workers suffer a penalty of around 15.3%. Having an agricultural *hukou* relates to around 10% decrease in earnings. In line with Card (1999), we find parental education positively relates to children's earnings. We also capture significant union and party membership premium. Interestingly, good English skill relates to around 20% payment premium. In terms of health capital, our estimates are similar to those in Asadullah and Xiao (2020). Being taller and perceiving oneself healthy both positively relate to working income. Among all industry sectors, Hong Kong, Macau or Taiwan funded enterprises and foreign funded enterprises enjoy the largest payment premium.

Table 5 compares OLS estimates with estimates from the two instrumental methods and the GPSM method. Overall, the OLS estimates across the six years are rather stable, remaining between 12.4-14.7% in the baseline specification and between 5.8-7.4% in the full specification. Turning to the 2SLS estimates, we report the first stage F-statistic along with the coefficient of interest. All F-statistics except one are strongly significant, indicating our instruments jointly have significant explanatory power for the years of education after controlling for other covariates. The 2SLS coefficients in the baseline specification are inflated dramatically, rendering virtually unreliable estimates. The 2SLS estimates in the full speci-

[6]Although referred to as 'marriage discrimination', the gender wage gap may well be the result of individual worker's choice over a lifecycle, instead of workplace discrimination (Polachek, 2019, 2007; Mincer and Polachek, 1974).

Table 4: OLS regressions of ln hourly wage with baseline and full specifications

	Baseline specification		Full specification	
Dependent variable: ln hourly wages	(1)	(2)	(3)	(4)
Years of education	0.134***	(0.002)	0.066***	(0.003)
Years of experience	0.022***	(0.003)	0.022***	(0.003)
Experience squared	-0.001***	(0.000)	-0.000***	(0.000)
Female (yes=1)	-0.262***	(0.012)	-0.064	(0.044)
Ethnicity (minority=1)	-0.280***	(0.025)	0.008	(0.030)
Marital status				
De facto			0.354***	(0.096)
Married			0.234***	(0.030)
Re-married			0.362***	(0.093)
Separated			0.123	(0.110)
Divorced			0.146*	(0.078)
Widowed			0.063	(0.139)
Female×*De facto*			-0.077	(0.135)
Female×Married			-0.153***	(0.044)
Female×Re-married			-0.225*	(0.137)
Female× Separated			0.084	(0.173)
Female× Divorced			0.037	(0.106)
Female× Widowed			-0.026	(0.174)
hukou (agricultural=1)			-0.103***	(0.018)
Mother's years of education			0.012***	(0.002)
Father's years of education			0.009***	(0.002)
Union membership (yes=1)			0.100***	(0.018)
Party membership (yes=1)			0.078***	(0.022)
Good English skill (at/above the average level=1)			0.197***	(0.020)
Health capital				
Height (in *cm*)			0.007***	(0.001)
Self-reported physical health				
Bad			-0.178***	(0.032)
Good			0.118***	(0.019)
Body Mass Index				
Underweight (BMI$<$ 18.5)			-0.047	(0.029)
Overweight (25 $\leqslant$BMI$<$ 30)			0.040**	(0.018)
Obese (BMI$\geqslant$ 30))			0.066	(0.049)
Sectors of industry				
State owned enterprise			0.535***	(0.027)
Collectively owned enterprise			0.400***	(0.034)
Privately owned enterprise			0.453***	(0.023)
Hong Kong, Macau or Taiwan funded enterprise			0.623***	(0.084)
Foreign funded enterprise			0.763***	(0.052)
2011	-0.001	(0.026)		
2012	0.249***	(0.020)	0.238***	(0.021)
2013	0.431***	(0.020)	0.420***	(0.022)
2015	0.673***	(0.022)	0.680***	(0.025)
2017	0.816***	(0.021)	0.816***	(0.024)
Province FE	Yes		Yes	
N	24821		15892	
Adjusted R-squared	0.349		0.476	

Notes: the data include waged workers from the Chinese General Social Survey in 2010, 2011, 2012, 2013, 2015, and 2017. 'Years of experience' equals age minus the sum of years of education and 6 (the legal school entry age in China).

fication are 1.5-4.5 times larger than the corresponding OLS estimates, with a significantly shrank sample. Turning to Lewbel method, we first implement a Breusch-Pagan test and find strong evidence of heteroskedasticity, which justifies the applicability of the Lewbel method. As expected, Lewbel estimates lie between OLS estimates and 2SLS estimates (Asadullah and Xiao, 2020).

Table 5: Marginal returns to education - OLS, 2SLS, Lewbel, and GPSM

Dependent variable: ln hourly wages	(1) All years	(2) All years	(3) 2010	(4) 2010	(5) 2011	(6) 2011	(7) 2012	(8) 2012	(9) 2013	(10) 2013	(11) 2015	(12) 2015	(13) 2017	(14) 2017
OLS	0.134***	0.066***	0.144***	0.068***	0.147***	0.074***	0.136***	0.062***	0.124***	0.058***	0.127***	0.072***	0.131***	0.069***
	(0.002)	(0.003)	(0.005)	(0.007)	(0.007)	(0.010)	(0.005)	(0.006)	(0.004)	(0.007)	(0.005)	(0.008)	(0.005)	(0.007)
N	24821	15892	4344	3445	2227	1838	4852	4075	4740	3098	3921	2406	4737	2868
Adj R-squared	0.349	0.476	0.295	0.419	0.301	0.410	0.314	0.433	0.287	0.462	0.267	0.390	0.291	0.486
2SLS	0.414***	0.217***	0.455***	0.259***	0.365***	0.215**	0.385***	0.266***	0.388***	0.107	0.450***	0.222	0.379***	0.211*
	(0.022)	(0.040)	(0.055)	(0.090)	(0.077)	(0.090)	(0.041)	(0.063)	(0.041)	(0.075)	(0.067)	(0.177)	(0.052)	(0.108)
First-stage F-stat	180.2***	48.19***	31.88***	13.61***	14.96***	13.20***	53.97***	22.96***	40.76***	9.034***	18.85***	1.942	30.60***	6.824***
N	6383	4567	845	697	369	314	1480	1274	1501	1102	1130	794	1058	700
Adj R-squared	0.0202	0.339	.	0.265	0.0695	0.253	.	0.246	.	0.356	.	0.277	0.0377	0.252
Lewbel	0.370***	0.190***	0.380***	0.242***	0.221***	0.137***	0.352***	0.249***	0.218***	0.115***	0.297***	0.059	0.134***	0.043
	(0.020)	(0.030)	(0.036)	(0.050)	(0.042)	(0.050)	(0.036)	(0.052)	(0.025)	(0.035)	(0.043)	(0.060)	(0.031)	(0.046)
Heteroskedasticity test	572***	572***	87.49***	87.49***	20.69***	20.69***	47.63***	47.63***	138.1***	138.1***	110***	110***	204.1***	204.1***
N	4929	4567	700	697	319	314	1282	1274	1110	1102	803	794	715	700
Adj R-squared	0.133	0.351	0.0752	0.0416	0.164	0.0848	0.0536	0.0676	0.192	0.136	0.127	0.110	0.151	0.0534
Generalized Propensity Score Matching - marginal effects corresponding to different levels of education														
Compulsory	0.163***	0.176***	0.172***	0.177***	0.167***	0.173***	0.160***	0.160***	0.148***	0.167***	0.144***	0.156***	0.152***	0.179***
	(0.002)	(0.002)	(0.009)	(0.011)	(0.021)	(0.024)	(0.009)	(0.010)	(0.008)	(0.010)	(0.010)	(0.012)	(0.009)	(0.011)
Upper secondary	0.181***	0.185***	0.188***	0.182***	0.174***	0.168***	0.171***	0.159***	0.166***	0.175***	0.154***	0.157***	0.185***	0.204***
	(0.002)	(0.003)	(0.013)	(0.012)	(0.025)	(0.034)	(0.012)	(0.015)	(0.013)	(0.016)	(0.012)	(0.017)	(0.011)	(0.015)
University	0.191***	0.191***	0.197***	0.185***	0.179***	0.168***	0.177***	0.160***	0.177***	0.181***	0.159***	0.159***	0.203***	0.219***
	(0.003)	(0.004)	(0.016)	(0.013)	(0.030)	(0.040)	(0.015)	(0.017)	(0.016)	(0.021)	(0.015)	(0.021)	(0.014)	(0.020)
N	24821	15892	4344	3445	2227	1838	4852	4075	4740	3098	3921	2406	4737	2868

Notes: odd-numbered columns control for years of experience, experience-squared, gender, ethnicity, and yearly fixed effects. Even-numbered columns additionally control for marital status, interaction terms of gender and marital status, *hukou*, parental education, union membership, party membership, whether an individual report having 'average or above' level of English skills, self-perceived health condition, height, Body Mass Index categories, sectors of industry, and location fixed effects. The *First-stage F-stat* represents the F statistics of each first-stage estimation of the 2SLS method. The *Heteroskedasticity test* is done using Breusch-Pagan method. GPSM coefficients are marginal effects corresponding to the last year of compulsory/upper-secondary/university education. Robust standard errors in parentheses for OLS, 2SLS, and Lewbel. Bootstrapped standard errors with 200 repetitions are reported in parentheses for GPSM. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Lastly, we look at the GPSM estimates. As previously mentioned, GPSM estimates two arrays of coefficients (overall effect estimates and marginal effect estimates, respectively) corresponding to various levels of the treatment variable (see Figure 2). To facilitate comparison, we present a representative set of the marginal effect coefficients in Table 5. From column (2) of Table 5, we see that one additional year of compulsory education raises the return by 16.3%, one additional year of upper secondary education raises the return by 18.1%, and one additional year of university education raises the return by 19.1%. One possible explanation for the larger GPSM coefficients, relative to OLS estimates, is that GPSM reduces heterogeneity and captures more cleanly the actual effect of education. To further visualize the changes in marginal returns to education across the six survey years, we plot the marginal effects in Figure 3. Interestingly, although our OLS results do not support the 'decreasing trend' between 2010 and 2015 found by Asadullah and Xiao (2020), the GPSM results are in line with their findings, as can be seen from Table 5 and Figure 3. Yet the low returns to education in year 2015 is reverted in 2017. One possible explanation is that there was indeed a decrease in returns to education in China, but the Chinese government stepped in and initiated the "supply-chain reform" and the "Made in China 2025" vision which seem to have immediately boosted the labor market. From Figure 3, we also see that returns to university education peaked in 2017, widening the payment gap relative to below-tertiary educations. The diffential returns to different levels of education (i.e., returns to university education>returns to upper secondary education>returns to compulsory education) are consistent with findings in a recent meta-analysis on returns to education in China (Churchill and Mishra, 2018).

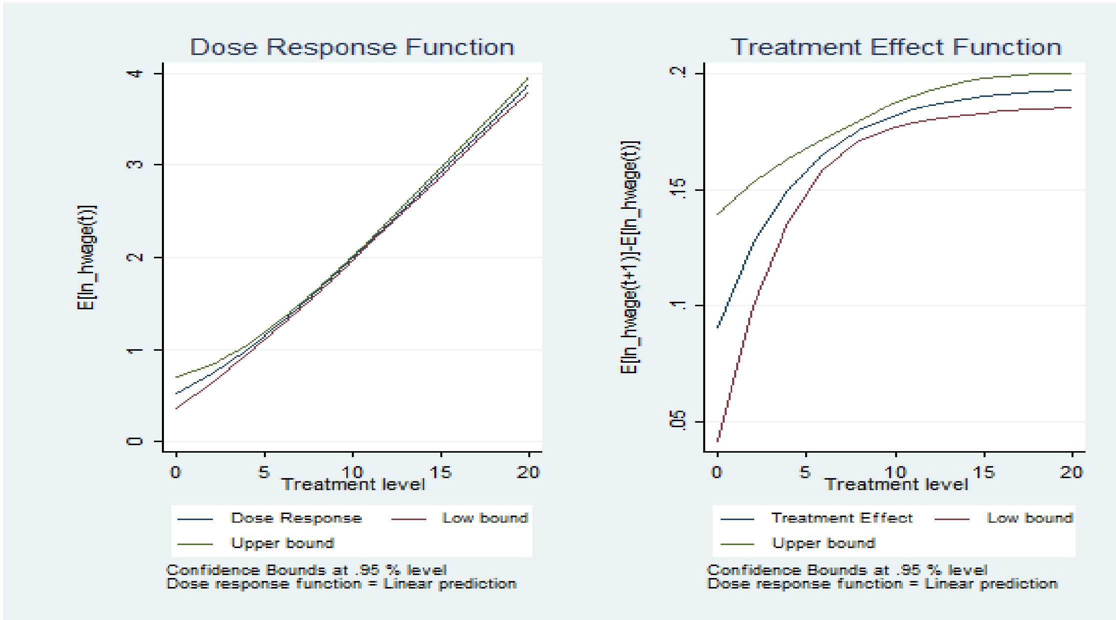
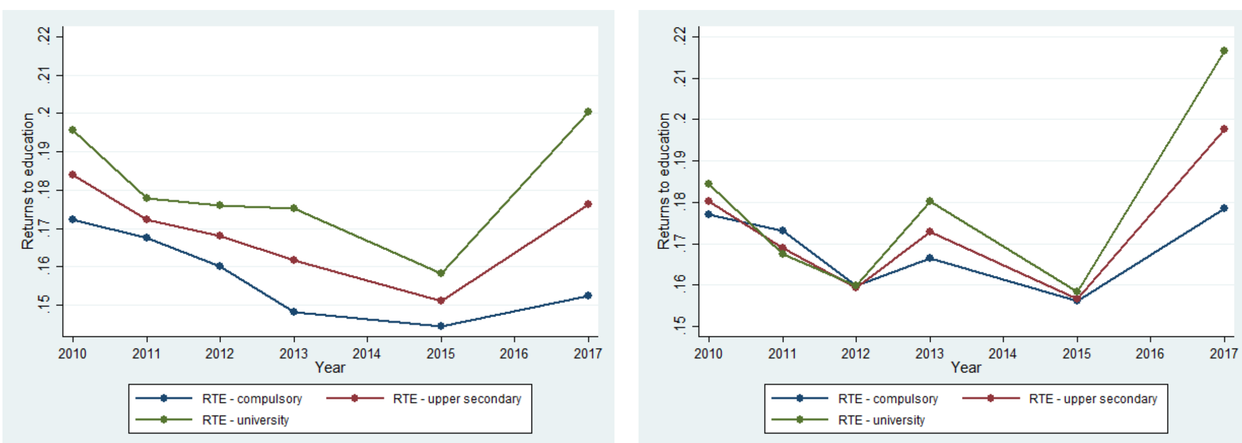


Figure 2: GPSM results for all years

Note: the returns to education is estimated, with full specification, for all the working age people. The full estimating equation controls for all the control variables reported in Table 2. Treatment level is the years of education. Left panel represents overall treatment effect in response to different years of education. Right panel represents marginal treatment effect in response to one unit change in the years of education.



(a) Baseline specification

(b) Full specification

Figure 3: Overall returns to education

Note: the returns to education is estimated for all the working age people. The baseline estimating equation only controls for experience, experience-squared, gender, and ethnicity. The full estimating equation controls for all the control variables reported in Table 2.

6 Discussions and Policy implications

With the ongoing massification of higher education in China, understanding recent changes of returns to education in China is of great policy interest. In this paper, we estimate the returns to education in China with a nationally representative and recently updated dataset. To gauge returns to different levels of education, we apply the GPSM method and compare its estimates to OLS and two IV estimates. We find that returns to university education are higher than those to upper secondary education, both of which are higher than returns to compulsory education. Time-wise, returns to education decreased between 2010 and 2015 but reverted in 2017. The GPSM estimates are closer to OLS estimates, compared to the IV estimates. Overall, GPSM seems to outperform the more popular instrumental methods in complementing the OLS results. It also allows more flexibility in estimating returns to different levels of education.

7 Acknowledgement

This paper is part of the Global Labor Organization (GLO) Virtual Young Scholars Program. The survey data in this paper is from the research project "Chinese General Social Survey (CGSS)" of the National Survey Research Center (NSRC), Renmin University of China.

References

Altonji, J. G. and Dunn, T. A. (1995). The effects of school and family characteristics on the return to education. Technical report, National Bureau of Economic Research.

Antonovics, K. and Town, R. (2004). Are all the good men married? uncovering the sources of the marital wage premium. *American Economic Review*, 94(2):317–321.

Asadullah, M. N. and Xiao, S. (2020). The changing pattern of wage returns to education in post-reform china. *Structural Change and Economic Dynamics*, 53:137–148.

Baum, C. L. and Ford, W. F. (2004). The wage effects of obesity: a longitudinal study. *Health economics*, 13(9):885–899.

Bia, M., Leombruni, R., Messe, P., et al. (2009). Young in-old out: A new evaluation based on generalized propensity score. *LABORatorio R. Revelli*.

Bia, M. and Mattei, A. (2008). A stata package for the estimation of the dose-response function through adjustment for the generalized propensity score. *The Stata Journal*, 8(3):354–373.

Byron, R. P. and Manaloto, E. Q. (1990). Returns to education in china. *Economic development and cultural change*, 38(4):783–796.

Card, D. (1999). The causal effect of education on earnings. In *Handbook of Labor Economics*, volume 3, pages 1801–1863. Elsevier Science.

Case, A., Paxson, C., and Ableidinger, J. (2004). Orphans in africa: Parental death, poverty, and school enrollment. *Demography*, 41(3):483–508.

Churchill, S. A. and Mishra, V. (2018). Returns to education in china: a meta-analysis. *Applied Economics*, 50(54):5903–5919.

Dai, L. and Martins, P. S. (2020). Does vocational education pay off in china? instrumental-variable quantile-regression evidence. Technical report, GLO Discussion Paper.

Goldin, C. and Katz, L. F. (2007). The race between education and technology: The evolution of us educational wage differentials, 1890 to 2005. Technical report, National Bureau of Economic Research.

Guardabascio, B. and Ventura, M. (2014). Estimating the dose–response function through a generalized linear model approach. *The Stata Journal*, 14(1):141–158.

Gunderson, M., Lee, B. Y., and Wang, H. (2016). Union pay premium in china: an individual-level analysis. *International Journal of Manpower*.

Hirano, K. and Imbens, G. W. (2004). The propensity score with continuous treatments. *Applied Bayesian modeling and causal inference from incomplete-data perspectives*, 226164:73–84.

Jamison, D. T. and Van der Gaag, J. (1987). Education and earnings in the people's republic of china. *Economics of Education Review*, 6(2):161–166.

Johnson, E. N. and Chow, G. C. (1997). Rates of return to schooling in china. *Pacific Economic Review*, 2(2):101–113.

Juhn, C. and McCue, K. (2017). Specialization then and now: Marriage, children, and the gender earnings gap across cohorts. *Journal of Economic Perspectives*, 31(1):183–204.

Lewbel, A. (2012). Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models. *Journal of Business & Economic Statistics*, 30(1):67–80.

Li, H. (2003). Economic transition and returns to education in china. *Economics of education review*, 22(3):317–328.

Ma, X. and Iwasaki, I. (2021). Does communist party membership bring a wage premium in china? a meta-analysis. *Journal of Chinese Economic and Business Studies*, pages 1–40.

Mason, G. (1996). Graduate utilisation in british industry: the initial impact of mass higher education. *National Institute Economic Review*, 156(1):93–103.

Meara, K., Pastore, F., and Webster, A. (2020). The gender pay gap in the usa: a matching study. *Journal of Population Economics*, 33(1):271–305.

Meng, X. and Kidd, M. P. (1997). Labor market reform and the changing structure of wage determination in china's state sector during the 1980s. *Journal of Comparative Economics*, 25(3):403–421.

Mincer, J. and Polachek, S. (1974). Family investments in human capital: Earnings of women. *Journal of political Economy*, 82(2, Part 2):S76–S108.

Polachek, S. (2007). Earnings over the lifecycle: The mincer earnings function and its applications (iza discussion papers no. 3181). *Institute for the Study of Labor, Bonn.*

Polachek, S. W. (2019). Equal pay legislation and the gender wage gap. *IZA World of Labor.*

Ren, W. and Miller, P. W. (2012). Changes over time in the return to education in urban china: Conventional and oru estimates. *China Economic Review*, 23(1):154–169.

Schultz, T. P. (2002). Wage gains associated with height as a form of health human capital. *American Economic Review*, 92(2):349–353.

Song, Y. (2014). What should economists know about the current chinese hukou system? *China Economic Review*, 29:200–212.

Tao Yang, D. (1997). Education and off-farm work. *Economic development and cultural change*, 45(3):613–632.

Uysal, S. D. (2015). Doubly robust estimation of causal effects with multivalued treatments: an application to the returns to schooling. *Journal of Applied Econometrics*, 30(5):763–786.

Wang, L. (2013). How does education affect the earnings distribution in urban china? *Oxford Bulletin of Economics and Statistics*, 75(3):435–454.

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.

A Appendix

Table 6: OLS regressions for returns to education

Dep. var.: ln hourly wages	(1)	(2)	(3)	(4)
All years	0.152***	0.134***	0.080***	0.066***
	(0.002)	(0.002)	(0.003)	(0.003)
N	24821	24821	18669	15892
Adj R-squared	0.267	0.349	0.465	0.476
2010	0.157***	0.144***	0.083***	0.068***
	(0.004)	(0.005)	(0.006)	(0.007)
N	4344	4344	3593	3445
Adj R-squared	0.285	0.295	0.407	0.419
2011	0.160***	0.147***	0.085***	0.074***
	(0.006)	(0.007)	(0.009)	(0.010)‡
N	2227	2227	1950	1838
Adj R-squared	0.278	0.301	0.395	0.410
2012	0.152***	0.136***	0.076***	0.062***
	(0.003)	(0.005)	(0.006)	(0.006)
N	4852	4852	4227	4075
Adj R-squared	0.285	0.314	0.417	0.433
2013	0.138***	0.124***	0.068***	0.058***
	(0.004)	(0.004)	(0.007)	(0.007)
N	4740	4740	3221	3098
Adj R-squared	0.265	0.287	0.447	0.462
2015	0.138***	0.127***	0.087***	0.072***
	(0.004)	(0.005)	(0.007)	(0.008)
N	3921	3921	2595	2406
Adj R-squared	0.245	0.267	0.385	0.390
2017	0.145***	0.131***	0.080***	0.069***
	(0.004)	(0.005)	(0.007)	(0.007)
N	4737	4737	3083	2868
Adj R-squared	0.270	0.291	0.463	0.486

Notes: the data include waged workers from the Chinese General Social Survey in 2010, 2011, 2012, 2013, 2015, and 2017. Specification (1) only controls for years of education. Specification (2) additionally controls for years of experience, experience-squared, gender, ethnicity, and yearly fixed effects. Specification (3) controls for extra and conventionally controlled factors (i.e., marital status, *hukou*, sectors of industry, and location fixed effects). Specification (4) adds pertinent covariates that are sometimes uncontrolled in the literature (i.e. an interaction of gender and marital status, parental education, union membership, party membership, whether an individual report having 'average or above' level of English skills, self-perceived health condition, height, and Body Mass Index categories). ‡ English skills and self-perceived health condition are not controlled for year 2011 due to data unavailability.

Table 7: Balance tests pre- and post- GPS matching (Year 2017)

Variables	Diff.	S.E.	t-stat	Diff.	S.E.	t-stat	Diff.	S.E.	t-stat	Diff.	S.E.	t-stat
	Panel A: treatment level [0,9]						**Panel B: treatment level (9,12]**					
	Unadjusted			Adjusted			Unadjusted			Adjusted		
Experience	-13.464	0.305	**-44.178**	-1.626	0.041	**-40.053**	1.735	0.448	**3.874**	-0.326	0.077	**-4.258**
Experience squared	-627.859	14.893	**-42.158**	-81.204	1.856	**-43.749**	123.152	21.556	**5.713**	-11.022	3.794	**-2.905**
Gender	-0.004	0.014	**-0.269**	0.002	0.002	**0.806**	0.084	0.018	**4.747**	0.017	0.003	**5.301**
Marital status	-0.133	0.012	**-11.074**	-0.012	0.002	**-6.000**	0.006	0.015	**0.417**	-0.004	0.003	**-1.535**
Ethnicity	-0.044	0.008	**-5.681**	-0.008	0.001	**-6.603**	0.032	0.010	**3.346**	0.003	0.002	**1.619**
hukou	-0.459	0.012	**-36.837**	-0.055	0.002	**-30.530**	0.098	0.017	**5.634**	0.010	0.003	**3.193**
Mother education	4.519	0.121	**37.513**	0.512	0.017	**29.709**	-0.373	0.171	**-2.183**	0.049	0.031	**1.608**
Father education	4.377	0.123	**35.568**	0.509	0.017	**29.140**	-0.591	0.173	**-3.424**	0.034	0.031	**1.090**
Union membership	0.228	0.010	**22.633**	0.028	0.002	**16.612**	-0.024	0.013	**-1.805**	-0.008	0.002	**-3.079**
Party membership	0.171	0.009	**18.889**	0.016	0.001	**11.619**	0.019	0.012	**1.658**	0.000	0.002	**-0.142**
BMI: underweight	0.038	0.008	**4.824**	0.004	0.001	**3.041**	0.005	0.010	**0.462**	0.000	0.002	**-0.123**
BMI: overweight	-0.015	0.012	**-1.247**	0.000	0.002	**0.033**	-0.003	0.015	**-0.192**	-0.002	0.003	**-0.691**
BMI: obese	0.002	0.005	**0.351**	0.001	0.001	**0.870**	-0.002	0.006	**-0.333**	-0.001	0.001	**-0.918**
Good English skills	0.294	0.010	**29.798**	0.032	0.001	**22.052**	0.060	0.013	**4.531**	0.008	0.003	**3.244**
Self-reported heath: bad	-0.099	0.008	**-12.078**	-0.017	0.001	**-13.645**	0.037	0.010	**3.575**	0.001	0.002	**0.480**
Self-reported heath: good	0.164	0.013	**12.349**	0.024	0.002	**11.297**	-0.038	0.017	**-2.279**	0.002	0.003	**0.718**
Height (cm)	3.063	0.223	**13.711**	0.379	0.035	**10.867**	-1.330	0.281	**-4.727**	-0.199	0.051	**-3.876**
Sector: state owned enterprise	0.302	0.014	**21.058**	0.033	0.002	**18.610**	-0.052	0.019	**-2.680**	-0.004	0.003	**-1.478**
Sector: collectively owned enterprise	0.035	0.008	**4.358**	0.004	0.001	**4.379**	-0.013	0.010	**-1.256**	0.000	0.001	**-0.162**
Sector: privately owned enterprise	0.150	0.018	**8.561**	0.024	0.002	**10.867**	-0.163	0.022	**-7.296**	-0.009	0.003	**-3.279**
Sector: Hong Kong, Macau or Taiwan funded enterprise	0.007	0.003	**2.751**	0.001	0.000	**2.016**	0.005	0.003	**1.390**	0.001	0.000	**1.485**
Sector: foreign funded enterprise	0.035	0.006	**5.990**	0.003	0.001	**4.541**	0.014	0.007	**1.853**	0.002	0.001	**2.449**
Location: central	-0.087	0.012	**-7.474**	-0.012	0.002	**-6.724**	-0.013	0.015	**-0.859**	-0.002	0.003	**-0.966**
Location: west	-0.148	0.012	**-12.402**	-0.024	0.002	**-12.961**	0.032	0.015	**2.108**	0.004	0.003	**1.601**
Location: northeast	-0.022	0.010	**-2.316**	-0.003	0.001	**-1.907**	0.021	0.012	**1.811**	0.003	0.002	**1.377**
	Panel C: treatment level (12,15]						**Panel D: treatment level (15,16]**					
	Unadjusted			Adjusted			Unadjusted			Adjusted		
Experience	8.709	0.529	**16.468**	0.526	0.083	**6.316**	13.293	0.455	**29.223**	1.269	0.085	**14.992**
Experience squared	405.939	25.553	**15.886**	27.293	4.403	**6.199**	580.742	22.290	**26.055**	54.953	4.420	**12.432**
Gender	-0.031	0.022	**-1.444**	-0.009	0.004	**-2.261**	-0.059	0.020	**-3.027**	-0.015	0.004	**-3.824**

Continued on next page

Table 7 – continued from previous page

Variables	Diff.	S.E.	t-stat	Diff.	S.E.	t-stat	Diff.	S.E.	t-stat	Diff.	S.E.	t-stat
Marital status	0.090	0.018	**4.980**	0.004	0.003	**1.349**	0.148	0.016	**8.995**	0.009	0.003	**2.952**
Ethnicity	0.033	0.012	**2.772**	0.004	0.002	**1.655**	0.010	0.011	**0.913**	-0.001	0.002	**-0.643**
hukou	0.277	0.021	**13.294**	0.020	0.003	**6.170**	0.425	0.018	**23.304**	0.021	0.003	**6.577**
Mother education	-2.542	0.203	**-12.515**	-0.125	0.031	**-3.974**	-4.853	0.173	**-28.114**	-0.345	0.030	**-11.310**
Father education	-2.353	0.205	**-11.464**	-0.094	0.034	**-2.799**	-4.475	0.176	**-25.376**	-0.339	0.033	**-10.369**
Union membership	-0.163	0.016	**-10.298**	-0.014	0.003	**-5.199**	-0.219	0.014	**-15.458**	-0.012	0.003	**-4.648**
Party membership	-0.098	0.014	**-6.991**	-0.005	0.002	**-2.056**	-0.180	0.013	**-14.353**	-0.009	0.002	**-4.298**
BMI: underweight	-0.033	0.012	**-2.841**	0.001	0.002	**0.417**	-0.037	0.011	**-3.488**	-0.001	0.002	**-0.690**
BMI: overweight	0.004	0.018	**0.230**	-0.003	0.003	**-0.963**	0.023	0.016	**1.428**	0.002	0.003	**0.458**
BMI: obese	-0.017	0.008	**-2.132**	-0.002	0.001	**-1.777**	0.010	0.007	**1.423**	0.000	0.001	**-0.096**
Good English skills	-0.124	0.016	**-7.703**	-0.002	0.002	**-0.803**	-0.399	0.014	**-29.613**	-0.023	0.002	**-11.337**
Self-reported heath: bad	0.059	0.013	**4.760**	0.009	0.003	**3.215**	0.081	0.011	**7.119**	0.010	0.003	**3.850**
Self-reported heath: good	-0.119	0.020	**-5.882**	-0.014	0.004	**-3.675**	-0.146	0.018	**-7.947**	-0.016	0.004	**-4.366**
Height (cm)	-1.533	0.341	**-4.490**	-0.090	0.062	**-1.454**	-2.296	0.309	**-7.428**	-0.023	0.060	**-0.389**
Sector: state owned enterprise	-0.177	0.022	**-8.206**	-0.016	0.003	**-5.485**	-0.223	0.019	**-11.583**	-0.018	0.003	**-6.312**
Sector: collectively owned enterprise	-0.025	0.011	**-2.210**	-0.003	0.002	**-1.939**	-0.028	0.010	**-2.720**	-0.004	0.002	**-2.532**
Sector: privately owned enterprise	-0.091	0.025	**-3.603**	-0.010	0.004	**-2.634**	-0.037	0.023	**-1.638**	-0.010	0.004	**-2.634**
Sector: Hong Kong, Macau or Taiwan funded enterprise	-0.007	0.004	**-1.775**	-0.001	0.000	**-2.760**	-0.006	0.003	**-1.783**	0.000	0.000	**-1.045**
Sector: foreign funded enterprise	-0.014	0.008	**-1.724**	-0.001	0.001	**-0.676**	-0.047	0.007	**-6.409**	-0.002	0.001	**-2.216**
Location: central	0.049	0.018	**2.759**	0.004	0.003	**1.339**	0.109	0.016	**6.844**	0.013	0.003	**4.009**
Location: west	0.104	0.018	**5.734**	0.007	0.004	**2.085**	0.121	0.016	**7.374**	0.009	0.003	**2.716**
Location: northeast	0.025	0.014	**1.735**	0.001	0.003	**0.357**	-0.016	0.013	**-1.187**	0.000	0.003	**0.050**
	Panel E: treatment level (16,19]											
	Unadjusted			Adjusted								
Experience	14.033	1.158	**12.116**	1.180	0.292	**4.037**						
Experience squared	586.774	56.069	**10.465**	53.144	14.386	**3.694**						
Gender	-0.056	0.047	**-1.198**	-0.016	0.012	**-1.402**	This					
Marital status	0.100	0.039	**2.550**	-0.007	0.010	**-0.727**	is					
Ethnicity	0.036	0.025	**1.407**	0.004	0.007	**0.569**	intentionally					
hukou	0.465	0.045	**10.264**	0.005	0.011	**0.490**	left					
Mother education	-5.497	0.439	**-12.519**	-0.346	0.101	**-3.428**	blank.					
Father education	-5.405	0.447	**-12.096**	-0.315	0.104	**-3.040**						
Union membership	-0.250	0.034	**-7.318**	-0.002	0.008	**-0.194**						
Party membership	-0.450	0.030	**-15.145**	-0.044	0.006	**-7.292**						

Continued on next page

Table 7 – continued from previous page

Variables	Diff.	S.E.	t-stat	Diff.	S.E.	t-stat	Diff.	S.E.	t-stat	Diff.	S.E.	t-stat
BMI: underweight	-0.063	0.025	**-2.480**	-0.004	0.006	**-0.643**						
BMI: overweight	0.031	0.039	**0.796**	-0.005	0.010	**-0.520**						
BMI: obese	0.015	0.017	**0.920**	-0.002	0.004	**-0.464**						
Good English skills	-0.679	0.034	**-20.288**	-0.061	0.007	**-8.234**						
Self-reported heath: bad	0.057	0.027	**2.098**	0.006	0.007	**0.778**						
Self-reported heath: good	-0.084	0.044	**-1.917**	-0.012	0.011	**-1.105**						
Height (cm)	-3.051	0.738	**-4.135**	-0.124	0.178	**-0.698**						
Sector: state owned enterprise	-0.381	0.043	**-8.783**	-0.026	0.009	**-2.952**		This				
Sector: collectively owned enterprise	0.031	0.023	**1.361**	0.001	0.005	**0.136**		is				
Sector: privately owned enterprise	0.147	0.051	**2.901**	-0.004	0.011	**-0.359**		intentionally				
Sector: Hong Kong, Macau or Taiwan funded enterprise	-0.027	0.008	**-3.467**	0.000	0.002	**0.019**		left blank.				
Sector: foreign funded enterprise	-0.059	0.017	**-3.559**	-0.003	0.003	**-1.110**						
Location: central	0.166	0.038	**4.362**	0.016	0.009	**1.725**						
Location: west	0.174	0.039	**4.444**	0.016	0.010	**1.561**						
Location: northeast	0.055	0.031	**1.773**	0.012	0.007	**1.661**						