Educational Migration in China

Naiyuan Hu [*] Lin Ma[*]

This version: October 31, 2023
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 Chinese city only has 6 universities. Even worse, the bottom quarter of the cities, many with millions of population, has no more than a single college. The uneven distribution of educational resources could lead to dire consequences for both individuals and society at large. Access to colleges shapes the fate of students: those born in unlucky locations with scarce resources must endure the ordeal of long-term migration at a young age to seek a higher education; deterred by such costs, many of the talented students forgo such 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? Are there any aggregate losses coming from the over-concentration of colleges? 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 general equilibrium dynamic spatial model with life-cycle elements to analyze the impact 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, and the costs of obtaining a degree. 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, all the skilled and unskilled workers supply labor, consume, and move to other labor markets subject to age-specific migration frictions throughout their life cycle until their retirement and exit from the model.

Educational resources exert their long-run spatial impacts through several channels in the model. First, locations with better educational resources feature lower costs of education and directly benefit the 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 locational advantages 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. The positive supply of skilled workers could also spill over to the nearby areas. 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. China is an exciting case to focus on in our context: it is a country with highly concentrated educational resources, as highlighted earlier. The large spatial variations in educational resources are particularly valuable econometrically, as they allow us to structurally estimate the costs of higher education 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%. 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 results could be achieved with a more even distribution of educational resources.

To understand the aggregate and the distributional effects of college distribution, the core empirical question is how to map the observed distributions of resources (number of colleges, teachers, etc) to the unobserved costs of attending college in a location. To this end, we exploit the prediction of our model to discipline the education costs with the observed student distribution structurally. In particular, the predicted migration probability matrix of the students summarizes the attractiveness of a location to college seekers that depends on the underlying transportation network, migration policy, expected option value, and educational resources. Conditional on all the other elements shaping the migration probability, our model provides a natural mapping from location-specific educational costs to the observed distribution of college seekers. Together with some functional form assumptions that map the observed resources to the educational costs, we can infer the location-specific educational costs as a function of resources and estimate the parameters using non-linear least squares.

The estimated education cost suggests diminishing returns to college concentration. In the median city in terms of educational resources, a 10 percent increase in teachers leads to a 2.34 percent reduction in education costs. However, the return to more colleges quickly recedes in better-endowed locations. For example, at the 90th percentile of cities, a 10 percent increase in teachers only reduces costs by 0.06 in level, while at the 10th percentile, the same increment leads to a 5 times larger reduction. The shape of the cost function is identified through the distribution of students. In the data, the number of college students increases significantly with a small increment of colleges in the left-tail of the distribution. This pattern directly suggests the existence of a sizeable latent student population that would seek higher education had the resources been available. Subsequently, the data pattern also implies that the educational costs associated with a scarcity of colleges must be exceedingly high in those locations. The shape of the cost function also foreshadows many of our quantitative results: educational investment has a higher return in places that are relatively lacking in educational resources, and therefore, an over-concentration of colleges might carry a sizable negative consequence.

We find that college expansion between 2005 and 2015 led to a limited increase in the welfare level at the aggregate level and had a negligible effect on the aggregate skill ratio. 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 only changed by 0.02 percent, while the aggregate skill ratio only increased by 0.29 percent by 2015. The lukewarm 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 cost function, the average costs of attending college 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 1% increase in college teachers is xx% 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 xx college teacher. We find that the aggregate welfare effects of this simple "equal growth" scheme are roughly five times larger, and the impacts on skill ratio more than ten times higher than the observed expansion at 3.57%.

Lastly, we show that the unequal distribution of educational resources is responsible for up to 20% of the observed spatial inequality in skill composition. Moreover, equalizing educational resources is roughly 25% as effective as equalizing fundamental productivity in reducing spatial inequality. 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 6% to 20% 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 31% to 80% along the transition path to the steady state, as compared to the baseline model. In this sense, an evenly distributed educational resource is 25% as effective in reducing spatial inequality as an evenly distributed 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, Chaudhuri, and McLaren (2010); Allen and Arkolakis (2014); Ahlfeldt, Redding, Sturm, and Wolf (2015); Caliendo, Dvorkin, and Parro (2019); Caliendo, Opromolla, Parro, and Sforza (2021); Kleinman, Liu, and Redding (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 policymaking 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, Liu, and Wu, 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 life-cycle model within a dynamic spatial framework where we model individuals' educational choices endogenously. The goal of our research is to examine the impacts of unevenly distributed educational resources on individuals' migration choices, the composition of the labor market in each region, and its overall impact on inequality over time.

The model builds on the quantitative model introduced by Caliendo et al. (2019). In their work, they examine how equilibrium allocations are influenced by factors such as individual mobility frictions, trade costs, geographical variables, and input-output linkages. It's worth noting that, in their model, agents are assumed to be immortal and make labor market decisions every period. In contrast, our model introduces heterogeneous cohorts. In our model, individuals make education decisions in the early stages and persist with their chosen skill type throughtout the model's duration.

We will first introduce our main quantitative framework to characterize endogenous migration and educational choice. Following this, we will outline the equations that define our equilibrium conditions.

2.1 Setup

The economy has N locations, separated by bilateral migration costs $D^{j}_{n'n}$, which depend on the origin n and destination n' of migration. Migration costs increase with individuals' age, aligning with the observed decrease in migration rates for various reasons. These migration costs are time-invariant and are perceived as a disutility. Labor markets are local in each location, but labor is mobile. Our focus is on solving for the equilibrium labor distribution and examining individuals' education decisions.

Demographics The economy is inhabited by successive generations, each spanning J periods.

Initially, each cohort comprises $\tilde{L}$ individuals and is replaced by a new cohort every J periods. The new cohort mirrors the distribution of the old cohort, ensuring that the labor force distribution remains unaffected by cohort replacement. However, the labor force distribution will be influenced by the migration decisions of the new cohort during their first period. The entire population in this economy equals $J\tilde{L}$.

Individuals go through two life stages: "young adulthood," which lasts for one period, and "late adulthood," spanning $J-1$ periods. Individuals are identified by their skill level, which can be either skilled or unskilled ($e = \{l, h\}$), as well as their current location (n) and cohort (j). Each individual experiences origin-specific educational shocks denoted as $z(i)$ and destination preference shocks denoted as $\boldsymbol{\epsilon_{n'}}(i)$. In the first stage, individuals make choices between pursuing education and entering the labor force. Those who enter the labor force are all employed. Those who pursue education consume home production. An individual with skill type e in location n supplies a unit of labor inelastically and earns a competitive market wage w_n^e. These individuals spend their entire wage income on consumption and decide how to distribute their income across a basket of goods from all locations.

We denote the indirect flow utility for an individual in cohort j, with an education level of e and location n, as $U^j(e, n)$.

Technology The production technology in our model closely follows the framework outlined by Armington (1969) and is combined with Constant Elasticity of Substitution (CES) preferences. We make the assumption that each location specializes in the production of a different good, and consumers have a preference for consuming a variety of goods. In each location, a perfectly competitive market prevails. The production technology is assumed to be constant elasticity of substitution, and it requires both skilled and unskilled workers. The output for location n is given by:

$$y_{nt} = a_n \left[\chi^{\frac{1}{\eta}} \left(a_n^{-\omega} l^l{}_{nt} \right)^{\frac{\eta-1}{\eta}} + (1-\chi)^{\frac{1}{\eta}} \left(a_n^{\omega} l_{nt}^h \right)^{\frac{\eta-1}{\eta}} \right]^{\frac{\eta}{\eta-1}},$$

where a_n denotes the productivity in location n. We allow productivity complementarity in the production process, denoted by ω, as introduced by Burstein and Vogel (2017). Productivity complementarity implies that employing high-skilled labor is more effective in production. The coefficient governing this productivity difference between skilled and unskilled labor serves as a proxy for capital complementarity in the model. It's worth noting that we abstract the model from capital's direct role in production, making our setup parsimonious yet versatile. Wage in each location for each skill level, denoted as w_{nt}^e, is determined endogenously through local labor market clearing. Wage disparities between locations are one of the factors that drive labor migration.

Additionally, individuals' migration decisions are influenced by the expected future wages, which contribute to the option value of migration.

Timeline and Decisions At the start of each cohort's lifetime, individuals in the new cohort born in location n make sequential decisions regarding education and study location during their "young adult" stage. These forward-looking individuals, after observing education preference shocks $\{z\}$, make decisions about education. Following their education choice, they select the location from the set of all possible locations $n \in \mathcal{N}$, considering bilateral migration costs specific to their cohort group $D^{j}_{n'n}$, education cost $\{F_n\}$ (if they choose to study) and personal preference. After making these decisions, individuals move to their chosen location at the start of the period, where they either study or work and consume. In subsequent periods, they draw location choice shocks to make migration decisions and move at the start of the next period.

Figure 1 shows the timeline for individual entering and exiting 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 who is born in location n draws an preference for education and choose whether to become a skilled worker or not. After this decision, the agent also draws preference shocks for all locations, and selects a place. Those who opt to work remain as unskilled workers and earn corresponding wages. Those who decide to study consume home production and incur costs to attend college F_n, with costs varying across locations based on the availability of educational resources. If the location is rich in educational resources, it takes less effort for a student to get to the same level of skill. Apart from the cost of education, individuals also bear the disutility of migration. The decision of where to go to school affects the future payoff. Individuals weigh the decision to attend a local school, which offers post-graduation opportunities in their immediate area, against the option of studying in other locations. Opting for the latter involves paying additional migration costs, but it has potential to result in better job prospects and closer proximity to labor markets.

The location of the school affects their payoff in three ways, firstly, locations with better educational resources feature lower costs of education. Secondly, the distance between home and the chosen location affects migration costs in the first period, influencing their immediate payoff. Thirdly, the labor market conditions in the destination city and the proximity to major labor markets influence the expected value in the future and in turn affect today's decision. In later periods, individuals can make migration decisions at the beginning of each period. These decisions depend on real wage in the destination labor market, migration cost, the preference shock drawn by the individual at the beginning of this period, and the future option value in the destination. When individual reaches the last period of their working life, they base their decision solely on current utility, moving or staying in the place where they can maximize the utility from consumption $U^{J}_{t}(n, e)$ and preference shock $\epsilon_{n'}$.

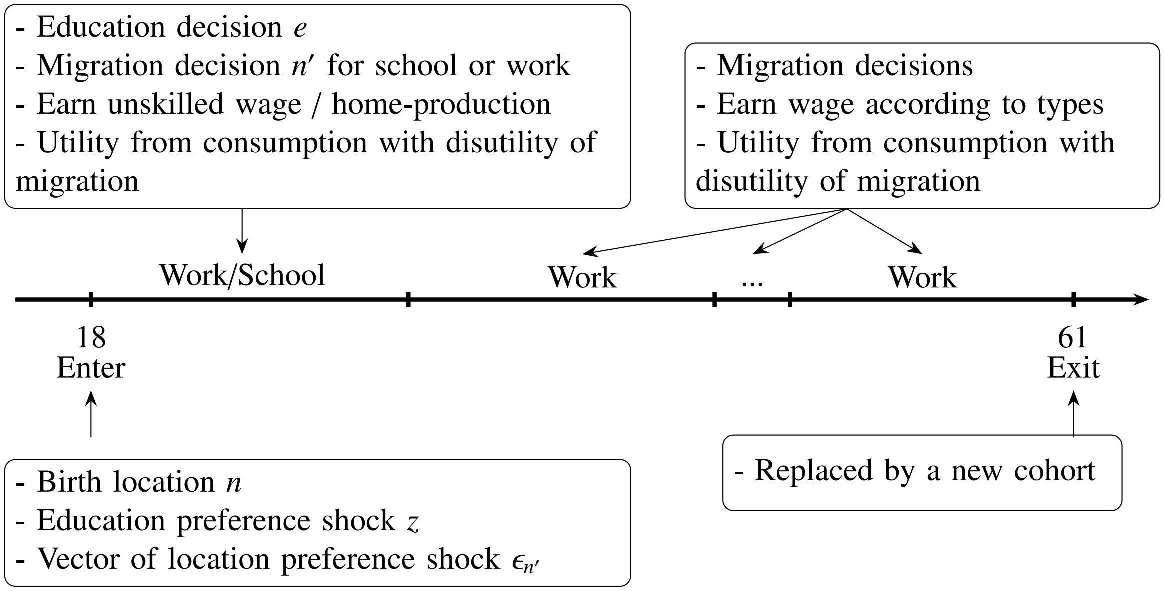


Figure 1: Estimated Education Costs

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

Dynamic Choice Problem In the model, individuals make education decisions and select their migration destinations while they are in the model. These forward-looking individuals discount the future at rate $\beta \geq 0$. We start with the cohorts in the middle and present the value function. The value function for an individual in cohort j, with education level e, living in location n, is equal to the current flow of utility in that location plus the option value to move into any other location in the next period, we have

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

when $1 < j < J$. The migration costs are differed by each cohort denoted by $D_{n'n}^j$. In the last period, the value function does not include the option value since individual can no longer move after this period. Thus, we have when $j = J$,

$$V_t^J = U_t^J(n, e)$$

At the very beginning, individuals who were born in location n, make education decisions and then make migration decisions. Here we use $V_t^1(n, e)$ to denote the first period utility omitting education cost and migration cost[1].

$$V_t^1(n,e) = U_t^1(n,e) + E_{\epsilon_{n'}} \max_{n'} \left\{ \beta V_{t+1}^2(n',e) - D_{n'n}^2 + \kappa \epsilon_{n'} \right\},$$

Conditional on the education choice e, individual chose the optimal location n' as the destination of the first period.

$$V_t^0(n',e) = \max_{\{n'\}} \{ V_t^1(n',e) - D_{n'n}^1 - 1_{\{e=h\}} F_{n'} + \kappa \epsilon_{n'} \}$$
$$\max_{\{e\}} \{ V_t^0(n,e) + \psi z \}$$

where $F_{n'}$ is the cost of education in terms of utility. This cost only occurs when the individual chooses to become a skilled labor later in their life. This nested discrete choice allows us to disentangle the heterogeneous preference of location and skill level. Individuals make decisions sequentially, they first decide on the skill type given the ability shock they receive. The education preference shock, z, only enters the value function in the first period. In our model, the preference shock for education can be viewed as the ability of deriving joy from learning. This preference shock, which contributes to an individual's utility, can also help explain why some individuals choose to study in the first period despite facing high opportunity costs. The preference shock captures an individual's intrinsic joy or inclination for learning, providing an additional layer of insight into their choices.

Individuals derive utility from consumption:

$$U_t^j(n,e) \equiv \ln C_t^j(n,e),$$

where $C_t^j(n,e)$ is consumption index for workers with education level e in location n at time t. The consumption index is over a basket of goods from all locations

$$C_t^j(n,e) = \left(\sum_{i \in \mathcal{N}} (c_{int}^e)^{\frac{\sigma}{\sigma-1}} \right)^{\frac{\sigma}{\sigma-1}}, \quad \sigma > 1, \quad e \in \{l,h\},$$

where c_{int}^e is the consumption of i's good in destination n at time t and σ is the constant elasticity of substitution (CES) between each varieties. We denote the ideal price index by $P_{nt} = \left(\sum_i p_{nit}^{1-\sigma} \right)^{\frac{1}{1-\sigma}}$, where p_{nit} is the price of goods purchased from location n for consumption in location i at time t.

[1]Omitting migration and education cost can simplify the notation, since we won't need to track the destination at this point, but they do present when the agent make choices

Workers' indirect utility depends on the income(I^e_{nt}) and the price index(P_{nt}):

$$U^j_t(n,e) = \begin{cases} b_{nt} & \text{if } j = 1 \text{ and } e = h \\ \ln I^j_{e,nt} - \ln P_{nt}, & \text{otherwise} \end{cases}$$

$$I^j_{ent} = w^e_{nt}$$

"For individuals who are working, their income is derived from the wage they earn by inelastically supplying one unit of labor. However, if an individual is studying in the first period ($e = h$ and $j = 1$), their consumption is sourced from home production. Due to the Armington assumption, the share of income at location n spent on goods supplied by location i at time t is given by,

$$\alpha_{nit} \equiv \frac{(p_{nit})^{1-\sigma}}{\sum_{o=1}^{N}(p_{not})^{1-\sigma}},$$

where p_{nit} equals to the marginal cost of producing plus the transportation cost τ_{ni} at time t, $p_{ni} = \tau_{ni}\frac{1}{a_i}\left[a_i^{-\omega\eta}\chi {w^l_{it}}^{1-\eta} + a_i^{\omega\eta}(1-\chi){w^h_{it}}^{1-\eta}\right]^{\frac{1}{1-\eta}}$, a_i is location-specific productivity.
The market clearing condition can be written in labor income, the sum of labor income earned by both skilled and unskilled labor at all ages in location n at time t can be written as the total income spent on goods from all location, given by,

$$\sum_j \sum_e 1_{\{j\neq 1 and e\neq h\}} L^j_t(n,e) w^e_{nt} = \sum_j \sum_e \sum_o 1_{\{j\neq 1 and e\neq h\}} \alpha_{ont} L^j_t(o,e) w^e_{ot}.$$

we relabel the working labor as $\mathbb{L}$, and rewrite the condition,

$$\sum_e \mathbb{L}_t(n,e) w^e_{nt} = \sum_o \sum_e \alpha_{ont} \mathbb{L}_t(o,e) w^e_{ot} \tag{1}$$

Solving education and migration decisions We assume the idiosyncratic shocks of the location taste $\epsilon \geq 0$ and education taste $z \geq 0$ are drawn from type-I extreme distributions, with κ and ψ parameterize the importance of the shocks. The shocks are measured in terms of utility and are additive. The location preference ϵ is standard in dynamic discrete choice models, such as in (Artuç et al., 2010) and (Caliendo et al., 2019). The taste shock for education is a new shock introduced by our model. The assumption on the distribution of the ability shocks allow as to include two shocks and construct the value function by summing up the two shocks sequentially and still yield a closed form solution to the migration probability.

Under this assumption, we can solve the aggregate migration flow and the probability of educational choice.The expected life time value for worker with skill $e \in l, h$ being in location n, in period j is given by,

$$V_t^j(n,e) = U_t^j(n,e) + \kappa \log \sum_{o \in \mathcal{N}} \exp\left(\beta V_{t+1}^{j+1}(o,e) - D_{on}^{j+1}\right)^{1/\kappa}. \tag{2}$$

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,

$$\pi_t^j(n'|n,e) = \frac{\exp(\beta V_{t+1}^{j+1}(n',e) - D_{n'n}^{j+1})^{1/\kappa}}{\sum_{o \in \mathcal{N}} \exp(\beta V_{t+1}^{j+1}(o,e) - D_{on}^{j+1})^{1/\kappa}} \tag{3}$$

At the very beginning, the probability of individual's migration and education level choice can be written as

$$\begin{aligned}
&\pi_{nt}^0(n',e|n) = (1) \times (2) \qquad (4)\\
&(1) = \frac{\exp(V_t^1(n',e) - D_{n\prime n}^1 - 1_{e=h} F_{n't})^{1/\kappa}}{\sum_{o \in \mathcal{N}} \exp\left(V_t^1(o,e) - D_{n'n}^1 - 1_{e=h} F_{ot}\right)^{1/\kappa}}\\
&(2) = \frac{\exp(V_t^0(n,e))^{1/\psi}}{\sum_{e'} \exp\left(V_t^0(n,e')\right)^{1/\psi}}.
\end{aligned}$$

The first part summarizes the choices between locations conditional on education decision, e. The second part summarizes the decision to study and become a skilled labor later in life.

The movement of labor and its composition is summarized by migration probabilities and the initial distribution. Formally, the migration flows are expressed as:

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

We assume that the last cohort exiting the model is replaced in all locations by a new young cohort entering. The migration probability vary by age denoted by j. Furthermore, individuals who opt to become skilled workers do not work during the first period; instead, they participate in home production and consume their own produce.

2.2 Market clearing and Equilibrium

Profit maximization firms in each location n determine labor input by skill type. Individuals' education level choice and migration choice determine labor supply. Wage at each location, for each individual type are adjusted to clear the labor market.

Equilibrium Given the path of exogenous parameters, including location-specific productivities A_n, 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: $\{\pi_t(n', e'|n)$ and $\pi_t^j(n'|n, e)\}_{t=0}^{\infty}$

(ii) value functions for each location, education type and age $\{V_t^j(n, e)\}_{t=0}^{\infty}$

(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 2, the population flow condition 3, the educational level condition 4 and the goods market clearing condition 1 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 stay unchanged, the individual migration still exists, while the net inflows by cohorts and skill types equal to zero.

We solve the model in levels[2], the equations that characterizes the steady state are in the appendix.

3 Parameterization

3.1 Data Description

We conduct the quantitative exercise on prefecture cites in China. The number of cities is an intersect between the set of cities with available educational resource data (number of teachers in higher education) and the set of cities with available estimates of bilateral transportation cost from Ma and Tang (2022). The quantitative analysis of our model requires the following sets of data for each location: 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

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

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 percent of the total population of China. 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 city. Since we allow the initial population to iterate in the model, the distribution significantly affects the simulation. To address this issue, we use only the educational level and age distribution data from each city 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 a grid based on age and educational level. We only include individuals who are relevant to the model, specifically those aged between 18 and 61. These individuals are then disaggregated into 11 cohorts. We assume that individuals with an education level below high school, or including high school, are categorized as unskilled workers, while individuals with degrees above high school are considered 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.

China 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, including gross domestic product (GDP) and data on education resources. In our analysis, we specifically utilize GDP data from the year 2005 and information on the number of teachers in higher educational institutions for the years 2005 and 2015.

The geographic linkages in the model are summarized by bilateral trade costs and migration costs, which are from Ma and Tang (2022). They comprehensively document the quality of transportation infrastructure in China overtime and estimate trade costs and migration costs from a spatial model. We use their estimations for year 2015, and keep the costs unchanged.

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 are 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 Parameters in the model

In Table 1, we show 2 sets of parameters: The upper panel includes parameters that are either taken from the literature or calibrated directly from the data. The lower panel includes parameters that are calibrated by solving the model to match the observed data.

Table 1: **Parameters**

Symbol	Description	Value	Source
β	Discount rate	0.85	
σ	Elasticity of substitution	4	
κ	Elsaticity of migration	2.55	$3 \times \beta$
ω	Productivity skill complementarity	0.5	Burstein and Vogel (2017)
$\{\tau_{mn} \| m, n = 1, ..., N\}$	Bilateral transportation cost		Ma and Tang (2022)
$\{D_{mn} \| m, n = 1, ..., N\}$	Bilateral migration cost		Ma and Tang (2022)
χ	Weight of input share for low skilled labor	0.93	Calibrated using 2005 wage bill
$\{A_n \| n = 1, ..., N\}$	Productivity in each location		Calibrated using 2005 GDP
ψ	Elasticity of educational migration	5.74	Calibrated with college share
β_1 and β_2	Parameters transform educational cost	$\beta_1 = -5.36$ $\beta_2 = 2$	Calibrated with migration probability using NLS
D_j	Magnitude of migration cost by cohort		Calibrated using 2015 migration matrix
$\bar{D}$	Magnitude of migration cost	0.92	Calibrated using 2015 stay rate

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 estimation methods used.

We directly assign the following parameters of the model: we use a four-yearly discount factor β of 0.85, implying a yearly interest rate of roughly 4% [3]. We assume a elasticity of substitution, σ, equals to 4, as in Tombe and Zhu (2019). We choose κ equals to 2.55 as our baseline, suggested by Kleinman et al. (2023), where they suggest the elasticity of migration should be three times to the discount rate. We assume a value for the productivity skill complementarity of $\omega = 0.5$ as in

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

Burstein and Vogel (2017). We use the set of estimated trade costs and migration costs directly from Ma and Tang (2022).

We utilize the employment data in the 2005 mini-census, along with the 2005 prefectural-level GDP data from the China Statistical Yearbook to calibrate the productivity for each location. Specifically, we first guess productivity levels for each location and the weight of input share for skilled and unskilled labor. Then we compute the output in the model, using the information of labor distribution and wage bills. We iterate on the productivity, A_n, and input share, χ, until our model estimated disaggregated output is matched with the local GDP data.

Education is costly, we estimate the education cost using the data on educational resources. We obtain the number of teachers as a proxy for the abundance of educational resource. The number of teachers in each locations along with the conditional migration flow will be used to estimate the education costs.

To better suit our model, we assign a scalar uniformly to scale the migration costs. We estimate the migration cost scalar by matching the share of individuals choose to 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 from the 2015 mini-census and calibrate the migration cost scalar by different cohorts. We discuss the estimation process below.

3.3 Estimation

The remaining parameters are estimated using a two-step process. Firstly, we structurally estimate the migration costs for each cohort, and education costs. Following this, we employ the selected parameters and the estimated costs to simulate the model, aligning it with the population's stay rate and skill ratio in 2015, thereby backing out the migration cost scalar that applies to all agents in the model and the parameter related to education preferences shocks.

Estimation of migration cost by cohort We estimate the migration cost for each cohort to account for the observed decrease in migration probability as individuals age. Migration is costly in the model, we assume different bilateral costs of migration. Furthermore, we allow the bilateral migration cost to be different for different cohorts, by adding a cohort-specific migration cost scalar. The cohort-specific migration cost scalar captures the Recall the value function and migration flow for individual of cohort j, with skill type e, in location n.

$$V_t^j(n,e) = U_t^j(n,e) + \kappa \log \sum_{o \in \mathcal{N}} \exp\left(\beta V_{t+1}^{j+1}(o,e) - \bar{D}_j D_{on}\right)^{1/\kappa}.$$

$$\pi_t^j(n'|n,e) = \frac{\exp(\beta V_{t+1}^{j+1}(n',e) - \bar{D}_j D_{n'n})^{1/\kappa}}{\sum_{o\in\mathcal{N}} \exp(\beta V_{t+1}^{j+1}(o,e) - \bar{D}_j D_{on})^{1/\kappa}}$$

We rearrange the value function, and write it into a function of migration flow.

$$V_t^j(n,e) = U_t^j(n,e) - \kappa \log \pi_t^j(n'|n,e) + \beta V_{t+1}^{j+1}(n',e) - \bar{D}_j D_{n'n} \tag{5}$$

We arrange Equation 5, and take difference between the log probability of migrating to other location n' from n and the log probability of staying in location n given cohort j and skill type e.

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

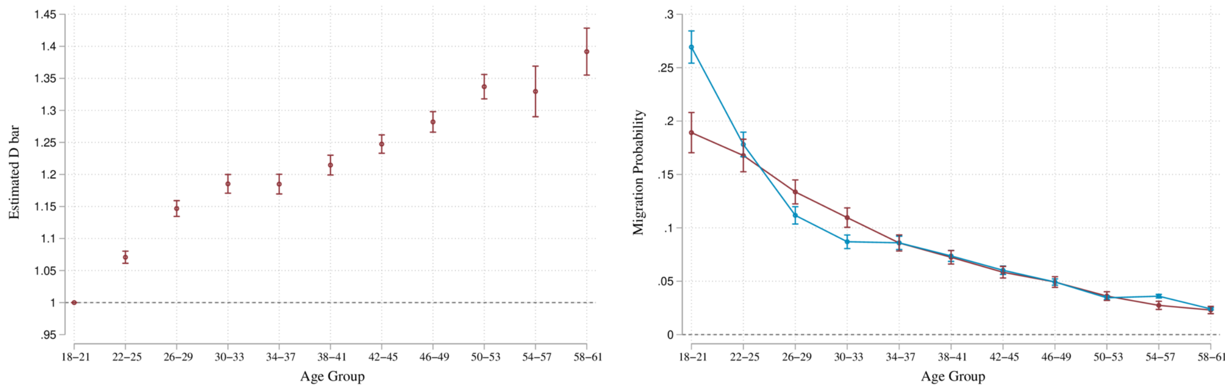
We observe that if we switch n with n', it can further eliminate the expected value and level us with migration cost.

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

Lastly, we take difference of the equation between each cohort and cohort $j = 1$. We observe the data on migration flows from the 2015 mini-census. In this way, the migration cost for each cohort can be estimated.

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

Figure 2 shows the estimated migration cost shifters by cohort where the first cohort serves as a reference point and is normalized to 1. The confidence intervals are computed using bootstrapping. The migration cost for the last cohort in the model is 1.4 times higher than the youngest cohort, suggesting it is harder for older cohort to migrate unless the destination has high real income despite the large migration cost. The right panel of Figure 2 shows the data and model simulated migration probability by cohort. Generally speaking, the model does well, particularly for cohorts in their later ages. The estimated migration cost shifters can well replicate the reality. Consistent with the data, the migration probability decrease for the first cohort to slightly above 3 percent just before the agent exit the model. The model slightly over predicts the migration rate for agents between 18-21, suggesting a close to 27 percent migration rate for the first cohort.



(a) Cost of Migration by Cohort

(b) Probability of Migration by Cohort

Figure 2: Migration by Cohort

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

Estimation of education cost outside of the model We estimate the education cost using number of teachers in each location. The cost function is estimated structurally using non-linear least square. We assume the cost function to be a function of a location's number of teachers in higher educational institutes, with ε_n being the error term.

$$F_n = \beta_2 \times \exp(\beta_1 \times \text{Num.teacher}_n) + \varepsilon_n.$$

We derive the relationship between education costs and conditional migration probability using the migration flow and value functions in the model. The log difference between the probability of individuals migrate from location n to location n' and the probability of individuals stay in location n conditional on getting higher education can be written into a function of expected value function and origin-destination migration costs.

$$\pi_t^{0(1)}(n', e'|n) = \frac{\exp\left(V_t^1(n', e) - D_{n'n}^1 - F_{n't}\right)}{\sum_{o\in\mathcal{N}} \exp\left(V_t^1(o, e) - D_{on}^1 - F_{ot}\right)^{1/\kappa}}$$

$$V_t^0(n, e) = \kappa \log \sum_{o\in\mathcal{N}} \exp\left(V_t^1(o, e) - D_{on} - F_o\right)^{1/\kappa}$$

$$\log\left(\frac{\pi_t^{o(1)}(n', h|n)}{\pi_t^{o(1)}(n, h|n)}\right) = \frac{1}{\kappa}\left(V_t^1(n\prime, h) - V_t^1(n, h) - \tilde{D}_{n'n} - \tilde{D}_{nn}\right),$$

$$\text{where } \tilde{D}_{n'n} = D_{n'n} + F_{n't}.$$

Apply the same trick, the expected value can be eliminated from the equation by switching n and n'. We lose some observations in this process, since for some location pairs we do not observe bilateral migration flows given the age and education level we restricted. But the rest of the observations is sufficient for us to estimate the cost function. We are able to estimate parameters in education cost function using:

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

We plug in the functional form and solve for β_1 and β_2.

The cost function we estimated shows diminishing returns with respect to college concentration. Figure 3 illustrates the estimated educational cost and the number of teachers in 2015. The estimated educational cost decreases drastically when the number of teachers is low, while as resources concentrate, the education 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. In the median city with 2,000 teachers, a 10 percent increase in teachers leads to a 2.34 percent reduction in education costs. However, the return on investment in additional colleges quickly diminishes in better-endowed locations, as indicated by the curvature of the education cost function. For example, at the 90th percentile of cities in terms of education resources, a 10 percent increase in teachers only reduces costs by 0.06 in level, while at the 10th percentile, the same increment leads to a reduction as large as 0.3. This pattern strongly indicates the presence of a substantial number of potential students who would pursue higher education if resources were more accessible. Additionally, the data pattern suggests that in areas with a limited number of colleges, educational costs are likely to be very high. 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 con-

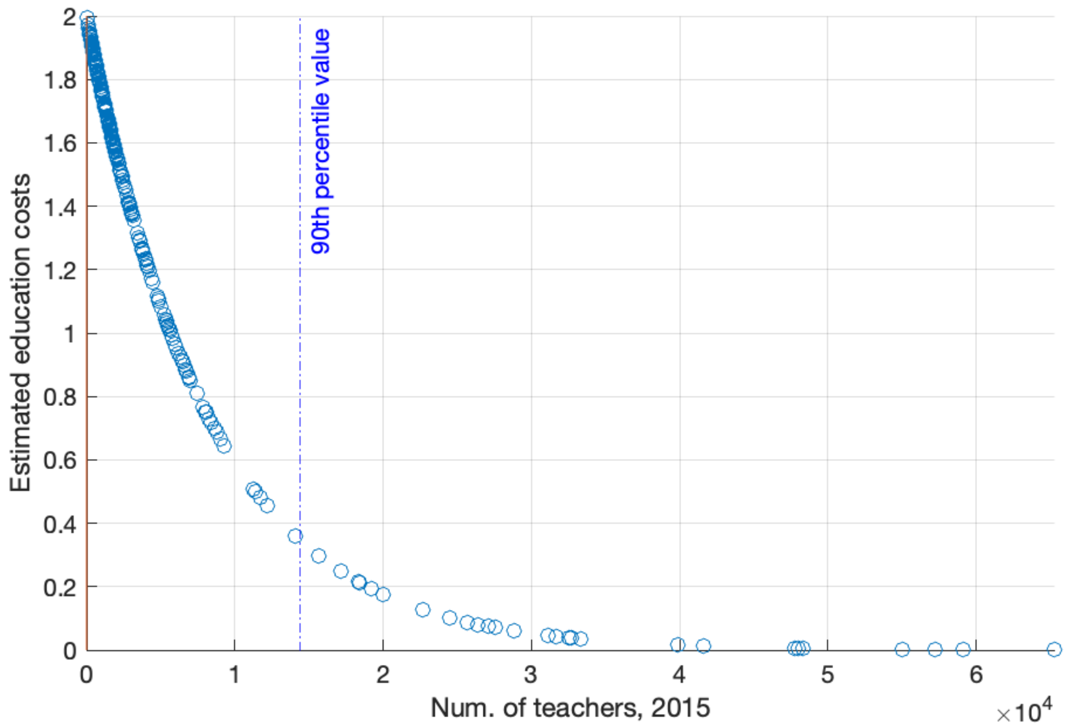


Figure 3: Estimated Education Costs

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

centration of colleges could have a notable negative impact. We elaborate this point in a general equilibrium framework in the counterfactual exercise.

Estimation Procedure Lastly, we are left with two parameters that need to be estimated in the model. We use a two-layer nested Nonlinear Least Squares procedure. In the outer loop, we choose overall migration cost shifter,$\bar{D}$, to match the overall stay rate

$$\sum_{e}\sum_{j>1}\sum_{n\in\mathcal{N}}(\pi^{j}(n|n,e)*L_{n}^{j})+\sum_{e}\sum_{n\in\mathcal{N}}(\pi^{0}(n,e|n)*L_{n}^{0}),$$

in the data. We calculate the 5-year stay rate from the One Percent Population Survey in 2015, and transform it into a 4-year stay rate[4], assuming the same stay rate each year. In the inner loop conditional on $\bar{D}$, we choose $\{\psi\}$, education elasticity, to match the share of individuals go to

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

college when they are young. We borrow migration cost directly from Ma and Tang (2022), but we scale up the migration cost by different cohorts and adjust the benchmark cohort. The identification of the benchmark adjustment scalar $\{\bar{D}\}$ rely 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 choose to stay in the current location instead of pursuing higher wages in other places. The identification of parameter $\{\psi\}$ governs the education choices. We match the model predicted share of college students at the steady state to the ones we observed in the 2015 census. The share of college student in the first cohort is given by

$$\sum_{n' \in \mathcal{N}} \sum_{n \in \mathcal{N}} (\pi_n^0(n', h|n) * L_n^0) / \sum_{n \in \mathcal{N}} L_n^0.$$

We simulate the model at the steady state to estimate the parameters, $\bar{D}$ and ψ. The estimation strategy yields an estimate for migration cost shifter, $\bar{D} = 0.92$, and education elasticity, $\psi = 5.74$[5].

Figure 4 illustrates the destination choices of individuals attending college. The estimated cost of obtaining education is plotted on the x-axis. 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 well generate this pattern without directly targeting the distribution of migrant college student when estimating the model parameters.

4 Counterfactual Simulations

We employ our model to analyze the distribution of educational resources and assess the impacts of college expansion. Additionally, we carry out counterfactual policy experiments to evaluate the results, with a specific focus on changes in the overall skill composition and spatial inequality.

4.1 College Expansion

We first examine both the aggregate and distributional impacts of allocation of education resources through the lify-cycle model. The first experiment examines directly the impact of the college expansion which happened in the beginning of this century. We set the college resources

[5] 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.

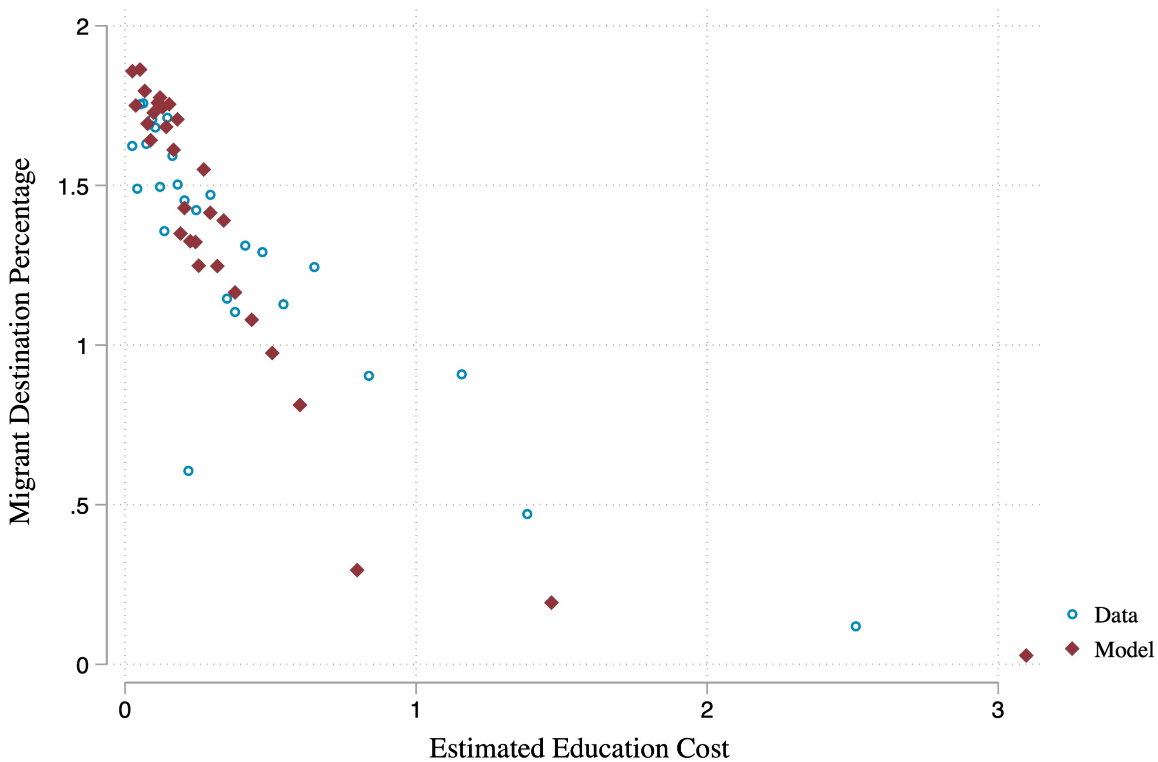
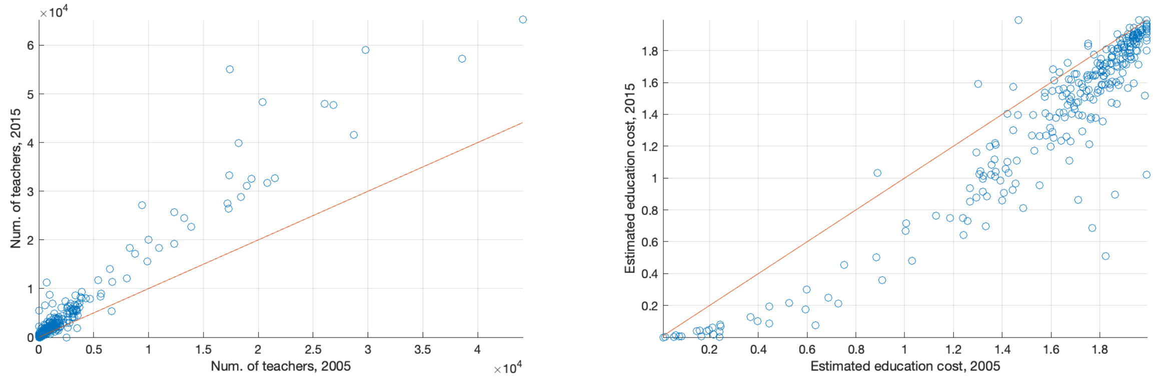


Figure 4: Migrants Destination Percentage

Note: This figure shows the binned scatter plot of estimated education cost and the destination choice for both data and model prediction. The number of bins is 30. The education cost is plotted on the x-axis, the percentage of students chose the place as the destination conditional on going to college in a place different than the home location is plotted on the y-axis. Each dot represents a group of observations with similar estimated education costs.



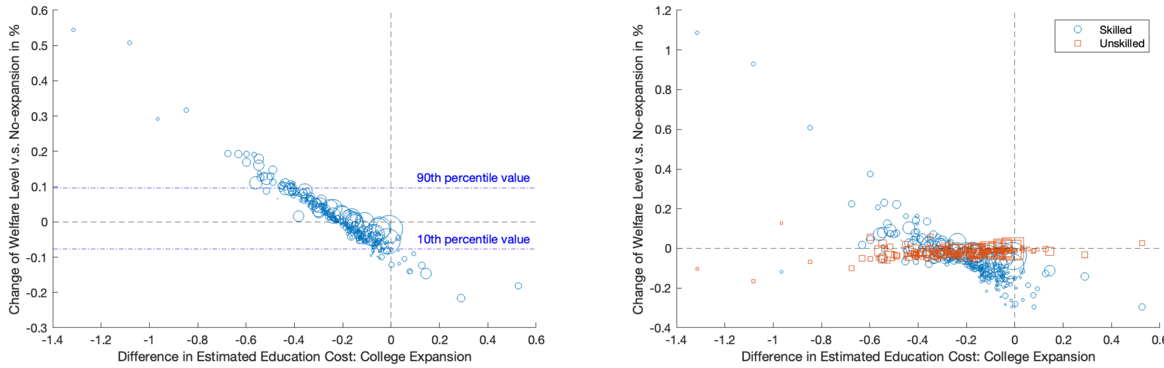
(a) Education Resources (b) Estimated Education Costs

Figure 5: Factual College Expansion

Note: Figure 5a shows the factual educational resources in terms of number of teachers hired in higher education institutes by locations. Figure 5b shows estimated education costs by locations. The number of teachers hired in higher education institutes comes from China Statistical Yearbook in 2005 and 2015. The red line represents the 45 degree line.

fixed at the beginning year of our analysis (2005), such that the cost of obtaining higher education does not vary over time. By comparing the scenario when there is no expansion with the factual scenario, we shed light on the distributional impact of college expansion. Figure 5 shows the actual expansion happened in place in 2005 to 2015 where we observe a disproportional allocation of resources in those places that already have a higher share of resources. Using the estimated cost, Figure 5b indicates that the concentration of the resources does not significantly reduce costs. The most significant change in education costs occurs in cities with median educational resources. Furthermore, the cost function we estimated suggests that a slight increase in areas with fewer resources leads to a substantial drop in the education cost. However, in the factual expansion, these areas receive minimal new resources. The increase of the top 10 percent of cities in terms of number of teachers is approximately 29 times higher than the bottom 10 percent. Overall, there is an 84.23% increase in the number of teachers, which leads to a 12.77% reduction in estimated education cost.

Table 2 panel A reports the overall welfare impact and distributional impact of college expansion. By the year 2015, the population-weighted welfare gains are around 0.02% and the skill ratio increased 0.29%. Both skilled and unskilled workers experience welfare loss. The overall welfare slightly increased, due to the level of welfare of skilled worker is higher than the unskilled worker. We see the labor composition shifts more to skilled labor, even though the welfare level of skilled labor decreased. The college expansion decreases the cost of becoming a skilled worker, attracting more individuals choose to go to college. Thus, the expansion creates a supply shock of skilled worker. The impact of the shock would eventually leads to a 0.16% welfare loss of the skilled



(a) Change in Welfare

(b) Change in Welfare by skill level

Figure 6: College Expansion Impacts

Note: Figure 6a shows the welfare change by locations. Figure 6b shows the welfare change of skilled and unskilled by loations. All changes are compared against the scenario without college expansion. Each dot represents a location. The size of the dots indicates the initial level of education resources, in terms of number of teachers in higher education institutes.

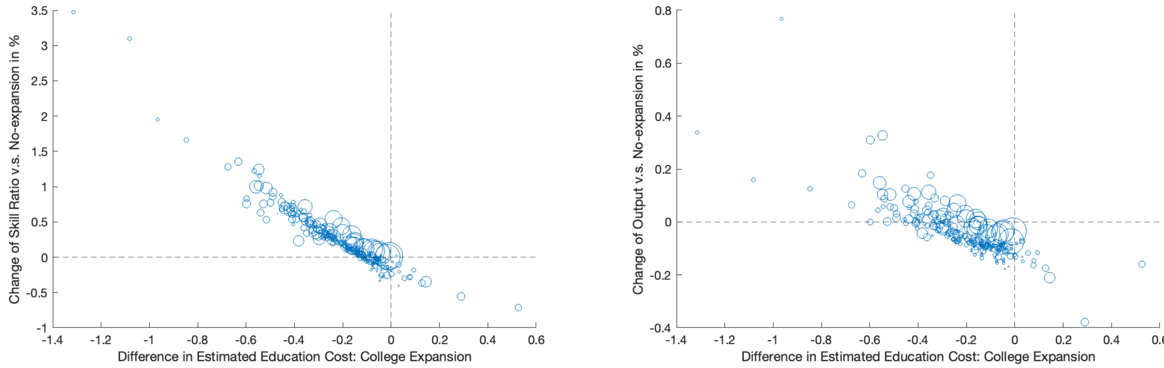
worker and a 0.16% gains of unskilled worker, when the economy transits to steady state. The impact of college expansion does not evenly affect all locations. The dispersion of overall welfare by locations measured by the coefficient of variation increased by 0.16% compared to no expansion.

Figure 6 and Figure 7 further present the impacts by locations. Each bubble represents a location, the size of the bubble represents the initial value of the educational resources before the college expansion. The welfare impact are mainly driven by the welfare change of skilled workers, as shown in Figure 6b, where the welfare change for unskilled workers is close to 0. Welfare improvement is greatest in areas that start with a moderate level of resources. In these locations, the reduction in education costs is most significant and, as a result, they can attract more high-skill workers.

The limited response to college expansion is expected. This is due to the diminishing returns associated with the concentration of educational resources, resulting in minimal changes in the average cost of attending college. Thus, we seek alternative allocations to better allocate educational resources across space.

4.2 Alternative Allocations

The factual college expansion results suggests that the allocation of educational resources has varying effects on different locations. To analyze this, we first calculate the elasticity of welfare with respect to college expansion, examining each prefecture individually. More specifically, we increment the actual number of college teachers by 10% in each location and then assess the change



(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 output chage by loations. All changes are compared against the scenario without college expansion. Each dot represents a location. The size of the circle indicates the initial level of education resources, in terms of number of teachers in higher education institutes.

in welfare compared to a scenario with no expansion. We categorize the prefectures based on their initial educational resources. Notably, the overall impact of a 10% increase in college teachers is significantly greater in the bottom 10% of prefectures than in the top 10%, when compared to the no expansion scenario. In the Appendix, Figure **??** displays the aggregate welfare impact for each location based on the location-specific shock experiment.

We then conduct another simulation where we keep the total increment of education resources unchanged and distribute the increment evenly across all locations. We call this simulation *"equal growth"*. In this case, all prefectures equally receive an additional 2600 college teachers. Figure 8 shows the counterfactual cost of education in this scenario. The evenly distributed resources keeps the relative rank of education costs untouched while substantially lower the cost of the prefectures where the costs are originally high. Unsurprisingly, the *"equal growth"* scheme leads to a even higher skill ratio change. Since the average education cost is 27% lower, comparing to the no expansion case, and 16.4% lower than the factual expansion case. Table 2 Panel B shows the aggregate welfare of this scheme is five times higher than the actual expansion program and the skill ratio is more that 10 times higher than the observed expansion at 3.57%.

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

Panel A: Impacts of College Expansion					
		ΔWelfare	ΔDispersion	10th percentile	90th percentile
Overall	$T = 3$	0.02%	0.16%	-0.08%	0.10%
	S.S.	0.03%	0.17%	-0.07%	0.02%
Skilled	$T = 3$	-0.02%	0.11%	-0.18%	0.05%
	S.S.	-0.16%	0.88%	-0.52%	-0.08%
Unskilled	$T = 3$	-0.01%	0.11%	-0.04%	0%
	S.S.	0.16%	0.10%	-0.12%	0.14%
		ΔSkill Ratio	ΔDispersion	10th percentile	90th percentile
Overall	$T = 3$	0.29%	0.07%	-0.13%	0.69%
	S.S.	1.65%	0.82%	0.59%	2.68%
Panel B: Impacts of Equal Growth					
		ΔWelfare	ΔDispersion	10th percentile	90th percentile
Overall	$T = 3$	0.17%	1.08%	-0.26%	0.36%
	S.S.	0.16%	3.11%	-0.61%	0.03%
Skilled	$T = 3$	0.23%	-0.95%	-0.46%	0.04%
	S.S.	-1.00%	2.15%	-1.54%	-1.11%
Unskilled	$T = 3$	-0.13%	1.97%	-0.75%	0.28%
	S.S.	0.95%	3.71%	0.05%	0.84%
		ΔSkill Ratio	ΔDispersion	10th percentile	90th percentile
Overall	$T = 3$	3.57%	-3.65%	2.28%	5.82%
	S.S.	8.34%	-17.37%	5.59%	13.09%

Notes: This table illustrates the changes in the levels and dispersion of both welfare and labor composition in both 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 educational resources grow equally in all locations. 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.

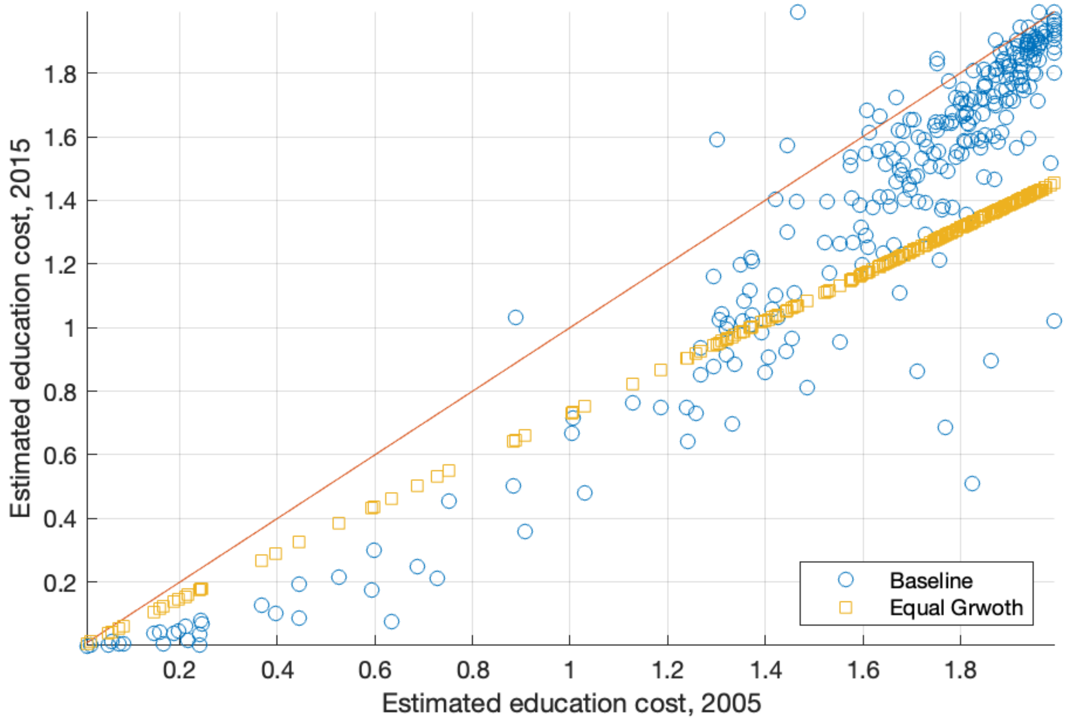


Figure 8: Factual College Expansion

Note: This figure shows the estimated education costs for college expansion and counterfactual "Equal Growth" in 2005 and 2015. The additional education resources are distributed evenly across all locations. Each dot represents a location.

Equal growth generates a bigger supply shock of skilled worker. This benefits unskilled workers, thus generating a higher welfare impact than college expansion. Since unskilled workers are a large share of labor force. Figure 8 demonstrates this point by showing the welfare impact by skilled type and location, with the size of the dots indicating the population. On the x-axis, we show the estimated educational cost under this "equal growth" scenario.

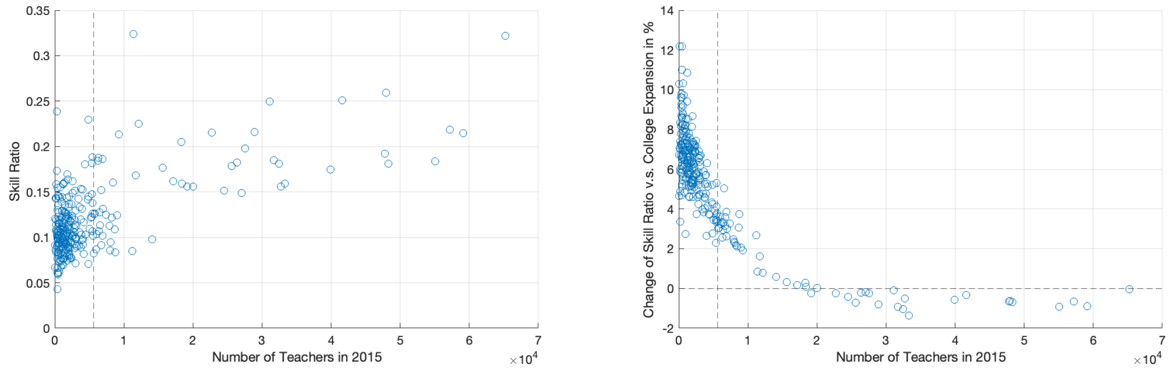
4.3 Equalization

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 20% of the observed spatial inequality in skill composition. 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

Figure 9: Euqal Growth

Note: This figure shows the welfare change in the "Equal Growth" scenario by skill level in each location comparing to the no expansion scenario.



(a) Skill Ratios by location

(b) Change in Skill Ratios

Figure 10: Education resources equalization

Note: Figure 10a shows the skill ratio by locations. Figure 10b shows the skill ratio chage by locations. All changes are compared against the scenario with college expansion. Each dot represents a location.

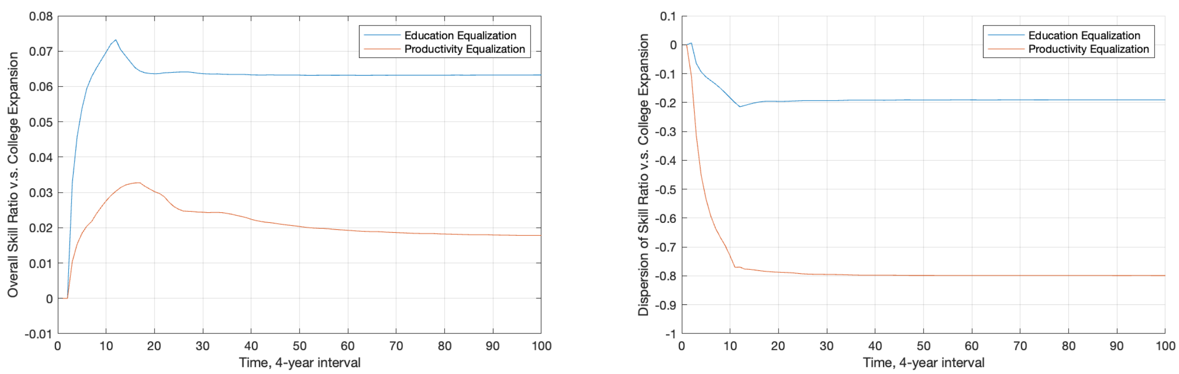
3, the skill ratio dispersion decreases by approximately 6% to 20% compared to the baseline results along the transition path. Figure 10 illustrates the changes in skill ratios in various locations under this equalization scenario. In Figure 10a, the x-axis represents the number of teachers in the actual college expansion scenario, while the y-axis depicts skill ratios across locations in the equalization scenario. The dashed vertical line indicates the number of teachers in the 'Equal college' scenario. Meanwhile, Figure 10b displays the changes in skill ratios in comparison to the college expansion scenario, plotted on the y-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 31% to 80% along the transition path to the steady state. Figure 11 illustrates the skill ratio change dispersion relative to the baseline college expansion over time for both educational equalization and productivity equalization. An evenly distributed educational resource is 25% as effective at reducing spatial inequality as an evenly distributed productivity.

Table 3: Summary of Counterfactual Exercises: Skill Ratio

Panel A: Impacts of Education Equalization					
		ΔSkill Ratio	ΔDispersion	10^{th} percentile	90^{th} percentile
Overall	$T = 3$	3.27%	-6.48%	0.53%	8.18%
	S.S.	6.33%	-19.07%	2.86%	12.48%
Panel B: Impacts of Productivity Equalization					
		ΔSkill Ratio	ΔDispersion	10^{th} percentile	90^{th} percentile
Overall	$T = 3$	1.04%	-31.43%	-7.65%	37.85%
	S.S.	1.73%	-79.88%	-0.95%	33.40%

Notes: This table illustrates the changes in the levels and dispersion of labor composition in both 2015 and the steady state, compared to the college expansion benchmark. In Panel A, we present scenarios reflecting education equalization, while Panel B depicts a situation where productivity are equalized to the mean in all location. The changes are calculated in comparison to a scenario in which we maintain the factual college expansion. Welfare is weighted by the population, and we also present a dispersion measure using the coefficient of variation.



(a) Overtime Skill Ratios

(b) Overtime Skill Ratio dispersions

Figure 11: Education Equalization and Productivity Equalization

Note: Figure 11a shows the overtime overall skill ratios. Figure 11b shows the dispersion of skill ratios overtime. All changes are compared against the scenario with college expansion.

5 Concluding Remarks

This paper integrates educational choices into a dynamic spatial model to examine how location-specific educational resources affect spatial inequality in China. 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 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 five-fold increase in the welfare impact, and the impact on skill ratio is more than ten times higher than the observed expansion.

Furthermore, we show that the unequal distribution of educational resources accounts for as much as 20% of the observed spatial disparity in skill composition. Equalizing educational resources is roughly 25% as effective as equalizing fundamental productivity.

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 (2022). The distributional impacts of transportation networks in china. *Available at SSRN 4118287*.

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:

$$
\begin{aligned}
p_{ni} &= \tau_{ni}\frac{1}{a_i}\left[a_i^{-\omega(\eta-1)}\chi {w_{it}^{l}}^{1-\eta} + a_i^{\omega(\eta-1)}(1-\chi){w_{it}^{h}}^{1-\eta}\right]^{\frac{1}{1-\eta}} \\
\alpha_{ni} &= \frac{p_{ni}^{1-\sigma}}{\sum_o p_{no}^{1-\sigma}} \\
P_n &= \left(\sum_i p_{ni}^{1-\sigma}\right)^{\frac{1}{1-\sigma}} \\
\mathbb{L}(n,h)w_n^h + \mathbb{L}(n,l)w_n^l &= \sum_o \alpha_{on}\left(\mathbb{L}(o,h)w_o^h + \mathbb{L}(o,l)w_o^l\right) \\
V^j(n,e) &= U^j(n,e) + \kappa \log \sum_{o\in\mathcal{N}} \exp\left(\beta V^{j+1}(o,e) - D_{on}^{j+1}\right)^{1/\kappa} \\
V^J(n,e) &= U^J(n,e) \\
V^0(n,e) &= \kappa \log \sum_{o\in\mathcal{N}} \exp\left(V^1(o,e) - D_{o,n}^1 - 1_{e=h}F_o\right)^{1/\kappa} \\
\pi^j(n'|n,e) &= \frac{\exp(\beta V^{j+1}(n',e) - D_{n'n}^{j+1})^{1/\kappa}}{\sum_{o\in\mathcal{N}}\exp(\beta V^{j+1}(o,e) - D_{on})^{1/\kappa}} \\
\pi^0(n',e|n) &= \frac{\exp\left(U^1(n',e) - D_{n'n}^1 - 1_{e=h}F_{n'}\right)^{1/\kappa}}{\sum_{o\in\mathcal{N}}\exp(U^1(o,e) - D_{on}^1 - 1_{e=h}F_o)^{1/\kappa}} \times \frac{\exp(V^0(n,e))^{1/\psi}}{\sum_{e'}\exp\left(V^0(n,e')\right)^{1/\psi}} \\
L_n^0 &= L_n^J \\
L^1(n',e') &= \sum_n L_n^0 \pi(n',e'|n) \\
L^j(n',e) &= \sum_n L^{j-1}(n,e)\pi^{j-1}(n',e|n) \quad j = 2,...,J \\
\mathbb{L}(n,e) &= \mathbb{L}_{n,h} + \mathbb{L}_{n,l}
\end{aligned}
$$

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 $\{L_t^j(n,e)\}$ for a long enough period.

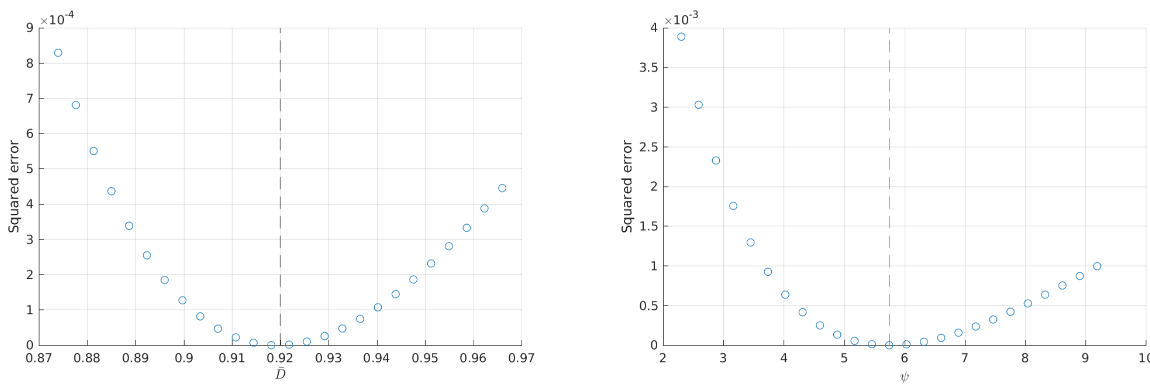
2. Solve for $\{w_{nt}^h, w_{nt}^l\}$ using the market clearing condition along with the relationship between w_{nt}^h and w_{nt}^l.

3. Compute recursively the value functions $\{V_t^j(n, e)\}$ at each location and time period for each individual type using equation 2.

4. Compute migration probability $\{\pi_t n', e'|n\}$ and $\{\pi_t^j(n'|n, e)\}$ using equation 4 and equation 3.

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

A.3 Additional figures



(a) $\bar{D}$ Objective function

(b) ψ Objective function

Figure A.1: Local plot for Nested NLS Estimation

Note: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 dotted line indicates the estimated value for $\bar{D}$ andψ.

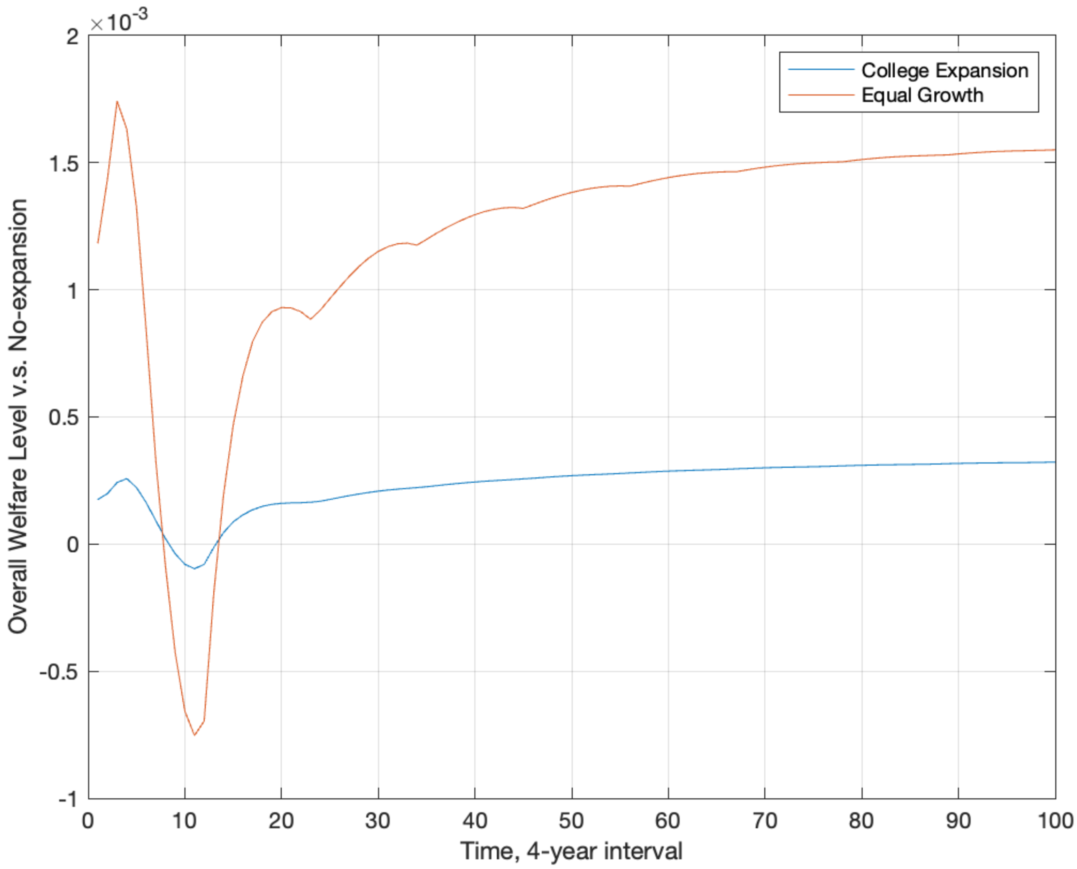


Figure A.2: Welfare change overtime

Note: This figure shows the population-weighted welfare change overtime for college expansion and equal growth scenarios comparing to the no expansion scenario.

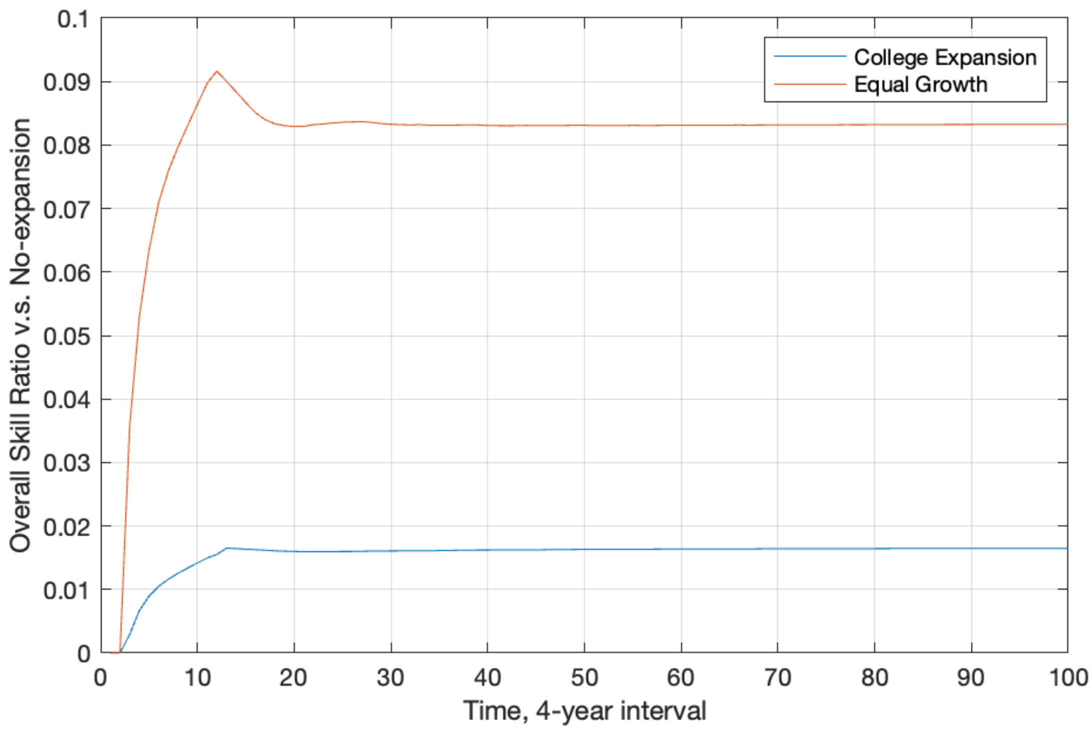


Figure A.3: Skill ratio change overtime

Note: This figure shows the skill ratio change overtime for college expansion and equal growth scenarios comparing to the no expansion scenario.

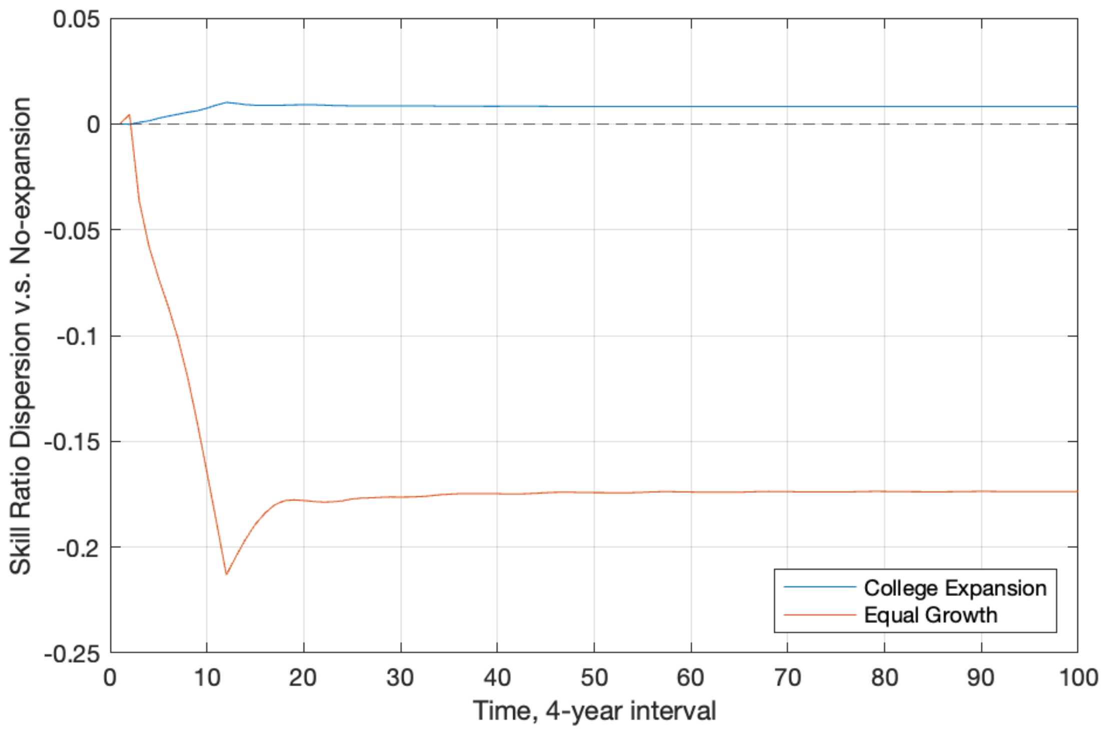


Figure A.4: Skill ratio dispersion change overtime

Note: This figure shows the dispersion change of skill ratio overtime for college expansion and equal growth scenarios comparing to the no expansion scenario.

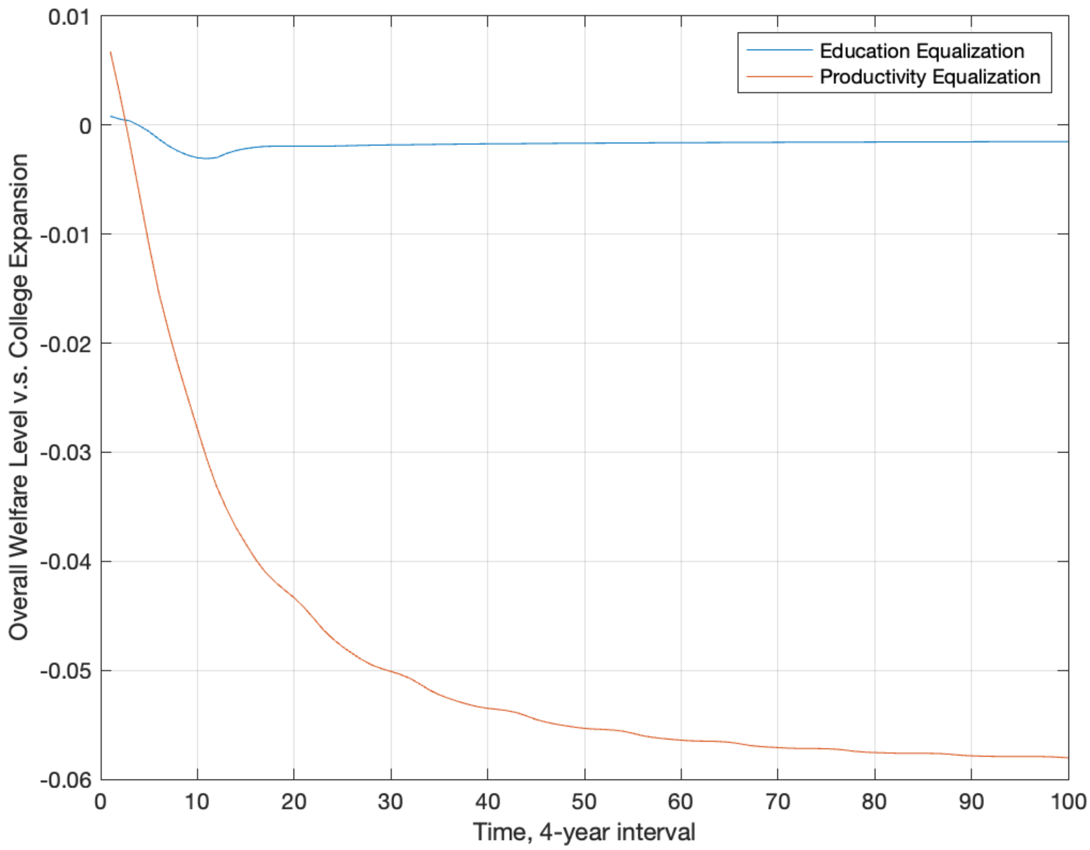


Figure A.5: Welfare change overtime

Note: This figure shows the population-weighted welfare change overtime for education equalization and productivity equalization scenarios comparing to the no expansion scenario.

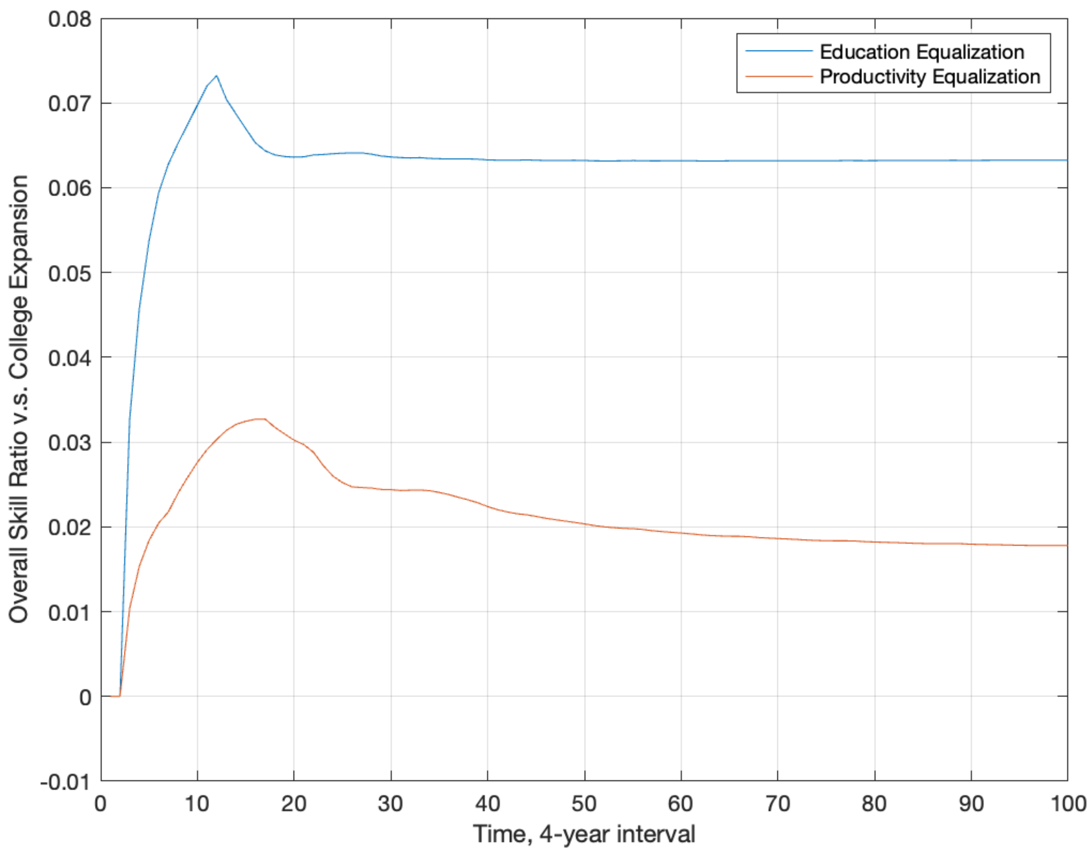


Figure A.6: Skill ratio change overtime

Note: This figure shows the skill ratio change overtime for education equalization and productivity equalization scenarios comparing to the no expansion scenario.

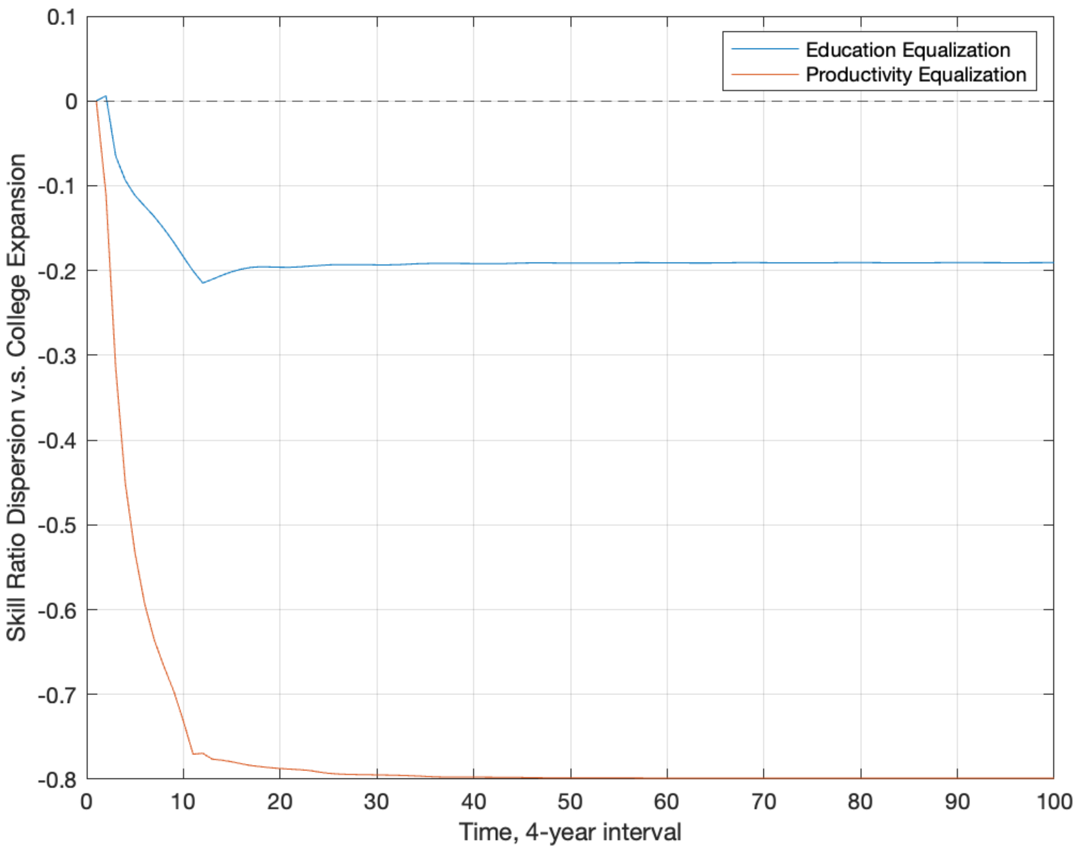


Figure A.7: Skill ratio dispersion change overtime

Note: This figure shows the dispersion change of skill ratio overtime for education equalization and productivity equalization scenarios comparing to the no expansion scenario.