Educational Migration

Naiyuan Hu [*] Lin Ma[*]

June 20, 2024

Click here for the latest version

Abstract

Educational resources are distributed unevenly across space and could contribute to spatial inequality. We develop a dynamic spatial model with life-cycle elements to study the impacts of location-specific educational resources. In the model, individuals determine whether and where to attend college, weighing on the distance to home, the expected option value of education, and the educational resources in the destination. Locations with more colleges attract more students. Moreover, as mobility costs increase with age, many college graduates stay in the city of their alma mater, leading to long-term changes in skill composition. We quantify the model to the context of China and structurally estimate the cost of obtaining a college degree in each location. We show that the college expansion between 2005 and 2015 had minimal impacts on welfare and skill composition, as it diverts resources towards the locations already well-endowed with colleges. More evenly distributed colleges could improve aggregate welfare and reduce spatial inequality at the same time.

Keywords: Spatial Economics; Economic Geography; Migration; Life-cycle
JEL codes: F12; O11; R12

[*]School of Economics, Singapore Management University

1 Introduction

The spatial distribution of colleges is highly uneven within a country. For example, in the context of China, educational hubs such as Beijing could host as many as 77 universities, while the average prefecture only has 6. Even worse, around a quarter of the cities, many with millions of population, have no more than a single college. The uneven distribution of educational resources could lead to dire consequences for both individuals and society. Access to colleges shapes the fate of young people: those born in unlucky locations with scarce resources must endure the ordeal of long-distance migration at a young age to seek a higher education; deterred by such costs, many talented students forgo the opportunities and remain as unskilled workers throughout their lifetime. The impacts of colleges do not stop at the student population either. Seeking higher education is one of the main motivations for migration (see, e.g., Dustmann and Glitz, 2011). In fact, it is the *only* migration spell for many college graduates in their lifetime: they settle in the city of their alma mater. As a result, access to educational resources could influence one's lifetime location choices. At the aggregate level, the forces mentioned above affect the skill composition of a location in the long run, exerting their impacts on skill premium, population, and economic prosperity for many years to come. To what extent is the uneven distribution of educational resources responsible for the observed spatial inequality? What is the optimal distribution of colleges across space? The answers to these questions not only arouse academic interests but also carry long-lasting policy implications. Answering these questions, however, requires careful modeling of individuals' educational and location choices over one's life cycle. Such a model is currently lacking in the literature, and we seek to fill the gap with this paper.

We developed a multi-sector general equilibrium dynamic spatial model with life-cycle elements to analyze the geography of educational resources. The model consists of overlapping generations of individuals that live for many periods. Upon entry into the model, individuals endogenously choose education levels, weighing the expected return to higher education against the costs of doing so. Conditional on seeking higher education, they then determine where to attend college. The locational choice of students depends on the distance to home, the option value of being a skilled worker in the destination, the opportunity costs of obtaining a degree, and non-monetary location characteristics such as the abundance of college resources. Upon graduation, individuals enter the labor market as skilled workers. The young workers who forgo higher education directly enter the labor market as unskilled workers. In each period, skilled and unskilled workers supply labor, consume, and move to other labor markets, subject to age-specific migration frictions throughout their life cycle until retirement and exit from the model.

Educational resources exert their long-run spatial impacts through several channels in the model. First, locations well-endowed with educational resources increase the appeal of higher education, encouraging more skill upgrading among local students. Moreover, the abundance of colleges also attracts potential students from all over the country, particularly individuals in nearby locations with relatively low migration costs. The implications of educational resources are also long-lasting. Given the considerable migration costs, many college students choose to stay in the location of their alma mater throughout their lifetime, pushing up the skill ratio of these locations persistently and contributing to local productivity through agglomeration forces. Lastly, the distribution of colleges also affects unskilled workers through general equilibrium effects: a relative abundance of skilled workers in one area pushes up the demand for unskilled workers, indirectly benefiting them.

We quantify and estimate the model in the context of China, a country with highly concentrated educational resources. The large spatial variations in educational resources are particularly valuable econometrically, as they allow us to structurally estimate the locations' attractiveness as educational migration destinations through the lens of our model. Moreover, China also offered an interesting policy experiment. Along with rapid economic development and urbanization, China initiated a large-scale college expansion program rarely seen worldwide. Between 2005 and 2015, the spending on college education increased by 466 percent, and the number of college teachers expanded by 84 percent. As a result, college enrollment increased from 5 million to around 14 million during a short period. We use our model to evaluate the aggregate and distributional impacts of the policy change; we also carry out counterfactual simulations to study if better aggregate and distributional results could be achieved with alternative distribution of educational resources.

To understand the spatial impacts of college distribution, the core empirical question is how the distribution of colleges affects educational migration — the destination choices of potential students. Our model offers a useful lens through which we can structurally map the observed educational resources, such as the number of college teachers and the quality of colleges, to the observed educational migration patterns. In particular, the predicted migration probability matrix of the students summarizes the attractiveness of a location to college seekers, which depends on the underlying geography, migration policy, expected option value, and educational resources. Controlling all the other elements shaping the migration probability, our model then provides a natural mapping from location-specific educational resources to the observed educational migration patterns. With some functional form assumptions, we can then infer a location's attractiveness for potential college students, which we call the "education appeals" of a location. We assume that the education appeals as a function of the quantity and quality of colleges, as well as natural amenities such as climate and terrain, and estimate these parameters using non-linear least squares. The estimated education attractiveness suggests diminishing returns to college concentration. In the city with a median level of college teachers, a 10 percent increase in teachers leads to a 2.87 percent increase in education appeals. However, the return to more colleges quickly recedes in better-endowed locations. For example, in the 90th percentile of cities, a 10 percent increase in teachers only increases its attractiveness by one-eighth of the same increment at the 10th percentile.

We find that college expansion between 2005 and 2015 led to a mild increase in the welfare and skill ratios at the aggregate level. To evaluate the effects of college expansion, we compare the baseline simulation of a transition path with the observed expansion to a counterfactual one without the college expansion. The comparison suggests that aggregate welfare increased by 0.52 percent and the aggregate skill ratio by 0.56 percent by 2015. Both skilled and unskilled workers benefited from the college expansion, while the average welfare gain among the skilled workers (0.82 percent) is substantially larger than that among the unskilled ones (0.24 percent). The college expansion increased the skilled ratio in most locations, leading to higher productivity, thus benefiting the unskilled workers. The gain in productivity is particularly rewarding to skilled workers through productivity-skill complementarity in the model. Eventually, it outweighs the negative impacts of an increased supply of college-educated workers on welfare, leading to considerable gains for college-educated workers. The mild response to college expansion comes as no surprise. The expansion program favored locations already well-endowed with educational resources while leaving the initially poorly-endowed locations behind. Given the high curvature of the education appeal function, the accessibility to higher education, especially in poor and remote locations, barely changed.

How can we better allocate educational resources across space? We answer this question in several ways. We first compute the welfare elasticity of college expansion prefecture-by-prefecture and find that the return to college investment is substantially higher in poorly endowed locations. The aggregate return to a 10% increase in college teachers is 2.6 times higher at the bottom 10% of the prefectures than those at the top 10%. We then simulate another counterfactual in which we allocate all the additional educational resources observed in the actual expansion program *equally* among all the prefectures. In this case, all prefectures receive an additional 2600 college teachers. We find that the aggregate welfare effects of this simple "equal growth" scheme are roughly 1.8 times larger, and the impacts on skill ratio more than 1.94 times higher than the observed expansion at 1.01%.

Lastly, we show that the unequal distribution of educational resources is responsible for up to 14% of the observed spatial inequality in skill composition. Moreover, equalizing educational resources is roughly 14% as effective as equalizing fundamental productivity in reducing spatial inequality in the short run. To understand the impact of college distribution on the observed spatial inequality, we first simulate one counterfactual called "equal college", in which we eliminate all spatial variation in educational resources and redistribute the existing stock equally across locations. Compared to the baseline result, the skill ratio dispersion across locations drops by around 3.74% to 0.47% along the transition path towards the long-run steady state. To benchmark the effect of equalization of educational access, we compute another counterfactual in which the location fundamental productivity, the usual culprit of spatial inequality (and structural residual that absorbs locational differences) in the quantitative models, is equalized across locations. In the productivity-equalizing world, the dispersion of skill ratio declines by 26.33% to 84.8% along the transition path to the steady state, as compared to the baseline model. In this sense, an evenly distributed educational resource is 14% as effective in reducing spatial inequality as an evenly distributed productivity in the short run, however, in the long run the skill ration inequality will be largely driven by the inequality in the productivity.

This paper is related to several strands of the literature. Firstly, our study is closely related to a broad literature on quantitative spatial and dynamic discrete choice models, such as Artuç et al. (2010); Allen and Arkolakis (2014); Ahlfeldt et al. (2015); Caliendo et al. (2019, 2021); Kleinman et al. (2023), as surveyed in Redding and Rossi-Hansberg (2017). We contribute to this literature in several ways. We are the first to introduce educational choice into the dynamic spatial framework to highlight the interlinkages between educational resources, transportation infrastructure, and geographical fundamentals. We show that the distribution of colleges not only directly affects student distribution but also shapes long-term skill composition in general equilibrium. We also introduce a structural interpretation to map the observed student distribution to the educational resources; through our estimation, we highlight the curvature of the education cost function, which could inform policy-making and researchers interested in educational issues.

Our research relates to the recent dynamic models with life-cycle assumptions, including Allen and Donaldson (2022), Eckert and Peters (2022), Takeda (2022), Komissarova (2022), and Suzuki (2023). We contribute to this literature by modeling and examining the impact of inter-generational linkages on decisions made over one's lifespan, specifically focusing on education attainment. We are also the first to estimate age-specific migration costs using these models, and our results, as discussed later, reveal a shape increase in migration costs as people age.

Our work is also broadly related to the quantitative spatial works that focus on China, such as Tombe and Zhu (2019), Fan (2019), Ma and Tang (2020). While this strand of work focuses on elements specific to the context of China, such as the hukou restrictions, we are among the first to introduce a dynamic structure to study the migration decisions over one's lifetime. In contrast to Fan (2019), where skill type is taken as given, and migration is primarily driven by wage incentives and amenities, our approach introduces a nuanced perspective.

Finally, a few works have studied the impact of college expansion, including the impacts on innovation and trade (Ma, 2023), the impacts on knowledge spillovers (Li et al., 2020), and the impact on human capital on productivity (Che and Zhang, 2018). This paper contributes to the literature on the impacts of college expansion by estimating the education cost in each location and simulating the counterfactual college expansion scenarios and their economic impacts. Specifically, we estimate the education cost of China in all prefectures using a detailed migration matrix and uncover the highly uneven educational costs that lead to a suppression of the talents who can otherwise become skilled workers.

The remainder of this paper proceeds as follows. In Section 2, we present the quantitative spatial model. In Section 3, we calibrate the model parameters. In particular, we estimate the education cost base on the endogenous education migration model. In Section 4, we present the counterfactual simulations. In Section 5, we conclude.

2 Model

We develop a dynamic spatial life-cycle model emphasizing individuals' endogenous education choices. The model allows us to examine how the spatial distribution of educational resources affects individuals' migration choices, the composition of the labor market in each region, and its overall impact on inequality over time. Our model extends the quantitative model introduced by Caliendo et al. (2019) to cover heterogeneous cohorts and endogenous education choices.

2.1 Life-Cycle and Individual Decisions

The economy has $n = 1, 2, \cdots, N$ locations, $s = 1, 2, \cdots, S$ sectors, inhabited by overlapping generations of individuals that live up to J periods. Each cohort comprises $\tilde{L}$ individuals, and the total population in this economy equals $J\tilde{L}$. Each period t, individual in location n derive flow utility from consuming a composite good from all locations:

$$U_t(n) = \log\left[C_t(n)\right], \quad \text{where: } C_t(n) = \left(\sum_{i=1}^{N}\left[\prod_{s=1}^{S}\left(c_t^s(n,i)\right)^{\zeta^s}\right]^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}, \quad \sigma > 1, \tag{1}$$

where $C_t(n)$ is the consumption consumption, which is a nested composite good across all locations and sectors. The variable $c_t^s(n, i)$ denotes the goods from sector s produced in location i and consumed at location n. The parameter σ captures the elasticity of substitution of goods across locations, and $\zeta^s > 0$ and $\sum_{s=1}^{S} \zeta^s = 1$ captures the weight of sector s in the consumption bundle.

Timeline and Decisions $\tilde{L}$ newborn individuals enter the model every period. The spatial distribution of the newborn is the same as the old cohort exiting the model to ensure the stability of population distribution in a steady state. At the start of one's lifetime, an individual first observes an idiosyncratic education preference shock, z, and, based on which, decides whether to pursue higher education and become a skilled worker. Individuals maximize the expected lifetime utility at the birthplace n by choosing their skill type, $e = \{l, h\}$, where "l" stands for unskilled ,and "h" for skilled workers:

$$V_t^0(n) = \max_{\{e\}}\{V_t^0(n,e) + \psi z\}. \tag{2}$$

In this expression above, $V_t^0(0)$ is the unconditional lifetime utility of a newborn in location n at time t, and $V_t^0(n,e)$ is the lifetime utility conditional on skill type e. The education preference shock, z, can be equally interpreted as heterogeneity in one's ability to learn — individuals with a higher z are innately better at learning and, therefore, have higher probabilities of attending college. We assume z follows a standard GEV-I distribution with zero mean and use ψ to capture the (inverse) elasticity of education choice that will be estimated later. Individuals can only choose education levels at the beginning of their lifetime, and those who forgo higher education will stay as unskilled workers for their entire lifetime.[1]

Conditional on pursuing higher education ($e = h$), the individual chooses the optimal location n' for college. For an individual born in location n, the decision is summarized as:

$$V_t^0(n,h) = E_{\epsilon_n} \max_{\{n'\}}\{V_t^1(n',h) - D^1_{n'n,t} + F_{n',t} + \kappa\epsilon_{n'}\}. \tag{3}$$

The location choice hinges on four factors. The first factor is the distance between their home location and the destination, $D^1_{n'n,t}$, which captures the roles of geography and transportation networks in education choice. The second factor is the option value of being a skilled worker in n', $V_t^1(n',h)$. The option value, which we will discuss in detail later, reflects two forces: the current wage for skilled workers in the destination and, recursively, the value of moving to other high-paying locations in the future. Everything else being equal, the potential students prefer to attend colleges in or close to locations with high-paying jobs for skilled workers to save on post-graduation migration costs. The third factor that affects one's location choice is the destination-specific "education appeals", denoted as $F_{n',t}$. We use this term to capture the non-monetary and unobserved appeals of a location for higher education. Conditional on a home location, two destinations with the same distance and option value might still differ in their attractiveness to students due to the abundance of college resources, the prestige of their colleges, and natural amenities such as temperature and precipitation. Lastly, as usual in the dynamic discrete choice literature, the location choice also depends on a vector of i.i.d idiosyncratic shocks, $\{\epsilon'_n\}_{n'=1}^N$ that follows a Type-I GEV shock with zero mean, and κ is the inverse of the migration elasticity.

The option value of a skilled worker in location n' at age $j = 1, 2, \cdots, J$, is defined recursively

[1] In the data, continuing education plays a minor role in China. According to the 2015 Annual Report of the Ministry of Education, only 292 out of 2,852 colleges in China offer continuing education. The number further declined to 265 in 2020.

as:

$$V_t^j(n,h) = U_t^j(n,h) + E_{\epsilon_{n'}} \max_{n'} \left\{ \beta V_{t+1}^{j+1}(n',h) - D_{n'n,t+1}^{j+1} + \kappa \epsilon_{n'} \right\}, \tag{4}$$

The above formulation is the same as a standard dynamic migration model: individual workers migrate every period to maximize the expected lifetime utility, considering distance, future payoff, and idiosyncratic location shocks.

We highlight two new features in our model. First, we directly model the opportunity cost of attending college. At $j = 1$, a college student cannot supply labor and relies on "home production" for consumption. The "home production" can be broadly interpreted as part-time and substance-level jobs that do not interfere significantly with the labor market clearing conditions or financial transfers from the government or family that we abstract away from the model. After graduation at $j > 1$, the educated worker earns the skilled wage rate. In summary, the flow indirect utility function for a skilled individual is as follows:

$$U_t^j(n,h) = \begin{cases} b_{nt} & \text{if } j = 1 \\ \ln w_{nt}^h - \ln P_{nt}, & \text{otherwise} \end{cases}, \tag{5}$$

where b_{nt} is the utility derived from home production, w_{nt}^h is the skilled wage, and P_{nt} is the ideal price index at location n, time t. The second new feature is directly linked to the life-cycle elements of the model. In the last period of an individual's life cycle at $j = J$, their value functions no longer include any future option and, therefore, simplify to:

$$V_t^J(n,h) = U_t^J(n,h). \tag{6}$$

The recursive problem of the unskilled workers ($e = l$) adopts a similar and simplified recursive structure following equations (4), (5) and (6), with three changes: 1) the skill type is no longer h but l, and 2) equation (4) applies to $j = 0, 1, \cdots, J$, due to the absence of college choice in equation (3), and 3) the unskilled workers start supplying labor at age $j = 1$ (when the skilled workers at the same age attend college), and therefore their flow utility is:

$$U_t^j(n,l) = \ln w_{nt}^l - \ln P_{nt}, \forall j \tag{7}$$

where w_{nt}^l is the wage for unskilled workers.

Figure 1 summarizes the timeline for individuals in the model. Individuals enter the model at age 18 and exit at age 61, spanning 11 periods with 4-year intervals. In the first period, an agent born in location n draws a preference for education and chooses whether to become a skilled

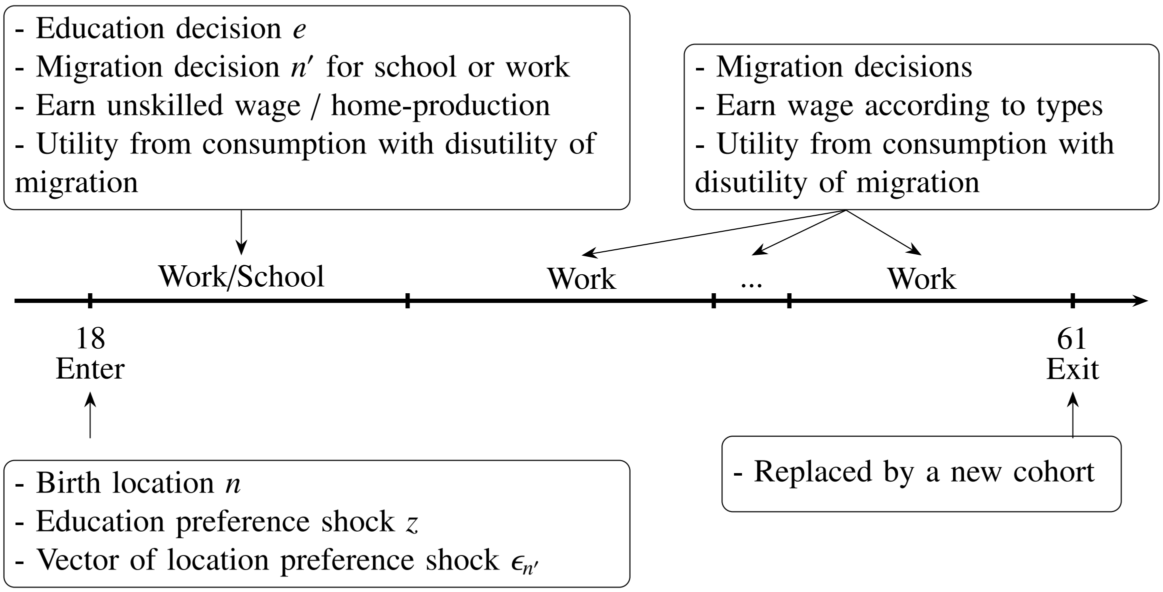


Figure 1: Decision Timeline

Note: This figure shows the timeline of individual decisions. A detailed description can be found in the text.

worker. Conditional on education, the agent draws preference shocks for all locations and selects a place to attend college. Those who opt out of college become unskilled workers. After $j = 1$, skilled and unskilled workers make similar recursive migration decisions until they exit the model at $j = J$.

Solving the Individual Decisions Under the assumption that both the idiosyncratic shocks of the location taste $\epsilon \geq 0$ and education taste $z \geq 0$ coming from type-I extreme distributions, we can solve the aggregate migration flow and the probability of educational choice.The expected lifetime value for worker with skill $e \in \{l, h\}$ being in location n, in period j, and $1 < j < J$, is given by,

$$V_t^j(n, e) = U_t^j(n, e) + \kappa \log \sum_{n'=1}^{N} \exp\left(\beta V_{t+1}^{j+1}(n', e) - D_{n'n,t+1}^{j+1}\right)^{1/\kappa}. \tag{8}$$

Therefore, the probability of the migration decisions for individuals at period j, with education level e, live in location n, moving to location n' is given by,

$$\pi_t^j(n'|n, e) = \frac{\exp(\beta V_{t+1}^{j+1}(n', e) - D_{n'n,t+1}^{j+1})^{1/\kappa}}{\sum_{n''=1}^{N} \exp(\beta V_{t+1}^{j+1}(n'', e) - D_{n''n,t+1}^{j+1})^{1/\kappa}}, \tag{9}$$

At the very beginning of one's lifetime, the probability of an individual's migration and educational choice can be written as

$$
\begin{aligned}
\pi_t^0(n', e|n) &= (1) \times (2) \\
(1) &= \frac{\exp(V_t^1(n', e) - D^1_{n\prime n,t} - 1_{e=h}F_{n't})^{1/\kappa}}{\sum_{n''=1}^{N} \exp\left(V_t^1(n'', e) - D^1_{n''n,t} - 1_{e=h}F_{n''t}\right)^{1/\kappa}} \\
(2) &= \frac{\exp(V_t^0(n, e))^{1/\psi}}{\sum_{e'} \exp\left(V_t^0(n, e')\right)^{1/\psi}}.
\end{aligned}
\tag{10}
$$

The first part is the location choice conditional on education decision, e. The second part summarizes the decision to study and become a skilled laborer later in life, given the birthplace n.

Lastly, migration probabilities and the initial distribution summarize the population movement and its composition. Formally, the migration flows are expressed as:

$$
\begin{aligned}
L^0_{nt} &= L^J_{nt-1} \\
L^1_t(n', e) &= \sum_n L^0_{nt}\pi^0_t(n', e|n) \\
L^j_t(n', e) &= \sum_n L^{j-1}_{t-1}(n, e)\pi^{j-1}_{t-1}(n'|e, n) \quad j = 2, ..., J \\
\mathbb{L}_t(n, e) &= \sum_{j=2}^{J} L^j_t(n, e) + \mathbb{1}_{e=l}L^1_t(n, e).
\end{aligned}
$$

In the set of equations above, the first one means that the newborn population in each location equals the retiring population in the previous period at the same location. The second equation describes the educational migration flows, and the third one describes the worker's movement. The last equation computes the total population across ages by location and education level.

2.2 Production, Aggregation, and Closing the Model

Production The production side in the model follows the Armington (1969) framework. We assume that each location specializes in producing goods in S sectors. Consumers demand all varieties of goods from all locations and sectors. In each location and sector, a perfectly competitive market prevails. Production technology is assumed to have a constant elasticity of substitution and requires both skilled and unskilled workers. The output in location n, sector s is given by:

$$
y^s_{nt} = A^s_n \left\{ \chi^{s\frac{1}{\eta}} \left[(A^s_n)^{-\omega} l^l{}_{nt} \right]^{\frac{\eta-1}{\eta}} + (1 - \chi^s)^{\frac{1}{\eta}} \left[(A^s_n)^{\omega} l^h_{nt} \right]^{\frac{\eta-1}{\eta}} \right\}^{\frac{\eta}{\eta-1}},
$$

where l^l_{nt} and l^h_{nt} are unskilled and skilled labor inputs respectively, to produce y^s_{nt} units of good. The parameter $A^s_n = \bar{A}^s_n \mathbb{L}_t(n,h)^\phi$ denotes the endogenous productivity in location n for sector s, where $\bar{A}^s_n$ is the fundamental productivity, $\mathbb{L}_t(n,h)$ is the total number of working skilled labor, and the parameter ϕ governs the agglomeration forces. We allow productivity-skill complementarity in the production process, denoted by ω, as introduced by Burstein and Vogel (2017). Productivity-skill complementarity implies that high-skilled labor is more effective in high-productivity locations.[2] The parameter χ^s is the share of input from skilled in production. The parameter η is the elasticity of substitution of skilled and unskilled worker in the production process. Workers can freely move across sectors. Wage in each location for each skill level, denoted as w^e_{nt}, is determined endogenously through local labor market clearing.

Market Clearing Conditions From the utility function, it is straightforward to see that the ideal price index can be expressed as

$$P_{nt} = \prod_{s=1}^{S} \left(\frac{\left(\sum_{i=1}^{N} p_{nit}^{1-\sigma} \right)^{\frac{1}{1-\sigma}}}{\zeta^s} \right)^{\zeta^s}$$

where p^s_{nit} is the price of goods in sector s purchased from location i for consumption in location n at time t. The Armington assumptions directly imply that p^s_{nit} equals the marginal cost of producing times the iceberg transportation cost τ_{ni} at time t,

$$p^s_{nit} = \tau_{ni} \frac{1}{A^s_i} \left[(A^s_i)^{-\omega(\eta-1)} \chi^s \left(w^l_{it} \right)^{1-\eta} + (A^s_i)^{\omega(\eta-1)} (1-\chi^s) \left(w^h_{it} \right)^{1-\eta} \right]^{\frac{1}{1-\eta}}.$$

Subsequently, conditional on sector s, the share of income at location n spent on goods supplied by location i at time t is given by:

$$\alpha^s_{nit} \equiv \frac{(p^s_{nit})^{1-\sigma}}{\sum_{o=1}^{N} (p^s_{not})^{1-\sigma}}.$$

The market clearing condition in the labor market implies that the sum of labor income earned by both skilled and unskilled labor at all ages in location n at time t is the same as the total income

[2] As we abstract away from physical capital formation, the coefficient governing this productivity-skill complementarity serves as a proxy for capital-skill complementarity in the model.

spent on goods from all locations:

$$\sum_j \sum_e \mathbb{1}_{\{j\neq 1,\ e\neq h\}} L_t^j(n,e) w_{nt}^e = \sum_j \sum_e \sum_{n'} \mathbb{1}_{\{j\neq 1,\ e\neq h\}} \alpha_{n'nt}^1 \zeta L_t^j(n',e) w_{n't}^e + \\ \sum_j \sum_e \sum_{n'} \mathbb{1}_{\{j\neq 1,\ e\neq h\}} \alpha_{n'nt}^2 (1-\zeta) L_t^j(n',e) w_{n't}^e.$$

Rearranging the terms and denote the working labor for each location, skill-type as $\mathbb{L}_t(n,e)$, the market clearing condition can also be written as:

$$\sum_e \mathbb{L}_t(n,e) w_{nt}^e = \sum_s \sum_{n'} \sum_e \alpha_{n'nt}^s \mathbb{L}_t(n',e) w_{n't}^e. \tag{11}$$

Sequential Equilibrium Given the path of exogenous parameters, including location-specific productivities A_n^s, and an initial distribution of workers $L(n,e;0)$, the recursive competitive equilibrium is defined by the paths of:

i. individuals' migration and educational choices for each location, education level, and age group: $\{\pi_t(n',e'|n)$ and $\pi_t^j(n'|n,e)\}_{t=0}^{\infty}$
ii. value functions for each location, education type, and age group $\{V_t^j(n,e)\}_{t=0}^{\infty}$
iii. the distribution of workers across location, educational type, and ages $\{L_t^j(n,e)\}_{t=0}^{\infty}$, and
iv. wages $\{w_{nt}^e\}_{t=0}^{\infty}$,

such that the value function (8), the population flow condition (9), the educational level condition (10) and the goods market clearing condition (11) are satisfied.

Steady State of the Equilibrium A steady state in this economy suggests no aggregate variables change over time. The labor composition in all locations stays unchanged; the individual migration still exists, while the net inflows by cohorts and skill types are equal to zero.

We solve the model in levels[3], the equations that characterize the steady state are in the Appendix.

3 Parameterization

We conducted the quantitative exercise on 273 prefecture cities in China. The sample of cities in the paper is the intersection between the cities with available educational resource data (number of teachers in higher education) and the cities with available estimates of bilateral transportation cost from Ma and Tang (2024).

[3]The model can be solved using hat algebra, but due to the lack of available data on internal trade flows at the prefecture-pair level in China, we have to solve the model in levels.

We model two broadly defined sectors $S = 2$ to distinguish the skill intensity in the production function. We first rank all the sectors from the industrial classification for national economic activities by their payment share to skilled workers observed in the *2005 One Percent Population Survey*, and classify the sectors above the median level as the "skilled-intensive sector", and those below the median as the "unskilled-intensive sector".

3.1 Data Description

We solve the model in levels and therefore require the following data: the initial distribution of labor by age, skill type, and location; wages and employment of skilled and unskilled workers in all regions; to calibrate share of high-skilled labor input in the production function, we need to calculate the share of the aggregate labor income earned by skilled labor; we need the bilateral trade costs and migration costs to address the bilateral frictions in the model; to calibrate the magnitude of the migration cost, we need the migration probability cognitional on age, skill-type, origin, and destination in 2015; to estimate parameters describing individual's education choice, we need the overall skill-ratio of individuals who are currently in higher education institutes in 2015.

The primary source of our data is the 2005 One Percent Population Survey (also known as the mini-census). The Survey was conducted by the National Bureau of Statistics of China, covering 1.31% of China's total population. This survey provides a comprehensive view of the population's demographic and socioeconomic characteristics. We use the 2005 One Percent Population Survey as our starting point to construct the initial labor distribution in the model. However, the 2005 One Percent Population Survey lacks a proportional representation of observations for the population size in each prefecture. To address this issue, we use only the educational level and age distribution data from each prefecture and scale the prefecture's total population using the prefecture-level population data from the 2005 statistical yearbook. This approach allows us to establish a population distribution by age, education level, and location that closely represents the Chinese demography. We further standardize the distribution by creating an age and educational level grid. We only include individuals aged between 18 and 61, evenly spaced into 11 cohorts. We assume that individuals with an education level at or below high school are unskilled workers, and the other are skilled workers. To maintain a stable labor supply, we simplify the distribution of each age group within the population by assuming that each cohort has an equal total population. At this point, we have obtained comprehensive information on the labor distribution within 273 prefecture cities, broken down into 11 cohorts and two skill types.

In our analysis, we specifically utilize GDP data from 2005 and information on the number of teachers in higher educational institutions for the years 2005 and 2015. *China City Statistical Yearbook* ("Yearbook" hereafter) is a valuable resource for accessing detailed location-specific data. Published annually by the National Bureau of Statistics of China, it provides essential information on the prefecture level.

The geographic linkages in the model are summarized by bilateral trade costs and migration costs, which are from Ma and Tang (2024). They comprehensively document the quality of transportation infrastructure in China over time and estimate trade costs and migration costs from a spatial model.

Finally, we utilize data from the 2015 mini-census to gather detailed information on migration probabilities. This mini-census covers 1.55% of the population in mainland China. Survey participants were asked about their previous residential locations 5 years ago if they did not currently reside at the surveyed address. By combining this information with their current residential data and individual characteristics such as age and educational level, we can create a bilateral migration flow between cities in China that represents a snapshot of labor movement within the economy.

3.2 Estimation Procedure

In Table 1, we show two sets of parameters: The upper panel includes parameters that are either taken from the literature or calibrated directly without solving the model. The lower panel includes parameters calibrated by inverting the model.

We directly assign the following parameters of the model: we use a four-yearly discount rate β of 0.85, implying a yearly interest rate of roughly 4 percent[4]. We assume an elasticity of substitution in the utility function, $\sigma = 4$, as in Tombe and Zhu (2019). We choose $\kappa = 3\beta = 2.55$, following Kleinman et al. (2023). The productivity skill complementarity of ω equals 0.5, as in Burstein and Vogel (2017). We use the estimated trade directly from Ma and Tang (2024). The migration costs in our model vary across ages, while those in Ma and Tang (2024) do not have an age dimension. We therefore assume the following functional form to map the migration frictions in our model to those in Ma and Tang (2024):

$$D^{j}_{n'nt} = \bar{D} \times \bar{D}^{j} \times D_{n'nt}. \tag{12}$$

In the expression above, $D_{n'nt}$ is the estimated migration cost from n to n' at time t in Ma and Tang (2024). $\{D^{j}\}$ is an age-specific shifter, and $\bar{D}$ is an overall shifter that governs the magnitude of the migration matrix. We will estimate $\{\bar{D}, \bar{D}^{j}\}$ together with the other parameters, as discussed in detail below.

We quantify the remaining parameters in the following three steps. The first group of parameters is estimated using the structural equations of the model without any simulation. Conditional

[4]The discount rate is $0.96^4 \approx 0.85$

Table 1: **Parameters**

Symbol	Description	Value	Source	Step
β	Discount rate	0.85	Yearly interest rate of 4%	
σ	Elasticity of substitution	4		
κ	Inverse elsaticity of migration	2.55	Kleinman et al. (2023)	
ω	Productivity skill complementarity	0.5	Burstein and Vogel (2017)	
η	Elasticity of substitution in production	1.4		
$\{\tau_{mn}\}_{m,n=1}^{N,N}$	Bilateral transportation cost		Ma and Tang (2024)	
$\{D_{mn}\}_{m,n=1}^{N,N}$	Bilateral migration cost		Ma and Tang (2024)	
$\boldsymbol{\beta}$	Parameters map observed educational resources to educational appeals		2015 Migration probability	1
D^j	Magnitude of migration cost by cohort		2015 Migration matrix	1
χ	Weight of input share for low skilled labor	0.93	2005 Wage bill	2
$\{A_n\}_{n=1}^{N}$	Productivity in each location		2005 GDP share	2
ψ	Inverse elasticity of educational migration	2.72	2015 College share	3
$\bar{D}$	Magnitude of migration cost	0.93	2015 Stay rate	3

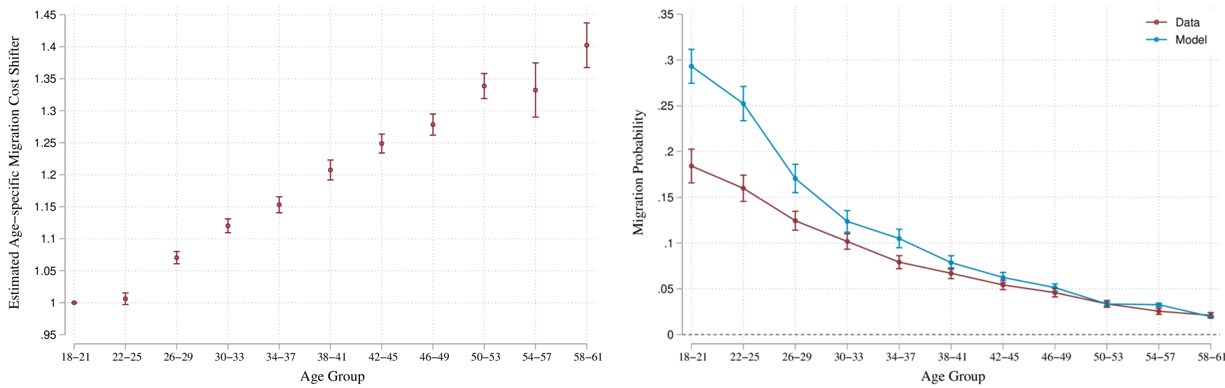
Notes: This table displays the parameters' estimated values along with the source materials used in calibration or relevant literature. The first and second columns show the symbolic representations used in the model and their respective descriptions. The third column provides the parameter values when available, and the last column describes the source citation or the source of the target used during the calibration.

on the first group, the second step estimates several parameters by solving the initial static equilibrium of the model. Lastly, the third step solves the dynamic model in the steady state to estimate the last group of parameters. In the rest of this section, we provide more details of the three-step estimation.

3.2.1 Step 1: Structural Estimation without Solving the Model

In the first step, we estimate two objects: 1) the age-specific migration costs, D^j, and 2) the parameters that map the location-specific education appeals, F_{nt}, to observables using structural equations from the model. This step does not require solving any part of the model.

Age-Specific Migration Costs We estimate the migration cost for each age group to account for the observed variations in migration probability as individuals age. To estimate $\bar{D}^j$, first recall the



(a) Cost of Migration by Age Group

(b) Probability of Migration by Cohort

Figure 2: Migration by Cohort

Note: Figure 2a shows estimated migration cost shifters by age group with 95% bootstrapped confidence interval. Figure 2b shows the migration probability by age group from both the model and the data. The migration probabilities from the data are plotted in red with confidence intervals computed by bootstrapping. The migration probabilities by cohort predicted by the model are plotted in blue, with confidence intervals calculated using weighted standard deviation.

value function and migration flow for individual of cohort j, with skill type e, in location n.

$$V_t^j(n,e) = U_t^j(n,e) + \kappa \log \sum_{n'=1}^{N} \exp\left(\beta V_{t+1}^{j+1}(n',e) - D_{n'n,t+1}^{j+1}\right)^{1/\kappa}$$

$$\pi_t^j(n'|n,e) = \frac{\exp(\beta V_{t+1}^{j+1}(n',e) - D_{n'n,t+1}^{j+1})^{1/\kappa}}{\sum_{n''=1}^{N} \exp(\beta V_{t+1}^{j+1}(n'',e) - D_{n''n,t+1}^{j+1})^{1/\kappa}}.$$

Rearranging the terms, we can write the value function as a function of migration flow:

$$V_t^j(n,e) = U_t^j(n,e) - \kappa \log \pi_t^j(n'|n,e) + \beta V_{t+1}^{j+1}(n',e) - D_{n'n,t+1}^{j+1}. \tag{13}$$

We arrange Equation (13), and take the difference between the log probability of migrating to another location n' from n and the log probability of staying in location n given cohort j and skill type e.

$$\log\left(\frac{\pi_t^j(n'|n,e)}{\pi_t^j(n|n,e)}\right) = \frac{\beta}{\kappa}\left(V_{t+1}^{j+1}(n,e) - V_{t+1}^{j+1}(n',e)\right) - \frac{D_{n'n,t+1}^{j+1}}{\kappa},$$

Similar to estimating trade costs, double-differencing the migration flows leaves us with a simple equation linking migration probabilities to migration costs. We omit the t subscript for simplification:

$$\log\left(\frac{\pi^j(n'|n,e)}{\pi^j(n|n,e)}\right)+\log\left(\frac{\pi^j(n|n',e)}{\pi^j(n'|n',e)}\right)=-\frac{\bar{D}^{j+1}\left(D_{n'n}+D_{nn'}\right)}{\kappa}.$$

Lastly, the difference in the equation between each age group and the initial age group ($j=1$) leads to the estimate of $\bar{D}^j$:

$$\underbrace{\frac{\log\left(\frac{\pi^j(n'|n,e)}{\pi^j(n|n,e)}\right)+\log\left(\frac{\pi^j(n|n',e)}{\pi^j(n'|n',e)}\right)}{\log\left(\frac{\pi^{j'}(n'|n,e)}{\pi^{j'}(n|n,e)}\right)+\log\left(\frac{\pi^{j'}(n|n',e)}{\pi^{j'}(n'|n',e)}\right)}}_{\text{data}}=\widetilde{D}^{j+1},$$

where $\widetilde{D}^{j+1}=\bar{D}^{j+1}/\bar{D}^1$ is the age-specific shifter, normalized by the first age group. We observe the data on migration flows from the 2015 mini-census.

Figure 2 shows the estimated migration cost shifters by age group, where the first age group serves as a reference point and is normalized to 1.[5] We find that migration costs increase significantly as people age. For example, the migration costs among the oldest individuals are 1.4 times higher than those of the youngest cohort. The increasing migration costs by age comes from the pattern in the data that younger individuals are much more likely to migrate than older ones. The right panel of Figure 2 shows the data and model-simulated migration probability by age group. Generally speaking, the model does well, particularly for cohorts of later ages.

The Education Appeals The education appeals at each location, F_n, capture a location's non-monetary attributes that appeal to students in their destination choices. The educational appeals are unobservable; therefore, we need to map them to observable characteristics at a location and then back out the mappings using the structure of the model. To be concrete, we first hypothesize that the abundance of college resources, the prestige and rank of colleges, and the natural amenities of a location could affect a student's choice, conditional on distance and future returns. Based on this logic, we first assume the following functional form for F_n

$$F_{nt}=\beta_2\times\exp(\beta_1\times\text{Num.teacher}_{nt}+\boldsymbol{X}\boldsymbol{\beta})+\varepsilon_{nt}. \tag{14}$$

In the expression above, we first map F_{nt} to the number of teachers in each location and time, a proxy for the abundance of educational resources. The educational resource is also the key variable through which we will conduct counterfactual analyses later. In addition to the main dependent variable, we also include a wide range of time-invariant control variables in the vector $\boldsymbol{X}$. These control variables include the number of elite universities measured as those that fall

[5]The 95% confidence intervals are computed using bootstrapping.

under Project 985 in a location as a proxy for the prestige and ranking of colleges.[6]. Lastly, we also include a prefecture's average elevation, slope, temperature, and precipitation as proxies for natural amenities. ε_n is the error term that varies across locations. The parameters of interest are β_1, β_2, and the factor loadings in the vector $\boldsymbol{\beta}$.

Our model provides a structural relationship between the education migration flows and educational appeals, based on which we can estimate the parameters of interest. We derive this structural equation using the migration flows and value functions. We first denote the educational migration probability as:

$$\pi_t^{0(1)}(n'|n,e) = \frac{\exp\left(V_t^1(n',e) - D^1_{n'n} - F_{n't}\right)^{1/\kappa}}{\sum_{n''=1}^{N} \exp\left(V_t^1(n'',e) - D^1_{n''n} - F_{n''t}\right)^{1/\kappa}}.$$

Conditional on pursuing higher education, the log difference between the probability of students moving from location n to n' and the probability of them staying in n can be written as:

$$\log\left(\frac{\pi_t^{o(1)}(n',h|n)}{\pi_t^{o(1)}(n,h|n)}\right) = \frac{1}{\kappa}\left(V_t^1(n',h) - V_t^1(n,h) - \tilde{D}_{n'n} - \tilde{D}_{nn}\right),$$
$$\text{where } \tilde{D}_{n'n} = D^1_{n'n} + F_{n't}.$$

Lastly, similar to the estimation of the migration cost, applying a double-differencing, we arrive at the structural equation based on which to back out the parameters of interest:

$$\log\left(\frac{\pi_t^{o(1)}(n',h|n)}{\pi_t^{o(1)}(n,h|n)}\right) + \log\left(\frac{\pi_t^{o(1)}(n,h|n')}{\pi_t^{o(1)}(n',h|n')}\right) = \frac{-1}{\kappa}\left(2F_{n't} + 2F_{nt} + D_{nn'} + D_{n'n}\right).$$

In the expression above, the dependent variable is a function of the observed educational migration probabilities, $\pi_t^{o(1)}(n',h|n)$. The independent variables on the right-hand side of the equation depend on the observed migration costs and the education appeals. Plugging in the functional form of F_{nt} defined in Equation (14), the equation above is the structural relationship based on which to estimate $\{\beta_1, \beta_2, \boldsymbol{\beta}\}$. Due to the non-linear nature of the equation, we use non-linear least squares to carry out the estimation. We use the observed educational migration flows and the number of college teachers in 2015 to back out F_{nt}.

The appeal function we estimated shows strong diminishing returns with respect to college concentration. Figure 3 illustrates the estimated educational appeals and the number of teachers in 2015. The estimated educational appeal increases drastically when the number of teachers is low,

[6] A total of 39 elite colleges was included in the sponsorship scheme initiated by the central government in May 1998. These colleges were widely regarded as the elite colleges in China. Within our sample period, the number of 985 universities remains unchanged.

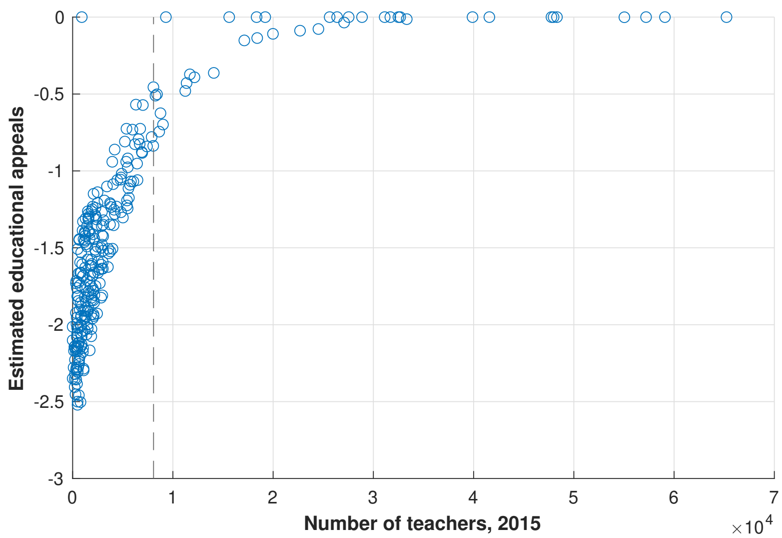


Figure 3: Estimated Education Costs

Note: This figure shows the estimated educational appeals and number of teachers in 2015. The estimated education appeals are plotted on the y-axis. The number of teachers hired in higher education institutes is plotted on the x-axis. Each dot represents a location in the model. The number of teachers in higher education institutes comes from the *China City Statistical Yearbook*. The dotted line indicates the cities in the 90th percentile regarding the number of teachers.

while as resources concentrate, the educational cost changes minimally. The median city, in terms of number of teachers hired in higher education institutes, has approximately 2,000 teachers, while the city in the top 10% has 7.4 times as many teachers as a median city. In a median city with 2,000 teachers, a 10 percent increase in teachers leads to a 2.87 percent increase in education appeals. However, the return on investment in additional colleges quickly diminishes in better-endowed locations, as indicated by the curvature of the education appeal function. For example, at the 90th percentile of cities regarding education resources, a 10 percent increase in teachers only increases costs by 0.13. In contrast, at the 10th percentile, the same increment leads to an increment of 8 times larger. This pattern strongly indicates many potential students would pursue higher education if resources were more accessible. Additionally, the data pattern suggests that educational appeals are likely to be very low in areas with a limited number of colleges. The form of the cost function also anticipates several of our numerical findings: investing in education yields a greater benefit in areas with relatively fewer educational resources. Therefore, having an excessive concentration of colleges could have a notable negative impact. We elaborate on this point in a general equilibrium framework in the counterfactual exercise.

The estimated appeals also increase the quality of colleges, as measured by the number of 985 colleges. If the median city with average amenities has one more Project 985 University, the appeal increased by 100%.

3.2.2 Step 2: Static Model Inversion at the Initial Equilibrium

In the second step, we invert the model at the initial static equilibrium to recover the location-specific fundamental productivity and the production function. The initial equilibrium corresponds to the year 2005 ($t = 1$). Model inversion at the initial equilibrium only requires solving the static market clearing conditions using the initial population distribution but not the forward-looking migration decisions. As a result, static inversion has a clear advantage in computational load and data requirements as we do not need any information regarding the migration costs to compute the initial equilibrium.

To calibrate $\{\chi^s\}$, the parameter governing the weight of input share for low-skilled workers in sector s, we target the payment share to unskilled workers in each sector. The payment share data is inferred using the 2005 mini-census that recorded the industry of employment, education attainment, and income at the individual level. We target the prefecture-sector-level output share in the data to back out the location-sector-fundamental productivity, $\{A_n^s\}$. We combine the information from the *China City Statistical Yearbook* and the 2005 mini-census to compute the target data. The former source provided the total output of each prefecture in 2005, and the mini-census allows us to compute the within-prefecture share of output by sector. As we target output share by sector, we normalize $A_1^s = 1$.

To better suit our model, we uniformly assign a scalar, $\bar{D}$, to scale the migration costs. We estimate the migration cost scalar by matching the share of individuals who stay in their current location from the 2015 census and the model predicted value. Furthermore, to address the observed decreasing migration probability over ages, we utilize the migration flow by age group from the 2015 mini-census and calibrate the migration cost scalar, D^j, for different age groups. We discuss the detailed estimation process below.

3.2.3 Step 3: Dynamic Model Inversion at Steady State

In the last step, we have two more migration-related parameters to be calibrated. The first is ψ, the inverse elasticity of educational migration, and the second is $\bar{D}$, the overall magnitude of the migration cost matrix. Both parameters shape the individuals' migration behavior in the dynamic model; therefore, we need to invert the model to dynamic equilibrium to estimate these parameters. In particular, conditional on all the other parameters estimated in the previous steps, with a guess of $\{\psi, \bar{D}\}$, we solve the model's steady state and compute the target moments used to calibrate these parameters.

We use a two-layer nested Nonlinear Least Squares procedure to perform the estimation. In the outer loop, we choose the overall migration cost shifter,$\bar{D}$, to match the overall stay rate to the

2015 stay rate in the data. To be specific, we compute the stay rate as follows:

$$\text{Stay Rate} = \frac{\sum_e \sum_{j=2}^{J} \sum_{n=1}^{N} (\pi^j(n|n,e) \times L_n^j) + \sum_e \sum_{n=1}^{N} (\pi^0(n,e|n) \times L_n^0)}{\text{Total Population}},$$

where $\pi^j(n|n,e)$ indicates the probability that an individual stays in location n with skill type e and age group j. We collect the 5-year stay rate from the One Percent Population Survey in 2015 and transform it into a 4-year stay rate[7] to align with our model period, assuming the same stay rate each year. The identification of the benchmark adjustment scalar, $\bar{D}$, relies on the stay rate of individuals. Migration costs can affect the value one can get from moving. Intuitively, the higher the migration cost, the higher the portion of individuals who choose to stay in their current location instead of pursuing higher wages elsewhere.

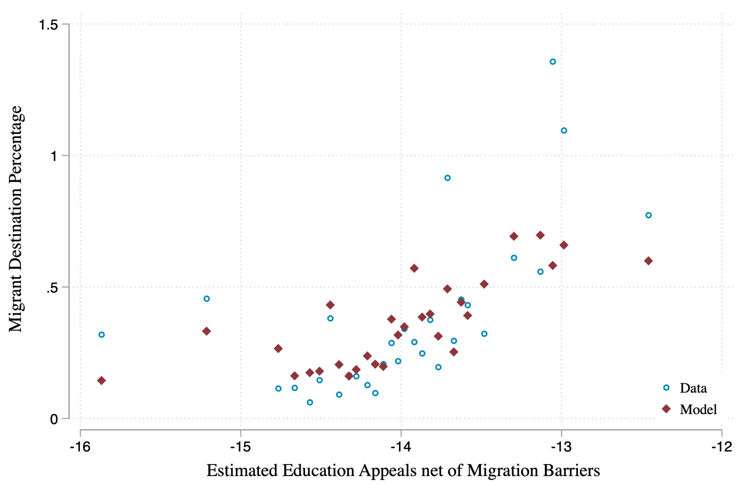


Figure 4: Migrants Destination Percentage
Note: This figure shows the binned scatter plot of the destination choice against the estimated education appeals net of the migration barrier costs for both data and model prediction. The number of bins is 30. The education cost plus the migration barrier is plotted on the x-axis and the percentage of students who chose the place as the destination conditional on going to college in a place different from the home location is plotted on the y-axis. Each dot represents a group of observations with similar estimated destination-specific costs.

In the inner loop conditional on $\bar{D}$, we choose the inverse of education elasticity, ψ, to match the share of individuals who go to college when they are young in 2015. The identification of parameter ψ governs the education choices. Given the expected lifetime utility of being a skilled and unskilled worker, a larger ψ indicates the individual is less sensitive to the value difference between being a skilled worker and an unskilled worker. We match the model-predicted share of

[7]Five-year stay rate is 91%, we transform it by assuming the same one-year stay rate: Four-year stay rate = (Five-year stay rate)$^{4/5}$

college students at the steady state to the ones we observed in the 2015 mini-census. The share of college students in the first cohort is given by

$$\frac{\sum_{n'=1}^{N}\sum_{n=1}^{N}(\pi_n^0(n',h|n)\times L_n^0)}{\sum_{n=1}^{N}L_n^0}.$$

The estimation strategy yields an estimate for the migration cost shifter, $\bar{D}=0.93$, and education elasticity, $1/\psi=0.38$[8]. Our estimation procedure fits the data fairly well. Figure 4 illustrates the destination choices of individuals. The estimated cost of obtaining an education is plotted on the x-axis, which includes $F_n + D_{n'n}$. We focus on the age group 18-21 who have chosen to attend college away from their hometown and compute the distribution of this group of people. The percentage of each location chosen as the migration destination is plotted on the y-axis. When the cost of education is lower, there is a greater likelihood that individuals will choose that location for college, given that they are part of the population pursuing higher education and moving. Our model can generate this pattern by only targeting the first moment of the distribution of migrant college students when estimating the model parameters.

4 Counterfactual Simulations

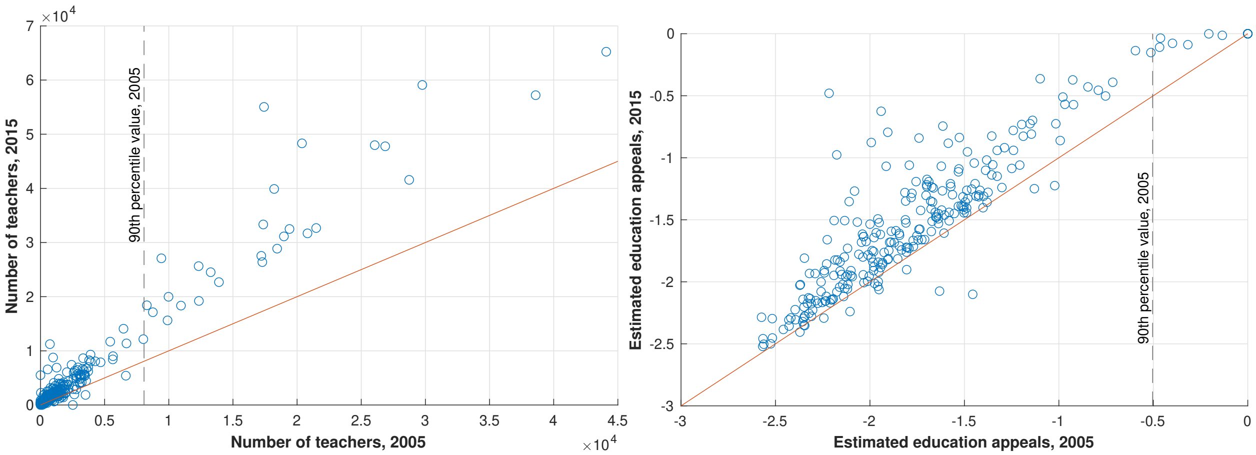
In this section, we evaluate the spatial impacts of educational resources. We carry out several counterfactual analyses with alternative distributions of higher education institutions and show how welfare and skill ratios respond to the changes in educational resources.

4.1 College Expansion

We first examine the impact of the factual expansion of the educational resources implemented by the central government between 2005 and 2015. Within a short span of time, the college expansion saw an 84 percent increase in the number of college teachers hired in higher education institutes and a 466 percent increase in college enrollments.

Through the lens of our model, the college expansion affects the educational appeals of a location through Equation (14). Our baseline simulation incorporated the changes in educational resources over time, including the college expansion between 2005 and 2015. To evaluate the impact of the factual expansion, we conducted a "no expansion" counterfactual, in which the number of college teachers is fixed at the initial year of our analysis (2005). In this counterfactual analysis, the educational appeals of a location no longer evolve over time ($F_{nt} = F_n, \forall n, t$) because

[8] Appendix A.3 displays the objective function for estimating $\bar{D}$ and ψ with each parameter varying while keeping the other parameters at their estimated values.



(a) Education Resources

(b) Estimated Education Appeals

Figure 5: Factual College Expansion

Note: Figure 4.1 shows the factual educational resources regarding the number of teachers hired in higher education institutes by location. Figure 5b shows estimated education appeals by location. The number of teachers hired in higher education institutes comes from the *China City Statistical Yearbook* in 2005 and 2015. The red line represents the 45-degree line.

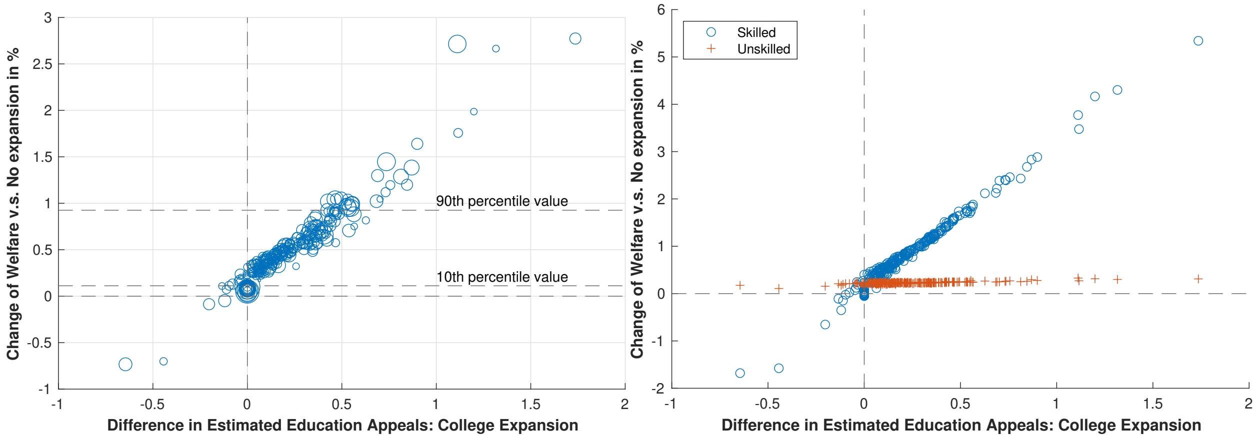
the number of teachers is the only location characteristics that change over time in Equation (14). Comparing the “no expansion” counterfactual with the baseline simulation uncovers the aggregated and distributional impact of college expansion.

The factual expansion of education resources is highly uneven across locations. Figure plots a location’s number of teachers in 2005 on the x-axis against that in 2015 on the y-axis, and the red solid line is the 45-degree line. As is evident in the figure, places with a high concentration of educational resources also received a disproportional allocation of additional resources. The additional resources allocated to the top 10 percent of prefectures in terms of number of teachers in 2005 is approximately 26 times higher than the additional resources allocated to the bottom 10 percent. The uneven distribution is partly due to cost considerations — expanding existing colleges with more teachers is arguably cheaper than creating a new college with the same number of teachers. The pattern of concentration is partially reflected in the changes in educational appeals, as shown in Figure 5b. In locations already well endowed with colleges, such as Beijing, the additional college teachers do not significantly lead to higher appeals among the students. Furthermore, the curvature of the appeal function shown in Figure 3 suggests that an increase in areas with few resources could substantially increase their education appeals. However, in the factual expansion, these areas receive minimal new resources. The most significant increase in educational appeals occurs in locations with medium-level educational resources.

The college expansion between 2005 and 2015 led to moderate welfare improvements in all locations. Panel A of Table 2 reports the overall welfare and distributional impact of factual college expansion. By the year 2015, the aggregate welfare increased by 0.52% due to the expansion of education resources. Both skilled and unskilled workers experience welfare improvement on average, and the impact on skilled workers is substantially larger than on unskilled ones (0.82% v.s. 0.24%). The expansion of college also shifted the labor force towards skilled workers, leading to a 0.56% increase in the aggregate skill ratio. The aggregate impacts of college expansion are channeled through the educational decisions at the individual level in the model. The college expansion increases the educational appeal in most locations, leading to a higher expected lifelong welfare of the skilled worker and, thus, a higher skill ratio. The higher skill ratio increases location-specific productivity through agglomeration forces, resulting in overall welfare gains. Skilled workers receive the lion's share of the benefits associated with higher productivity due to the positive productivity-skill complementarity introduced through the parameter ω. The impact of the college expansion is also long-lasting, eventually leading to a 0.72% welfare improvement for the skilled worker and a 0.35% gain for unskilled workers when the economy transits to a steady state.

Locations that experienced a larger increase in educational appeals also see higher welfare impacts, mainly through its skilled workers. As shown in Figure 6a, the welfare changes are positively correlated with the changes in the educational appeals. The dispersion of welfare changes is also substantial: the 90th percentile of the welfare change is around 0.9 percent, roughly nine times higher than those at the 10th percentile (0.1 percent). The spatial variation mainly comes from the welfare changes among the skilled workers, as shown in Figure 6b. The locations that benefited the most are often those with moderate resources initially. As a result, the spatial inequality, measured by the coefficient of variation of weighted welfare, also experienced a moderate decline due to the college expansion program. The mechanism at the prefecture level is the same as those at the aggregate level outlined above – higher educational resources lead to higher skill ratios, which benefit the local economy through agglomeration and the skilled workers through productivity-skill complementarity.

While the welfare impacts of the college expansion are non-trivial, the further concentration of educational resources in well-endowed locations could lead to sub-optimal welfare impacts. For example, while Beijing received a considerable increase in educational resources during the expansion, the impacts of its educational attractiveness and subsequently, welfare, is limited. Thus, we seek alternative allocations to allocate educational resources across space better in the next subsection.

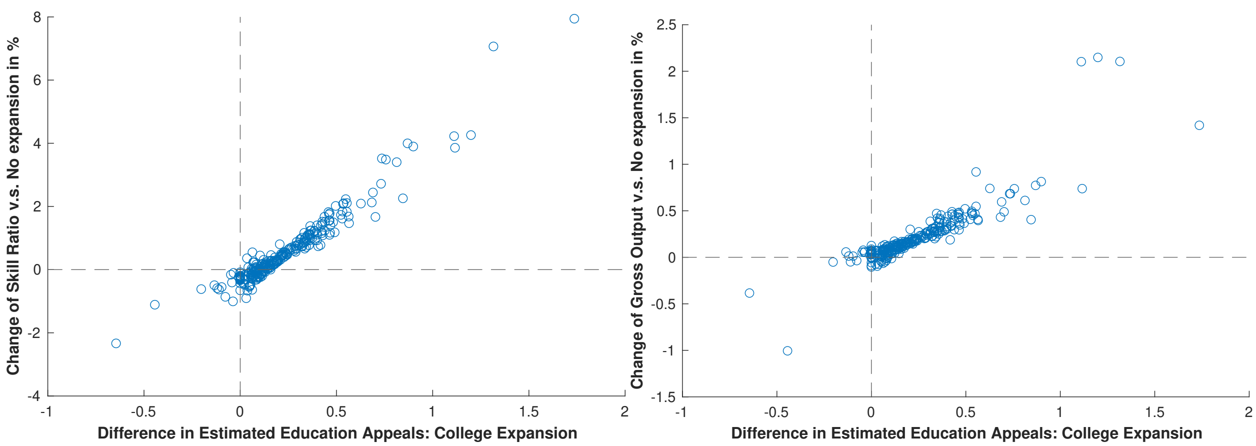


(a) Change in Welfare

(b) Change in Welfare by skill type

Figure 6: College Expansion Impacts

Note: Figure 6a shows the welfare change by locations. The size of the dot indicates the size of the population. Figure 6b shows the welfare change of skilled and unskilled by locations. All changes are compared against the scenario without college expansion. Each dot represents a location.



(a) Change in Skill Ratios

(b) Change in Output

Figure 7: College Expansion Impacts

Note: Figure 7a shows the skill ratio change by locations. Figure 7b shows the gross output change by location. All changes are compared against the scenario without college expansion. Each dot represents a location.

Table 2: Summary of Counterfactual Exercises: Welfare Change and Skill Ratio

Panel A: Impacts of College Expansion					
		ΔWelfare	ΔDispersion	10th percentile	90th percentile
Overall	T = 3	0.52%	-0.18%	0.11%	0.93%
	S.S.	0.44%	-0.11%	0.26%	0.65%
Skilled	T = 3	0.82%	-0.03%	0.05%	1.66%
	S.S.	0.72%	0.01%	0.03%	1.62%
Unskilled	T = 3	0.24%	-0.16%	0.2%	0.25%
	S.S.	0.35%	-0.26%	0.3%	0.38%
		ΔSkill Ratio	ΔDispersion	10th percentile	90th percentile
Overall	T = 3	0.56%	-0.29%	-16.38%	113.83%
	S.S.	3.87%	0.61%	0.53%	7.79%
Panel B: Impacts of Equal Expansion					
		ΔWelfare	ΔDispersion	10th percentile	90th percentile
Overall	T = 3	0.94%	-3.71%	0.51%	1.4%
	S.S.	0.83%	-2.16%	0.59%	1.06%
Skilled	T = 3	1.38%	-7.08%	0.65%	2.47%
	S.S.	1.25%	-7.43%	0.6%	2.38%
Unskilled	T = 3	0.52%	-0.39%	0.36%	0.44%
	S.S.	0.7%	-0.71%	0.6%	0.71%
		ΔSkill Ratio	ΔDispersion	10th percentile	90th percentile
Overall	T = 3	1.01%	-1.35%	-14.71%	115.87%
	S.S.	6.4%	-1.41%	2.32%	10.59%
Panel C: Impacts of Proportionate to Population Expansion					
		ΔWelfare	ΔDispersion	10th percentile	90th percentile
Overall	T = 3	0.98%	-2.88%	0.36%	1.49%
	S.S.	0.69%	-1.6%	0.46%	1%
Skilled	T = 3	1.47%	-4.9%	0.36%	2.55%
	S.S.	1.08%	-4.99%	0.4%	2.52%
Unskilled	T = 3	0.51%	-0.37%	0.36%	0.43%
	S.S.	0.58%	-0.57%	0.48%	0.59%
		ΔSkill Ratio	ΔDispersion	10th percentile	90th percentile
Overall	T = 3	1.17%	-1.69%	-14.41%	113.3%
	S.S.	5.81%	-3.38%	1.96%	11.41%

Notes: This table illustrates the changes in the levels and dispersion of both welfare and labor composition in the model simulated 2015 and the steady state, compared to the no-expansion benchmark. In Panel A, we present scenarios reflecting factual college expansion, while Panel B depicts a situation where the additional educational resources between 2015 and 2005 are distributed equally in all locations. In Panel C, we reallocate the additional resources proportionate to the initial population. The changes are calculated in comparison to a scenario in which we maintain educational resources at the 2005 level. Welfare is weighted by the population, and we also present a dispersion measure using the coefficient of variation.

4.2 Alternative Allocations

The factual college expansion results suggest that the allocation of educational resources has varying effects on different locations. To analyze this, we calculate the welfare elasticity with respect to college expansion, examining each prefecture individually. More specifically, we increase the actual number of college teachers by 10% in each location and then assess the change in local welfare compared to a scenario with no expansion. We categorize the prefectures based on their initial educational resources. Notably, the regional impact of a 10% increase in college teachers is 2.6 times greater in the bottom 10% of prefectures than in the top 10% when compared to the no expansion scenario.

We then conduct another simulation where we keep the total increment of education resources unchanged and distribute the additional teachers evenly across all locations. We call this counterfactual simulation *"equal expansion"*. In this case, all prefectures equally receive an additional 2600 college teachers. Figure 8 shows the counterfactual education appeals in this scenario. The evenly distributed resources keep the relative rank of education appeals untouched while substantially increasing the appeals of the prefectures with initially low appeals.

The *"equal expansion"* scheme leads to a substantially larger increase in welfare and skill ratio. Table 2 Panel B shows the equal expansion generates an aggregate welfare gain that is 1.8 times higher than the actual expansion program (0.94% v.s. 0.52%). The *"equal expansion"* scheme is also much more effective at increasing the overall skill ratio: the aggregate skill ratio in the counterfactual increased by 1.01%, as compared to the 0.56% increment in the baseline model. The abundance of educational resources is particularly effective in places where the initial number of teachers was low; subsequently, the welfare improvement for skilled workers from improved productivity is also more pronounced. The expansion of college programs also benefits unskilled workers through agglomeration. Figure 9 demonstrates this point by showing the welfare impact by skill type and location. On the x-axis, we show the estimated educational appeal under this "equal expansion" counterfactual. In this scenario, as shown in Figure 8, the locations with low initial education resources experienced the largest increase in education appeals and, subsequently, higher overall welfare.

We also conduct an expansion scheme, where we allocate the additional resources according to the population distribution. The impacts on education appeals for each location are shown in Figure 8. The *"population expansion"* yields a larger increase in educational appeal as compared to the factual expansion. The average educational costs decreased by 17.5% compared to the factual expansion. As a result, the overall welfare improvement is 0.98%, higher than both the factual expansion and *"equal expansion"*.

Table 3: Summary of Counterfactual Exercises: Skill Ratio

Panel A: Impacts of Education Equalization					
		ΔSkill Ratio	ΔDispersion	10^{th} percentile	90^{th} percentile
Overall	$T = 3$	1.53%	-3.72%	-1%	7.12%
	S.S.	4.36%	-0.47%	-1.74%	15.24%
Panel B: Impacts of Productivity Equalization					
		ΔSkill Ratio	ΔDispersion	10^{th} percentile	90^{th} percentile
Overall	$T = 3$	10.59%	-26.33%	0.39%	53.56%
	S.S.	30.72%	-84.8%	20.54%	112.81%

Notes: This table illustrates the changes in the levels and dispersion of labor composition in model simulated 2015 and the steady state, compared to the **college expansion** benchmark. In Panel F, we present scenarios reflecting educational resources equalization, while Panel G depicts a situation where fundamental productivities in both sectors are equalized to the average respectively in all location. The changes are calculated in comparison to a scenario in which we maintain the factual college expansion. We present a dispersion measure using the coefficient of variation.

4.3 Equalization

The final counterfactual exercise we conduct involves equalizing all educational resources to examine the impact of this even distribution of resources on spatial inequality in skill composition.

In the 'Equal college' scenario, we demonstrate that the unequal distribution of educational resources contributes to as much as 14% of the observed spatial inequality in skill composition in the short run and persist. In this exercise, educational costs are standardized across all locations while keeping the overall resource level constant. Consequently, skill ratios are adjusted accordingly. As indicated in Table 3, the skill ratio dispersion decreases by approximately -3.72% to -0.47% compared to the baseline results along the transition path. Figure 10 illustrates the changes in skill ratios in various locations under this equalization scenario. In Figure 10a, the x-axis represents the skilled intensive sector productivity, while the y-axis depicts the change in skill ratios comparing to the factual expansion in the equalization scenario. The dashed vertical line indicates the 10th and 90th percentile productivity value. Meanwhile, Figure 10b displays the number of teachers in 2015, plotted on the x-axis. Clearly apart from the educational resources, fundamental productivity also determines the labor compositions.

To assess the impact of this hypothetical resource equalization, we also standardize fundamental productivity across all locations. We set the fundamental productivity to the average estimated level across all locations. Table 3 demonstrates that equalizing productivity can significantly reduce skill ratio dispersion. The dispersion of skill ratio declines by 26.33% to 84.8% along the transition path to the steady state. An evenly distributed educational resource is 14% as effective at reducing spatial inequality as an evenly distributed productivity in the short run. However,in the

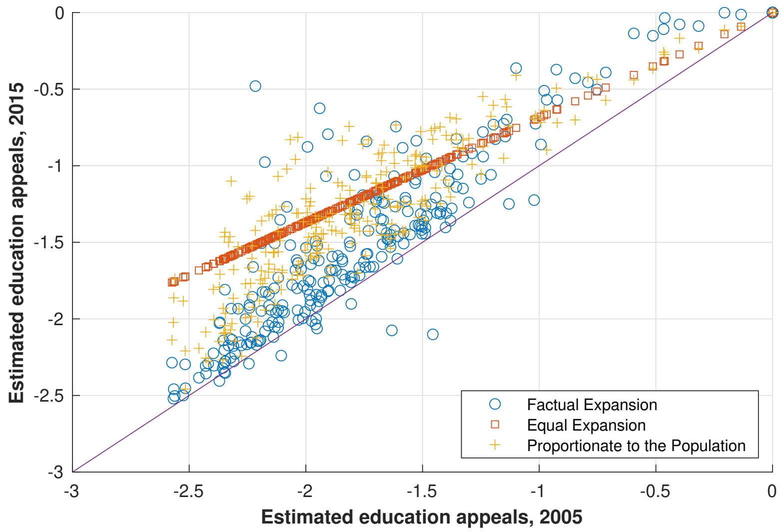


Figure 8: Factual College Expansion and Alternatives

Note: This figure shows the estimated education appeals for college expansion and alternative counterfactual in 2005 and 2015. Each dot represents a location.

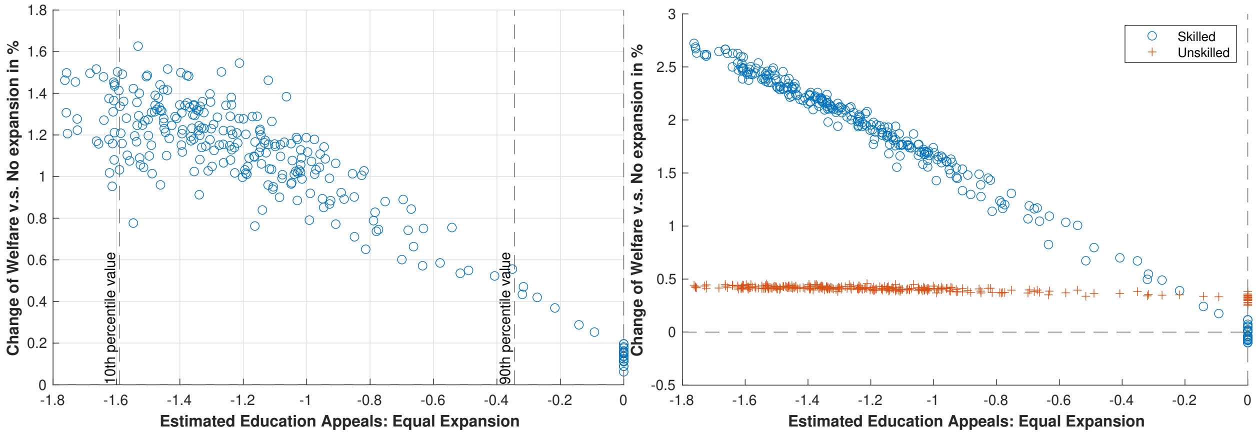
long run, the skill ratio has a strong correlation with the skill-intensive sector productivity.

5 Concluding Remarks

This paper integrates educational choices into a dynamic spatial model to examine how location-specific educational resources affect spatial inequality. We build a dynamic spatial model with overlapping generations. The individuals in the model make decisions on education, including, whether and where to attend college. We use the model to estimate the cost of higher education in each prefecture and perform counterfactual policy experiments to determine if more evenly distributed resources could lead to better outcomes. We quantify the model to mirror China and structurally estimate the cost of obtaining a college degree in each prefecture.

We find diminishing returns to college concentration in estimated education costs. Initially, educational costs decrease significantly with a small increase in resources. However, as resources continue to concentrate, the reduction in costs becomes less proportional compared to earlier stages. This suggests an over-concentration of colleges might carry a sizable negative consequences.

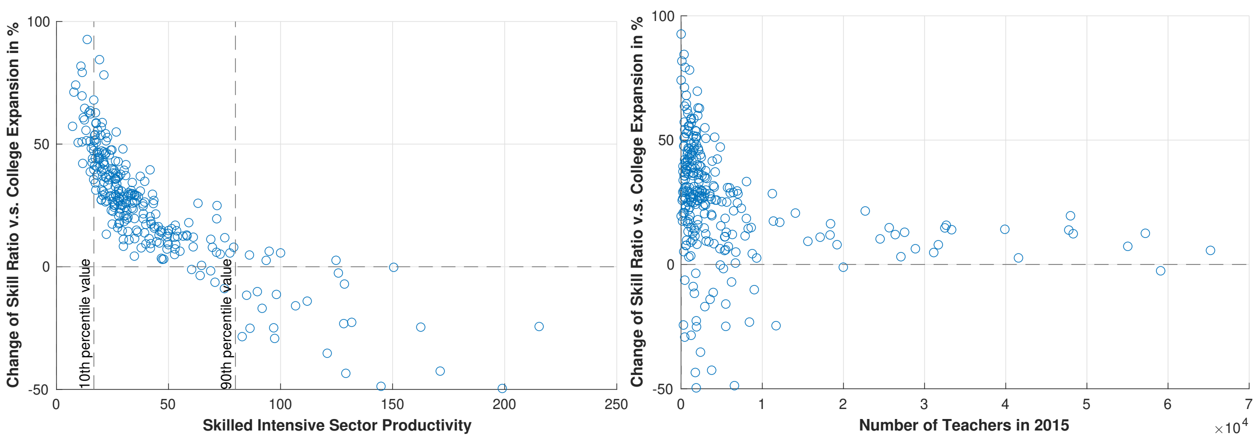
We also find that the real college expansion has a negligible effect on overall welfare and the skill ratio. This expansion disproportionately allocates resources to already well-endowed locations, with little impact on less-endowed areas. When we simulate a scenario in which additional resources are evenly distributed across regions, we observe a doubled increase in both the welfare impact and the impact on skill ratio comparing to the observed expansion.



(a) Change in Welfare (b) Change in Welfare by skill type

Figure 9: Equal Expansion Impacts

Note: Figure 9a shows the welfare change by locations. Figure 9b shows the welfare change of skilled and unskilled by locations. All changes are compared against the scenario without college expansion. Each dot represents a location.



(a) Productivity (b) Resources

Figure 10: Education resources equalization

Note: Figure 10a plots the change of skill ratio against the productivity by locations. Figure 10b shows the change of skill ratio against factual educational resources in 2015. All changes are compared against the scenario with college expansion. Each dot represents a location.

Furthermore, we show that in the short run, equalizing educational resources is roughly 14% as effective as equalizing fundamental productivity. However in the long run, the labor distribution would be largely explained by productivity difference in skill-intensive and unskilled-intensive sectors.

Finally, this study abstracts government's decisions on investment in educational resources. Allowing local government optimizing investment in educational investment, the model could explore intriguing topics such as the competition among local governments and the evaluation of policies aimed at attracting skilled talent. An important direction for future work is to incorporate government's role into analyses. Furthermore, the spatial general equilibrium life-cycle model developed in this paper can be adapted to investigate other individual decisions such as marriage and fertility.

References

Ahlfeldt, G. M., S. J. Redding, D. M. Sturm, and N. Wolf (2015). The economics of density: Evidence from the berlin wall. *Econometrica 83*(6), 2127–2189.

Allen, T. and C. Arkolakis (2014). Trade and the topography of the spatial economy. *The Quarterly Journal of Economics 129*(3), 1085–1140.

Allen, T. and D. Donaldson (2022). Persistence and path dependence: A primer. *Regional Science and Urban Economics 94*, 103724.

Armington, P. S. (1969). A theory of demand for products distinguished by place of production (une théorie de la demande de produits différenciés d'après leur origine)(una teoría de la demanda de productos distinguiéndolos según el lugar de producción). *Staff Papers-International Monetary Fund*, 159–178.

Artuç, E., S. Chaudhuri, and J. McLaren (2010). Trade shocks and labor adjustment: A structural empirical approach. *American economic review 100*(3), 1008–1045.

Burstein, A. and J. Vogel (2017). International trade, technology, and the skill premium. *Journal of Political Economy 125*(5), 1356–1412.

Caliendo, L., M. Dvorkin, and F. Parro (2019). Trade and labor market dynamics: General equilibrium analysis of the china trade shock. *Econometrica 87*(3), 741–835.

Caliendo, L., L. D. Opromolla, F. Parro, and A. Sforza (2021). Goods and factor market integration: A quantitative assessment of the eu enlargement. *Journal of Political Economy 129*(12), 3491–3545.

Che, Y. and L. Zhang (2018). Human capital, technology adoption and firm performance: Impacts of china's higher education expansion in the late 1990s. *The Economic Journal 128*(614), 2282–2320.

Dustmann, C. and A. Glitz (2011). Migration and education. In *Handbook of the Economics of Education*, Volume 4, pp. 327–439. Elsevier.

Eckert, F. and M. Peters (2022). Spatial structural change.

Fan, J. (2019, July). Internal geography, labor mobility, and the distributional impacts of trade. *American Economic Journal: Macroeconomics 11*(3), 252–88.

Kleinman, B., E. Liu, and S. J. Redding (2023). Dynamic spatial general equilibrium. *Econometrica 91*(2), 385–424.

Komissarova, K. (2022). Location choices over the life cycle: The role of relocation for retirement.

Li, J., S. Liu, and Y. Wu (2020). Identifying knowledge spillovers from universities: quasi-experimental evidence from urban china. *Available at SSRN 3621422*.

Ma, L. and Y. Tang (2020). Geography, trade, and internal migration in china. *Journal of Urban Economics 115*, 103181.

Ma, L. and Y. Tang (2024). The distributional impacts of transportation networks in china. *Journal of International Economics*, 103873.

Ma, X. (2023). College expansion, trade and innovation: Evidence from china. *International Economic Review*.

Redding, S. J. and E. Rossi-Hansberg (2017). Quantitative spatial economics. *Annual Review of Economics 9*, 21–58.

Suzuki, Y. (2023). Local shocks and regional dynamics in an aging economy. *Available at SSRN 4360965*.

Takeda, K. (2022). The geography of structural transformation: effects on inequality and mobility.

Tombe, T. and X. Zhu (2019). Trade, migration, and productivity: A quantitative analysis of china. *American Economic Review 109*(5), 1843–1872.

A Appendix

A.1 Steady State in Levels

The steady state of the equilibrium is described by the following system of equations:

$$
\begin{aligned}
p^s_{ni} &= \tau_{ni}\frac{1}{A^s_i}\left[A_i^{s-\omega(\eta-1)}\chi^s {w^l_{it}}^{1-\eta} + A_i^{s\omega(\eta-1)}(1-\chi){w^h_{it}}^{1-\eta}\right]^{\frac{1}{1-\eta}} \\
\alpha^s_{ni} &= \frac{{p^s_{ni}}^{1-\sigma}}{\sum_o {p^s_{no}}^{1-\sigma}} \\
P_n &= \left(\frac{P^{s1}_n}{\zeta}\right)^{\zeta}\left(\frac{P^{s2}_n}{1-\zeta}\right)^{1-\zeta} \\
\mathbb{L}(n,h)w^h_n + \mathbb{L}(n,l)w^l_n &= \sum_s\sum_o \alpha^s_{on}\left(\mathbb{L}^s(o,h)w^h_o + \mathbb{L}^s(o,l)w^l_o\right) \\
V^j(n,e) &= U^j(n,e) + \kappa\log\sum_{o\in\mathcal{N}}\exp\left(\beta V^{j+1}(o,e) - D^{j+1}_{on}\right)^{1/\kappa} \\
V^J(n,e) &= U^J(n,e) \\
V^0(n,e) &= \kappa\log\sum_{o\in\mathcal{N}}\exp\left(V^1(o,e) - D^1_{o,n} - 1_{e=h}F_o\right)^{1/\kappa} \\
\pi^j(n'|n,e) &= \frac{\exp(\beta V^{j+1}(n',e) - D^{j+1}_{n'n})^{1/\kappa}}{\sum_{o\in\mathcal{N}}\exp(\beta V^{j+1}(o,e) - D_{on})^{1/\kappa}} \\
\pi^0(n',e|n) &= \frac{\exp\left(U^1(n',e) - D^1_{n'n} - 1_{e=h}F_{n'}\right)^{1/\kappa}}{\sum_{o\in\mathcal{N}}\exp(U^1(o,e) - D^1_{on} - 1_{e=h}F_o)^{1/\kappa}} \times \frac{\exp(V^0(n,e))^{1/\psi}}{\sum_{e'}\exp\left(V^0(n,e')\right)^{1/\psi}} \\
L^0_n &= L^J_n \\
L^1(n',e') &= \sum_n L^0_n \pi(n',e'|n) \\
L^j(n',e) &= \sum_n L^{j-1}(n,e)\pi^{j-1}(n'|e,n) \quad j = 2,...,J \\
\mathbb{L}(n,e) &= \mathbb{L}_{n,h} + \mathbb{L}_{n,l}
\end{aligned}
$$

A.2 Algorithm to solve for the path

The economy evolves according to the law of motion of labor distribution. The following algorithm describe the process.

1. Guess $\{L^j_t(n,e)\}$ for a long enough period.

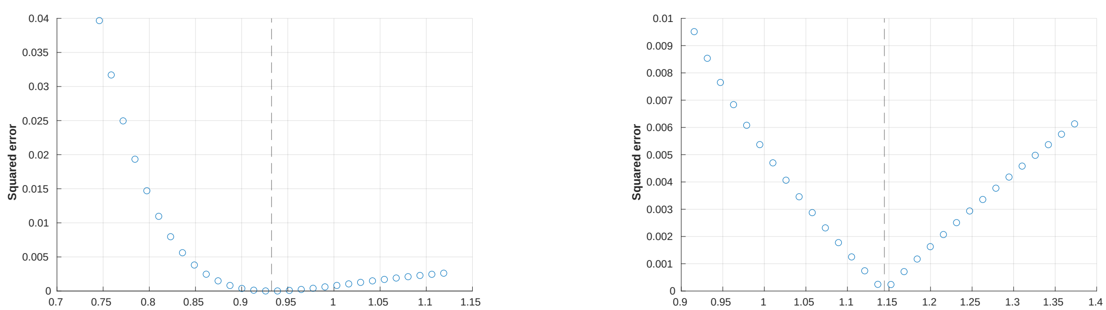
2. Solve for $\{w^h_{nt}, w^l_{nt}\}$ using the market clearing condition along with the relationship between w^h_{nt} and w^l_{nt}.

3. Compute recursively the value functions $\{V_t^j(n, e)\}$ at each location and time period for each individual type using equation 8.

4. Compute migration probability $\{\pi_t n', e'|n\}$ and $\{\pi_t^j(n'|n, e)\}$ using equation 10 and equation 9.

5. Update $\{L^j(n, e)\}$ using the law of motion, repeat until converge.

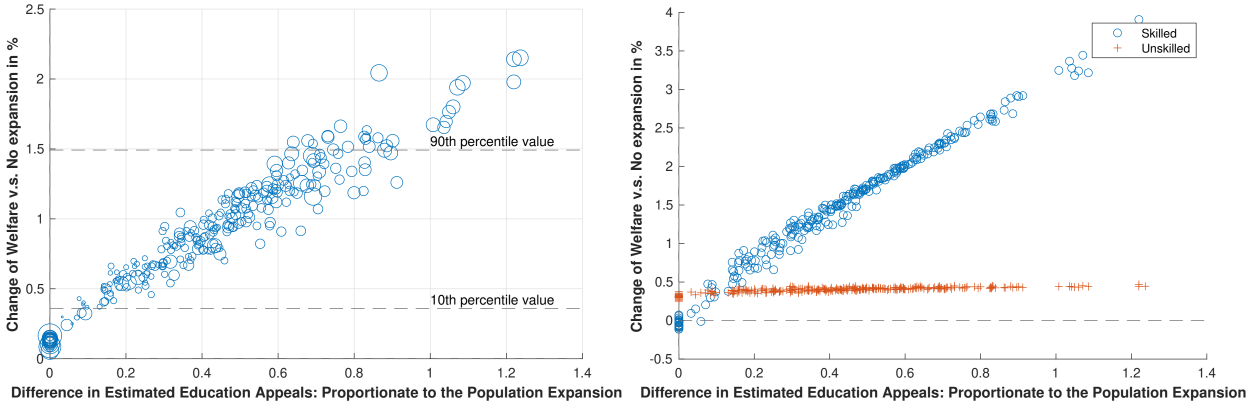
A.3 Additional figures



(a) $\bar{D}$ (b) ψ

Figure A.1: Local plot for Nested NLS Estimation

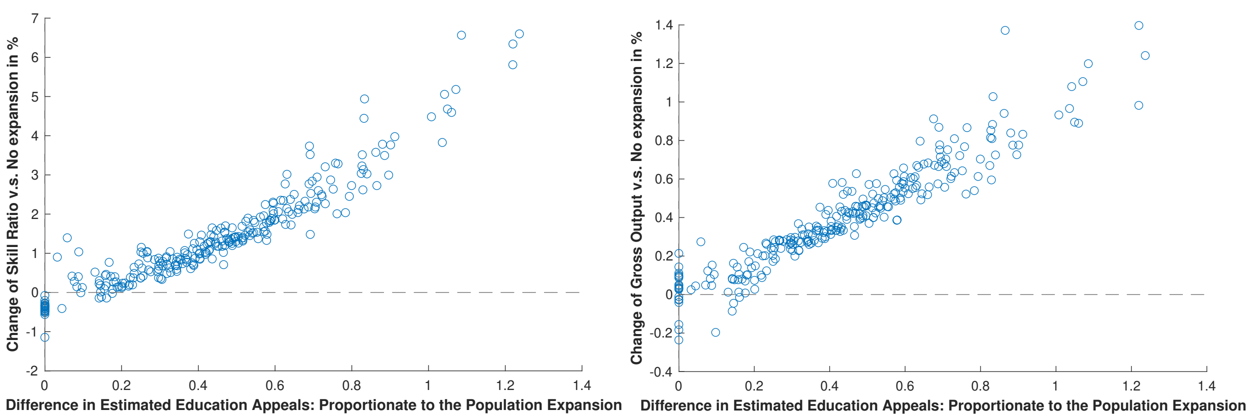
Note:The panels in this figure displays the objective function for estimating $\bar{D}$ and ψ with each parameter varying while keeping the other parameters at their estimated values. The dashed line indicates the estimated value for $\bar{D}$ andψ.



(a) Change in Welfare (b) Change in Welfare by skill type

Figure A.2: Proportionate to the Population Expansion

Note: Figure A.2a shows the welfare change by locations. Figure A.2b shows the welfare change of skilled and unskilled by locations. All changes are compared against the scenario without college expansion. Each dot represents a location.



(a) Change in Skill Ratios (b) Change in Output

Figure A.3: Proportionate to the Population Expansion

Note: Figure A.3a shows the skill ratio change by locations. Figure A.3b shows the gross output change by locations. All changes are compared against the scenario without college expansion. Each dot represents a location.