

PUEY UNGPHAKORN INSTITUTE FOR ECONOMIC RESEARCH

Heterogeneous Returns to Education across Hukou-Migration Subgroups in China

by

Juan Huang and Weerachart T. Kilenthong

November 2023 Discussion Paper No. 210

The opinions expressed in this discussion paper are those of the author(s) and should not be attributed to the Puey Ungphakorn Institute for Economic Research.

Heterogeneous Returns to Education across Hukou-Migration Subgroups in China

Juan Huang

Hezhou University

and

Weerachart T. Kilenthong

University of the Thai Chamber of Commerce

Abstract

This paper uses the China Household Income Project 2018 dataset to estimate returns to education for various Hukou-migration subgroups. We overcome the endogeneity problem of years of schooling using an instrument based on the Great Expansion of Higher Education policy. Our results indicate that the highest returns are for urban native workers (27.4%), followed by urban Hukou-converted (25.0%) and rural native workers (14.7%). In contrast, the returns to education for rural-urban migrant workers are insignificant. Further analyses suggest that Hukou conversion significantly increased the returns to education for rural-origin people by enabling them access to better job opportunities.

Keywords: returns to education; Hukou system; migration; China

JEL Classification: I24 I26 J15 J61

1. Introduction

Education plays a crucial role in shaping an individual's socioeconomic outcomes, with higher educational attainment often associated with better job prospects and higher earnings. Concerning China, several researchers have analyzed how education influences income. Their findings consistently show a significant connection between the level of education and earnings through returns to education (e.g., Appleton et al., 2005; Dai et al., 2022; Huang et al., 2022; Li, 2003; Yang, 2005).

The Hukou registration system in China has a significant implication for labor mobility and, therefore, should strongly influence returns to education. The Hukou system is a dual residence registration system that categorizes the total population according to the Hukou type (urban Hukou or rural Hukou) and the Hukou registration place (local or non-local Hukou based on the administrative unit).¹ Based on different Hukou statuses, people have unequal access to social benefits and entitlements to state resources, such as education, health care, job opportunities, and other welfare entitlements (Akgüc et al., 2014; Dong & Hao, 2011). Regardless of an individual location, their Hukou status remains unchanged unless they undergo a formal Hukou conversion. This paper uses the original Hukou type (rural or urban Hukou) and current location (rural or urban area) to identify an individual as a migrant. In particular, a migrant is an individual whose original Hukou was rural but located in an urban area during the survey. Combining the current Hukou and migrant status leads to four Hukou-migration subgroups, namely urban native workers (UNW), urban Hukouconverted workers (UHCW), rural-urban migrant workers (RUMW), and rural native workers (RNW).

Understanding the returns to education across Hukou-migration subgroups, especially rural-urban migrant workers and urban Hukou-converted workers, will guide the development of Hukou system reform and urbanization policies. These issues are central to the processes and outcomes of Hukou system reform and urbanization. They are interconnected with the questions of achieving an equitable Hukou system and enabling various Hukou subgroups to share the development brought about by urbanization equally.

This paper estimates returns to education in China using the Great Expansion of Higher Education (GEHE) policy as the source of the instrument to correct endogeneity biases, following Huang et al. (2022). However, with a different instrument, we can estimate returns to education across Hukou-migration subgroups, while Huang et al. (2022) could not. More specifically, our instrument is the ratio between the university enrollment and population in each province each year matched with appropriate years

¹ For more details, see Section 2.1.

for each cohort when the individual was 18, which was exogenously affected by the GEHE policy. It is worthy of emphasis that, using the CHIP2018 dataset, our instrument and theirs led to comparable estimates of overall returns to education in China (about 13%). In addition, we found that the returns to education for the UNW are the highest at 27.4%, followed by the UHCW at 25.0% and the RNW at 14.7%. On the other hand, the returns to education for the RUMW subgroup are the lowest and not statistically significant.

These findings are consistent with the results of Gao and Li (2022), who found that urban workers have larger returns to education than rural workers, regardless of Hukou, when using spouse education as the instrument. Similarly, using the CHIP2007 data, Wang et al. (2015) found that urban workers (UNW and UHCW) have higher returns than rural migrants (RUMW) when using father education and occupation as the instruments. However, these studies primarily analyzed the differences in the returns to education between urban and rural areas and treated urban native and urban Hukouconverted workers as a whole. They did not distinguish them by Hukou and work location as ours. There is a lack of literature investigating the returns to education for various Hukou-migration subgroups. This is particularly true for urban Hukouconverted workers, a new specific group of migrants during the urbanization process and Hukou system reform in China. Understanding the returns to education of this subgroup should shed light on the impact of Hukou conversion on labor market outcomes.

The paper is organized as follows. Section 2 introduces the contexts of the China Hukou system and the Great Expansion of Higher Education (GEHE). Section 3 describes the data and explains key variables. The empirical specification and strategy are presented in section 4. Section 5 reports the empirical results, and section 6 presents the conclusion and discussion.

2. Contexts: Hukou system and the Great Expansion of Higher Education Policy

2.1. Hukou-Migration Subgroups

China established a Hukou registration system that classifies Chinese citizens by birth since 1958. Each family usually has a registration certificate called a Hukou book, which records the Hukou information of all family members, including the Hukou registration place and Hukou type (urban or rural)². The Hukou registration place records the address to which the individual is affiliated. This information then determines whether the individual is in rural or urban Hukou. Under this system, individuals must have the same type of Hukou and Hukou registration place as their parents. The Hukou book, which records the attributes of a household, has been dubbed

"China's No. 1 Document" since it has the omnipotent power to determine many important aspects of life (Tian, 2003), which defines a person's relationship to the state and eligibility for a range of benefits.

Regardless of an individual location, their Hukou status remains unchanged unless they undergo a formal Hukou conversion. Converting rural Hukou to urban Hukou is the most popular Hukou conversion and the most critical conversion and challenging to obtain.³ This Hukou conversion process is known as Nongzhuanfei, which has undergone a long transformation from strict restrictions during Mao Zedong's era to gradual liberalization in the early 1980s and further liberalization in the late 20th century. Getting a higher education, purchasing a house, marriage, and employment are the most common channels for converting Hukou (Song, 2014). Even although the relaxation of the Hukou system has allowed nearly 300 million rural Hukou people to

² The initial Hukou types marked the division of occupation in the Chinese economy (i.e., being engaged in agricultural or non-agricultural production). However, as the system evolved, the distinction between Hukou types was not necessarily linked to the holder's occupation. Instead, as the main institutional pillar behind the deep rural-urban divide in China. Nowadays, people prefer to use Hukou to distinguish between people from rural or urban areas. Therefore, the agricultural and non-agricultural Hukou are usually called rural and urban Hukou, respectively, which is the convention employed in this paper.

³ Therefore, Hukou conversion in this paper refers to the conversion of rural Hukou to urban Hukou.

migrate to urban areas (NBSC, 2022),⁴ under the current rule of Hukou conversion, only a minority of rural-urban migrants can qualify for the Hukou conversion (Chan, 2009; Khanna et al., 2021).

This paper uses the original Hukou type (rural or urban Hukou) and current location (rural or urban area) to identify an individual as a migrant. In particular, a migrant is an individual whose original Hukou was rural but located in an urban area during the survey. Combining the current Hukou and migrant status leads to four Hukou-migration subgroups, namely urban native workers (UNW), urban Hukou-converted workers (UHCW), rural-urban migrant workers (RUMW), and rural native workers (RNW) as shown in Table 2.1.

Table 2.1: Major Hukou subgroups by current Hukou type and migrant status

	Migrate to urban areas	No migration
Rural Hukou	Rural-Urban Migrant Workers (RUMW)	Rural Native Workers (RNW)
Urban Hukou	Urban Hukou-Converted Workers (UHCW)	Urban Native Workers (UNW)

Hukou-migration status matters for returns to education because local Hukou registration defines a person's right to engage in many activities and eligibility for services a specific local government provides. For example, if someone has a Beijing Hukou, they generally cannot access a public pre-university school in Shanghai since they are a non-local Shanghai Hukou. Hukou type distinguishes between urban and rural areas. People with urban Hukou were entitled to housing, employment, food rations, education, health care, and other benefits provided by the state (Cheng & Selden, 1994), while people with rural Hukou were required to be self-reliant and contribute to the country without state support (Fu & Ren, 2010; Naughton, 2007). For example, suppose someone has a rural Beijing Hukou; in that case, they generally cannot access a public pre-university school in urban Beijing since they are a rural Beijing Hukou.

In addition, educational resources in China are highly unequal, with a strong preference for urban Hukou residents (Ding & Lehrer, 2012; Golley & Kong, 2013;

⁴ Data source: National Bureau of Statistics of China (NBSC).

Wang & Gao, 2013). Especially in compulsory education,⁵ teachers in rural schools have lower educational attainment, are less professionally accomplished, and have access to worse facilities than their counterparts in urban schools (see Figure 2.1). Chinese citizens can only access public schools in their Hukou registration regions, which means that rural Hukou holders generally can only gain access to rural schools. Hence, the children of rural-urban migrants primarily receive their education through rural schools, urban informal education institutions, or a mixture of the two. Unfortunately, these institutions usually have poor facilities and low-quality teachers (Goodburn, 2016).



Figure 2.1. Educational resources between urban and rural areas in 2013. Data source: Educational Statistics Yearbook of China. Panel A: Classrooms with computer and internet connection per 10 students. Panel B: Education facilities per student (100 RMB). Panel C: Fraction of full-time teachers with a bachelor's degree or higher. Panel D: Fraction of full-time teachers with a senior title of professional post.

2.2. The Great Expansion of Higher Education Policy

Our identification strategy is based on the Great Expansion of Higher Education (GEHE) policy. In China, higher education is not directly affected by the Hukou system. Students who fulfill the requirements are eligible to choose any university, regardless of their Hukou status. However, the limited supply of university enrollment was an issue in China until 1999, when the government implemented the GEHE policy. Under the policy, the Ministry of Education announced a considerable increase in accessible university seats in response to the economic downturn and growing number of

⁵ China implements a nine-year compulsory education policy.

unemployed young people caused by the 1997 Asian Financial Crisis (Dai et al., 2022). As Dai et al. (2022) reported, this unexpected growth amounted to a 47% and 38% increase in university seats in 1999 and 2000, respectively. The rapid expansion of enrollment has continued to this day, as shown in Figure 2.2. The annual enrollment has increased from 1.08 million in 1998 to nearly 10.15 million in 2022.



Figure 2.2. University student enrollment in China. Data source: NBSC

Previous studies generally agreed that the GEHE policy has improved China's overall education level by roughly one year around the cut-off point (e.g., Dai et al., 2022; Huang et al., 2022). However, Huang et al. (2022) indicated that the GEHE policy had a significant impact only on urban Hukou residents. Rural Hukou residents perhaps could not benefit from this policy due to the poor quality of compulsory education and the high cost of high school education.

The supply of university seats has risen steeply since 1999 and varied across provinces, as shown in Figure 2.3 below. We argue that this exogenous variation in university enrollment can help identify the returns on education. The raw data on university enrollment is only available from 1987 to 2017, and the respondent's birth year in our data started from 1961 to 1995. This limitation causes a minor problem for the oldest respondents, born between 1961 and 1968 and turned 18 years old between 1979 and 1986. Unfortunately, there is no university enrollment data during those years. Hence, we assume that the enrollment during those years was the same as the 1987 level. This assumption should be harmless because the university enrollment during those years was almost constant (see Figure 2.3).



Figure 2.3. University student enrollment in each province each year. Data source: China Statistical Yearbook

3. Data and Key Variables

The primary data come from the China Household Income Project (CHIP) 2018, which drew from the large sample pool of the Integrated Urban and Rural Regular Household Survey of the National Bureau of Statistics of China in 2019. The latter covers 160,000 households in all 31 provinces in mainland China. Due to the extreme regional imbalance in China's economic development, with the eastern region being the most developed and the central and western regions lagging, the CHIP project has stratified the CHIP sample according to the three regions based on a systematic sampling method. The final CHIP sample covers 15 provinces. Figure 3.1 presents the location of the 15 provinces (shaded areas). Even though the survey covered only 15 out of 31 provinces, those provinces accounted for 64.5% of the country's population and 66.1% of the nation's GDP in 2018 (NBSC, 2018). That is, the dataset should represent Chinese labor markets quite well.

The CHIP2018 dataset comprises two separate samples: the urban and rural household samples. By identifying individual original and current Hukou status and their current location (urban or rural), our final sample consists of four Hukou-migration subgroups, namely urban native workers (UNW), urban Hukou-converted workers (UHCW), rural-urban migrant workers (RUMW), and rural native workers (RNW).





This paper considers only workers who reported their employment status as employed. Moreover, we exclude individuals born before 1961, potentially affected by the Cultural Revolution's political events in China (only 0.77% of the total sample excluded). As a result, our main sample consisted of individuals aged between 23 and 57 who reported being in good health during the survey. Table A.1 in Appendix A provides summary statistics on the demographic and socioeconomic characteristics of people with different Hukou statuses.

This paper uses hourly earnings, which represent a unit price of labor, as the primary dependent variable unless stated otherwise. This choice responds to a concern that the returns to education using earnings without adjusting for working hours may be downward biased because workers with higher education levels may work fewer hours on average (Li, 2003).⁶ This dataset contains annual earnings, working months, working days per month, and working hours per day information for each worker. This information first allows us to calculate monthly earnings using annual earnings and working months. We then divided monthly earnings by the product of working days per month and working hours per day to get hourly earnings. This dataset also contains the maximum number of years an individual attended formal education, which is vital to estimating returns to education.

⁶ Our sample also shows a negative correlation between years of schooling and working hours, with a correlation coefficient of -0.27.

4. Empirical specification and strategy

This paper estimates returns to education using the following linear model:

$$W_i = \alpha + \beta_1 S_i + \gamma X_i + \varepsilon_i \tag{1}$$

where W_i is the earnings of individual *i*, S_i is the schooling years, X_i represents other individual characteristics that potentially influence personal earnings, including age, age squared, gender, marital status, public sector dummy, professional job dummy, industry fixed effects, regional fixed effects, party membership, ethnic minority dummy, parents' education, parents' professional job dummy, parents' public sector dummy, parents' party membership, and ε_i is a random error term, unless stated otherwise. Estimating this model using the ordinary least square (OLS) faces an endogeneity bias due to either the unobservability of innate ability or credit constraints.

This paper overcomes the endogeneity problem using an instrumental variable approach, the two-stage least square (TSLS), with the ratio between the university enrollment and population in each province each year matched with appropriate years for each cohort when the individual was 18 as the instrument. This instrument is defined based on the GEHE policy following the concept proposed by Huang et al. (2022). Similarly, we assume that the policy was an exogenous shock relevant to individuals 18 years of age after 1998 since Chinese students usually make college decisions when they are about 18 (Dai et al., 2022). Ours differs from the ones in Huang et al. (2022) in that our instrument is based on the panel data of university enrollments and population across all provinces.⁷ More specifically, our instrument for an individual is the ratio between the university enrollment and population for each province when the sample was 18 years old. Ideally, for each individual, the ratio should have been matched with their provincial residence when they were 18. Unfortunately, the CHIP dataset contains only the provincial location at 14 years of age. Therefore, we must

⁷ We recalculated the enrollment data for Sichuan Province and Chongqing Municipality from 1979 to 1996 by assuming that the enrollment ratio between the two provinces was constant as the 1997 ratio during that period since Chongqing Municipality was separated from Sichuan Province in 1997. The same method is also applied to Guangdong Province's and Hainan Province's data before 1987, as Hainan Province became an independent provincial administrative unit separated from Guangdong Province in 1988.

assume that each individual did not migrate across provinces between 14 and 18 years old. An advantage of our instrument over the ones in Huang et al. (2022) is that ours has sufficient variations for sub-Hukou analysis, while a weakness is that it relies on the assumption that university enrollment has been driven mainly by the GEHE policy conditional on the provincial population.

Note that part of our control variables, education, jobs, and political capital of one of the parents, is to deal with the endogeneity problem concerning migration by controlling. In particular, we use a series of dummy variables indicating that either one of the parents completed high school or above, has worked in professional occupations, has worked in the public sector, and has been a member of the Communist Party of China. Previous research has shown that these variables strongly influence individual migration decisions (e.g., Du, 2018; Zhao, 2023). However, there are still many more unobserved factors affecting the migration decision. Therefore, this attempt may partially mitigate the endogeneity problem at best but cannot solve the problem completely.

5. Empirical results

We first present the returns to education for China using our and Huang et al. (2022) instruments in Table 5.1. To better compare with Huang et al. (2022), all the estimations in this Table used the same set of control variables as theirs, which is a subset of our controls. The estimated returns to education using both instruments are comparable; the overall returns were about 13.0% and 13.1% for our and Huang et al. (2022) instruments, respectively. The return similarity indicates that our instrument can perform equally well as Huang et al. (2022). In addition, Kleibergen-Paap F-statistics (Kleibergen & Paap, 2006) from the first-stage estimation for all cases are well above 10, the popular threshold proposed by Stock and Yogo (2002). This finding implies that our instrument is relevant and sufficiently strong to avoid weak instrument bias.

Table 5.1: Returns to education for China using both our instrument and the Huang et al. (2022)

	Overall		Male		Female	
	Our IV	Huang's	Our IV	Huang's	Our IV	Huang's
Second-stage: Schooling years	0.130***	0.131***	0.107***	0.144**	0.152***	0.133***

	(0.016)	(0.034)	(0.024)	(0.060)	(0.022)	(0.035)
First-stage:						
Uni stud enroll/pop	3.243***		2.938***		3.576***	
	(0.211)		(0.284)		(0.313)	
P*U12		-0.938		0.411		-2.613**
		(0.787)		(1.065)		(1.171)
BT*U12		-0.088**		-0.015		-0.178***
		(0.034)		(0.047)		(0.051)
P*BT*U12		-0.030		-0.069*		0.013
		(0.031)		(0.041)		(0.046)
Controls	yes	yes	yes	yes	yes	yes
Diagnostic Tests:						
F-Statistics	236.43	22.79	106.78	7.66	130.97	21.50
Endo. (P-value)	0.000	0.028	0.017	0.094	0.000	0.047
Overident. (P-value)		0.206		0.607		0.302
N	15501	15501	8879	8879	6622	6622

Note: Robust standard errors are in parentheses. *, **, *** denote significance at 10%, 5%, and 1%, respectively. To better compare with Huang et al. (2022), all estimations in this Table used monthly earnings as the dependent variable and the same control variables as theirs, including a post-1980 birth year dummy, post-1980 birth year trend dummy, urban Hukou at age 12 dummy, age, gender, and regions. Following Huang et al. (2022), this Table restricted the analysis to individuals born in 1970 or later to exclude confounding effects of the "Compulsory Education Law."

Note that the returns to education using the instruments of Huang et al. (2022) and the CHIP data are moderately lower than the results in the original paper using the China Household Finance Survey data for males (14.4% versus 16.5%) but the opposite is true for females (13.3% versus 12.4%). One potential explanation is that our dataset contains more rural male workers but fewer rural female workers relative to Huang et al. (2022). In particular, the proportions of male workers with rural Hukou (RNW and RUMW) in our sample and Huang et al. (2022) were 62.0% and 55.1%, respectively, while they were 38.0% and 48.4% for females. On the other hand, previous studies have shown that the returns to education in rural China are significantly lower than in urban areas (e.g., Gao & Li, 2022; Wang et al., 2015).

We next present our main results, the returns to education for each Hukou subgroup, in Table 5.2. Given that migrant workers are likelier to hold longer-hours jobs, we use hourly earnings as the primary outcome to better reflect the unit price of labor. The first column in Table 5.2 gives the overall results for China, and the second to fifth columns present the estimated results for urban native workers, urban Hukou-converted workers, rural native workers, and rural-urban migrant workers, respectively.

The TSLS estimation results show that the overall return to education in China was 17.2%⁸, and urban native workers have the highest returns to education with a rate of 27.4%, followed by the urban Hukou-converted workers with a 25.0% return, and the rural native workers have a 14.7% return. Again, based on the F-statistics from the first-stage estimation, the instrument is sufficiently strong for all cases. Therefore, we should not need to worry about weak instrument bias. In addition, the rejection of Durbin-Wu-Hausman tests indicates that the OLS estimates differ from the instrumented ones for all but the RUWM subgroup. See the OLS results in Table A.2 in the appendix. In particular, our results suggest that the OLS estimates for all but the RUWM subgroup are downward biased, implying that highly unobserved ability individuals may not be able to pursue additional education due to financial constraints.

	Overall	UNW	UHCW	RNW	RUMW	
Second-stage:						
Schooling years	0.172***	0.274***	0.250***	0.147***	0.048	
	(0.020)	(0.060)	(0.078)	(0.052)	(0.048)	
First-stage:						
Uni stud enroll/pop	2.747***	1.833***	1.691***	2.274***	2.238***	
	(0.200)	(0.293)	(0.407)	(0.358)	(0.434)	
Controls	yes	yes	yes	yes	yes	
Diagnostic Tests:						
F-Statistics	189.24	39.15	17.26	40.27	26.60	
Endo. (P-value)	0.000	0.000	0.013	0.020	0.923	
N	15832	3628	1940	5698	4566	

Tab	le 5.2	: Returns	to e	ducation	for	each	Hu	kou	sut	ogrou	ıр.
-----	--------	-----------	------	----------	-----	------	----	-----	-----	-------	-----

Note: Robust standard errors are in parentheses. *, **, *** denote significance at 10%, 5%, and 1%, respectively.

We also confirm that the differential returns to education of the original urban (UNW) and the current rural Hukou (RNW and RUMW) holders are statistically significant, as formally confirmed by Fisher's Permutation test results in Table 5.3. This finding is consistent with Gao & Li (2022), who found that urban workers have larger returns to education than rural workers regardless of Hukou, with a maximum gap of around 8% when using spouse education as the instrument. Similarly, using the CHIP2007 data, Wang et al. (2015) found that urban workers (UNW and UHCW) have higher returns than rural migrants (RUMW), with a maximum gap of around 8% when

⁸ The overall returns to education reported here is different from the one in Table 5.1 due to the difference in sample size and control variables.

using father education and occupation as the instruments. However, both papers did not separate workers by Hukou and work location as ours.

On the other hand, the returns to education for UHCW and UNW are indistinguishable, as confirmed by Fisher's Permutation test result in Table 5.3. This result indicates that Hukou conversion increases the returns for workers originally in rural Hukou. This finding is consistent with the general belief in China that migration and education are critical tools for rural Hukou residents to escape the rural poverty trap (Zhang, 2017).

	UNW	UHCW	RNW	RUMW
UNW		0.426	0.018	0.000
UHCW	0.426		0.105	0.001
RNW	0.018	0.105		0.084
RUMW	0.000	0.001	0. 084	

Table 5.3: P-values from Fisher's Permutation test for each pair of Hukou subgroups

Note: Fisher's Permutation test was used to test the difference of coefficients between each pair, and the test was performed by permuting 1000 times.

One puzzling result is that the returns to education for rural Hukou holders who worked in urban areas or rural-urban migrant workers (RUMW) are lowest and not statistically significant. One potential explanation is that the returns to education are heterogeneous across occupational sectors, with the public sector being the highest,⁹ and only 12% of RUMW worked in the public sector (see Table A.1 in the appendix). As shown in Table 5.4 below, the returns to education in the public sector are about 32.1%, more than double that of the non-public sector (at about 12.4%). Another potential explanation is that most of them might be low-skilled workers, so they can not convert their Hukou and are typically sorted into jobs with relatively low education premiums, such as non-professional occupations. As shown in Table 5.4, the returns to education for non-professional occupations are significantly smaller than for

⁹ The public sector includes government and party agencies, public institutions, solely state-owned (state-holding) enterprises and collective enterprises.

professional occupations.¹⁰ While only 70.5% of UHCW worked in non-professional occupations, 81.5% of RUMW worked there.

	Sectors		Occupation						
	Public	Non-public	Profession	Non-profession					
Second-stage:									
Schooling years	0.321***	0.124***	0.330**	0.157^{***}					
	(0.064)	(0.020)	(0.137)	(0.019)					
First-stage:									
Uni_stud_enroll/pop	1.827***	3.332***	1.155***	3.252***					
	(0.342)	(0.242)	(0.414)	(0.228)					
Controls:	yes	yes	yes	yes					
Diagnostic Tests:									
F-Statistics	28.51	189.43	7.80	204.19					
Endo. (P-value)	0.000	0.000	0.013	0.000					
Ν	3607	12225	3300	12532					

Table 5.4: Returns to education for various occupational sectors and occupation

Note: Robust standard errors are in parentheses. *, **, *** denote significance at 10%, 5%, and 1%, respectively.

We also performed two additional estimations for robustness tests: (i) dropping samples who were 18 years old in 1987 or before and (ii) changing the reference age to 19. The first is to respond to a concerning issue about the absence of data regarding university enrollment from 1979 to 1986. The second is to deal with the possibility that some students may enter the university later at 19. The estimation results are generally robust to those changes, as shown in Table A.3 and Table A.4 in the appendix.

6. Conclusion and Discussion

This paper estimated the returns to education across Hukou-migration subgroups in China using the CHIP2018 data. Using the ratio between the university enrollment and population in each province each year matched with appropriate years for each cohort when the individual was 18 as the instrument, we have shown that the returns to education for urban native workers (UNW) are the highest at 27.4%, followed by urban Hukou-converted workers (UHCW) at 25.0% and rural native workers (RNW) at

¹⁰ We identified an individual who answered the occupation question that he/she was a person in charge of the unit (department) or professional and technical personnel, as having a professional occupation, while the other answers were grouped into a non-professional occupation.

14.7%. In contrast, the returns for rural-urban migrant workers (RUMW) are relatively low and insignificant.

One striking feature of the results is an enormous difference in the returns to education between migrants who could convert their Hukou and those who could not. This finding suggests that Hukou's status strongly influences labor market opportunity. One supporting evidence is that urban Hukou-converted workers (UHCW) are more likely than rural-urban migrant workers (RUMW) to work in high-return sectors and occupations, namely the public sector and professional occupations. In other words, Hukou conversion significantly increased the returns to education for rural-origin people by enabling them access to better job opportunities. On the other hand, most urban Hukou-converted workers (UHCW), roughly 86.4% of them, converted their Hukou after 15, which suggests that most of them must have completed the compulsory schooling (if they did) in the rural areas. This finding indicates that the differential of the returns is unlikely to stem directly from unequal education opportunities. However, it is plausible that better education increases the likelihood of Hukou conversion and subsequentially improves income. We still need more evidence before closing this channel, for sure.

Since the Hukou conversion helps improve labor-market opportunities, Hukou reform should be more comprehensive, especially by further relaxing restrictions on Hukou conversion, which can help utilize rural human capital more efficiently. In addition, efforts to mitigate Hukou-based discrimination in the labor market should be implemented, and policies should promote equal access to well-paying occupations, regardless of Hukou status.

This study has several limitations. First, although we addressed the issue of endogeneity in migration by controlling for individual parents' backgrounds, many potential unobserved factors remain unaccounted for. Two factors may contribute to the issue of endogeneity. One factor arises from the unobserved ability since they must undergo a rigorous screening process to qualify for Hukou conversion, which requires good education and excellent working skills. Another factor comes from their destination of migration. Usually, developed cities are the preferred choice of migrants. The more developed the city, the higher the wage level, likely to lead to higher returns to education.

Another limitation of this study is that we only observed the registration province of the respondent's Hukou when the samples were 14 instead of 18, which is the time of university decision. As a result, we have to assume that the samples did not change their Hukou registration province and did not move across provinces between 14 and 18 years old. Fortunately, only 1.8% of urban Hukou-converted workers migrated across provinces between the ages of 14 and 18. Therefore, this assumption should not cause a meaningful problem. Nevertheless, future research with an alternative dataset containing the information at the right age is encouraged.

Declaration of Competing Interest

None.

Acknowledgment

We thank the Survey and Research Centre for the Chinese Household Income Project (CHIP) and the China Institution for Income Distribution of Beijing Normal University for data access. Juan Huang thanks the University of Thai Chamber of Commerce for financially supporting her Ph.D. study. We are grateful to Dusanee Kesavayuth, Sasiwimon Warunsiri Paweenawat, Suparee Boonmanunt, and Sartja Duangchaiyoosook for their helpful comments. All mistakes are our own.

REFERENCES

- Akgüc, M., Liu, X., & Tani, M. (2014). Expropriation with Hukou Change: Evidence from
 a Quasi-Natural Experiment. *China Economic Review*.
 https://doi.org/10.2139/ssrn.2543883
- Appleton, S., Song, L., & Xia, Q. (2005). Has China crossed the river? The evolution of wage structure in urban China during reform and retrenchment. *Journal of Comparative Economics*, 33(4), 644–663. https://doi.org/10.1016/J.JCE.2005.08.005

- Chan, K. W. (2009). The Chinese Hukou system at 50. *Eurasian Geography and Economics*, 50(2), 197–221. https://doi.org/10.2747/1539-7216.50.2.197
- Cheng, T., & Selden, M. (1994). The Origins and Social Consequences of China's Hukou System*. *The China Quarterly*, *139*, 644–668. https://doi.org/10.1017/S0305741000043083
- Dai, F., Cai, F., & Zhu, Y. (2022). Returns to higher education in China evidence from the 1999 higher education expansion using a fuzzy regression discontinuity. *Applied Economics Letters*, 29(6), 489–494. https://doi.org/10.1080/13504851.2020.1871465
- Ding, W., & Lehrer, S. (2012). Incentives and the quality of teachers and schools. In H. Emily & P. Albert (Eds.), *Education and Reform in China* (pp. 191–204). Routledge. https://doi.org/10.4324/9780203960950-13
- Dong, Y., & Hao, L. (2011). The key issues in measuring the return on investment in education in China. *Journal of Southwestern University (Social Science Edition)*, *37*(1), 115–121.
- Du, H. (2018). Rich dad, poor dad: the impact of family background on educated young people's migration from peripheral China. *Journal of Youth Studies*, 21(1), 90–110. https://doi.org/10.1080/13676261.2017.1343939
- Fu, Q., & Ren, Q. (2010). Educational Inequality under China's Rural–Urban Divide: The Hukou System and Return to Education. *Environment and Planning A: Economy and Space*, 42(3), 592–610. https://doi.org/10.1068/a42101
- Gao, X., & Li, M. (2022). Differences between returns to education in Urban and rural China and its evolution from 1989 to 2019. PLOS ONE, 17(10). https://doi.org/10.1371/journal.pone.0274506
- Golley, J., & Kong, S. T. (2013). Inequality in Intergenerational Mobility of Education in China. China & World Economy, 21(2), 15–37. https://doi.org/10.1111/j.1749-124X.2013.12013.x
- Goodburn, C. (2016). Educating Migrant Children: The effects of rural-urban migration on access to primary education. *Spotlight on China*, 365–380. https://doi.org/10.1007/978-94-6209-881-7_21

- Huang, B., Tani, M., Wei, Y., & Zhu, Y. (2022). Returns to education in China: Evidence from the great higher education expansion. *China Economic Review*, 74, 101804. https://doi.org/10.1016/j.chieco.2022.101804
- Khanna, G., Liang, W., Mobarak, A. M., & Song, R. (2021). *The Productivity Consequences* of Pollution-Induced Migration in China. https://doi.org/10.3386/w28401
- Kleibergen, F., & Paap, R. (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics*, *133*(1), 97–126. https://doi.org/10.1016/j.jeconom.2005.02.011
- Li, H. (2003). Economic transition and returns to education in China. *Economics of Education Review*, 22(3), 317–328. https://doi.org/10.1016/S0272-7757(02)00056-0
- Naughton, B. (2007). The Chinese Economy. MIT Press.
- Song, Y. (2014). What should economists know about the current Chinese hukou system? *China Economic Review*, *29*, 200–212. https://doi.org/10.1016/j.chieco.2014.04.012
- Stock, J. H., & Yogo, M. (2002). Testing for Weak Instruments in Linear IV Regression.
- Tian, B. (2003). *Zhongguo diyi zhengjian (China's No.1 Document)*. Guangdong Renmin Chubanshe.
- Wang, D., & Gao, M. (2013). Educational equality or social mobility: The value conflict between preservice teachers and the Free Teacher Education Program in China. *Teaching and Teacher Education*, 32, 66–74. https://doi.org/10.1016/J.TATE.2013.01.008
- Wang, W., Liu, T., & Chen, M. (2015). Urban-rural differences in returns to education and income transfer mechanism. *Journal of Hunan Agricultural University (Social Sciences)*, 16(1), 90–95.
- Yang, D. T. (2005). Determinants of schooling returns during transition: Evidence from Chinese cities. *Journal of Comparative Economics*, 33(2), 244–264. https://doi.org/10.1016/j.jce.2005.03.010

- Zhang, H. (2017). Opportunity or new poverty trap: Rural-urban education disparity and internal migration in China. *China Economic Review*, 44, 112–124. https://doi.org/10.1016/j.chieco.2017.03.011
- Zhao, M. (2023). Returning home for a better job? Return migration, family background, and labor market outcomes among recent college graduates in China. *Chinese Journal* of Sociology, 9(3), 408–428. https://doi.org/10.1177/2057150X231191397

Appendix A. Additional Tables

		Overall	UNW	UHCW	RNW	RUMW
Depend variable	Hourly earnings	26.61	33.41	30.57	21.94	25.37
Education variable	Schooling years	10.54	13.04	12.19	8.612	10.23
Occupation	Profession	0.208	0.314	0.295	0.131	0.185
	Non-profession	0.792	0.686	0.705	0.869	0.815
Sector of job	Public	0.228	0.463	0.426	0.098	0.119
	Non-public	0.772	0.537	0.574	0.902	0.881
Industry	Indus1	0.039	0.015	0.021	0.067	0.032
	Indus2	0.334	0.204	0.202	0.471	0.321
	Indus3	0.627	0.780	0.777	0.463	0.647
Regions	East provinces	0.374	0.384	0.354	0.322	0.440
	Central provinces	0.358	0.366	0.297	0.406	0.319
	West provinces	0.268	0.250	0.349	0.273	0.241
Parents background	High school education	0.169	0.358	0.195	0.069	0.132
	Profession job	0.075	0.163	0.080	0.042	0.044
	Public sector	0.068	0.163	0.074	0.034	0.031
	Party	0.126	0.239	0.171	0.077	0.079
Other controls	Age	40.27	40.62	41.47	41.02	38.55
	Male	0.592	0.555	0.552	0.644	0.575
	No minority	0.940	0.942	0.942	0.927	0.952
	Party	0.133	0.227	0.275	0.078	0.067
	Married	0.878	0.878	0.909	0.862	0.884
Observations		15832	3628	1940	5698	4566

Table A.1: Demographic and socioeconomic characteristics of each Hukou-migration subgroup

Table A.2: OLS estimates returns to education by Hukou

Overall	UNW	UHCW	RNW	RUMW
0.061***	0.067***	0.064***	0.035***	0.040***
(0.002)	(0.005)	(0.006)	(0.004)	(0.004)
yes	yes	yes	yes	yes
15832	3628	1940	4566	5698
	Overall 0.061*** (0.002) yes 15832	Overall UNW 0.061*** 0.067*** (0.002) (0.005) yes yes 15832 3628	OverallUNWUHCW0.061***0.067***0.064***(0.002)(0.005)(0.006)yesyesyes1583236281940	OverallUNWUHCWRNW0.061***0.067***0.064***0.035***(0.002)(0.005)(0.006)(0.004)yesyesyesyes15832362819404566

			IF 0	0	
	Overall	UNW	UHCW	RNW	RUMW
Second-stage:					
Schooling years	0.182***	0.338***	0.234***	0.198***	0.011
	(0.023)	(0.083)	(0.068)	(0.061)	(0.053)
First-stage:					
Uni stud enroll/pop	2.724***	1.545***	2.031***	2.270***	2.247***
``	(0.216)	(0.300)	(0.425)	(0.405)	(0.490)
Controls	yes	yes	yes	yes	yes
Diagnostic Tests:					
F-Statistics	158.66	26.45	22.82	31.39	21.01
Endo. (P-value)	0.000	0.000	0.012	0.002	0.407
N	12570	2914	1551	4207	3898

Table A.3: Returns to education by Hukou with dropping the missing enrollment data

Table A.4: Returns to education by Hukou with the measures for the IV at respondent age 19

	Overall	UNW	UHCW	RNW	RUMW
Second-stage:					
Schooling years	0.172***	0.264***	0.242***	0.150***	0.055
	(0.019)	(0.053)	(0.069)	(0.055)	(0.045)
First-stage:					
Uni_stud_enroll/pop	2.903***	2.069***	1.902***	2.156***	2.334***
	(0.197)	(0.292)	(0.399)	(0.350)	(0.422)
Controls	yes	yes	yes	yes	yes
Diagnostic Tests:					
F-Statistics	217.19	50.35	22.69	37.94	30.57
Endo. (P-value)	0.000	0.000	0.008	0.023	0.958
N	15832	3628	1940	5698	4566

Note: Robust standard errors are in parentheses. *, **, *** denote significance at 10%, 5%, and 1%, respectively.

Appendix B. Estimation Results with More Detailed

Table B.1: OLS estimates returns to education by Hukou

	cation by Huk	ou		
Overall	UNW	UHCW	RNW	RUMW
0.061***	0.067***	0.064***	0.035***	0.040***
(0.002)	(0.005)	(0.006)	(0.004)	(0.004)
0.098***	0.120***	0.161**	0.089***	0.077**
(0.021)	(0.041)	(0.063)	(0.032)	(0.036)
0.245***	0.275***	0.296***	0.099***	0.251***
(0.016)	(0.027)	(0.038)	(0.028)	(0.027)
0.041**	0.031	0.052	-0.198***	-0.138***
(0.016)	(0.026)	(0.036)	(0.033)	(0.033)
0.052***	0.047***	0.031	0.040***	0.093***
(0.006)	(0.013)	(0.019)	(0.010)	(0.011)
-0.001***	-0.001***	-0.000*	-0.001***	-0.001***
(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	Overall 0.061*** (0.002) 0.098*** (0.021) 0.245*** (0.016) 0.041** (0.016) 0.052*** (0.006) -0.001**** (0.000)	OverallUNW 0.061^{***} 0.067^{***} (0.002) (0.005) 0.098^{***} 0.120^{***} (0.021) (0.041) 0.245^{***} 0.275^{***} (0.016) (0.027) 0.041^{**} 0.031 (0.016) (0.026) 0.052^{***} 0.047^{***} (0.006) (0.013) -0.001^{***} -0.001^{***} (0.000) (0.000)	Overall UNW UHCW 0.061*** 0.067*** 0.064*** (0.002) (0.005) (0.006) 0.098*** 0.120*** 0.161** (0.021) (0.041) (0.063) 0.245*** 0.275*** 0.296*** (0.016) (0.027) (0.038) 0.041** 0.031 0.052 (0.016) (0.026) (0.036) 0.052*** 0.047*** 0.031 (0.006) (0.013) (0.019) -0.001*** -0.000* (0.000)	OverallUNWUHCWRNW 0.061^{***} 0.067^{***} 0.064^{***} 0.035^{***} (0.002) (0.005) (0.006) (0.004) 0.098^{***} 0.120^{***} 0.161^{**} 0.089^{***} (0.021) (0.041) (0.063) (0.032) 0.245^{***} 0.275^{***} 0.296^{***} 0.099^{***} (0.016) (0.027) (0.038) (0.028) 0.041^{**} 0.031 0.052 -0.198^{***} (0.016) (0.026) (0.036) (0.033) 0.052^{***} 0.047^{***} 0.031 0.040^{***} (0.006) (0.013) (0.019) (0.010) -0.001^{***} -0.001^{***} -0.000^{***} (0.000) (0.000) (0.000) (0.000)

Male	0.277***	0.202***	0.289***	0.388***	0.380***
	(0.012)	(0.024)	(0.033)	(0.020)	(0.021)
Non-minority	0.010	-0.063	0.061	0.029	0.060
-	(0.025)	(0.050)	(0.069)	(0.035)	(0.047)
Party	0.066***	0.089***	0.100**	-0.077**	0.002
-	(0.019)	(0.031)	(0.040)	(0.036)	(0.042)
Indus2	0.162***	0.291***	0.255**	0.335***	0.365***
	(0.032)	(0.097)	(0.115)	(0.039)	(0.060)
Indus3	0.064**	0.245**	0.208*	0.223***	0.292***
	(0.031)	(0.095)	(0.111)	(0.039)	(0.058)
Central provinces	-0.076***	-0.232***	-0.164***	0.018	0.003
-	(0.014)	(0.027)	(0.040)	(0.022)	(0.024)
West provinces	-0.054***	-0.146***	-0.126***	-0.102***	-0.021
-	(0.015)	(0.030)	(0.039)	(0.024)	(0.026)
Parents education	0.050***	-0.006	0.004	0.061	0.015
	(0.019)	(0.031)	(0.045)	(0.039)	(0.032)
Parents profession job	0.009	0.017	-0.015	-0.040	-0.003
	(0.026)	(0.037)	(0.066)	(0.048)	(0.052)
Parents public sector	-0.007	-0.024	0.052	-0.001	-0.011
	(0.028)	(0.037)	(0.069)	(0.056)	(0.064)
Parents party	0.001	0.008	-0.036	0.041	-0.056
	(0.020)	(0.032)	(0.046)	(0.037)	(0.039)
Ν	15832	3628	1940	4566	5698

	Overall		Male		Female	
	1 st -stage	2 nd -stage	1 st -stage	2 nd -stage	1 st -stage	2 nd -stage
Schooling years		0.130***		0.107***		0.152***
		(0.016)		(0.024)		(0.022)
Age	-0.443***	0.144***	-0.263	0.139***	-0.677***	0.153***
	(0.142)	(0.034)	(0.189)	(0.045)	(0.213)	(0.053)
Age2	0.003**	-0.001***	0.001	-0.001***	0.006**	-0.002***
	(0.002)	(0.000)	(0.002)	(0.001)	(0.002)	(0.001)
Central provinces	-0.626***	-0.017	-0.600***	-0.005	-0.675***	-0.041
	(0.056)	(0.019)	(0.074)	(0.025)	(0.087)	(0.029)
West provinces	-0.509***	-0.076***	-0.572***	-0.100***	-0.439***	-0.054*
	(0.065)	(0.020)	(0.085)	(0.028)	(0.099)	(0.030)
Urban*age12	3.258***	-0.285***	3.216***	-0.277***	3.313***	-0.281***
	(0.054)	(0.057)	(0.074)	(0.082)	(0.079)	(0.080)
Post-1980	-4.295***	0.459**	-3.618***	0.340	-5.107***	0.598*
	(0.881)	(0.209)	(1.175)	(0.274)	(1.327)	(0.322)
Post-1980 trend	-0.228***	0.021*	-0.187***	0.014	-0.278***	0.030*
	(0.046)	(0.011)	(0.062)	(0.014)	(0.069)	(0.017)
Male	-0.071	0.333***				
	(0.048)	(0.012)				
Uni_stud_enroll/pop	3.243***		2.938***		3.576***	
	(0.211)		(0.284)		(0.313)	
Diagnostic Tests:						
F-Statistics		236.43		106.78		130.97
Endo. (P-value)		0.000		0.017		0.000

Table B.2: Returns to education for China using our instrument and the Huang et al. (2022)'s controls

Ν	15501	15501	8879	8879	6622	6622
Note: Robust standard	errors are in	parentheses.	*, **, ***	denote signific	ance at 10%	, 5%, and 1%,
respectively. To better	compare wi	th Huang et a	ıl. (2022), a	all estimations i	n this Table	e used monthly
earnings as the depende	ent variable a	and the same of	control vari	ables as theirs, i	ncluding a p	post-1980 birth
year dummy, post-198	0 birth year	trend dummy	y, urban H	ukou at age 12	dummy, ag	ge, gender, and

regions. Following Huang et al. (2022), this Table restricted the analysis to individuals born in 1970

Table B.3: Returns to education for China using Huang et al. (2022)'s strategy

and later to exclude confounding effects of the "Compulsory Education Law."

	Overall		Male		Female	
	1 st -stage	2 nd -stage	1 st -stage	2 nd -stage	1 st -stage	2 nd -stage
Schooling years		0.131***		0.144**		0.133***
		(0.034)		(0.060)		(0.035)
Age	0.026	0.144***	0.120	0.135***	-0.089	0.149***
-	(0.139)	(0.034)	(0.186)	(0.047)	(0.210)	(0.052)
Age2	-0.002	-0.001***	-0.003	-0.001**	-0.001	-0.002***
	(0.002)	(0.000)	(0.002)	(0.001)	(0.002)	(0.001)
Central	-0.788***	-0.017	-0.744***	0.023	-0.860***	-0.058
provinces						
	(0.055)	(0.030)	(0.072)	(0.048)	(0.084)	(0.037)
West provinces	-0.820***	-0.076**	-0.848***	-0.069	-0.796***	-0.069*
	(0.061)	(0.032)	(0.081)	(0.054)	(0.094)	(0.036)
Urban*age12	2.760***	-0.287**	1.763*	-0.400**	3.896***	-0.215*
	(0.764)	(0.116)	(1.031)	(0.199)	(1.141)	(0.120)
Post-1980	0.422	0.459**	0.182	0.333	0.855	0.593*
	(0.860)	(0.210)	(1.137)	(0.286)	(1.309)	(0.314)
Post-1980 trend	0.022	0.021*	0.014	0.013	0.040	0.029*
	(0.045)	(0.011)	(0.060)	(0.015)	(0.068)	(0.016)
Male	-0.070	0.333***				
	(0.048)	(0.012)				
P*U12	-0.938		0.411		-2.613**	
	(0.787)		(1.065)		(1.171)	
BT*U12	-0.088**		-0.015		-0.178***	
	(0.034)		(0.047)		(0.051)	
P*BT*U12	-0.030		-0.069*		0.013	
	(0.031)		(0.041)		(0.046)	
Diagnostic Tests:						
F-Statistics		22.79		7.66		21.50
Endo. (P-value)		0.028		0.094		0.047
Overident. (P-		0.206		0.607		0.302
value)						
Ν	15501	15501	8879	8879	6622	6622

Note: Robust standard errors are in parentheses. *, **, *** denote significance at 10%, 5%, and 1%, respectively. To better compare with Huang et al. (2022), all estimations in this Table used monthly earnings as the dependent variable and the same control variables as theirs, including a post-1980 birth year dummy, post-1980 birth year trend dummy, urban Hukou at age 12 dummy, age, gender, and regions. Following Huang et al. (2022), this Table restricted the analysis to individuals born in 1970 and later to exclude confounding effects of the "Compulsory Education Law."

Table B.4: The first stage results of returns to education for each Hukou subgroup using TSLS (dependent variable: schooling years)

Overall	UNW	UHCW	RNW	RUMW

Uni_stud_enroll/pop	2.747***	1.833***	1.691***	2.274***	2.238***
	(0.200)	(0.293)	(0.407)	(0.358)	(0.434)
Married	-0.339***	-0.157	0.183	-0.690***	-0.222
	(0.076)	(0.131)	(0.221)	(0.120)	(0.142)
Professional	1.514***	1.292***	1.587***	0.724***	1.297***
	(0.057)	(0.086)	(0.131)	(0.101)	(0.107)
Public sector	2.030***	1.207***	1.384***	1.245***	1.525***
	(0.058)	(0.084)	(0.127)	(0.125)	(0.126)
Age	0.011	0.093*	0.031	-0.309***	-0.158***
	(0.028)	(0.050)	(0.077)	(0.044)	(0.057)
Age2	-0.001***	-0.002***	-0.002**	0.003***	0.000
	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)
Male	-0.036	-0.151*	0.086	0.154**	0.112
	(0.044)	(0.082)	(0.118)	(0.066)	(0.079)
Non-minority	-0.006	-0.022	-0.135	-0.183	-0.179
	(0.086)	(0.161)	(0.228)	(0.117)	(0.187)
Party	1.721***	1.394***	1.692***	1.535***	1.626***
	(0.068)	(0.096)	(0.135)	(0.125)	(0.160)
Indus2	0.394***	1.312***	0.084	0.147	0.267
	(0.107)	(0.380)	(0.399)	(0.123)	(0.235)
Indus3	1.370***	1.735***	0.717*	0.724***	0.734***
	(0.105)	(0.374)	(0.382)	(0.124)	(0.230)
Central provinces	-0.556***	-0.461***	-0.671***	-0.405***	-0.485***
	(0.050)	(0.098)	(0.150)	(0.071)	(0.091)
West provinces	-0.595***	-0.418***	-1.034***	-0.683***	-0.337***
	(0.057)	(0.112)	(0.151)	(0.084)	(0.102)
Parents education	1.464***	1.104***	0.708^{***}	0.871***	0.869***
	(0.066)	(0.100)	(0.155)	(0.136)	(0.125)
Parents profession job	0.339***	0.187	0.268	0.190	0.281
	(0.091)	(0.117)	(0.207)	(0.178)	(0.208)
Parents public sector	0.032	-0.325***	-0.191	0.154	0.614***
	(0.094)	(0.118)	(0.220)	(0.189)	(0.236)
Parents party	0.429***	0.304***	0.153	0.368***	0.237
	(0.071)	(0.101)	(0.155)	(0.120)	(0.147)
Ν	15832	3628	1940	5698	4566

Table B.5: The second stage results of returns to education for each Hukou subgroup using TSLS (dependent variable: log(hourly earning))

(dependent variable: log(nourly carining))							
	Overall	UNW	UHCW	RNW	RUMW		
Schooling years	0.172***	0.274***	0.250***	0.147***	0.048		
	(0.020)	(0.060)	(0.078)	(0.052)	(0.048)		
Married	0.127***	0.157***	0.134*	0.136***	0.038		
	(0.023)	(0.051)	(0.081)	(0.046)	(0.039)		
Professional	0.078**	0.018	-0.001	0.032	0.265***		
	(0.035)	(0.082)	(0.129)	(0.050)	(0.070)		
Public sector	-0.189***	-0.145*	-0.121	-0.271***	-0.026		
	(0.046)	(0.081)	(0.113)	(0.077)	(0.081)		
Age	0.076***	0.046***	0.037	0.079***	0.077***		
	(0.008)	(0.018)	(0.026)	(0.028)	(0.021)		
Age2	-0.001***	-0.000	-0.000	-0.001***	-0.001***		
-	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)		
Male	0.280***	0.167***	0.220***	0.351***	0.315***		

	(0.013)	(0.031)	(0.041)	(0.024)	(0.023)
Non-minority	0.010	-0.074	0.023	0.053	0.043
	(0.028)	(0.066)	(0.093)	(0.042)	(0.051)
Party	-0.125***	-0.161*	-0.175	-0.246***	0.035
-	(0.041)	(0.092)	(0.137)	(0.092)	(0.092)
Indus2	0.118***	-0.000	0.321**	0.177***	0.064
	(0.045)	(0.177)	(0.158)	(0.056)	(0.092)
Indus3	-0.092*	-0.146	0.130	-0.012	-0.029
	(0.053)	(0.188)	(0.161)	(0.069)	(0.099)
Central provinces	0.002	-0.128**	-0.053	0.095***	-0.015
-	(0.021)	(0.053)	(0.086)	(0.034)	(0.037)
West provinces	0.039	-0.042	0.095	0.126**	-0.019
	(0.024)	(0.055)	(0.115)	(0.053)	(0.036)
Parents education	-0.117***	-0.222***	-0.115	-0.023	0.023
	(0.036)	(0.079)	(0.071)	(0.061)	(0.051)
Parents profession job	-0.030	0.003	-0.009	-0.072	-0.013
	(0.029)	(0.049)	(0.076)	(0.057)	(0.058)
Parents public sector	-0.008	0.018	0.099	0.004	0.026
	(0.028)	(0.049)	(0.073)	(0.063)	(0.073)
Parents party	-0.047**	-0.054	-0.029	-0.019	-0.058
	(0.022)	(0.043)	(0.051)	(0.043)	(0.045)
F-Statistics	189.24	39.15	17.26	40.27	26.60
Endo. (P-value)	0.000	0.000	0.013	0.020	0.923
N	15832	3628	1940	5698	4566

	Overall	UNW	UHCW	RNW	RUMW
Uni_stud_enroll/pop	2.724***	1.545***	2.031***	2.270***	2.247***
	(0.216)	(0.300)	(0.425)	(0.405)	(0.490)
Married	-0.552***	-0.424***	0.093	-0.891***	-0.363**
	(0.083)	(0.140)	(0.243)	(0.130)	(0.150)
Professional	1.509***	1.240***	1.583***	0.779***	1.301***
	(0.062)	(0.092)	(0.142)	(0.117)	(0.114)
Public sector	2.038***	1.131***	1.274***	1.499***	1.686***
	(0.064)	(0.092)	(0.136)	(0.145)	(0.135)
Age	0.335***	0.475***	0.412***	-0.165**	0.141*
	(0.041)	(0.072)	(0.112)	(0.064)	(0.078)
Age2	-0.006***	-0.008***	-0.007***	0.001	-0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Male	-0.130***	-0.130	0.098	0.009	0.017
	(0.049)	(0.088)	(0.129)	(0.075)	(0.087)
Non-minority	0.017	0.069	-0.058	-0.204	-0.213
	(0.096)	(0.173)	(0.259)	(0.139)	(0.202)
Party	1.730***	1.163***	1.732***	1.705***	1.698***
	(0.076)	(0.104)	(0.146)	(0.159)	(0.177)
Indus2	0.456***	1.364***	-0.167	0.193	0.231
	(0.140)	(0.449)	(0.503)	(0.162)	(0.299)
Indus3	1.429***	1.708***	0.418	0.772***	0.732**
	(0.138)	(0.442)	(0.487)	(0.162)	(0.293)
Central provinces	-0.637***	-0.497***	-0.747***	-0.495***	-0.497***
	(0.057)	(0.106)	(0.165)	(0.083)	(0.100)

Table B.6: The first stage results of returns to education by dropping the missing enrollment data using TSLS (dependent variable: schooling years)

West provinces	-0.581***	-0.428***	-0.996***	-0.693***	-0.287***
-	(0.064)	(0.121)	(0.167)	(0.101)	(0.111)
Parents education	1.407***	1.150***	0.681***	0.781***	0.827***
	(0.068)	(0.101)	(0.157)	(0.136)	(0.126)
Parents profession job	0.354***	0.237**	0.328	0.138	0.307
	(0.091)	(0.114)	(0.208)	(0.176)	(0.207)
Parents public sector	0.042	-0.256**	-0.113	0.122	0.578**
	(0.096)	(0.119)	(0.222)	(0.189)	(0.238)
Parents party	0.479***	0.396***	0.203	0.408***	0.190
	(0.072)	(0.103)	(0.159)	(0.123)	(0.150)
Ν	12570	2914	1551	4207	3898

Table B.7: The second stage results in returns to education by dropping the missing enrollment data using TSLS (dependent variable: log(hourly earning))

	Overall	UNW	UHCW	RNW	RUMW
Schooling years	0.182***	0.338***	0.234***	0.198***	0.011
	(0.023)	(0.083)	(0.068)	(0.061)	(0.053)
Married	0.159***	0.241***	0.127	0.217***	0.024
	(0.026)	(0.068)	(0.081)	(0.061)	(0.044)
Professional	0.062	-0.069	0.001	0.013	0.319***
	(0.038)	(0.107)	(0.116)	(0.061)	(0.076)
Public sector	-0.237***	-0.244**	-0.108	-0.352***	0.027
	(0.051)	(0.103)	(0.095)	(0.104)	(0.097)
Age	0.044***	-0.029	0.024	0.036	0.096***
-	(0.011)	(0.039)	(0.039)	(0.028)	(0.018)
Age2	-0.000*	0.001	0.000	-0.000	-0.001***
-	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)
Male	0.291***	0.182***	0.208***	0.362***	0.311***
	(0.015)	(0.037)	(0.044)	(0.027)	(0.023)
Non-minority	0.029	-0.098	-0.010	0.095*	0.086
	(0.030)	(0.078)	(0.102)	(0.050)	(0.057)
Party	-0.135***	-0.199*	-0.164	-0.330***	0.112
	(0.046)	(0.107)	(0.126)	(0.119)	(0.103)
Indus2	0.135***	-0.104	0.128	0.184***	0.204**
	(0.052)	(0.225)	(0.150)	(0.065)	(0.104)
Indus3	-0.065	-0.239	-0.045	-0.023	0.153
	(0.061)	(0.239)	(0.146)	(0.081)	(0.111)
Central provinces	0.003	-0.104	-0.089	0.145***	-0.054
	(0.024)	(0.068)	(0.089)	(0.045)	(0.040)
West provinces	0.028	-0.015	0.065	0.147**	-0.035
	(0.026)	(0.070)	(0.107)	(0.063)	(0.037)
Parents education	-0.119***	-0.290***	-0.102	-0.049	0.057
	(0.038)	(0.107)	(0.064)	(0.065)	(0.053)
Parents profession job	-0.037	-0.014	-0.012	-0.084	-0.002
	(0.029)	(0.055)	(0.077)	(0.060)	(0.059)
Parents public sector	-0.005	0.031	0.089	-0.009	0.030
	(0.029)	(0.054)	(0.071)	(0.067)	(0.076)
Parents party	-0.054**	-0.084	-0.033	-0.051	-0.050
	(0.024)	(0.053)	(0.051)	(0.049)	(0.046)
F-Statistics	158.66	26.45	22.82	31.39	21.01
Endo. (P-value)	0.000	0.000	0.012	0.002	0.407
N	12570	2914	1551	4207	3898

	Overall	UNW	UHCW	RNW	RUMW
Uni_stud_enroll/pop	2.903***	2.069***	1.902***	2.156***	2.334***
	(0.197)	(0.292)	(0.399)	(0.350)	(0.422)
Married	-0.367***	-0.183	0.159	-0.706***	-0.243*
	(0.076)	(0.131)	(0.221)	(0.120)	(0.142)
Professional	1.511***	1.293***	1.591***	0.722***	1.293***
	(0.057)	(0.086)	(0.131)	(0.101)	(0.107)
Public sector	2.027***	1.204***	1.380***	1.246***	1.525***
	(0.058)	(0.084)	(0.127)	(0.124)	(0.126)
Age	0.014	0.107**	0.046	-0.323***	-0.159***
	(0.028)	(0.049)	(0.076)	(0.043)	(0.055)
Age2	-0.001***	-0.002***	-0.002**	0.003***	0.001
	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)
Male	-0.039	-0.154*	0.080	0.154**	0.111
	(0.044)	(0.081)	(0.118)	(0.066)	(0.079)
Non-minority	-0.012	-0.027	-0.140	-0.186	-0.180
	(0.086)	(0.161)	(0.228)	(0.117)	(0.187)
Party	1.717***	1.390***	1.690***	1.534***	1.620***
	(0.068)	(0.096)	(0.135)	(0.125)	(0.160)
Indus2	0.391***	1.298***	0.072	0.146	0.267
	(0.107)	(0.380)	(0.400)	(0.123)	(0.235)
Indus3	1.364***	1.716***	0.704*	0.724***	0.732***
	(0.105)	(0.374)	(0.384)	(0.124)	(0.230)
Central provinces	-0.547***	-0.438***	-0.648***	-0.408***	-0.484***
	(0.050)	(0.098)	(0.150)	(0.071)	(0.091)
West provinces	-0.579***	-0.387***	-1.003***	-0.688***	-0.331***
	(0.057)	(0.112)	(0.150)	(0.084)	(0.102)
Parents education	1.455***	1.105***	0.709***	0.860***	0.863***
	(0.066)	(0.100)	(0.155)	(0.136)	(0.125)
Parents profession job	0.346***	0.191	0.278	0.193	0.292
	(0.091)	(0.117)	(0.207)	(0.178)	(0.208)
Parents public sector	0.038	-0.317***	-0.190	0.157	0.617***
	(0.094)	(0.118)	(0.220)	(0.190)	(0.236)
Parents party	0.435***	0.306***	0.159	0.375***	0.236
	(0.071)	(0.101)	(0.155)	(0.120)	(0.147)
N	15832	3628	1940	5698	4566

Table B.8: The first stage results of returns to education with the measures for the IV at respondent age 19 using TSLS (dependent variable: schooling years)

Table B.9: The second stage results of returns to education with the measures for the IV at respondent age 19 using TSLS (dependent variable: log(hourly earning))

Tespondent age 19 using	respondent age 19 using TSLS (dependent variable, log(nourly earning))							
	Overall	UNW	UHCW	RNW	RUMW			
Schooling years	0.172***	0.264***	0.242***	0.150***	0.055			
	(0.019)	(0.053)	(0.069)	(0.055)	(0.045)			
Married	0.127***	0.156***	0.136*	0.138***	0.039			
	(0.023)	(0.050)	(0.080)	(0.047)	(0.039)			
Professional	0.078**	0.030	0.011	0.030	0.256***			

	(0.033)	(0.074)	(0.117)	(0.051)	(0.066)
Public sector	-0.189***	-0.132*	-0.110	-0.275***	-0.037
	(0.043)	(0.072)	(0.102)	(0.080)	(0.077)
Age	0.076***	0.045***	0.036	0.080***	0.079***
	(0.008)	(0.017)	(0.026)	(0.029)	(0.020)
Age2	-0.001***	-0.000	-0.000	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Male	0.280***	0.165***	0.221***	0.351***	0.314***
	(0.013)	(0.031)	(0.040)	(0.024)	(0.023)
Non-minority	0.010	-0.074	0.022	0.054	0.044
	(0.028)	(0.065)	(0.092)	(0.043)	(0.051)
Party	-0.126***	-0.147*	-0.163	-0.251***	0.024
	(0.039)	(0.082)	(0.124)	(0.095)	(0.087)
Indus2	0.118***	0.013	0.322**	0.177***	0.062
	(0.045)	(0.171)	(0.157)	(0.056)	(0.092)
Indus3	-0.092*	-0.128	0.136	-0.015	-0.035
	(0.053)	(0.180)	(0.158)	(0.071)	(0.098)
Central provinces	0.002	-0.135***	-0.059	0.096***	-0.012
	(0.020)	(0.049)	(0.079)	(0.035)	(0.036)
West provinces	0.039*	-0.048	0.085	0.129**	-0.016
	(0.023)	(0.051)	(0.104)	(0.054)	(0.035)
Parents education	-0.117***	-0.211***	-0.110*	-0.026	0.017
	(0.034)	(0.072)	(0.066)	(0.062)	(0.049)
Parents profession job	-0.031	0.005	-0.007	-0.073	-0.015
	(0.028)	(0.048)	(0.075)	(0.057)	(0.058)
Parents public sector	-0.008	0.014	0.097	0.004	0.021
	(0.028)	(0.047)	(0.072)	(0.063)	(0.072)
Parents party	-0.047**	-0.051	-0.028	-0.020	-0.059
	(0.022)	(0.041)	(0.049)	(0.044)	(0.045)
F-Statistics	217.19	50.35	22.69	37.94	30.57
Endo. (P-value)	0.000	0.000	0.008	0.023	0.958
Ν	15832	3628	1940	5698	4566

 N
 15852
 5026
 1940
 5076
 4500

 Note: Robust standard errors are in parentheses. *, **, *** denote significance at 10%, 5%, and 1%, respectively.
 second
 second

1 at	De B .10. Returns	10 0	uucation	101	various	occu	pational	sectors
Tab	Jo R 10. Roturne	to e	ducation	for	various	occur	national	sectors

Table B.10: Returns to	education for various	occupational sec	ctors	
	Non-p	ublic sector	Pub	lic sector
	1 st -stage	2 nd -stage	1 st -stage	2 nd -stage
Schooling years		0.124***		0.321***
		(0.020)		(0.064)
Married	-0.378***	0.113***	-0.189	0.136**
	(0.087)	(0.025)	(0.157)	(0.054)
Professional	1.461***	0.141***	1.550***	-0.144
	(0.073)	(0.035)	(0.090)	(0.102)
Age	0.014	0.074***	0.106*	0.047***
	(0.032)	(0.009)	(0.058)	(0.017)
Age2	-0.001***	-0.001***	-0.002***	-0.000
	(0.000)	(0.000)	(0.001)	(0.000)
Male	0.020	0.323***	-0.207**	0.167***
	(0.050)	(0.015)	(0.091)	(0.034)
Non-minority	0.072	0.016	-0.260	0.061
	(0.098)	(0.031)	(0.172)	(0.060)
Party	2.021***	-0.146***	1.431***	-0.225**
	(0.099)	(0.050)	(0.093)	(0.098)

Indus2	0.363***	0.110**	0.171	0.424**
	(0.111)	(0.045)	(0.346)	(0.168)
Indus3	1.334***	-0.067	1.162***	0.116
	(0.110)	(0.053)	(0.334)	(0.181)
Central provinces	-0.697***	0.019	-0.071	-0.151***
-	(0.056)	(0.023)	(0.110)	(0.040)
West provinces	-0.742***	0.019	-0.085	-0.010
-	(0.064)	(0.027)	(0.124)	(0.044)
Parents education	1.541***	-0.072*	1.244***	-0.220**
	(0.084)	(0.038)	(0.105)	(0.088)
Parents profession job	0.548***	-0.077**	0.091	0.060
	(0.122)	(0.037)	(0.129)	(0.050)
Parents public sector	0.211	-0.027	-0.083	0.011
	(0.132)	(0.037)	(0.130)	(0.050)
Parents party	0.552***	-0.045*	0.241**	-0.047
	(0.090)	(0.027)	(0.111)	(0.044)
Uni_stud_enroll/pop	3.332***		1.827***	
	(0.242)		(0.342)	
F-Statistics		189.43		28.51
Endo. (P-value)		0.000		0.000
N	12225	12225	3607	3607

TT 11 D 11 D (1	C	•		
Table B 11. Returns to	\mathbf{n}	education	tor	various	occumat	nong
1 abic D.11. Returns t	v	cuucuuon	101	various	occupat	nons

	Non-	Non-profession		ofession
	1 st -stage	2 nd -stage	1 st -stage	2 nd -stage
Schooling years		0.157***		0.330**
		(0.019)		(0.137)
Married	-0.386***	0.134***	-0.159	0.119*
	(0.086)	(0.025)	(0.169)	(0.063)
Public sector	1.975***	-0.159***	2.008***	-0.530*
	(0.070)	(0.043)	(0.107)	(0.277)
Age	0.008	0.078***	0.070	0.060***
	(0.032)	(0.009)	(0.063)	(0.020)
Age2	-0.001***	-0.001***	-0.002***	-0.000
	(0.000)	(0.000)	(0.001)	(0.000)
Male	0.038	0.311***	-0.445***	0.227***
	(0.048)	(0.015)	(0.106)	(0.071)
Non-minority	0.066	0.008	-0.260	0.066
	(0.097)	(0.031)	(0.177)	(0.076)
Party	1.820***	-0.109**	1.485***	-0.338
	(0.083)	(0.043)	(0.115)	(0.208)
Indus2	0.421***	0.121**	0.271	0.148
	(0.112)	(0.048)	(0.370)	(0.152)
Indus3	1.315***	-0.082	1.515***	-0.248
	(0.110)	(0.054)	(0.364)	(0.255)
Central provinces	-0.491***	0.022	-0.729***	0.012
	(0.056)	(0.021)	(0.115)	(0.116)
West provinces	-0.513***	0.046*	-0.839***	0.109
	(0.063)	(0.024)	(0.130)	(0.137)
Parents education	1.477***	-0.085**	1.413***	-0.366*
	(0.078)	(0.036)	(0.127)	(0.200)
Parents profession job	0.513***	-0.018	0.216	-0.075

Parents public sector	(0.113) 0.077 (0.112)	(0.036) -0.018 (0.033)	(0.154) -0.089 (0.174)	(0.066) 0.031 (0.066)
Parents party	0.502***	-0.074***	0.181	0.013
	(0.083)	(0.026)	(0.133)	(0.056)
Uni_stud_enroll/pop	3.252***		1.155***	
	(0.228)		(0.414)	
F-Statistics		204.19		7.80
Endo. (P-value)		0.000		0.013
Ν	12532	12532	3300	3300