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

DISCUSSION PAPER SERIES

IZA DP No. 12071

Returns to Higher Education Subjects and Tiers in China: Evidence from the China Family Panel Studies

Lili Kang
Fei Peng
Yu Zhu

JANUARY 2019

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12071

Returns to Higher Education Subjects and Tiers in China: Evidence from the China Family Panel Studies

Lili Kang
Shanghai Lixin University of Accounting and Finance

Fei Peng
Shanghai Lixin University of Accounting and Finance

Yu Zhu
University of Dundee and IZA

JANUARY 2019

IZA – Institute of Labor Economics

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

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

www.iza.org

ABSTRACT

Returns to Higher Education Subjects and Tiers in China: Evidence from the China Family Panel Studies*

Using the recent China Family Panel Studies, we identify the subjects studied by college (2–3 years) graduates and university (4–5 years) graduates. For the university graduates, 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, hukou status at age 12, and the mother's age and education to estimate university applicants' simultaneous choice of subject and tier of prestige of higher education institutions (HEIs). Using the doubly robust Inverse Probability Weighted Regression Adjustment method to account for selection – on observables – into subjects and tiers, our treatment effect estimates suggest that pooled OLS and random-effect models substantially underestimate the effect of attending universities that are more prestigious for graduates of both genders in law, economics, and management (LEM). We also demonstrate that the recent massive expansion of the higher education sector resulted in reduced returns to HE for all graduates, except for graduates who studied LEM or Other non-STEM (sciences, technology, engineering and math/medicine) subjects at the most prestigious universities. The results are robust to treating subjects as predetermined for the selection into HEIs by tiers of prestige.

JEL Classification: I23, I26

Keywords: returns to university tier and subjects, China, inverse probability weighted regression adjustment, higher education expansion

Corresponding author:
Yu Zhu
School of Business
University of Dundee
3 Perth Road
Dundee, DD1 4HN
United Kingdom
E-mail: yuzhu@dundee.ac.uk

* 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).
We thank the Institute of Social Science Survey at Peking University for allowing access to the China Family Panel Studies' data and participants of the Asian and Australian Society of Labour Economics 2018 Conference for comments. All errors are ours.

1. Introduction

There is an enormous amount of literature on the returns to education. Most of the empirical literature has focused 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, and this phenomenon is partly due to data availability limitations.

Returns to different university subjects [known as *college majors* in the United States (US)] vary substantially [see e.g., Paglin and Rufolo (1990), Arcidiacono (2004) and Altonji et al (2012) in the US and Walker and Zhu (2011) for the United Kingdom (UK)]. Similarly, attending universities that are more selective (prestigious) has been associated with higher wages and earnings.

Although various studies have indicated that the development and expansion of the Chinese education system in general, and the higher education (HE) system in particular, has played a role in China's remarkable economic growth over the last 4 decades, we observed little agreement regarding its relative contribution. Using a growth accounting model that treats human capital as the opportunity cost foregone, Whalley and Zhao (2013) suggested that increases in human capital contributed to as much as 38.1% of economic growth in China from 1978–2008, and that percentage was even higher from 1999–2008, a period of rapid HE expansion which witnessed an increase of annual enrollment by more than fourfold. Other growth accounting studies, however, have pointed to a more modest contribution of average human capital accumulation to per capita GDP growth: 10%–15%, see for example, Wang and Yao (2003) and Zhu (2012). Liao et al. (2017) suggested that rural–urban migration accounts for approximately 11% of per capita output in China from 1981–2007, of which 60% can be attributed to education-based migration.

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 we observed surprisingly little empirical evidence in the literature. In this paper, we address two specific research questions. First, to what extent, if any, do the returns to HE in China vary by the subject studied and tiers of prestige of the institution attended? Second, how does the recent HE expansion affect the returns to HE differentially by subject and tiers?

We make the following contributions to the empirical literature on returns to HE. Firstly, according to our review of the literature, this is the first attempt to estimate the treatment effects of HE subjects and tiers of prestige on earnings in China. Secondly, we show that OLS substantially underestimates the effect of attending universities that are more prestigious, 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 observe that our results are robust to treating subjects as predetermined for the selection into higher education institutions (HEIs) by tiers of prestige. This result is consistent with the tracking choice made by students in the Chinese education system into the science or arts subjects during the Senior High School restricting the subsequent HE subject choice at the time of university application.

The remainder of the paper is organized as follows. Section 2 is the review the relevant literature, with special reference to China. Section 3 is a discussion of the institutional background. Section 4 is a presentation of the data. Section 5 is a discussion 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

The economic literature on returns to HE types can be classified into two strands. The first strand concerns returns to subjects (or majors, in the US literature) while typically holding the prestige tier constant, and the second concern is returns to prestige tier (a.k.a selectivity) while typically treating subjects as given. Both strands are dominated by descriptive analyses. In this brief review, we focus on the studies that have attempted to estimate the causal effect of the subject or prestige type.

Altonji et al. (2012) surveyed the empirical literature on the demand for and returns to college major, allowing for the effects of high school curriculum, and observed that most of the studies are from the US and the UK, which we presume is because of data availability.

Paglin and Rufolo (1990) highlighted the importance of mathematical ability in determining field choice for US college students, which is consistent with earnings maximization by major. They observed that graduates with above average Graduate Record Exam quantitative scores for their undergraduate field tend to switch to fields requiring higher average scores. Using a dynamic model of college and major choice that allows for switching and dropout, Arcidiacono (2004) focused on ability sorting across majors in the National Longitudinal Study of the Class of 1972 (NLS72). Although 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, based on 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. By contrast, Hamermesh and Donald (2008) demonstrated that overlooking non-response bias in survey data leads to overestimation of the earnings differentials across college majors in the US.

O’Leary and Sloane (2005) conduct one of the first UK studies that focuses on the heterogeneous returns to broad and narrowly defined subjects by using the Labour Force Survey. Their results highlight the substantial variation in returns across degree subjects and by gender in the UK. Using a survey of a cohort of young UK graduates linked to administrative records of academic attainment and family background, Chevalier (2011) documents large heterogeneity in mean wages between subjects and an even larger variation by unobserved individual characteristics within subjects.

As for the causal studies on the effect of college selectivity on earnings, most studies have used matching methods that assume selection on observables only [see for example, Chevalier and Conlon (2003) and Hussain et al. (2009) for the UK, both of whom observe a modest return to attending universities that are more selective of approximately 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 students who also applied to but were rejected by the same set of colleges, they observe little evidence of returns to attending colleges that are more prestigious in the US for students with the same ability. Following the Dale and Krueger method, Broecke (2012) compares UK students who satisfied the conditional offers for their first-choice to students who applied to the same universities but attended their second-choice universities due to not fulfilling the conditions of their preferred offer. He finds that one standard deviation in selectivity leads to a 7% increase in earnings in the UK.

Walker and Zhu (2018) attempt to estimate the treatment effect of university selectivity and subjects for the UK. Using the IPWRA approach to allow for selection on observables into subjects and institution types, they find strong differences in returns to selectivity by subject. Belfield et al. (2018a, 2018b) apply the IPWRA method to the new Longitudinal Education Outcomes administrative dataset to account for variation in course selectivity and student characteristics to estimate the relative and absolute labor market returns to different degrees in the UK, respectively.

This paper follows the same IPWRA methodology, which is discussed in Section 5. Compared with 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 body of literature, mostly published in Chinese. The literature review by Feng (2012) concludes that most of studies from the perspective of education or sociology show that inequality has been 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 younger cohorts. Using 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 HE expansion increases monthly wage income by 21%, compared with an OLS estimate of 8%.

Few Chinese studies have examined the choice of university subjects and tiers in China. One exception is Sheng (2017), who shows that although there is little difference in subject choice across social class, secondary school students from high-income families in Beijing 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, which have a duration of 3 years. Students in the Senior High Schools are streamed into the science or arts tracks for the last 2 years of upper secondary education (OECD 2016).

After 12 years of schooling, secondary school graduates can apply to colleges and universities through a centralized admissions system that proceeds sequentially in tiers based on the standardized National College Entrance Examinations (*gaokao*).[1] Using real admissions data from a top school in a top university in China, Wu and Zhong (2014) document substantial variation at the provincial level on the sequencing of the submission of the school and subject preference by applicants. Colleges and universities in China can be classified into three tiers in descending order of prestige and entry requirements: Key Universities (mostly national project 985 and project 211 universities), Ordinary Universities, and vocational training colleges. The duration of study for the Key Universities and Ordinary Universities is 4–5 years, and these programs lead to a bachelor's degree. The duration of study for the vocational training colleges is 3 years, and these programs lead to a college diploma.[2]

Admissions in the second tier start after the assignments in the first tiers are finalized, and so forth. Each applicant submits a lexicographic list to the provincial student placement office that indicates their HEI (i.e., colleges and universities) preferences and preferences within each HEI, their preference regarding subjects. Important for this analysis, university applicants must consider the tier of the HEI and the subject at a given HEI simultaneously, which defines a university course. These considerations are an important feature of the Chinese educational system that must be considered in the econometric analysis.

[1] The description of the Chinese College Admissions system relies heavily on Zhu (2014), who 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 occurred in 2013.

[2] There are nearly 2,600 regular HEIs (i.e., excluding, e.g., adult education) in China employing 1.6 million academic staff as of 2016 (Ministry of Education 2018).

Another important institutional feature of China is the *hukou*, or household registration system, which classifies people into rural and urban status at birth, usually according to the mother's *hukou* status. Education resources at the primary and secondary level are highly unequal and in favor of urban residents in China, who have better access to HEIs than their rural counterparts, especially to the most prestigious universities and colleges. 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 that 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 also exploit *hukou* status at age 12 as a key determinant of the choice of HE types available to the HE applicants in the formal analysis.

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 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 subjects of study at each level of post-secondary education, including the Senior High School, college, undergraduate and postgraduate levels. For university graduates, the CFPS also asks them in the first wave (i.e., the 2010 survey) about the tier of prestige, for example, national key and key.

Our sample consists of all individuals aged 20–60 years whose highest qualification is Senior High School, College, or University in the first wave of the CFPS, which was conducted in 2010. We exclude individuals whose highest qualification is Junior High School or below, because HE choice is irrelevant to them. We also exclude the small number of respondents with a master or PhD level qualification because there might be important unobservables that 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 labor 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 limitations, we use a 3 by 3 grouping of HE types, namely, three subjects consisting of STEM (sciences, technology, engineering and math/medicine), LEM, and Other Subjects,[3] and three 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, the sample is 2,813 distinct individuals, and 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.

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 dominate those of Senior High School graduates, except toward the very end of the career, where the earnings become a bit noisy due to small cells and possible early retirement. Notably, 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.

[3] 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.

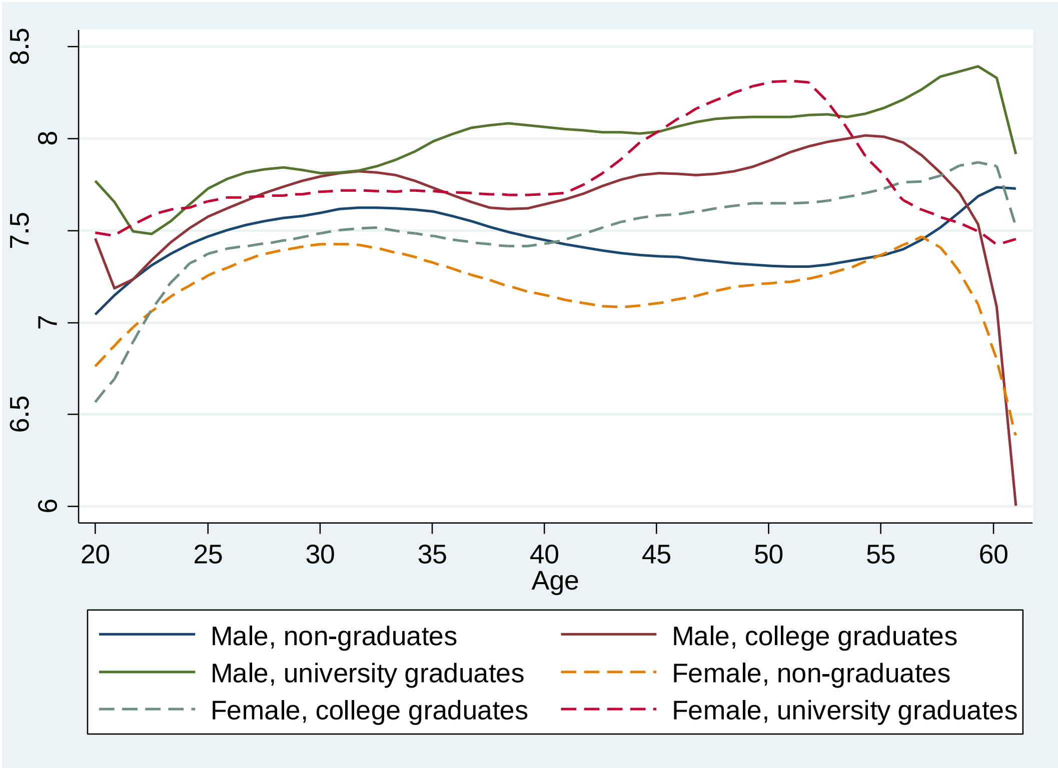

Figure 1: Smoothed age-earnings profiles by level of qualifications and gender

Table 1A shows the relative frequencies by gender. Notably 46.7% of males in the sample are graduates, and 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 reports the summary statistics for key variables in Wave 1 by gender. The mean real monthly earnings in January 2009 constant price are RMB 2,443 and 1,849 for men and women, respectively. Compared with men, women are almost 3 years younger and more likely to have had an advantageous background as proxied by a non-agricultural *hukou* at age 12 and mother's education level. Although women are more likely to live in urban areas than men, 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 men and women.[4] Male college graduates earn 0.29 log points more than their Senior High School counterparts and 0.11 log points less than male Ordinary University graduates. Although Key Universities and Ordinary Universities take the same time to complete, we observe a staggering 0.25 log points' earnings difference among male university graduates. Female college graduates earn 0.25 log points more than their Senior High

[4] For simplicity, we interpret a slope coefficient *b* in the log wage equation in terms of log points, which approximate a 100*b* 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.

School counterparts and 0.19 log points less than male Ordinary University graduates. The earnings premium for attending Key Universities for women is 0.22 log points. Notably, STEM graduates have the lowest earnings at college level, but the highest earnings at Key University level for both genders, implying a higher return to selectivity to study those subjects. Finally, the gender difference in earnings is more or less constant across all education levels and types: approximately 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^{***}
All Colleges	*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^{***}
All Ordinary Universities	*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^{**}
All Key Universities	*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. * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

5. Methodology

We begin with a 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 this study with the 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 the students' ability, because high ability students earn more on average regardless of HE attainment. Unfortunately, the CFPS does not contain information on standardized test scores, for example, 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 because 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 an 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 exploit discontinuities associated with admission requirements. For the purpose of our analysis, the choice set is so large (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 available (see Kirkeboen *et al* (2016) for an exception).

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

In particular, we explore the "doubly robust" IPWRA estimator (see Wooldridge 2007, Wooldridge 2010 chapter 13, and Imbens and Wooldridge 2009). The IPWRA estimates the ATE of any HE type and allows for selection into a particular HE type by using multinomial logit model in the first step. In the second step, this estimator then estimates an OLS regression of log earnings by using the reweighted data, using the inverse of the predicted probabilities from the first step as the weights. In other words, the 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.

IPWRA is *doubly* robust in the sense that only one of the two steps must 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. Although 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, 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, as well as survey years. We deliberately choose this parsimonious specification, which later facilitates creating the IPWRA model.[5] However, the coefficients on HE types in POLS and RE models are robust to specifications allowing for provincial fixed-effects. Notably, 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 nine treatment groups.

The RE and POLS estimates are largely indistinguishable for both genders. Thus, in the following, we comment on the RE results. For men, attending a 3-year college yields a return between 24%–30%, 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 have a return between 33% and 41%, again with the lowest returns for STEM. Men attending the most prestigious Key Universities have a return between 49% and 64%, with substantially lower returns for graduates studying Other Subjects than STEM or LEM. Thus, we observe that returns to attending HEIs that are more prestigious vary by subject. A model that fails to allow for the interaction effects is likely to yield biased estimates.

For women, we observed returns similar to those for men, except for Ordinary University STEM graduates and Key-University Other Subjects graduates, which are substantially higher.

For men, having a non-agricultural *hukou* at age 12 is associated with 7% higher returns; living in urban areas is associated with approximately 14% higher monthly earnings compared with living in rural areas; and living in the more developed East Region is associated with approximately 36% higher earnings than living in the central or western regions. Moreover, men's earnings are in general increasing in mother's

[5] Controlling for occupation, industry, and ownership types of employers will lead to endogeneity bias, because all these are post-education choices that, in principle, simultaneously affect earnings.

education qualifications, although the coefficients are not always statistically significant. For women, we observe no significant link between *hukou* status at 12, but the wage premiums for living in urban areas or the East Region are even higher than for men.

Table 3: OLS and RE estimates by gender

	Men		Women	
	Pooled OLS	RE	Pooled OLS	RE
	(1)	(2)	(3)	(4)
College STEM	0.254***	0.239***	0.217***	0.221***
	(0.049)	(0.048)	(0.064)	(0.061)
College LEM	0.305***	0.295***	0.229***	0.228***
	(0.044)	(0.045)	(0.047)	(0.046)
College Other	0.295***	0.281***	0.289***	0.290***
	(0.069)	(0.070)	(0.076)	(0.072)
OrdinUG STEM	0.346***	0.333***	0.517***	0.516***
	(0.093)	(0.086)	(0.088)	(0.086)
OrdinUG LEM	0.391***	0.386***	0.398***	0.405***
	(0.094)	(0.088)	(0.113)	(0.104)
OrdinUG Other	0.434***	0.411***	0.449***	0.438***
	(0.074)	(0.079)	(0.087)	(0.082)
KeyUG STEM	0.621***	0.642***	0.606***	0.609***
	(0.084)	(0.087)	(0.154)	(0.148)
KeyUG LEM	0.608***	0.624***	0.579***	0.579***
	(0.106)	(0.112)	(0.096)	(0.094)
KeyUG Other	0.454***	0.485***	0.530***	0.534***
	(0.120)	(0.122)	(0.077)	(0.079)
Age	0.023*	0.038***	0.028*	0.036**
	(0.012)	(0.012)	(0.015)	(0.014)
Age sq	-0.000	-0.000***	-0.000	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)
Non-agricultural *hukou*	0.080**	0.073**	0.002	-0.000
at age 12	(0.036)	(0.035)	(0.041)	(0.039)
Mother Primary Edu	0.072*	0.071*	0.026	0.030
	(0.039)	(0.038)	(0.052)	(0.050)
Mother Junior High Edu	0.081	0.100**	0.140***	0.142***
	(0.052)	(0.049)	(0.052)	(0.051)
Mother Senior High Edu	-0.018	-0.014	0.124*	0.139**
	(0.055)	(0.054)	(0.064)	(0.062)
Mother College+ Edu	0.209*	0.240**	0.192*	0.212**
	(0.116)	(0.115)	(0.101)	(0.100)
Urban	0.141***	0.141***	0.237***	0.228***
	(0.046)	(0.045)	(0.071)	(0.067)
East	0.357***	0.363***	0.444***	0.436***
	(0.031)	(0.031)	(0.037)	(0.035)
Constant	6.515***	6.243***	6.011***	5.895***
	(0.231)	(0.222)	(0.269)	(0.247)
Observations	3,402	3,402	2,395	2,395
R^2	0.157		0.195	

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

Table 4 focuses on the corresponding IPWRA estimates of the ATEs of the HE types relative to Senior High School graduates.[6] 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 key determinants of educational choices (see e.g., Berger 1988).

Table 4: IPWRA ATE, by gender

	Men (1)	Women (2)
College STEM	0.268***	0.241***
	(0.044)	(0.051)
College LEM	0.311***	0.226***
	(0.041)	(0.037)
College Other	0.289***	0.284***
	(0.060)	(0.057)
OrdinUG STEM	0.470***	0.600***
	(0.051)	(0.069)
OrdinUG LEM	0.483***	0.396**
	(0.079)	(0.154)
OrdinUG Other	0.428***	0.385***
	(0.068)	(0.099)
KeyUG STEM	0.684***	0.720***
	(0.058)	(0.121)
KeyUG LEM	0.912***	0.832***
	(0.107)	(0.055)
KeyUG Other	0.616***	0.548***
	(0.081)	(0.051)
Observations	3,335	2,355

Note: Robust standard errors in parentheses. Clustering at the individual level for pooled OLS. * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$. OrdinUG and KeyUG stand for Ordinary and Key Universities, respectively. Other controls in the outcome (wage) equation include age, age squared, *hukou* at age 12, dummies for mother's education, and dummies for urban, East Region, and survey years. The full set of results in both the outcome and treatment equations (except for dummies for survey years) are presented in Tables A1 and A2 in the Appendix.

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

Table A3 in the Appendix tabulates the choice of HE subjects and HE tiers by mother's education, and 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, at least for *urban* hukou holders (notably, very few agricultural *hukou* holders had highly educated mothers). For instance, as the mother's education increases from illiteracy to university degree, the chance of an urban resident not going to HE decreases from 67% to 17% and that for going to the most selective Key Universities increases from 4% to 23%. There is also suggestive evidence that rural and urban residents might have different preferences regarding HE subjects, with the former more likely to choose STEM and the latter more inclined to choose LEM. These patterns are consistent with the literature that suggests that students with advantaged backgrounds from urban areas are over-represented in the most selective universities in China.

The small differences in sample sizes (less than 2% for both genders) between OLS/RE and the IPWRA estimates reflect that observations 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 observe good matches from the control group of non-HE graduates.

Compared with the RE estimates, the IPWRA returns to LEM graduates are much higher, by 0.29 and 0.25 log points for men and women, respectively, but only if they attend the most prestigious Key Universities. This result implies that ignoring the endogeneity of HE types is likely to lead to underestimation of the returns to attending universities that are more selective, 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 although our model might be satisfactory at predicting the selection into different tiers of HEIs, we were unable to precisely predict the subject choice due to an absence of information on prior attainment in different subjects. Ideally, we would prefer to have the respondents' grades in key subjects (e.g., math, physics, chemistry, Chinese, and history) from 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 estimate the effect of HE tiers ***conditional*** on subjects by using RE and IPWRA. Effectively, we are treating subjects as exogenous while focusing on HE selectivity. This method is a useful simplifying assumption of the reality, because streaming occurs at the beginning of Year 11 in China, when students must choose between the STEM and Arts track. By the time the students apply to colleges and universities at the end of Year 12, the subject choice is limited to some extent by the streaming choice made approximately 2 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. Regarding our main results, RE severely underestimates the returns to attending HEIs that are more prestigious for LEM graduates by as much as 0.32 log points, and accounting for endogeneity of HE tiers is observed to make little difference for graduates in Other Subjects. Moreover, the difference between RE and IPWRA is almost 0.16 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 0.33 log points. Notably, the IPWRA estimates for STEM at either Ordinary or Key Universities exceed their RE counterparts by more than 0.13 log points.

Overall, we conclude that our main results are robust to treating subjects as predetermined in HE choices. The results are reassuring regarding studies in the literature that have focused 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

	RE			IPWRA		
	STEM	LEM	Other	STEM	LEM	Other
	(1)	(2)	(3)	(4)	(5)	(6)
College	0.229***	0.299***	0.263***	0.253***	0.319***	0.264***
	(0.048)	(0.046)	(0.070)	(0.046)	(0.043)	(0.050)
Ordinary university	0.316***	0.386***	0.377***	0.474***	0.513***	0.436***
	(0.088)	(0.089)	(0.079)	(0.049)	(0.078)	(0.057)
Key university	0.647***	0.621***	0.510***	0.678***	0.937***	0.560***
	(0.087)	(0.114)	(0.125)	(0.059)	(0.113)	(0.081)
Observations	2,442	2,435	2,087	2,442	2,394	2,049

Note: Robust standard errors in parentheses, * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$. Other controls in the outcome (wage) equation include age, age squared, hukou at age 12, dummies for mother's education, and dummies for urban, East Region, and survey years.

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	0.211***	0.221***	0.266***	0.264***	0.254***	0.299***
	(0.061)	(0.046)	(0.072)	(0.052)	(0.037)	(0.057)
Ordinary university	0.505***	0.398***	0.397***	0.641***	0.397***	0.381***
	(0.088)	(0.103)	(0.082)	(0.076)	(0.146)	(0.108)
Key university	0.592***	0.568***	0.521***	0.762***	0.900***	0.563***
	(0.145)	(0.097)	(0.078)	(0.142)	(0.055)	(0.052)
Observations	1,428	1,624	1,449	1,420	1,615	1,424

Note: Robust standard errors in parentheses, * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$. Other controls in the outcome (wage) equation include age, age squared, hukou at age 12, dummies for mother's education, and dummies for urban, East Region, and survey years.

6.3. Heterogeneous effect of the HE expansion

Finally, we explore the heterogeneous effect of the HE expansion by HE tiers and subjects to address the second research question. The HE sector in China experienced an unpresented expansion from 1999–2018, with an annual enrollment that increased from approximately 1 million to 6 million. The most dramatic growth occurred between 1999 and 2001: approximately 40% annual growth per 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	0.284***	0.063	0.297***	0.120**
	(0.052)	(0.058)	(0.046)	(0.054)
College LEM	0.307***	0.140**	0.338***	0.161**
	(0.041)	(0.063)	(0.034)	(0.064)
College Other	0.381***	0.022	0.295***	0.104
	(0.064)	(0.078)	(0.065)	(0.082)
Ordinary STEM	0.554***	-0.013	0.662***	-0.032
	(0.072)	(0.091)	(0.050)	(0.073)
Ordinary LEM	0.423***	0.301***	0.426***	0.324***
	(0.066)	(0.107)	(0.091)	(0.113)
Ordinary Other	0.497***	0.157	0.473***	0.191
	(0.074)	(0.101)	(0.063)	(0.167)
KeyUG STEM	0.804***	0.402***	0.916***	0.464***
	(0.095)	(0.103)	(0.067)	(0.088)
KeyUG LEM	0.681***	0.430***	0.825***	1.007***
	(0.104)	(0.113)	(0.053)	(0.190)
KeyUG Other	0.489***	0.609***	0.543***	1.065***
	(0.093)	(0.152)	(0.068)	(0.118)
Female	-0.239***	-0.216***		
	(0.029)	(0.040)		
Observations	4,025	1,772	3,260	1,434

Note: Robust standard errors in parentheses, * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$. Other controls in the outcome (wage) equation include age, age squared, hukou at age 12, dummies for mother's education, and dummies for urban, East Region, and survey years.

We classified people born in August 1979 or before as the pre-expansion cohort and people born in September 1979 or later as post-expansion.[7] Approximately 30% of the sample are post-expansion. Due to the small sample sizes, we pool gender in Table

[7] 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.

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

Although the RE and IPWRA results are remarkably similar for the pre-expansion cohorts, they are very different for the post-expansion cohorts. These results suggest 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 a 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 have declined for most of the HE types, except for the LEM and Other Subjects graduates from Key Universities. In the case of Other Subjects graduates from the most prestigious universities, the returns to the degree (relative to Senior High School Diploma) almost doubled from 0.54 to 1.07 log points. This result is a staggering increase of 0.53 log points. For graduates studying LEM subjects at Key Universities, the increase in returns is approximately 0.18 log points.

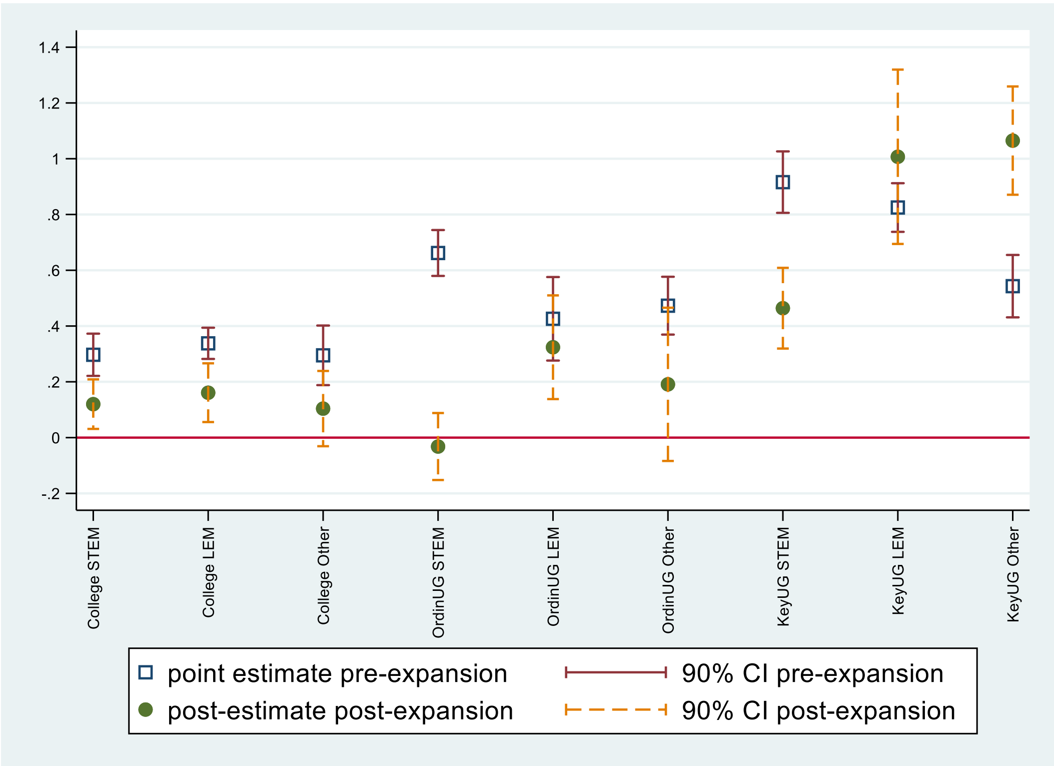

Figure 2: Changes in IPWRA estimates, by HE type and time period

Figure 2 visualizes the changes in returns by HE types resulting from the HE expansion. The hollow squares represent the pre-expansion point estimates, and the solid circles indicate the post-expansion point estimates. The solid and dashed spikes with caps represent the corresponding 90% confidence intervals. Although 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. Notably, seven out of the nine HE types experience 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. Moreover, the decline in returns to LEM at the college level is also significant at the 10% level. Only LEM and Other (non-STEM) Subjects at Key Universities have increases in returns post-expansion, although only Other (non-STEM) Subjects at Key Universities is statistically significant at the 10% level.

One possible explanation of our finding is that although 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 more socioeconomically advantaged backgrounds might have benefited disproportionately from the expansion.

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

7. Concluding Remarks

According to our review of the literature, no study 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 CFPS, we identify the subjects studied and tiers of HE prestige. We take advantage of the rich information on the respondent's school cohort, *hukou* status at age 12, and the mother's age and education to estimate the ***simultaneous*** choice of subject and tier of prestige of HEIs by Senior High School graduates. These factors are significant determinants of HE types defined by the 3 by 3 combinations of subjects and tiers.

By allowing for all possible combinations of university tiers and subjects in the students' HE choice set (despite our modest sample size limiting the number of groups we could 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 for university courses, as defined by subject at specific institutions.

Using the doubly robust IPWRA method to account for selection on observables in subjects and tiers, we observe strong evidence that POLS and RE substantially underestimate the treatment effect of attending universities that are more prestigious for graduates of both genders in LEM. These findings are robust to treating subjects as predetermined for the selection into HEIs by tiers of prestige. To a large extent, this result reflects the reality that subject choice at the time of HE applications is heavily restricted by the academic streaming in Senior High Schools, which occurs 2 years before HE enrollment. 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 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 limitations worth highlighting. First, the sample size is small, especially for the post-expansion analysis. Secondly, the absence 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 to estimate the causal effect of the returns to HE tiers and subjects, which have important policy implications and are of wide public interest in China. Further causal studies are required before an enhanced understanding of this important topic is possible.

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, Sibieta, Luke, Vignoles, Anna, Walker, Ian, Zhu, Yu, 2018a. The relative labor market returns to different degrees. *Department for Education Research Report* No 787.

Belfield, Chris, Britton, Jack, Buscha, Franz, Dearden, Lorraine, Dickson, Matt, van der Erve, Laura, Sibieta, Luke, Vignoles, Anna, Walker, Ian, Zhu, Yu, 2018b. The impact of undergraduate degrees on early-career earnings. *Department for Education Research Report* No. 808.

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.

Chevalier, Arnaud, 2011. Subject choice and earnings of UK graduates. Economics of Education Review 30, 1187-1201.

Dai, Fengyan, Cai, Fang, Zhu, Yu, 2018. Returns to Higher Education in China - Evidence from the Great Higher Education Expansion using Fuzzy Regression Discontinuity. *IZA Discussion Paper* No. 11735.

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.

Feng, Le-an, 2012. Who has more accessibility to Higher Education? Education Research Monthly, 2012 (8), 12-15 (*in Chinese*).

Hamermesh, Daniel S., Donald, Stphen G., 2008. The effect of college curriculum on earnings: An affinity identifier for non-ignorable non-response bias. Journal of Economics 144, 479-491.

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, Xinzheng, 2017. Skill complementarities and returns to higher education: Evidence from college enrollment expansion in China. China Economic Review 46, 10-26.

Liao, Pei-Ju, Wang, Ping, Wang, Yin-Chi, Yip, Chong Lee, 2017. Educational Choice, Rural-urban Migration and Economic Development: The Role of *Zhaosheng* in China. 2017 Meeting Papers 738, Society for Economic Dynamics.

Ministry of Education of the PRC, 2018. Educational Statistics in 2016. http://en.moe.gov.cn/Resources/Statistics/edu_stat_2016/2016_en01/. Accessed 2 April, 2018.

O'Leary, Nigel C. and Sloane, Peter, J., 2005. The return to a university education in Great Britain. National Institute Economic Review No. 193, 75-89.

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.

Wang, Yan, Yao, Yudong, 2003. Sources of China's economic growth 1952–1999: Incorporating human capital accumulation. China Economic Review, 14, 32−53.

Walker, Ian, Zhu, Yu, 2018. University selectivity and the relative returns to higher Education: Evidence from the UK. Labour Economics 53, 230-249.

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, 2nd Ed.

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

Wu, Binzhen, Zhong, Xiaohan, 2014. Matching mechanisms and matching quality: Evidence from a top university in China. Games and Economic Behavior 84, 196-215.

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.

Zhu, Xiaodong, 2011. Understanding China’s growth: Past, present and future. Journal of Economic Perspectives 26(4), 103-124.

Appendix

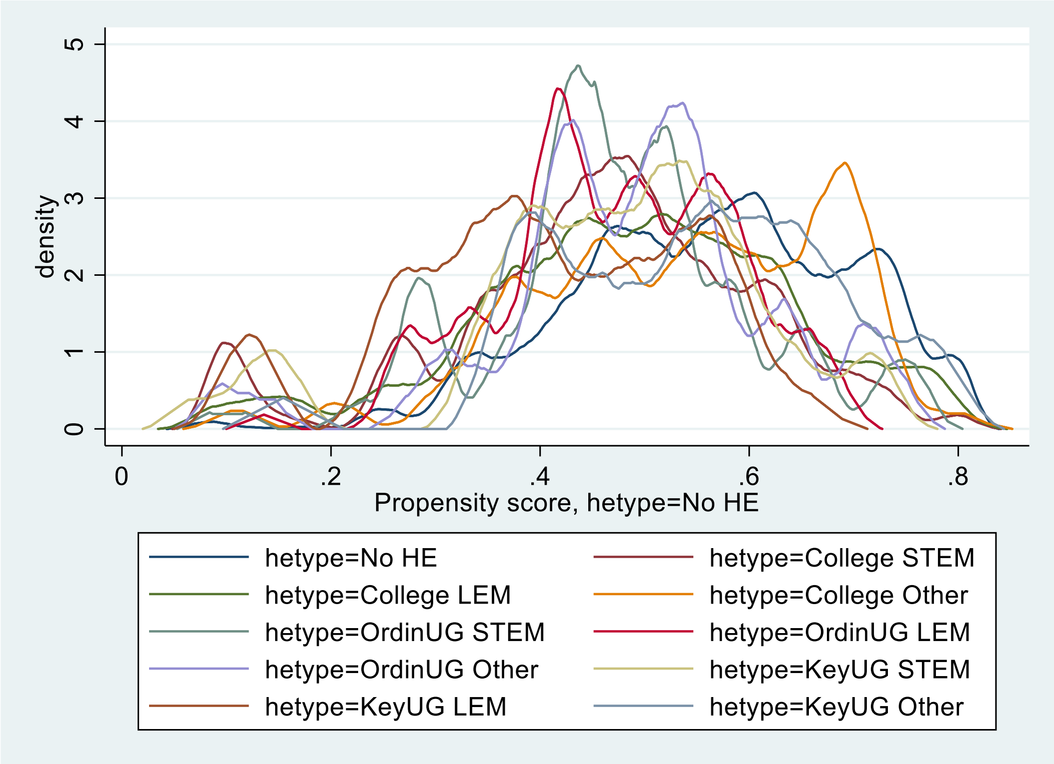

Figure A1: Overlap Plots, Men

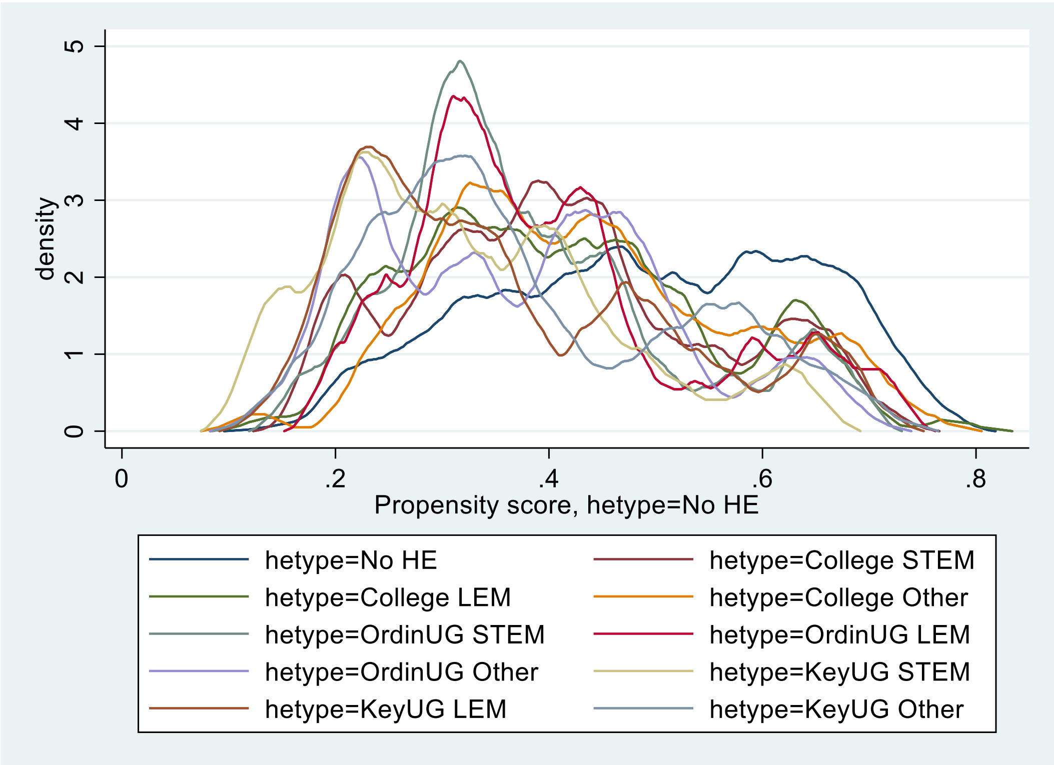

Figure A2: Overlap Plots, Women

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	0.033**	0.086	0.002	-0.020	0.062	-0.010	0.118**	0.246***	0.045	-0.118*
	(0.014)	(0.057)	(0.037)	(0.059)	(0.081)	(0.076)	(0.055)	(0.055)	(0.076)	(0.071)
Age sq	-0.000**	-0.001	0.000	0.000	-0.000	0.000	-0.001**	-0.003***	0.000	0.001*
	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Non-agr	0.032	0.122	0.095	0.148	0.213	0.114	-0.078	0.455***	0.125	0.340
hukou at 12	(0.043)	(0.113)	(0.068)	(0.136)	(0.273)	(0.174)	(0.122)	(0.148)	(0.155)	(0.207)
Mother	0.079*	-0.004	-0.044	-0.032	-0.153	0.069	-0.200	0.156	1.047***	0.096
Primary Edu	(0.042)	(0.120)	(0.094)	(0.104)	(0.216)	(0.176)	(0.156)	(0.138)	(0.189)	(0.186)
Mother Junior	0.154***	0.144	-0.009	-0.229*	-0.541	-0.145	0.060	0.001	1.281***	-0.026
High Edu	(0.051)	(0.122)	(0.110)	(0.137)	(0.583)	(0.378)	(0.153)	(0.163)	(0.249)	(0.294)
Mother Senior	-0.034	0.116	-0.163	0.154	-0.217	-0.175	0.222	-0.248	1.301***	1.168***
High Edu	(0.083)	(0.177)	(0.107)	(0.270)	(0.395)	(0.211)	(0.317)	(0.186)	(0.193)	(0.313)
Mother	0.075	-0.015	0.238**	1.410	0.190	0.263*	0.783	-0.215	1.331***	0.351*
College/Uni	(0.168)	(0.204)	(0.103)	(0.968)	(0.914)	(0.141)	(0.769)	(0.237)	(0.398)	(0.197)
Urban	0.071	0.166	0.350***	0.212**	0.213	-0.008	0.045	-0.359**	-2.211***	0.279
	(0.046)	(0.127)	(0.118)	(0.104)	(0.327)	(0.246)	(0.112)	(0.162)	(0.459)	(0.270)
East	0.305***	0.519***	0.480***	-0.021	0.632***	0.499***	0.499***	0.402***	0.564***	0.077
	(0.036)	(0.086)	(0.071)	(0.112)	(0.189)	(0.145)	(0.160)	(0.145)	(0.196)	(0.194)
Constant	6.488***	5.465***	7.047***	7.621***	5.709***	7.509***	5.093***	2.928***	7.084***	9.650***
	(0.275)	(1.002)	(0.781)	(1.172)	(1.377)	(1.439)	(1.177)	(1.035)	(1.390)	(1.483)
Treatment eq.										
School cohort		0.026*	0.021*	-0.041**	0.027	0.018	0.013	0.072***	0.009	0.027
		(0.014)	(0.012)	(0.019)	(0.018)	(0.021)	(0.020)	(0.019)	(0.033)	(0.034)
Mother's birth		0.018	-0.016*	0.026	-0.002	-0.006	0.007	-0.012	0.027	-0.033
year		(0.012)	(0.010)	(0.016)	(0.017)	(0.017)	(0.021)	(0.015)	(0.032)	(0.022)
Non-agricul		-0.064	0.111	-0.575***	-0.660***	0.041	-0.733***	0.096	-0.292	0.195
hukou at 12		(0.136)	(0.131)	(0.214)	(0.197)	(0.200)	(0.233)	(0.259)	(0.285)	(0.345)
Mother		0.455***	0.403***	-0.190	1.041***	0.285	0.779***	0.672***	1.056***	0.853**
Primary Edu		(0.163)	(0.144)	(0.259)	(0.242)	(0.237)	(0.286)	(0.254)	(0.401)	(0.379)
Mother Junior		0.515***	0.311*	0.654***	0.924***	0.497*	0.208	-0.616	0.552	0.157
High Edu		(0.180)	(0.186)	(0.248)	(0.277)	(0.292)	(0.378)	(0.411)	(0.476)	(0.495)

Mother Senior	0.984***	0.940***	0.860**	1.814***	1.015***	0.811*	0.333	1.949***	0.247
High Edu	(0.214)	(0.203)	(0.337)	(0.309)	(0.292)	(0.434)	(0.402)	(0.417)	(0.644)
Mother	2.775***	2.255***	2.113***	1.785**	0.268	2.616***	2.800***	3.382***	1.641
College/Uni	(0.407)	(0.410)	(0.581)	(0.796)	(1.079)	(0.641)	(0.538)	(0.647)	(1.121)
Urban	0.527***	0.823***	0.145	0.778***	2.661***	0.650**	0.746**	2.144***	0.021
	(0.170)	(0.181)	(0.210)	(0.259)	(0.585)	(0.303)	(0.316)	(0.731)	(0.415)
East	-0.570***	-0.259**	-0.292*	-0.203	-0.544***	-0.996***	0.390*	-0.658**	0.113
	(0.126)	(0.113)	(0.169)	(0.176)	(0.185)	(0.240)	(0.217)	(0.275)	(0.331)
Constant	-87.795***	-11.329	28.069	-53.152***	-30.592	-42.240**	-122.870***	-76.107***	5.965
	(14.475)	(12.363)	(18.735)	(17.373)	(20.250)	(21.362)	(22.453)	(26.431)	(33.143)

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 years.

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	0.021	0.028	0.009	0.060	0.057	-0.077	0.147**	0.115	-0.078	-0.165**
	(0.018)	(0.047)	(0.029)	(0.053)	(0.084)	(0.068)	(0.071)	(0.155)	(0.072)	(0.065)
Age sq	-0.000	0.000	0.000	-0.000	-0.000	0.001	-0.002*	-0.001	0.002*	0.002***
	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
Non-agr	-0.013	-0.177*	-0.026	0.099	0.279**	0.023	-0.031	0.546**	0.495***	-0.177*
hukou at 12	(0.050)	(0.101)	(0.074)	(0.115)	(0.114)	(0.269)	(0.131)	(0.245)	(0.107)	(0.106)
Mother	0.077	0.317***	0.045	-0.245*	0.039	-0.547	-0.465	-0.816***	0.006	0.366***
Primary Edu	(0.054)	(0.118)	(0.090)	(0.148)	(0.177)	(0.587)	(0.348)	(0.175)	(0.222)	(0.121)
Mother Junior	0.104*	0.364**	-0.052	0.311**	0.115	-0.293	0.130	-0.007	0.030	0.393***
High Edu	(0.061)	(0.143)	(0.105)	(0.151)	(0.173)	(0.400)	(0.178)	(0.125)	(0.245)	(0.145)
Mother Senior	0.123	0.440***	0.011	0.104	0.049	0.087	-0.073	-0.786**	-0.154	0.316
High Edu	(0.081)	(0.162)	(0.110)	(0.225)	(0.204)	(0.333)	(0.205)	(0.394)	(0.281)	(0.215)
Mother	0.155	0.578*	0.077	-0.002	0.207	0.395	-0.055	-0.591	-0.390*	0.623***
College/Uni	(0.149)	(0.297)	(0.196)	(0.296)	(0.210)	(0.446)	(0.234)	(0.543)	(0.226)	(0.208)
Urban	0.403***	0.088	0.230**	0.028	-0.448***	0.166	0.692	0.311*	0.153	0.552***
	(0.082)	(0.152)	(0.109)	(0.143)	(0.169)	(0.346)	(0.649)	(0.180)	(0.215)	(0.122)
East	0.332***	0.590***	0.659***	0.329***	0.621***	0.595***	0.267*	0.776***	0.157	0.314***
	(0.043)	(0.096)	(0.064)	(0.115)	(0.120)	(0.135)	(0.142)	(0.108)	(0.108)	(0.096)
Constant	6.203***	5.750***	6.429***	5.727***	5.865***	9.203***	4.282***	4.509	7.831***	9.653***
	(0.320)	(0.829)	(0.515)	(0.917)	(1.442)	(1.680)	(1.498)	(2.759)	(1.341)	(1.340)
Treatment eq.										
School cohort		-0.003	0.019	0.020	0.106***	0.033	0.074***	0.028	0.097***	0.008
		(0.018)	(0.016)	(0.021)	(0.024)	(0.025)	(0.023)	(0.039)	(0.033)	(0.031)
Mother's birth		0.033**	-0.006	-0.003	-0.072***	-0.001	-0.019	0.028	-0.031	0.008
year		(0.015)	(0.012)	(0.017)	(0.018)	(0.021)	(0.020)	(0.029)	(0.031)	(0.024)
Non-agricul		-0.139	0.158	-0.315*	-0.379	0.606**	-0.317	-0.425	0.994***	-0.765***
hukou at 12		(0.172)	(0.135)	(0.187)	(0.263)	(0.252)	(0.218)	(0.320)	(0.356)	(0.258)
Mother		0.349	0.615***	0.746***	0.588*	0.081	0.630**	0.856	0.117	0.395
Primary Edu		(0.216)	(0.175)	(0.232)	(0.353)	(0.305)	(0.304)	(0.694)	(0.448)	(0.403)
Mother Junior		0.835***	0.922***	1.091***	1.622***	0.796***	0.569	1.802***	0.315	1.547***
High Edu		(0.239)	(0.193)	(0.252)	(0.391)	(0.290)	(0.355)	(0.681)	(0.435)	(0.365)

Mother Senior	1.144***	1.364***	0.786**	1.828***	0.978***	1.634***	2.568***	1.394***	2.148***
High Edu	(0.270)	(0.224)	(0.354)	(0.469)	(0.349)	(0.354)	(0.649)	(0.424)	(0.374)
Mother	1.089**	1.536***	0.679	2.465***	-0.258	2.687***	4.050***	2.320***	2.355***
College/Uni	(0.540)	(0.414)	(0.765)	(0.661)	(1.058)	(0.488)	(0.759)	(0.592)	(0.719)
Urban	0.328	0.366*	-0.092	0.413	0.793	0.910**	1.398*	-0.400	0.823*
	(0.246)	(0.207)	(0.234)	(0.347)	(0.488)	(0.359)	(0.741)	(0.455)	(0.434)
East	-0.471***	0.243**	-0.384**	-0.213	-0.351*	-0.957***	0.395	0.996***	-0.001
	(0.159)	(0.120)	(0.179)	(0.229)	(0.213)	(0.215)	(0.323)	(0.284)	(0.238)
Constant	-61.119***	-28.941*	-34.366	-73.006**	-66.468***	-111.785***	-115.734***	-135.868***	-36.276
	(20.464)	(15.577)	(23.651)	(29.561)	(21.969)	(22.792)	(38.185)	(28.627)	(26.200)

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 years.