Dai, Li; Martins, Pedro S.

Working Paper

Does vocational education pay off in China? Instrumental-variable quantile-regression evidence

GLO Discussion Paper, No. 495

Provided in Cooperation with:
Global Labor Organization (GLO)

Suggested Citation: Dai, Li; Martins, Pedro S. (2020) : Does vocational education pay off in China? Instrumental-variable quantile-regression evidence, GLO Discussion Paper, No. 495, Global Labor Organization (GLO), Essen

This Version is available at:
https://hdl.handle.net/10419/214871

Does vocational education pay off in China? Instrumental-variable quantile-regression evidence*

Li Dai†
Queen Mary University of London

Pedro S. Martins‡
Queen Mary University of London
& IZA & GLO

Abstract

As China's firms upgrade their position in the quality ladder, vocational education may become more important. In this paper, we study returns to secondary vocational education in China paying attention to individual heterogeneity. We use instrumental variables based on geographical and longitudinal changes in enrolment to address the selection between the two types of education. We find that vocational education provides a wage premium vis-à-vis academic education of over 30% but which applies only for individuals at the middle of the conditional wage distribution.

Keywords: Human capital, vocational education, quantile treatment effects.
JEL Codes: I26, I25, J24, J31, C36.

*We thank comments from Martha Prevezer, Klaus Zimmermann, Corrado Giulietti and participants at a 2018 Queen Mary University of London workshop and the 2019 GLO-Renmin University conference. Dai thanks financial support from the China Scholarship Council. All errors are our own.

†Corresponding author. Email:li.dai@qmul.ac.uk. Address: School of Business and Management, Queen Mary University of London, 327 Mile End Road, London E1 4NS, United Kingdom. Phone: +44 (0)20 7882 8570

‡Email: p.martins@qmul.ac.uk. Address: School of Business and Management, Queen Mary University of London, 327 Mile End Road, London E1 4NS, United Kingdom. Web: https://sites.google.com/site/pmrsmartins/.

1 Introduction

Formal education is one of the most important forms of investment that both households and governments can make. Consequently, younger generations around the world - including in emerging economies - tend to spend more and more time in school, at the same time that households and governments make significant financial investments in their education. In this context, questions about the most appropriate types of schooling are increasingly more important.

The voluminous research on returns to different tracks of education seeks to provide evidence on this question. However, these studies face several challenges. On the one hand, students may self-select into vocational or academic education, given different entry requirements and education/career prospects that each stream offers. In this case, Ordinary least squares analyses would generate biased estimates and potentially misleading conclusions. On the other hand, each track of education will be subject to heterogeneity in the quality of education and in labour market demand for its graduates. For example, a UK review of vocational education (Wolf, 2011) indicates that although academic education is generally preferred over vocational education, apprenticeships at prestigious firms are subject to greater demand compared to popular courses at the best universities. Although estimation of the mean effect of a type of education provides useful information, it may overlook heterogeneity in its returns.

Our study contributes to the literature by providing evidence on the current and recent returns to vocational education in comparison to academic education in China. Given its size, China has a large number of graduates from different types of education (as shown in Figure 1). Our study also contributes to policy making regarding the allocation of scarce public funding in education and across its different streams. This is particularly relevant as China's firms seek to upgrade their position in the quality ladder and the schooling profiles of their workers may be an increasingly important driver of such process.

(Figure 1 here)

We believe this is the first study to model heterogeneous effects of the education track on wages over the distribution and to do so in the presence of endogeneity. We exploit a subtler version of policy leverage to instrument individual's choice of education track in the context of our IV-QR model. Our instrument is the provincial and longitudinal variation in secondary enrolment quotas, with respect to compulsory schooling graduation numbers. This

approach draws on the literature that examines the role of education provision in individual educational attainment and estimation of labour market outcomes (Card, 1993; Duflo, 2001; Carneiro et al., 2018).

Our OLS results show that secondary vocational graduates receive higher wages compared to their academic counterparts both on average and along the conditional wage distribution. However, we also show that these vocational premiums are partly driven by positive selection into vocational education among the individuals. Our first-stage results of 2SLS estimation show a positive correlation between an individual's secondary vocational education completion, on the one hand, and upper secondary vocational enrolment in relation to lower secondary graduation in the individual's region and year of study, on the other hand. The relationship is also of economic and statistical significance.

After controlling for selection with the instrumental variable above, we find that, on average, secondary vocational education does not generate any wage premium in comparison to academic education. However, our instrumental variable quantile regression (IV-QR) results show evidence of variations in the vocational effects across the conditional wage distribution. Specifically, individuals with the lowest or the highest earnings potential do not benefit from secondary vocational education vis-à-vis academic education. On the other hand, vocational education does generate a significant wage premium of 31%-37% for individuals with average earnings potential. This pattern also applies to different sub-samples of workers, namely those with urban household registrations (hukou) and when also considering lower secondary education.

The remaining of the paper is organised as follows. Section 2 discusses the Chinese vocational education system and the labour market implications of different tracks of education. The research design and the data are described in Section 3. The econometric results are presented in the Section 5 and the robustness checks in Section 6. The final section concludes.

2 Research context

2.1 Literature review

Existing research on whether vocational or academic education offers better labour market prospects provides mixed findings. Some studies credit vocational education for higher employability (O'Reilly et al., 2015) and as a useful active labour market policy (Eichhorst et al.,

2015). In addition, the benefits of vocational education extend beyond labour market outcomes. Traditionally, vocational education and training are associated with skills and quality of work, leading to respectful job titles in many societies (Jäger, 2016; Hutchinson and Kettlewell, 2015). Moreover, vocational education keeps low achievers in school (Brunello and Rocco, 2015), increasing their skills and preventing them from becoming 'NEETs' ('not in employment, education or training').

In contrast, there are also studies that associate vocational education with poor labour market outcomes. Evidence of wage penalties suffered by vocational education graduates is reported in several papers (Brunello and Rocco, 2015; Psacharopoulos and Patrinos, 2018). Some indicate a trade-off between short-term gains and long-term losses, as the skills acquired in vocational education facilitate quicker school-to-work transitions, but may become obsolete over time, resulting in poorer career prospects (Hanushek and Wössmann, 2006).

We also highlight the increasing interest on the heterogeneity in education returns and different tracks of education. In a encyclopedic review of return to education studies over the past two decades, Psacharopoulos and Patrinos (2018) conclude that there is a wage premium of academic education over vocational education. However, in the presence of individual selection, those who opt for vocational education may possess different ability and preferences from those choosing academic education (Eichhorst et al., 2015). This creates an empirical challenge, as individuals are not randomly assigned to either track of education.

To address the resulting selection, Meer (2007) applies multinomial logit first-stage regression to form counterfactuals for those in the vocational track. Using the 1999 wave of the US National Education Longitudinal Survey, the study shows that individuals with vocational qualifications would earn 3.7% less if they had been on the academic track. On the other hand, the income of academic graduates would be 18.0% less if they had taken the vocational track. However, unobserved characteristics and circumstances (Ota and Moffatt, 2007), as well as expectations for future earnings (Wilson et al., 2005), would matter when making the choice between academic and vocational education. Therefore, matching by observed characteristics in Meer (2007) may be insufficient. Oosterbeek and Webbink (2007) applies difference-in-differences to evaluate the long-term wage effect of a one-year extension of vocational education in the Netherlands in 1975, and finds no marginal effect on income. The study tests the parallel trend assumption by demonstrating that the wage-cohort relation

is similar between the treatment and control groups before and after the extension.

Instrumental variables are another frequently used method to estimate the return to vocational education. For instance, Fersterer et al. (2008) use exogenous changes in apprentice training duration that result from the failure of host firms in Austria. Their results show that the IV estimates of the marginal return to an additional year of apprenticeship are comparable with the OLS estimates, which suggests that selection into apprenticeship may be limited. Cappellari (2004) shows that, with multiple IVs derived from family characteristics, academic graduates are more likely to be in the low-pay group than the vocationally educated. As always, the validity of instrumental variables may be discussed: firm failure may be a concern if the failing firms were likely to hire less skilled apprentices (and to offer lower wages). Fersterer et al. (2008) defend the validity of their identification using a sample of small firms where failures happened suddenly, excluding firms in which employment decline occurred before firm exit. The instrumental variable used in Cappellari (2004) for attending general high school is having a grandparent with at least a high school degree, under the assumption that intergenerational schooling decision loses its relevant in shaping students' choices when the students are between 14 and 19, conditional on its impact on students at age 14.

Another topic of interest is the heterogeneity in the return to vocational education. Although the estimation of mean return provides useful information, the analysis of how labour market rewards education for individuals with different characteristics can increase our understanding of the role of human capital. Some studies investigate heterogeneous treatment effects by splitting samples by observables, e.g. gender, age cohort, etc. Another useful, and potential more intuitive, approach is to examine heterogeneity by unobserved characteristics across the conditional distribution (Koenker and Bassett Jr, 1978). Martins and Pereira (2004) investigate the heterogeneous effects of an additional year of schooling on wages using European and US data, showing a consistent pattern across most of the 16 countries covered that individuals at upper quantiles receive higher marginal returns to schooling. However, this study does not differentiate between academic and vocational education nor considers the endogeneity of schooling.

Heterogeneity in vocational education provision and labour market outcomes has been observed in many studies (Wolf, 2011; Dearden et al., 2002). However, only two papers that we know examine the heterogeneity in the return to one track of education over the other.

McIntosh and Morris (2016) uses quantile regression to estimate the marginal return to an extra year of vocational education. Based on the UK Labour Force Survey, the study shows that marginal returns to vocational education generally increase when moving up along the quantiles of the conditional wage distribution. Balestra and Backes-Gellner (2017) use the same approach to examine heterogeneous wage effects of vocational and academic education across conditional wage quantiles. Their results show that vocational education is better rewarded than academic education at lower quantiles. However, both McIntosh and Morris (2016) and Balestra and Backes-Gellner (2017) did not address the issue of individual selection into different tracks of education.

2.2 The case of China

According to Pepper (2000), the idea that 'bad students go to vocational schools' prevails among Chinese families and even education practitioners. Jiang and Xu (2005) use a probit model and find that vocational schools in China tend to enroll students from poor social and economic backgrounds. This is supported by Hannum et al. (2011), who shows that educational opportunity and selection by exams largely determine students' transition to secondary and tertiary education in the context of China. Ling (2015) points to discrimination against vocational education through narratives from vocational graduates in China. The study indicates that vocational education, together with the hukou policy[1], serve as a policy leverage to prevent assimilation of immigrants' children into the local labour market. Specifically, children of immigrants are offered priority in admission into vocational education, which prepares them for the occupations that are predominantly taken by their migrant parents.

There is little evidence that compares the returns to vocational and academic education in China. Among the limited quantitative evidence, Li et al. (2012) regresses wages on a list of qualification dummies in a Becker-Mincer wage model. They find that the return to vocational high school is 19-21% higher than compulsory or lower education, and 4-5% higher than academic high school. At tertiary level, on the contrary, tertiary academic education offers a wage premium of 35-40% over tertiary vocational education. However, Li et al. (2012)

[1]The hukou, or household registration, policy is an institution by which every individual is registered as a resident of a particular administrative area since birth. Geographic migration (and social mobility, e.g. agricultural to non-agricultural) is largely limited by the hukou. Although the constraints placed by the hukou system have now been largely eliminated, they still relate to the entitlement of local social benefits. Migrant workers without the local hukou are not fully entitled to the benefits provided by governments to local residents, including free compulsory education.

did not address the potential selection into different types of education.

We also provide an overview of the institutional setting of the education system in China. Vocational education forms a parallel track to academic education, yet possesses unique features of its own. Historically, Chinese vocational education experienced a volatile course of development as the government's priorities shifted between vocational and academic education for decades. According to Tsang (2000), the shifts largely result from the 'two-line struggles' between conservatives and reformers within the ruling Communist party, which are reflected on three education dilemmas: education for political/ideological development vs economic development, education for equality vs efficiency, and positive vs negative attitude towards the intellectuals and the skilled.

In addition to the volatile development of vocational education over time, regional discrepancies in vocational education provision are substantial. Hansen and Woronov (2013) show large regional differences in education provision in China as more developed areas receive disproportionately more and better education resources than poorer areas. Another important feature of the Chinese vocational education system is that the relationship between vocational education provision and labour market demand is largely negotiated locally (OECD, 2010). Vocational education provision has been closely aligned with local development priorities for decades. Moreover, the dynamics in urbanisation and industrialisation across China lead to variation in vocational education availability (Lin et al., 2004). Figure 2 presents the density of secondary vocational schools (SVSs) across China and its changes over time, measured as the number of SVSs per 10,000 residents in a province and year. When China began shifting from a planned economy to a market-oriented one in 1987, the Western region possessed a higher SVS density compared to the central and coastal regions. Since then, the density in the Western region has gradually declined while those in the central provinces have increased. Up until 2007, provincial variation in the SVS density is still significant.

(Figure 2 here)

As for the structure of the educational system, students who complete the nine-year compulsory education (six years of primary education plus three years of lower secondary education) can choose between academic education at secondary academic schools and vocational education at secondary vocational schools, or enter the labour market (Figure 3). Entering into academic or vocational track is largely determined by a student's performance in upper

secondary entrance exams. The total number of slots, or enrolment quotas, provided by secondary academic schools and secondary vocational schools are set preceeding the exams on the basis of local government's forecasts of labour market demand and development priorities. The quotas and the score structure of students determine the score requirements for entering academic education and secondary vocational education. Typically, the scores required to enter academic education are above those for entering vocational education. In addition, academic admission precedes vocational admission. Students who fail to meet the requirements of academic education enter the pools for vocational admission. Students whose scores fall below the vocational education requirement can either resit high school entrance exams in the following year or drop out. An OECD review of Chinese vocational education (OECD, 2010) shows that, in 2007, 74% of compulsory education graduates entered upper secondary education. Within this group, 43% were enrolled in vocational education, and the remaining in academic education.

(Figure 3 here)

As a legacy of the communist era, planning is still embedded in Chinese educational policy, as demonstrated by many studies, e.g. in Zheng (2018). Enrolment quotas are a frequently-used tool to adjust the supply of skilled workers to serve particular development priorities. For example, the country's 12th 'pivotal guidelines for social and economic development', *Five-year Plan for National Economic and Social Development (2011-2015)*, prioritises the supply of skilled workers to meet the labour market demand derived from progressing urbanisation and industrialisation. The Ministry of Education responds to this guideline by publishing a key policy paper, *National Guideline for Intermediate and Long-term Educational Reform and Development (2010-2020)*, which includes increasing the ratio of vocational enrolment in relation to academic enrolment, and building more vocational schools nationwide. Another example accounted in Dai et al. (2018) shows that the Chinese government increased tertiary education enrolment by 47% in 1999 immediately after the Asian financial crisis, in a response to rising youth unemployment and economic slowdown. In 2000 and 2001, the growth rate in enrolment declined to 38% and 22%, but remained at two-digit rates for nearly a decade. Our study takes advantage of the changes in enrolment of different education tracks that result from policy shifts. We contend that time and provincial enrolment adjustments act as a subtler version of educational reforms but a potentially powerful shifter of education choices,

with significant potential for econometric identification. We will discuss this view in more detail in Section 4.

3 Research design

Our study investigates whether there is a wage premium or penalty associated with secondary vocational education, when compared to academic education, in China. We also investigate whether the wage premium/penalty varies among people of heterogeneous earnings potential. Our econometric model extends the conventional Becker-Mincer wage equation by adding a dummy variable for whether an individual graduated from vocational or academic track before entering the labour market plus year and provincial fixed effects. Our model is as follows:

$$Y_i = \alpha_0 + \beta_1 VOC_i + \beta_2 S_i + \beta_3 Exp_i + \beta_4 Exp_i^2 + \beta_5 Female_i + \theta X_i + \mu_i, \tag{1}$$

where Y is the natural logarithm of the hourly wage of individual i; S denotes the years of schooling an individual completed. Job market experience are captured by the quadratics of working years, Exp_i. $Female_i$ is a dummy variable for gender. X_i contains survey year and province fixed effects. μ_i is the error term, which captures unobserved factors that affect wages. VOC_i is a dummy variable that denotes whether the highest educational qualification that an individual obtained is vocational (when it takes value 1) or academic (0). The coefficient for VOC, β_1, is the key parameter of interest in our study. The 'treatment' is individuals who finished secondary vocational education. In our study, we consider two alternative 'control' groups (individuals with the VOC_i dummy switched off): the first 'control' group contains only upper secondary academic graduates, and the second includes lower secondary graduates *plus* upper secondary academic graduates.

Given the fact that an individual is not randomly assigned to either vocational or academic education, we use instrumental variables as in the studies on the causality of education provision and individual schooling decision (e.g. Card (1993); Duflo (2001); Carneiro et al. (2018)). We use geographical and longitudinal variations in upper secondary academic and vocational enrolment quotas in relation to lower secondary graduation to instrument each individual's likelihood of attending vocational education. The enrolment data are matched to

the individual by the year and province that apply to each case. We use the two-stage least squares (2SLS) estimator to retrieve the mean effect of vocational education on hourly wages. We then apply the instrumental variable quantile regression (IV-QR) estimator to model heterogeneous vocational treatment effects for individuals with different earnings potential in the presence of endogeneity.

4 Data and variables

The data used in our study are pooled from nine cross-sectional data sets from China's General Social Survey (CGSS). Initiated in 2003, the CGSS is designed as annual survey on more than 10,000 households across China. The sampling strategy ensures that the data collected are nationally representative (Bian and Li, 2012). Questions include both subjective and objective items on individual characteristics and social networks. The surveys include Asian Social Survey and International Social Survey Program modules in the 2006 and 2008 waves, respectively Bian and Li (2012), which enable internationally comparative studies. CGSS data are widely used in sociology and economics research, including to inform policy making.

The study sample is kept to those of working age (16 to 65 years old) and in employment at the time of survey. Individuals whose hourly wage fell within the top and bottom one percentiles are dropped. In terms of education, we categorise the highest qualification into either vocational or academic education. The individual's years of schooling is reported in the 2006 and 2008 waves of CGSS data sets, while the highest educational qualifications are reported in all years. We use the reported years of schooling for 2006 and 2008. For the other years when the schooling year data are not available, we use the minimal years of schooling required for each educational qualification[2].

The data do not provide information on the individual's education history. As from upper secondary to tertiary level, individuals in one education track can progress to the other education track, we consider only the individuals who at most complete upper secondary education in this study for whom we are certain of their vocational/academic status. We also keep only non-migrant workers in the sample as migrants in the 2003-2008 waves of the CGSS data did not report their whereabouts in the year when they were expected to finish

[2]The years of schooling is calculated as the years required for certain educational qualifications by the Ministry of Education: primary education (6 years), junior secondary education (9 years), senior secondary education: academic or vocational (12 years), tertiary education: academic (16 years), tertiary education: vocational (15 years), postgraduate: master's (20 years), postgraduate: doctoral (24 years).

compulsory education, which would prevent us from constructing IVs for them. However, given the significant number of migrant workers in China, we also investigate the vocational effects on earnings for the 2010-2015 sub-samples of CGSS data, in which migrant workers' location information is available .

Table 1 provides the descriptive statistics of our study sample. The first column contains all observations in the study sample. The mean age of our observations is 43 years, with an average of 9 years of education. Female workers constitute 47% of the sample. 55% of respondents hold household registrations in rural areas (rural hukou). The majority, or 87%, in the sample are married. Members of the Communist Party make up 17% of all respondents. 44% are employed in agricultural sectors. The average job experience is 28 years, with mean hourly wages of CNY 7 yuan.[3] Columns (2)-(4) report the summary statistics for cohorts with different education qualifications as the highest attainment. In comparison to the sample average, the compulsory graduates are more likely to be paid lower hourly rates, be males and married, have rural hukou, and work in the agricultural sector while less likely to be members of the Communist Party.

(Table 1 here)

Individuals taking the academic track also differ significantly from vocational graduates, as shown in column (5). The secondary academic graduates in our samples are approximately 4 years older than their vocational counterparts. Secondary academic graduates on average have slightly less year of schooling, yet 5 years more job experience. Secondary academic graduates are more likely to hold rural hukou and to work in the agricultural sector. The summary statistics also show that secondary vocational graduates on average receive higher hourly wages than their academic counterparts. Figure 4 shows that higher earnings received by vocational graduates are observed at all quantiles of the unconditional wage distribution.

(Figure 4 here)

The instrumental variables used in this study are the ratios of the enrolment in upper secondary academic or vocational education over lower secondary graduation by province and year. Figure 5 shows the changes in the ratios between 1986 and 2007, as well as the proportion of upper secondary graduates in our study sample that entered the vocational track in particular years. The ratio for academic enrolment fluctuated between 1986 and 1997, followed by a steadily increase until 2007.

[3]See Duan and Martins (2020) for a recent study on wage determination in China using a different data set.

(Figure 5 here)

In contrast, the ratio of vocational enrolment over lower secondary graduation increased from 1986 to 1995 before it began to fall until 2000. The ratio remained constant in the first three years of the 21st century, and increased dramatically since 2004. The vertical bars show the proportions of upper secondary graduates in the study sample who entered the vocational track in each year. The change in the proportion of students generally coincides with the change in the vocational enrolment level. The proportion generally increased between 1996 and 1997, before it began to fall. It fluctuated between 1998 and 2003, and picked up again since.

Because the administration of secondary schools is largely the responsibility of provincial governments, there are also substantial provincial variations in the enrolment ratios. We provide an example of four provinces/metropolis to illustrate the correlation between the vocational enrolment and schooling choices, which is included in Figure 6. All four graphs indicates a potential correlation between vocational enrolment and individual's likelihood of choosing the vocational track.

(Figure 6 here)

We reckon that the ratio of upper secondary academic and vocational enrolment over lower secondary graduation would not only affect the track choice for the individuals who were to enter upper secondary education, but also influence the schooling decisions of all lower secondary graduates who decided between employment and upper secondary education. Figure 7 indicates a correlation between the ratio of upper secondary enrolment over lower secondary graduation *and* the proportion of lower secondary graduates that entered upper secondary vocational education nationally. Consistently, the graph indicates a correlation between enrolment ratios and individual's schooling decision.

(Figure 7 here)

The graphs for the whole nation and for four provinces/metropolis imply that the vocational enrolment level may be a potential shifter of individual schooling choice, consequently a relevant IV for schooling. In the next section, we will test the validity of the longitudinal and provincial variations in the academic and vocational enrolment ratios as IVs for individual's likelihood of graduating from vocational education in relation to graduating from academic education. We will also examine the likelihood of graduating from vocational education in

relation to graduating from academic education *and* completing only lower secondary education.

The IVs are matched to each individual by the year when and the province where an individual was due to progress from lower secondary education to upper secondary education. The greater the enrolment levels in a particular stream in a given year and province, the higher the availability of such option, and the more likely that the individual will choose it in his or her education. In addition, the enrolment at a higher level is also likely to affect those who tend to drop out of education after completing the lower level of education. Moreover, such availability will not have a direct effect on the wages of that individual (the exclusion restriction) as the latter will depend on the overall supply of graduates from different (younger and older) cohorts, when supply availability levels may have been very different.

5 Results

Here we provide empirical evidence of labour market returns to vocational education vis-à-vis academic education. We also investigate potential heterogeneity in relative returns among individuals with different earnings potential. We test the validity of academic and vocational school enrolment as instrumental variables for individual's likelihood of entering the vocational track, and provide both 2SLS estimates of wage differentials between vocational and academic qualification holders and the instrumental variable quantile regression (IV-QR) results of the vocational education effect across the wage distribution.

Columns (1), (4) and (7) in Table 2 contain the OLS estimates of the return to vocational education vis-à-vis academic education for the whole sample, the sub-sample of workers with rural household registration (hukou) and the sub-sample of workers with urban household registration, respectively. The OLS results for the whole sample show a wage premium of 20% is received by vocational graduates, which is statistically significant at the 1% level. For workers with urban hukou, vocational education is associated with 15% higher wages. For workers with rural hukou, there is no statistically significant earnings differential.

(Table 2 here)

The OLS results show a positive wage effect of vocational status at the average for the whole sample (and sub-samples). Given the heterogeneity in vocational education provision and in the labour market demand for vocational graduates as discussed in the previous

sections, we investigate the variations in the vocational effect on earnings along the conditional wage distribution. The quantile regression results are reported in Table 3. For the whole sample, the results indicate variations in the effect of vocational education on earnings. Specifically, the vocational premium ranges between 13% and 26% and generally decreases when moving up the conditional wage distribution. This is consistent with the findings by Balestra and Backes-Gellner (2017). The vocational premium at the bottom decile is 63% higher than that at the top decile.

(Table 3 here)

This pattern of positive but declining vocational premium is applicable to the sub-sample of workers with urban hukou. The largest vocational premium of 19% is received by the individuals at the 2nd decile of the conditional wage distribution, whereas the lowest, of 11%, is received by those at the 8th decile. Compared with the whole sample, the variation in the vocational wage premium in the urban sub-sample is smaller. In contrast, for the workers with rural hukou, the quantile regression results show no statistically significant earnings differential between academic graduates and vocational graduates at all deciles of the conditional wage distribution.

As already mentioned, individuals are not randomly assigned to either track of education. Earlier evidence indicates substantial selection, as students may self-select based on unobserved characteristics like preferences or motivation, or may also be screened into either vocational or academic education subject to their performance in upper secondary entrance exams. Excluding this information may bias the OLS estimates. In the remaining of the section, we address the issue of selection in estimation using instrumental variable approaches.

For the average vocational treatment effect, Columns (2), (5) and (8) of Table 3 report the 2SLS estimates and the first-stage results for the whole sample, the urban sub-sample and the rural sub-sample, respectively. In these cases, we use the ratio of vocational enrolment over lower secondary graduation as the single IV. For the whole sample, the first-stage results in Column (2) supports to the validity of the IV. Specifically, the marginal effect of an increase in the ratio of vocational enrolment over lower secondary graduates (in the relevant province and year) is positive and statistically significant at the 1% level. The 2SLS estimates of the return to vocational education in comparison to academic education is not significantly different from zero. For the urban sub-sample, the marginal effect of vocational enrolment

level is larger than for the whole sample, and is statistically significant. Similar to the whole sample, the 2SLS estimate for the urban sample shows no significant earnings differential between academic and vocational graduates.

In addition, we add the ratio of academic enrolment over lower secondary graduation as a second IV. The first-stage results show that only the vocational enrolment acts as a relevant IV, for the whole sample and the urban sample. The 2SLS estimates of the vocational effect are consistent with the model with a single IV. Our 2SLS results show no significant earnings difference between vocational and academic graduates. In comparison to the OLS results of positive vocational premium, there is evidence of positive selection into vocational education for our sample of those who at most completed secondary education. The individuals with higher unobserved ability, for instance those that are highly motivated, are more likely to enter the vocational track than their counterparts with lower unobserved ability. Although this seems to contradict the idea that students with lower ability tend to enter vocational tracks, recall that our findings are only applicable to those who finished only secondary education. We reckon for this cohort upper secondary vocational education provides ready-to-use skills, occupationally-relevant knowledge and attitudes, which make secondary vocational graduates more successful in the labour market in comparison to secondary academic graduates.

When it comes to the heterogeneity in the premiums, our QR results support the variation in the effects of vocational education. Similar to OLS estimates, the QR estimates are also subject to endogeneity bias. What's more, the magnitude of the bias may vary if the selection pattern is not constant along the conditional wage distribution. In other words, the variation in the vocational effect may result from changing selection bias, instead of heterogeneous effects of vocational education.

To address this issue of selection in quantile regression, we apply the IV-QR estimator(Chernozhukov and Hansen, 2005, 2008). The results are reported in Table 4. Our estimates based on the whole sample and the urban sample show that after addressing selection, vocational education generates earnings premium for individuals at the 5th and the 6th deciles. Specifically, the vocational premium is between 31% and 37% for the whole sample, and between 30% and 34% for the urban sub-sample. On the other hand, for individuals at both the lower and upper deciles of the distribution, there is no wage differential between vocational graduates and academic graduates.

(Table 4 here)

To summarise, our OLS results show that for the whole sample (and also the urban subsample), secondary vocational graduates earn a wage premium compared to their academic counterparts. This premium applies both on average and across the conditional wage distribution (in which case it appears to decrease along the distribution).

However, our evidence also indicates that these results are partly attributable to the positive selection into the vocational track. Once we address this selection using our instrumental variables, we find that there is neither a wage premium nor a penalty associated with vocational education. This is the case on average and also for those with lower or higher earnings potential. However, for individuals with average earnings potential, vocational education leads to 30% - 37% higher hourly wages.

6 Robustness checks

6.1 Compulsory graduates

Earlier in this section, we compared the earnings profile of the upper secondary vocational graduates against their academic counterparts as the academic track is an alternative to the vocational track. In addition to the secondary academic track, another alternative for the would-be vocational graduates may be entering the labour market after completing lower secondary education (compulsory education). Given the fact that lower secondary education shares similar orientation and pedagogy with upper secondary academic education, we use both the lower secondary graduates and the upper secondary academic graduates as potential counterfactuals (vocational dummy equal to zero) and estimate the earnings effect of vocational education.

The IVs used for the vocational status (as opposed to academic and compulsory status) remain the ratios of upper secondary academic and vocational enrolment over lower secondary graduation. We reckon that the change in academic enrolment in relation to vocational enrolment may affect the availability of places in either track of education, which consequently influence individual's choice of educational tracks. The changes in upper secondary enrolment in relation to lower secondary graduation may also increase or decrease an individual's chance of being admitted into upper secondary education. Given the fact that the vocational admission happens after the academic admission and provides a 'final' chance for lower secondary

graduates to stay in school, an increase in vocational enrolment in relation to lower secondary graduation may encourage more people who would otherwise drop out of schools to attend vocational education. This correlation has been displayed in Figure 7 in Section 4.

Table 5 contains the OLS and 2SLS estimates of the average effect of vocational education on earnings. The OLS results in columns (1), (4) and (7) show that vocational education is associated with wage premiums of 20%, 16% and 10% for the whole sample, the urban sub-sample and the rural sub-sample, respectively. All are statistically significant at the 1% level. The QR results in Table 6 show that the vocational premium ranges between 15% and 27% for the whole sample, and between 11% and 19% for the urban sample. Again, the vocational premium is larger at lower deciles. For the rural sample, the vocational premium is only significant at the 6th decile.

(Table 5 here)

(Table 6 here)

When addressing the issue of selectivity, our first-stage results show that the vocational enrolment ratio has a positive effect on an individual's likelihood of attending vocational education for the whole sample and the urban sub-sample, in a relationship with economic and statistical significance. The 2SLS estimates with the vocational enrolment ratio as an IV indicate that there is no wage premium for upper secondary vocational graduates in comparison to compulsory and upper secondary academic graduates. In terms of heterogeneity in the vocational effect, the IV-QR results in Table 7 show that the urban workers at the 7th decile receive 18% higher wages compared to their academic counterparts.

(Table 7 here)

6.2 Migrant workers

Given the substantial share of internal migrant workers in the Chinese labour force, we perform additional analysis by including these individuals as well. In the previous analysis, we use all waves of CGSS survey data from 2003 to 2015. However, the data from 2003 to 2008 do not provide information on the source provinces of migrant workers, which makes it infeasible for this additional analysis. We use instead the data collected between 2010 and 2015.

Migrant workers here are defined as only inter-provincial migrants, namely those whose province of birth differs from the province of current residence/work. Although the number

of workers moving within a province is substantial, the data do not allow us to trace such migration. The province where a migrant worker completed compulsory education is determined by comparing the year when the individual was due to complete compulsory education and the year of last migration (the data on the year of first migration or migrant patterns are not available). Specifically, if a migrant worker first migrated before he/she was due to complete compulsory education, the host province is assigned as the province where he/she completed compulsory education; otherwise, the source province is used.

We modify the econometric model by including a migrant dummy. As in our main results, the vocational status equals 0 if an individual graduated from upper secondary academic education and 1 if from upper secondary vocational education. The OLS and 2SLS results are reported in Table 8. Column (1) shows that the OLS estimate of the return to vocational education vis-à-vis academic education is 11%. The first-stage results in Column (2) and (3) show again a statistically significant correlation between the vocational enrolment level and individual's likelihood of attending vocational education. However, the model is weakly identified after the inclusion of migrant workers. We reckon that it is plausible that a migrant worker might move between provinces at school age. Without the data on first migration and migrant pattern, our IVs and identification strategy are not suitable for migrant workers.

(Table 8 here)

7 Conclusion

As China progresses in its economic development, greater attention should be paid to its education system and its interplay with the labour market. This paper contributes to this research agenda by estimating the returns to secondary vocational education vis-á-vis academic education. In addition to estimating the average effect of vocational education on hourly wages, we investigate the heterogeneity in the vocational effect for individuals with different earnings potential, as indicated by their standing in the conditional wage distribution.

Our results, using China's General Social Survey, show there is significant vocational wage premium both on average and at each decile of the conditional wage distribution. However, this premium is largely driven by positive selection into the vocational track. We reckon that, for those who at most completed secondary education, the individuals with higher unobserved ability, e.g. motivation, prefer vocational education over academic education. Our IV results

show there is no statistically significant earnings differential between secondary vocational graduates and secondary academic graduates. However, we also find that workers with average earnings potential receive a 30%-34% vocational premium. When extending the comparison to both upper and lower secondary graduates (and partialling out the effect of years of schooling), we find an average 18% premium.

The evidence of the return to vocational education can inform students and their families in schooling decisions. Our findings show that vocational education benefits students as much as academic education for the individuals who at most complete secondary education. Vocational education proves to be a more rewarding alternative for individuals with average earnings potential. For individuals with either low or high earnings potential, we do not find evidence of a similar premium. We also note that, although the human capital investment through formal education matters in labour market outcomes, there are a substantial proportion of earnings differentials that could still not be explained by education qualifications alone.

Despite this, our study provides evidence again the conventional idea that vocational education is inferior to academic education, as least in our study sample. Secondary academic education acts largely as a transitioning stage between compulsory education and tertiary education. In contrast, secondary vocational education possesses a clear objective of preparing its graduates for a particular occupation or trade. The knowledge, skills, and attitude acquired through secondary vocational education enable an uninterrupted shift from school to work and a smooth transition into adulthood. This feature of vocational education is important in a context of youth unemployment and 'NEETs'.

In this study, we also document a correlation between enrolment quotas and individual schooling decisions. Our first-stage results show that the ratio of the upper secondary vocational enrolment over the number of lower secondary graduates acts as a valid IV for the individual's likelihood of entering the vocational track. An increase in vocational enrolment involves more slots, or opportunities, to attend vocational education for those who could otherwise choose the academic track, as well as for the lower secondary graduates who would otherwise enter the labour market directly. This is consistent to earlier findings in China that the government can adjust the skill structure of its workforce or provide incentives for individuals to stay in formal education by changing the number of enrolment in different types of education. This practice may also affect the labour market equilibrium in the short-run, as in

the case of university expansion after the 1997 Asian financial crisis. Given the vast number of students in China, the policy leverage of enrolment adjustment bears significant social and economic implications.

References

Arias, O., Hallock, K. F., and Sosa-Escudero, W. (2002). Individual heterogeneity in the returns to schooling: instrumental variables quantile regression using twins data. In *Economic Applications of Quantile Regression*, pages 7–40. Springer.

Balestra, S. and Backes-Gellner, U. (2017). Heterogeneous returns to education over the wage distribution: Who profits the most? *Labour Economics*, 44:89–105.

Bian, Y. and Li, L. (2012). The chinese general social survey (2003-8) sample designs and data evaluation. *Chinese Sociological Review*, 45(1):70–97.

Brunello, G. and Rocco, L. (2015). The effects of vocational education on adult skills and wages. *OECD ilibrary*, (168).

Cappellari, L. (2004). High school types, academic performance and early labour market outcomes. *SSRN working paper*, (1048).

Card, D. (1993). Using geographic variation in college proximity to estimate the return to schooling. Working Paper 4483, National Bureau of Economic Research.

Carneiro, P. M., Liu, K., and Salvanes, K. G. (2018). The supply of skill and endogenous technical change: evidence from a college expansion reform. *NHH Department of Economics Discussion Paper*, (16).

Chernozhukov, V. and Hansen, C. (2002). Inference for distributional effects using instrumental quantile regression. *MIT Department of Economics Working Paper*, (02-20).

Chernozhukov, V. and Hansen, C. (2005). An IV model of quantile treatment effects. *Econometrica*, 73(1):245–261.

Chernozhukov, V. and Hansen, C. (2008). Instrumental variable quantile regression: A robust inference approach. *Journal of Econometrics*, 142(1):379–398.

Dai, F., Cai, F., and Zhu, Y. (2018). Returns to Higher Education in China: Evidence from the 1999 Higher Education Expansion Using Fuzzy Regression Discontinuity. *IZA Discussion Paper*, (11735).

Dearden, L., Ferri, J., and Meghir, C. (2002). The effect of school quality on educational attainment and wages. *Review of Economics and Statistics*, 84(1):1–20.

Doksum, K. (1974). Empirical probability plots and statistical inference for nonlinear models in the two-sample case. *Annals of Statistics*, pages 267–277.

Duan, W. and Martins, P. S. (2020). Rent sharing in China: Magnitude, heterogeneity and drivers. GLO Discussion Paper Series 448.

Duflo, E. (2001). Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment. *American Economic Review*, 91(4):795–813.

Eichhorst, W., Rodríguez-Planas, N., Schmidl, R., and Zimmermann, K. F. (2015). A road map to vocational education and training in industrialized countries. *Industrial and Labor Relations Review*, 68(2):314–337.

Fersterer, J., Pischke, J.-S., and Winter-Ebmer, R. (2008). Returns to apprenticeship training in Austria: Evidence from failed firms. *Scandinavian Journal of Economics*, 110(4):733–753.

Hannum, E., An, X., and Cherng, H.-Y. S. (2011). Examinations and educational opportunity in China: Mobility and bottlenecks for the rural poor. *Oxford Review of Education*, 37(2):267–305.

Hansen, M. H. and Woronov, T. E. (2013). Demanding and resisting vocational education: a comparative study of schools in rural and urban China. *Comparative Education*, 49(2):242–259.

Hanushek, E. A. and Wössmann, L. (2006). Does educational tracking affect performance and inequality? Differences-in-differences evidence across countries. *Economic Journal*, 116(510):C63–C76.

Hutchinson, J. and Kettlewell, K. (2015). Education to employment: complicated transitions in a changing world. *Educational Research*, 56(1):114–115.

Jäger, M. (2016). Dual vocational education and training as an option in development cooperation,survey of experts on behalf of the donor committee for dual vocational education and training. Final report, Zurich: dcDVET.

Jiang, Q. and Xu, G. (2005). Project-based curriculum research of vocational education based on tasks. *Vocational and Technical Education*, 22.

Koenker, R. and Bassett Jr, G. (1978). Regression quantiles. *Econometrica*, pages 33–50.

Li, Y. A., Whalley, J., Zhang, S., and Zhao, X. (2012). The higher educational transformation of China and its global implications. In *The Globalization of Higher Education*, pages 135–162. Springer.

Lin, J. Y., Wang, G., and Zhao, Y. (2004). Regional inequality and labor transfers in China. *Economic Development and Cultural Change*, 52(3):587–603.

Ling, M. (2015). Bad students go to vocational schools!?: Education, social reproduction and migrant youth in urban China. *The China Journal*, (73):108–131.

Martins, P. S. and Pereira, P. T. (2004). Does education reduce wage inequality? Quantile regression evidence from 16 countries. *Labour Economics*, 11(3):355–371.

McIntosh, S. and Morris, D. (2016). Labour market returns to vocational qualifications in the labour force survey. *CVER Research Paper*, 6.

Meer, J. (2007). Evidence on the returns to secondary vocational education. *Economics of Education Review*, 26(5):559–573.

OECD (2010). *Learning for Jobs OECD Reviews of Vocational Education and Training Options for China*. OECD, Paris, France.

Oosterbeek, H. and Webbink, D. (2007). Wage effects of an extra year of basic vocational education. *Economics of Education Review*, 26(4):408–419.

O'Reilly, J., Eichhorst, W., Gábos, A., et al. (2015). Five characteristics of youth unemployment in Europe: Flexibility, education, migration, family legacies, and EU policy. *Sage Open*, 5(1):1–19.

Ota, M. and Moffatt, P. G. (2007). The within-household schooling decision: a study of children in rural Andhra Pradesh. *Journal of Population Economics*, 20(1):223–239.

Pepper, S. (2000). *Radicalism and education reform in 20th-century China: The search for an ideal development model*. Cambridge University Press.

Psacharopoulos, G. and Patrinos, H. A. (2018). Returns to investment in education: A decennial review of the global literature. *World Bank Policy Research Working Paper*, (8402).

Tsang, M. C. (2000). Education and national development in China since 1949: Oscillating policies and enduring dilemmas. *China Review*, pages 579–618.

Wilson, K., Wolfe, B., and Haveman, R. (2005). The role of expectations in adolescent schooling choices: Do youths respond to economic incentives? *Economic Inquiry*, 43(3):467–492.

Wolf, A. (2011). Review of vocational education. *London: DfE*.

Zheng, C. (2018). *Study on the evolution of China's entrance examination and enrolment policy (1977-2017)*. PhD dissertation, Tianjin Normal University.

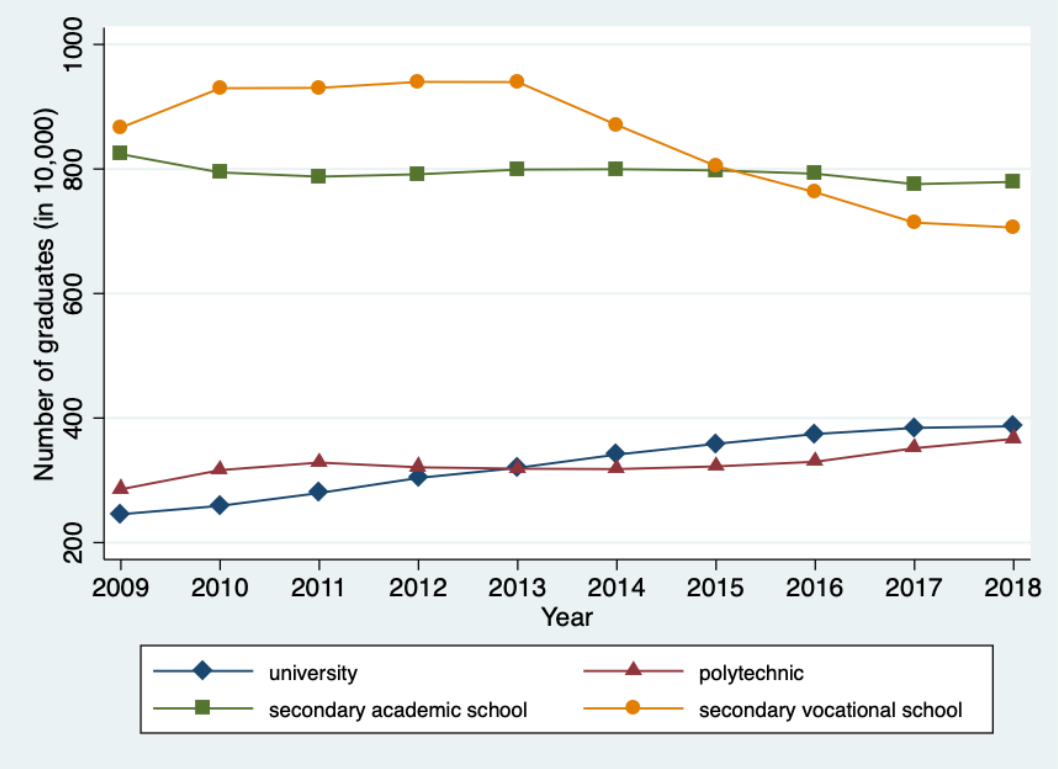

Figure 1: Number of graduates from different tracks and levels of education over the year
Note: the figure shows the number of graduates from academic or vocational education at the secondary or tertiary levels between 2009 and 2018. Data source: National Bureau of Statistics. Compiled and graphed by authors.

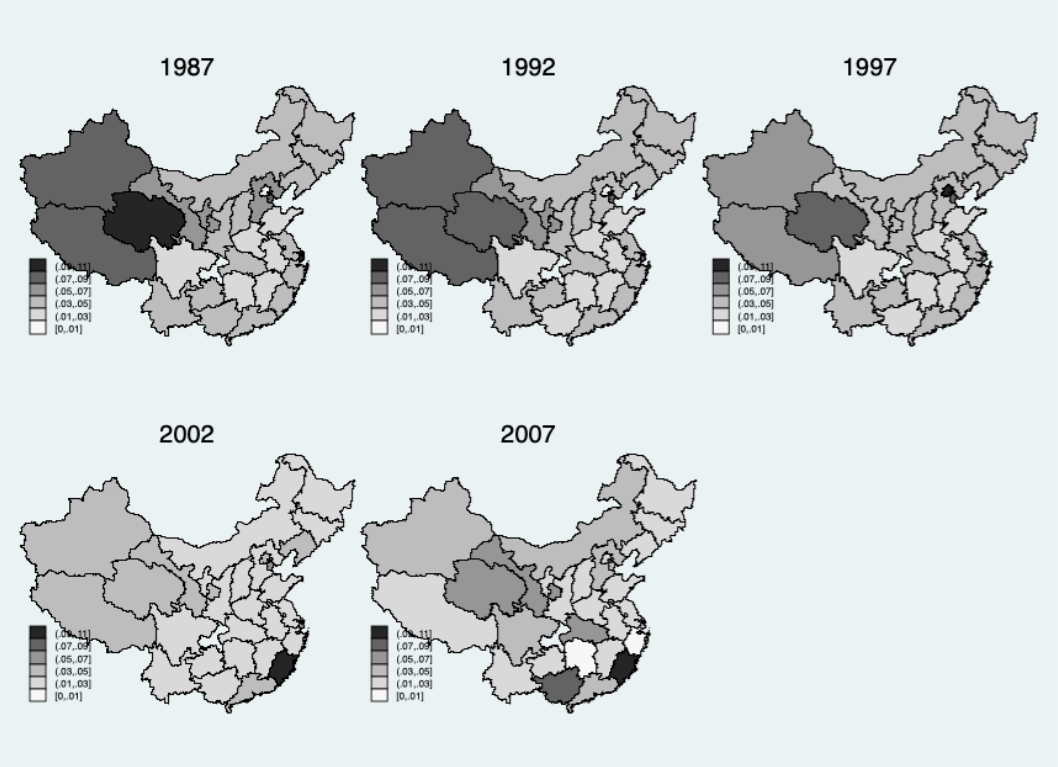

Figure 2: Density of secondary vocational school across China
Note: the figure shows the density of secondary vocational schools (SVSs) by population across Chinese province between 1987 and 2007. The density is calculated as the number of SVSs for every 10 thousand residents. Data source: China Educational Yearbook (1987-2007). Compiled and graphed by authors.

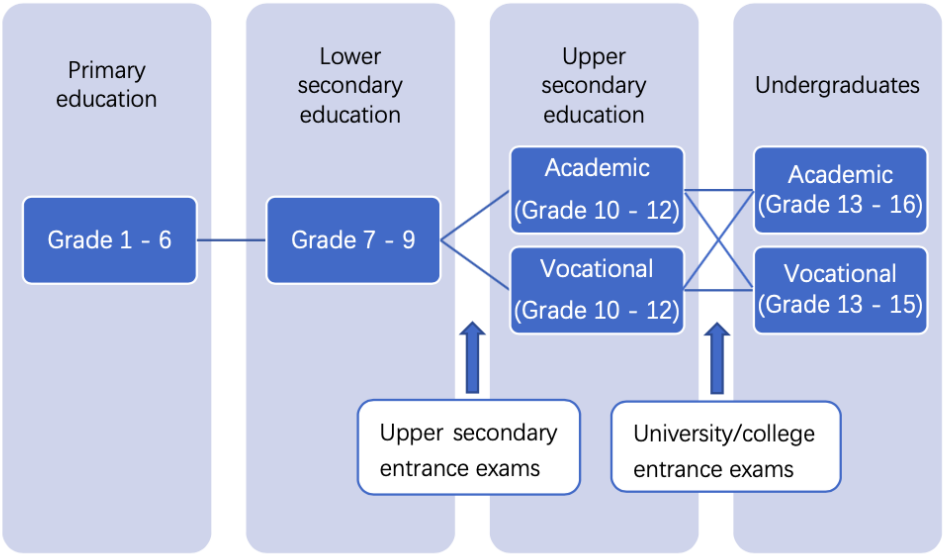

Figure 3: Structure of Chinese education system administered by Ministry of Education
Note: the figure shows the education system administered by Ministry of Education in China. The postgraduate level is not shown here. The track of education diverse at upper secondary level. Both secondary vocational and academic education graduates can choose between tertiary vocational and academic education. Student application, performance at entrance exams and enrolment quota set by schools together decide admission outcome. Ministry of Human Resource and Social Security administers another track for adult education and training.

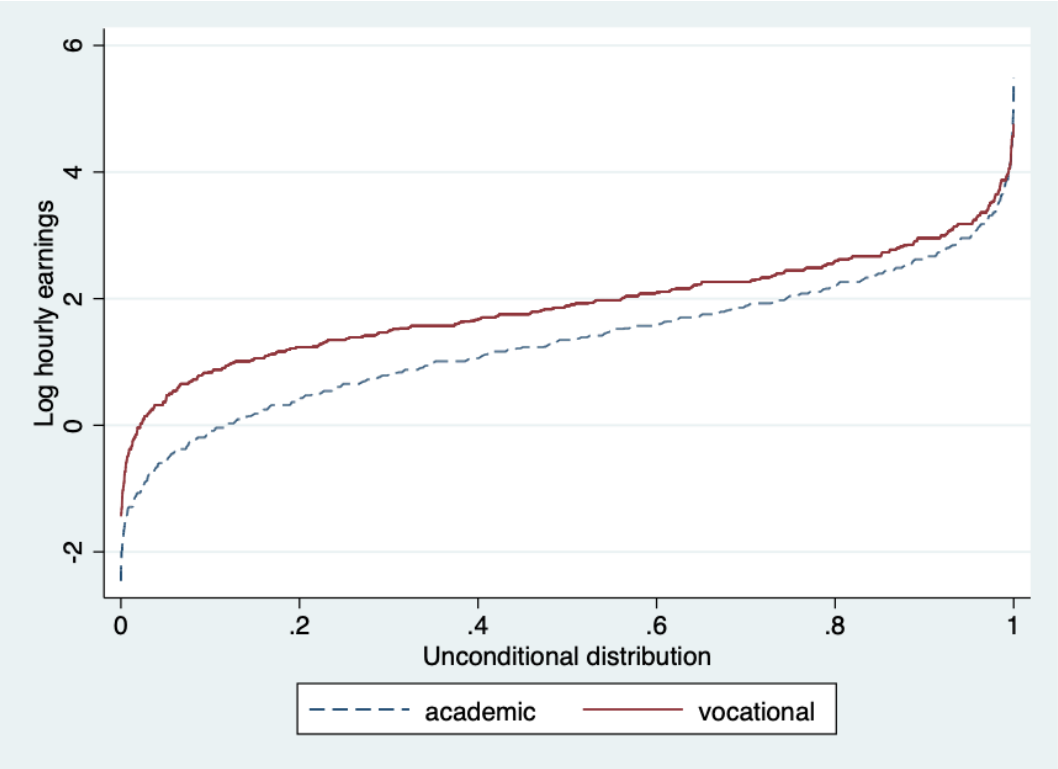

Figure 4: Unconditional distribution of log hour wage by educational track
Note: the figure shows the distribution of the natural logarithm of annual wage along unconditional wage quantiles by educational track. Data source: China General Social Survey (2003-2015). Compiled by authors.

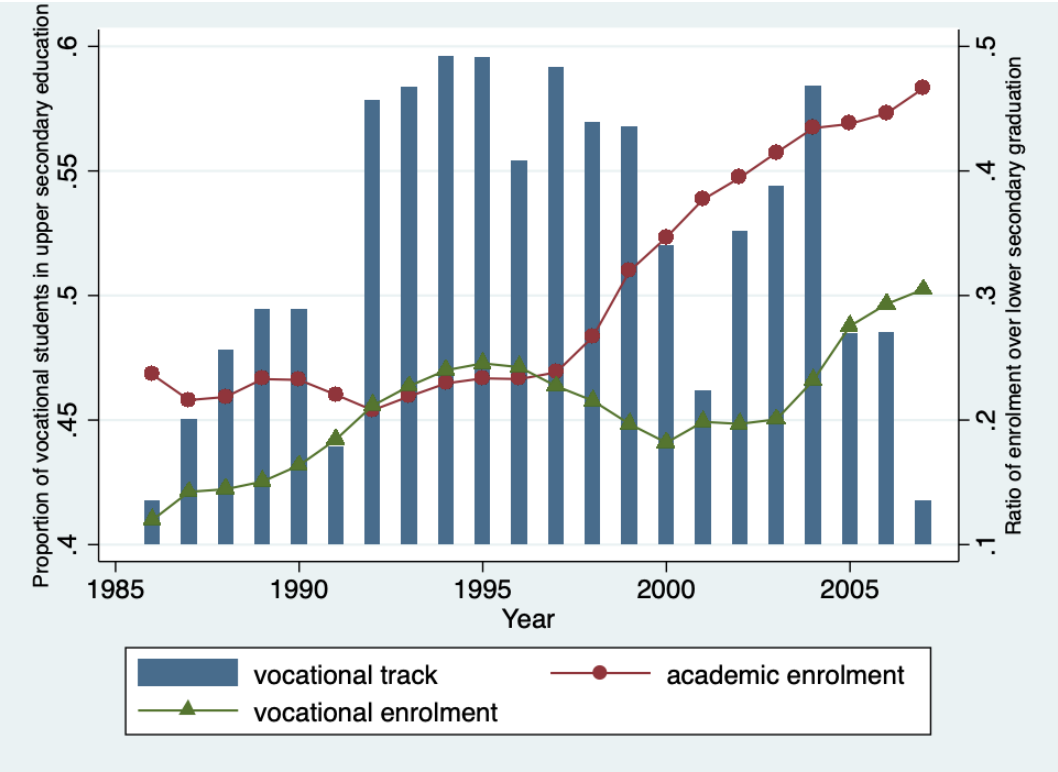

Figure 5: Enrolment quota and individual schooling decision (upper secondary)
Note: The figure shows the ratios of the enrolment quota by upper secondary academic and vocational schools over the number of graduates of lower secondary education between 1986 and 2007, and the proportion of upper secondary students in the vocational track. The data are aggregated at national level. Data source: China Educational Yearbook (1987-2008) and China General Social Survey (2003-2005). Compiled and graphed by authors.

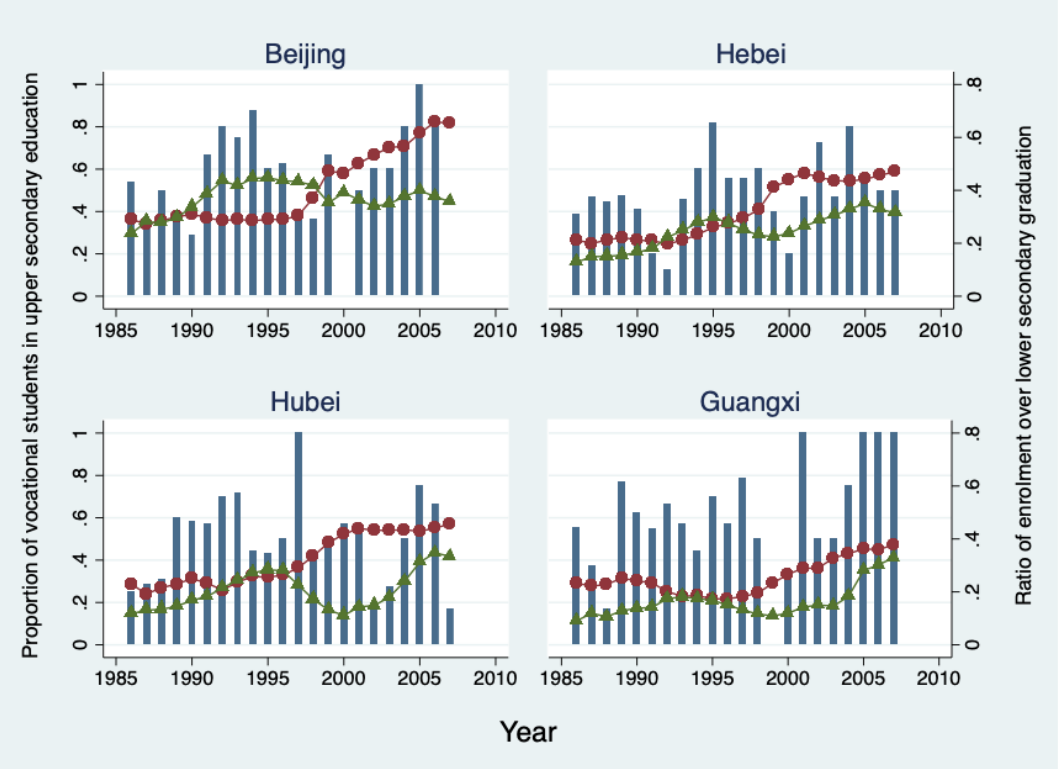

Figure 6: QR and IVQR estimates

Note: The figure shows the ratios of the enrolment quota by upper secondary academic and vocational schools over the number of graduates of lower secondary education between 1986 and 2007, and the proportion of upper secondary students in the vocational track in each province. We choose four provinces/metropolis in different stages of social and economic development, which include Beijing (the capital city), Hebei and Hunan (central provinces), and Guangxi (western province). The data are aggregated at national level. Data source: China Educational Yearbook (1987-2008) and China General Social Survey (2003-2005). Compiled and graphed by authors.

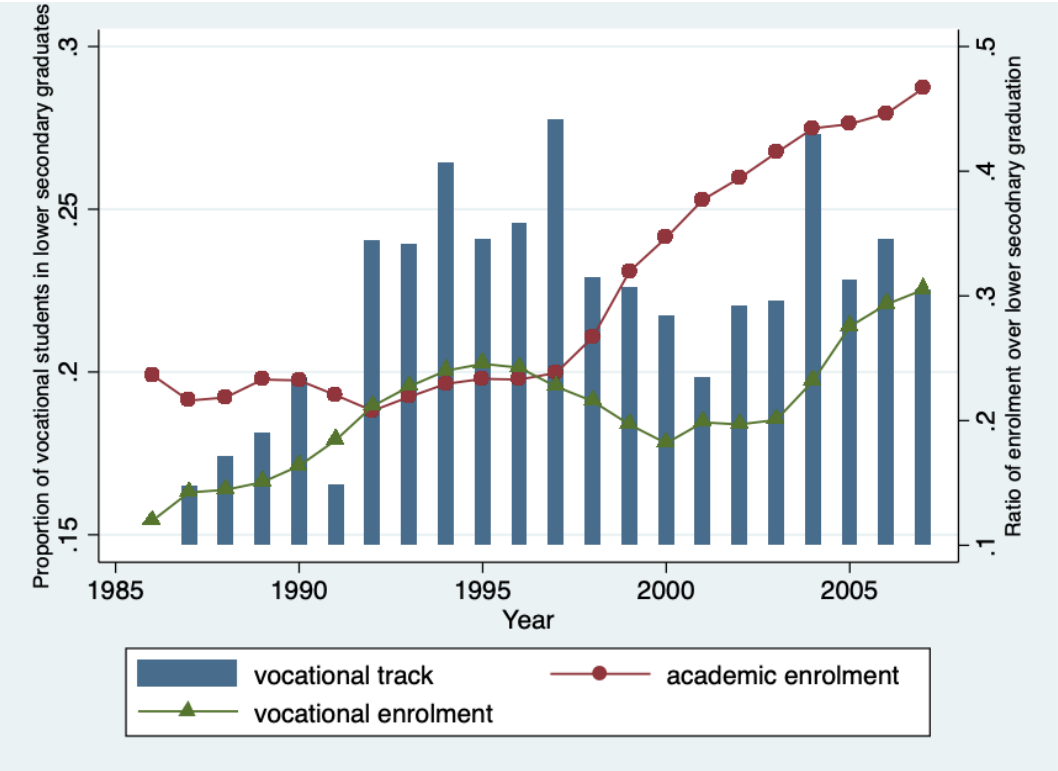

Figure 7: Enrolment quota and individual schooling decision (overall secondary)
Note: The figure shows the ratios of the enrolment quota by upper secondary academic and vocational schools over the number of graduates of lower secondary education between 1986 and 2007, and the proportion of lower secondary graduates that entered the vocational track. The data are aggregated at national level. Data source: China Educational Yearbook (1987-2008) and China General Social Survey (2003-2005). Compiled and graphed by authors.

Table 1: Summary statistics by educational track

	(1) All		(2) Compulsory		(3) Academic		(4) Vocational		(5) (3)-(4)	
	mean	sd	mean	sd	mean	sd	mean	sd	b	t
Age	43.24	12.00	44.41	11.66	41.94	11.94	37.71	12.54	4.23***	(17.67)
Female	0.47	0.50	0.48	0.50	0.43	0.50	0.45	0.50	-0.02	(-1.79)
Hourly salary	6.57	9.15	5.60	8.54	8.33	10.24	9.34	9.66	-1.01***	(-4.36)
Job experience	28.33	12.75	30.75	12.12	24.09	11.99	19.75	12.56	4.34***	(18.08)
Year of schooling	8.91	2.42	7.66	1.70	11.85	0.60	11.96	0.70	-0.11***	(-8.45)
Rural hukou	0.55	0.50	0.67	0.47	0.30	0.46	0.22	0.41	0.08***	(9.90)
Married	0.87	0.34	0.90	0.30	0.81	0.39	0.76	0.43	0.05***	(6.49)
Communist's party	0.17	0.38	0.16	0.37	0.20	0.40	0.21	0.40	-0.00	(-0.21)
Agricultural sectors	0.44	0.50	0.55	0.50	0.19	0.39	0.08	0.27	0.11***	(14.23)
Observations	39,217		27,655		7,443		4,119		11,562	

Note: the table shows descriptive statistics by the track of education and the t-test results. The compulsory column show the summary statistics for individuals with compulsory or lower education. The columes for academic and vocational are for those who completed secondary academic or vocational education, respectively. T-test is performed on the statistics for academic and vocational graduates. Age is averaged by the mean. The proportion of female, married, education qualificaitons, communist, non-agricultural hukou are reported. The nominal value of hourly wage are reported with unit of one Chinese Yuan (CNY). The significance levels of t-test are * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Data source: 2003-2015 Chinese General Social Survey.

Table 2: OLS and 2SLS results

	Whole sample			Urban hukou			Rural hukou		
	OLS	IV	IV	OLS	IV	IV	OLS	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Vocational education	0.20	-0.22	-0.23	0.15	-0.03	-0.05	0.07	-1.02	-1.06
	(0.02)***	(0.41)	(0.41)	(0.02)***	(0.36)	(0.36)	(0.05)	(2.18)	(2.20)
Year of schooling	0.07	0.08	0.08	0.06	0.07	0.07	0.05	0.11	0.11
	(0.01)***	(0.03)***	(0.03)***	(0.01)***	(0.03)***	(0.03)***	(0.04)	(0.07)	(0.07)
Experience	0.02	0.03	0.03	0.02	0.03	0.03	0.02	-0.00	-0.00
	(0.00)***	(0.01)***	(0.01)***	(0.00)***	(0.01)***	(0.01)***	(0.01)***	(0.02)	(0.02)
Experience^2	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
	(0.00)***	(0.00)*	(0.00)*	(0.00)***	(0.00)*	(0.00)*	(0.00)***	(0.00)	(0.00)
Female	-0.19	-0.20	-0.20	-0.20	-0.19	-0.19	-0.34	-0.31	-0.31
	(0.02)***	(0.03)***	(0.03)***	(0.02)***	(0.03)***	(0.03)***	(0.04)***	(0.09)***	(0.09)***
Province F.E.	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year F.E.	YES	YES	YES	YES	YES	YES	YES	YES	YES
First-stage									
Vocational enrolment		0.70	0.71		0.90	0.92		0.31	0.31
		(0.19)***	(0.20)***		(0.23)***	(0.24)***		(0.36)	(0.36)
Academic enrolment			0.03			0.06			0.00
			(0.18)			(0.22)			(0.40)
Kleibergen-Paap LM statistics		13.49	13.51		15.14	15.22		0.79	0.79
P> Chi^2(LM)		0.00	0.00		0.00	0.00		0.38	0.68
Kleibergen-Paap Wald F statistics		13.37	6.69		14.96	7.52		0.75	0.37
Hanson J statistics			0.42			0.66			2.66
P> Chi^2(Hanson J)		.	0.52		.	0.42		.	0.10
Observations	7,691	3,271	3,271	5,784	2,360	2,360	1,905	909	909

Note: The table includes the OLS and 2SLS estimates and the first-stage results for the whole sample and the sub-samples of workers with urban hukou and of workers with rural hukou. Vocational education is a dummy variable with 1 for upper secondary vocational graduates and 0 for upper secondary academic graduates. The instrumental variables are a) the ratio of the upper secondary vocational enrolment number over the lower secondary graduate number (vocational enrolment), and b) the ratio of the upper secondary academic enrolment number over the lower secondary graduate number (academic enrolment). Standard errors in parentheses. Significance level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Data source: China General Social Survey 2003-2015.

Table 3: Quantile regression results

	Whole sample		Urban hukou		Rural hukou	
	Coefficient	95% CI	Coefficient	95% CI	Coefficient	95% CI
	(1)	(2)	(3)	(4)	(5)	(6)
1st decile	0.26	[0.18,0.34]	0.18	[0.10,0.26]	0.01	[-0.17,0.19]
2nd decile	0.22	[0.16,0.28]	0.19	[0.14,0.25]	0.05	[-0.09,0.18]
3rd decile	0.20	[0.16,0.25]	0.18	[0.13,0.23]	0.10	[-0.02,0.22]
4th decile	0.19	[0.15,0.24]	0.17	[0.12,0.21]	0.06	[-0.05,0.18]
5th decile	0.19	[0.15,0.23]	0.16	[0.11,0.20]	0.07	[-0.04,0.18]
6th decile	0.18	[0.14,0.23]	0.15	[0.11,0.20]	0.05	[-0.06,0.16]
7th decile	0.17	[0.12,0.21]	0.14	[0.09,0.19]	0.05	[-0.06,0.17]
8th decile	0.13	[0.08,0.18]	0.11	[0.06,0.15]	0.07	[-0.06,0.21]
9th decile	0.16	[0.10,0.21]	0.14	[0.08,0.19]	0.08	[-0.08,0.24]
Observations	7,691		5,784		1,905	

Note: The table contains the quantile regression estimates of the vocational status (vis-á-vis the academic status) on earnings. Vocational education is a dummy variable with 1 for upper secondary vocational graduates and 0 for upper secondary academic graduates. 95% confidence intervals in brackets. Data source: China General Social Survey 2003-2015.

Table 4: Instrumental variable quantile regression results

	Whole sample		Urban hukou		Rural hukou	
	Coefficient	95% CI	Coefficient	95% CI	Coefficient	95% CI
	(1)	(2)	(3)	(4)	(5)	(6)
1st decile	0.09	[-0.25,0.56]	0.04	[-0.48,0.45]	-0.44	[-1.08,0.26]
2nd decile	0.13	[-0.36,0.41]	0.10	[-0.37,0.50]	-0.33	[-0.97,0.27]
3rd decile	0.14	[-0.25,0.39]	0.10	[-0.21,0.45]	-0.24	[-0.95,0.16]
4th decile	0.24	[-0.11,0.50]	0.27	[-0.08,0.43]	-0.21	[-0.88,0.14]
5th decile	0.37	[0.14,0.56]	0.34	[0.06,0.44]	0.01	[-0.64,0.65]
6th decile	0.31	[0.02,0.54]	0.30	[0.04,0.48]	0.15	[-0.83,0.57]
7th decile	0.18	[-0.15,0.46]	0.19	[-0.01,0.42]	0.06	[-1.06,0.64]
8th decile	0.09	[-0.07,0.45]	0.22	[-0.11,0.48]	-0.10	[-0.85,0.55]
9th decile	0.11	[-0.16,0.57]	0.07	[-0.12,0.53]	-0.62	[-1.11,0.33]
Observations	3,271		2,360		909	

Note: The table contains the instrumental variable quantile regression estimates of the vocational status (vis-á-vis the academic status) on earnings. Vocational education is a dummy variable with 1 for upper secondary vocational graduates and 0 for upper secondary academic graduates. The instrumental variable is for the endogenous vocational status is the ratio of the number of upper secondary vocational enrolment over the number of lower secondary graduates. 95% confidence intervals in brackets. Data source: China General Social Survey 2003-2015.

Table 5: OLS and 2SLS results for robustness check

	Whole sample			Urban hukou			Rural hukou		
	OLS	IV	IV	OLS	IV	IV	OLS	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Vocational education	0.20	-0.26	-0.26	0.16	0.05	0.03	0.10	-0.75	-1.10
	(0.02)***	(0.55)	(0.55)	(0.02)***	(0.42)	(0.42)	(0.04)**	(2.19)	(2.23)
Year of schooling	0.09	0.18	0.18	0.08	0.11	0.12	0.04	0.19	0.24
	(0.00)***	(0.09)**	(0.09)**	(0.01)***	(0.07)*	(0.07)*	(0.01)***	(0.32)	(0.33)
Experience	0.02	0.03	0.03	0.01	0.04	0.04	0.01	0.01	0.01
	(0.00)***	(0.01)***	(0.01)***	(0.00)***	(0.01)***	(0.01)***	(0.00)***	(0.02)	(0.02)
Experience^2	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
	(0.00)***	(0.00)***	(0.00)***	(0.00)***	(0.00)***	(0.00)***	(0.00)***	(0.00)	(0.00)
Female	-0.26	-0.29	-0.29	-0.24	-0.24	-0.24	-0.40	-0.38	-0.38
	(0.01)***	(0.02)***	(0.02)***	(0.01)***	(0.02)***	(0.02)***	(0.02)***	(0.03)***	(0.03)***
Province F.E.	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year F.E.	YES	YES	YES	YES	YES	YES	YES	YES	YES
First-stage									
Vocational enrolment		0.39	0.39		0.64	0.65		0.15	0.15
		(0.09)***	(0.09)***		(0.16)***	(0.16)***		(0.09)	(0.09)
Academic enrolment			-0.01			0.04			0.02
			(0.08)			(0.14)			(0.09)
Kleibergen-Paap LM statistics		19.74	19.77		17.32	17.39		2.70	2.73
P> Chi^2(LM)		0.00	0.00		0.00	0.00		0.10	0.26
Kleibergen-Paap Wald F statistics		19.68	9.85		17.20	8.63		2.67	1.35
Hanson J statistics			0.05			0.72			2.25
P> $Chi^2(HansonJ)$			0.82			0.40			0.13
Observations	18,116	7,537	7,537	10,144	3,660	3,660	7,965	3,871	3,871

Note: The table includes the OLS and 2SLS estimates and the first-stage results for the whole sample and the sub-samples of workers with urban hukou and of workers with rural hukou. Vocational education is a dummy variable with 1 for upper secondary vocational graduates and 0 for lower secondary graduates *plus* upper secondary academic graduates. The instrumental variables are a) the ratio of the upper secondary vocational enrolment number over the lower secondary graduate number (vocational enrolment), and b) the ratio of the upper secondary academic enrolment number over the lower secondary graduate number (academic enrolment). Standard errors in parentheses. Significance level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Data source: China General Social Survey 2003-2015.

Table 6: Quantile regression results for robustness check

	Whole sample		Urban hukou		Rural hukou	
	Coefficient	95% CI	Coefficient	95% CI	Coefficient	95% CI
	(1)	(2)	(3)	(4)	(5)	(6)
1st decile	0.27	[0.20,0.35]	0.19	[0.12,0.27]	0.08	[-0.10,0.25]
2nd decile	0.21	[0.15,0.27]	0.18	[0.13,0.24]	0.11	[-0.02,0.24]
3rd decile	0.19	[0.14,0.25]	0.18	[0.13,0.23]	0.16	[0.04,0.28]
4th decile	0.18	[0.13,0.23]	0.17	[0.13,0.22]	0.09	[-0.03,0.20]
5th decile	0.18	[0.13,0.22]	0.15	[0.11,0.20]	0.09	[-0.01,0.20]
6th decile	0.17	[0.13,0.21]	0.14	[0.10,0.19]	0.11	[0.00,0.21]
7th decile	0.16	[0.12,0.21]	0.14	[0.10,0.19]	0.08	[-0.02,0.19]
8th decile	0.15	[0.10,0.19]	0.11	[0.06,0.16]	0.09	[-0.04,0.21]
9th decile	0.16	[0.09,0.22]	0.12	[0.06,0.19]	0.09	[-0.06,0.24]
Observations	18,116		10,144		7,965	

Note: The table contains the quantile regression estimates of the vocational status (vis-á-vis the academic status) on earnings. Vocational education is a dummy variable with 1 for upper secondary vocational graduates and 0 for lower secondary graduates *plus* upper secondary academic graduates. 95% confidence intervals in brackets. Data source: China General Social Survey 2003-2015.

Table 7: Instrumental variable quantile regression results for robustness check

	Whole sample		Urban hukou		Rural hukou	
	Coefficient	95% CI	Coefficient	95% CI	Coefficient	95% CI
	(1)	(2)	(3)	(4)	(5)	(6)
1st decile	0.20	[-0.04,0.48]	0.04	[-0.28,0.23]	-0.02	[-0.42,0.26]
2nd decile	0.05	[-0.11,0.27]	-0.02	[-0.24,0.25]	-0.01	[-0.36,0.21]
3rd decile	0.09	[-0.15,0.32]	0.07	[-0.21,0.30]	-0.13	[-0.37,0.13]
4th decile	0.13	[-0.08,0.27]	0.18	[-0.09,0.37]	-0.01	[-0.35,0.20]
5th decile	0.16	[-0.02,0.32]	0.17	[-0.05,0.40]	-0.10	[-0.53,0.25]
6th decile	0.13	[-0.02,0.35]	0.19	[-0.01,0.35]	-0.10	[-0.37,0.10]
7th decile	0.13	[-0.04,0.29]	0.18	[0.01,0.32]	-0.13	[-0.42,0.24]
8th decile	0.06	[-0.06,0.20]	0.08	[-0.04,0.36]	-0.12	[-0.38,0.27]
9th decile	0.14	[-0.11,0.33]	0.23	[-0.10,0.50]	-0.04	[-0.55,0.30]
Observations	7,537		3,660		3,871	

Note: The table contains the instrumental variable quantile regression estimates of the vocational status (vis-á-vis the academic status) on earnings. Vocational education is a dummy variable with 1 for upper secondary vocational graduates and 0 for lower secondary graduates *plus* upper secondary academic graduates. The instrumental variable for the endogenous vocational status is the ratio of the number of upper secondary vocational enrolment over the number of lower secondary graduates. Standard errors in parentheses. 95% confidence intervals in brackets. Data source: China General Social Survey 2003-2015.

Table 8: OLS and 2SLS results with migrant workers

	OLS	IV	IV
	(1)	(2)	(3)
Vocational education	0.11	-0.58	-0.41
	(0.03)***	(0.77)	(0.66)
Experience	0.03	0.03	0.03
	(0.01)***	(0.02)*	(0.01)*
Experience^2	-0.00	-0.00	-0.00
	(0.00)***	(0.00)	(0.00)
Female	-0.27	-0.26	-0.26
	(0.03)***	(0.04)***	(0.04)***
Migrant	0.13	0.09	0.09
	(0.03)***	(0.05)*	(0.04)**
Province F.E.	YES	YES	YES
Year F.E.	YES	YES	YES
First-stage			
Vocational enrolment		0.50	0.53
		(0.22)**	(0.23)**
Academic enrolment			0.28
			(0.24)
Kleibergen-Paap LM statistics		4.62	5.86
P> Chi^2(LM)		0.03	0.05
Kleibergen-Paap Wald F statistics		4.54	2.88
Hanson J statistics			0.22
P> Chi^2(Hanson J)			0.64
Observations	4,259	2,061	2,061

Note: The table includes the OLS and 2SLS estimates and the first-stage results for the whole sample and the sub-samples of workers with urban hukou and of workers with rural hukou. Migrant workers are included in the sample. Vocational education is a dummy variable with 1 for upper secondary vocational graduates and 0 for upper secondary academic graduates. Year of schooling is dropped for collinearity. The instrumental variables are a) the ratio of the upper secondary vocational enrolment number over the lower secondary graduate number (vocational enrolment), and b) the ratio of the upper secondary academic enrolment number over the lower secondary graduate number (academic enrolment). Significance level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Data source: China General Social Survey 2010-2015.

Appendices

A Quantile regression methods

The quantile regression is a useful tool to model heterogeneous treatment effect on outcome variables, conditional on covariates (Chernozhukov and Hansen, 2008). Koenker and Bassett Jr (1978) define θth regression quantile as the vector of $\beta(\theta)$ under the asymmetric least absolute deviation loss, which could be efficiently solved through convex linear programming:

$$\min_{b \in R^K} [\sum_{t \in t: y_t > x_t b} \theta |y_t - x_t b| + \sum_{t \in t: y_t > x_t b} (1-\theta)|y_t - x_t b|] \tag{2}$$

In terms of interpretation, Chernozhukov and Hansen (2002) show that the outcomes could be expressed as follows in the latent outcome model, $Y_d = q(d, x, u_d), u_d \sim u(0,1)$, where the u_d is called rank variables and is responsible for heterogeneity in outcomes given covariates x and treatment status d. In the conventional Becker-Mincer framework, the stochastic term ϵ is viewed as individual unobserved ability or proneness Doksum (1974) that associates with earning potential. Correspondingly, the conditional distribution of wages reflects heterogeneous unobserved ability.

The quantile regression method developed by Koenker and Bassett Jr (1978) is applicable in scenarios where the independent variable of interest is exogenously determined. However, as indicated by Arias et al. (2002), variation of effects of education along the conditional distribution of earnings may result from the endogeneity bias that also varies across earning quantile. Chernozhukov and Hansen (2005) put forward instrumental variable quantile regression (IVQR) method to tackle the endogeneity problem across regression quantile. They prove that regression quantiles could be successfully identified with the imposition of rank invariance or rank similarity, in additional to a series of 2SLS-analogue assumptions, in the presence of endogeneity. In our case, the assumption of rank invariance indicates that an individual?s ranking in conditional wage distribution is invariant to his or her schooling decision conditional on the instrumental variables. A subtle version of rank invariance is rank similarity, where the rank variable may change across treatment status, d, by unsystematic, or 'noisy', variation. Rank similarity requires u_d to be identically distributed (Balestra and

Backes-Gellner, 2017), which indicates that, in the setting of our study, individuals choose between vocational and academic tracks without full knowledge of potential earnings. With rank invariance or similarity, the quantile variable could be successfully identified under the condition of monotone likelihood ratio condition, which requires the instrumental variable increases the relative joint likelihood of receiving treatment and having a weakly higher outcome (Chernozhukov and Hansen, 2008).