College Education, Enrollment Location, and the Geographical Mobility of Young Adults*

Shu Cai† *Jinan University*

Xiaozhou Ding‡ *Dickinson College*

Rui Du§ *Oklahoma State University*

Abstract

Understanding the determinants of young adults' migration patterns is crucial, and college education plays a central role in this context. However, the distinct impacts of college education and the location of college enrollment have not been sufficiently distinguished. Using the college seat expansion in China during the late 1990s as a natural experiment, combined with cross-province migration patterns derived from censuses, we show that both in-province and out-of-province college expansion increase college attendance rates. However, only out-of-province expansion increases the likelihood of attending college outside the home province. Our two-step instrumental variable approach reveals a strong effect of college enrollment location *per se* on migration, contrasted by a negative effect of college education *per se*. We explore how differences in sensitivity to distant job opportunities, perceptions of housing costs, and access to local social welfare among young adults can shed light on our findings.

Keywords: College education, enrollment location, migration, higher education expansion
JEL Classification Codes: I23, J61, O15, R23

*This paper has benefited from comments from Jingting Fan, Min Fang, Suqin Ge, Yilin Hou, Zibin Huang, Shumin Carl Lin, Le Wang, John Winters, Chunbin Xing, Junsen Zhang, and seminar and conference participants at the SEA Annual Meeting, AEA Annual Meeting, SRSA Annual Meeting, NARSC Annual Meeting, CES North America Conference, and Renmin University of China. Shu Cai acknowledges financial support from the China Natural Science Foundation (Project No. 72173056) and the Fundamental Research Funds for the Central Universities (Project No. 23JNQMX29). All remaining errors are our own.

†Institute for Economic and Social Research, Jinan University, 601 W Huangpu Da Dao, Tianhe District, Guangzhou, 510632, China. Email: shucai@jnu.edu.cn.

‡Department of Economics, Dickinson College, Althouse Hall 112, 28 N College St, Carlisle, PA 17013, USA. Email: dingx@dickinson.edu.

§Department of Economics, Oklahoma State University, 347 Business Building, Stillwater, OK 74078, USA. Email: rui.du@okstate.edu.

1 Introduction

The mobility of human capital plays a crucial role in explaining regional disparities in economic development. Regions that successfully attract highly educated individuals often prosper as these individuals contribute to increases in local productivity, innovation, knowledge dissemination, and skill enhancement (e.g., Glaeser, Scheinkman, and Shleifer, 1995; Moretti, 2004, 2011; Bosetti, Cattaneo, and Verdolini, 2015; Ma, 2023). Conversely, regions losing such talents often face the issue of brain drain (Beine, Docquier, and Rapoport, 2001, 2008).

Given this, both national and local policymakers have a keen interest in the factors influencing the location choices of young college graduates (Groen, 2011). The mobility of these graduates has significant implications for aggregate productivity, labor market dynamics, human capital accumulation, innovation spillovers, and the spatial distribution of skills and economic activities (Au and Henderson, 2006; Li, Whalley, and Xing, 2014; Bryan and Morten, 2019; Tombe and Zhu, 2019; Howard, 2020; Kong, Zhang, and Zhang, 2020; Rong and Wu, 2020; Diamond and Gaubert, 2022; Fabre, 2023).

Chinese universities primarily rely on government funding.[1] Recognizing the importance of a local supply of college-educated individuals, local governments frequently enact various higher education policies, including financial support, subsidies for public institutions, and investment in infrastructure and resources, all designed to promote college attendance and retain college-educated talents within their regions.

In addition to local efforts, nationwide centralized policies are in place to address regional disparities in local human capital. A notable example is China's province-based college admission quota rationing system (Yang, 2021). These policies operate on the underlying assumption that graduates will likely reside in the location where they received their college education.

Indeed, an extensive literature has established a robust link between attending college in a specific location and residing there afterward (Bound et al., 2004; Groen, 2004; Abel and Deitz, 2012; Fitzpatrick and Jones, 2016; Winters, 2020; Huang, Xing, and Cui, 2022). This implies that the geographical distribution of educational institutions and the spatial disparity in educational opportunities can influence the mobility and educational demand among the youth (Kelchtermans and Verboven, 2010; Gibbons and Vignoles, 2012; Fu et al., 2022; Ishimaru, 2023).

Such dynamics can intensify spatial inequality in human capital. Regions with scarce opportunities may witness a talent drain as graduates move to areas with better educational and career prospects (Fabre, 2023). These regions, grappling with the challenge of attracting and retaining skilled individuals, face increasing hurdles in fostering economic growth and development.

[1] According to the China Educational Finance Statistical Yearbook, Chinese colleges at all tiers rely on governments, including central, provincial, and city governments, for more than half of their financing.

On the other hand, it is important to note that the pronounced migration effect of college enrollment location could be counterbalanced by the broader influence of higher education itself. Unlike earlier schooling with localized effects, higher education equips individuals with transferable skills and knowledge, opening doors to broader labor markets. This expanded range of opportunities and the resultant higher mobility of college graduates can potentially outweigh the location effect (Wozniak, 2010; Machin, Pelkonen, and Savanes, 2012; Malamud and Wozniak, 2012; Grogger and Hanson, 2015; Weiss, 2015; Kerr et al., 2017; Ding, 2021).

The migration-enhancing effect of college education *per se* underscores the role of general human capital, contrasting with the concept of "learning-by-doing" through distant college attendance (Malamud and Wozniak, 2012). As a result, the positive influence of college education on migration could potentially neutralize the benefits for regions that make substantial investments in educating their young population.

The intricate and nuanced relationship between college education and enrollment location has sparked significant interest in understanding their relative contributions to young individuals' migration decisions later in life. This distinction is important, not only for differentiating between those who pursue college education for location preference and those for geographical labor market mobility but also for informing higher education policies aimed at fostering and retaining high-skilled workers through financial support for students or subsidies for public institutions.

For example, if the migration choices of college-educated individuals are primarily driven by the location of their college enrollment, rather than the direct impact of the education itself (the general human capital channel), it could justify local government investments in public colleges and could also support the use of nationwide college seat quotas to alleviate regional disparities in human capital.

Despite its importance, empirical research that effectively disentangles these two effects remains limited. Existing studies typically estimate a combined effect of college education and enrollment location (Machin, Pelkonen, and Savanes, 2012; Malamud and Wozniak, 2012; Ding, 2021).[2] Our paper seeks to fill this critical gap by independently establishing credible causal estimates for the impact of college education *per se* and the impact of college enrollment location *per se* on young adults' migration decisions.

An immediate challenge in this effort is effectively tracking individual mobility. To address this, we make a unique contribution by utilizing data from China's population census to identify individuals' mobility patterns during the early stages of their careers. Through careful analysis of

[2] Machin, Pelkonen, and Savanes (2012) examine the impact of years of schooling on migration using the Norwegian compulsory education reform as an identification strategy; Malamud and Wozniak (2012) investigate the effect of college education on out-of-state migration among U.S. men, leveraging the Vietnam draft lottery as exogenous variation; Ding (2021) explores the relationship between college attendance and cross-province migration among Chinese young adults.

China's 2010 population census data, we managed to ascertain three key locations for a college-educated individual: birthplace, college enrollment location, and current location of residence.

Moreover, identifying the causal effects is challenging due to potential sorting bias and the intertwined nature of education and location preferences. First, individuals with higher unobserved abilities are more likely to seek higher education and respond to distant labor market opportunities. This creates a misleading relationship between college education and the propensity for migration. Second, those attending non-local colleges may have less attachment to their hometowns, leading to a greater inclination to seek job opportunities elsewhere. Third, unobserved social welfare and transfer payments may disproportionately compensate individuals based on their education level, influencing their migration decisions (Notowidigdo, 2020). Furthermore, college enrollment decisions may not be made in isolation but may factor in preferences for location of residence.

To address endogeneity issues, studies have leveraged natural experiment shocks such as foreign remittances, military drafts, school reforms, and more. These have helped isolate exogenous variations in skill acquisition (Malamud and Wozniak, 2012; Aydemir, Kirdar, and Torun, 2021; Fan and Li, 2023; Allende, Luksic, and Navarrete, 2024) and individual mobility (Edwards and Ureta, 2003; Yang, 2008; Howell, 2023, 2024).

Our analysis, however, focuses on a different aspect: the college seat supply mechanism. We exploit the variation in province-level college enrollment, a result of China's massive higher education expansion in the late 1990s. This expansion led to the annual admission of millions of students and provided a plausibly exogenous shock to individual educational decisions.

Given China's province-based college seat quota and admission system, we use province-level college enrollment and the initial distribution of admitted students across provinces to construct indicators that capture college expansion shocks at the provincial level.[3] These constructed regional college seat supply shocks created rich spatiotemporal variations in college enrollment that allow us to identify the effect on the migration decisions of young adults.

To establish causality, we employ a two-step instrumental variable (IV) approach. In the first step, we use our constructed college seat supply shocks to predict individual propensities of attending local colleges and attending college out-of-province. We find that both in- and out-of-province college seat expansion positively influence an individual's likelihood of college attendance. However, only the expansion of out-of-province college seats leads to a higher propensity for individuals to attend colleges outside their home province.

In the second step, we use the predicted probabilities of college education and out-of-province college attendance as instruments for individuals' actual decisions regarding college attendance

[3]China's administrative divisions are organized into provinces, prefectures, counties, townships, and villages, with provinces being the focus of this study due to data availability. For comparison, a province in China can be loosely equated to a state in the U.S.

and out-of-province college enrollment. Our identification strategy hinges on a key assumption that province-level college seat supply shocks influence individuals' later-life migration solely through their effects on individuals' decisions regarding whether and where to attend college. This assumption is unlikely to be violated in our research design because the unanticipated macrogeographic-level college seat supply shocks, driven by a nationwide college expansion policy, are effectively exogenous to unobserved determinants of individuals' residential preferences.

Our IV estimation results reveal that attending a college outside one's province is a strong predictor of future migration patterns. However, once we isolate the effect of distant college attendance, we find that college education itself negatively impacts migration propensities. This finding contradicts existing literature that reports higher mobility among more educated individuals.

Our heterogeneity analysis suggests that the greater mobility among non-college-educated individuals, especially males, could be attributed to their perception of larger regional wage differentials. These individuals also seem to face less mobility friction. This is likely due to their lower sensitivity to housing costs during migration, a consequence of a culture of migrant remittances and a preference for purchasing homes in their hometowns. We show that regional differences in the effect of college enrollment location appear to be explained by spatial differences in wages, housing costs, and *hukou* restrictions, along with the social welfare tied to them.[4]

Overall, our work breaks new ground by disentangling the distinct effects of college education and enrollment location, revealing the underlying mechanisms driving postcollege location choices of the youth. This offers valuable insights for policymakers and sheds light on labor market dynamics, regional development, and talent retention strategies.

The paper proceeds as follows. Section 2 provides an overview of the institutional background. In Section 3, we present the data utilized in this study, detailing its sources and characteristics. Section 4 introduces our empirical strategy. Section 5 presents the estimation results. Section 6 discusses the results of the heterogeneity analysis. Section 7 concludes.

2 Background

2.1 The Great Migration and Spatial Skill Sorting

The economic transition in China has had a significant impact on migration patterns. Figure 1 plots the country's migration population from 1982 to 2017, based on data obtained from the *2018*

[4]The *hukou* system is a household registration system implemented in mainland China. It officially designates an individual as a permanent resident of a specific area. A household registration record contains essential identifying information such as the person's name, date of birth, and place of permanent resident. See Section 6.3 below for more details. For comprehensive reviews of the *Hukou* system in China, refer to Cheng and Selden (1994), Chan (2015), and Young (2013).

China Migration Population Development Report. Over the past four decades, China has witnessed a remarkable surge in population mobility. In 1982, the number of migrants stood at a mere 6.6 million, but by 2010, it had surpassed 200 million. Although the upward trend peaked in 2015, it has remained relatively stable since then. This period has been characterized by a substantial flow of migration from rural areas to more developed urban regions or from the interior to the coast.

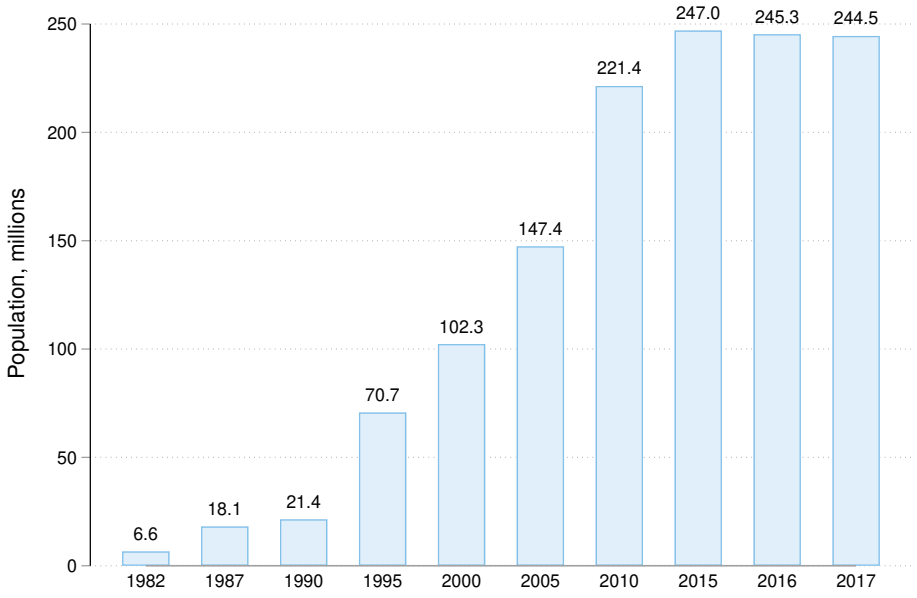

Figure 1: China's migration population, 1982-2017

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

China's significant migration phenomenon can be traced back to several factors. A key driver is the shift from agriculture to manufacturing, a trend fueled by increased agricultural productivity due to the "household responsibility system" reforms initiated in 1979. This shift was further propelled by the emergence of Township and Village Enterprises (TVEs) in the mid-1980s, which created a heightened demand for urban labor.[5] China's subsequent accession to the World Trade Organization (WTO) in 2001 and the massive urbanization in the following decades sparked a boom in the manufacturing and construction sectors, intensifying migration flows.

Remarkably, the surge in mobility within China has predominantly involved young adults starting their careers. Data from the 2010 population census documents a migration of approximately 1.4 million individuals aged 18 to 24, making up around 22% of this demographic's total population. This massive migration could be influenced by the dynamic interplay of evolving preferences among skill groups and shifts in location-specific attributes such as wages, rents,

[5]TVEs in China are market-oriented public enterprises administered by local governments and primarily located in townships and villages. Li (1993) provides a comprehensive examination of TVEs in China.

amenities, and the extent of mobility restrictions (Diamond and Gaubert, 2022). Moreover, on the supply side, the geography of higher education and nationwide college expansion have played significant roles in driving the spatial sorting of young adults.

2.2 Higher Education Expansion and College Seat Quotas

The higher education sector in China was small and recovering from the Cultural Revolution when the "reform and opening up" era began in 1978. Between 1978 and 1999, college admissions increased modestly from 0.4 million to 1 million.

To alleviate the escalating unemployment situation due to the state-owned enterprise reform and meet the demand for skill, China implemented a large-scale expansion of higher education in the late 1990s, commonly referred to as college expansion.[6] To illustrate its impact, we present the annual number of new college students and the college enrollment rate from 1990 to 2012 in Figure 2, based on data compiled in the *China Statistical Yearbooks*.[7] The red curve in the figure represents the yearly new college enrollment in China. The blue curve illustrates the national college enrollment rate.

Before the expansion of higher education in 1999, national college enrollment in China remained relatively stable, at around one million. However, the number of newly enrolled college students surged by over 40% between 1998 and 1999, when the country started the college expansion (Li, Whalley, and Xing, 2014). This trend of expansion continued in the years that followed, culminating in an enrollment figure of nearly 7 million students by 2012. Moreover, we observed a distinct discontinuity in the enrollment rate in 1999. During the pre-expansion period, the enrollment rate fluctuated around 30%. However, following the college expansion, the enrollment rate experienced a pronounced acceleration, reaching approximately 80% immediately after the expansion.

Since the resumption of the National Higher Education Entrance Examination, or *Gaokao*, in 1977, China has adopted a province-based college admission system. This system, managed by the Ministry of Education in partnership with local governments and colleges (Bollinger, Ding, and Lugauer, 2022; Feng and Xia, 2022), utilizes both sequential (immediate acceptance) and parallel admission mechanisms, with variations across provinces (Chen, Jiang, and Kesten, 2020). Each receiving province allots annual quotas to the sending provinces, determining the number of students to be enrolled. In this system, college candidates compete with their peers within their respective provinces of residence for college seats.

The significant growth of higher education, coupled with a province-based admission system,

[6] In the remainder of this paper, the terms "higher education expansion," "college seat expansion," and "college expansion" will be used interchangeably.

[7] The enrollment rate is defined as the ratio of new college enrollments to the number of high school graduates.

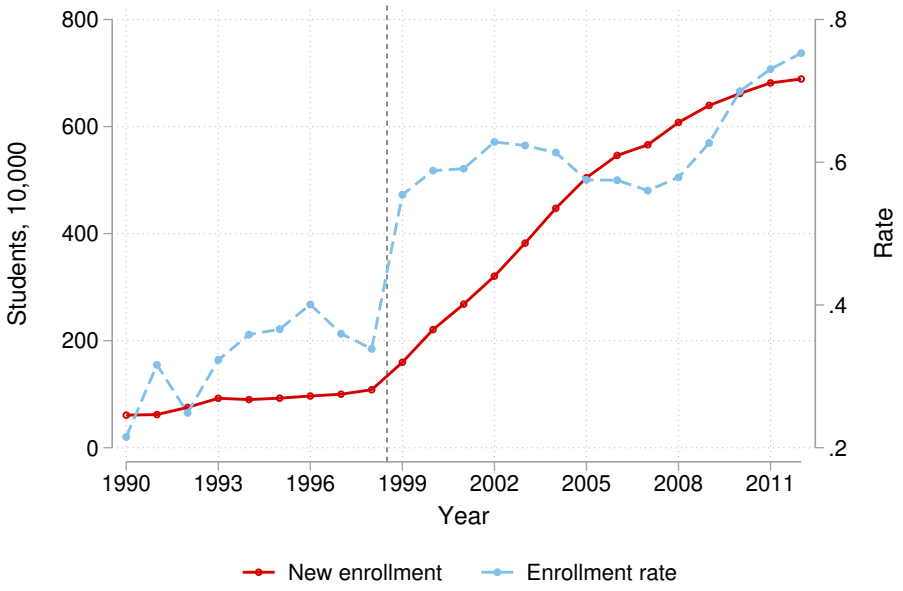

Figure 2: The expansion of higher education in China

Notes: This figure presents the trends in national college enrollment and the corresponding enrollment rate. The enrollment rate is derived by dividing the number of college enrollments by the number of high school graduates. The data used for this analysis is sourced from the China Educational Statistical Yearbooks and the National Bureau of Statistics.

brings about rich spatiotemporal variations in college enrollment. This scenario provides a compelling case for our study.

3 Data and Measures

3.1 Population Census

The main data for our empirical analysis is derived from the 2010 decennial census, which provides a comprehensive count of residents across each province, capturing their current city and province of residence, as well as their province of residence five years before the survey date.

Moreover, the censuses provide information on individuals' birthplaces. This allows us to ascertain if an individual was born in their current city of residence, a different city within the same province, or a different province. Therefore, our study primarily focuses on the cross-province migration patterns of young adults. We classify an individual as a migrant if they currently reside in a province different from their birth province and have lived there for at least six months.

The censuses also include individual-level characteristics such as age, gender, marital status, and ethnicity. Importantly, they provide information on individuals' completed level of education, allowing us to identify different skill groups within the population. We focus on respondents in

the 2010 census with at least a high school education, as their education and migration decisions are likely influenced by college expansion.

3.2 College Enrollment Location

Measuring the locations of college attendance is vital for our analysis; however, this information is not directly available in the census data. Therefore, identifying these locations requires making reasonable assumptions based on several key pieces of information in the survey.

First, our focus is on individuals who have achieved a college degree, including both four-year colleges and vocational colleges that typically span 3 years (Wu and Ye, 2018). We operate under the assumption that these individuals start their college education at 18, aligning with the standard college entry age in China's education system (Wang, 2010).

Second, a typical college student begins their freshman year in September and completes their senior year in June, four years later. To account for variations, individuals born before September 1st start school one year earlier, so we estimate their year of college attendance as the birth year plus eighteen, while for those born after September 1st, it is the birth year plus nineteen.

Third, China's census typically commences on November 1st of the census year. This timing allows us to gather information about individuals' places of residence at two specific time points: October 2005 and October 2010.

Based on the available information and necessary assumptions, we focus on cohorts who started college between 2003 and 2005, as their college attendance locations can be inferred. These cohorts began their studies between September 2003 and September 2005 and completed them between June 2006 and June 2009. By leveraging their reported locations of residence in October 2005, as captured in the 2010 census data, we can infer where they attended college. We excluded the year 2002 from our analysis, given that some vocational college students may conclude their studies within three years.

Determining the locations of college attendance is challenging for cohorts who attended college before 2002 or after 2005, leading us to exclude these cohorts from our analysis. For instance, the cohort that started college in 2001 finished their studies in June 2005, potentially altering their residence location before October 2005 – the time point at which we know their residence. Consequently, accurately determining their college attendance locations poses difficulties. Similarly, for the cohort that began college in September 2006, they completed their studies in June 2010, and the recorded "current residence location" in October 2010 may differ from their college enrollment location due to possible post-graduation changes in their residence.

Based on this sample, we identify those who pursued higher education outside their birth province by contrasting their birth province with their college attendance province, as determined

from the survey.

It is important to note that our approach to identifying college enrollment locations yields a sample of young individuals in the early stages of their careers. This is a dynamic period marked by high labor market mobility (Topel and Ward, 1992) and marriage market sorting (Fan and Zou, 2021), both of which contribute to the spatial distribution of skilled workers.

3.3 City-Level Characteristics

Although China's higher education admission system is mainly organized at the provincial level, there is significant heterogeneity among cities within each province, which can influence post-college migration choices. We match individuals with the characteristics of their birth city following a procedure elaborated in the subsequent section. These characteristics are sourced from the *China City Statistic Yearbooks*. Specifically, we collected data on city-level population, GDP per capita, sectoral composition, and the city's shortest distance to the coast. Furthermore, we collected data on urban employment and GDP per capita at the province level and computed their respective growth rates.

3.4 Summary Statistics

We compile a dataset for our baseline analysis by aligning each individual with their birthplace's city-level and province-level attributes. We use a specific procedure to determine an individual's birth city based on their birthplace and place of *hukou* registration. The procedure considers three scenarios based on the survey question about one's birthplace, including individuals born in their current county of residence, a different county within their current province of residence, or a different province. Depending on the scenario and the individual's *hukou* registration, we managed to identify the birth city for 63,498 individuals in our sample, constituting 18.9 percent of the census population aged between 23 and 26. The details of the sample construction procedure and data assembling are available in Section A.1 in the appendix.

Table 1 presents descriptive statistics for the key variables employed in our empirical analysis where we group individual-level characteristics in Panel A and regional characteristics in Panel B. We categorize individuals into three groups: those without a college education ($College = 0$), those with a college education ($College = 1$), and individuals who pursued their college education outside their birth province ($CollegeOut = 1$). It is important to note that while the first two groups are mutually exclusive, the third group is a subsample of the second group. Migrants constitute a comparable proportion among college-educated individuals (around 8.9%) and non-college-educated individuals (about 12.2%). However, a significant majority (91%) of those who attended college out-of-province migrated post-education, offering initial insights into the

influence of college enrollment location on subsequent migration patterns.

Moreover, college-educated individuals exhibit a roughly balanced proportion of males and females overall. However, among college graduates who pursued education outside their province and non-college-educated, a higher proportion of males is evident. Ethnicity and age remain relatively consistent across the groups, though the non-college-educated cohort has the largest average household size. Notably, disparities emerge in the characteristics of the birth cities and provinces of individuals. Non-college-educated individuals tend to originate from smaller, economically less developed cities compared to their college-educated counterparts. Conversely, those who attended college out-of-province tend to come from regions with relatively lower economic development levels.

Overall, these observed descriptive differences underscore the potential presence of selection bias, stemming from the intricate interplay of education, migration, and location attributes. Failing to account for these factors could compromise the accuracy of our estimated effects of interest.

4 Empirical Strategy

4.1 Baseline Model

Our baseline estimation equation aims to establish a relationship between individual-level migration outcomes documented in the 2010 census and decisions made regarding college attendance and enrollment location. Specifically, we consider variations of the following main specification:

$$M_{ijt} = \beta_0 + \beta_1 College_{ijt} + \beta_2 CollegeOut_{ijt} + \beta_3 \mathbf{X}_{ijs} + \phi_j + \phi_t + \varepsilon_{ijt}, \tag{1}$$

where M_{ijt} is a binary indicator variable representing the migration status of individual i born in province j and year t, which turns on if individual i's current province of residence is not their birth province j. The variables $College_{ijt}$ and $CollegeOut_{ijt}$ are dummy variables indicating whether individual i attended college (including both in- and out-of-province college attendance) and whether their college education was obtained out-of-province, respectively. $\mathbf{X}_{ijs}$ denotes a vector of pre-determined characteristics about individual i and their birth province j in their college freshman year s if they were to attend college (i.e., the age of 18 as per our assumption). Additionally, ϕ_j and ϕ_t represent birth-province fixed effects and birth-year cohort fixed effects. ε_{ijt} is the error term. We cluster the standard errors by province to account for arbitrary within-province autocorrelation. Our regressions are weighted using the number of high school graduates at the provincial level.

Our main parameters of interest are β_1 and β_2. Specifically, β_1 captures the effect of college education on the probability of migration, while β_2 captures the effect of college enrollment

Table 1: Summary statistics

	(1) $College$=0	(2) $College$=1	(3) $CollegeOut$=1
Panel A. Individual Characteristics			
Migrant	0.122	0.089	0.910
	(0.327)	(0.284)	(0.286)
$College$	0.000	1.000	1.000
	(0.000)	(0.000)	(0.000)
$CollegeOut$	0.000	0.091	1.000
	(0.000)	(0.288)	(0.000)
Male	0.541	0.484	0.562
	(0.498)	(0.500)	(0.496)
Han ethnicity	0.964	0.963	0.954
	(0.185)	(0.189)	(0.210)
Age	24.081	24.135	24.142
	(0.915)	(0.910)	(0.919)
Household size	4.221	3.699	3.964
	(1.956)	(1.805)	(1.600)
Panel B. City and Province Characteristics			
City population (million)	6.203	6.339	6.120
	(4.791)	(4.565)	(5.149)
City GDP per capita (10,000)	1.469	1.884	1.195
	(1.172)	(1.441)	(0.881)
City tertiary employment share (%)	36.695	38.179	36.442
	(7.503)	(8.326)	(6.535)
City distance to coast (100km)	3.964	3.435	4.425
	(3.434)	(3.392)	(3.338)
Province urban employment (million)	0.176	0.193	0.165
	(0.067)	(0.083)	(0.054)
Province GDP per capita (10,000)	1.405	1.603	1.197
	(0.765)	(0.962)	(0.585)
Provincial capital city	0.127	0.174	0.106
	(0.333)	(0.379)	(0.308)
Municipality	0.060	0.101	0.043
	(0.237)	(0.302)	(0.203)
Province GDP growth rate	0.172	0.171	0.169
	(0.038)	(0.038)	(0.039)
Province employment growth rate	0.055	0.053	0.049
	(0.048)	(0.051)	(0.041)
Observations	32,118	31,380	2,858

Note: This table presents the summary statistics for the main sample. Standard deviations are in parentheses.

location on postcollege migration choice.

4.2 Endogeneity Concerns

Both $College_{ijt}$ and $CollegeOut_{ijt}$ are likely to suffer from endogeneity issues since individuals often make simultaneous decisions regarding education and migration. We first incorporate birth-province fixed effects into our model. These fixed effects account for unobserved time-invariant disparities across provinces that could influence both migration and education choices, such as local geography and natural advantages. By employing birth-province fixed effects, we estimate the model based on individual deviations from province-specific means over time, rather than relying on cross-province variations in individual education decisions.

Furthermore, individuals born in different years may encounter time-specific shocks that affect both their migration and education decisions. Such shocks can arise from various sources, including targeted policies aimed at attracting college graduates to specific locations, education policy reforms, or macroeconomic fluctuations. To address the potential influence of these time-specific shocks, we incorporate birth-year cohort fixed effects into our analysis. The inclusion of these fixed effects allows us to isolate the variations in individual outcomes relative to those born in the same year.

We include additional controls such as ethnicity, gender (Ge, Isaac, and Miller, 2019), household size, and age, along with the squared term of age (Schwartz, 1976) to account for individual differences that may influence both migration and education decisions.[8]

To account for demand-side mechanisms, such as the economic conditions and the trade-based channels examined by Fan and Li (2023), we incorporate birth-city-level controls such as population, GDP per capita, employment composition, and the city's proximity to the coast. Moreover, we control for time-varying characteristics of the birth province to capture macrogeographic-level confounding factors. These include variables such as urban employment and its growth rate, GDP per capita and its growth rate at the province level, and a dummy variable indicating whether an individual was born in a provincial capital or a direct-administered municipality.[9]

While our model incorporates a conservative array of fixed effects and controls, concerns remain about potential bias in the OLS estimates. Such bias could arise from unobserved individual heterogeneity affecting both migration and educational choices.

A notable source of bias might be individuals with elevated innate abilities, as they are more inclined to pursue a college education, especially in prominent colleges located outside their birth

[8]Defining cohorts by academic year captures age to some extent, but some individual age variations within cohorts remain unaccounted for. Given this, age is not entirely absorbed by the cohort fixed effects.

[9]A direct-administered municipality (*Zhixiashi*), commonly known as a municipality, is one of four types of province-level divisions in China. It represents the highest level of city classification in China and includes cities such as Beijing, Shanghai, Tianjin, and Chongqing during our study period.

provinces if local options are limited. Those with heightened labor market mobility, often tied to higher abilities, face fewer constraints in seeking job opportunities beyond their birth provinces. In contrast, individuals deeply attached to their hometowns tend to opt for local college attendance and exhibit reduced migration tendencies.

Moreover, our estimation's accuracy could be affected by measurement errors, which stem from our methods of pinpointing college enrollment locations and the assumptions used to deduce this information.

Thus, mitigating potential bias requires employing suitable instruments to isolate exogenous variations in college education decisions that are uncorrelated with unobserved factors driving migration choices.

4.3 A Shift-Share Design: Regional College Expansion Shocks

Following the literature that exploits college expansion as a natural experiment (e.g., Feng and Xia, 2022), we leverage policy-driven variations in the changes of college enrollment quotas across provinces to produce plausibly exogenous variations in individual choices about college attendance and college enrollment location. We begin by calculating the share of college students from each origin province j who enrolled in destination province k in the year 2000.[10] This computation is based on population census data and can be achieved using the following equation:

$$s_{jk,t_0} = \frac{N_{jk,t_0}}{N_{k,t_0}} \tag{2}$$

where t_0 is the initial period, specifically the year 2000, which follows the significant expansion of higher education in 1999. N_{k,t_0} denotes the total number of college enrollments in destination province k during the initial period, while N_{jk,t_0} represents the total number of college enrollments from origin province j to destination province k within the same initial period. This equation allows us to derive a comprehensive set of origin-destination pairs containing the shares s_{jk,t_0}.

Subsequently, we employ the shares s_{jk,t_0} to predict the intensity of in- and out-of-province college seat supply shocks. To achieve this, we multiply the initial share of students from the origin province j in each destination province k by the annual enrollment of destination province k in year t, denoted E_{kt}. By doing so, we capture the scale of college expansion in destination province k that influenced the rising college freshmen from the original province j in year t had the distribution of college seats remained constant since the massive expansion of higher education in 1999.

To account for the variations in incoming college student populations across different provinces,

[10] Our analysis utilizes the year 2000 as the initial period. This is because the 2000 census is the earliest survey providing detailed birthplace information.

we net out scale effects by dividing the predicted college expansion intensities faced by the origin province j in year t by the number of high school graduates in province j during the same year. Specifically, the aforementioned procedures can be expressed using the following equations:

$$IR_{jt} = \frac{E_{jt} \times s_{jj,t_0}}{H_{jt}}, \tag{3}$$

$$OR_{jt} = \frac{\sum_{k \neq j} E_{kt} \times s_{jk,t_0}}{H_{jt}}, \tag{4}$$

where IR_{jt} and OR_{jt} are in- and out-of-province college enrollment rates for origin province j in year t, respectively. These predicted enrollment rates essentially capture the in- and out-of-province college expansion intensity faced by the rising college freshmen in origin province j in year t. E_{kt} is the college new enrollment of destination province k in year t. H_{jt} is the number of high school graduates of origin province j in year t. s_{jk,t_0} is the share of college students from origin province j who enrolled in destination province k in the year 2000.

We obtained the province-level college enrollment data and the number of high school graduates from the *Educational Statistics Yearbook of China*. Our analysis focuses on the period from 2003 to 2005, for which we can infer individuals' college enrollment locations from the information reported in the census.

To illustrate the spatial variation in exposure to province-level college expansion shocks, we present variations in IR_{jt} (predicted in-province enrollment rate) and OR_{jt} (predicted out-of-province enrollment rate) for the year 2005 across each province in figures 3 and 4.

Notably, Figure 3 documents substantial heterogeneity in the in-province college expansion across provinces. Regions such as Tibet, the northeastern provinces, and the southeastern coastal areas had notably higher exposure to in-province college expansion compared to other provinces.

In contrast, Figure 4 shows that peripheral provinces, particularly those with geographical constraints like being landlocked, had significantly higher exposure to out-of-province college expansion. These spatial discrepancies in college expansion shocks are expected to influence diverse college education choices and subsequent post-college location decisions.

4.4 A Two-Step IV Approach

We harness plausibly exogenous fluctuations in province-level college expansions to separate the variations in individual decisions about college attendance and enrollment location from unobserved migration preference determinants. Specifically, we adopt a two-step IV approach.

Our first step is to estimate the probabilities of college attendance and out-of-province college enrollment using probit regressions. Specifically, we estimate the following regressions that relate

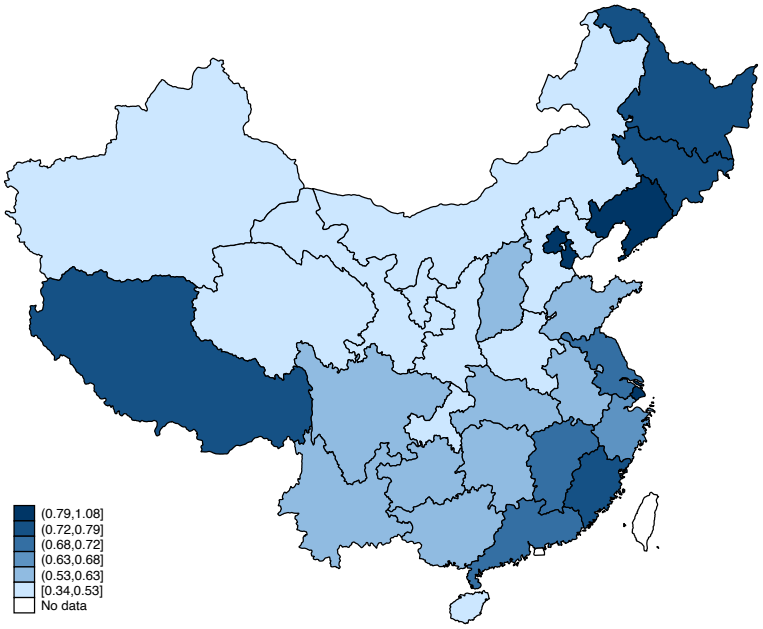

Figure 3: In-province college enrollment rate, 2005

Notes: This figure shows the in-province enrollment rate for each province in the year 2005. The enrollment rates are calculated by dividing the predicted in-province enrollment by the number of high school graduates in 2005, following Equation (3): $IR^{2005} = \frac{E_j^{2005} \times s_{jj,2000}}{H_j^{2005}}$.

binary outcome variables indicating college attendance and out-of-province college enrollment to intensity levels of in- and out-of-province college expansion:

$$College_{ijt} = \Phi(\alpha_0 + \alpha_1 IR_{it} + \alpha_2 OR_{it} + \alpha_3 \mathbf{X}_{is} + \eta_j + \eta_t + u_{ijt}), \tag{5}$$

$$CollegeOut_{ijt} = \Phi(\theta_0 + \theta_1 IR_{it} + \theta_2 OR_{it} + \theta_3 \mathbf{X}_{is} + \gamma_j + \gamma_t + v_{ijt}), \tag{6}$$

where we incorporate birth-province fixed effects (η_j and γ_j) and birth-year cohort fixed effects (η_t and γ_t) into the models. The remaining variables are defined similarly as those in equation (1).

In the second phase, we employ a two-stage least squares (2SLS) methodology to estimate equation (1), where we use the predicted probabilities $\widehat{College}$ and $\widehat{CollegeOut}$ obtained by estimating equations (5) and (6) as instruments for actual decisions concerning college attendance ($College$) and out-of-province college enrollment ($CollegeOut$), allowing for a more nuanced exploration of individual-level variation in contrast to province-level college expansion shocks. This refined approach fosters more precise identification of the causal effects of interest (Wooldridge, 2010).

Our identification approach hinges on the assumption that macrogeographic-level college expansion shocks are uncorrelated with unobserved determinants of individual location preferences and thus exclusively influence individual post-education migration decisions through their impact on concrete choices related to college attendance and enrollment location.

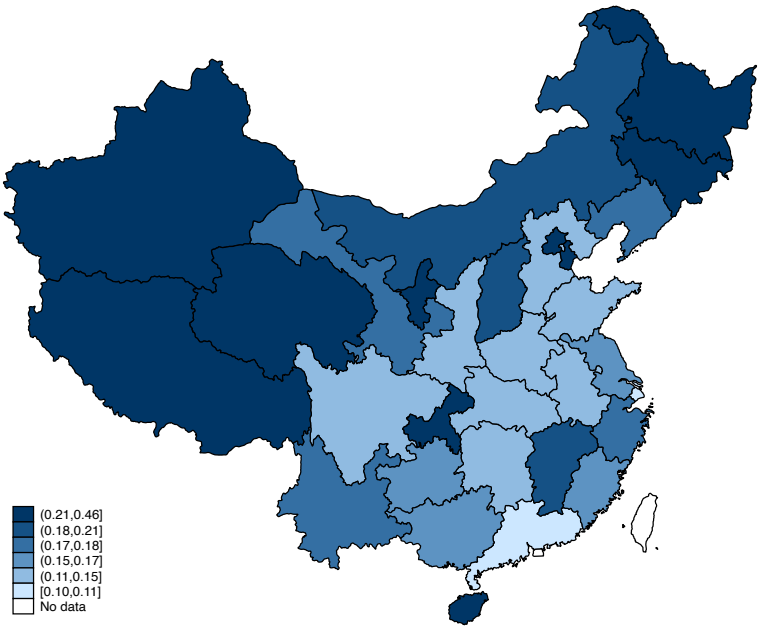

Figure 4: Out-of-province college enrollment rate, 2005

Notes: This figure shows the out-of-province enrollment rate for each province in the year 2005. The out-of-province enrollment rates are calculated by dividing the predicted out-province enrollment by the number of high school graduates in 2005, following Equation (4) $OR^{2005} = \frac{\sum_{k \neq j} E_k^{2005} \times s_{jk,2000}}{H_j^{2005}}$.

4.5 The Validity of the Instruments

We expect that our constructed college enrollment rates can effectively capture the exogenous variations in provincial-level college expansion. These variations are unlikely to be influenced by individual decisions on college attendance and enrollment location, thus mitigating the concern about reverse causality in the first stage.

Two key factors lend credibility to this claim. First, our analysis leverages the initial college seat distribution that followed the 1990s college expansion policy. This distribution is unlikely to be influenced by individuals' education choices in subsequent periods.

Second, college candidates declare their enrollment preferences shortly after the College Entrance Exam (Gaokao), and just before the official release of enrollment statistics. This timing ensures that the formulation of province-level college seat quotas remains uninfluenced by candidates' preferences.

We anticipate that individuals are more likely to attend college following substantial expansions in higher education opportunities. Furthermore, exposure to significant out-of-province college expansion shocks may increase the probability of individuals attending college out-of-province.

Furthermore, the exogenous nature of college expansion shocks at the macrogeographic level ensures that they are largely unrelated to the unobserved determinants of individual migration de-

Table 2: Province-level characteristics and college expansion shocks

	(1) IR	(2) OR	(3) IR	(4) OR
Population (million)	0.012	0.002	0.014	0.004*
	(0.009)	(0.003)	(0.009)	(0.003)
Provincial GDP growth rate	-0.049	-0.083**	-0.097	-0.055
	(0.116)	(0.036)	(0.193)	(0.089)
Provincial employment growth rate	-0.039	0.003	-0.096	-0.017
	(0.083)	(0.014)	(0.085)	(0.015)
Province FE	✓	✓	✓	✓
Year FE			✓	✓
Observations	155	155	155	155
R^2	0.879	0.958	0.897	0.967

Note: This table presents the OLS estimates showing the relationship between local fundamentals and college enrollment rates at the province level. Dependent variables are in-province enrollment rate (IR) in columns 1 and 3 and out-of-province enrollment rate (OR) in columns 2 and 4. They are calculated from using the 2000 population census origin-destination province college students shares and provincial total high school graduates from 2001 to 2005. We include all 31 provinces in mainland China. Robust standard errors are in parentheses.

cisions. This argument is strengthened by the construction of IR_{jt} (in-province college expansion intensity) and OR_{jt} (out-of-province college expansion intensity), which incorporate information on the initial-period college seat distribution and province-level total college enrollment – factors beyond the control of individual decision-making.

We perform a regression analysis to examine the possibility of potential selection into province-level college expansion based on location-specific fundamentals. The model investigates potential associations between the intensity of in-province and out-of-province college expansion and province-level fundamental characteristics such as population size, GDP growth rate, and employment growth rate, including province fixed effects and year fixed effects as additional controls.

The results of our estimation are detailed in Table 2. Our analysis reveals little evidence of selection into college expansion according to local fundamental factors. This result strengthens our baseline assumption that province-level college seat supply shocks, driven by higher education expansion, can be considered effectively exogeneous.

5 Baseline Results

5.1 College Expansion and College Education Choice

In this section, we initiate the first step of our two-step IV approach by utilizing in- and out-of-province college expansion shocks to predict individual decisions regarding college attendance and college enrollment location. Specifically, we estimate equations (5) and (6) using probit model estimation. The results are presented in Table 3. In column 1, we employ the college attendance status, $College$, as the outcome variable, while in columns 2 and 3, we utilize the out-of-province college attendance, $CollegeOut$, as the outcome variable.

Column 1 of Table 3 reveals a positive correlation between both in- and out-of-province college expansions and the likelihood of college attendance, despite the coefficient estimate for out-of-province college expansion being somewhat less precise. More specifically, a one-percentage-point rise in in-province college enrollment increases the probability of an average young adult attending college by 0.13 percentage points (marginal effect at the mean). Similarly, the marginal effect of out-of-province enrollment rate on the probability of college attendance is 0.05 percentage points.

Moving to column 2, the coefficient estimate for in-province college expansion turns insignificantly negative. This suggests a negligible or slightly negative impact of in-province college expansion on the likelihood of attending an out-of-province college. This result is in line with expectations, as individuals often encounter increased opportunity costs when considering education outside their home provinces, particularly when there is an ample supply of college seats within their home provinces.

On the other hand, the coefficient estimate for out-of-province expansion remains positive and statistically significant. This suggests that out-of-province college expansion shocks encourage individuals to pursue college education outside their home provinces. Specifically, a one-percentage-point increase in the out-of-province college enrollment rate corresponds to an approximate 0.18-percentage-point increase in the probability of attending college out-of-province (marginal effect at the mean).

In column 3, we focus on individuals with a college education and reproduce the analysis from column 2. The point estimate for out-of-province expansion derived from this subset closely aligns with those in column 2 but with greater magnitude. The influence of in-province college expansion remains negative and statistically insignificant, with its magnitude being roughly twice that derived from the full sample estimation.

These results strengthen the idea that the choice to attend an out-of-province institution is primarily driven by out-of-province college expansion. Although in-province college expansion potentially influences college attendance decisions for young adults, it may not significantly impact choices regarding enrollment location. If anything, a substantial supply of college seats within their home province could motivate individuals to pursue their college education locally.

We check the robustness of our analysis by employing linear probability estimation in Table

Table 3: The effect of college expansion on college attendance choices, probit estimates

	(1) $College$	(2) $CollegeOut$	(3) $CollegeOut$
In-province college expansion (IR)	0.331*	-0.274	-0.423
	(0.199)	(0.210)	(0.259)
Out-of-province college expansion (OR)	0.128	2.396*	2.805*
	(0.544)	(1.383)	(1.617)
Han ethnicity	-0.080***	-0.013	0.016
	(0.025)	(0.031)	(0.033)
Male	-0.138***	0.088***	0.179***
	(0.016)	(0.022)	(0.023)
Household size	-0.084***	-0.009	0.031***
	(0.005)	(0.007)	(0.008)
Age	0.544	-1.333**	-2.042***
	(0.470)	(0.674)	(0.737)
Age square	-0.011	0.027*	0.041***
	(0.010)	(0.014)	(0.015)
City population	-0.012***	-0.002	0.004
	(0.004)	(0.009)	(0.009)
City GDP per capita	0.106***	-0.093***	-0.145***
	(0.014)	(0.025)	(0.031)
City tertiary employment share	0.004**	0.002	0.001
	(0.002)	(0.002)	(0.003)
City distance to the coast	-0.022**	0.005	0.014
	(0.011)	(0.021)	(0.023)
Province urban employment	-0.319	-1.941	-1.615
	(0.977)	(1.504)	(1.761)
Province GDP per capita	-0.078	0.252***	0.269***
	(0.050)	(0.080)	(0.097)
Provincial capital city	0.121***	-0.154**	-0.255***
	(0.036)	(0.067)	(0.076)
Municipality	-0.019	0.226	0.199
	(0.251)	(0.556)	(0.488)
Province GDP growth rate	0.176	0.341	0.439
	(0.216)	(0.299)	(0.337)
Province employment growth rate	-0.007	-0.210	-0.237
	(0.162)	(0.182)	(0.187)
Observations	63,498	63,498	31,380
Marginal effect of IR	0.126*	-0.025	-0.064
	(0.076)	(0.019)	(0.039)
Marginal effect of OR	0.049	0.022*	0.425*
	(0.207)	(0.126)	(0.245)

Note: This table shows the probit regression results of in- and out-of-province enrollment rate on college attendance and out-of-province college attendance using the 2010 census. The dependent variable in column 1 is a dummy variable indicating whether an individual attended college. The dependent variable in columns 2 and 3 is a dummy variable indicating whether an individual attended college outside the birth province. Column 3 keeps college individuals only. All regressions control for birth province and birth year fixed effects and are weighted by high school graduates size. Robust standard errors are clustered at the birth province level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.1. The results align with those obtained from the probit model estimation.

5.2 The Effect of College Education and the Effect of Enrollment Location

In the second phase of our two-step IV approach, we utilize the predicted values $\widehat{College}$ and $\widehat{CollegeOut}$, derived from the estimations of equations (5) and (6), as instruments for individuals' actual decisions regarding college attendance ($College$) and out-of-province college enrollment ($CollegeOut$). Subsequently, we estimate our baseline model, as defined in equation (1), using a 2SLS estimation with these constructed instruments.

The results are reported in Table 4. In column 1 of panel A, we report the first-stage estimates for the outcome variable $College$. Our findings indicate that a one-percentage-point increase in the predicted probability of college attendance $\widehat{College}$ corresponds to an average increase of 2.1 percentage points in the actual probability of attending college.

We find a weak negative correlation between the predicted probability of out-of-province college enrollment, denoted as $\widehat{CollegeOut}$, and the likelihood of actual college attendance, denoted as $\widehat{College}$. This result can be attributed to the inherent definitions of the variables $College$ and $CollegeOut$. For instance, when a student opts for in-province college attendance, the variable $College$ is set to 1. However, this requires that the out-of-province college attendance, $CollegeOut$, must be set to 0, leading to a negative correlation by construction.

In column 2 of panel A, we present the first-stage estimates for the outcome variable $CollegeOut$. A one-percentage-point increase in $\widehat{CollegeOut}$ corresponds to a 1.1-percentage-point rise in the actual propensity for out-of-province college enrollment. Interestingly, an increase in $\widehat{College}$ diminishes the actual likelihood of attending college out-of-province, again possibly due to the inherent definitions of the variables $College$ and $CollegeOut$. Our first-stage F-statistic, standing at 16.82, mitigates concerns about weak instruments.

Turning to column 3 of panel A, we provide the reduced-form estimates. These estimates reveal that a one-percentage-point increase in $\widehat{College}$ is associated with a 0.8-percentage-point decrease in an average person's probability of migrating to another province after college. On the other hand, a one-percentage-point increase in $\widehat{CollegeOut}$ leads to an approximate 3.6-percentage-point increase in the probability of postcollege migration.

We present the second-stage estimates in panel B of Table 4. In column 1, we report the findings from the 2SLS estimation. Consistent with the reduced-form estimates, our 2SLS estimates reveal that the pursuit of a college education itself diminishes the likelihood of postcollege migration. Specifically, a one-percentage-point increase in the probability of attending college leads to a reduction of approximately 0.2 percentage points in the probability of postcollege migration, net of the effect of college enrollment location.

Table 4: Estimating the college education effect and the college location effect, 2SLS estimates

Dependent variable:	(1) $College$	(2) $CollegeOut$	(3) Migration
Panel A: First-stage & RF estimates			
$\widehat{College}$	2.114***	-0.131**	-0.792***
	(0.215)	(0.063)	(0.193)
$\widehat{CollegeOut}$	-1.009*	1.133***	3.601***
	(0.523)	(0.187)	(0.481)
Observations	63,498	63,498	63,498
Kleibergen-Paap F	16.82	16.82	-
Dependent variable: Migration	IV	OLS	
Panel B: Second-stage estimates			
$College$	-0.189**	-0.102***	
	(0.0735)	(0.0178)	
$CollegeOut$	3.011***	0.859***	
	(0.485)	(0.0143)	
Observations	63,498	63,498	

Notes: This table presents the main estimation results of Equation (1). Columns 1 and 2 in panel A show the first-stage results of 2SLS estimation and column 3 reports the reduced-form (RF) estimates. Column 1 uses a college attendance dummy as the dependent variable; Column 2 uses an out-province college enrollment dummy as the dependent variable. Column 3 uses a postcollege migration dummy as the dependent variable. Panel B presents our second-stage estimates. The dependent variable is a postcollege migration dummy variable. All regressions control for birth-province fixed effects, birth-year fixed effects, and individual and location characteristics. Regressions are weighted by the number of high school graduates. Robust standard errors are clustered at the birth province level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

In contrast to the weak negative effect of college education itself, we observe a significantly positive impact of out-of-province college enrollment on postcollege migration propensity. Our analysis shows that a one-percentage-point increase in the likelihood of out-of-province college enrollment leads to an approximate 3.0-percentage-point rise in the probability of migration after college, net of the effects of college education itself.

Our results reveal a significant effect of college enrollment location, with a magnitude seemingly notably larger than those reported in existing literature. For instance, Winters (2020) observed a modest 0.41 percentage point increase in the likelihood of individuals residing in their birth state post-education for every one percentage point rise in home-state enrollment rates.

However, caution is warranted when making direct comparisons due to our use of different metrics and individual-level analyses, as opposed to regional-level. It is crucial to note that our study assesses the impact of college enrollment location on migration patterns post-graduation, while simultaneously netting out the influence of college education *per se*.

For comparison, we present the OLS estimation results for the second-stage analysis in column 2 of panel B, offering a comparative perspective. These OLS estimates align with the outcomes derived from the 2SLS estimation in column 1, albeit exhibiting variations in both the magnitudes and statistical significance of the point estimates.

Specifically, the negative college education effect is stronger in the IV specification, although its precision diminishes slightly. Furthermore, the IV specification uncovers a markedly stronger positive effect of college enrollment location. These discrepancies suggest that the OLS estimates of the effects of college education and college enrollment location might be biased towards zero due to measurement errors or the selection according to unobserved individual heterogeneity within the birth province.

5.3 Robustness Checks

We conducted a series of sensitivity tests based on the IV specification in column 1 of Table 4, panel B to examine the robustness of our main findings. The results are reported in Table 5.

In column 1, we evaluated the sensitivity of the coefficient estimates to the weights applied in the baseline model. This was achieved by re-running the baseline estimation without the use of any weights. The results from this analysis were largely consistent with our initial estimates.

Our IV approach primarily hinges on college expansion shocks at the provincial level, leaving room for potential sorting bias at the city level. This could cast a shadow of doubt on our baseline findings. To address this issue, we examined the potential impact of unobserved heterogeneity in column 2, driven by the migration preferences of young individuals towards larger cities, as underscored by Xing and Zhang (2017).

Table 5: Robustness checks

	(1) No weight	(2) No Tier-1 cities	(3) No provincial capitals
College	-0.163**	-0.200**	-0.253***
	(0.0726)	(0.0881)	(0.0706)
CollegeOut	3.170***	3.122***	2.212***
	(0.507)	(0.598)	(0.402)
Observations	63,498	60,058	53,982
Kleibergen-Paap F	17.09	12.01	28.42

Note: This table presents the results of robustness checks. All the results are based on the IV estimation. Column 1 does not include weight in our specifications. Column 2 drops individuals born in tier-1 cities, including Beijing, Shanghai, Guangzhou, and Shenzhen. Column 3 drops individuals who were born in provincial capitals. All regressions control for birth province and birth year fixed effects and the same variables as those in Table 3. Regressions in columns 2 and 3 are weighted by the local total number of high school graduates. Robust standard errors are clustered at the birth province level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

In particular, we excluded individuals born in Tier-1 cities–Beijing, Shanghai, Guangzhou, and Shenzhen–from the sample. Following this, we replicated the baseline estimation using this refined sample. Reassuringly, the estimates derived from this reduced sample are closely aligned with our baseline results.

In column 3, we excluded individuals born in provincial capitals, given these cities often attract intra-provincial migrants. This measure was taken to ascertain if our primary findings were dependent on specific samples, especially those from economically developed cities. The results of this analysis, reassuringly, mirrored our baseline estimates.

6 Understanding the Estimated Migration Effects

Our baseline results suggest a weak negative influence of college education on the migration of young adults in later life. This finding contradicts previous empirical studies that document a correlation between higher education and increased mobility (Malamud and Wozniak, 2012; Weiss, 2015). How do we reconcile college education *per se* weakly lowering the propensity to migrate, while strongly predicting migration when considering distant college attendance?

When individuals decide on their place of residence, they aim to maximize the net present value of their expected future utility. Both the location of college enrollment and the attainment of a college education can influence an individual's utility through two primary channels: preference and income.

To identify the college education effect per se, we compare two individuals: one who attended

college within their home province and another who did not attend college, both born in the same province and year. Accounting for individual and birth location characteristics, we can attribute any variations in migration tendencies between these two individuals to their differing educational attainment. This difference in education leads to distinct preferences for specific location characteristics and varying expected income if they were to migrate.

To discern the effect of the location of college enrollment *per se*, we compare two college-educated individuals. Both were born in the same province and year, but one attended college within their home province, and the other attended college outside the province.

Accounting for variations in individual characteristics and birth location, we can attribute any differences in migration choices between these two individuals to their unique preferences for specific location characteristics as well as disparities in expected income at potential migration destinations, both of which stem from their residential experiences during college either at home or outside the home province.

Building on the previously discussed rationale, this section delves into the heterogeneity of the effects of college education and enrollment location on later-life mobility. We consider a set of factors that could potentially capture variations in young adults' expected income and preferences for location characteristics to help explain our baseline findings.

6.1 Migration and Wage Growth

Wages play a crucial role in individuals' location choices (Sjaastad, 1962; Wozniak, 2010). Through the income channel, individuals make decisions regarding whether to stay in their birthplace or migrate based on the wage differential between their birth location and other locations. This wage differential can reflect individual preferences for productivity, amenities, and other location characteristics (Hoogstra, van Dijk, and Florax, 2017).

The observed negative impact of college education *per se* on migration in our baseline result may be explained by the higher wage growth anticipated by non-college-educated individuals upon migration, which leads to a preference for distant job opportunities beyond their home provinces. This trend could be due to a more pronounced spatial mismatch between the demand and supply of low-skilled labor in the country. If this is the case, a stronger negative effect of college education would be expected in provinces where the wage gains from migration are more substantial for non-college-educated individuals.

By contrast, the impact of college enrollment location should remain unaffected by the expected wage growth upon migration for non-college-educated individuals, given their lack of college attendance. However, a more pronounced positive effect of college enrollment location is anticipated in provinces with larger potential wage gains from migration for college-educated

Table 6: Heterogeneous effects of college attendance and college enrollment location

Dependent variable: Migration (=1)	(1)	(2)	(3)	(4)	(5)
$College$	-0.117*	-0.130*	-0.126*	-0.184***	-0.138**
	(0.0628)	(0.0724)	(0.0670)	(0.0680)	(0.0646)
$CollegeOut$	1.907***	2.369***	0.454	3.031***	2.997***
	(0.710)	(0.477)	(0.542)	(0.476)	(0.483)
$College \times \Delta w^L$	-0.632***				
	(0.214)				
$CollegeOut \times \Delta w^L$	3.718*		‘		
	(2.037)				
$College \times \Delta w^H$		-0.349*			
		(0.187)			
$CollegeOut \times \Delta w^H$		2.633*			
		(1.598)			
$College \times \Delta hp$			-0.0929***		
			(0.0200)		
$CollegeOut \times \Delta hp$			1.105***		
			(0.358)		
$College \times \Delta hukou$				-0.0103	
				(0.0444)	
$CollegeOut \times \Delta hukou$				0.218***	
				(0.0841)	
$College \times$Male					-0.0896***
					(0.0278)
$CollegeOut \times$Male					-0.0244
					(0.225)
Observations	63,498	63,498	63,498	63,354	63,498
Kleibergen-Paap F	5.883	8.349	8.280	8.452	8.979
p-values (Anderson-Rubin test)	0.023	0.020	0.011	0.043	0.007

Notes: This table presents the results of the heterogeneity analysis, where the dependent variable is a binary indicator representing out-of-province migration. Δw^H and Δw^L are migration wage differentials for college-educated individuals and non-college-educated individuals calculated using equation (7). Δhp is the migration housing price gap calculated using equation (8). $\Delta hukou$ represents the *hukou* reform index differentials for each province using equation (9). All specifications include birth-province and birth-year fixed effects, as well as the same set of control variables as in equation (3). The regressions are weighted by the number of high school graduates at the province level. Robust standard errors are clustered at the birth-province level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

individuals. This effect could be attributed to the influence of the college residential experience, which strengthens individuals' connections to their college enrollment location through enhanced social networks, employer relationships, and access to location-specific amenities (Winters, 2011; Huang, Xing, and Cui, 2022).

To test the hypotheses, we utilize data from the 2005 population census. The data provides monthly earnings information for both educational groups. We focus on prime-age individuals (aged 16-60) who were employed and had positive earnings in 2005. Since the cohorts in our main analyses all graduated after 2005, this data measures the expected wage of both non-college-educated and college-educated individuals prior to their job search.[11]

Let w_j^L and w_j^H denote the average monthly wages for non-college-educated and college-educated individuals in province j, respectively. Using census data, we calculate M_{jk}, the number of migrants from province j now residing in province k. The total number of migrants from province j, represented as M_j, is obtained by summing the number of migrants from province j to all other provinces, excluding province j itself, i.e., we have $M_j = \sum_{k=1,j\neq k}^{31} M_{jk}$.

We use the weight M_{jk}/M_j to estimate the probability of migration from province j to province k. These weights enable us to calculate the weighted average of prevailing wage levels in potential destination provinces for individuals born in province j.

The expected wage differential upon migration is measured by calculating the difference between the weighted average expected wage gains from migrating from province j to other provinces and the prevailing wage w_j in province j. Specifically, the wage differential upon migration is calculated as follows:

$$\Delta w_j = \sum_{k\neq j} w_k \times \frac{M_{jk}}{M_j} - w_j \tag{7}$$

A positive migration wage differential, Δw_j, intuitively reflects the degree to which the wage in one's home province is lower than in potential migration destinations. A larger Δw_j indicates higher wages in other provinces, suggesting better job opportunities and career prospects.

Figures A.1 and A.2 map the spatial distribution of estimated wage differentials for two groups upon migration: non-college-educated (Δw_j^L) and college-educated individuals (Δw_j^H). Darker shades indicate higher migration wage differentials. The figures show similar patterns, with individuals from gridlocked provinces expecting the most significant wage gains from migration. In contrast, those from Beijing, Shanghai, Guangdong, and coastal provinces anticipate minimal

[11] In 2005, higher wages were predominantly concentrated in major cities such as Beijing, Shanghai, and Guangzhou, as well as in coastal provinces like Jiangsu, Zhejiang, and Fujian. Conversely, inland provinces and northwestern regions exhibited lower wage levels for non-college-educated individuals. Notably, provinces with substantial ethnic populations tended to offer higher wages to non-college-educated individuals, possibly due to favorable policies and a thriving tourism industry.

or even negative wage gains from migration.

The figures also reveal a notable difference. High wage premiums for migration for non-college-educated workers are predominantly found in western and central China, excluding provinces like Shanxi and Inner Mongolia which are rich in resources. For those with a college education, provinces with high wage premiums are only found in central China, suggesting potential subsidies in western provinces to retain such workers.

Building upon our baseline specification as outlined in equation (1), we introduce interaction terms involving the variables $College_{ijt}$ and $CollegeOut_{ijt}$ with Δw_j. To mitigate endogeneity concerns, we employ $\widehat{College}_{ijt} \times \Delta w_j$ and $\widehat{CollegeOut}_{ijt} \times \Delta w_j$ as instruments for the newly introduced interaction terms $College_{ijt} \times \Delta w_j$ and $CollegeOut_{ijt} \times \Delta w_j$. By analyzing the coefficients of these interaction terms, we capture the heterogeneity in the effects of college education and college location on migration, which is determined by the wage differentials associated with migration.

The results are presented in columns 1 and 2 of Table 6. The F-statistics from the first stage of the IV estimation are below the conventional acceptable levels but we report small p-values from the Anderson-Rubin weak IV tests, which mitigates the concerns about weak instruments.

Our findings continue to demonstrate that pursuing a college education itself tends to discourage individuals from migrating after college, whereas distant college residential experience increases the likelihood of migration later in life.

Examining the coefficient estimates for interaction terms reveals a more pronounced negative impact of college education in provinces where residents see greater wage gains for low-skill migration. This observation aligns with existing evidence indicating that less educated individuals in China tend to be more responsive to changes in labor demand (Luo and Xing, 2016).

However, no significant difference in the moderating effect of wage potential through migration is found between high-skill and low-skill individuals. This could be due to a high correlation between the wage differential measures for different skill groups, or perfect substitution between high- and low-education workers in the Chinese context.

6.2 Housing Cost

Bilal and Rossi-Hansberg (2021) emphasize the importance of viewing individuals' location choices as "location asset" investments. Each location provides benefits such as job opportunities, potential human capital returns, and amenities. However, it also incurs costs, including housing expenses. Locations with productivity advantages and superior amenities, which command higher rents, are typically associated with higher costs, such as elevated housing prices.

We argue that the importance of location as an asset feature should differ among skill groups.

College-educated individuals notably experience the tangible impact of housing costs upon migration, given their better chances of acquiring the local *hukou* system in their migration destinations. This system tends to favor the college-educated, allowing them to purchase property and access local public benefits, including public school enrollment for their children (Zheng, Hu, and Wang, 2016; Ding and Itoh, 2023).

However, non-college-educated individuals, without access to the local *hukou* and consequently local housing and public benefits in migration destinations, typically choose to send their earnings back to their home province and invest in properties there (Tao et al., 2015). This is often the case as low-skill migrants' children are usually left in their hometowns, under the care of grandparents or other relatives (Ai and Hu, 2016; Gao et al., 2023).

Given the distinct institutional background, individuals without a college education may face lower initial migration costs, particularly those related to housing in destination provinces. This is in contrast to their college-educated counterparts. Therefore, we anticipate a more pronounced negative impact of college education *per se* among individuals born in provinces with lower housing costs. This is especially true for low-skill individuals who perceive higher returns to migrating to expensive cities but are not significantly affected by the upfront housing costs associated with migration.

However, it is complex to ascertain whether the impact of the college enrollment location itself amplifies or diminishes with varying housing cost disparities between the home province and migration destinations. This complexity arises from the fact that the college residential experience could update individuals' beliefs about the returns and costs associated with specific locations.

To investigate the heterogeneous effects of college education and enrollment location on migration according to regional differences in housing cost, we collected data on the average annual sale prices of commodity housing in 36 major cities from 2002 to 2005.[12] This period precedes the graduation of our sample cohorts. The data, obtained from the National Bureau of Statistics, includes housing price data from 4 direct-administered municipalities (Beijing, Tianjin, Shanghai, and Chongqing), 27 provincial capitals, and 5 other large cities.

To measure the housing price gap at the province level, we used a calculation method similar to that for determining potential wage growth upon migration. This method employs the migration share (M_{jk}/M_j), the ratio of migrants from each origin province j to destination province k, as weights. The housing price gap (Δhp_j) for individuals born in province j who might migrate is then calculated by comparing the weighted average of *ex-ante* housing prices in all other destination provinces with the average *ex-ante* housing price in their home province, which is

[12]In China, commodity housing refers to housing sold at market price, not as a subsidized benefit to employees with resale restrictions.

given by the equation as follows:

$$\Delta hp_j = \sum_{k \neq j} hp_k \times \frac{M_{jk}}{M_j} - hp_j \tag{8}$$

where hp_k represents the average housing prices for destination province k, while hp_j represents the average housing price in the birth province j. A positive Δhp_j indicates a lower *ex-ante* average housing price in the birth province j compared to other potential migration destinations. This suggests that individuals born in province j could perceive both high returns (e.g., higher wages, nicer amenities, and other positive externalities) and high costs (e.g., expensive housing) associated with potential migration.

Figure A.3 illustrates the spatial distribution of the *ex-ante* housing price differential (Δhp_j) for potential migrants, excluding Tibet due to data unavailability. Developed regions like Beijing, Shanghai, Guangdong, and coastal provinces exhibit the most negative housing differentials, indicating higher home housing prices compared to other provinces. In contrast, gridlocked provinces such as Guangxi, Hunan, and Jiangxi, show the largest positive housing price differential, indicating lower home housing prices compared to other provinces.

To test our hypothesis, we incorporated the interaction term between the housing price gap for potential migrants and two key variables: the college education indicator (College) and the out-of-province college enrollment indicator (CollegeOut), into our baseline model. The results, presented in column 3 of Table 6, show a statistically significant coefficient estimate of -0.093 for $College \times \Delta hp_j$, suggesting a more noticeable mobility difference between college-educated and non-college-educated individuals from provinces with a larger housing cost gap between destination and home provinces. This result aligns with the hypothesis that non-college-educated individuals are less affected by housing cost increases in migration destinations, but are equally, if not more, responsive to migration returns than their college-educated counterparts.

Furthermore, a notably positive coefficient estimate of 1.105 for $CollegeOut \times \Delta hp_j$ indicates that the influence of college enrollment location on post-college migration increases for those born in provinces with a higher gap in housing prices between destination provinces and home province. This suggests that individuals recognize the value of the "location asset" once they reside in a specific location during college. The benefits of staying in college enrollment locations after graduation, such as higher wages, local alumni networks, human capital spillovers, efficient labor market matching, urban amenities, and the experience of living in a large, expensive city, outweigh the costs, including higher housing expenses (Moretti, 2004; Zheng, 2016; Roca and Puga, 2017; Charles, Hurst, and Notowidigdo, 2018).

6.3 Household Registration (*Hukou*) Reforms

China's mass migration is intricately linked to the reforms of the *hukou* system. Historically, individuals were confined to their *hukou*-registered location. However, staggered *hukou* reforms in the late 1990s enabled migrants to acquire local *hukou*, leading to a significant reduction in province-level migration costs between 2000 and 2005 (Tombe and Zhu, 2019). This spurred substantial migration flows (Fang and Huang, 2022).

Despite policy relaxations in the 1980s and 1990s promoting regional migration, access to local benefits such as healthcare, education, and housing remained tied to one's *hukou* status. These recent reforms, which had a significant impact on the labor market, tend to favor college-educated individuals. This has implications for the heterogeneity in the effects of college education and enrollment location on geographical labor mobility.

The link between *hukou* ownership and migration patterns is complex for young adults entering the labor market. College graduates experience both advantages and constraints due to the ease of *hukou* transfer. On one hand, they can typically transfer their household registration to their post-graduation residence, granting them local public benefits and reducing migration friction. This may increase their likelihood of migrating to their first job location after college. On the other hand, those who attend college locally may be less likely to move post-graduation due to the benefits tied to a local *hukou*. As a result, geographical mobility among college-educated individuals may be limited, even after considering the effect of college location.

For high-school graduates without a college degree, the labor market presents a different scenario. Despite being allowed to migrate, they typically cannot change their place of *hukou* registration. This restriction could implicitly limit their permanent migration, as they lack access to the welfare benefits available to local residents and college-educated migrants. However, non-college-educated individuals often seek temporary work in distant locations with abundant labor market opportunities. Often, they remit earnings back to their families in their hometowns, particularly during the early years of their careers before marriage (Brown and Jimenez-Soto, 2015). As their chances of obtaining a *hukou* are low, their migration decisions are less influenced by the *hukou* system's benefits. Consequently, they face fewer *hukou*-related mobility costs, resulting in greater geographical mobility compared to their college-educated counterparts during the early stages of their careers.

To examine the heterogeneous effects of college education and enrollment location on migration across provinces with differing *hukou* stringency levels, we need a measure of province-level differences in *hukou* stringency for those who might migrate. For this purpose, we utilize the *hukou* reform index (R) compiled by Fan (2019). A higher *hukou* reform index indicates a lower *hukou* stringency. Since the index was constructed at the city level, we first use the 2000 population census to estimate the share of the migrant population in each city c of its province j. This allows

us to construct a weight reflecting the relative size of out-of-province migrants for city c:

$$\omega_{cj} = \frac{\text{Total number of migrants in city } c}{\text{Total number of migrants in province } j}$$

We then calculate the exposure of individuals born in province j to the *hukou* reforms in other provinces, relative to the *hukou* stringency in province j itself. This exposure measure is denoted as $\Delta hukou_{jt}$ and is given by the following equation:

$$\Delta hukou_{jt} = \sum_{k \neq j} R_{kt} \times \frac{M_{jk}}{M_j} - R_{jt} \tag{9}$$

Using the exposure measure of the difference in *hukou* reforms in potential migration destinations and home provinces, we estimate a variant of the IV estimation of equation (1). Specifically, we introduce interaction terms between the variables $College$, $CollegeOut$, and $\Delta hukou_{jt}$. We then construct instruments for these new interaction terms using the interactions between predicted college education choices and the *hukou* reform index.

The results, presented in column 4 of Table 6, show the continued negative influence of college education and the positive effect of college location. Interestingly, the effect of college education itself does not significantly vary with the *hukou* stringency difference between potential migration destinations and home provinces. This aligns with our hypothesis that low-skill individuals are less affected by geographic differences in *hukou* stringency, given that the ease of *hukou* transfer is mainly targeted at college graduates.

However, we find that individuals who attended college out-of-province are more likely to migrate post-college if they perceive more lenient *hukou* restrictions in potential migration destinations. This highlights the lasting impact of the residential experience during college years on future migration. College-educated individuals are more likely to recognize the benefits of favorable policies, such as relaxed *hukou*-related regulations or *hukou* sponsorship opportunities for college-educated talents, if they have residential experience in their college enrollment location.

6.4 Gender Differences

Male college graduates in China are more likely to relocate from their home province (Huang, Xing, and Cui, 2022). This trend is potentially more pronounced among non-college-educated individuals due to cultural norms and the impact of childbirth on women's labor supply.

Moreover, the gender wage gap is less among college graduates due to the rising demand for skilled labor (Hughes and Maurer-Fazio, 2002). Consequently, non-college-educated males may be more attracted to better job opportunities in other provinces, which could lead to a larger migration effect of college education *per se*. However, the effect of the college enrollment location, resulting

from the residential experience during college, should be independent of gender differences.

To test our hypotheses, we incorporated interaction terms between the gender dummy variable ($Male$), the college education dummy ($College$), and the dummy variable for out-of-province college attendance ($CollegeOut$), into equation (1). The results, presented in column 5 of Table 6, reveal a significantly higher mobility disparity between college-educated and non-college-educated males than that between their female counterparts. This suggests that non-college-educated males may perceive higher wage potential through migration and face lower migration costs.

Interestingly, the coefficient for the interaction term between out-of-province college attendance and males ($CollegeOut \times Male$) is insignificant, indicating a gender-neutral effect of college enrollment location. This confirms that the college residential experience similarly influences both males' and females' ties to their college enrollment location and consequently, their later-life propensity to migrate.

6.5 Discussion

Our study challenges the common belief that higher education leads to increased geographic mobility. Contrary evidence suggests that in China, the observed enhanced mobility among non-college-educated individuals (i.e., the negative effect of college education *per se*, net of the influence of college enrollment location) mainly arises from varying sensitivities to factors such as wage growth expectations, location-specific rents and costs, and *hukou* restrictions among college-educated and non-college-educated young adults.

Our findings indicate that non-college-educated individuals show a higher sensitivity to distant job opportunities while being less responsive to housing costs at migration destinations. This behavior is tied to their aim to secure high-paying jobs and send income back to their hometowns, coupled with lower short-term migration costs due to skill-favoring *hukou* regulations. Interestingly, the negative effect of college education on mobility is more pronounced among non-college-educated males, possibly due to traditional norms, societal conventions, and labor market conditions.

The impact of the college enrollment location *per se* is primarily due to the "experience good" feature of a specific place. College graduates form strong bonds with their college location due to their firsthand experience of living there during their studies. This exposure updates their perceptions about the location assets' rents (e.g., labor market dynamics, amenities, social networks), fostering a preference to stay in that area after college.

Our research has significant implications for policy discussions surrounding higher education investment and migration restriction policies. We highlight the potential role of targeted investments in college expansion at the local level in improving local college enrollments, with

effects that might not dissipate through migration. However, it is still crucial for future research to thoroughly consider both the benefits and costs associated with public college investment. Factors such as income tax revenue, college tuition revenue, subsidies, and college financing should be carefully evaluated to fully justify the implementation of such policies (Kennan, 2015).

Our findings underscore the intricate issue of college seat distribution across regions, revealing a tradeoff between efficient human capital sorting and equal access to higher education. The significant effect of the college enrollment location *per se* suggests that an imbalanced distribution of college seats can lead to a spatial concentration of college-educated individuals. This exacerbates spatial inequality in the supply of skilled labor (Huang, Xing, and Cui, 2022; Fabre, 2023) and results in welfare losses due to inefficient sorting (Fajgelbaum and Gaubert, 2020).

An uneven distribution also poses challenges for regions with few elite education opportunities to halt the brain drain, given the high costs of implementing favorable policies such as monetary incentives, local *hukou* privileges, and housing subsidies. However, an equal distribution of higher education seats may hinder efficient student sorting and reduce overall welfare (Yang, 2021).

Finally, our estimates suggest that China's skill-favoring *hukou* system amplifies the influence of college education and enrollment location on migration decisions. This underscores the *hukou* system's restrictive effect on college graduates, binding them to certain locations through location-specific welfare, thereby impeding efficient allocation.

It is worth noting that, while the *hukou* system does not limit the temporary migration of non-college-educated youth, it may significantly reduce their chances of permanent migration later in life. Policymakers could alleviate mobility barriers for non-college-educated youth by easing *hukou* restrictions.

7 Conclusion

This study leverages the higher education expansion in China's late 1990s to investigate the effects of college education and enrollment location on young adults' migration decisions. Our findings reveal that higher education expansion, regardless of location, increases college attendance likelihood. However, only out-of-province expansion raises the probability of attending a non-home province college.

Contrary to the prevailing evidence, our two-step IV estimates suggest that college education itself, net of the effect of college enrollment location, does not encourage migration. This is attributed to non-college-educated individuals' higher sensitivity to high-wage job opportunities and lower sensitivity to upfront housing costs upon migration.

In line with existing literature, we find that the choice of college enrollment location significantly influences young adults' migration tendencies. The response of higher-educated individuals

to opportunities in distant labor markets, cost of living, productivity benefits, amenities, and migration regulations are tied to contextual factors perceived differently by college students due to their college residential experience.

Our paper provides valuable insights for governments considering college expansion policies as a strategic approach to reallocating human capital within specific regions. Policymakers must consider a variety of interconnected factors, such as migration policies, public finance, housing costs, labor market conditions, demographic characteristics, and local amenities (Zheng, 2016; Bilal and Rossi-Hansberg, 2021) when formulating policies that promote efficient spatial allocation of human capital and enhance overall welfare.

Future research should evaluate the welfare impact of education and migration policies, considering both their benefits and costs. Our study focuses on young adults due to data availability. This suggests opportunities for further investigation into how migration decisions respond to changing policy environments such as *hukou* reforms and talent initiatives. These dynamics are crucial for understanding the lifelong effects of college education and enrollment location on internal migration patterns.

References

Abel, Jaison R, and Richard Deitz. 2012. "Do Colleges and Universities Increase Their Region's Human Capital?" *Journal of Economic Geography*, 12(3): 667–691.

Ai, Hongshan, and Junmin Hu. 2016. "Psychological Resilience Moderates the Impact of Social Support on Loneliness of "Left-Behind" Children." *Journal of Health Psychology*, 21(6): 1066–1073.

Allende, Claudio, Juan Diego Luksic, and Nicolas Navarrete. 2024. "Better Together? Student's Benefits of Educational Market Integration." *Journal of Development Economics*, 167: 103233.

Au, Chun-Chung, and J Vernon Henderson. 2006. "How Migration Restrictions Limit Agglomeration and Productivity in China." *Journal of Development Economics*, 80(2): 350–388.

Aydemir, Abdurrahman, Murat G Kirdar, and Huzeyfe Torun. 2021. "The Effect of Education on Geographic Mobility: Incidence, Timing, and Type of Migration." *IZA Discussion Paper*.

Beine, Michel, Frédéric Docquier, and Hillel Rapoport. 2001. "Brain Drain and Economic Growth: Theory and Evidence." *Journal of Development Economics*, 64(1): 275–289.

Beine, Michel, Fréderic Docquier, and Hillel Rapoport. 2008. "Brain Drain and Human Capital Formation in Developing Countries: Winners and Losers." *Economic Journal*, 118(528): 631–652.

Bilal, Adrien, and Esteban Rossi-Hansberg. 2021. "Location as An Asset." *Econometrica*, 89(5): 2459–2495.

Bollinger, Christopher, Xiaozhou Ding, and Steven Lugauer. 2022. "The Expansion of Higher Education and Household Saving in China." *China Economic Review*, 71: 101736.

Bosetti, Valentina, Cristina Cattaneo, and Elena Verdolini. 2015. "Migration of Skilled Workers and Innovation: A European Perspective." *Journal of International Economics*, 96(2): 311–322.

Bound, John, Jeffrey Groen, Gábor Kézdi, and Sarah Turner. 2004. "Trade in University Training: Cross-State Variation in the Production and Stock of College-Educated Labor." *Journal of Econometrics*, 121(1-2): 143–173.

Brown, Richard P C, and Eliana Jimenez-Soto. 2015. "Chapter 20 - Migration and Remittances." In *Handbook of the Economics of International Migration*. Vol. 1, , ed. Barry R Chiswick and Paul W Miller, 1077–1140. North-Holland.

Bryan, Gharad, and Melanie Morten. 2019. "The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia." *Journal of Political Economy*, 127(5): 2229–2268.

Chan, Kam Wing. 2015. "Five Decades of the Chinese Hukou System." In *Handbook of Chinese Migration*. 23–47. Edward Elgar Publishing.

Charles, Kerwin Kofi, Erik Hurst, and Matthew J Notowidigdo. 2018. "Housing Booms and Busts, Labor Market Opportunities, and College Attendance." *American Economic Review*, 108(10): 2947–2994.

Cheng, Tiejun, and Mark Selden. 1994. "The Origins and Social Consequences of China's Hukou System." *The China Quarterly*, 139: 644–668.

Chen, Yan, Ming Jiang, and Onur Kesten. 2020. "An Empirical Evaluation of Chinese College Admissions Reforms through a Natural Experiment." *Proceedings of the National Academy of Sciences*, 117(50): 31696–31705.

Diamond, Rebecca, and Cecile Gaubert. 2022. "Spatial Sorting and Inequality." *Annual Review of Economics*, 14: 795–819.

Ding, Kangzhe, and Ryo Itoh. 2023. "JUE Insight: The Impact of the School Admission Restriction Policy on the Housing Market in Shanghai." *Journal of Urban Economics*, 103568.

Ding, Xiaozhou. 2021. "College Education and Internal Migration in China." *China Economic Review*, 69: 101649.

Edwards, Alejandra Cox, and Manuelita Ureta. 2003. "International Migration, Remittances, and Schooling: Evidence from El Salvador." *Journal of Development Economics*, 72(2): 429–461.

Fabre, Anaïs. 2023. "The Geography of Higher Education and Spatial Inequalities." *Working Paper*.

Fajgelbaum, Pablo D, and Cecile Gaubert. 2020. "Optimal Spatial Policies, Geography, and Sorting." *The Quarterly Journal of Economics*, 135(2): 959–1036.

Fang, Min, and Zibin Huang. 2022. "Migration, Housing Constraints, and Inequality: A Quantitative Analysis of China." *Labour Economics*, 78: 102200.

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

Fan, Jingting, and Ben Zou. 2021. "The Dual Local Markets: Family, Jobs, and the Spatial Distribution of Skills." *Jobs, and the Spatial Distribution of Skills (September 1, 2021)*.

Fan, Jingting, and Lei Li. 2023. "Skill-Biased Imports, Skill Acquisition, and Migration." *IZA Discussion Paper*.

Feng, Shuaizhang, and Xiaoyu Xia. 2022. "Heterogeneous Firm Responses to Increases in High-skilled Workers: Evidence from China's College Enrollment Expansion." *China Economic Review*, 73: 101791.

Fitzpatrick, Maria D, and Damon Jones. 2016. "Post-Baccalaureate Migration and Merit-Based Scholarships." *Economics of Education Review*, 54: 155–172.

Fu, Chao, Junjie Guo, Adam J Smith, and Alan Sorensen. 2022. "Students' Heterogeneous Preferences and the Uneven Spatial Distribution of Colleges." *Journal of Monetary Economics*, 129: 49–64.

Gao, Xuwen, Wenquan Liang, Ahmed Mushfiq Mobarak, and Ran Song. 2023. "Migration Restrictions Can Create Gender Inequality: The Story of China's Left-Behind Children." National Bureau of Economic Research.

Ge, Suqin, Elliott Isaac, and Amalia Miller. 2019. "Elite Schools and Opting In: Effects of College Selectivity on Career and Family Outcomes." *Unpublished Manuscript*.

Gibbons, Stephen, and Anna Vignoles. 2012. "Geography, Choice and Participation in Higher Education in England." *Regional Science and Urban Economics*, 42(1-2): 98–113.

Glaeser, Edward L., JoséA. Scheinkman, and Andrei Shleifer. 1995. "Economic Growth in a Cross-Section of Cities." *Journal of Monetary Economics*, 36(1): 117 – 143.

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

Groen, Jeffrey A. 2011. "Building Knowledge Stocks Locally: Consequences of Geographic Mobility for the Effectiveness of State Higher Education Policies." *Economic Development Quarterly*, 25(4): 316–329.

Grogger, Jeffrey, and Gordon H. Hanson. 2015. "Attracting Talent: Location Choices of Foreign-Born PhDs in the United States." *Journal of Labor Economics*, 33(S1): S5–S38.

Hoogstra, Gerke J., Jouke van Dijk, and Raymond J. G. M. Florax. 2017. "Do Jobs Follow People or People Follow Jobs? A Meta-Analysis of Carlino–Mills Studies." *Spatial Economic Analysis*, 12(4): 357–378.

Howard, Greg. 2020. "The Migration Accelerator: Labor Mobility, Housing, and Demand." *American Economic Journal: Macroeconomics*, 12(4): 147–179.

Howell, Anthony. 2023. "Impact of a Guaranteed Minimum Income Program on Rural–Urban Migration in China." *Journal of Economic Geography*, 23(1): 1–21.

Howell, Anthony. 2024. "Rural Road Stimulus and the Role of Matching Mandates on Economic Recovery in China." *Journal of Development Economics*, 166: 103211.

Huang, Mian, Chunbing Xing, and Xiaoyong Cui. 2022. "Does College Location Affect the Location Choice of New College Graduates in China?" *China & World Economy*, 30(3): 135–160.

Hughes, James, and Margaret Maurer-Fazio. 2002. "Effects of Marriage, Education and Occupation on the Female / Male Wage Gap in China." *Pacific Economic Review*, 7(1): 137–156.

Ishimaru, Shoya. 2023. "Geographic Mobility of Youth and Spatial Gaps in Local College and Labor Market Opportunities." Working Paper.

Kelchtermans, Stijn, and Frank Verboven. 2010. "Program Duplication in Higher Education Is Not Necessarily Bad." *Journal of Public Economics*, 94(5-6): 397–409.

Kennan, John. 2015. "Spatial Variation in Higher Education Financing and the Supply of College Graduates." *NBER Working Paper*.

Kerr, Sari Pekkala, William Kerr, Çağlar Özden, and Christopher Parsons. 2017. "High-Skilled Migration and Agglomeration." *Annual Review of Economics*, 9: 201–234.

Kong, Dongmin, Bohui Zhang, and Jian Zhang. 2020. "Higher Education and Corporate Innovation." *Unpublished Manuscript*.

Li, Bingkun. 1993. "Retrospect and Prospect of TVEs during the Fifteen Years of Reform and Opening Up." *Management World*, 5: 156–165.

Li, Shi, John Whalley, and Chunbing Xing. 2014. "China's Higher Education Expansion and Unemployment of College Graduates." *China Economic Review*, 30: 567 – 582.

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

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

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

Ma, Xiao. 2023. "College Expansion, Trade, and Innovation: Evidence from China." *International Economic Review*.

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

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

Notowidigdo, Matthew J. 2020. "The Incidence of Local Labor Demand Shocks." *Journal of Labor Economics*, 38(3): 687–725.

Roca, Jorge De La, and Diego Puga. 2017. "Learning by Working in Big Cities." *The Review of Economic Studies*, 84(1 (298)): 106–142.

Rong, Zhao, and Binzhen Wu. 2020. "Scientific Personnel Reallocation and Firm Innovation: Evidence from China's College Expansion." *Journal of Comparative Economics*, 48(3): 709 – 728.

Schwartz, Aba. 1976. "Migration, Age, and Education." *Journal of Political Economy*, 84(4).

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

Tao, Li, Eddie C.M. Hui, Francis K.W. Wong, and Tingting Chen. 2015. "Housing Choices of Migrant Workers in China: Beyond the Hukou Perspective." *Habitat International*, 49: 474–483.

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

Topel, Robert H, and Michael P Ward. 1992. "Job Mobility and the Careers of Young Men." *The Quarterly Journal of Economics*, 107(2): 439–479.

Wang, Xiufang. 2010. *Education in China since 1976.* McFarland.

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

Winters, John V. 2011. "Why Are Smart Cities Growing? Who Moves and Who Stays." *Journal of Regional Science*, 51(2): 253–270.

Winters, John V. 2020. "In-State College Enrollment and Later Life Location Decisions." *The Journal of Human Resources*, 55(4): 1400–1426.

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

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

Wu, Xueping, and Yiqun Ye. 2018. "An Introduction to Technical and Vocational Education in China." *Technical and Vocational Education in China*, 1–43. Singapore:Springer Singapore.

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

Yang, Dean. 2008. "International Migration, Remittances and Household Investment: Evidence from Philippine Migrants' Exchange Rate Shocks." *The Economic Journal*, 118(528): 591–630.

Yang, Yu. 2021. "Place-Based College Admission, Migration and the Spatial Distribution of Human Capital: Evidence from China." Unpublished Manuscript.

Young, Jason. 2013. *China's Hukou System.* Palgrave Macmillan UK.

Zheng, Liang. 2016. "What City Amenities Matter in Attracting Smart People?" *Papers in Regional Science*, 95(2).

Zheng, Siqi, Wanyang Hu, and Rui Wang. 2016. "How Much Is A Good School Worth in Beijing? Identifying Price Premium with Paired Resale and Rental Data." *The Journal of Real Estate Finance and Economics*, 53: 184–199.

Appendix

A.1 Sample Selection and Data Assembling

To compile the final dataset for our analysis, we must align each individual with the corresponding city-level and province-level attributes of their birthplaces. Given that population censuses report an individual's birth province, further matching province-level characteristics with individuals is straightforward.

However, identifying an individual's birth city is challenging. The only city-level location information we know is one's current city of residence and the city of *hukou* registration. Note that our interest lies in geographical mobility so we must exclude *hukou* migration cases where one has had a change in *hukou* location but this may not correspond to geographical migration.

We employ a specific procedure to ascertain an individual's birth city, drawing on information about their birthplace and *hukou* registration. The survey question about one's birthplace presents three options: (1) The current county of residence; (2) A different county within the current province of residence (county code is unknown); (3) A different province other than the current one of residence (province code is known).

1. When the response to the aforementioned survey question is (1), an individual's birthplace aligns with their current province and city of residence. In this case, their birth city is identified as the city of their *hukou* registration.

2. When the response to the aforementioned survey question is (2), these individuals were born in their current province of residence but in a different city.

   (a) When an individual's *hukou* registration city differs from their current city of residence, we identify their birth city as the city of their *hukou* registration. This applies to a total of 60,622 individuals. We hypothesize that these individuals have relocated within their birth province without changing their *hukou* since birth, a scenario most likely among young adults. As our study is centered on cross-province geographical migration, these individuals are not classified as (cross-province) migrants in our analysis.

   (b) However, for individuals who have registered their *hukou* in the current city of residence, their birth city remains undetermined. Consequently, these individuals are excluded from the study. This group comprises 8,133 individuals.

3. When the response to the aforementioned survey question is (3), an individual's birth province differs from their current province of residence.

   (a) When an individual's *hukou* registration province matches their birth province, we infer that their birth city is the city of their *hukou* registration. It is important to note that an individual might originate from a city within their birth province that is different from their *hukou* registration city. However, such instances are uncommon among young adults.

   (b) If an individual's *hukou* registration is within the same province as their current residence, it implies that the individual has relocated and subsequently updated their *hukou* registration. As a result, the city of their *hukou* registration cannot be considered

their birth city. We have identified 1,601 such cases. Due to the inability to determine their birth city, these individuals have been omitted from our sample.

The selection procedure outlined above yields a sample of 63,498 individuals. Collectively, these individuals constitute 18.9 percent of the census population aged between 23 and 26.

A.2 The Expected Wage Premiums for Migration

Figures A.1 and A.2 map the spatial distribution of estimated wage differentials for two groups upon migration: non-college-educated (Δw_j^L) and college-educated individuals (Δw_j^H). Darker shades indicate higher migration wage differentials. The figures show similar patterns, with individuals from gridlocked provinces expecting the most significant wage gains from migration. In contrast, those from Beijing, Shanghai, Guangdong, and coastal provinces anticipate minimal or even negative wage gains from migration.

The figures also reveal a notable difference. High wage premiums for migration for non-college-educated workers are predominantly found in western and central China, excluding provinces like Shanxi and Inner Mongolia which are rich in resources. For those with a college education, provinces with high wage premiums are only found in central China, suggesting potential subsidies in western provinces to retain such workers.

A.3 College Expansion and College Education Choice: OLS Estimates

In this section, we initiate the first step of our two-step IV approach by utilizing in- and out-of-province college expansion shocks to predict individual decisions regarding college attendance and enrollment location using linear probability estimation. The results are presented in Table A.1. In column 1, we employ the college attendance status, $College$, as the outcome variable, while in columns 2 and 3, we utilize the out-of-province college attendance, $CollegeOut$, as the outcome variable. Columns 1 and 2 use the full sample; Column 3 focuses on individuals with a college education. The results align with those obtained from the probit model estimation presented in Table 3.

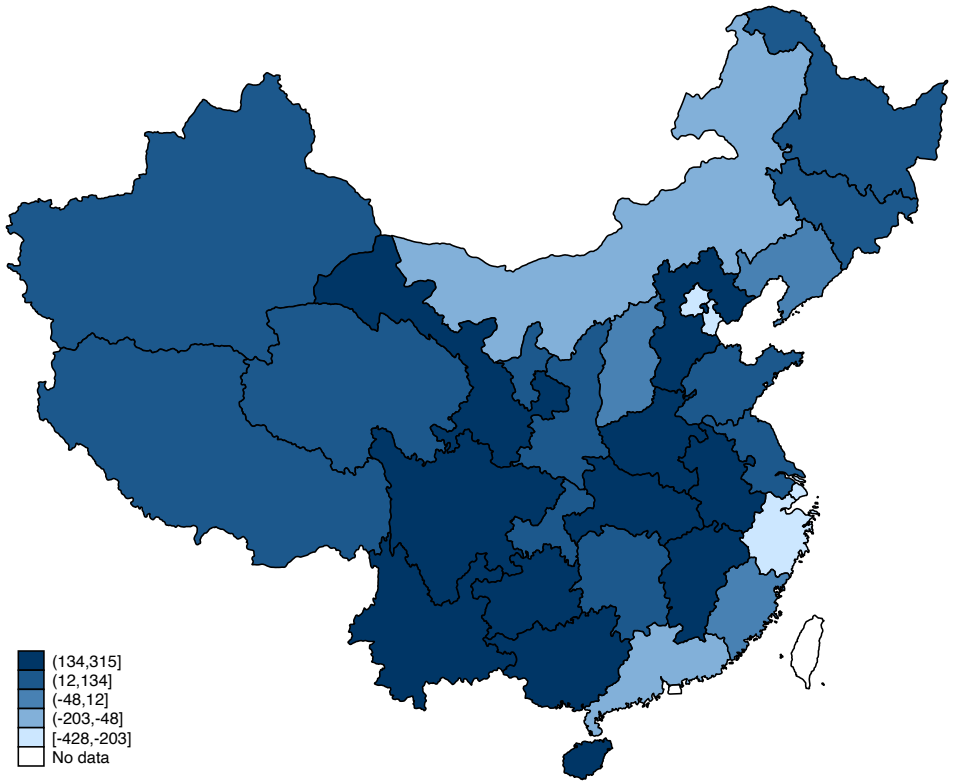

Figure A.1: The expected migration wage differentials for low-skilled workers (Δw_j^L)

Notes: This map shows the difference in the monthly earnings (yuan) between the home province j and the expected wage if an individual moves out of that province. We first calculate the share of low-skilled migrants (non-college-educated) from the home province in different destination provinces, $\frac{M_{jk}^L}{M_j^L}$. We then calculate the expected migration wage by averaging the destination low-skilled workers' wages weighted by migrant shares $E(w_j^L|\text{migrate}) = \sum_{k\neq j} w_k^L \times \frac{M_{jk}^L}{M_j^L} - w_j^L$. Lastly, we calculate the difference between the two, $\Delta w_j^L = \sum_{k\neq j} w_k^L \times \frac{M_{jk}^L}{M_j^L} - w_j^L$. All wages are from the 2005 population census.

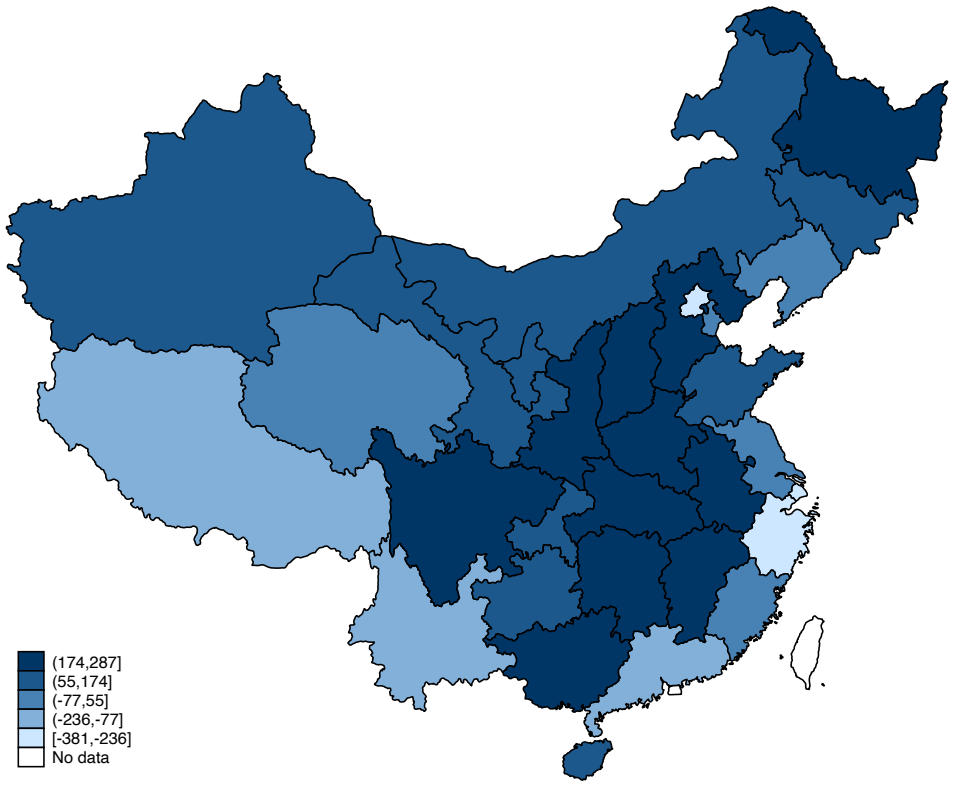

Figure A.2: The expected migration wage differentials for high-skilled workers (Δw_j^H)

Notes: This map shows the difference of the monthly earnings (yuan) between the home province j and the expected wage if an individual moves out of that province. We first calculate the share of high-skilled migrants (college-educated) from the home province in different destination provinces, $\frac{M_{jk}^H}{M_j^H}$. We then calculate the expected migration wage by averaging the destination high-skilled workers' wages weighted by migrant shares $E(w_j^H|\text{migrate}) = \sum_{k\neq j} w_k^H \times \frac{M_{jk}^H}{M_j^H} - w_j^H$. Lastly, we calculate the difference between the two, $\Delta w_j^H = \sum_{k\neq j} w_k^H \times \frac{M_{jk}^H}{M_j^H} - w_j^H$. All wages are from the 2005 population census

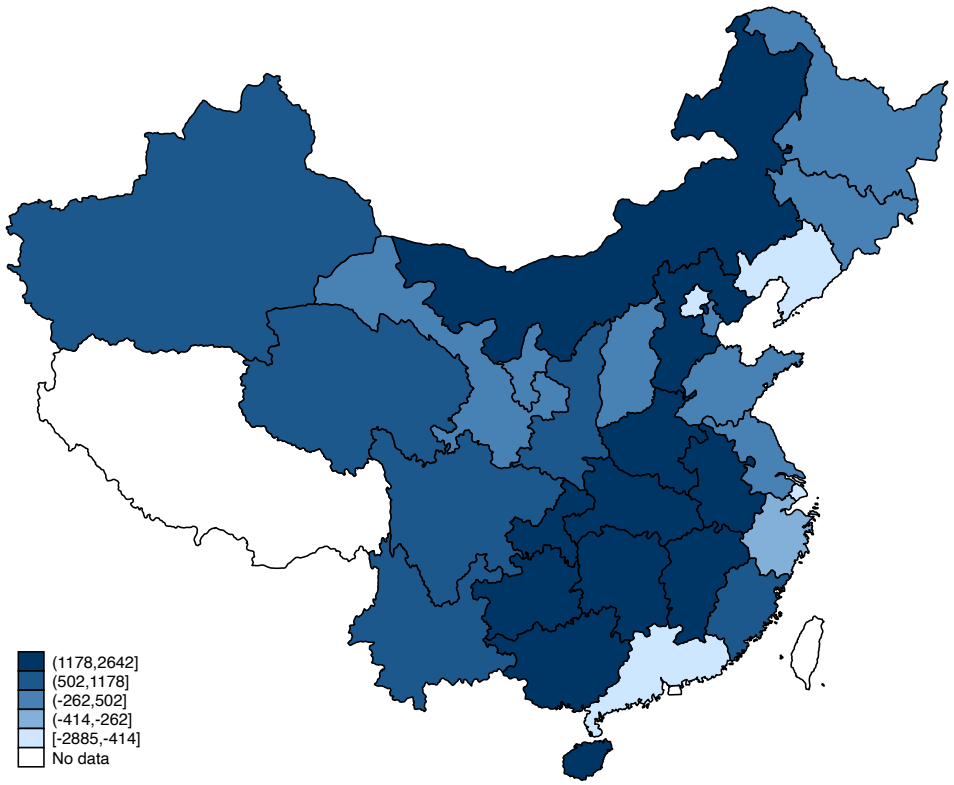

Figure A.3: The spatial distribution of average housing price differentials (Δhp_j), 2002-2005

Notes: This map shows the spatial distribution of the difference in the average commodity housing sale price per square meter (yuan/m^2) between the home province and outside the home province. The details about the construction of an average migration destination province's housing price are discussed in Section 6. The author's calculation is based on the city housing price data from the National Bureau of Statistics and the 2005 population census.

Table A.1: The effect of college expansion on college attendance choices, OLS estimates

	(1) $College$	(2) $CollegeOut$	(3) $CollegeOut$
In-province enrollment rate	0.115	-0.0214	-0.0542
	(0.0759)	(0.0177)	(0.0383)
Out-province enrollment rate	0.102	0.224*	0.407
	(0.205)	(0.123)	(0.277)
Han ethnicity	-0.0306***	-0.00139	0.00257
	(0.00939)	(0.00340)	(0.00612)
Male	-0.0527***	0.00814***	0.0283***
	(0.00605)	(0.00226)	(0.00526)
Household size	-0.0311***	-0.000592	0.00599***
	(0.00223)	(0.000581)	(0.00192)
Age	0.212	-0.125*	-0.331**
	(0.180)	(0.0707)	(0.136)
Age square	-0.00444	0.00249*	0.00663**
	(0.00371)	(0.00143)	(0.00274)
City population	-0.00451**	5.63e-05	0.00151
	(0.00170)	(0.000931)	(0.00175)
City GDP per capital	0.0401***	-0.00424	-0.0111*
	(0.00480)	(0.00283)	(0.00647)
City tertiary employment share	0.00154**	0.000343	0.000528
	(0.000688)	(0.000236)	(0.000626)
Distance to coast	-0.00864*	0.000803	0.00398
	(0.00436)	(0.00230)	(0.00443)
Province urban employment	-0.140	-0.0974	-0.0731
	(0.354)	(0.0802)	(0.146)
Province GDP per capita	-0.0260	0.0151**	0.0163
	(0.0182)	(0.00622)	(0.0127)
Capital city	0.0478***	-0.0183***	-0.0504***
	(0.0139)	(0.00601)	(0.0131)
Municipality	-0.00671	0.0203	0.0124
	(0.0896)	(0.0852)	(0.101)
Province GDP growth rate	0.0649	0.0362	0.0821
	(0.0831)	(0.0306)	(0.0601)
Province employment growth rate	-0.00143	-0.0160	-0.0294
	(0.0614)	(0.0109)	(0.0213)
Observations	63,498	63,498	31,380
R^2	0.057	0.018	0.053

Note: This table shows the OLS results of in- and out-of-province enrollment rate on college attendance and out-of-province college attendance using the 2010 census. The dependent variable in column 1 is a dummy variable indicating whether an individual attended college. Dependent variables in columns 2 and 3 are dummy variables indicating whether an individual attended college outside the birth province. Column 2 includes the full sample and column 3 uses a subsample of college graduates only. All regressions control for birth-province, birth-cohort-year, and birth-region-by-cohort fixed effects and are weighted by the number of high school graduates at the province level. Robust standard errors are clustered by birth province. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.