

Kang, Lili; Peng, Fei; Zhu, Yu

**Working Paper**

## Returns to higher education subjects and tiers in China - Evidence from the China Family Panel Studies

GLO Discussion Paper, No. 238

**Provided in Cooperation with:**  
Global Labor Organization (GLO)

*Suggested Citation:* Kang, Lili; Peng, Fei; Zhu, Yu (2018) : Returns to higher education subjects and tiers in China - Evidence from the China Family Panel Studies, GLO Discussion Paper, No. 238, Global Labor Organization (GLO), Maastricht

This Version is available at:  
<http://hdl.handle.net/10419/180905>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Returns to higher education subjects and tiers in China - Evidence from the China Family Panel Studies

by

Lili Kang<sup>a,b</sup>, Fei Peng<sup>a</sup> and Yu Zhu<sup>c,\*</sup>

<sup>a</sup> Shanghai Lixin University of Accounting and Finance, China

<sup>b</sup> Institute of Economics, School of Social Sciences at Tsinghua University, China

<sup>c</sup> University of Dundee, UK

## Abstract

Using the recent China Family Panel Studies, we are able to identify the subjects studied of both college (2-3 years) graduates and university (4-5 years) graduates. For the latter group, we can further distinguish universities by the tier of selectivity (i.e. Key and Ordinary Universities). We take advantage of the rich information on the respondent's school cohort and *hukou* status at age 12 and the mother's age and education to estimate the simultaneous choice of subject and tier of prestige of higher education institutions (HEI) faced by university applicants. Using the doubly robust Inverse Probability Weighted Regression Adjustment (IPWRA) method to account for selection (on observables) into subjects and tiers, our treatment effect estimates suggest that OLS substantially underestimate the effect of attending more prestigious universities, for graduates of both genders in Law, Economics and Management (LEM). We also show that the recent massive expansion of the higher education sector have reduced returns to HE for all graduates, except for those studying LEM or Other non-STEM subjects at the most prestigious universities. Our results are robust to treating subjects as predetermined for the selection into HEIs by tiers of prestige.

**JEL code:** I23, I26

**Keywords:** Returns to university tier and subjects, China, Inverse Probability Weighted Regression Adjustment, Higher Education expansion

---

\* **Corresponding author.** School of Business, University of Dundee, 3 Perth Road, Dundee, DD1 4HN, UK. [yuzhu@dundee.ac.uk](mailto:yuzhu@dundee.ac.uk).

**Funding sources:** Fei Peng thanks the Shanghai Young Eastern Scholarship (grant number: QD2015049) for financial support. Lili Kang thanks the financial aids from China National Social Science Fund (grant number: 14BJL028) and Shanghai Shuguang Program (grant number: 15SG53).

**Acknowledgement:** We thank the Institute of Social Science Survey (ISSS) at Peking University for allowing access to the CFPS data. All errors are ours.

## 1. Introduction

There is an enormous literature on the returns to education. Most of the empirical literature focuses on the returns to different quantities of education, typically measured by years of schooling or different levels of academic or vocational qualifications. In general, inadequate attention has been paid to the returns to different types of education, partly due to data availability issues.

It is by now well established that returns to different university subjects (*majors* in the US) vary substantially (see e.g. Paglin and Rufolo (1990), Arcidiacono (2004) and Altonji et al (2012) for the US and Walker and Zhu (2011) for the UK). Similarly, there is a general consensus that attending more selective (prestigious) universities is associated with higher wages and earnings.

The development and expansion of the Chinese education system in general, and the higher education (HE) system in particular, has been a key contributing factor of China's remarkable economic growth over the last 4 decades. Using a growth accounting model which treats human capital as the opportunity cost foregone, Whalley and Zhao (2013) suggest that increases in human capital contributed to as much as 38.1% of economic growth in China over 1978-2008, and even higher for 1999-2008, a period of rapid HE expansion which witnessed an increase of annual enrolment by more than 4-fold. Whether and to what extent returns to HE vary by subjects and tiers of prestige are of enormous interest to policy makers and the general public in China, yet there is surprisingly little empirical evidence on this.

We contribute to the empirical literature on returns to higher education in the following ways. Firstly, to the best of our knowledge, this is the first attempt to estimate the treatment effects of *both* HE subjects and tiers of prestige on earnings in China. Secondly, we show that OLS substantially underestimates the effect of attending more prestigious universities, especially for graduates of both genders in Law, Economics and Management (LEM). Moreover, our results indicate that after the massive HE expansion, returns to HE have declined for all graduates, except for those studying LEM or other subjects at the most prestigious universities. Lastly, we find that our results are robust to treating subjects as predetermined for the selection into higher education institutions (HEIs) by tiers of prestige. This is consistent with the fact that the tracking choice made by students in Chinese education system into the Science or Arts subjects

during the Senior High School does restrict the HE subject choice at the time of university applications.

The rest of the paper is organized as follows. Section 2 reviews the relevant literature, with special reference to China. Section 3 discusses the institutional background. Section 4 presents the data while Section 5 discusses the Inverse Probability Weighted Regression Adjustment (IPWRA) methodology. The empirical results are presented and discussed in Section 6. Finally, Section 7 concludes.

## **2. Literature Review**

It is well known that graduate earnings vary substantially by the subjects they study or the prestige (selectivity) tier of HEIs they attend. Numerous university league tables (e.g. the Complete University Guide for the UK) publish average graduate starting salaries by subjects on an annual basis. Recently, the UK Department for Education starts to publish official statistics of graduates employment and earnings outcomes up to 5 years after graduation for all subjects by university based on linked administrative data (Department for Education 2017).

The economic literature on returns to HE types can be classified into two strands: the first is concerned with returns to subjects (or majors as in the US literature) while typically holding the prestige tier constant, and the second is concerned with returns to prestige tier (also known as selectivity) while typically treating subjects as given. Both strands are dominated by descriptive analysis. In this brief review we will focus on the handful of studies which attempt to estimate the causal effect of either subject or prestige type.

Paglin and Rufolo (1990) highlight the importance of mathematical ability in determining field choice for US college students, which is consistent with earnings maximization by major. In particular, they find that graduates with above average Graduate Record Exam (GRE) quantitative scores for their undergraduate field tend to switch to fields requiring higher average scores.<sup>1</sup>

---

<sup>1</sup> Altonji et al. (2012) survey the literature on the demand for and returns to college major, allowing for the effects of high school curriculum.

Using a dynamic model of college and major choice which allows for switching and dropout, Arcidiacono (2004) focuses on ability sorting across majors in the National Longitudinal Study of the Class of 1972 (NLS72). While individuals appear to make the initial choice about college and major according to the course-specific expected earnings, they update their decisions by dropping out or changing the course, in the light of new information about their preferences and ability while in college. Moreover, he finds large earnings premiums for natural science and business majors even after controlling for selection, with preferences for majors playing a key role.

As for the causal studies on the effect of college selectivity on earnings, most studies use matching methods which assume selection on observables only, see e.g. Chevalier and Conlon (2003) and Hussain et al. (2009) for the UK, both of which find a modest return to attending more selective universities of around 6% for one standard deviation increase in HEI quality. Dale and Krueger (2002) is the first attempt to allow for selection on unobservables. By matching students who were accepted, with those also applied to but were rejected, by the same set of colleges, they find little evidence of returns to attending more prestigious colleges in the US for students with the same ability. Following the Dale and Krueger method, Broecke (2012) finds that one standard deviation in selectivity leads to a 7% increase in earnings in the UK. The identification relies on comparing students who satisfied the conditional offers for their first-choice to students who applied to the same universities but who attended second-choice universities due to failure to meet the conditions of their preferred offer.

Walker and Zhu (2018) attempt to estimate the treatment effect of both university selectivity and subjects for the UK. Using the Inverse Probability Weighted Regression Adjustment (IPWRA) approach to allow for selection on observables into both subjects and institution types, they find strong differences in returns to selectivity by subject. Belfield et al. (2018) also apply the IPWRA method on the new Longitudinal Education Outcomes (LEO) administrative dataset to account for variation in course selectivity and student characteristics, in order to estimate the relative labour market returns to different degrees in the UK.

This paper will follow the same IPWRA methodology, which will be discussed in details in Section 5. Compared to the more conventional matching method which can only be applied to a binary treatment, the main advantage of the IPWRA method is that

it allows for robust estimation of treatment effects when the number of treatments outnumber the number of potential instruments.

The recent HE expansion in China has been the topic of a growing literature, mostly published in Chinese. The literature review by Feng (2012) focusing on access equity indicates that most of the studies take the perspective of education or sociology and conclude that inequality has exacerbated following the expansion. In particular, students from disadvantaged (e.g. rural) backgrounds enroll disproportionately in lower-tier HEIs and/or less popular (lucrative) subjects. Allowing for complementariness among workers of different ages and qualifications, Li et al. (2017) show that the HE expansion has increased the college premium of older cohorts of graduates at the expense of more recent cohorts. Exploring the discontinuity in the months of births induced by the HE expansion, Dai et al. (2018) show that each additional year of university education induced by the 1999 higher education expansion increases monthly wage income by 21%, compared to an OLS estimate of 8%.

Few Chinese studies have examined the choice of university subjects and tiers in China. Using a sample of secondary school students from Beijing, Sheng (2017) shows that while there is little class difference in subject choice, students from high-income families are more likely to enter national key universities, which is the most prestigious tier in the Chinese education hierarchy.

### **3. Institutional Background**

In 1986, China introduced 9-year compulsory education starting at age 6, with 6 years of primary education and 3 years of Junior High Schools. Students who continue with their education after completing compulsory schooling, enter vocational schools or Senior High Schools, both lasting 3 years. Students in the latter type of schools are streamed into the Science or Arts tracks for the last two years of upper secondary education (OECD 2016).

After 12 years of schooling, secondary school graduates can apply to colleges and universities, through a centralised admissions system which proceeds sequentially in tiers on the basis of the standardized National College Entrance Examinations

(*gaokao*).<sup>2</sup> Colleges and universities in China can be classified into 3 tiers in descending order of prestige and entry requirements, i.e. Key universities (mostly National Project 985 and Project 211 universities), Ordinary universities, and vocationally training colleges. The duration of studies for the first two tiers is 4-5 years leading to a Bachelor's Degree upon the successful completion of study, whereas the duration of the last tier is 3 years leading to a college Diploma.<sup>3</sup>

Admissions in the second tier only starts once the assignment in the first tiers are finalized, and so on. Each applicant submits a lexicographic list to the provincial student placement office, indicating her preference over HEI (i.e. colleges and universities) and within each HEI, her preference over subjects. Importantly for our analysis, university applicants must consider both the tier of the HEI and the subject given HEI at the same time, which defines a university course. This is an important feature of the Chinese educational system which must be taken into account in the econometric analysis.

Another important institutional feature of China is the *hukou*, or household registration system, which classifies people as rural and urban status at birth, usually according to the mother's *hukou* status. It is widely known that education resources at primary and secondary level are highly unequal, in favour of urban residents in China, who enjoyed better access to HEIs, especially to the most prestigious ones. Intuitively, *hukou* status at that age is likely to determine whether the respondent attended an urban or rural secondary school, with systematic differences in the quality which might also be subject specific (e.g. rural secondary schools might struggle to recruit competent English teachers). In addition to family background variables such as mother's education, we will also exploit *hukou* status at age 12 as a key determinant of the choice of HE types faced by HE applicants in our formal analysis.

---

<sup>2</sup> The description of the Chinese College Admissions system relies heavily on Zhu (2014), which assesses the change from a sequential choice algorithm to the parallel choice algorithm. Our sample predates the national rollout of the parallel choice algorithm which took place in 2013.

<sup>3</sup> There are nearly 2600 regular HEIs (i.e. excluding adult education etc.) in China employing 1.6 million academic staff as of 2016 (Ministry of Education 2018).

#### 4. Data

This study is based on the China Family Panel Studies (CFPS), a biennial longitudinal nationally representative survey of Chinese families undertaken by the Institute of Social Science Survey (ISSS) of Peking University of China since 2010. The survey collects information on employment, income, education and health at individual, family and community levels (see Xie and Hu (2014) for details). Important for the purpose of this paper, the survey contains detailed information on the respondent's subject of studies at each level of post-secondary education, including Senior High School, College, Undergraduate and Postgraduate levels. For universities graduates, it also asks them in the first wave (i.e. the 2010 survey) about the tier of prestige, i.e. national key, key, etc.

Our sample consists of all individuals aged 20-60 whose highest qualification is Senior High School, College or University in the first wave of the CFPS which was conducted in 2010. We drop individuals whose highest qualification is Junior High School or below, as HE choice is irrelevant to them. We also exclude the small number of respondents with Master's or PhDs as there might be important unobservables which distinguish them from the rest of the graduates. Implicitly we want to model the choice of a Senior High School graduate between entering HE, or entering the labour market straight away, and if choosing the former option, between different HE subjects and different tiers of prestige or selectivity of the HEIs. Due to sample size limit, we choose a 3 by 3 grouping of HE types, namely 3 subjects consisting of STEM (Sciences, Technology, Engineering and Math/Medicine), LEM (Law, Economics and Management) and Other subjects,<sup>4</sup> and 3 institution tiers consisting of Colleges, Ordinary Universities and Key-Universities.

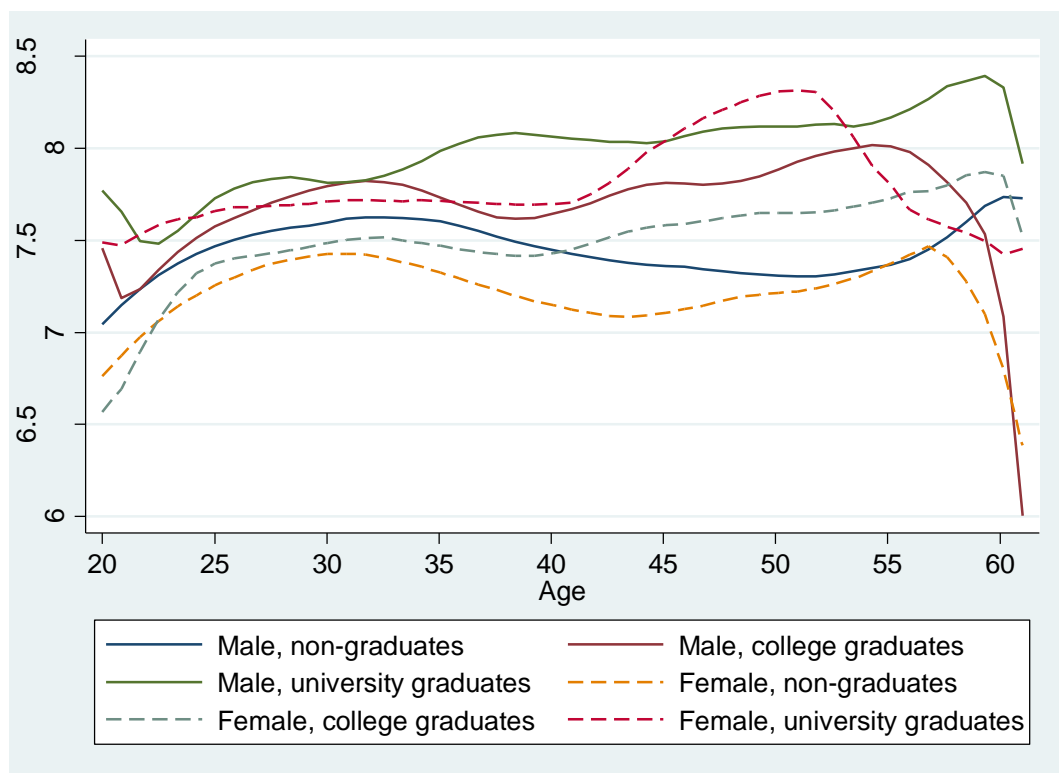
After excluding individuals with missing values on key variables including the outcome variable of monthly earnings and a handful of graduates with degrees from abroad, we end up with 2,813 distinct individuals, of which 1,173 (41.7%) are women. We take full advantage of the panel nature of the CFPS by including waves 2 and 3 (conducted in 2012 and 2014 respectively). However, only earnings, age, survey years and survey months are treated as time-varying in our analysis.

---

<sup>4</sup> This is derived from the 11 subjects reported: Sciences, Engineering, Agriculture, Medicine are grouped into STEM; Law, Economics and Management are grouped into LEM; and Philosophy, Education, Literature, History are grouped into Other Subjects.



Figure 1 shows the smoothed age-earnings profiles, derived from a kernel-weighted local polynomial regression of log real monthly incomes (in January 2009 prices) on age, by level of qualifications and gender using our panel sample. For each gender, the earnings of university graduates dominate those of college graduates, which in turn dominate those of Senior High School graduates, except towards the very end of the career where it becomes a bit noisy due to small cells and possible early retirement. It is also worth noting that people with higher qualifications tend to have higher earnings growth at the beginning of their career. However, these age-earnings profiles might be confounded by cohort and time effects.



**Figure 1: Smoothed age-earnings profiles, by level of qualifications and gender**

Table 1A shows the relative frequencies by gender. While 46.7% of males in the sample are graduates, 54.6% of females hold a college or university degree. The variation in the relative frequencies partly reflect the popularity of a certain subject-tier combination, with Key-University Other Subjects and Key-University STEM being the least common combination for men and women respectively.

Table 1B report the summary statistics for key variables in Wave 1 by gender. The mean real monthly earnings in January 2009 constant price are 2443 and 1849 RMB yuan for men and women respectively. Compared to men, women are almost 3

years younger and more likely to come from an advantageous background as proxied by a non-agricultural *hukou* at age 12 and mother's education level. While women are more likely to live in urban areas than men do, there is no difference in the probability of living in the East Region, which is the most economically developed region of China.

**Table 1A: Relative Frequencies, by gender**

HE types	Men	Women	Total
Senior High School	53.35	45.44	50.05
All HE, of which	46.65	54.56	49.95
College STEM	11.04	9.12	10.24
College LEM	12.74	16.20	14.18
College Other	4.57	7.16	5.65
OrdinUG STEM	4.70	4.01	4.41
OrdinUG LEM	4.21	4.60	4.37
OrdinUG Other	2.62	5.46	3.80
KeyUG STEM	3.41	1.88	2.77
KeyUG LEM	2.01	3.07	2.45
KeyUG Other	1.34	3.07	2.06
Total	100.00	100.00	100.00
Obs	1,640	1,173	2,813

**Table 1B: Summary Statistics, by gender**

	Men	Women	Total
Real monthly salary (Jan 2009 price)	2443	1849	2196
Age	37.3	34.6	36.2
School cohort	1972.1	1974.8	1973.2
Non-agricultural hukou at age 12	0.449	0.500	0.471
Mother's year of birth	1945.6	1948.4	1946.8
Mother's education Level (1-6)	2.13	2.42	2.25
Urban	0.802	0.872	0.831
East Region	0.441	0.443	0.442

Note: Distinct individuals in Wave 1. OrdinUG and KeyUG stand for Ordinary and Key universities, respectively.

Table 2 presents the mean log real monthly salaries by HE types and gender for the wage panel. The raw graduate wage premium is 0.36 log points for both men and

women.<sup>5</sup> Male college graduates earn 0.29 log points more than their Senior High School counterparts, but 0.11 log points less than male Ordinary University graduates. Although Key-Universities and Ordinary-Universities take the same time to complete, there is a staggering 0.25 log points' earnings difference among male university graduates. On the other hand, female college graduates earn 0.25 log points more than their Senior High School counterparts, but 0.19 log points less than male Ordinary University graduates. The earnings premium for attending Key Universities for women is 0.22 log points. It is also interesting to note that STEM graduates have both the highest and the lowest earnings for both genders, implying a higher return to selectivity to study those subjects. Finally, it turns out the difference in earnings is more or less constant across all education levels and types, at about 0.20 log points.

**Table 2: Mean log real monthly salaries, by HE types and gender**

HE types	Men	Women	Gender difference
Senior High School	7.44	7.22	0.22***
All HE, of which	7.81	7.58	0.23***
<i>All Colleges</i>	7.73	7.47	0.26***
College STEM	7.70	7.40	0.29***
College LEM	7.77	7.50	0.27***
College Other	7.71	7.48	0.23***
<i>All Ordinary Universities</i>	7.84	7.66	0.18***
OrdinUG STEM	7.81	7.76	0.05
OrdinUG LEM	7.87	7.62	0.25*
OrdinUG Other	7.84	7.62	0.23**
<i>All Key Universities</i>	8.09	7.88	0.21***
KeyUG STEM	8.17	7.96	0.21
KeyUG LEM	8.05	7.91	0.13
KeyUG Other	7.95	7.81	0.14
Total	7.62	7.42	0.19***

Note: Unweighted wage panel. OrdinUG and KeyUG stand for Ordinary and Key universities, respectively. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

<sup>5</sup> For simplicity, we interpret a slope coefficient  $b$  in the log wage equation in terms of log points, which approximate a  $100b$  percentage point change. So a 0.1 log point increase translates to an approximate 10 percentage point increase. The exact point change is given by  $100*(\exp(b)-1)$  where  $\exp$  denotes the exponential function.

## 5. Methodology

We begin with Pooled Ordinary Least Squares (POLS) estimation of attending different types of HEIs, defined as combinations of subjects and tiers of varying selectivity, as a benchmark, to connect ourselves to the existing literature that has depended heavily on least squares. Due to the limitation imposed by sample sizes, we use a 3 by 3 grouping of HE subjects and tiers. We also present the corresponding random-effect (RE) results by gender.

Graduates' earnings will, to some extent, reflect their ability, as high ability students will earn more on average regardless of their HE attainment. Unfortunately, the CFPS does not contain information on standardized test scores, e.g. the National University Entrance Exam (*gaokao*) scores which can be used as proxies for ability. However, even if we were able to control for HE entry scores, coefficients of HE types in the regression analysis may still not be regarded as providing causal estimates, as the OLS counterfactual depends on there being no unobservable confounders (i.e. selection only on observables) *and* a parametric functional form assumption.

There may be important unobserved confounders such as family background, institutional factors (e.g. *hukou* status in the Chinese context) and personality traits that affect individual's choice of HE types. The usual approach to this problem is to either search for instruments for the choices that individuals make, or to exploit discontinuities associated with admission requirements. For the purpose of our analysis, the choice set is so large (indeed the 3 by 3 grouping we adopted could be regarded as a realistic minimum) that it would be extremely difficult to find a sufficiently large number of instruments or discontinuities being available (see Kirkeboen *et al* (2016) for an exception).

In our analysis of the effect of HE types on monthly earnings, we need to allow for multiple treatments, where we assume that selection into each treatment is driven only by observables. While matching methods is only applicable to a single binary treatment, weighting methods can deal with multiple treatments by ensuring treatment groups are similar to the control group by weighting them accordingly. Under the conditional independence assumption (CIA) which implies that there is only selection on observables, this weighting method can yield causal estimates of the Average Treatment Effects (ATE).

In particular, we will explore the “doubly robust” inverse-probability weighted regression-adjustment (IPWRA) estimator (see Wooldridge 2007, Wooldridge 2010 chapter 13, and Imbens and Wooldridge 2009). IPWRA estimates the average treatment effect (ATE) of any HE types, allowing for selection into a particular HE type using multinomial logit model in the first step. In the second step, this estimator then estimates an OLS regression of log earnings using the reweighted data, using the inverse of the predicted probabilities from the first step as the weights. In other words, IPWRA weights observations in the sparse parts of the distribution *more* heavily. If the functional form is correct then the OLS estimates are unaffected by any weighting. However, the true functional form is rarely known in practice, and any misspecification is likely to yield biased estimates.

Indeed, IPWRA is *doubly* robust in the sense that only one of the two steps needs to be specified correctly. In other words, the estimates of the second step, the log earnings equation, are robust to misspecification in the weighting of the data conditional on the specification of the second step being correct; *and* the estimates of the second step are also robust to misspecification of the second step provided the multinomial logit weighting in the first step is correctly specified.

Walker and Zhu (2018) is a recent example of the application of the IPWRA approach to estimate the relative returns to HE types in the UK. While they are able to control for the selectivity of the courses a graduate attended using the detrended mean standardized university entry scores known as A-Levels, the fact that the UK Labour Force Survey does not contain information on family background means that their treatment effect estimates might still be biased due to potential selection on unobservables.

## 6. Empirical Results

### 6.1. Main results

Table 3 presents the POLS and RE estimates of the effect of various HE types on log real monthly earnings for each gender separately, controlling for age, age squared and living in urban areas or in the East Region. We deliberately choose this parsimonious specification, which facilitates the IPWRA model later on.<sup>6</sup> However, the coefficients on HE types are robust to specifications allowing for provincial fixed-effects. Note that we are estimating the absolute returns to different HE types, using the same Senior High School graduates with no HE credentials as the control group in all specifications while allowing for 9 treatment groups.

For men, attending a 3-year college yields a return between 25-31%, with STEM subjects having the lowest returns and LEM having the highest returns. However, the differences are statistically insignificant across subjects. Men attending 4-year Ordinary Universities enjoy a return between 34% and 43%, again with the lowest returns for STEM. Men attending the most prestigious Key Universities enjoy a return between 47% and 66%, with substantially lower returns for graduates studying Other Subjects than STEM or LEM. So, it appears that returns to attending more prestigious HEIs do vary by subject. A model which fails to allow for the interaction effects is likely to yield biased estimates.

For women, the returns appear to be quite similar to those for men, except for Ordinary University STEM graduates and Other Subjects Key university graduates, which are significantly higher. It is also worth noting the POLS and RE estimates are largely indistinguishable from each other.

For men, living in urban areas is associated with approximately 19% higher monthly earnings compared to living in rural areas, while living in the more developed East region is associated with roughly 35% higher earnings than living in central or western regions. For women, the wage premia for living in urban areas or the East Region are even higher.

---

<sup>6</sup> Controlling for occupation, industry and ownership types of employers will lead to endogeneity bias, as all these are post-education choices which in principle simultaneously affect earnings.

**Table 3: OLS and Random-effects (RE) Estimates, by gender**

	Men		Women	
	Pooled OLS	RE	Pooled OLS	RE
College STEM	0.266*** (0.049)	0.253*** (0.048)	0.231*** (0.064)	0.236*** (0.062)
College LEM	0.313*** (0.044)	0.304*** (0.044)	0.255*** (0.046)	0.258*** (0.045)
College Other	0.295*** (0.070)	0.280*** (0.071)	0.301*** (0.077)	0.303*** (0.074)
OrdinUG STEM	0.346*** (0.092)	0.335*** (0.085)	0.554*** (0.088)	0.553*** (0.087)
OrdinUG LEM	0.393*** (0.094)	0.389*** (0.088)	0.418*** (0.112)	0.430*** (0.103)
OrdinUG Other	0.434*** (0.074)	0.418*** (0.080)	0.472*** (0.085)	0.467*** (0.080)
KeyUG STEM	0.621*** (0.084)	0.651*** (0.086)	0.656*** (0.155)	0.662*** (0.150)
KeyUG LEM	0.610*** (0.108)	0.630*** (0.114)	0.613*** (0.095)	0.617*** (0.094)
KeyUG Other	0.466*** (0.122)	0.495*** (0.122)	0.566*** (0.075)	0.575*** (0.077)
Age	0.025** (0.012)	0.039*** (0.012)	0.023 (0.015)	0.030** (0.014)
Age sq	-0.000* (0.000)	-0.000*** (0.000)	-0.000 (0.000)	-0.000* (0.000)
Urban	0.187*** (0.043)	0.187*** (0.042)	0.249*** (0.069)	0.240*** (0.065)
East	0.348*** (0.032)	0.356*** (0.031)	0.429*** (0.038)	0.424*** (0.037)
Constant	6.622*** (0.275)	6.420*** (0.260)	6.429*** (0.408)	6.194*** (0.386)
Observations	3,402	3,402	2,395	2,395
$R^2$	0.156		0.193	

Note: Robust standard errors in parentheses. Clustering at the individual level for pooled OLS. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . OrdinUG and KeyUG stand for Ordinary and Key universities, respectively. Other controls include dummies for survey year and months.

Table 4 focuses on the corresponding IPWRA estimates of the ATEs of HE types, relative to Senior High School graduates.<sup>7</sup> The choice of HE types is estimated using a multinomial logit on the respondent's school cohort and *hukou* status at age 12, as well as his/her mother's age and educational qualification. These family background variables have been widely used in the economics of education literature as instruments for education types (see e.g. Berger 1988).

**Table 4: IPWRA Average Treatment Effects (ATE), by gender**

	(1) Men	(2) Women
College STEM	0.247*** (0.049)	0.227*** (0.051)
College LEM	0.283*** (0.045)	0.232*** (0.039)
College Other	0.284*** (0.067)	0.328*** (0.056)
OrdinUG STEM	0.445*** (0.050)	0.637*** (0.064)
OrdinUG LEM	0.491*** (0.069)	0.384*** (0.156)
OrdinUG Other	0.486*** (0.052)	0.453*** (0.080)
KeyUG STEM	0.699*** (0.074)	0.697*** (0.113)
KeyUG LEM	1.006*** (0.159)	0.777*** (0.053)
KeyUG Other	0.638*** (0.069)	0.581*** (0.051)
Observations	2,755	2,230

Note: Robust standard errors in parentheses. Clustering at the individual level for pooled OLS. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . OrdinUG and KeyUG stand for Ordinary and Key universities, respectively. Other controls include dummies for survey year and months.

Table A3 in the Appendix tabulates the choice of HE subjects and HE tiers by mother's education, separately for agricultural and non-agricultural *hukou* holders as defined at age 12. The patterns indicate a very strong and monotonic relationship between mother's educational qualifications and respondent's HE tiers, for *urban* *hukou* holders at least (note that there are very few agricultural *hukou* holders with highly educated mothers). For instance, as the mother's education increase from

<sup>7</sup> The estimation was implemented using the Stata routine *teffects ipwra*. Tables A1 and A2 in the Appendix presents the full set of results for men and women, respectively. Consistent with the existing literature, mother's education qualifications seem to play the most significant role in the choice of respondents' HE types.



illiteracy to university degree, the chance of an urban resident not going to HE drops from 67% to 17% while that for going to the most selective Key Universities increase from 4% to 23%. There is also suggestive evidence that rural and urban residents might have different preferences with regard to HE subjects, with the former more likely to choose STEM and the latter more inclined to do LEM. These patterns are consistent with the existing literature which suggests that students with advantaged backgrounds from urban areas are over-represented in the most selective universities in China.

The differences in sample sizes between OLS/RE and the IPWRA estimates reflect the fact that observations which are off *common-support* are dropped from the final outcome (wage) equations in IPWRA. However, the overlap plots for men and women respectively in Appendix Figures A1 and A2 suggest that for most graduates (i.e. the treated) we are able to find good matches from the control group of non-HE graduates.

Compared to the OLS estimates, the IPWRA returns to Law, Economics and Management (LEM) graduates are much higher, by 0.38 and 0.16 log points for men and women respectively, but only if they attend the most prestigious Key-Universities. This implies that ignoring the endogeneity of HE types is likely to lead to underestimation of the returns to attending more selective universities, especially the most prestigious Key Universities in China.

## ***6.2. Treatment Effects of HE Tiers Conditional on Subjects***

One concern with the analysis so far is that while our model might be doing a good job in predicting the selection into different tiers of HEIs, we were unable to precisely predict the subject choice due to a lack of information on prior attainment in different subjects. Ideally, we would like to have the respondents' grades on key subjects such as Maths, Physics, Chemistry, Chinese, History etc., from say the final year of Junior High School, which would help determine the selection into STEM or non-STEM tracks in the Senior High Schools and consequently the subject choice for HE. Unfortunately, this detailed information is typically not available in surveys like ours.

To overcome this problem, in the following analysis we will estimate the effect of HE tiers ***conditional*** on subjects using RE and IPWRA. Effectively we are treating subjects as exogenous while focusing on HE selectivity. This is a useful simplifying

assumption of the reality, as streaming takes place at the beginning of Year 11 in China when students have to choose between the STEM and Arts track. By the time of college/university application at the end of Year 12, the subject choice is limited to some extent by the streaming choice made some two years ago.

Table 5 presents the RE estimates in columns 1-3 and the IPWRA estimates in columns 4-6 of the effects of HE tiers for men. As for our main results, RE severely underestimates the returns to attending more prestigious HEIs for LEM graduates, by as much as 0.43 log points while accounting for endogeneity of HE tiers seems to make little difference for graduate in Other Subjects. Moreover, the difference between RE and IPWRA also exceed 0.13 log points for STEM and LEM graduates from Ordinary-Universities.

Table 6 shows the corresponding RE and IPWRA estimates of HE tiers for women conditional on subjects. RE underestimates the treatment effect of studying LEM at Key Universities by 0.25 log points. It is also noteworthy that the IPWRA estimates for STEM at either Ordinary or Key Universities exceed their RE counterparts by more than 0.12 log points.

Overall, we conclude that our main results are robust to treating subjects as predetermined in HE choices. This is reassuring for existing studies which focus on estimating the returns to university tiers, as long as subjects are controlled for.

**Table 5: RE and IPWRA Treatment Effects of HE Selectivity, conditional on subject, Men**

	STEM	RE LEM	Other	STEM	IPWRA LEM	Other
	(1)	(2)	(3)	(4)	(5)	(6)
College	0.235*** (0.049)	0.307*** (0.045)	0.272*** (0.071)	0.241*** (0.045)	0.310*** (0.042)	0.262*** (0.057)
Ordinary university	0.320*** (0.086)	0.387*** (0.089)	0.394*** (0.083)	0.472*** (0.053)	0.521*** (0.074)	0.485*** (0.045)
Key university	0.637*** (0.086)	0.621*** (0.115)	0.518*** (0.125)	0.687*** (0.060)	1.052*** (0.155)	0.604*** (0.066)
Observations	2,442	2,435	2,087	2,431	2,352	1,740

Note: Robust standard errors in parentheses, \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

**Table 6: RE and IPWRA Treatment Effects of HE types, conditional on subject, Women**

	RE			IPWRA		
	STEM	LEM	Other	STEM	LEM	Other
	(1)	(2)	(3)	(4)	(5)	(6)
College	0.225*** (0.062)	0.243*** (0.045)	0.285*** (0.074)	0.257*** (0.051)	0.262*** (0.038)	0.322*** (0.054)
Ordinary university	0.543*** (0.088)	0.405*** (0.101)	0.418*** (0.079)	0.673*** (0.073)	0.394*** (0.142)	0.486*** (0.078)
Key university	0.642*** (0.146)	0.586*** (0.095)	0.561*** (0.076)	0.766*** (0.141)	0.834*** (0.052)	0.623*** (0.049)
Observations	1,428	1,624	1,449	1,356	1,567	1,403

Note: Robust standard errors in parentheses, \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

### 6.3. Heterogeneous effect of the HE expansion

Finally, we explore the heterogeneous effect of the HE expansion by HE tiers and subjects. The HE sector in China experienced an unprecedented expansion in the decade since 1999, with annual enrolment increased from roughly 1 million to 6 million. The most dramatic growth took place between 1999 and 2001, with around 40% annual growth in each year.

**Table 7: RE and IPWRA Treatment Effects HE types, Pre- and post-expansion, pooled gender**

	RE		IPWRA	
	Pre-expansion (1)	Post-expansion (2)	Pre-expansion (3)	Post-expansion (4)
College STEM	0.303*** (0.052)	0.083 (0.058)	0.300*** (0.046)	0.119** (0.054)
College LEM	0.324*** (0.041)	0.156** (0.062)	0.337*** (0.034)	0.166*** (0.061)
College Other	0.380*** (0.063)	0.035 (0.078)	0.298*** (0.065)	0.126 (0.089)
Ordinary STEM	0.567*** (0.071)	0.012 (0.091)	0.648*** (0.053)	-0.021 (0.075)
Ordinary LEM	0.435*** (0.066)	0.325*** (0.106)	0.430*** (0.089)	0.337*** (0.095)
Ordinary Other	0.503*** (0.074)	0.187* (0.101)	0.480*** (0.070)	0.210 (0.178)
KeyUG STEM	0.819*** (0.095)	0.432*** (0.101)	0.904*** (0.067)	0.438*** (0.098)
KeyUG LEM	0.712*** (0.103)	0.462*** (0.112)	0.742*** (0.057)	1.079*** (0.192)
KeyUG Other	0.495*** (0.093)	0.661*** (0.151)	0.536*** (0.068)	0.842*** (0.150)
Female	-0.231*** (0.029)	-0.212*** (0.040)		
Observations	4,025	1,772	3,260	1,434

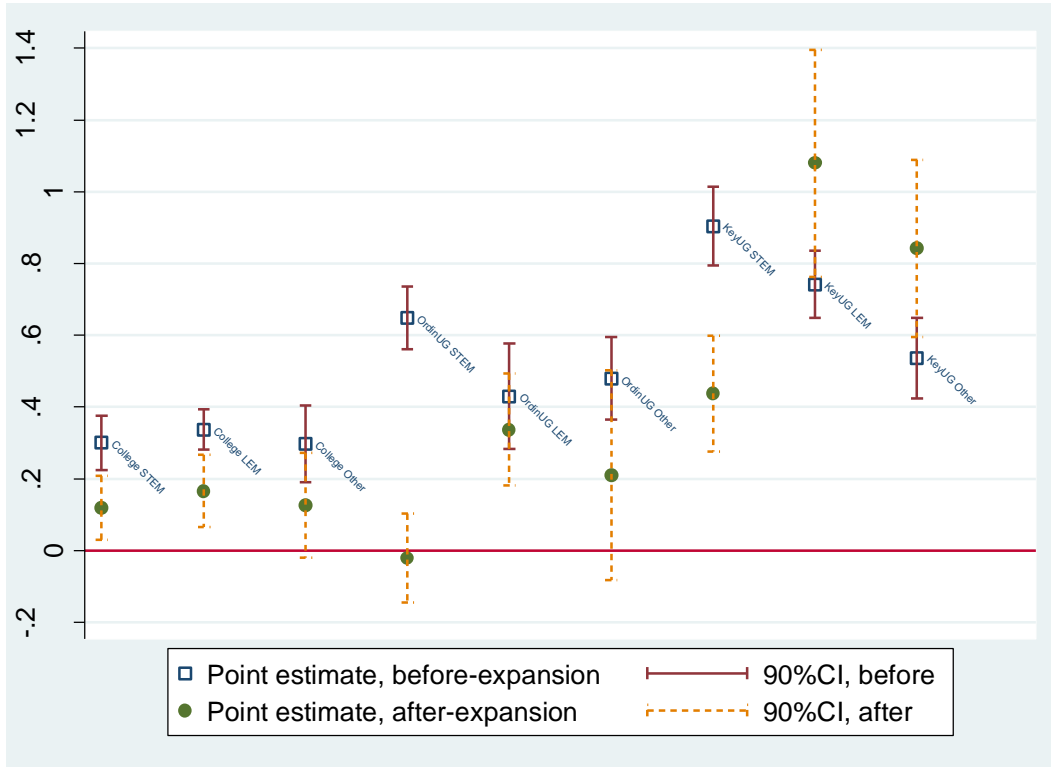
Note: Robust standard errors in parentheses, \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

We classified people born in August 1979 or before as the pre-expansion cohort and those born in September 1979 or later as post-expansion.<sup>8</sup> Only about 30% of our sample are post-expansion. Due to the small sample sizes, we choose to pool gender in

<sup>8</sup> This is consistent with Wu and Zhao (2010) who show that high school students account for the vast majority of 18-year olds and a small majority of 19-year olds among full-time students in the various censuses. In China, it is not uncommon for children to start school at an age later than 6, especially in rural areas, or to repeat grades. Cohorts who completed secondary school just before the HE expansion might also be tempted to retake the National University Entrance Exam to take advantage of the reform.

Table 7 which presents RE and IPWRA for pre- and post-expansion cohorts separately. We also drop the survey month dummies from the controls.

While RE and IPWRA results are remarkably similar for the pre-expansion cohorts, they look rather different for the post-expansion cohorts. It suggests that the difference between the treatment effect estimates and their RE counterparts is driven by the substantial HE expansion in recent years. Moreover, the HE expansion seems to have very heterogeneous effect on the returns to HE, depending on the subject and tier of prestige. A comparison of the IPWRA estimates in column 4 to column 3 suggests that the returns to HE seem to have declined for most graduates, with the exception of LEM and Other Subjects graduates from Key Universities. In the case of LEM graduates from the most prestigious universities, the returns to the degree (relative to Senior High School Diploma) jumped from 0.74 to 1.08 log points. This is a staggering increase of 0.34 log points. For graduates studying Other Subjects than LEM or STEM at Key Universities, the increase in returns is about 0.30 log points.



**Figure 2: Changes in returns to HE, by HE type and time period**

Figure 2 visualizes the changes in returns by HE types arising from the HE expansion. The hollow squares represent the point estimates pre-expansion while the solid circles indicate the point estimates post-expansion. The solid and dashed spikes with caps

represent the corresponding 90% confidence intervals. Whereas the returns to all types of HE are significantly positive before the expansion, we can no longer reject the null of a zero return at 10% significance for Other Subjects at Colleges, or STEM and Other Subjects at Ordinary Universities post-expansion. It turns out that 7 out of the 9 HE types have experienced a decline in returns after the expansion. In particular, the declines in returns to STEM subjects are statistically significant at the 10% level at all tiers of selectivity. Only LEM and Other (non-STEM) Subjects at Key Universities experience increases in returns post-expansion, even though they are not statistically significant at the 10% level. On the other hand, the decline in returns to LEM at college level is significant at the 10% level.

One possible explanation is that while the HE expansion has improved the overall access to colleges and universities, it might have intensified the competition to the most prestigious HEIs (see e.g. Feng 2012). Moreover, students from socio-economically more advantaged backgrounds might have benefited disproportionately from the expansion.

However, one should exercise caution in the interpretation of the post-expansion results. Firstly, the IPWRA results are rather imprecise for post-expansion cohorts due to small sample size. Secondly, almost 20% of the sample are off common support for both pre- and post-expansion cohorts. This could be partly due to the narrow age range for each subset, especially for the post-expansion cohorts. Indeed, graduates who entered HE in 1999 or later were only observed for a maximum of 7 years by the time of the 2010 survey.

## 7. Concluding remarks

To the best of our knowledge, no study so far has attempted to estimate the treatment effects of *combinations of* university subjects and tiers of prestige in China. Using the first three waves of the China Family Panel Studies, we are able to identify both the subjects studied and tiers of HE prestige. We take advantage of the rich information on the respondent's school cohort and *hukou* status at age 12 and the mother's age and education to estimate the *simultaneous* choice of subject and tier of prestige of higher education institutions faced by Senior High School graduates. These factors are shown to be significant determinants of HE types defined by the 3x3 combinations of subjects and tiers.

By allowing for all possible combinations of university tiers and subjects in the students' HE choice set (even though our modest sample size limits the number of groups we can accommodate in practice), we do not impose arbitrary restrictions on the sequencing or interactions of university subjects and tiers of prestige. This modelling strategy also fits well with the Chinese college admissions system under which students list preferences over university courses, as defined by subjects at specific institutions.

Using the doubly robust Inverse Probability Weighted Regression Adjustment (IPWRA) method to account for selection on observables into both subjects and tiers, we find strong evidence that OLS and RE substantially underestimate the treatment effect of attending more prestigious universities, for graduates of both genders in Law, Economics and Management (LEM). These findings are robust to treating subjects as predetermined for the selection into HEIs by tiers of prestige. To a large extent, this reflects the reality that subject choice at the time of HE applications are heavily restricted by the streaming in Senior High Schools which takes place two years before HE enrolment. Moreover, we present suggestive evidence that these findings might be partly driven by the recent substantial expansion of the HE sector in China, which could have intensified the competition for the most prestigious HEIs despite widening overall access. In particular, the returns to HE seem to have declined for most graduates due to the HE expansion, with the exception of LEM and Other non-STEM Subjects graduates from Key Universities.

This study has certain limitations which are worth highlighting. First of all, the sample size is still relatively small, especially for the post-expansion analysis. Secondly, the lack of measures of prior educational attainment from secondary schools, such as the actual academic tracks and subjects chosen, imposes limits on feasible identification strategies.

Nevertheless, this study represents an attempt at estimating the causal effect of the returns to HE tiers and subjects, which have important policy implications and are of wide public interest in China. More causal studies are needed in the future before we get a better understanding on this important issue.

## References:

- Altonji, Joseph G, Blom, Erica, Meghir, Costas, 2012. Heterogeneity in human capital investments: High School curriculum, college major, and careers. *Annual Review of Economics* 4, 185-223.
- Arcidiacono, Peter, 2004. Ability sorting and the returns to college major. *Journal of Econometrics* 121(1-2):343-75
- Belfield, Chris, Britton, Jack, Buscha, Franz, Dearden, Lorraine, Dickson, Matt, van der Erve, Laura, Sibiet, Luke, Vignoles, Anna, Walker, Ian, Zhu, Yu, 2018 The relative labour market returns to different degrees. Department for Education Research Report. <https://www.gov.uk/government/publications/undergraduate-degrees-relative-labour-market-returns>.
- Berger, Mark C., 1988. Predicted future earnings and choice of college major. *Industrial & Labor Relations Review* 41(3): 418-429.
- Broecke, Stijn, 2012. University selectivity and earnings: Evidence from the UK data on applications and admissions to university. *Economics of Education Review* 31, 96-107.
- Chevalier, Arnaud, Colon, Gavan, 2003. Does it pay to attend a prestigious university? IZA DP No. 848.
- Dai, Fengyan, Cai, Fang, Zhu, Yu, 2018. Returns to Higher Education in China - Evidence from the Great Higher Education Expansion using Fuzzy Regression Discontinuity, *mimeo*, presented to the 2018 China Meeting of Econometric Society.
- Dale, Stacy Berg, Krueger, Alan B., 2002. Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables. *Quarterly Journal of Economics* 117, 1491-1527.
- Department for Education, 2017. Graduate outcomes for all subjects by university. <https://www.gov.uk/government/statistics/graduate-outcomes-for-all-subjects-by-university>. Accessed 2<sup>nd</sup> April 2018.
- Feng, Le-an, 2012. Who has more accessibility to Higher Education? *Education Research Monthly*, 2012 (8), 12-15 (*in Chinese*).
- Hussain, Iftikhar, McNally, Sandra, Telhaj, Shqiponja, 2009. University quality and graduate wages in the UK. Centre for the Economics of Education (CEE) DP 99.



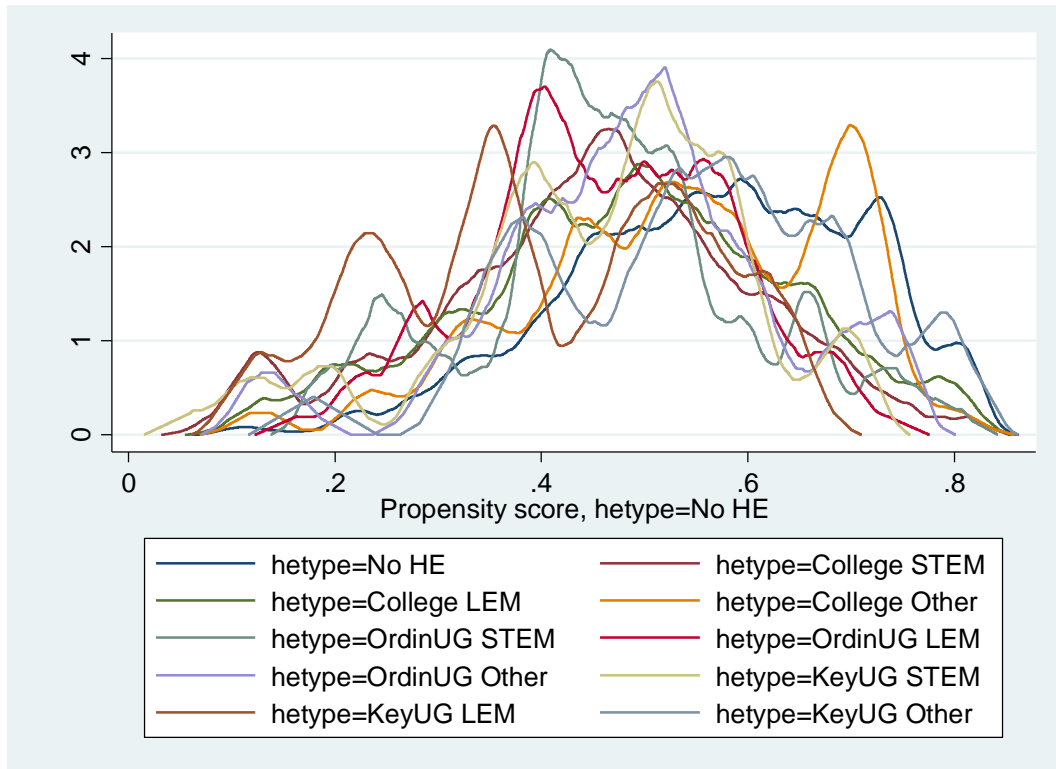
- Imbens, Guido W., Wooldridge, Jeffrey M., 2009. Recent developments of the econometrics of program evaluation. *Journal of Economic Literature* 47, 5-86.
- Kirkeboen, Lars J., Leuven, Edwin, Mogstad, Magne, 2016. Field of study, earnings, and self-selection. *Quarterly Journal of Economics* 131, 1057-1111.
- Li, Hongbin, Ma, Yueyuan, Meng, Lingsheng, Qiao, Xue, Shi, Xinzehng, 2017. Skill complementarities and returns to higher education: Evidence from college enrollment expansion in China. *China Economic Review* 46, 10-26.
- Ministry of Education of the PRC, 2018. Educational Statistics in 2016. [http://en.moe.gov.cn/Resources/Statistics/edu\\_stat\\_2016/2016\\_en01/](http://en.moe.gov.cn/Resources/Statistics/edu_stat_2016/2016_en01/). Accessed 2 April, 2018.
- OECD (Organisation for Economic Co-Operation and Development), 2016. Education in China – A Snap Shot. OECD, Paris.
- Paglin, Morton, Rufolo, Anthony M., 1990. Heterogeneous human capital, occupation choice, and male-female earnings differences. *Journal of Labor Economics* 8(1):123-144.
- Sheng, Xiaoming, 2017. Cultural capital, family background and education: choosing university subjects in China. *British Journal of Sociology of Education* 38(5), 721-737.
- Walker, Ian, Zhu, Yu, 2011. Differences by Degree: Evidence of the Net Financial Rates of Return to Undergraduate Study for England and Wales. *Economics of Education Review* 30 (6) 1177-1186.
- Walker, Ian, Zhu, Yu, 2018. University selectivity and the relative returns to higher Education: Evidence from the UK. *Labour Economics Special Issue* (in press). <https://doi.org/10.1016/j.labeco.2018.05.005>.
- Whalley, John, Zhao, Xiliang, 2013. The contribution of Human Capital to China's economic growth. *China Economic Policy Review* (CEPR), World Scientific Publishing Co. Ltd., vol. 2(01), 1350001-1-1.
- Wooldridge, Jeffrey M., 2007. Inverse probability weighted estimation for general missing data problems. *Journal of Econometrics* 141, 1281-1301.
- Wooldridge, Jeffrey M., (2010) *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge, MA, 2<sup>nd</sup> Ed.

Wu, Yaowu, Zhao, Quan, 2010. Higher Education expansion and employment of university graduates. *Economic Research* 2010 (9), 93-108 (*in Chinese*).

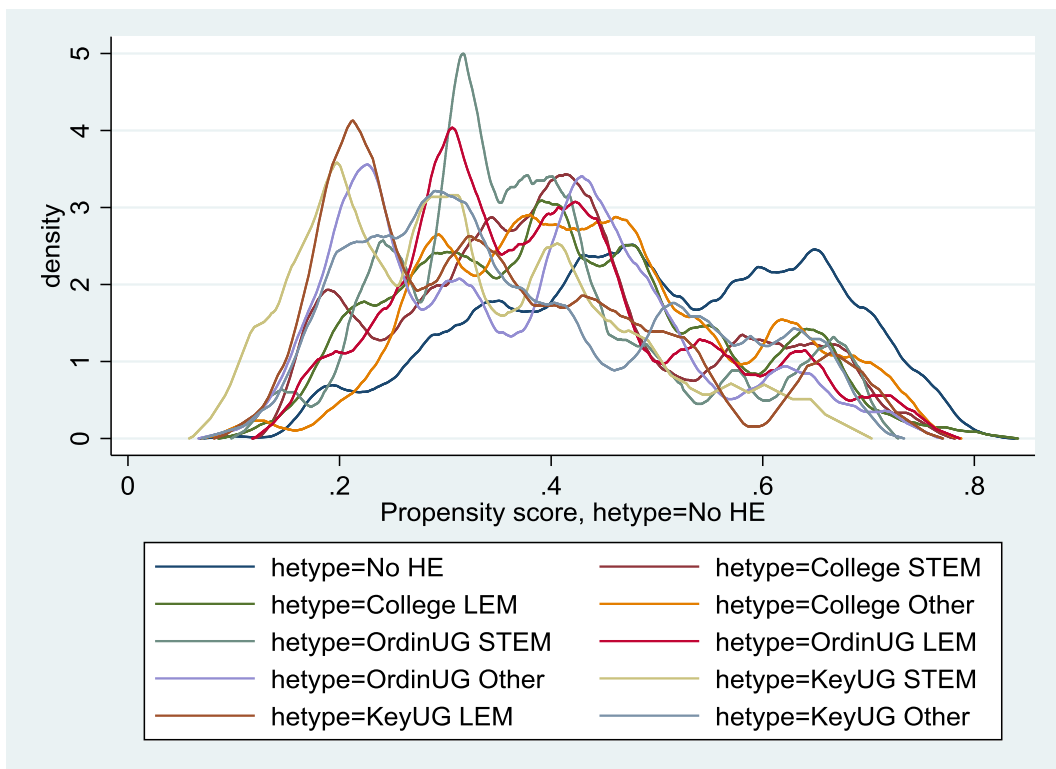
Xie, Yu, Hu, Jingwei, 2014. An introduction to the China Family Panel Studies (CFPS). *Chinese Sociological Review* 47, 3-29.

Zhu, Min, 2014. College admissions in China: A mechanism design perspective. *China Economic Review* 30, 618-631.

## Appendix



**Figure A1: Overlap Plots, Men**



**Figure A2: Overlap Plots, Women**

**Table A1: Full Set of IPWRA Treatment Effects of Table 4, Men**

Outcome (wage) eq.	No HE	College STEM	College LEM	College Other	OrdinUG STEM	OrdinUG LEM	OrdinUG Other	KeyUG STEM	KeyUG LEM	Key Other
Age	0.032* (0.016)	0.002 (0.069)	0.032 (0.039)	-0.011 (0.051)	0.077* (0.043)	0.025 (0.073)	-0.024 (0.038)	0.224*** (0.057)	0.189** (0.090)	-0.034 (0.083)
Age sq	-0.000** (0.000)	-0.000 (0.001)	-0.000 (0.000)	0.000 (0.001)	-0.001 (0.001)	-0.000 (0.001)	0.000 (0.000)	-0.002*** (0.001)	-0.002* (0.001)	0.001 (0.001)
Urban	0.106** (0.047)	0.234* (0.124)	0.375*** (0.124)	0.203** (0.089)	0.205 (0.277)	-0.299 (0.182)	0.055 (0.092)	-0.346** (0.158)	-2.270*** (0.792)	0.662*** (0.217)
East	0.270*** (0.039)	0.374*** (0.085)	0.441*** (0.088)	0.052 (0.135)	0.502*** (0.180)	0.441*** (0.148)	0.427*** (0.122)	0.462*** (0.163)	1.057*** (0.172)	0.310* (0.175)
Constant	6.547*** (0.317)	7.269*** (1.164)	6.369*** (0.784)	7.431*** (0.980)	5.261*** (0.771)	7.009*** (1.333)	7.783*** (0.823)	3.165*** (1.035)	5.475*** (1.578)	7.535*** (1.751)
Treatment eq.										
School cohort		0.024 (0.016)	0.017 (0.013)	-0.042** (0.021)	0.025 (0.019)	0.020 (0.023)	0.019 (0.022)	0.073*** (0.022)	-0.005 (0.037)	0.045 (0.035)
Non-agricul hukou at 12		-0.035 (0.153)	0.199 (0.144)	-0.626*** (0.242)	-0.693*** (0.215)	0.103 (0.222)	-0.857*** (0.257)	-0.044 (0.299)	-0.293 (0.332)	0.176 (0.366)
Mother's year of birth		0.020 (0.013)	-0.011 (0.011)	0.019 (0.018)	0.005 (0.018)	-0.006 (0.018)	0.002 (0.022)	-0.006 (0.018)	0.041 (0.037)	-0.044** (0.021)
Mother Primary Edu		0.490*** (0.181)	0.414*** (0.158)	0.065 (0.277)	1.134*** (0.255)	0.418* (0.253)	0.665** (0.304)	0.732*** (0.279)	0.685 (0.448)	0.968** (0.391)
Mother Junior High		0.516** (0.205)	0.237 (0.204)	0.895*** (0.274)	0.829*** (0.305)	0.602* (0.322)	0.184 (0.405)	-0.960* (0.517)	0.287 (0.537)	0.166 (0.537)
Mother Senior High		1.133*** (0.242)	0.983*** (0.222)	1.175*** (0.372)	2.045*** (0.328)	0.974*** (0.330)	0.685 (0.494)	0.434 (0.468)	1.933*** (0.444)	0.112 (0.779)
Mother College/Uni		2.533*** (0.426)	1.933*** (0.426)	2.232*** (0.579)	1.024 (1.078)	0.228 (1.084)	2.463*** (0.644)	2.703*** (0.586)	2.710*** (0.745)	1.607 (1.130)
Urban		0.573*** (0.189)	0.922*** (0.204)	0.068 (0.226)	0.766*** (0.272)	2.925*** (0.714)	0.693** (0.316)	0.637* (0.333)	1.994*** (0.736)	-0.023 (0.425)
East		-0.584*** (0.142)	-0.323** (0.127)	-0.299 (0.184)	-0.265 (0.188)	-0.558*** (0.208)	-1.075*** (0.259)	0.231 (0.246)	-0.726** (0.315)	-0.012 (0.360)
Constant		-89.007*** (16.270)	-15.865 (13.832)	42.753** (20.919)	-61.336*** (19.052)	-33.174 (22.473)	-44.072* (23.031)	-136.905*** (26.328)	-74.650** (30.891)	-6.551 (36.778)

Note: Robust standard errors in parentheses. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . OrdinUG and KeyUG stand for Ordinary and Key universities, respectively. Other controls include dummies for survey year and months.

**Table A2: Full Set of IPWRA Treatment Effects of Table 4, Women**

Outcome (wage) eq.	No HE	College STEM	College LEM	College Other	OrdinUG STEM	OrdinUG LEM	OrdinUG Other	KeyUG STEM	KeyUG LEM	Key Other
Age	0.022 (0.018)	0.023 (0.049)	0.007 (0.030)	0.032 (0.047)	0.032 (0.069)	-0.144 (0.126)	-0.016 (0.105)	-0.057 (0.143)	-0.072 (0.063)	-0.175*** (0.052)
Age sq	-0.000 (0.000)	0.000 (0.001)	0.000 (0.000)	-0.000 (0.001)	0.000 (0.001)	0.002 (0.001)	0.000 (0.002)	0.001 (0.002)	0.002* (0.001)	0.002*** (0.001)
Urban	0.400*** (0.077)	-0.100 (0.137)	0.242** (0.102)	0.095 (0.141)	-0.236* (0.129)	0.488 (0.334)	0.519 (0.491)	0.220 (0.376)	0.327 (0.207)	0.363*** (0.106)
East	0.323*** (0.046)	0.645*** (0.098)	0.659*** (0.069)	0.416*** (0.104)	0.562*** (0.107)	0.779*** (0.226)	0.339** (0.156)	0.435* (0.234)	-0.024 (0.091)	0.279*** (0.089)
Constant	6.213*** (0.319)	6.377*** (0.832)	6.592*** (0.542)	6.229*** (0.942)	6.241*** (1.351)	9.942*** (2.349)	7.129*** (1.329)	8.197*** (2.723)	8.070*** (1.248)	10.275*** (0.974)
Treatment eq.										
School cohort		0.002 (0.019)	0.017 (0.016)	0.011 (0.022)	0.105*** (0.024)	0.037 (0.026)	0.068*** (0.023)	0.029 (0.039)	0.093*** (0.035)	0.002 (0.033)
Non-agricul hukou at 12		-0.127 (0.180)	0.214 (0.139)	-0.326* (0.193)	-0.350 (0.265)	0.602** (0.260)	-0.323 (0.224)	-0.408 (0.321)	0.858** (0.348)	-0.794*** (0.274)
Mother's year of birth		0.034** (0.016)	-0.003 (0.013)	0.006 (0.018)	-0.069*** (0.018)	0.002 (0.021)	-0.014 (0.020)	0.027 (0.029)	-0.027 (0.032)	0.013 (0.025)
Mother		0.286 (0.223)	0.598*** (0.180)	0.804*** (0.242)	0.573 (0.356)	0.113 (0.318)	0.629** (0.308)	0.847 (0.691)	0.254 (0.453)	0.523 (0.414)
Primary Edu		0.724*** (0.246)	0.882*** (0.201)	1.101*** (0.264)	1.575*** (0.389)	0.763** (0.300)	0.562 (0.361)	1.733** (0.685)	0.486 (0.425)	1.522*** (0.383)
Mother Junior		1.159*** (0.279)	1.403*** (0.232)	0.821** (0.369)	1.833*** (0.465)	1.031*** (0.360)	1.652*** (0.358)	2.620*** (0.654)	1.555*** (0.414)	2.269*** (0.391)
High		1.102** (0.547)	1.602*** (0.422)	0.771 (0.775)	2.397*** (0.669)	-0.172 (1.073)	2.660*** (0.508)	3.849*** (0.774)	2.506*** (0.613)	2.534*** (0.730)
Mother Senior		0.318 (0.250)	0.436** (0.219)	-0.055 (0.242)	0.431 (0.347)	0.792 (0.493)	1.035*** (0.376)	2.093** (1.029)	-0.339 (0.454)	0.767* (0.439)
College/Uni		-0.408** (0.170)	0.261** (0.126)	-0.360* (0.187)	-0.136 (0.232)	-0.417* (0.229)	-0.874*** (0.223)	0.443 (0.323)	0.971*** (0.286)	0.010 (0.253)
Urban		-73.576*** (21.424)	-29.395* (16.354)	-35.343 (24.408)	-76.736** (29.904)	-79.562*** (22.465)	-110.869*** (23.327)	-118.501*** (39.543)	-134.729*** (29.192)	-33.079 (26.960)

Note: Robust standard errors in parentheses. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . OrdinUG and KeyUG stand for Ordinary and Key universities, respectively. Other controls include dummies for survey year and months.

**Table A3: Higher Education Subjects and Tiers Choice, by family background**

%	By HE Subjects				By HE Tiers			
	Senior High School	STEM	LEM	Other	Senior High School	College	OrdinUG	KeyUG
<b>By mother's education (share in parentheses), Agricultural <i>hukou</i> at 12, N=1,489 (52.9%)</b>								
Illiterate (45.8%)	65.1	11.3	13.9	9.7	65.1	23.6	7.5	3.8
Primary (30.8%)	49.2	17.7	18.5	14.6	49.2	30.7	12.6	7.4
Junior High (17.7%)	48.5	23.9	14.8	12.9	48.5	30.7	12.9	8.0
Senior High (5.2%)	40.3	22.1	24.7	13.0	40.3	33.8	20.8	5.2
College or above (0.5%)	0	42.9	14.3	42.9	0	57.1	14.3	28.6
<b>By mother's education (share in parentheses), Non-agricultural <i>hukou</i> at 12, N=1,324 (47.1%)</b>								
Illiterate (19.9%)	66.9	9.9	17.1	6.1	66.9	20.9	8.4	3.8
Primary (23.1%)	51.6	17.3	22.9	8.2	51.6	30.1	10.5	7.8
Junior High (28.0%)	39.9	19.7	28.0	12.4	39.9	37.4	16.2	6.5
Senior High (23.3%)	27.2	24.3	34.3	14.2	27.2	38.2	20.7	13.9
College or above (5.7%)	17.3	29.3	36.0	17.3	17.3	38.7	21.3	22.7

Note: Distinct individuals in Wave 1. OrdinUG and KeyUG stand for Ordinary and Key universities, respectively.