College Education and Internal Migration in China *

Xiaozhou Ding†

May 25, 2021

Abstract

In this paper, I examine the causal impact of college education on young adults' out-province migration in China using China Family Panel Studies 2010 wave data. I use the number of colleges at the province-year level to identify the effect of college attendance on young adults' later life location choice. 2SLS estimates suggest that attending college significantly increases the likelihood of residing in a different province later in life by 7.5 percentage points. A series of tests shows that the impact of college on migration is heterogenous to people's childhood location, gender, hukou origin, and occupation.

Keywords: College Education; Internal Migration; China.

JEL Classification Codes: I25; J61; R23.

*I am indebted to Chris Bollinger, Bill Hoyt, and Steve Lugauer for their guidance and support throughout the whole project. I am also grateful for the helpful comments from the editor Suqin Ge, two anonymous referees, Shu Cai, Rui Du, Chenxu Hu, Samuel Ingram, Yang Jiao, Esteban Lopez, Anh Nguyen, Dongxiao Niu, Ashley O'Donoghue, Yaxiang Song, Kenneth Tester, and seminar and conference participants at the CES North America Annual Conference, the UEA North American Meeting, the NARSC Annual Conference, the KEA Annual Conference, the SEA Annual Meetings, and University of Kentucky.

†Department of Economics, Dickinson College. 28 N College St, P.O. Box 1773, Carlisle, PA 17013. Email: dingx@dickinson.edu.

1 Introduction

Geographic mobility is an equalizing force for wages and employment (Moretti, 2011). Movement across geographical boundaries has large implications, including affecting housing prices (Garriga et al., 2017), improving intergenerational mobility (Derenoncourt, 2018; Nakamura et al., 2019), reducing welfare dependence (Hartley, Lamarche and Ziliak, 2017), reducing income inequality (Razin and Sadka, 2016), improving health status (Black et al., 2015), and increasing education attainment (Chen et al., 2020). Understanding what factors affect migration, therefore, is important for both researchers and policymakers.

Among those factors, it is of particular interest how college education affects migration because college-educated workers are essential for both the local and national economy (Moretti, 2004). Longstanding empirical results show mobility rises with education (Molloy, Smith and Wozniak, 2011), and empirical evidence found in the U.S. and Europe support this relationship (Malamud and Wozniak, 2012; Machin, Pelkonen and Savanes, 2012; Weiss, 2015). However, the connection between education and internal migration in developing countries is less researched and may differ due to a large share of a less-educated population. China is a large developing country that has experienced rapid economic development and a sharp increase in human capital. How does education affect mobility in China therefore is informative to other developing countries.[1] Several reasons can explain the lack of studies looking into this relationship. First, there is a lack of data with detailed information on an individual's timing of education and residence location throughout time, so it is very challenging to investigate how education affects people's mobility later in life. Second, the composition of long-distance migrants does not vary in education due to the small share of college-educated people. Data from the Rural Household Survey suggests the majority of rural migrants are people who completed education less than secondary school (de Brauw and Giles, 2015). Since most colleges are located in cities (Xing and Zhang, 2017), seeking tertiary education in order to move to cities became one popular approach to overcome mobility barriers for young people. Once college education is completed, it is much easier to leave the rural area, therefore this mechanical rural-urban education-migration relationship is less attractive to researchers in the past.

China's higher education has been growing since the 1990s and an unprecedented expansion was implemented in 1999. The enrollment rate for high school graduates jumped from less than thirty percent to more than sixty percent in 1999 and was kept at a high level and now is eighty percent. Nearly seven million college students graduate every year. In this paper, I use the number of colleges at the province level to estimate the impact of college education on *lifetime cross-province* migration in China, including both unskilled and skilled labor.

Understanding the link between college education and mobility in China is important in several aspects. First and foremost, China is of special interest because of its sheer size of the migrant population. The latest data shows in 2017 there are 244 million people residing in a place that is

[1] Luo and Xing (2016) find that less educated people in China are more responsive to regional demand shifts. Lu and Xia (2016) use 2010 Census data to show unconditionally college workers are more likely to move intra-provincial while people without a college degree are conducting more inter-provincial migration.

different from their registered location, while every year, the higher education system adds seven million college students to the labor market. How do these college graduates choose where to live is an important issue. Second, College workers are vital for local and national economic development and understanding how inter-province migration of educated workers is shaped by college education is important for designing local public policies. Increasing enrollment locally could retain more college students (Groen, 2004; Winters, forthcoming); however, if college-educated workers are more responsive to labor market shocks (Wozniak, 2010), brain drain could happen to some local economies. Educated workers are attracted by different amenities (Su, Tesfazion and Zhao, 2017) and also make an area more desirable to live (Shapiro, 2006). There are cases in recent years where some provinces in China have a sizable higher education system but suffer losses of college graduates. Rural-urban and urban-urban migration might go beyond provincial boundaries and take time, which is less examined in the literature. It is important for local governments to understand the mechanism between college education and lifetime mobility.

In this paper, I look into how college education affects people's later life location choices in China. An individual is a migrant if she is residing in a province that is different from her province at age twelve. Throughout this paper, I treat the age twelve province as the home province. An individual is college-educated if she attended college. Therefore, the effect includes both college education per se and the locational effect associated with college. Unlike previous literature, it includes regional movement across administrative boundaries rather than focusing only on rural-urban migration. Regressions of migration on education are biased due to omitted variables and self-selection. People who would be likely to migrate also tend to earn more education. I utilize the variation in the province level number of colleges to examine the impact of college education on inter-province migration of young adults in China, using the China Family Panel Studies (CFPS) 2010 wave. It is the ideal dataset for this study for several reasons. First, it records adult respondent's residence locations of childhood and survey year. Second, it has rich information on the current year and duration of each level of education so that I can calculate the year survey respondents graduated from high school and went to college. Once the college-going year is obtained, I am able to match the province-year number of colleges to each individual. The exclusion restriction requires later life location choice is affected by education only through college-going year number of colleges. Because college admission is province-based, if people could select into provinces with more colleges, then the identification is invalid. But due to hukou restrictions, people have to take the college entrance exam in their hukou province no matter where their secondary education takes place. Selecting into advantageous provinces is unlikely in China. The other potential threat to my identification strategy is the correlation between migration and college expansion policy. However, pre-trend analysis suggests there is no clear relationship between the size of the expansion and the migration patterns at the province level. If anything, it will bias my estimates downward since a better location will attract people to stay, not to move away.

The IV estimates show a causal explanation that college education makes people more likely to reside in a different province than the childhood province. In general, attending college increases

one's out-province migration propensity by 7.5 percentage points, which is a large effect considering the average migration rate is around 9 percent in the data. The result is also robust to different measures of migration and subsample tests. I discuss several explanations that lead to the main results. I first examine how male and female high school graduates are affected differentially. The result shows college-educated men are more likely to migrate out compared to their high school counterparts. Next, I examine how childhood location affects out-province migration through college education. Then I look into the differential impact of college in terms of within-province and out-province hukou change. These results indicate that urban residents are less likely to conduct long-distance moves possibly due to the higher opportunity cost of changing hukou, which is associated with advantageous urban amenities; college is more salient in determining out-province permanent migration. Lastly, I investigate whether college-educated people are more likely to migrate out of their home provinces due to education-occupation mismatch. The result indicates occupational sorting of skilled labor is an important factor making college people move.

The contribution of this paper is threefold. First, to my best knowledge, this paper is the first attempt to link college education and out-province lifetime migration in China and finds it increases inter-province mobility of young adults. Second, unlike existing studies focusing on unskilled and hukou migration, it includes both hukou and non-hukou, skilled and unskilled labor migration and answers the question of whether people are more likely to move across provincial boundaries in a general context. Third, it adds Chinese evidence to the literature on the determinants of migration and finds a consistent story compared to the U.S. and Europe.

The rest of the paper is organized as follows: Section 2 discusses related literature on internal migration. Section 3 introduces the institutional background of China's internal migration and higher education. Section 4 and 5 describe the data and empirical strategy. Section 6 presents the empirical results. Section 7 concludes.

2 Motivation and Related Literature

There are two strands of literature focusing on different aspects of migration, typically the determinants and impacts of migration.[2] The determinant side is related to questions concerning why and who moves, which is believed to be affected by push- and pull- factors. Sjaastad (1962) was the earliest paper proposing the idea that whether an individual will move depends on the difference between the present value of income and costs of moving, where the latter includes both money and non-money costs and the migration rate increases with education and diminish with age and distance moved (Schwartz, 1976). Factors such as income (Kennan and Walker, 2011), wealth constraints (Dustmann and Okatenko, 2014), historical linkage between host and home locations (Kinnan, Wang and Wang, 2018), number and gender of children (Huang, Lin and Zhang, 2019), tax incentives (Agrawal and Foremny, 2019; Kleven et al., 2019), trade shocks (Greenland, Lopresti and McHenry, 2019; Tombe and Zhu, 2019; Fan, 2019), welfare reform (Kaestner et al., 2003),

[2]For detailed discussions of internal migration, please refer to two survey papers Lucas (1997) and Greenwood (1997) that focus on developing and developed countries respectively.

health care (Alm and Enami, 2017), local amenities (Su, Tesfazion and Zhao, 2017), environmental conditions (Chen, Oliva and Zhang, 2017; Khanna et al., 2019), and cyclical factors (Saks and Wozniak, 2011) all contribute to the determinants of moving.

People always move for better jobs and higher living conditions. In theory, it is generally believed that mobility rises with education. Empirical evidence shows more educated people tend to move more and take longer-distance migration even though there is a secular decline in mobility in the U.S. for every demographic group (Molloy, Smith and Wozniak, 2011, 2017). Malamud and Wozniak (2012) use the national and state-level induction risk during the Vietnam War to identify both college attainment and veteran status for men observed in the 1980 Census and examine the causal effect of education on migration using variation in college attainment. They find that an additional year of college education increases the probability of living outside one's birth state by 0.03 to 0.09 percentage points, which could account for 10 to 25 percent of the probability of moving for men in their sample. Wozniak (2010) uses local labor market demand shocks to examine the impact of education on mobility and finds college students are more responsive to distant labor market shocks and hence more likely to move than less educated people. Machin, Pelkonen and Savanes (2012) use Norwegian data to examine how education affects mobility using the compulsory education reform. Weiss (2015) adopts a similar identification strategy but includes more European countries. Both of the two papers find more education leads to higher mobility in Europe.

Most research studying China's internal migration has been focused on the rural-urban population flow and contemporaneous hukou migration because of its unique characteristics in terms of both large rural-urban divide and mobility restrictions. These types of migration not only have impacts on migrants themselves but also leave intergenerational problems (Chen and Feng, 2017, 2019). Even though local residence benefits are limited to people without hukou, less educated people from rural places still have large incentives to move into cities. Zhang and Song (2003) find that rural-urban migration contributes dominantly to urban population growth and is attracted by economic growth, not vice versa. Their analysis also shows that the rural-urban income gap encourages interprovincial migration. Other papers that look into the wage differential, education opportunities, and discrimination against migrant workers (Wang, Zhang and Ni, 2015, Zhang et al., 2016; Pakrashi and Frijters, 2017). The probability of rural migrants returning to home is increasing in education attainment, and more educated people are less likely to choose migration over local non-farm employment in order to avoid the risks associated with migration and the cost of being separated from families (Zhao, 1999, 2002).

Given the prosperous development of research on internal migration in China and its outcomes, there is a missing piece that lacks investigation. How does education affect people's migration in China? Few studies look into education-based migration. Because of the fact that regular higher institutions in China are located in cities (Xing and Zhang, 2017) and hukou restrictions, education-based migration creates a huge incentive for skilled people to move (Liao et al., 2017). Pan (2016) also finds such incentives to stay in school disappear after the removal of selective mobility restrictions. But it still lacks answers to the question of whether college people are more

likely to conduct long-distance moves since rural-urban migration includes both within-province and cross-province movement. Given the great influx of college students induced by the higher education expansion,[3] the relationship between education and migration in China demands more research since the migration of college students affects the spatial distribution of human capital, and it has important implications on economic development and local labor market.

3 Internal Migration and Higher Education in China

In this section, I first discuss the institutional background of mobility in China and its related migration measures. Then I discuss the higher education system in China and its recent expansion.

Hukou is a unique characteristic of China's socioeconomic development and was created in the 1950s to restrain people from moving between rural and urban as well as different cities in order to facilitate production under a centrally-planned economy. It does not only classify residents into rural and urban areas but also into specific urban areas. Hukou status was attached to an individual when he/she was born, mainly following parent's hukou status. For example, a child born in Shanghai can be registered in Beijing if the father has a Beijing hukou. As China began the opening-up and reform policies in the late 1970s, an increasing amount of labor has been freed from rural areas, leading to an increasing number of people flowing into cities. Figure 1 shows the migration population in China for the last three decades. It reflects the relaxation of mobility restrictions which allows more people to move and the dramatic size of the population that engages in moving within China. The migration control policies were gradually relaxed to meet the demand and since the 1990s, people could find a job without converting hukou. Even though the restrictions of moving have been reduced after the economic reform and allowing people to live anywhere in China, major public services such as health care and public education are still only available to local hukou holders. Education, joining the army, getting a job, and buying an apartment are among the most common ways one can change hukou. However, in some cases, they do not necessarily result in hukou conversion, which could be due to the stringent hukou restriction or personal choice. Table 1 presents different measures of internal migration in China. As can be seen from this table, hukou measures of both lifetime and contemporaneous migration is smaller than more general migration. Therefore, given that general moving across geographical areas is more driven by market conditions and the interest of this paper is to look at how college education affects later life location choice, I focus on out-province lifetime migration in this paper. Throughout this paper, I define the main measure of a migrant based on the current residence location (survey location) and the childhood location at age twelve, which is the nearest year I have to the college year.

The National College Entrance Exam (NCEE) is held in June every year, and eligible students take the exam, receive scores, and then apply for colleges. The enrollment is mainly administered by

[3] It increased new college enrollment from 1.5 million in 1999 to 6.9 million in 2012. Most of them look for a job and stay in cities after graduation. Based on a 2013 survey, 12.6 percent of college graduates migrate to a third place which is different from the origin and college locations, 11.1 percent are return migrants, 9.1 percent stick to where the college locates, and 54.3 are stayers, who learn and get employed in the original area without any migration (Yue, 2014).

the Ministry of Education and coordinated between provincial and central governments. Admission is based on a provincial quota system, so students from the same province compete with each other. The number of students that can be enrolled for each province depends on how many colleges a province has and the allocation between in-province and out-province quota. Students who fail to get into college can retake the exam in another year. In 1999, the central government implemented the policy to expand its higher education size to enroll more students. This expansion has led to millions of students getting into college. The enrollment rate went up dramatically in the early years of expansion and soon stabilized around 80 percent. The number of colleges also went up at the same time to meet the demand for higher education. Figure 2 plots the total number of colleges in China for each year, and we can see a massive increase post 1999.

4 Data

4.1 Data Sources

The China Family Panel Studies (CFPS) is a widely used longitudinal survey conducted by the Institute of Social Science Survey (ISSS) of Peking University. It is the most comprehensive panel survey data covering contemporary China. The first wave started in 2010, which is the baseline sample. It records detailed information on household member residence location, education attainment level and graduation year, economic conditions such as wage, income, and non-economic outcomes, including beliefs and health status. Five provinces were selected for initial sampling, and other provinces were later sampled to ensure representativeness. These 25 provinces contain 95% of the total population of China, and closely resemble the full population of China (Xie and Hu, 2014).

I collect college data for each province and each year from the China Yearly Statistical Book. It records the total number of regular higher institutions for each province from 1987, including two-year colleges that offer professional undergraduate degrees, which vary both across provinces and over time. Provinces had a great deal of discretion in the design and administration of their higher education institutions, and both the quantity and quality varied widely. Figure 3 shows variations in the number of colleges for selected provinces over time. I take advantage of the massive policy change to estimate how college education affected young adults' migration decisions.

In order to account for local economic factors at the province level, I collect data from the China Yearly Statistical Book on provincial total urban population, GDP per capita, and the total number of non-farm employment.

4.2 Key Variables

The primary interest of outcome is migration status. Measuring mobility involves defining geographic units of origin and destination locations and the time period people must move between origin and destination locations (Molloy, Smith and Wozniak, 2011). However, defining migration in a Chinese context is more complicated. There are at least three dimensions of migration under a Chinese context: rural-urban, hukou, and general type of moving. Existing literature focuses on the

first two types in part because of the lack of information on survey respondent's residence location over time. However, CFPS enable me to examine general migration as well as the other two types due to the rich information on residence location. I define out-province migration as the respondent is residing in a province that is different from the province at the age of twelve. CFPS asks for very detailed information about family member's residence province at birth, age three, age twelve, and the time of the survey, I use province at age twelve as a proxy to measure province status before college-going year. Though moving between age twelve and age eighteen is possible and may be endogenous to college-going decisions, given hukou restriction in China, this is unlikely. In fact, for those migrant children who reside out of their hukou registered province during their teenage time, they have to go back to take the college entrance exam (Chen and Feng, 2017), which would create a problem if we simply use hukou province or birth province because it does not accurately measure the education and locational effects they had before college. Using the age twelve province is more consistent with the goal of this paper that captures how college education affects later life location choices.

CFPS also obtains detailed information on education level, duration, and completion (drop-out) year and converts corresponding education level to years of schooling. I first calculate the starting year of each education level by using the end year subtracting duration. I assume the college year is the same year when high school is completed for those people without college. For those people who do not finish high school, I assign eighteen to them, which is the year they became adults. Then I am able to match the college-going year with the number of colleges in that province to exploit the variation of access to college on one's college status.

4.3 Sample Selection

I first restrict the age of respondents when taking the survey to be between 25 and 40, so I focus on the lifetime migration decision of young adults. People without detailed birth province and hukou information are not included in the sample. Individuals who took the college entrance exam prior to 1987 are also dropped because the earliest information of higher education at the province level is available from 1987, which corresponds to the birth cohort of 1970 that was supposed to take the college entrance exam in the same year. I also drop the cohort born after 1986 because they are supposed to take the college entrance exam after 2004 and will not be expected to graduate until 2009 when the survey was conducted. Lastly, I include a full sample of individuals from both urban and rural as well as women and men. Table 2 shows the comparison between the CFPS sample and the 2010 census using the same criteria. In general, the two key variables are very similar in both data sets. CFPS is more conservative in migration rate and has a slightly larger share of college people compared to the census.

4.4 Summary Statistics

Table 3 reports summary statistics for major variables in terms of an individual's migration status and college attainment. There are large differences in education, gender, age, and father's

education for non-migrant and migrant, while other individual characteristics are similar. However, non-college and college-educated people differ in almost every aspect. From columns (1) and (2), we see that migrants are more educated than non-migrants. The share of people who ever attended college is 16.3% and 22.7% respectively and the difference is around six percentage points and statistically significant. The propensity of cross-province migration for non-college and college individuals are also different, as shown in column (4) and (5), 8.1% for people without any college experience and 11.7% for people attended college. Females are more likely to reside in a different province compared to males but slightly less likely to become college students. Migrants are younger and have fewer siblings, which is in part due to the one-child policy, but the difference is not statistically significant at conventional levels. The age gap between non-college and college people, however, is larger, which could be associated with retaking the college exam. Urban children are also more likely to attend college as 48.9% of them had urban hukou at age three. In contrast, only 13.2% of non-college people had urban hukou at age three. Migration before age twelve shows whether the individual has changed residence location between birth and age twelve as previous migration experience could potentially affect future migration propensity. There are no noticeable differences between non-migrant and migrant in terms of this variable, but college people are more likely to have changed location in childhood. All the differences in observables indicate the importance of controlling these variables in a regression setting.

5 Empirical Strategy

The ordinary least squares estimation of migration on education takes the specification:

$$Mig_{ij} = \beta_0 + \beta_1 College_{ij} + \boldsymbol{\beta}_2 \mathbf{X}_i + \boldsymbol{\beta}_3 \mathbf{Z}_{js} + \phi_j + \theta_t + \psi_{rt} + u_{ij}. \tag{1}$$

Mig_{ij} is equal to one if individual i from province j resides outside the childhood province j in 2010 and zero otherwise. This specification controls for many personal attributes and unobserved heterogeneity. $\mathbf{X}_i$ is a vector of variables controlling for individual characteristics such as gender, number of siblings, ethnicity, and urban hukou status at age three. $\mathbf{Z}_{js}$ represents a set of variables accounting for provincial economic and labor market conditions for individuals from province j whose college-going year is s, which includes population, GDP growth rate, and urban employment. ϕ_j is province fixed effect, and θ_t denotes birth year fixed effect. I also include region-cohort fixed effect ψ_{rt} which uses east, central, and west regions to further account for locational time-varying effect at regional level such as movements across regions.

Even though I can control for many observed attributes and unobserved fixed effects, OLS estimates may still be biased due to the remaining factors in the error term. For example, the individual ability may be positively correlated with preferences for moving and also makes the individual more likely get into college, leading β_1 to be overestimated. In this paper, I use an instrumental variable approach to identify the effect of college education on out-province migration. An instrument for college needs to meet two conditions: correlated with college-going decisions and

exogenous to later life location choices through the error term. I use the number of colleges at the province level to identify the variation in college education following the spirit of Currie and Moretti (2003), among others.[4]

There are two reasons why I choose the number of colleges at the province level as the instrument for college education. First, college admission in China is based on a province system. Students from the same province compete with each other. Colleges favor local students; therefore a student in a province with more colleges likely to have a higher chance of enrolling in a college. Second, unlike using admission scores (Fan et al., 2018) and enrollment rates (Bollinger, Ding and Lugauer, 2020), the number of colleges measures the supply side shock rather than the equilibrium effect of both demand and supply. Another advantage of this instrument is that it will better measure the college probability for a student in a year while changes in the number of colleges, in contrast, are better to capture the changes in college probability. In addition, it is measured by the number of regular higher institutions in province j of college-going year s, which varies both across provinces and over time. It has an additional dimension of variation compared to using college cohort and exposure to higher education to instrument for college status (Bollinger and Hu, 2017). Consequently, the estimated effect is identified off people from the same province born in the same year but have different college-going years.

Since college status is a binary endogenous variable, I use a two-step IV method to estimate the impact of college on migration. First, I estimate the binary response model of college attendance on a set of covariates and the number of colleges Num_{js} by probit:

$$College_{ij} = \Phi\left(\gamma_0 + \gamma_1 Num_{js} + \gamma_2 \mathbf{X}_i + \gamma_3 \mathbf{Z}_{js} + \eta_j + \delta_t + \zeta_{rt} + \varepsilon_i\right) \tag{2}$$

and predict the likelihood of each individual going to college $\widehat{G}$. After obtaining the fitted probabilities $\widehat{G}$, I then estimate Equation (1) via two-stage least squares instrumental variable approach using $\widehat{G}$ as the instrument, which delivers an efficient estimator (Wooldridge, 2010):

$$Mig_{ij} = \beta_0 + \beta_1 \widehat{College}_{ij} + \boldsymbol{\beta}_2 \mathbf{X}_i + \boldsymbol{\beta}_3 \mathbf{Z}_{js} + \phi_j + \theta_t + \psi_{rt} + u_{ij}, \tag{3}$$

where

$$\widehat{College}_{ij} = \alpha_0 + \alpha_1 \widehat{G}_{ij} + \alpha_2 \mathbf{X}_i + \alpha_3 \mathbf{Z}_{js} + \mu_j + \omega_t + \xi_{rt} + \nu_{ij}. \tag{4}$$

Next, I discuss the validity of my identification strategy. The exclusion restriction requires that the number of colleges in each province has no influence on later life migration choices except through the channels of college attendance. My instrument could fail to meet this restriction if young people attempted to exploit provincial variation in college probability by moving between provinces. The first violation is unlikely because of the aforementioned hukou restriction. One can only take the college entrance exam in the hukou province while changing hukou status is quite

[4] Currie and Moretti (2003) and Cowan and Tefft (2020) use the number of colleges at the county level to instrument for mother and adult's education attainment. Böckerman and Haapanen (2013) use polytechnic school expansion in Finland to examine the supply of college on regional mobility.

difficult. In my main analysis, out-province migration is defined as current residence province is different from the province at age twelve. If some parents want to switch from a province with a small number of colleges to a province with a large number of colleges so that their children have a better chance of attending college, it is not easy to achieve due to the hukou restriction.

The second threat is the potential correlation between the size of higher education expansion and local economic conditions. One can imagine that a place with more educational resources might also have better amenities which affects people's migration propensity. Above all, the identification strategy relies on the exogeneity of higher education expansion in China. That is the supply of colleges in a province is independent of other variables that could potentially affect future migration decisions. Following the spirit of Böckerman and Haapanen (2013) along with other work on educational reforms in Norway (Black, Devereux and Salvanes, 2005; Machin, Pelkonen and Savanes, 2012), I test the exogeneity of this reform by including province-level population size, GDP growth rate, and unemployment rate together with province and year fixed effects. The results in Table 4 show that after controlling for such fixed effects, only population size meaningfully predict the magnitude of college expansion.

Can local government use the construction of colleges as a way to mitigate population flow? To further examine the possibility, I use 1990 and 2000 China censuses to construct migration patterns at the province level and conduct a pre-trend test. To be more specific, for each year, I first calculate the 5-year migration rate based on provinces people migrate out. Then I calculate the 5-year in-migration rate based on destination provinces. Once I obtain these two measures, I calculate the net migration rate by subtracting the out-migration rate from the in-migration rate and estimate the slope of the change in net migration rate between 1990 and 2000 on the change in the number of colleges between 1999 and 2004. Figure 4 shows there is no clear evidence that the increase in the number of colleges selects into provinces that draw more migrants. In addition, my identification strategy lands on comparing people born in the same year from the same province but have different education attainment due to the variability of the number of colleges at different years in the same province. If better education amenities will make such locations more desirable to live, it is supposed to retain people to stay in home provinces, rather than move away.

The third potential threat comes from the hukou reforms. Various papers utilize hukou policy reforms at different administrative levels to examine if those changes affect people's mobility though there is no consensus on findings. Sun, Bai and Xie (2011) find little evidence of hukou reform on migration. Kinnan, Wang and Wang (2018) however find strong evidence on rural-urban migration. Fan (2019) uses city-level data and finds strong evidence of hukou reform on the share of local hukou holders and migration rates. If the number of colleges is correlated with hukou reform, for instance, local governments could endogenously relax hukou policies and build more colleges to attract more human capital, then people from provinces with more relaxed hukou policies are also more likely to go to college. I use data collected by Fan (2019) along with the number of colleges at province-year level to test if hukou reforms are correlated with college construction. Figure 5 presents the evidence that there is no clear positive correlation between these two reforms. Hukou

reforms do not select into those provinces with more aggressive college expansion. Again, if there exists any possible connection between these two reforms, it could only result in an underestimate of college's impact on cross-province migration.

6 The Impact of College Education on Cross-Province Migration in China

This section reports the main findings by estimating Equation (1) via ordinary least squares and Equation (3) through a two-step instrumental variable method. Based on the estimates, attending college leads people 7.5 percentage points more likely to reside in a province that is different from the age twelve province. The results are robust to a series of tests.

6.1 Main Estimates

I first present the probit regression estimates for Equation (2) in Table 5. The dependent variable is an indicator variable of college attendance, with one being having attended any college. As can be seen from the top row, the number of colleges significantly predicts an individual's college probability after controlling for the province of childhood, birth year, and childhood region by cohort fixed effect. The coefficient on male show that men are more likely to go to college than women, but it is not statistically different. More siblings will lead to a lower probability of college, which is consistent with quantity and quality trade-off theory (Becker and Lewis, 1973). Father's education and childhood urban status also positively affect college attendance. Local economic conditions such as employment opportunities are positively correlated with college attendance while GDP growth rate affects college probably in the opposite way though with less precision. The coefficient of population implies students from larger provinces are more likely to attend college.

With those estimates in hand, I am able to predict college probability $\widehat{G}$. I then estimate Equation (3). The results are presented in Table 6. Columns (1)-(4) show first-stage estimates which consistently reveal a positive and significant correlation between predicted college probability originated from college openings and college attendance. The F-statistics for the joint significance test of the instrument indicate that using the number of colleges to predict college likelihood has sufficient explanatory power for my identification.[5] Columns (5)-(8) present reduced-form estimates of college attendance on migration, that is, the direct regressions of the excluded instrument along with all controls on out-province migration status. The results indicate large impacts from college probability: increasing one's college probability by one percentage point raises out-province migration propensity by 0.096 percentage points.

The corresponding main results are presented in Table 7. Column (1) to (4) report OLS estimates for Equation (1) while I increasingly add more controls. Column (1) shows that college education

[5] Please see Andrews, Stock and Sun (forthcoming) for a discussion of weak instrument and first-stage F statistics. They derive if there is only one endogenous variable, the Kleibergen-Paap robust F statistic is equivalent to the MOP effective F statistic (Montiel Olea and Pflueger, 2013).

increases out-province migration by 3.8 percentage points. However, after controlling individual characteristics and home province conditions, individuals with some college experience on average are 3.3 percentage points more likely to move out of their childhood province, suggesting migration is a result of self-selection that both observed and unobserved heterogeneity affect individual's mobility.

The 2SLS estimates are presented in columns (5) through (8) of Table 7. The set of control variables have similar signs and magnitudes in both OLS and IV regressions. Males are less likely to migrate compared to females with a coefficient of -3.3 percentage points, possibly through the marriage channel. Chinese females are commonly moving to their husbands' locations; therefore we see the differences between gender (Davin, 2007). More siblings create a higher probability for individuals to migrate out of their childhood place. They can afford to leave because their siblings could share the responsibility to co-reside with parents and provide old-age support. This is consistent with what Ma and Wen (2016) found that the probability of co-residence is positively associated with relative education of the children when parents can provide help but negatively associated with education when parents need help. There is no significant impact from childhood urban hukou status. Push factors such as GDP growth rate and employment opportunities at the province level have a quantitatively large but not statistically significant impact on moving. However, the signs are consistent with the intuition that the more rapid local economy is growing in the home provinces, and the more local jobs are available, the less likely people will migrate out.

The IV estimate for college is larger than OLS and is statistically significant at 5% level. Without any controls, the IV estimate is close to the OLS. Once individual characteristics and provincial variables are controlled, attending college makes people 7.5 percentage points more likely to migrate out of their home province compared to non-college people.[6] It is a large effect considering the mean migration rate is only 9 percent, so that college nearly doubles the migration propensity. The explanation for such a difference between OLS and IV is that using college as a treatment variable, the IV recovers the local average treatment effect which a particular group of higher returns to college is picked up (Angrist and Krueger, 2001). While the estimated effect is smaller than what Malamud and Wozniak (2012) found in the U.S. that attendance and completion of college increase out-of-state migration for men by 20 and 24.1 percentage points, it is because I use a different sample that includes all male and female people. I discuss the heterogenous effect on gender among other factors in the next section. Before that, I present the results of robustness checks that further support the causal relationship between college education and out-province migration in China.

6.2 Robustness Checks

In this section, I conduct a series of robustness checks using different first-stage functions and subsamples to examine the impact of college education on out-province migration in China. The

[6]The result of using changes in the number of colleges is presented in Appendix Table A1 and similar to the main estimate using levels of the number of colleges. But it is more intuitive to interpret how the supply of colleges affects college probability using levels.

results support the main estimates.

First, I use a logit model in the first stage to estimate the likelihood of attending college using the variation in the number of colleges. The result is shown in column (1) of Table 8. People with college education are 7.4 percentage points more likely to migrate to a different province in their mid-ages, comparing 7.5 percentage points of main results, the estimate is slightly smaller.

Second, early college cohorts in my sample were assigned jobs by the government (i.e., bao fen pei) prior to 1994, while later cohorts searched for jobs on their own. To alleviate this concern, I dropped cohorts whose college years were before 1991 (the expected graduation year is 1994). The result in column (2) shows a 9.3 percentage-point effect on college attendance, which is consistent with the main estimate. A more balanced sample consisting of five years before and after the expansion (1994-2004) shows a similar estimate of 8.7 percentage points in column (3).

What is driving the impact of college on out-province migration in China? Is it because of the regional disparity that more educated people have been moving from the west and central areas to the coastal region and more developed provinces in east China attract more college students? Figure 6 shows the geography of cross-province migration in China based on the share of in-migrants to current native residents calculated from CFPS data. As we can see from this figure, Beijing, Shanghai, and Chongqing, as the largest municipalities, are very diversified and have been long attracting people to move in. Other dark regions such as Guangdong, Zhejiang, and Jiangsu are also east and coastal provinces. There is a common perception of "peacock flies southeast",[7] which people use to describe the direction of population flow of skilled labor in China. I conduct a series of tests to see if the migration pattern is shaped by people move from inland China to these provinces and the east region. I first drop people who lived in Beijing and Shanghai during childhood. The estimate slightly decreases to 6.8, indicating a consistent effect of college on out-province migration. I then drop people whose location at age twelve was in Chongqing which also yields a similar result. Lastly, I exclude people who were in the east region when they were young. The coefficient is raised to 8.5 percentage points, higher than the estimate from the whole sample. The impact of college education on out-province migration is more pronounced for people who are from inland China.

6.3 Heterogeneous Effects

The empirical results evidently reveal a causal relationship between college education and long-distance move in China. In this section, I propose and explore several heterogenous impacts behind education and out-province migration in China. First, I investigate the effect from only comparing high school graduates and college people. Next, I discuss how the hukou system distorts the impact of college on out-province migration. The estimates show that hukou lowers the propensity of out-province migration in general but has a differential impact on people with different education levels. Then I discuss the education-occupation match and its impact on migration.

[7] The *Peacock Flies Southeast* is the first narrative poem in Chinese history.

6.3.1 The Impact on High School Educated People and Gender

The main analysis contains all individuals because college expansion likely increases education at all levels (Xing, 2014). However, as previously discussed that the effect may be picked up by a small group who are more affected by the policy. To examine such heterogeneity, I limit my sample to people who at least attended high school. Table 9 shows that college people are 24.5 percentage points more likely to migrate out compared to people with only high school education, but it is not statistically significant. However, if I split the sample to men and women and separately estimate the impact of college education on out-province migration, the result is surprisingly close to what is found in the U.S. Attending college increases out of childhood province by 29.9 percentage points for men compared to 20.1 in Malamud and Wozniak (2012).[8] The impact for high school women is smaller and less precise. This set of results is also consistent with the phenomenon that many unskilled people who have education less than high school participate in out-province migration in China, which pushes the estimated average effect down. It implies that the effect is higher for marginal people who are more affected by college opportunities.

6.3.2 The Impact of Hukou

Rural/Urban Origins Rural-urban migration has been heavily examined in the literature and considered as a large contribution to the urbanization of China (Zhang and Song, 2003; Chan, 2008; Liao et al., 2017). Since most unskilled workers come from rural areas and make up the majority of the migration population, it draws a lot of attention where urban-urban migration has received less focus.[9] In contrast, attending colleges in cities can be seen as a significant migration experience for rural students, and prior migration experience could affect future migration propensity (Faggian et al., 2007). Therefore, it is interesting to examine the heterogenous effects of college on out-province migration in terms of rural/urban origins. Figure 7 shows there exists an increasing gap in college opportunities between rural and urban China. If college education affects mobility and access to college is easier in the urban area, we should see higher mobility among urban residents because inverse migration (urban to rural) is less likely and less observed in data. However, the results show the opposite. College education affects long-distance move for rural people more significantly than urban counterparts. The first two columns in Table 10 present the estimates after I split the sample into two groups which include people who were holding rural and urban hukou at ages 3 and 12 respectively. A rural student who receives college education is 8.3 percentage points more likely to move out of their home province compared to rural people without college experience. Urban college people also have a higher propensity to move out of their home provinces, but not statistically different compared to urban non-college people. This is in line with what Colas and Ge (2019) found that the positive influence of education on an individual's migration decision is much larger for rural individuals. One explanation is that migration experience spurs further migration. Since most colleges are located in cities, rural college-educated students have to first move to urban

[8] 20.1 is using national level induction risk. When they use both national and state levels, the estimate is 54.4.
[9] One paper Ye et al. (2016) looks into this issue with a focus on high-skilled labor migration in China.

areas to receive college education. They tend to have more and a different migration experience due to vastly changing social and economic conditions than their urban counterpart. Therefore, a much larger effect is found on the rural sample. Another potential explanation for this is due to the hukou system. The opportunity cost of converting from one urban city to another is higher than from rural to urban, since people have to give up amenities associated with the original hukou, which offsets the gain from college effect on long-distance moving. Another possibility is that the local labor market absorbs rural people within the same province first. The impact of the GDP growth rate speaks to such an effect. Though it is not statistically significant, it places a large negative effect on out-province migration for rural children, implying that on average rural children from a more developed province relocate to close urban areas within the same province.

Hukou Migration Why are urban people more likely to receive a college education but on average are less likely to migrate out? The suspect is that the hukou conversion cost hinders people's mobility (Fan, 2019). As discussed earlier in this paper, another commonly used measure of migration in China is hukou migration. Moving within China has been increasingly popular while changing hukou status is not as easy as moving between locations. If college probability is positively correlated with general migration, it should also affect hukou migration. Columns (3)-(4) in Table 10 follow this idea and report two estimates based on different types of hukou migration.

I construct two new outcome migration measures, HK_{in} and HK_{out}. They are obtained by matching one's current hukou location with birth hukou location. Notice this measure does not require people to stay at the currently registered hukou location. One could have a registered Beijing hukou when he or she was born and never changed it, but is living in Shanghai now. This individual in the main analysis will be defined as a migrant but will not be treated as a migrant in this case since the registered hukou location has not changed. As long as the two hukou cities do not match with each other within a provincial administrative boundary, this respondent then is an in-province hukou migrant. If the current hukou city is in a different province from the birth city province, he or she will be labeled as an out-province hukou migrant.

Column (3) only includes individuals who are staying or moving within birth hukou provinces. I do not include people who moved hukou out of their birth province. The dependent variable is HK_{in} which equals one if the respondent's hukou is associated with a different city. I add back those people who moved hukou out of their home province in column (4).

The coefficients for college in both regressions are positive, but the effect on out-province migration is more pronounced and precise. Because residing and working in a city is practically unconstrained while changing hukou status often involves additional cost, we see a higher and statistically significant impact of college on out-province hukou migration. It also suggests that college students are more responsive to distant labor market demand shocks (Wozniak, 2010). College does not affect in-province hukou change may be due to the possibility that cities expand by absorbing the surrounding rural area first in the early stage, and people know cities closer to their birthplace better than cities far away in another province. It is easier for less skilled people to

get a job and settle down in a more integrated market within a short distance range. Therefore we observe such a less significant impact of college on in-province hukou migration.

This set of results first identifies that different original location generates a differential impact of college on out-province migration. It shows the distortion role of hukou on internal migration, which lowers the propensity of urban people with college experience to overcome the barrier of moving. Second, college education affects out-province hukou conversion significantly in both economic and statistical perspectives. As discussed earlier, hukou reforms do not select into provinces with more rapid higher education expansion and explain very little of the variation in the number of colleges. In addition, hukou policies in the sample period do not discriminate against people based on their original province. If hukou policies favor college students more, then all else equal people with college education could be more able to obtain hukou both in and out of their birth province.

6.3.3 Occupational Sorting

Occupational sorting is an important factor shaping college-educated people's migration propensity. Literature has found that education-occupation mismatch has contributed to migration due to overeducation (Quinn and Rubb, 2005). People in occupations that require less than their level of schooling are more likely to migrate to have a better match, while people in occupations that require more than their level of schooling tend to stay. To test this hypothesis, I examine the occupational choice of college-educated people and whether overeducation leads to out-home province migration.

First, using occupation codes provided in CFPS, I classify individuals with an occupation into six categories. CFPS use the International Standard Classification of Occupations (ISCO-88), which classifies jobs into nine major groups based on the tasks and duties undertaken in the job. I combine "professionals" and "technicians and associate professionals" into a relatively skilled group. I treat "crafts and related trades workers", "plant and machine operators and assemblers", and "elementary occupations" as another relatively unskilled group. Then I use the main IV specification to examine the effect of college attendance on occupational choice. The estimated results are listed in Table 11. It shows college experience sorts individuals into certain types of occupations such as professionals, technicians, and clerks but discourages them from being service workers.

Next, I calculate the average years of education each occupation requires at home province in 2010. Then I create a new variable called *overeducation*, measuring the difference between an individual's self-reported years of education and the home province-occupation average years of education. I estimate a variant of equations focusing on the effect of the education-occupation match on out-province migration. The results are in Table 12. As seen from the first column, overeducation leads to a higher propensity of being a migrant out of home province, but the magnitude is small and insignificant. Once controlling for college education, this effect disappears, implying a potential selection of education on migration. The inclusion of interaction between college and overeducation in column (3) is both economically large (2.3 percentage points for one year overeducation of college people) and statistically significant, indicating that college-educated people tend to move out of provinces where their education is above the average. The last two columns control for occupations,

which further supports the idea that occupational sorting of skilled labor.

Last, I split the sample into rural and urban origins as people might suspect that the effect is driven by rural people who go to colleges in urban areas seeking occupations such as professionals and technicians. The results are in Appendix Table A2, showing very similar estimates compared to the pooled regression, and the effect of education-occupation interaction is not statistically different between the two groups. Rural college-educated individuals move out of their home provinces where their education is above the average with 2.1 percent. The estimate for urban people is 1.9 percent. If occupation controls are included, urban individuals respond more aggressively with 2.2 percent, while the effect for rural people is 1.3 percent, but both are noisier compared to the pooled sample results. This set of results shows regardless of the rural and urban origin, college education enables people to move beyond administrative boundaries and respond to more distant job opportunities (Wozniak, 2010).

6.4 Specification Tests

In this section, I perform two specification checks to further test the validity of the empirical strategy and the robustness of the main results. Existing literature relies on the assumption that people go to college from their birth province at age 18, which may not be true if moves took place prior to the college-going year. The estimate of using birth and current province to measure migration with the number of colleges varying at age twelve provinces is likely to be biased because it underestimates the impact of college.[10] Also, if the number of colleges at the province-college-year level is correctly reflecting the college-going opportunities and affects the migration decision, then fake province and college years should not have a causal impact on one's migration decision later in life.

I first examine the impact of college education on birth province migration. The estimate is presented in column (1) of Table 13. Attending college increases one's propensity to leave birth province by 3.4 percentage points. Compared to 7.5 in the main result, this is smaller and not significant. This is consistent with the hypothesis that using birth location to examine college's impact on migration likely underestimates the effect.

Next, I randomly assign a province to each individual and forward and postpone the college year up to two years respectively and match the corresponding number of colleges. Then I reestimate the main specification holding other variables constant. Columns (2) to (5) show the corresponding results. Overall the mismatch of the number of colleges at the province-year level fails to identify the effect of college on out-province migration. All estimates are not statistically significant and small in magnitudes. These tests suggest my estimated impacts from college are causal.

[10] For instance, consider an individual who was born in province A, age twelve province is B, and eventually ended up in province C. In this case, the estimates of both birth province migration and age twelve migration yield the same result. However, if after attending college, this individual returns to province A, using birth province to measure migration will underestimate the impact.

7 Conclusion

In this paper, I use variation in the number of colleges in each province and college year level to instrument for people's college attendance and analyze its impact on cross-province migration in China. Unlike existing literature, I include both rural-urban and urban-urban migration as well as hukou and non-hukou migration to see the general effect of college. The 2SLS estimates show that attending college will increase the probability of residing in a province that is different from childhood province by 7.5 percentage points. The effect is large as it nearly doubles the average cross-province migration rate in China but consistent with what has been found in the U.S. literature. The impact of college on the less developed area is more pronounced in making college-educated people move out of home province. A series of tests shows that out-province migration is heterogenous to people's childhood location, gender, and hukou status. Urban local hukou holders are not affected by college on moving over provincial boundaries, possibly due to the migration cost of giving up benefits associated with existing hukou. High skilled people are more likely to change hukou both within- and out- province, but a more significant impact is found in out-province hukou conversion. College men are significantly more likely to move out of home province and the estimate is very close to what Malamud and Wozniak (2012) found in the U.S. College education could affect migration through both direct and indirect channels. I show that college-educated individuals sort into certain occupations such as professionals and technicians and are more likely to migrate if their education is above the required level of their occupation in their home provinces. Nonetheless, there are many indirect ways that college could affect people's migration. For instance, college-educated people may be in a better financial position to cover the upfront cost of migration due to college wage premium. Therefore, they are more likely to move. As Colas and Ge (2019) found that the unobserved effects embedded in the coefficient associated with education contribute to a large share of the change in migration rates using a decomposition approach, it is important to examine other mechanisms that college education affects out-province migration in the future research. Though the evidence shows a clear causal relationship between college attendance and internal migration including both education and location effects, it remains a question to what extent the contribution to migration comes from education per se and college location. I leave this important work for future research. This paper adds to the missing piece of linking education and internal migration in a large developing country during a period of college expansion. It provides evidence that college education makes people not only more educated but also more mobile in a Chinese context.

References

Agrawal, David R. and Dirk Foremny, "Relocation of the Rich: Migration in Response to Top Tax Rate Changes from Spanish Reforms," *The Review of Economics and Statistics*, 2019, *101* (2), 214–232.

Alm, James and Ali Enami, "Do Government Subsidies to Low-Income Individuals Affect Interstate Migration? Evidence from the Massachusetts Health Care Reform," *Regional Science and Urban Economics*, 2017, *66* (December 2016), 119–131.

Andrews, Isaiah, James H. Stock, and Liyang Sun, "Weak Instruments in IV Regression: Theory and Practice," *Annual Review of Economics*, forthcoming.

Angrist, Joshua D. and Alan B. Krueger, "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments," *Journal of Economic Perspectives*, 2001.

Becker, Gary S. and H. Gregg Lewis, "On the Interaction between the Quantity and Quality of Children," *Journal of Political Economy*, 1973, *81* (2), S279–S288.

Black, Dan A., Seth G. Sanders, Evan J. Taylor, and Lowell J. Taylor, "The Impact of the Great Migration on Mortality of African Americans : Evidence from the Deep South," *American Economic Review*, 2015, *105* (2), 477–503.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes, "Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital," *Am. Econ. Rev.*, 2005, *95* (1), 437–449.

Böckerman, Petri and Mika Haapanen, "The Effect of Polytechnic Reform on Migration," *Journal of Population Economics*, 2013, *26* (2), 593–617.

Bollinger, Christopher and Chenxu Hu, "China's Higher Education Expansion and College Premium: IV Estimations," *Working paper*, 2017.

_ , Xiaozhou Ding, and Steve Lugauer, "The Expansion of Higher Education and Household Saving in China," *Working paper*, 2020.

Chan, Kam Wing, "Internal Labor Migration in China: Trends, Geographical Distribution and Policies.," *Proceedings of the United Nations Expert Group Meeting on Population Distribution, Urbanization, Internal Migration and Development*, 2008, (December 2007), 93–122.

Chen, Shuai, Paulina Oliva, and Peng Zhang, "The Effect of Air Pollution on Migration: Evidence from China," 2017.

Chen, Yi, Ziying Fan, Xiaomin Gu, and Li An Zhou, "Arrival of Young Talent: The Send-Down Movement and Rural Education in China," *American Economic Review*, 2020, *110* (11), 3393–3430.

Chen, Yuanyuan and Shuaizhang Feng, "Quality of Migrant Schools in China: Evidence from a Longitudinal Study in Shanghai," *Journal of Population Economics*, 2017, *30* (3), 1007–1034.

_ **and** _ , "The Education of Migrant Children in China's Urban Public Elementary Schools: Evidence from Shanghai," *China Economic Review*, 2019, *54* (January), 390–402.

Colas, Mark and Suqin Ge, "Transformations in China's Internal Labor Migration and Hukou System," *Journal of Labor Research*, 2019.

Cowan, Benjamin W. and Nathan Tefft, "College Access and Adult Health," *NBER Working Paper*, 2020, p. No. 26685.

Currie, Janet and Enrico Moretti, "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings," *The Quarterly Journal of Economics*, 2003, *118* (4), 1495–1532.

Davin, Delia, "Marriage Migration in China and East Asia," *Journal of Contemporary China*, 2007, *16* (50), 83–95.

de Brauw, Alan and John Giles, "Migrant Opportunity and the Educational Attainment of Youth in Rural China," *The Journal of Human Resources*, 2015.

Derenoncourt, Ellora, "Can You Move to Opportunity? Evidence from the Great Migration," 2018.

Dustmann, Christian and Anna Okatenko, "Out-Migration, Wealth Constraints, and the Quality of Local Amenities," *Journal of Development Economics*, 2014, *110*, 52–63.

Faggian, Alessandra, Philip Mccann, and Stephen Sheppard, "Some Evidence that Women are More Mobile than Men: Gender differences in U.K. Graduate Migration Behavior," *Journal of Regional Science*, 2007, *47* (3), 517–539.

Fan, Elliott, Xin Meng, Zhichao Wei, and Guochang Zhao, "Rates of Return to Four-Year University Education: An Application of Regression Discontinuity Design," *The Scandinavian Journal of Economics*, oct 2018, *120* (4), 1011–1042.

Fan, Jingting, "Internal Geography, Labor Mobility, and the Distributional Impacts of trade," *American Economic Journal: Macroeconomics*, 2019, *11* (3), 252–288.

Garriga, Carlos, Aaron Hedlund, Yang Tang, and Ping Wang, "Rural-Urban Migration, Structural Transformation, and Housing Markets in China," *NBER Working Paper Series*, 2017, pp. 1–47.

Greenland, Andrew, John Lopresti, and Peter McHenry, "Import Competition and Internal Migration," *Review of Economics and Statistics*, 2019, *101* (1), 44–59.

Greenwood, Michael J., "Chapter 12 Internal Migration in Developed Countries," *Handbook of Population and Family Economics*, 1997, *1* (PART B), 647–720.

Groen, Jeffrey A., "The Effect of College Location on Migration of College-Educated Labor," *Journal of Econometrics*, 2004, *121* (1-2), 125–142.

Hartley, Robert, Carlos Lamarche, and James P. Ziliak, "Welfare Reform and the Intergenerational Transmission of Dependence," *IZA Discussion Paper*, 2017, (10942).

Huang, Zibin, Lin Lin, and Junsen Zhang, "Fertility, Child Gender, and Parental Migration Decision: Evidence from One Child Policy in China," 2019.

Kaestner, Robert, Neeraj Kaushal, and Gregg Van Ryzin, "Migration Consequences of Welfare Reform," *Journal of Urban Economics*, 2003, *53* (3), 357–376.

Kennan, John and James R. Walker, "The Effect of Expected Income on Individual Migration Decisions," *Econometrica*, 2011, *79* (1), 211–251.

Khanna, Gaurav, Wenquan Liang, A. Mushfiq Mobarak, and Ran Song, "The Productivity Censequences of Pollution-Induced Migration in China," Technical Report 2019.

Kinnan, Cynthia, Shing yi Wang, and Yongxiang Wang, "Access to migration for rural households," *American Economic Journal: Applied Economics*, 2018, *10* (4), 79–119.

Kleven, Henrik, Camille Landais, Mathilde Muñoz, and Stefanie Stantcheva, "Taxation and Migration: Evidence and Policy Implications," 2019.

Liao, Pei-Ju, Ping Wang, Yin-Chi Wang, and Chong Kee Yip, "Educational Choice, Rural-Urban Migration and Economic Development," *NBER Working Paper Series*, 2017.

Lu, Ming and Yiran Xia, "Migration in the People's Republic of China," *ADBI Working Paper Series*, 2016, (593).

Lucas, Robert E., "Internal Migration in Developing Countries," *Handbook of Population and Family Economics*, 1997, *1*, 721–798.

Luo, Dongdong and Chunbing Xing, "Population Adjustments in Response to Local Demand Shifts in China," *Journal of Housing Economics*, 2016, *33*, 101–114.

Ma, Sen and Fangqi Wen, "Who Coresides With Parents? An Analysis Based on Sibling Comparative Advantage," *Demography*, 2016, *53* (3), 623–647.

Machin, Stephen, Paku Pelkonen, and Kjell G. Savanes, "Education and Mobility," *Journal of the European Economic Association*, 2012, *10* (2), 417–450.

Malamud, Ofer and Abigail Wozniak, "The Impact of College on Migration," *Journal of Human Resources*, 2012, *47* (4), 913–950.

Molloy, Raven, Christopher L Smith, and Abigail Wozniak, "Internal Migration in the United States," *Journal of Economic Perspectives*, 2011, *25* (3—Summer), 173–196.

_ , Christopher L. Smith, and Abigail Wozniak, "Job Changing and the Decline in Long-Distance Migration in the United States," *Demography*, 2017, *54* (2), 631–653.

Montiel Olea, José Luis and Carolin Pflueger, "A Robust Test for Weak Instruments," *Journal of Business and Economic Statistics*, 2013, *31* (3), 358–369.

Moretti, Enrico, "Estimating the Social Return to Higher Education: Evidence from Longitudinal and Repeated Cross-Sectional Data," *Journal of Econometrics*, 2004, *121* (1-2), 175–212.

_ , *Local Labor Markets*, Vol. 4, Elsevier B.V., 2011.

Nakamura, Emi, Jósef Sigurdsson, and Jón Steinsson, "The Gift of Moving: Intergenerational Consequences of a Mobility Shock," 2019.

Pakrashi, Debayan and Paul Frijters, "Migration and Discrimination in Urban China: A Decomposition Approach," *Review of Income and Wealth*, 2017, *63* (4), 821–840.

Pan, Yao, "The Impact of Removing Selective Migration Restrictions on Education: Evidence from China," *Journal of Human Resources*, 2016, (May 2016).

Quinn, Michael A. and Stephen Rubb, "The Importance of Education-Occupation Matching in Migration Decisions," *Demography*, 2005, *42* (1), 153–167.

Razin, Assaf and Efraim Sadka, "How Migration Can Change Income Inequality?," 2016.

Saks, Raven E. and Abigail Wozniak, "Labor Reallocation over the Business Cycle: New Evidence from Internal Migration," *Journal of Labor Economics*, 2011, *29* (4), 697–739.

Schwartz, Aba, "Migration, Age, and Education," *Journal of Political Economy*, 1976, *84* (4, Part 1), 701–719.

Shapiro, Jesse M., "Smart Cities: Quality of Life, Productivity, and the Growth Effects of Human Capital," *The Review of Economics and Statistics*, 2006, *88* (May), 324–335.

Sjaastad, Larry A., "The Costs and Returns of Human Migration," *Journal of Political Economy*, 1962, *70* (5), 80–93.

Su, Yaqin, Petros Tesfazion, and Zhong Zhao, "Where Are the Migrants from? Inter-vs. Intra-Provincial Rural-Urban Migration in China," *China Economic Review*, 2017.

Sun, Wenkai, Chongen Bai, and Peichu Xie, "The Impact of Hukou Reforms on Rural Labor Migration in China," *Economic Research Journal*, 2011, *1*, 28–41.

Tombe, Trevor and Xiaodong Zhu, "Trade, Migration, and Productivity: A Quantitative Analysis of China," *American Economic Review*, 2019, *109* (5), 1843–1872.

Wang, Chunchao, Chenglei Zhang, and Jinlan Ni, "Social Network, Intra-Network Education Spillover Effect and Rural-Urban Migrants' Wages: Evidence from China," *China Economic Review*, 2015, *35*, 156–168.

Weiss, Christoph T., "Education and Regional Mobility in Europe," *Economics of Education Review*, 2015, *49*, 129–141.

Winters, John V., "In-State College Enrollment and Later Life Location Decisions," *Journal of Human Resources*, forthcoming, pp. 0916–8255R2.

Wooldridge, Jeffrey M., *Econometric Analysis of Cross Section and Panel Data*, MIT press, 2010.

Wozniak, Abigail, "Are College Graduates More Responsive to Distant Labor Market Opportunities?," *Journal of Human Resources*, 2010, *45* (4), 944–970.

Xie, Yu and Jingwei Hu, "An Introduction to the China Family Panel Studies (CFPS)," *Chinese Sociological Review*, 2014, *47* (1), 3–29.

Xing, Chunbing, "Education Expansion, Migration and Rural-Urban Education Gap: A Case Study on the Effect of University Expansion," *China Economic Quarterly*, 2014, *13* (1), 207–232.

_ **and Junfu Zhang**, "The Preference for Larger Cities in China: Evidence from Rural-Urban Migrants," *China Economic Review*, 2017, *43*, 72–90.

Ye, Jingjing, Xiaokai Wu, and Jijun Tan, "Migrate to Skilled Cities: Human Capital Agglomeration and Urban-to-Urban Migration in China," *Emerging Markets Finance and Trade*, 2016, *52* (8), 1762–1774.

Yue, Changjun, "Gender and Migration: Evidence from College Graduates in China," in "Education's Role in Preparing Globally Competent Citizens" 2014.

Zhang, Kevin Honglin and Shunfeng Song, "Rural-Urban Migration and Urbanization in China: Evidence from Time-Series and Cross-Section Analyses," *China Economic Review*, 2003, *14* (4), 386–400.

Zhang, Li, Rhonda Vonshay Sharpe, Shi Li, and William A. Darity Jr, "Wage Differentials between Urban and Rural-Urban Migrant Workers in China," *China Economic Review*, 2016, *41*, 222–233.

Zhao, Yaohui, "Labor Migration and Earnings Differences: The Case of Rural China," *Economic Development and Cultural Change*, 1999, *47* (4), 767–782.

_ , "Causes and Consequences of Return Migration: Recent Evidence from China," *Journal of Comparitive Economics*, 2002, *30* (2), 376–394.

8 Tables

Table 1: Measures of Internal Migration in China

Origin	Destination	Migrants	Non-migrants	Migration Rate
Lifetime Measures				
Age 12 province	Current province	566	5,387	9.51%
Birth province	Current province	630	5,322	10.58%
Birth hukou province	Current hukou province	321	5,541	5.48%
Age 3 Rural	Current urban	1,082	2,425	30.85%
Contemporaneous				
Current hukou province	Current province	347	5,597	5.84%
Current hukou city	Current city	754	5,144	12.78%
Current hukou rural	Current urban	1,394	3,693	27.40%

Note: This table shows different measures of internal migration based on origin and destination location and time period. Author's calculation based on 2010 CFPS data. The sample consists of adults age between 25 and 40.

Table 2: Comparisons of CFPS and Census on Migration and Education

Measure	CFPS 2010	Census 2010
Birth province migration rate	10.58%	11.22%
Share of college people	16.55%	14.65%

Note: This table shows migration rates and college people shares in CFPS and census data. The sample consists of adults age between 25 and 40.

Table 3: Summary Statistics

	Non-migrant (1)	Migrant (2)	Diff (3)	Non-college (4)	College (5)	Diff (6)
Migrant				0.082	0.117	0.036**
				(0.273)	(0.321)	[3.098]
College	0.163	0.227	0.063**			
	(0.370)	(0.419)	[3.112]			
Male	0.522	0.410	-0.112***	0.510	0.525	0.015
	(0.500)	(0.492)	[-4.668]	(0.500)	(0.500)	[0.816]
Birth year	1977.1	1977.6	0.574**	1976.9	1978.2	1.275***
	(4.676)	(4.467)	[2.618]	(4.653)	(4.551)	[7.597]
Age	32.95	32.37	-0.577**	33.11	31.84	-1.274***
	(4.671)	(4.467)	[-2.634]	(4.650)	(4.542)	[-7.607]
College age	18.56	18.49	-0.066	18.24	20.06	1.812***
	(1.706)	(1.622)	[-0.824]	(0.970)	(3.120)	[17.179]
Han ethnicity	0.935	0.948	0.012	0.933	0.952	0.018*
	(0.246)	(0.223)	[1.114]	(0.249)	(0.214)	[2.265]
No. siblings	1.970	1.930	-0.040	2.114	1.239	-0.876***
	(1.465)	(1.470)	[-0.552]	(1.462)	(1.251)	[-18.497]
Father years of schooling	6.451	7.187	0.737***	5.916	9.461	3.545***
	(4.298)	(4.545)	[3.334]	(4.149)	(3.949)	[24.230]
Urban hukou at age 3	0.188	0.237	0.050*	0.132	0.489	0.357***
	(0.391)	(0.426)	[2.396]	(0.339)	(0.500)	[20.379]
Migration before age 12	0.0602	0.0632	0.003	0.0504	0.110	0.059***
	(0.238)	(0.244)	[0.254]	(0.219)	(0.313)	[5.416]
Observations	4,820	459	5,279	4,387	892	5,279

Note: This table shows summary statistics of major variables in the sample for the main analysis. Standard deviations are in parentheses. t-statistics of differences are in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Exogeneity Tests of College Expansion

	(1)	(2)	(3)	(4)	(5)
Population	0.005***	0.055***	0.028***	0.005***	0.007***
	(0.000)	(0.005)	(0.003)	(0.000)	(0.000)
GDP growth rate	1.698	44.056***	9.553	-8.542	26.143
	(9.878)	(9.560)	(10.850)	(13.319)	(32.450)
Unemployment rate	-10.797***	-6.625***	-0.564	1.136	7.398
	(3.517)	(2.533)	(2.098)	(2.611)	(5.819)
Province FE	N	Y	Y	N	N
Year FE	N	N	Y	Y	Y
Regional FE	N	N	N	Y	Y
R^2	0.219	0.608	0.832	0.743	0.771
Observations	618	618	618	618	299

Note: This table shows the results of the exogeneity test. Dependent variables in each column are the total enrollment of college students within the province. The first four columns use the sample from 1987 to 2008. The last column uses the sample after the expansion in 1999. Robust standard errors are in parentheses.

Table 5: Effects of the Number of Colleges on College Attendance

	(1)	(2)	(3)	(4)
No. college	0.039***	0.037***	0.019*	0.024***
	(0.009)	(0.008)	(0.010)	(0.009)
Male		0.032	0.024	0.026
		(0.040)	(0.041)	(0.041)
Han ethnicity		0.006	0.031	0.021
		(0.095)	(0.092)	(0.090)
No. siblings		-0.132***	-0.145***	-0.143***
		(0.030)	(0.029)	(0.029)
Father years of schooling		0.091***	0.090***	0.088***
		(0.007)	(0.008)	(0.008)
Urban hukou at age 3		0.730***	0.744***	0.744***
		(0.088)	(0.090)	(0.091)
Migration before age 12		0.037	0.058	0.080
		(0.102)	(0.110)	(0.108)
GDP growth rate			-0.252	-0.536
			(1.238)	(1.241)
Urban employment			0.419	0.636*
			(0.376)	(0.380)
Population			1.088	1.187
			(0.699)	(0.785)
Province FE			✓	✓
Cohort FE			✓	✓
Region-Cohort FE			✓	✓
Base GDP*Cohort FE				✓
Observations	5,441	5,441	5,441	5,441

Note: Dependent variables are college dummies that equal one if the respondent attended college. Robust standard errors are clustered at the age twelve province level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: 2SLS First-Stage and Reduced-Form Results

	First-stage				Reduced-form			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\widehat{G}$	1.193***	1.225***	1.277***	1.273***	0.076	0.100*	0.104*	0.096*
	(0.200)	(0.125)	(0.104)	(0.101)	(0.059)	(0.053)	(0.055)	(0.050)
Male		-0.004	-0.004	-0.003		-0.033***	-0.033***	-0.033***
		(0.007)	(0.007)	(0.007)		(0.009)	(0.009)	(0.009)
Han ethnicity		-0.000	-0.001	-0.001		-0.002	-0.002	-0.002
		(0.015)	(0.014)	(0.014)		(0.036)	(0.036)	(0.036)
No. siblings		0.004	0.006	0.006		0.008***	0.009***	0.008***
		(0.004)	(0.004)	(0.004)		(0.002)	(0.003)	(0.003)
Father years of schooling		-0.004	-0.004**	-0.004**		0.000	-0.000	0.000
		(0.002)	(0.002)	(0.002)		(0.001)	(0.001)	(0.001)
Urban hukou at age 3		-0.047	-0.058*	-0.058*		0.008	0.007	0.010
		(0.040)	(0.033)	(0.035)		(0.024)	(0.020)	(0.019)
Migration before age 12		-0.006	-0.006	-0.006		-0.033**	-0.033**	-0.035**
		(0.026)	(0.027)	(0.026)		(0.015)	(0.015)	(0.015)
GDP growth rate			0.017	0.021			-0.133	-0.128
			(0.270)	(0.278)			(0.119)	(0.121)
Urban employment			-0.039	-0.046			-0.026	-0.032
			(0.068)	(0.063)			(0.030)	(0.032)
Population			-0.043	-0.042			-0.009	-0.009
			(0.049)	(0.053)			(0.014)	(0.013)
Province FE			✓	✓			✓	✓
Cohort FE			✓	✓			✓	✓
Region-Cohort FE			✓	✓			✓	✓
Base GDP*Cohort FE				✓				✓
F statistic	35.49	96.76	149.69	159.65				
Observations	5,441	5,441	5,441	5,441	5,441	5,441	5,441	5,441

Note: This table presents the results of 2SLS estimation in the second step. Columns (1)-(4) show first-stage estimates and dependent variables are college attendance dummies. Columns (5)-(8) show reduced-form results and dependent variables are migration dummies. Robust standard errors are clustered at the age twelve province level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Effects of College Education on Out-Province Migration

	OLS (1)	OLS (2)	OLS (3)	OLS (4)	IV (5)	IV (6)	IV (7)	IV (8)
No. college	0.038*	0.033*	0.032*	0.033*	0.064	0.081*	0.081**	0.075**
	(0.021)	(0.018)	(0.018)	(0.018)	(0.053)	(0.044)	(0.041)	(0.037)
Male		-0.033***	-0.033***	-0.033***		-0.033***	-0.033***	-0.033***
		(0.009)	(0.009)	(0.009)		(0.009)	(0.009)	(0.009)
Han ethnicity		-0.002	-0.002	-0.002		-0.002	-0.002	-0.002
		(0.036)	(0.036)	(0.036)		(0.035)	(0.035)	(0.035)
No. siblings		0.007**	0.007**	0.007**		0.008***	0.008***	0.008***
		(0.003)	(0.003)	(0.003)		(0.002)	(0.002)	(0.002)
Father years of schooling		0.001	0.001	0.001		0.000	0.000	0.000
		(0.001)	(0.001)	(0.001)		(0.001)	(0.001)	(0.001)
Urban hukou at age 3		0.023	0.023	0.024		0.012	0.012	0.014
		(0.018)	(0.018)	(0.018)		(0.021)	(0.018)	(0.017)
Migration before age 12		-0.031**	-0.032**	-0.033**		-0.033**	-0.033**	-0.034**
		(0.015)	(0.015)	(0.015)		(0.015)	(0.015)	(0.014)
GDP growth rate			-0.137	-0.133			-0.135	-0.129
			(0.118)	(0.121)			(0.116)	(0.118)
Urban employment			-0.015	-0.020			-0.023	-0.029
			(0.029)	(0.030)			(0.031)	(0.032)
Population			0.004	0.002			-0.005	-0.006
			(0.012)	(0.012)			(0.013)	(0.013)
Province FE			✓	✓			✓	✓
Cohort FE			✓	✓			✓	✓
Region-Cohort FE			✓	✓			✓	✓
Base GDP*Cohort FE				✓				✓
Observations	5,441	5,441	5,441	5,441	5,441	5,441	5,441	5,441

Note: Dependent variables are dummies of cross-province migration (mean=0.09). Robust standard errors are clustered at the age twelve province level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: Robustness Checks

	(1) Logit	(2) Drop Bao fen pei	(3) Balanced sample	(4) Drop Beijing & Shanghai	(5) Drop Chongqing	(6) Drop East region
College	0.074**	0.093**	0.087*	0.068*	0.072*	0.085*
	(0.036)	(0.041)	(0.048)	(0.035)	(0.037)	(0.045)
Male	-0.033***	-0.038***	-0.037***	-0.034***	-0.033***	-0.039***
	(0.009)	(0.010)	(0.011)	(0.009)	(0.009)	(0.011)
Han ethnicity	-0.002	-0.010	-0.010	-0.002	-0.003	0.000
	(0.035)	(0.038)	(0.042)	(0.036)	(0.035)	(0.035)
No. siblings	0.008***	0.010***	0.011***	0.008***	0.008***	0.007**
	(0.002)	(0.003)	(0.004)	(0.002)	(0.002)	(0.003)
Father years of schooling	0.001	0.001	0.000	0.001	0.001	0.001
	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Urban hukou at age 3	0.015	0.010	0.015	0.017	0.013	0.011
	(0.016)	(0.016)	(0.022)	(0.017)	(0.017)	(0.018)
Migration before age 12	-0.034**	-0.041**	-0.042**	-0.037**	-0.034**	-0.046***
	(0.014)	(0.017)	(0.020)	(0.015)	(0.015)	(0.015)
GDP growth rate	-0.130	-0.078	-0.098	-0.127	-0.096	-0.085
	(0.118)	(0.111)	(0.115)	(0.121)	(0.115)	(0.125)
Urban employment	-0.028	-0.049	-0.122***	-0.028	-0.030	-0.024
	(0.031)	(0.036)	(0.035)	(0.032)	(0.031)	(0.064)
Population	-0.005	-0.010	-0.003	-0.002	-0.004	-0.005
	(0.013)	(0.017)	(0.028)	(0.012)	(0.013)	(0.020)
F statistic	157.536	141.684	100.331	130.495	156.879	152.595
Observations	5,441	4,408	3,409	5,093	5,420	4,302

Note: This table uses alternative specifications to examine the robustness of the main model. Column (1) uses the logit model in the first step predicting college probabilities. Column (2) drops cohorts whose college-going years were before 1991. Column (3) further drops cohorts whose college-going years were before 1994. Column (4) drops individuals whose age twelve provinces were Shanghai or Beijing. Column (5) drops people whose age twelve provinces were Chongqing. Column (6) drops people who lived in the east region in childhood. All specifications control the full set fixed effects in column (8) Table 7. Robust standard errors are clustered at the age twelve province level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 9: Effects of College on High School People and Gender

	Pooled (1)	Men (2)	Women (3)
College	0.245	0.299**	0.034
	(0.151)	(0.128)	(0.139)
Male	-0.041***		
	(0.010)		
Han ethnicity	-0.016	0.036	-0.042
	(0.055)	(0.036)	(0.059)
No. siblings	0.017	0.009	-0.002
	(0.011)	(0.008)	(0.015)
Father years of schooling	0.000	-0.005	0.010*
	(0.003)	(0.004)	(0.006)
Urban hukou at age 3	-0.003	-0.006	0.025
	(0.021)	(0.026)	(0.043)
Migration before age 12	-0.057**	-0.032	-0.071***
	(0.026)	(0.031)	(0.026)
GDP growth rate	-0.146	-0.194	0.162
	(0.161)	(0.184)	(0.267)
Urban employment	-0.074	-0.121*	0.009
	(0.056)	(0.070)	(0.061)
Population	-0.049	-0.074	0.040
	(0.064)	(0.053)	(0.058)
F statistic	18.14	21.347	34.459
Observations	1,917	1,007	903

Note: This table provides IV estimates for people with at least high school education. Column (2) uses only male people and column (3) uses only female people. All specifications control the full set fixed effects in column (8) Table 7. Robust standard errors are clustered at the age twelve province level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 10: Heterogeneous Effects of College on Migration by Rural and Urban Origins and Hukou

	Childhood in Rural (1)	Childhood in Urban (2)	In-province hukou migration (3)	Out-province hukou migration (4)
College	0.083*	0.028	0.037	0.142***
	(0.047)	(0.062)	(0.047)	(0.048)
Male	-0.026**	-0.035**	-0.032***	-0.039***
	(0.011)	(0.015)	(0.011)	(0.009)
Han ethnicity	0.016	-0.097	0.037**	0.001
	(0.037)	(0.068)	(0.019)	(0.027)
No. siblings	0.010***	-0.011	-0.004	0.008***
	(0.003)	(0.010)	(0.005)	(0.003)
Father years of schooling	0.000	0.003	0.001	-0.001
	(0.001)	(0.002)	(0.002)	(0.001)
Urban hukou at age 3			-0.057**	-0.011
			(0.028)	(0.027)
Migration before age 12	-0.012	-0.030*	0.257***	0.134***
	(0.021)	(0.018)	(0.050)	(0.034)
GDP growth rate	-0.098	0.082	-0.208	-0.008
	(0.154)	(0.241)	(0.150)	(0.081)
Urban employment	-0.027	-0.048	0.053*	0.007
	(0.033)	(0.111)	(0.031)	(0.040)
Population	-0.006	0.021	-0.030**	-0.009
	(0.016)	(0.035)	(0.012)	(0.033)
F statistic	59.132	87.694	181.21	116.775
Observations	4,223	1,012	5,078	5,369

Note: This table provides IV estimates for people from different childhood hukou origins and hukou migration. Columns (1) and (2) use people whose age 3 and 12 hukou were rural and urban respectively. Dependent variables are age twelve migration. Columns (3) and (4) look into how people change their hukou status. Dependent variables are corresponding hukou migration dummies. All specifications control the full set of fixed effects as in Table 7 column (8). Robust standard errors are clustered at the age twelve province level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 11: Occupational Sorting of College Education

	(1) Legislators, senior officials and managers	(2) Professionals and technicians	(3) Clerks	(4) Service workers	(5) Skilled agricultural and fishery workers	(6) Machine operators and other workers
College	0.057	0.366***	0.113**	-0.242***	0.141**	-0.109
	(0.042)	(0.082)	(0.043)	(0.069)	(0.053)	(0.080)
Observations	4,408	4,408	4,408	4,408	4,408	4,408

Note: This table presents IV results of how college education affects the choice of occupations. Dependent variables are occupation dummies. Personal demographical variables and provincial statistics in the college year are controlled. All specifications control the full set fixed effects in column (8) Table 7. Robust standard errors are clustered at the age twelve province level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 12: Effects of Education-Occupation Match on Migration

	(1)	(2)	(3)	(4)	(5)
Overeducation	0.001	-0.002	-0.006	0.000	-0.002
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
College		0.031	-0.011	0.021	-0.025
		(0.021)	(0.023)	(0.023)	(0.025)
Overeducation×College			0.023**		0.019*
			(0.009)		(0.009)
Occupation Controls				✓	✓
Observations	3,627	3,627	3,627	3,627	3,627

Note: This table shows the impact of years of overeducation on migration propensity. Personal demographical variables and provincial statistics in the college year are controlled. All specifications control the full set of fixed effects in column (8) Table 7. Robust standard errors are clustered at the age twelve province level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 13: Specification Checks

	Birth Migration (1)	-2 Years (2)	-1 Year (3)	+1 Year (4)	+2 Years (5)
College	0.034	0.031	0.027	0.012	0.006
	(0.048)	(0.036)	(0.039)	(0.037)	(0.042)
Observations	5,280	5,344	5,009	5,388	5,396

Note: This table shows the results of specification checks. In column (1), I use birth migration dummies as dependent variables. Column (2)-(5) randomly assign individuals to different provinces and move college year back 2 years and forward 2 years. All specifications control the personal and provincial characteristics as well as full set fixed effects. Robust standard errors are clustered at the age twelve province level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

9 Figures

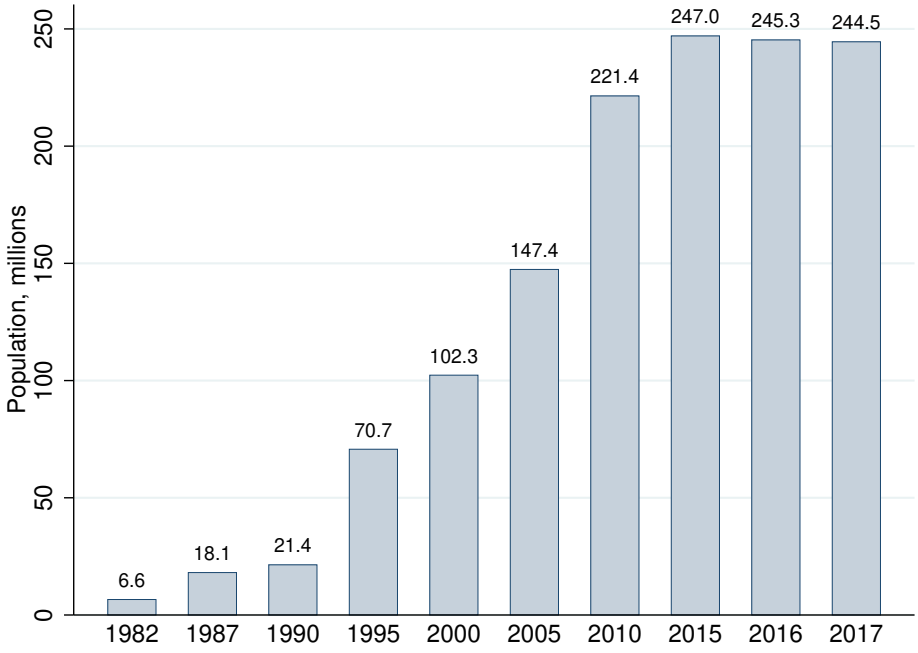

Figure 1: China's Migration Population, 1982-2017

Notes: This figure plots China's migration population from 1982 to 2017. Data is from the 2018 China Migration Population Development Report.

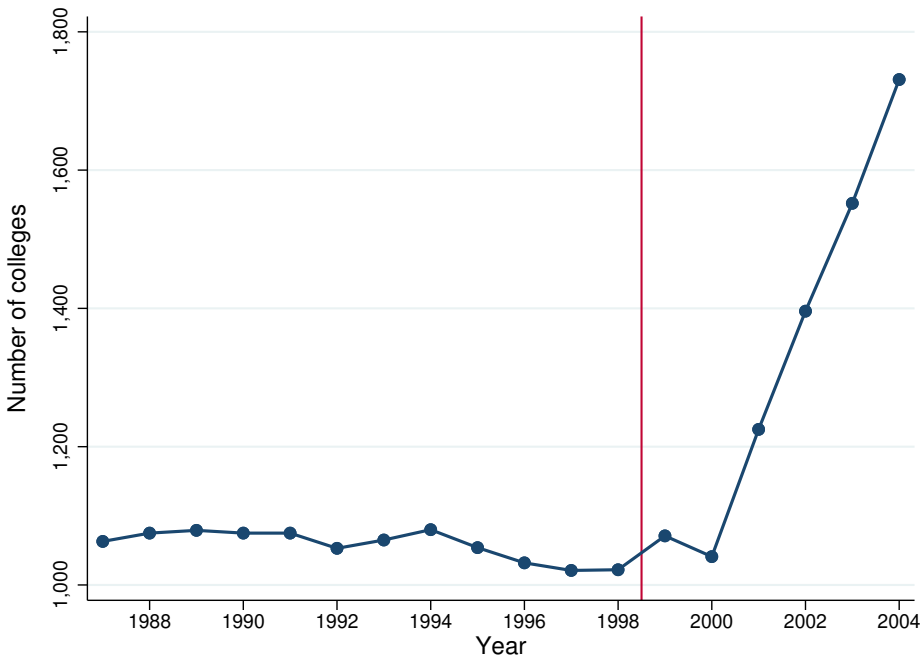

Figure 2: Trend of the Number of Colleges

Notes: This figure plots the number of regular higher institutions in China. Data is compiled from China Yearly Statistical Book.

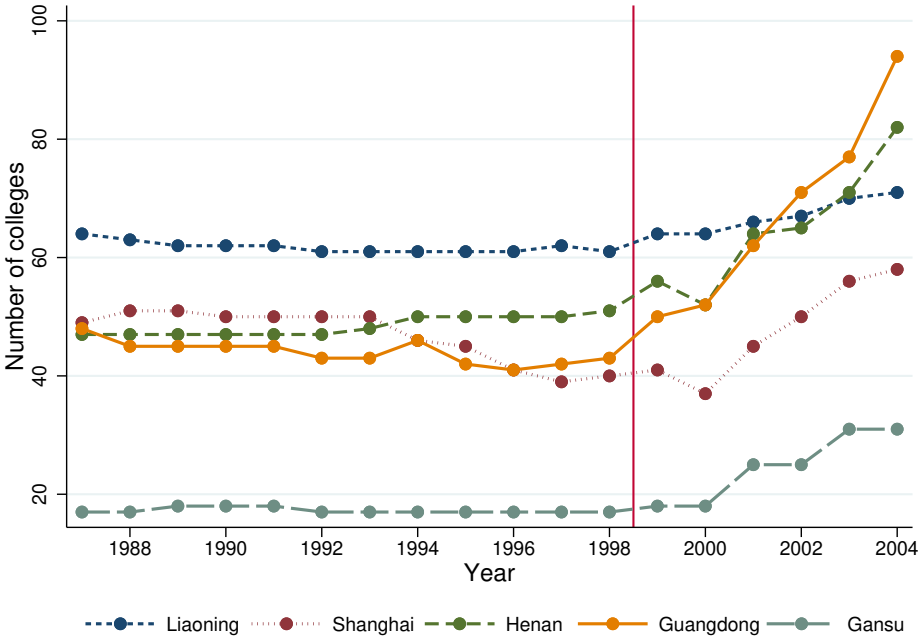

Figure 3: Variation in the Number of Colleges

Notes: This figure presents the number of colleges for selected provinces over time. Source: China Yearly Statistical Book.

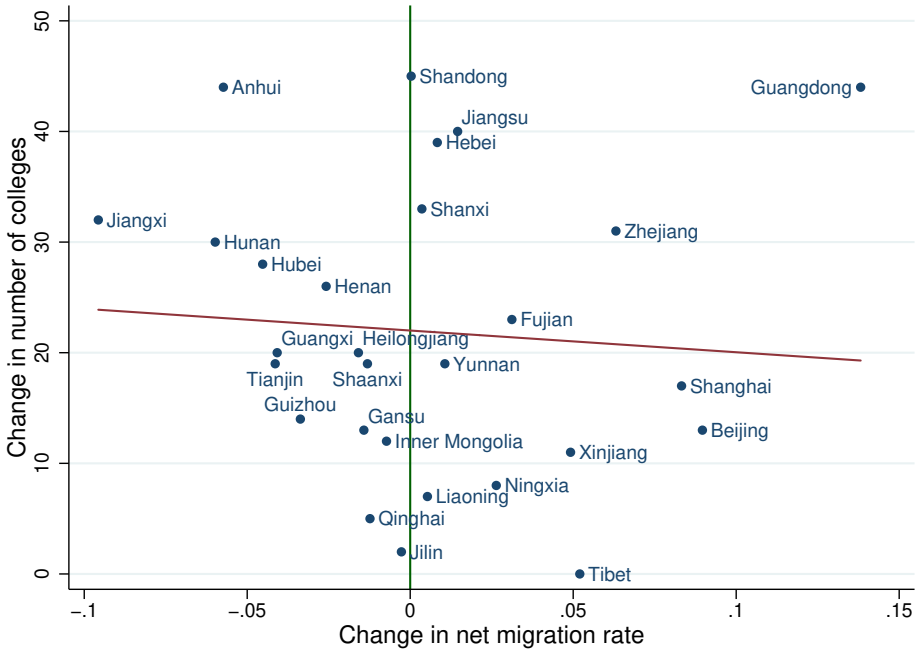

Figure 4: Pre-Trend Tests of Migration Rate Changes and Number of Colleges Changes

Notes: This figure presents the evidence that changes in the number of colleges at the province level are not targeting migration patterns. The net-migration rate is calculated by using in-migration rate subtracting out-migration rate. Out-migration is defined as people who live in a different province five years ago and based on the origin provinces. In-migration is similarly calculated but based on the current residence provinces. Data is obtained from censuses 1990 and 2000. Changes in the number of colleges are calculated as the difference between 2004 and 1999. Sichuan and Chongqing are dropped out because of the changed provincial boundary. Hainan is also dropped out because of no data. The fitted line has a slope of -19.63 with p-value 0.702.

Figure 5: The Number of Colleges and Hukou Reforms

Notes: This figure shows the scatter plot of the number of colleges and hukou reforms for province-year pairs between 1987 and 2004. Each dot represents a pair of the number of colleges and hukou reform indices. Reform indices are the means of city reform indices in a province, which are compiled by Fan (2019). I first run separate regressions of the two variables on a set of province and year fixed effects and obtain the residuals. Then I plot the residuals of the two variables against each other. The fitted line has a slope of -0.034 with a standard error 0.41. The p-value is 0.934 and R^2 is 0.

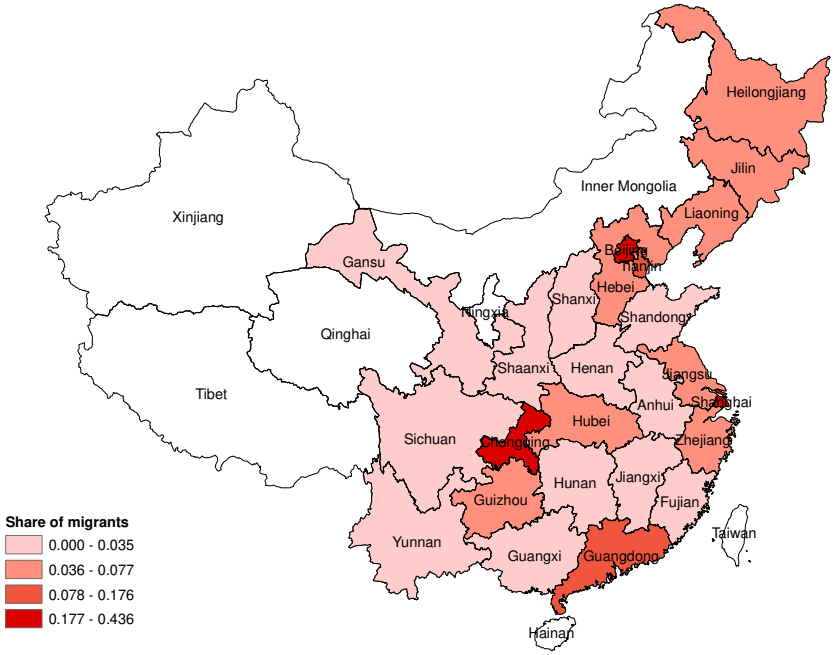

Figure 6: The Geography of Cross-Province Migration

Notes: This figure shows the map of cross-province migration in China based on the author's calculation of CFPS sample data. The sample consists people of age 25 to 40. The share is calculated as the number of people who are residing in the province that is different from the age twelve province, which is the same as the share of in-migrants for each province.

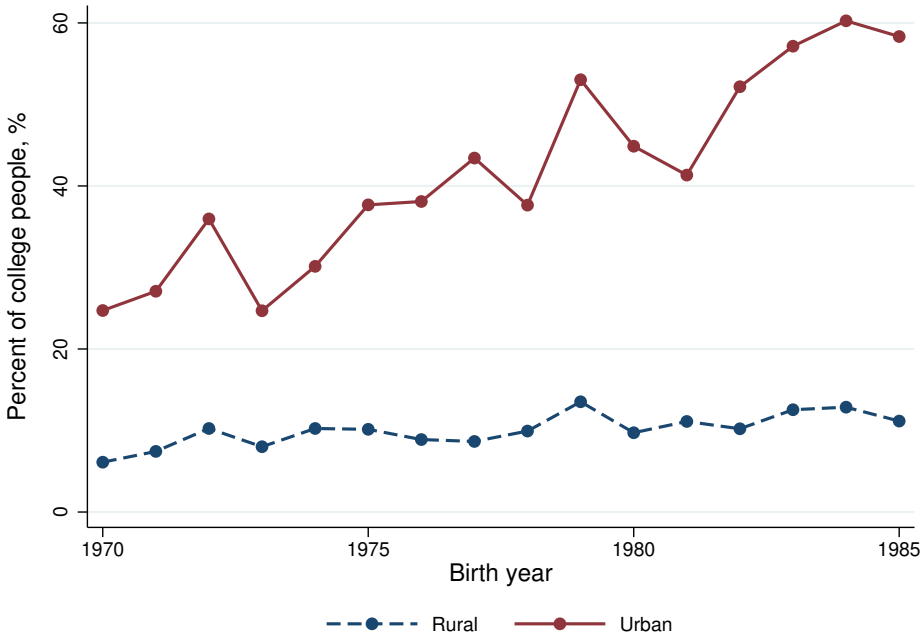

Figure 7: Share of College People by Rural/Urban hukou at Age 12

Notes: This figure shows the share of rural and urban college people in 2010 by their urban hukou status at age twelve using CFPS 2010 data.

Online Appendix
College Education and Internal Migration in China

A Additional Tables

Table A1: Results of Using Changes in the Number of Colleges as Exogenous Variation

	(1) Probit	(2) IV
Changes in the number of colleges	0.045***	
	(0.013)	
College		0.073*
		(0.041)
Male	0.037	-0.034***
	(0.040)	(0.009)
Han ethnicity	0.048	-0.002
	(0.086)	(0.035)
No. siblings	-0.142***	0.008***
	(0.028)	(0.002)
Father years of schooling	0.089***	0.000
	(0.008)	(0.001)
Urban hukou at age 3	0.739***	0.016
	(0.092)	(0.016)
Migration before age 12	0.104	-0.035**
	(0.103)	(0.015)
GDP growth rate	0.279	-0.123
	(1.118)	(0.117)
Urban employment	0.657	-0.028
	(0.426)	(0.031)
Population	1.402*	-0.005
	(0.770)	(0.013)
Observations	5,426	5,426

Note: This table shows the results of using changes in the number of colleges in the first-stage probit regression and the IV regression. Robust standard errors are clustered at the age twelve province level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A2: Effects of Education-Occupation Match on Migration by Rural and Urban

	(1) Rural	(2) Urban	(3) Rural	(4) Urban
College	-0.010	-0.016	-0.013	-0.052
	(0.029)	(0.042)	(0.033)	(0.043)
Overeducation	-0.005	-0.003	-0.002	0.002
	(0.003)	(0.008)	(0.003)	(0.008)
Overeducation×College	0.021*	0.019	0.013	0.022
	(0.012)	(0.012)	(0.012)	(0.014)
Male	-0.017	-0.019	-0.024**	-0.008
	(0.010)	(0.021)	(0.011)	(0.021)
Han ethnicity	0.007	-0.062	0.008	-0.065
	(0.033)	(0.064)	(0.032)	(0.063)
No. siblings	0.005	-0.010	0.007	-0.012
	(0.004)	(0.013)	(0.004)	(0.014)
Father years of schooling	0.001	0.006*	0.000	0.005
	(0.002)	(0.003)	(0.002)	(0.003)
Migration before age 12	0.001	-0.028	-0.006	-0.026
	(0.030)	(0.032)	(0.031)	(0.032)
Occupation controls			✓	✓
Observations	2,900	726	2,900	726

Note: This table shows the estimates of supplemental analysis of Table 12 column (3). I split the sample by hukou status at age 3. Robust standard errors are clustered at the age twelve province level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.