Higher Education Policies and Spatial Inequality Across US States

Sang Min Lee and Bipul Verma*

December 19, 2024

Link to the latest version

Abstract

To what extent have the federal and state higher education policies—which have affected the supply and demand of college education—contributed to the growing cross-state wage inequality since 1980? To address this question, we develop and estimate a quantitative spatial model with endogenous college education and migration. The model features skill-biased technical changes and agglomeration effects, which means that a state's average wage rises as more residents attain college degrees or as college-educated workers migrate into the state. Counterfactual analyses reveal that tuition discounts and public appropriations by the federal and state governments have become more progressive. In the baseline, the standard deviation of log mean wages across states increased from 0.087 in 1980 to 0.117 in 2019. In contrast, had the higher education policies remained at their 1980 levels, the standard deviation would have risen by 73% more to 0.139 by 2019. Agglomeration effects played a significant role in this result: if they had been half as strong as estimated, the standard deviation would have increased by 17% more (instead of 73%) in the counterfactual than in the baseline.

JEL Codes: E24, I23, R11

Keywords: College education, higher education policies, wage inequality, quantitative spatial model

*Lee: Department of Economics, Darla Moore School of Business, University of South Carolina (email: sangmin.lee@moore.sc.edu); Verma: Department of Economics, University of Minnesota; and Federal Reserve Bank of Minneapolis (email: verma176@umn.edu). The views expressed in this paper are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Minneapolis or the Federal Reserve System. Lee acknowledges financial support from the DMSB Research Grants and the Excel Funding Program at the University of South Carolina. The authors have no other sources of research support, financial relationships, or conflicts of interest to disclose.

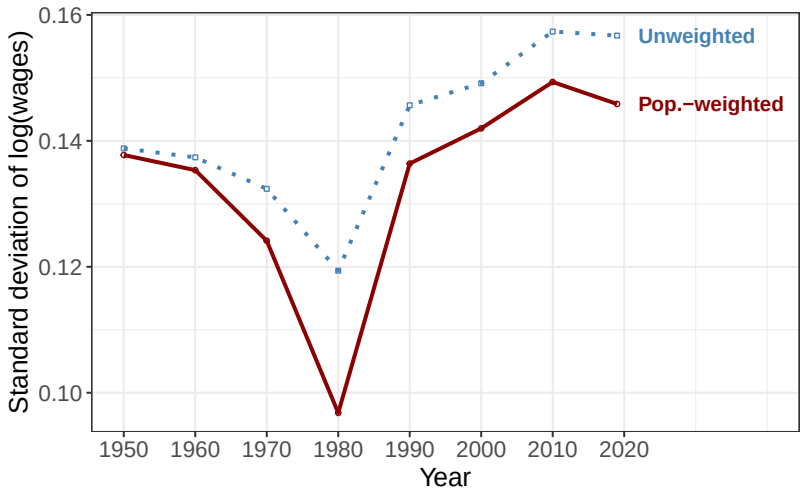
1 Introduction

Prior to 1980, wage inequality across 50 US states and the District of Columbia had been declining. However, post-1980, this convergence has disappeared (Figure 1).[1] To mitigate the rising spatial inequality, several industrial policies have been recently proposed, such as Invest in America, Build Back Better Regional Challenge, and Chips for America. In contrast, higher education policies are relatively underutilized as a tool to address rising geographic inequality, despite the increasing wage premium for college graduates, or equivalently, high-skilled workers.

We ask: "To what extent have the federal and state higher education policies—which have affected the supply and demand of college education—contributed to the growing cross-state wage inequality since 1980?" By developing and estimating a quantitative spatial model with endogenous college education, we demonstrate that the higher education policies have become more progressive. In the baseline, the standard deviation of log mean wages across states increased from 0.087 in 1980 to 0.117 in 2019.[2] In contrast, had the policies remained at their 1980 levels, the standard deviation would have risen by 73% more to 0.139 by 2019.

Our main contribution is to uncover the quantitative significance of college education in

Figure 1: Standard Deviation of Log Mean Wages Among US States Across Time



* Source: Decennial Census and American Community Survey (ACS).
* Note: "Pop.-weighted" denotes the weighted standard deviation with total state employment in 1980 as weights.

[1] Figure 1 illustrates the pattern of sigma convergence, which is assessed by the variance of wages across states. In Appendix A, we also show that beta convergence, where initially poorer states exhibit faster subsequent wage growth than richer states, was present prior to 1980, but disappeared post-1980.

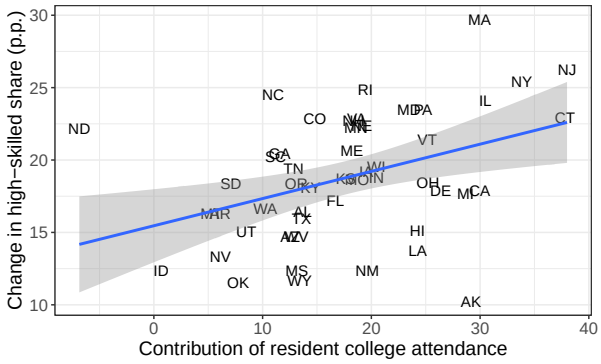
[2] The values of the standard deviation differ from those in Figure 1 mainly because, for the model parameterization, data were processed to align with the model's definitions and structure, whereas the Figure presents raw values directly from the data.

understanding the rise in spatial inequality by building a novel and tractable spatial model with supply-and-demand structures of higher education. This contrasts with the existing literature's focus on the migration of high-skilled workers. The literature has documented very well that the immediate causes behind the rising inequality are the increasing cross-region disparity in college-educated workers' wages and in their share in the labor force.[3] To explain these, recent papers have highlighted the agglomeration of high-skilled labor through migration (e.g., Ganong and Shoag 2017; Giannone 2017; Eckert 2019; and Rubinton 2020).

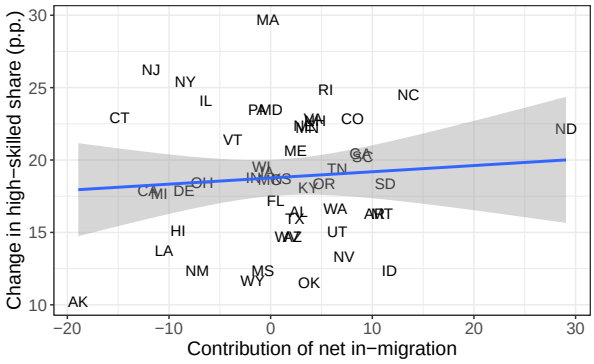
Our focus on college education is grounded in empirical evidence, not merely theoretical interest. We have discovered that for a state-level analysis of high-skilled agglomeration, college attendance by residents is quantitatively more important than the net migration of high-skilled workers.[4] Figures 2a and 2b break down the growth in the share of college graduates in a state (y-axis) into the part due to the growth in the residents' college attainment (x-axis of (a)) and the portion due to the net migration of the high-skilled (x-axis of (b)). The figures show that the former is correlated with the growth of the high-skilled share, while the latter is not. This suggests that the barriers to interstate migration are large enough that college education is likely a key factor behind the growing spatial inequality.

Figure 2: Decomposition of the Growth in the High-Skilled Share from 1980 to 2019

(a) Residents' College Attendance Component



(b) Net Migration Component



* Source: IPUMS-USA (Ruggles et al. 2021): Decennial Census and ACS
* Note: The high-skilled share of a state is defined by (the number of college-graduated workers born in the state + the number of college-graduated workers migrated to the state)/(the size of the labor force). Changes are expressed in percentage points (p.p.).

[3] In Appendix A, we demonstrate that the same holds in our data.

[4] This is not to claim that the assumptions of other papers are incorrect. Most papers have focused on granular geographical units, such as Metropolitan Statistical Areas. However, to examine the role of regional skill composition, a state-level analysis is particularly relevant due to the prevalence of state colleges. To illustrate the difference with an example, an individual who was born in Bakersfield, California, went to college in Los Angeles, and is working in San Jose would count as a migrant to San Jose with granular geography, but the hypothetical person is a resident of California with state-level analysis.

In our model, which captures the supply and demand factors of the higher education market highlighted by the existing literature (e.g., Gordon and Hedlund 2022), individuals decide whether and where to go to college, taking into account the tuition they pay and the future wages they earn. Then, they choose where to work. A representative college in a state incurs costs based on the total enrollment and charges markups based on the degree of its market power. The federal and state governments raise labor income taxes and provide tuition discounts to individuals and production subsidies to colleges.

Another key feature of our model is the agglomeration effects in the spirit of Diamond (2016) and Davis and Dingel (2020), where a state's high-skilled productivity rises when the workforce share of the high-skilled or the total population of the state increases.[5] Our model also incorporates skill-biased technical change (SBTC, Katz and Murphy 1992), where the productivity growth of the low-skilled and that of the high-skilled differ due to variations in the evolution of their fundamental productivities or due to the agglomeration effects.

In the presence of the agglomeration effects and SBTC, a state's mean wage increases when it accumulates high-skilled workers either through residents' college attendance or through the net migration of college-educated workers. Theoretically, policies promoting higher education can impact wage growth in either direction. For example, higher tuition subsidies provided by a state government to its in-state students can decrease the state wages, possibly due to the leakage (out-migration) of college students after graduation, along with increased tax burdens associated with the policy. Therefore, how the policies impact the rising inequality is a quantitative question, and we address it with our model.

To parameterize the model, we compile the data on individuals' education- and work-related choices from sources, such as the Decennial Census and the ACS, and the data on colleges' balance sheets from the Integrated Postsecondary Education Data System (IPEDS) by the National Center for Education Statistics (NCES). Notably, the model primitives regarding higher education policies directly map to actual policy variables in the colleges' balance sheets. Consistent with the higher education literature, we find that tuition discounts have been rising and that public appropriations have been stagnating. This implies that the policies, in general, have been increasing the demand while decreasing the supply in the college education market.

Using the data, we first estimate the elasticities that govern how individuals' choices respond to changes in the equilibrium prices by minimizing the distance between the choice probabilities in the model and those in the data. Furthermore, we estimate the strength of the agglomeration effects using the model equation that exposes the relationship between

[5] As pointed out by Giannone (2017), the agglomeration effects rationalize the data pattern that a region with faster growth in high-skilled wages also exhibits faster accumulation of high-skilled labor.

the skill premium (i.e., the wage gap between low- and high-skilled workers) and the number of low- and high-skilled workers in a state. Then, using the structural equations of the estimated model, we obtain the degree of each college's market power and the fundamental productivities—productivities net of the agglomeration effects—for the low- and high-skilled workers.

With the parameterized model, we run the counterfactual exercises to assess the role of policies in the rise of cross-state wage inequality. In the baseline, the inequality measured by the standard deviation of log mean wages across states weighted by the population in 1980 increased from 0.087 in 1980 to 0.117 in 2019.[6] Counterfactual exercises reveal that if tuition discounts provided to students and production subsidies provided to colleges by the federal and state governments had remained at the 1980 levels, the inequality would have increased faster to 0.139. This implies that the inequality would have risen 73% faster.

For this result, the agglomeration effects—which is friction in our model since individuals do not internalize the impact of their choices on high-skilled productivity—played a significant role. If we set the parameters for the agglomeration effects as halves of the estimated levels, the inequality would have increased 17% faster in the counterfactual than in the baseline. In the extreme case of the no-agglomeration case, the counterfactual exhibited a standard deviation of 0.118, implying that higher education policies had a minimal impact. The agglomeration effects act as an amplifying force. A higher number of college graduates in a state, all else being equal, increases the skill premium in the state. Therefore, for the average wage of the region, not only are higher weights given to high-skilled wages (which tend to be higher than low-skilled wages), but also the skill gap of wages increases, which further increases the mean wage. In addition, higher skill premium induces higher rates of college attendance by future generations.

By studying how each of the four policy factors—federal tuition discounts, state tuition discounts, federal college production subsidies, and state college production subsidies—affects a state's wage growth, we discover a significant difference between the federal and state policies. Higher discounts or subsidies provided to a state by the federal government lead to higher future wage growth, since the state becomes more high-skilled intensive. However, more generous state higher education policies do not necessarily lead to the wage growth in the state. The difference arises due to how federal and state tax revenues are raised. The increased tax burdens due to more generous federal policies are shared by all states, whereas those due to more generous state policies are levied on the state residents. Therefore, as for the federal policies in a state, the benefits accrue to the state, whereas the costs are shared by all 51 states (50 states plus the D.C.). In contrast, changes in state policies significantly

[6] Using the unweighted standard deviation provides the same qualitative results.

impact the tax burdens, affecting the education and migration choices of individuals. How a state should alter its higher education policy to achieve wage growth depends on multiple factors, such as whether it is more prone to the out-migration of high-skilled workers. Counterintuitively, for some states in some periods of time, it is optimal to cut back on their higher education policies for wage growth. This characteristic also reflects the additional value of our model: it is capable of identifying optimal policies while accounting for equilibrium effects associated with the policies.

Despite not being our main focus, we also assess the extent to which policies affected the increase in college tuition, which received significant interest from both policymakers and academic researchers. We demonstrate that in the absence of policy changes post-1980, the median college tuition across US states would have increased by 6% from 1980 to 2019 instead of 36% in the baseline.[7] This result implies that the changes in the policies explain approximately 83% of the tuition hike post-1980, which echoes the result of Gordon and Hedlund (2022).

The rest of the paper is organized as follows. Section 1.1 reviews the related literature. Section 2 outlines the model. In Section 3, we discuss the data and model estimation. Section 4 discusses the counterfactual exercises. Lastly, we conclude in Section 5.

1.1 Related Literature

This paper relates to three broad strands of literature: the evolution of spatial inequality in the US, structural models of higher education, and the empirical analysis of the impact of higher education policies on educational attainment.

Prior to 1980, there was a strong wage convergence across US states, first documented in the seminal work by Barro and Sala-i Martin (1992). However, recent evidence shows the stoppage of convergence after 1980 (e.g., Ganong and Shoag 2017; Giannone 2017; Rubinton 2020; and Eckert 2019). In contrast to the prevailing narrative based on the migration of college-educated workers, our paper provides a novel empirical finding that residents' college education, rather than migration, is a key factor behind the overall increase in the share of college graduates in a state. We further contribute to the literature by developing a quantitative spatial model with endogenous education to assess the role of higher education policies. Notably, this shift in emphasis from mobility-based explanations to local skill production is policy-relevant. Given the importance of higher education, policies designed to improve in-state college enrollment and completion may be more effective than policies

[7] We adopt a broad notion of tuition incorporating both tuition and the sales of services (e.g., room and board).

targeted at attracting out-of-state talents in increasing long-run human capital stocks and, subsequently, wage levels.

Our paper also contributes methodologically to the literature on the structural models of higher education. Cai and Heathcote (2022) and Gordon and Hedlund (2022) have studied the underlying causes behind the upward trajectory of tuition costs since the 1980s. Kennan (2015) shows how cross-state differences in college subsidies influence the migration decisions of college graduates, thereby shaping the spatial distribution of human capital. Epple, Romano, and Sieg (2006), Fu (2014), and Epple et al. (2017) develop equilibrium models of college markets to study the competition and interaction among heterogeneous colleges and students. Our model complements the existing models in two ways. First, our model focuses on long-term economic growth aspects of higher education by incorporating skill-biased technical change, agglomerations, and increasing skill premium (e.g., Acemoglu 2007; Autor 2014; and Diamond 2016). This enables us to study the impact of higher education policies on wage growth. To illustrate further, previous studies emphasized the role of rising skilled wages in increased college attendance. Yet, wages themselves are not fixed. They depend on the skill composition of the labor force, which is, in turn, influenced by college attendance. This interdependence suggests that labor and higher education markets must be analyzed within a unified, general equilibrium framework, such as ours. Second, the spatial details of our model, where individuals can migrate for education and work, allow us to realistically study the impact of policies, as they can account for leakage or inflow of local talents.

Lastly, our work relates to the empirical literature on the role of federal and state higher education policies in educational attainments, to which we contribute through a multi-region, general-equilibrium model. For example, Deming and Walters (2018) document how state- and federally-funded financial aid programs and changes in state appropriations to public universities affect college enrollment and completion rates. More recently, Burland et al. (2023) note the importance of policy interventions in shaping the educational attainment of local populations.

2 Model

For brevity, we first lay out a simplified version of the model in its steady state without time notation. Later, we discuss the structure of the full model and how we match its transition path to the data.

2.1 Environment

$\bar{L}$ measure of individuals populate N locations (regions). The set of locations is denoted by $\mathcal{N}$. Location n is characterized by the level of amenities it provides for the residents, $A_a(n)$, fixed housing supplies, $H(n)$, and moving costs to and from the different locations. Throughout the rest of the paper, we suppress the set notation for conciseness whenever there is no confusion. For example, $\sum_n$ denotes $\sum_{n\in\mathcal{N}}$.

Individuals live for two periods: young and old. When young, they make choices about whether and where to attend college. There are E types of college education, and the set of education types is denoted by $\mathcal{E}$. Before turning old, the individual decides where to migrate for work. When old, they inelastically supply labor, pay for the college education ex-post, and raise their young child.

In each location, a representative firm produces consumption goods by aggregating labor with different education types. The consumption goods market is perfectly competitive.

In each region and for each education type, there is a semi-monopolistic college. Each college maximizes the weighted sum of the logarithm of its profits and the number of students enrolled. The latter reflects the non-profit nature of higher education. Both national (federal) and regional governments raise revenue through labor income taxes and utilize it to provide tuition discounts to individuals or to subsidize the production cost of college education.

2.2 Preferences

When young, individuals make choices about their education type, education location, and work location. When old, they inelastically supply labor, pay for the college education ex-post, and raise their young child.

We first lay out the old individual's static problem. An individual who was born in region $n^b \in \mathcal{N}$, chose education type $e \in \mathcal{E}$ and education location $n^e \in \mathcal{N}$, and is working in region $n^w \in \mathcal{N}$ solves the following.

$$U(n^b, e, n^e, n^w) = \max_{H,C} \; A_a(n^w) + \log\left[\left(\frac{H}{\alpha}\right)^{\alpha}\left(\frac{C}{1-\alpha}\right)^{1-\alpha}\right]$$
$$\text{s.t. } C + r(n^w)H \leq (1-\psi(e,n^w))\,w(e,n^w) - (1-d(n^b,e,n^e))\tau(e,n^e),$$

where $A_a(n^w)$ denotes the amenities provided by location n^w that makes the place more desirable to live and work; H and C represent the consumption of housing services and goods, respectively; $r(n^w)$ is the rental rate; $w(e,n^w)$ is the wages for individuals with education type e at location n^w; $\psi(e,n^w)$ is the proportional income tax rates; $\tau(n^b,e,n^e)$

denotes the gross tuition before discounts; and $d(n^b, e, n^e)$ is the tuition discounts provided by the governments. Note that the tuition discounts depend on the individual's birthplace. This feature allows the model to capture in-state tuition discounts.[8]

The solution to the problem is as below:

$$\begin{aligned} & U(n^b, e, n^e, n^w) \\ =& A_a(n^w) + \log\left[(1-\psi(e,n^w))w(e,n^w) - (1-d(n^b,e,n^e))\tau(e,n^e)\right] - \alpha\log(r(n^w)) \\ \equiv& A_a(n^w) + \log\left(DI(n^b,e,n^e,n^w)\right) - \alpha\log(r(n^w)) \end{aligned} \tag{1}$$

The utility of an individual is the sum of the amenities and the logarithm of real disposable income after taxes and net tuition, defined by $DI(\cdot)$, minus the rental rates multiplied by its income share. The old individual gives birth to one child, whose birthplace is the parent's workplace. We abstract from the incentive of choosing a work location based on children's utility to focus on our main mechanisms.

Now, let's turn to the forward-looking problem of young individuals. Let $V(e^p, n^b)$ denote the value of the individual who is born to a parent with education type e^p and in a location n^b. They make choices on (i) education type, e; (ii) education location, n^e; and (iii) work location, n^w, to maximize their utility. The value of the individual is given by

$$\begin{aligned} & V(e^p, n^b) = \\ & \mathbb{E}_{\epsilon_e}\Bigg[\max\Bigg\{\mathbb{E}_{\epsilon_{n^e}}\Big[\max_{n^e}\Big\{\mathbb{E}_{\epsilon_{n^w}}\big[\max_{n^w}\big\{U(n^b,e,n^e,n^w) - m_{n^w}(e,n^e,n^w) + \sigma_{n^w}\epsilon_{n^w}\big\}\big] \\ & - m_{n^e}(n^b,e,n^e) + \sigma_{n^e}\epsilon_{n^e}\Big\}\Big] \\ & - m_e(e^p,n^b,e) + \sigma_e\epsilon_e\Bigg\}\Bigg], \end{aligned}$$

where $m_{n^w}(e, n^e, n^w)$ represents the utility costs of moving to work location n^w given the individual's education type and location, (e, n^e). $m_{n^e}(n^b, e, n^e)$ denotes the utility costs of obtaining education at location n^e for individuals with type $(n^b,\ e)$. $m_e(e^p, n^b, e)$ represents the utility costs of choosing education type e for individuals born in n^b to a parent with education type e^p. Regarding the idiosyncratic preference shocks to work location, education location, and education type ($\sigma_{n^w}\epsilon_{n^w}$, $\sigma_{n^e}\epsilon_{n^e}$, and $\sigma_e\epsilon_e$), ϵ_{n^w}, ϵ_{n^e}, and ϵ_e are drawn i.i.d. from a Gumbel distribution with the location parameter of zero and the scale parameter one, that

[8] Alternatively, we could model tuition discounts to depend on parental education type (e^p) as well (i.e., $d(e^p, n^b, e, n^e)$). This alternative specification reflects the reality better in that tuition discounts can depend on parental income or parental education type—in the model, parental education level reflects parental income. However, we do not have data on tuition discounts based on parental education type, so we abstract from it.

is, from the distribution with the cumulative distribution function, $F(x) = \exp(-\exp(-x))$. Constants σ_{n^w}, σ_{n^e}, and σ_e govern the variance of the shocks.

The maximization problem can be divided based on each choice:

$$V(e^p, n^b) = \mathbb{E}_{\epsilon_e}\left[\max_e \left\{V_e(n^b, e) - m_e(e^p, n^b, e) + \sigma_e \epsilon_e\right\}\right]. \tag{2}$$

$$V_e(n^b, e) = \mathbb{E}_{\epsilon_{n^e}}\left[\max_{n^e} \left\{V_{n^e}(n^b, e, n^e) - m_{n^e}(n^b, e, n^e) + \sigma_{n^e} \epsilon_{n^e}\right\}\right].$$

$$V_{n^e}(n^b, e, n^e) = \mathbb{E}_{\epsilon_{n^w}}\left[\max_{n^w} \left\{U(n^b, e, n^e, n^w) - m_{n^w}(e, n^e, n^w) + \sigma_{n^w} \epsilon_{n^w}\right\}\right].$$

Due to the assumptions of extreme-value-distributed idiosyncratic preference shocks, the choice problem can be solved in closed forms (McFadden 1972). The choice probabilities and the corresponding values in the reverse order of choices are as below. Note that γ denotes the Euler-Mascheroni constant.

$$Pr[n^w = \tilde{n}^w | n^b, e, n^e] = \frac{\exp\left(\frac{U(n^b, e, n^e, \tilde{n}^w) - m_{n^w}(e, n^e, \tilde{n}^w)}{\sigma_{n^w}}\right)}{\sum_{\hat{n}^w} \exp\left(\frac{U(n^b, e, n^e, \hat{n}^w) - m_{n^w}(e, n^e, \hat{n}^w)}{\sigma_{n^w}}\right)} \tag{3}$$

$$V_{n^e}(n^b, e, n^e) = \sigma_{n^w} \log\left[\sum_{\hat{n}^w} \exp\left(\frac{U(n^b, e, n^e, \hat{n}^w) - m_{n^w}(e, n^e, \hat{n}^w)}{\sigma_{n^w}}\right)\right] + \sigma_{n^w}\gamma \tag{4}$$

$$Pr[n^e = \tilde{n}^e | n^b, e] = \frac{\exp\left(\frac{V_{n^e}(n^b, e, \tilde{n}^e) - m_{n^e}(n^b, e, \tilde{n}^e)}{\sigma_{n^e}}\right)}{\sum_{\hat{n}^e} \exp\left(\frac{V_{n^e}(n^b, e, \hat{n}^e) - m_{n^e}(n^b, e, \hat{n}^e)}{\sigma_{n^e}}\right)} \tag{5}$$

$$V_e(n^b, e) = \sigma_{n^e} \log\left[\sum_{\hat{n}^e} \exp\left(\frac{V_{n^e}(n^b, e, \hat{n}^e) - m_{n^e}(n^b, e, \hat{n}^e)}{\sigma_{n^e}}\right)\right] + \sigma_{n^e}\gamma \tag{6}$$

$$Pr[e = \tilde{e} | e^p, n^b] = \frac{\exp\left(\frac{V_e(n^b, \tilde{e}) - m_e(e^p, n^b, \tilde{e})}{\sigma_e}\right)}{\sum_{\hat{e}} \exp\left(\frac{V_e(n^b, \hat{e}) - m_e(e^p, n^b, \hat{e})}{\sigma_e}\right)} \tag{7}$$

$$V(e^p, n^b) = \sigma_e \log\left[\sum_{\hat{e}} \exp\left(\frac{V_e(n^b, \hat{e}) - m_e(e^p, n^b, \hat{e})}{\sigma_e}\right)\right] + \sigma_e\gamma \tag{8}$$

With the choice probabilities, we can derive the law of motion for the mass of individuals, $M(e^p, n^b, e, n^e, n^w)$. Since we are laying out the steady state, $M(\cdot)$ stays constant and is given

by

$$M(\tilde{e}^p, \tilde{n}^b, \tilde{e}, \tilde{n}^e, \tilde{n}^w)$$
$$= \left\{ \sum_{e^p, n^b, n^e} M(e^p, n^b, \tilde{e}^p, n^e, \tilde{n}^b) \right\} Pr[e' = \tilde{e} | \tilde{e}^p, \tilde{n}^b] Pr[n^{e'} = \tilde{n}^e | \tilde{n}^b, \tilde{e}] Pr[n^{w'} = \tilde{n}^w | \tilde{n}^b, \tilde{e}, \tilde{n}^e]. \tag{9}$$

Note that in the previous period, the education type and work location of individuals (e, n^w) equal the parental education type and birth location for the next generation $(\tilde{e}^p, \tilde{n}^b)$.

2.3 Technology: Firm

A representative firm in location n aggregates labor with different education types to produce consumption goods. The consumption goods market is perfectly competitive. The goods are traded without costs across regions. The production function of the firm in region n is given by

$$F_n(\{L(e,n)\}_e) = \left[\sum_e \{A_p(e,n) L(e,n)\}^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}}, \tag{10}$$

where $A_p(e,n)$ denotes a skill-augmenting productivity, and ρ represents the elasticity of substitution across different education types. The skill-augmenting productivity depends endogenously on the workforce composition and exogenously on productivity terms, $\bar{A}_p(e,n)$:

$$A_p(e,n) = \bar{A}_p(e,n) \left(\frac{\sum_{e \neq 0} L(e,n)}{L(0,n)} \right)^{\eta_r^e} \left(\sum_e L(e,n) \right)^{\eta_s^e}, \tag{11}$$

where η_r^e and η_s^e govern the elasticity of how the local skill-augmenting productivity $A_p(e,n)$ responds to the increase in the share of the education type relative to that of non-college-educated workers ($e = 0$) and that in the total population, respectively. These reflect the agglomeration effects pointed out by the existing literature (e.g., Diamond 2016 and Davis and Dingel 2020, among many others). As Giannone (2017) highlighted, the agglomeration effect can rationalize the observation that regions with increasing high-skilled share also exhibit faster growth in the high-skilled wages. The exogenous productivity term, $\bar{A}_p(e,n)$, represents the residual productivity that is not explained by the agglomeration. A differential increase in both the exogenous productivity and the agglomeration effect reflect skill-biased technical change (Katz and Murphy 1992).

Letting the total production of the firm be $Y(n)$, the firm's cost minimization problem

yields the following:

$$W(e,n)L(e,n) = \left(\frac{W(e,n)}{A_p(e,n)}\right)^{1-\rho} Y(n). \tag{12}$$

2.4 Technology: Colleges

There is no consensus in the literature on how to model the objective of colleges. For instance, Cai and Heathcote (2022) model it as profits, Fu (2014) as a function of quality (abilities) of enrolled students and profits, Epple, Romano, and Sieg (2006) and Gordon and Hedlund (2022) as a function of the quality of students, education expenditures per student, and total enrollment.[9]

We model a college as an entity that maximizes a function of profits and total enrollment each period by choosing the level of tuition.[10] Specifically, the objective of the college of education type e in location n is given by

$$\phi(e,n)\log\left(\Pi_{e,n}\left(\tau\left(e,n\right)\right)\right) + \left(1-\phi(e,n)\right)\log\left(E_{e,n}\left(\tau(e,n)\right)\right), \tag{13}$$

where $\phi(e,n) \in [0,1]$ represents the weight on the logarithm of profits, and $\Pi_{e,n}\left(\tau\left(e,n\right)\right)$ and $E_{e,n}\left(\tau(e,n)\right)$ denote the profits and enrollment given the college's tuition choice $\tau(e,n)$ and all the other equilibrium prices and allocations. Note that the total enrollment term reflects an altruistic motive of colleges.

The college produces a unit of education services, which turns a young individual into an old individual with education type e, by incurring the exogenous marginal costs of $MC(e,n)$. The costs are in terms of the consumption goods. Furthermore, the government subsidizes the production costs by a fraction of $s(e,n)$.

The college maximizes its objective subject to the constraint that it does not incur a loss: $\tau(e,n) \geq (1-s(e,n))MC(e,n)$. The college is owned by absentee landlords, so the profits do not accrue to the individuals.[11]

The choice of tuition affects the objective through intra-temporal and inter-temporal effects. First, the level of tuition affects the contemporaneous enrollment and markups. Second, it affects the future stock of workers with different education types—not just in its own region, but in all regions due to migration—which affects the current education choices.

[9] For more discussion on the objective of a college, see the discussion section of Fu (2014). When we expand our scope to other education levels, Singleton (2019) models an objective of a charter school as a function of total enrollment and profits.

[10] The four aforementioned papers assume heterogeneous student qualities; however, we abstract from this dimension and focus on enrollment and profits to parsimoniously study our research question.

[11] An alternative is to follow the approach by Desmet and Rossi-Hansberg (2014) assuming that a portfolio owns colleges of all types in all regions, and the profits are returned to individuals as dividends. To parsimoniously focus on our main mechanisms, we adopted the absentee-landlord assumption.

For tractability and realism of the model, we assume that a college is myopic and does not internalize the second inter-temporal effect.[12] Then, the optimality condition is as follows. For each $e \in \mathcal{E}$ and $n \in \mathcal{N}$,

$$\begin{aligned} \underbrace{\frac{\tau(e,n) - (1-s(e,n))MC(e,n)}{\tau(e,n)}}_{\text{Markup}} &= \underbrace{-\phi(e,n)\frac{\partial \log\left(\tau(e,n)\right)}{\partial \log\left(E_{e,n}(\tau(e,n))\right)}}_{\phi(e,n) \text{ times the inverse of price elasticity of demand}} \\ &\equiv -\phi(e,n)\frac{1}{\mathscr{E}_{e,n}(\tau(e,n))}. \end{aligned} \tag{14}$$

Note that if there is only profit motive for college operation, i.e., $\phi(e,n) = 1$, the above condition implies a standard solution for the monopolist's problem, where markup (Lerner index) equals negative of the inverse of the price elasticity of demand. When $\phi(e,n) = 0$, reflecting the case where college only has the output maximizing motive, then the pricing rule is the same with the perfect competition case.

The total enrollment is given by

$$\begin{aligned} E_{e,n}(\tau(e,n)) &= \sum_{\tilde{e}^p,\tilde{n}^b} \left[\left\{ \sum_{e^p,n^b,n^e} M(e^p,n^b,\tilde{e}^p,n^e,\tilde{n}^b) \right\} Pr[e'=e|\tilde{e}^p,\tilde{n}^b] Pr[n^{e'}=n|\tilde{n}^b,e] \right] \\ &\equiv \sum_{\tilde{e}^p,\tilde{n}^b} M_s(\tilde{e}^p,\tilde{n}^b,e,n), \end{aligned}$$

where $M_s(\tilde{e}^p,\tilde{n}^b,e,n)$ represents the measure of students enrolling in (e,n)-college who are born to parents with education type $\tilde{e}^p$ in location $\tilde{n}^b$. Note that the choice probabilities for e' and $n^{e'}$ depend on $\tau(e,n)$.

Defining the share of enrolled students of type $(\tilde{e}^p,\tilde{n}^b)$ at the current tuition level as $S_{e,n}(\tilde{e}^p,\tilde{n}^b) \equiv \frac{M_s(\tilde{e}^p,\tilde{n}^b,e,n)}{E_{e,n}(\tau(e,n))}$, the tuition elasticity of college enrollment is given by

$$\begin{aligned} \mathscr{E}_{e,n}(\tau(e,n)) \equiv& \frac{\partial \log\left(E_n(\tau(e,n))\right)}{\partial \log\left(\tau(e,n)\right)} \\ =& \sum_{\tilde{e}^p,\tilde{n}^b} S_{e,n}(\tilde{e}^p,\tilde{n}^b) \left\{ \frac{\partial \log\left(Pr[e'=e|\tilde{e}^p,\tilde{n}^b] Pr[n^{e'}=n|\tilde{n}^b,e]\right)}{\partial \log\left(\tau(e,n)\right)} \right\} \\ \equiv& \sum_{\tilde{e}^p,\tilde{n}^b} S_{e,n}(\tilde{e}^p,\tilde{n}^b) \mathscr{E}_{e,n}(\tau(e,n)|\tilde{e}^p,\tilde{n}^b). \end{aligned}$$

[12] We are assuming a representative college for a single education type in a region, which represents a collection of "infinitesimal" colleges in reality. It is unrealistic that each college in the real world makes its tuition choice considering the future stock of workers of different types.

The elasticities of the product of choice probabilities with respect to tuition are as follows.

$$\begin{aligned}
&\mathscr{E}(\tau(e,n)|\tilde{e}^p,\tilde{n}^b) \\
&= \left\{ \frac{1}{\sigma_e}(1 - Pr[e' = e|\tilde{e}^p,\tilde{n}^b])Pr[n^{e'} = n|\tilde{n}^b,e] + \frac{1}{\sigma_{n^e}}\left(1 - Pr[n^{e'} = n|\tilde{n}^b,e]\right) \right\} \\
&\times \sum_{\tilde{n}^w} \left\{ Pr[n^w = \tilde{n}^w|\tilde{n}^b,e,n] \frac{-(1 - d(\tilde{n}^b,e,n))\tau(e,n)}{(1 - \psi(e,\tilde{n}^w))w(e,\tilde{n}^w) - (1 - d(\tilde{n}^b,e,n))\tau(e,n)} \right\}
\end{aligned}$$

2.5 Government

There are two types of governments: federal and state, denoted by subscripts f and n, respectively. (A state is a location in our model.) Each government collects labor income taxes and utilizes them to provide tuition discounts to individuals and to subsidize college production.

The individual with education type e and work location n pays labor income taxes to both governments.

$$\psi(e,n) = \psi_f(e,n) + \psi_n(e,n).$$

The tuition discounts and production subsidies are funded by both types of governments. For a college type with (e,n),

$$d(n^b,e,n) = d_f(n^b,e,n) + d_n(n^b,e,n),\ \forall n^b \in \mathcal{N}.$$

$$s(e,n) = s_f(e,n) + s_n(e,n).$$

Each government runs a balanced budget each period. The federal government's budget constraint is

$$\begin{aligned}
&\sum_{e,n} \psi_f(e,n)w(e,n)M_r(e,n) \\
&= \sum_{e,n} \left[\sum_{\tilde{n}^b} \left\{ d_f(\tilde{n}^b,e,n)\tau(e,n) \sum_{\tilde{e}^p} M_s(\tilde{e}^p,\tilde{n}^b,e,n) \right\} + s_f(e,n)MC(e,n) \sum_{\tilde{e}^p,\tilde{n}^b} M_s(\tilde{e}^p,\tilde{n}^b,e,n) \right],
\end{aligned} \tag{15}$$

where the measure of residents with education type e in work location n is denoted by $M_r(e,n) \equiv \left(\sum_{e^p,n^b,n^e} M(e^p,n^b,e,n^e,n) \right)$.

The budget constraint for the state government in location n is given by

$$\begin{aligned}&\sum_{e}\psi_n(e,n)w(e,n)M_r(e,n)\\&=\sum_{e}\left[d_n(n,e,n)\tau(e,n)\sum_{\tilde{e}^p}M_s(\tilde{e}^p,n,e,n)+s_n(e,n)MC(e,n)\sum_{\tilde{e}^p,\tilde{n}^b}M_s(\tilde{e}^p,\tilde{n}^b,e,n)\right]\end{aligned}\tag{16}$$

Unlike the federal government, which raises taxes from all workers across all education types and work locations and disburses them in all regions, the state government provides tuition discounts and production subsidies only to its constituents (in-state students and colleges in the state).

2.6 Housing

Every period, region n has a fixed housing supply of $H(n)$, which is owned by the absentee landlord. The rental rates are determined by an optimality condition from the individual's utility maximization, which states that α share of all disposable incomes in a region equals the total spending on housing services:

$$r(n)H(n)=\alpha\left[\sum_{n^b,e,n^e}\left\{DI(n^b,e,n^e,n)\sum_{e^p}M(e^p,n^b,e,n^e,n)\right\}\right].\tag{17}$$

2.7 Equilibrium

The steady-state equilibrium is defined below. The model primitives are as follows.

- Amenities: $\{A_a(n)\}_n$
- Housing stock: $\{H(n)\}_n$
- Exogenous skill-augmenting productivities: $\left\{\bar{A}_p(e,n)\right\}_{e,n}$
- Tax rates on labor income: $\{\psi_f(e,n),\psi_n(e,n)\}_{e,n}$
- Tuition discounts: $\left\{d_f(n^b,e,n),d_n(n^b,e,n)\right\}_{n^b,e,n}$
- College production subsidies: $\{s_f(e,n),s_n(e,n)\}_{e,n}$
- College production costs: $\{MC(e,n)\}_{e,n}$
- College's market power: $\{\phi(e,n)\}_{e,n}$

- Utility costs of education: $\left\{m_e(e^p, n^b, e)\right\}_{e^p,n^b,e}$
- Utility costs of moving for education: $\left\{m_{n^e}(n^b, e, n^e)\right\}_{n^b,e,n^e}$
- Utility costs of moving for work: $\{m_{n^w}(e, n^e, n^w)\}_{e,n^e,n^w}$

The model parameters are the variance parameters for preference shocks, $\{\sigma_e, \sigma_{n^e}, \sigma_{n^w}\}$, and the production parameters governing the agglomeration effects, $\{\eta_r^e, \eta_s^e\}_e$.

The endogenous prices and allocations are as follows.

- Wages: $\{w(e, n)\}_{e,n}$
- (Gross) tuition rates: $\{\tau(e, n)\}_{e,n}$
- Rental rates: $\{r(n)\}_n$
- Value functions: $\left\{V(e^p, n^b), V_e(n^b, e), V_{n^e}(n^b, e, n^e)\right\}_{e^p,n^b,e,n^e}$
- Choice probabilities: $\left\{Pr[n^{w'} = n^w|n^b, e, n^e], Pr[n^{e'} = n^e|n^b, e], Pr[e' = e|e^p, n^b]\right\}_{e^p,n^b,e,n^e}$
- Firm's labor demand: $\{L(e, n)\}_{e,n}$
- Firm's output: $\{Y(n)\}_n$
- Steady-state distribution of individuals: $\left\{M(e^p, n^b, e, n^e, n^w)\right\}_{e^p,n^b,e,n^e,n^w}$

Given the model primitives and parameters, the equilibrium is prices and allocations that satisfy the following conditions. The optimality conditions, the government budget constraints, and the law of motion are as follows.

- Wages, tuition rates, rental rates, value functions, and choice probabilities satisfy individuals' optimality conditions and the law of motion of population, provided in Equations (1) and (3) to (9).
- Wages, firms' labor demand, and output of a location satisfy firm's optimality conditions in Equations (11) and (12).
- The distribution of individuals, choice probabilities of individuals, and tuition rates satisfy Equation (14).
- Wages, distribution of individuals, and tuition levels satisfy the federal and state governments' budget constraints in Equations (15) and (16).

Market clearing conditions are below.

- Consumption goods market clears: for all $n \in \mathcal{N}$,

$$
\begin{aligned}
&Y(n) \\
&= \underbrace{(1-\alpha)\left[\sum_{e^p,n^b,e,n^e}\left\{DI(n^b,e,n^e,e)M(e^p,n^b,e,n^e,n)\right\}\right]}_{\text{(Consumption)}} + \underbrace{\sum_{e^p,n^b,e}\left\{MC(e,n)M_s(e^p,n^b,e,n)\right\}}_{\text{(Costs of college production)}} \\
&+ \underbrace{r(n)H(n) + \sum_e \Pi_{e,n}(\tau(e,n))}_{\text{(Rent and college profits)}} + \underbrace{\sum_e \{\psi(e,n)w(e,n)M_r(e,n)\}}_{\text{(Tax payments)}} \\
&- \underbrace{\sum_{e^p,n^b,e}\left\{d(n^b,e,n)\tau(e,n)M_s(e^p,n^b,e,n) + s(e,n)MC(e,n)M_s(e^p,n^b,e,n)\right\}}_{\text{(Tax benefits)}} \\
&+ \underbrace{\sum_{e^p,n^b,e,n^e}\left\{\left(1-d(n^b,e,n^e)\right)\tau(e,n^e)M(e^p,n^b,e,n^e,n)\right\}}_{\text{(Tuition paid by residents)}} \\
&- \underbrace{\sum_{e^p,n^b,e}\left\{\left(1-d(n^b,e,n)\right)\tau(e,n)M_s(e^p,n^b,e,n)\right\}}_{\text{(Tuition received by colleges)}}
\end{aligned}
\tag{18}
$$

- College market clears: for all $n \in \mathcal{N}$ and $e \in \mathcal{E}$, $E_{e,n}(\tau(e,n)) = \sum_{\tilde{e}^p,\tilde{n}^b} M_s(\tilde{e}^p,\tilde{n}^b,e,n)$.
- Labor market clears: for all $n \in \mathcal{N}$, $\sum_{e^p,n^b,n^e} M(e^p,n^b,e,n^e,n) = L(e,n)$.
- Housing market clears: for all $n \in \mathcal{N}$, Equation (17) holds.

2.8 Supply and Demand of College Education

Since our main theoretical contribution is incorporating endogenous college education into a quantitative spatial model, this section discusses how our model incorporates the key supply and demand factors highlighted by the existing literature. For this purpose, we link our model ingredients to those of Gordon and Hedlund (2022). They break down the rise in college tuition—which is the equilibrium price of the college market—from 1987 to 2010 into supply-side and demand-side factors.

Regarding the supply-side factors, the production costs of colleges have increased due to Baumol's (1967) cost disease. Baumol points out that general productivity growth increases wages, but since productivity growth happens at a different pace across sectors, the production costs of stagnant sectors increase. A string quartet is a prime example of the stagnant sector with increasing production costs (Baumol, Bowen, et al. 1993)—it always requires

four musicians to perform. Higher education also belongs to the category of the stagnant sector—one instructor is needed per class. Increasing $MC(e, n)$ over time in our model can incorporate the cost disease.

Another supply factor is declining state and federal funding, especially for public institutions. This channel can be studied in our model by assessing the impact of changes in the subsidy rates for the production of colleges ($s(e, n)$).

Regarding the demand-side factors, the first is the Bennett (1987) Hypothesis, in which "increases in financial aid ... have enabled colleges and universities blithely to raise their tuitions." Our model can capture the hypothesis through increasing tuition discount rates ($d(n^b, e, n)$) along with the college's market power, which is governed by $\phi(e, n)$. In the model, an increase in demand through the rise in subsidy incurs the tuition increase. Furthermore, if we let $\phi(e, n)$ vary over time, the model can incorporate the degree of colleges' market power that could change over time. Despite these capabilities, the model cannot capture some aspects of demand-side public higher education policies, such as their impact on defaults on student loans or on relaxing the households' financial constraints.

The second factor related to demand is the rise of parental income. A child's education is a normal good, where the demand increases with the rise in parental income. Our model can indirectly capture this channel through the disutility of education ($m_e(e^p, n^b, e)$), i.e., when $m_e(\cdot)$ is lower for the type e^p with higher incomes. However, this formulation is a reduced form in nature. Our model cannot fully reflect all the channels through which parental income affects children's higher education choices, such as inter-vivo transfer for tuition. This abstraction is for the tractability of the model. If parental wealth or income were to directly enter into the child's (young individual's) optimization problem, we would need to track the distribution of parents' wealth or income over time. Our approach is in line with other spatial models (e.g., those surveyed by Redding and Rossi-Hansberg (2017)) whose tractability arises from the model features that render tracking the distribution of state variables unnecessary.

The third demand-side factor, skill premium, or equivalently, ex-ante returns on college enrollment, is reflected by the fact that a young individual takes into account wages for their choices. Furthermore, the changes in the skill premium are affected by the skill-biased technical change in our model. The growth rates of skill-augmenting productivity, $A_p(e, n)$, can differ across education types.

2.9 Full Model and Transition Paths

Until now, we have outlined a simplified version of the model in its steady state for exposition. In this section, we outline the full model and discuss how we can map the transition path to the data. The rise in spatial inequality and tuition costs after 1980 is direct evidence that the economy was not in a steady state. To fit our model to the transition path while maintaining tractability, we make the following assumptions for the full model.

Each time period is 10 years—we are parameterizing the model with the decennial data from 1980, 1990, 2000, 2010, and 2019. Each individual lives five periods: young (ages 13–22), old_1 (ages 23–32), old_2 (ages 33–42), old_3 (ages 43–52), and old_4 (ages 53–62), denoted by y, o_1, o_2, o_3, and o_4.[13] Individuals make choices on their education type, education location, and work location when they are young. We assume that individuals are myopic. They think that the prices they observe when they are young are the prices that they will face when in $\{o_i\}_{i=1,2,3,4}$.[14] Individuals attend college when they are young if they choose to do so. They pay their college tuition ex-post when they are old. We assume that individuals borrow from the outside world at a risk-free rate i to pay for their tuition when they are young and pay down their loan with four equal payments when they are in $\{o_i\}_{i=1,2,3,4}$. Regardless of education type, the individual enters the labor force at age o_1, and they earn $w_t(e, n^w)$ each time period. In this formulation, for example, a young individual makes their choices based on the prices in 1970 and enters the labor force in 1980. To account for the loss of working years due to college education, we assume that the unit of labor supplied by a college-educated worker is nine-tenths of that of a non-college-educated worker.[15] Lastly, we assume that there is exogenous population growth. An exogenous measure of children are born to a parent of each type, the process which we will describe in more detail later in this subsection.

To describe how the above timeline and assumptions are consistent with the model exposition based on the steady-state of the simplified version, let $U_t(n^b, e, n^e, n^w)$ denote the per-period utility of an individual in $\{\text{old}_i\}_{i=1,2,3,4}$ with a type (n^b, e, n^e, n^w). Let β denote the 10-year discount rates. Suppose that a young individual in time t is making choices related to education and work. Since the individual is myopic, they expect that the per-

[13] The individuals finish college education at age 22 and retire at age 62, which is the earliest age to claim retirement benefits.

[14] Without this simplifying assumption, we have to consider individuals' expectations about future prices. Even without the myopia assumption, we need to make assumptions regarding expectations, such as whether the individual has perfect foresight.

[15] This is to reflect that college education takes four years. Since we are utilizing decennial data, it's not possible to model the loss of working years as individuals not working for certain periods. Therefore, we alternatively assume that the effective labor of college-educated individual is nine-tenths of their non-college-educated counterpart to reflect that four years out of 40 working years are lost.

period utility of when they are $\{\text{old}_i\}_{i=1,2,3,4}$ in $\{t+1,...,t+4\}$ to be constant at $U_t(\cdot)$. Therefore, they evaluate their future lifetime utility—after incurring the utility costs of $m_e(\cdot)$, $m_{n^e}(\cdot)$, $m_{n^w}(\cdot)$—for the choice of (e, n^e, n^w) as $(1+\beta+\beta^2+\beta^3)U_t(\cdot)$. Therefore, the lifetime utility function in Equation (1) can be regarded as the scaled version of the "perceived" lifetime utility. The actual lifetime utility that an individual receives will be $U_{t+1}(\cdot)+\beta U_{t+2}(\cdot)+\beta^2 U_{t+3}(\cdot)+\beta^3 U_{t+4}(\cdot)$.

Two additional assumptions are as follows. First, the tuition paid in each period of $\{o_i\}_{i=1,2,3,4}$ adds up to tuition charged by a college. Let $\tau_s(e,n,t)$ denote the split payment of tuition at $t+1,...,t+4$ for college attendance in t. Then,

$$\left(\frac{1}{1+i}+\frac{1}{(1+i)^2}+\frac{1}{(1+i)^3}+\frac{1}{(1+i)^4}\right)\tau_s(e,n,t)=\tau_t(e,n),$$

where $\tau_t(e,n)$ is the tuition charged by a college for its service in t. In this formulation, the tuition term in $U_t(\cdot)$ in the steady-state exposition in the model should be interpreted as $\tau_s(e,n,t)$.

Second, the governments' budget constraints are period-by-period. In each period, the governments tax old generations to provide tuition discounts and subsidies to colleges for the young.[16]

Given the assumptions, the law of motion and the labor market clearing condition are as follows. Let $\overline{M}_{c,t}(e^p,n^b)$ denote an exogenous measure of children born at time t to a parent of education type $e=e^p$ and work location $n^w=n^b$—the parent's education type and work location are the child's parental education type and birth location. Since a young individual is born when an old individual turns from o_2 to o_3, the measure of young individuals at time $t+1$ of type $(\tilde{e}^p,\tilde{n}^b)$ is given by: $M_{y,t+1}(\tilde{e}^p,\tilde{n}^b)=\overline{M}_{c,t+1}(\tilde{e}^p,\tilde{n}^b)\sum_{e^p,n^b,n^e}M_{o2,t}(e^p,n^b,\tilde{e}^p,n^e,\tilde{n}^b)$. The law of motion for the measure of individuals who are $\{o_i\}_i=1,2,3,4$ is given below.

$$\begin{aligned}&M_{o1,t+1}(e^p,n^b,n,n^e,n^w)\\ =&M_{y,t}(e^p,n^b)Pr_t[e'=e|e^p,n^b]Pr_t[n^{e'}=n^e|e^p,n^b,e]Pr_t[n^{w'}=n^w|e^p,n^b,e,n^e],\text{ and}\end{aligned}$$

$M_{o2,t+1}(\cdot)=M_{o1,t}(\cdot)$, $M_{o3,t+1}(\cdot)=M_{o2,t}(\cdot)$, and $M_{o4,t+1}(\cdot)=M_{o3,t}(\cdot)$

Lastly, due to college-educated workers' labor supply being nine-tenths of those of non-college-educated workers, labor market clearing condition is provided by the following. For

[16] We assume that the discounts are equally applied to the split repayments for the student loan. The old generation with college education repays $(1-d(n^b,e,n))\tau_s(e,n,t)$ each period.

all $n \in \mathcal{N}$,

$$L(e,n) = \begin{cases} \sum_{e^p,n^b,n^e} M(e^p,n^b,e,n^e,n) & \text{, if } e=0, \text{ and} \\ 0.9 \sum_{e^p,n^b,n^e} M(e^p,n^b,e,n^e,n) & \text{, if } e \neq 0, \end{cases}$$

where $e = 0$ denotes the option of not going to any colleges.

The dynamics of equilibrium allocations and prices are driven by the changes in the model primitives. All the primitives outlined in the definition of the equilibrium (Section 2.7), excluding the housing stock, are time-varying. For the rest of the paper, we introduce time subscripts for the time-varying model objects.

3 Parameterization

We estimate our model using the data for 50 US states and District of Columbia and for five decennial years (1980, 1990, 2000, 2010, 2019).[17] There are two education types, $e = 1$ representing individuals who obtained a four-year college degree or higher, and $e = 0$ denoting the other individuals.

3.1 Data

To estimate the model, we compile data from multiple sources, including those from the IPUMS-USA (Ruggles et al. 2021)—specifically, the Decennial Census (1980, 1990, and 2000) and the five-year American Community Survey (ACS, 2008-2012 for 2010 and 2017-2021 for 2019)—as well as data from the National Center for Education Statistics (NCES), such as the Integrated Postsecondary Education Data System (IPEDS) and the Baccalaureate and Beyond Longitudinal Study (B&B). Table 1 lists the model objects that we obtain from the data.

Regarding tuition, we adopt a broad notion of tuition incorporating both tuition (tuition fees paid by students along with the scholarships and grants they received) and the sales of services (e.g., room and board).[18] As for the tuition discount rates, since we cannot observe how the discounts are allocated across students of different types in IPEDS, we make the following measurement assumption. Tuition discounts provided by the state government are given only to students born in the state and who go to college in the same state. That is, we assume that $d_{n,t}(n^b, e = 1, n) = 0$, if $n^b \neq n$. For the federal government's tuition dis-

[17] We utilized the data from 2019 instead of 2020 due to the Covid-19 pandemic.

[18] The reason for this choice is as follows. The sales of services account for a significant fraction of colleges' revenues. If we adopt the narrow notion of tuition, then given the total production costs, the colleges are incurring losses.

Table 1: Data Observables

Variable	Data	Source
$w_t(e,n)$	Mean yearly wages ($\times 10$) for each education type and each state	Census/ACS
$\alpha = 0.325$	Housing expenditure as a share of annual expenditures (average across 1990, 2000, 2010, 2019)	BLS CEX
$r_t(n)$	State fixed effects from the hedonic price regression for rents following Eeckhout, Pinheiro, and Schmidheiny (2014)	IPUMS-NHGIS
$\psi_{f,t}(e,n)/\psi_{f,t}(0,1)$	Federal labor income tax rates relative to type $(e,n) = (0,1)$	NBER-Taxsim
$\psi_{n,t}(e,n)/\psi_{n,t}(0,n)$	State labor income tax rates relative to type $(e,n) = (0,n)$	NBER-Taxsim
$Pr_t[n^w \vert n^b, e=0, n^e]$	CP (choice probabilities) for work-related migration of non-college-educated workers	Census/ACS
$Pr_t[n^w \vert n^b, e=1, n^e]$	CP for work-related migration of college-educated workers	B&B
$Pr_t[n^e \vert n^b, e]$	CP for education-related migration	Census/ACS
$Pr_t[e \vert e^p, n^b]$	CP for education type	B&B
$M_{y,t}(e^p, n^b)$	Population of young individuals	Census/ACS
$\tau_t(e,n)$	Gross tuition and fees per enrolled student	IPEDS
$MC_t(e,n)$	Expenditure per enrolled student	IPEDS
$d_{n,t}(n^b,e,n)$	Discounts on tuition and fees given by the state and local governments divided by tuition	IPEDS
$d_{f,t}(n^b,e,n)$	Discounts on tuition and fees given by the federal government divided by tuition	IPEDS
$s_{n,t}(e,n)$	Appropriation given to colleges from the state government divided by the total expenditure	IPEDS
$s_{f,t}(e,n)$	Appropriation given to colleges from the federal government divided by the total expenditure	IPEDS

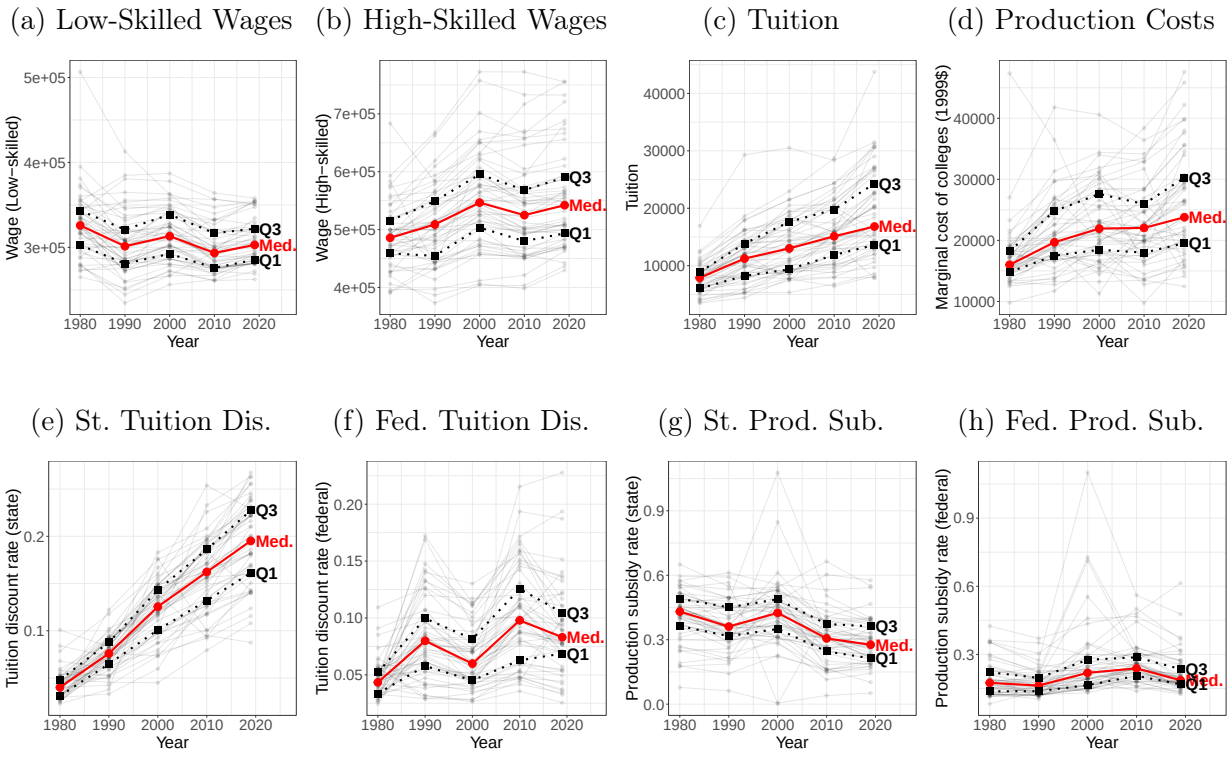
counts, we assume that students receive the same discount regardless of their birth location: $d_{f,t}(n^b, e=1, n) = \bar{d}_{f,t}(e=1,n)\ \forall n^b \in \mathcal{N}$.

Two additional remarks are in order. First, notably, primitives related to college education, such as the discount and subsidy rates, are directly obtained from the balance sheets of colleges in IPEDS. This enables the model to directly assess the effects of higher education policies on rising inequality. Second, in our model, taxes are only used for higher education policies, and governments run balanced budgets. Hence, we need to scale down the

tax rates in the data. This is the reason why we utilize data on the relative tax incidence across types, not the levels. Assuming that the proportion of tax revenue used for higher education is independent of the population from whom the taxes are raised, the relative tax incidence, $\left\{\frac{\psi_{f,t}(e,n)}{\psi_{f,t}(0,1)}\right\}_{e,n,t}$ and $\left\{\frac{\psi_{n,t}(e,n)}{\psi_{n}(0,n)}\right\}_{e,n,t}$, are obtained from labor income tax rates from NBER-Taxsim.[19] We construct the tax rates in the following way. The expenditures are the right-hand sides of Equations (15) and (16) are obtained from data. Then, we choose $\{\psi_{f,t}(0,1)\}_t$ and $\{\psi_{n,t}(0,n)\}_{n,t}$ so that the government's budget constraints are satisfied. Other details on data construction are provided in Appendix C.

Figure 3 describes the first quartile, median, and third quartiles across states over time for the prices (wages and tuition) and the model primitives related to higher education

Figure 3: Wages, Tuition, and College-Related Primitives Across States and Years



* Note: A grayed-out line represents observations for a state. Production costs represent $\{MC_t(e=1,n)\}_{n,t}$. St. and Fed. Tuition Dis. denote state and federal tuition discount rates, $\{d_{n,t}(n,e=1,n)\}_{n,t}$ and $\{d_{f,t}(n^b,e=1,n)\}_{n^b,n,t}$, respectively. For the calculation of the tuition discounts from data, we assumed that the state tuition discounts are only provided to students born in the state and that federal discounts for a college in a state are equally distributed across students regardless of their birth location. St. and Fed. Prod. Sub. denote state and federal production subsidy rates, $\{s_{n,t}(e=1,n)\}_{n,t}$ and $\{s_{f,t}(e=1,n)\}_{n,t}$, respectively.

[19]For states without state income taxes, we assume that the relative tax incidence between college-educated and non-college-educated workers is equal to that in the federal income taxes.

(marginal costs, discounts, and production subsidies) directly observable from data. From the figure, we see the following general patterns, which are in line with the existing literature: skill premiums increased (Figures 3a and 3b); tuition fees significantly increased (Figure 3c); production costs for colleges increased (Baumol's cost disease, Figure 3d); tuition discounts provided by the state and federal governments increased (related to the Bennett Hypothesis, Figures 3e and 3f); and state and federal funding for colleges decreased (Figure 3g) and stagnated (Figure 3h), respectively.

3.2 Preference Parameters

We estimate our model in two steps. In this subsection, we estimate preference parameters. Then, in the next subsection, we obtain parameters and primitives related to firms and colleges using the structural equations from the model. Before the estimation of the preference parameters, we need to parameterize the utility costs as a function of data observables. The utility costs of work-related migration are given below:

$$m_{n^w}(e, n^e, n^w) = \delta_{n^w,1} dist(n^e, n^w) + \delta_{n^w,2} \mathbb{1}(n^e \neq n^w) + \delta_{n^w,3} \mathbb{1}(e \neq 0) \mathbb{1}(n^e \neq n^w),$$

where $\delta_{n^w,1}$ captures the moving costs that depend on the distance between the education location and work location; $\delta_{n^w,2}$ represents the degree of home bias (the extra utility derived from working in the state that an individual is educated in); $\delta_{n^w,3}$ captures the unobservable factors affecting the relative mobility of high-skilled labor—it is a well-known fact that high-skilled labor is more mobile; and $\mathbb{1}(\cdot)$ denotes an indicator function.

The utility costs of moving for education are given by

$$m_{n^e}(n^b, e, n^e) = \begin{cases} 0, & \text{if } e = 0 \ \& \ n^b = n^e, \\ \infty, & \text{if } e = 0 \ \& \ n^b \neq n^e, \text{ and} \\ \delta_{n^e,1} dist(n^b, n^e) + \delta_{n^e,2} \mathbb{1}(n^b \neq n^e) & \text{if } e \neq 0. \end{cases}$$

Given this formulation, if an individual does not obtain a college education (i.e., $e = 0$), they do not pay the utility costs, and the individual's education location equals their birth location due to the infinite cost of moving to other states for college education. For individuals who choose to go to college (i.e., $e = 1$), their utility costs of moving are zero if they choose an in-state college. Otherwise, they pay the utility costs based on the distance ($\delta_{n^e,1}$) plus the home-bias term ($\delta_{n^e,2}$).

The utility costs of education are below.

$$m_e(e^p, n^b, e) = \mathbb{1}(e \neq 0)\left\{\delta_{e,1} + \delta_{e,2}\mathbb{1}(e^p = 1)\right\}.$$

The disutility cost of education is zero if an individual chooses not to go to college, $\delta_{e,1}$ if they go to college and their parent is non-college-educated, and $\delta_{e,1} + \delta_{e,2}$ if they attend college and their parent has a college degree. $\delta_{e,2}$ can reflect that the children of college-educated parents—who also earn higher income—are more likely to attend college.

With the parameterized utility costs and other data observables, we estimate the model so that the choice probabilities of the individuals (Equations (3), (5) and (7)) in the model are closest to those in the data. Since the sum of probabilities equals one, we difference-out the choice probabilities and use the log odds ratio for our estimation. Specifically, from Equations (3), (5) and (7), we can derive the following estimating equation with added error terms.

$$\begin{aligned}&\log\left(\frac{Pr_t[n^w = \tilde{n}^w|n^b, e, n^e]}{Pr_t[n^w = 1|n^b, e, n^e]}\right)\\ =&\frac{1}{\sigma_{n^w}}\left\{U_t(n^b, e, n^e, \tilde{n}^w) - U_t(n^b, e, n^e, 1) - m_{n^w}(e, n^e, \tilde{n}^w) + m_{n^w}(e, n^e, 1)\right\} + u_{n^b,e,n^e,\tilde{n}^w,t}\end{aligned} \tag{19}$$

$$\begin{aligned}&\log\left(\frac{Pr_t[n^e = \tilde{n}^e|n^b, e]}{Pr_t[n^e = 1|n^b, e]}\right)\\ =&\frac{1}{\sigma_{n^e}}\left\{V_{n^e,t}(n^b, e, \tilde{n}^e) - V_{n^e,t}(n^b, e, 1) - m_{n^e}(n^b, e, \tilde{n}^e) + m_{n^e}(n^b, e, 1)\right\} + u_{n^b,e,\tilde{n}^e,t}\end{aligned} \tag{20}$$

$$\begin{aligned}&\log\left(\frac{Pr_t[e = 1|e^p, n^b]}{Pr_t[e = 0|e^p, n^b]}\right)\\ =&\frac{1}{\sigma_e}\left\{V_{e,t}(n^b, 1) - V_{e,t}(n^b, 0) - m_e(e^p, n^b, 1) + m_e(e^p, n^b, 0)\right\} + u_{e^p,n^b,1,t}\end{aligned} \tag{21}$$

By minimizing the sum of the squared errors in the Equations (19) to (21), we obtain the variance terms for idiosyncratic shocks of the extreme-valued distribution (σ_{n^w}, σ_{n^e}, and σ_e) along with the parameters for the utility costs ($\delta_{n^w,1}$, $\delta_{n^w,2}$, $\delta_{n^w,3}$, $\delta_{n^e,1}$, $\delta_{n^e,2}$, $\delta_{e,1}$, and $\delta_{e,2}$). The estimation results are in Table 2.

Two additional remarks are in order for the estimation. First, we estimate the Equation (19) with a time-location fixed effect term for $\tilde{n}^w$ and obtain the location-specific amenity term $A_{a,t}(\tilde{n}^w)$ from this fixed effect term. Second, the estimating equations are interrelated. For example, estimates from Equation (19) are used to derive $V_{n^e,t}(\cdot)$ through Equation (4).

Table 2: Estimation Results for Preference Parameters

Parameter	Description	Estimate (Std. Error)
σ_{n^w}	Dispersion parameter: work location	0.176 (0.002)
σ_{n^e}	Dispersion parameter: education location	0.212 (0.002)
σ_e	Dispersion parameter: college attendance	0.590 (0.085)
$\delta_{n^w,1}$	Effect of bilateral distance on $m_{n^w}(\cdot)$	1.09×10^{-4} (1.01×10^{-6})
$\delta_{n^w,2}$	Effect of home bias ($\mathbb{1}(n^e \neq n^w)$) on $m_{n^w}(\cdot)$	1.259 (0.012)
$\delta_{n^w,3}$	Effect of college education on $m_{n^w}(\cdot)$	-0.274 (0.007)
$\delta_{n^e,1}$	Effect of bilateral distance on $m_{n^e}(\cdot)$	1.37×10^{-4} (1.53×10^{-6})
$\delta_{n^e,2}$	Effect of home bias ($\mathbb{1}(n^b \neq n^e)$) on $m_{n^e}(\cdot)$	0.920 (0.018)
$\delta_{e,1}$	Effect of education disutility on $m_e(\cdot)$	1.280 (0.105)
$\delta_{e,2}$	Effect of parental education ($\mathbb{1}(e^p = 1)$) on $m_e(\cdot)$	-0.975 (0.143)

3.3 Production Parameters: Colleges

Regarding the colleges' market power terms ($\phi_t(e = 1, n)$), we recover them from Equation (14). The markup is directly observable in the data, and the tuition elasticity of college demand is calculated with the estimated preference parameters, $\hat{\sigma}_e$ and $\hat{\sigma}_{n^e}$, and data moments. We set the market power term time-varying since we observe a significant increase in the gap between the markup and the inverse of the elasticity over time. The summary statistics for the obtained college's market power parameters are in Table 3, which demonstrate that, in general, colleges' market power has been increasing over time. This suggests that the Bennett Hypothesis could have played an important role in the tuition hike.

Table 3: Summary Statistics for 51 (50 States + D.C.) Colleges' Market Power Terms Times Hundred ($100 \times \phi_t(e = 1, n)$) across Years

Year	Mean	Q1	Q2	Q3
1980	0.20	0.13	0.18	0.25
1990	0.25	0.17	0.23	0.31
2000	0.84	0.35	0.51	0.89
2010	0.62	0.38	0.59	0.70
2019	0.69	0.42	0.57	0.82

3.4 Production Parameters: Firms

In the firm's production function, there are two sets of parameters: ρ, which governs the substitutability of high- and low-skilled labor, and $\{\eta_r^0, \eta_r^1, \eta_s^0, \eta_s^1\}$, which control the strength of the agglomeration effects. To estimate the parameters, we difference out Equation (12) and plug in Equation (11).

$$
\begin{aligned}
&\log\left(\frac{W_t(1,n)}{W_t(0,n)}\right) \\
&= -\frac{1}{\rho}\log\left(\frac{L_t(1,n)}{L_t(0,n)}\right) + \frac{\rho-1}{\rho}\log\left(\frac{A_{p,t}(1,n)}{A_{p,t}(0,n)}\right) \\
&= \left\{-\frac{1}{\rho} + \frac{\rho-1}{\rho}(\eta_r^1 - \eta_r^0)\right\}\log\left(\frac{L_t(1,n)}{L_t(0,n)}\right) + \frac{\rho-1}{\rho}(\eta_s^1 - \eta_s^0)\log\left(L_t(0,n) + L_t(1,n)\right) \\
&+ \frac{\rho-1}{\rho}\log\left(\frac{\bar{A}_{p,t}(1,n)}{\bar{A}_{p,t}(0,n)}\right)
\end{aligned}
\tag{22}
$$

In the above equation, we can see how the model can rationalize the data observation that there is a positive correlation between the skill intensity, $\left(\frac{L_t(1,n)}{L_t(0,n)}\right)$, and the skill premium, $\left(\frac{W_t(1,n)}{W_t(0,n)}\right)$, in the data. If there are no agglomeration effects, i.e., $\eta_r^0 = \eta_r^1 = \eta_s^0 = \eta_s^1 = 0$, the data pattern can be rationalized through the regional differences in the underlying productivities, $\left(\frac{\bar{A}_{p,t}(1,n)}{\bar{A}_{p,t}(0,n)}\right)$. With the agglomeration effects, the data observation can be supported with the differences in the skill composition, $\left(\frac{L_t(1,n)}{L_t(0,n)}\right)$, or total population, $(L_t(o,n) + L_t(1,n))$, even without any significant differences in the underlying productivities. We proceed with the estimation of the agglomeration effect; however, we conduct the counterfactual exercises for both types of models, with and without the agglomeration effects.

To estimate Equation (22), we assume that the regional underlying productivity consists of a nationwide term and an idiosyncratic shock: $\bar{A}_{p,t}(e,n) = \tilde{A}_{p,t}(e) + \epsilon_{p,t}(e,n)$. Then, Equation (22) can be re-expressed as below:

$$
\begin{aligned}
&\log\left(\frac{W_t(1,n)}{W_t(0,n)}\right) \\
&= \left\{-\frac{1}{\rho} + \frac{\rho-1}{\rho}(\eta_r^1 - \eta_r^0)\right\}\log\left(\frac{L_t(1,n)}{L_t(0,n)}\right) + \frac{\rho-1}{\rho}(\eta_s^1 - \eta_s^0)\log\left(L_t(0,n) + L_t(1,n)\right) \\
&+ \frac{\rho-1}{\rho}\log\left(\frac{\tilde{A}_{p,t}(1)}{\tilde{A}_{p,t}(0)}\right) + \frac{\rho-1}{\rho}\log\left(\frac{\epsilon_{p,t}(1,n)}{\epsilon_{p,t}(0,n)}\right)
\end{aligned}
\tag{23}
$$

Accounting for $\frac{\rho-1}{\rho}\log\left(\frac{\tilde{A}_{p,t}(1)}{\tilde{A}_{p,t}(0)}\right)$ with the time fixed effects and assuming that $\frac{\rho-1}{\rho}\log\left(\frac{\epsilon_{p,t}(1,n)}{\epsilon_{p,t}(0,n)}\right)$ follows normal distribution with mean zero, we run the OLS estimation for the above. Since ρ, η_r^0, and η_r^1, and ρ are not separately identified, we externally calibrate $\rho = 1.41$ following Katz and Murphy (1992) and assume $\eta_r^0 = 0$. The same applies to ρ, η_s^0, and η_s^1, and we set $\eta_s^0 = 0$. The estimation results are in Table 4.

The estimation is consistent with the model. Young individuals are myopic, which implies that the wages of the current period did not affect the education and work choices of the

Table 4: Estimation Results for Production Parameters

Parameter	Description	Estimate (Std. Error)
η_r^1	Agglomeration: skill ratio	2.67 (0.05)
η_s^1	Agglomeration: total population	0.25 (0.02)

previous period's young individuals, and the labor supply of old individuals is inelastic. Therefore, there is no simultaneity bias for Equation (23). However, in reality, the estimates for the agglomeration terms are likely to be upward-biased if there are supply shocks and the supply is sufficiently inelastic.[20] Therefore, for quantitative analyses, we explore three cases: no ($\eta_r^1 = \eta_s^1 = 0$), medium ($\eta_r^1 = 1.34$, and $\eta_s^1 = 0.13$), and high ($\eta_r^1 = 2.67$, and $\eta_s^1 = 0.25$) agglomeration effects.

With the values of ρ, η_r^1, and η_s^1, we obtain the total and underlying skill-augmenting productivities, $A_{p,t}(e,n)$ and $\bar{A}_{p,t}(e,n)$, respectively, using Equation (12) and Equation (11). Equation (12) can be re-expressed as

$$\begin{aligned}\log\left(A_{p,t}(e,n)\right) =& \frac{\rho}{\rho-1}\log(w_t(e,n)) + \frac{1}{\rho-1}L_t(e,n) \\ &- \frac{1}{\rho-1}\log\left(w_t(0,n)L_t(0,n) + w_t(1,n)L_t(1,n)\right).\end{aligned} \tag{24}$$

Using the above equation, the total productivity, $A_{p,t}(e,n)$, is derived since all the other prices and allocations are directly observable from the data. Utilizing Equation (11), we can obtain the underlying productivity, $\bar{A}_{p,t}(e,n)$.

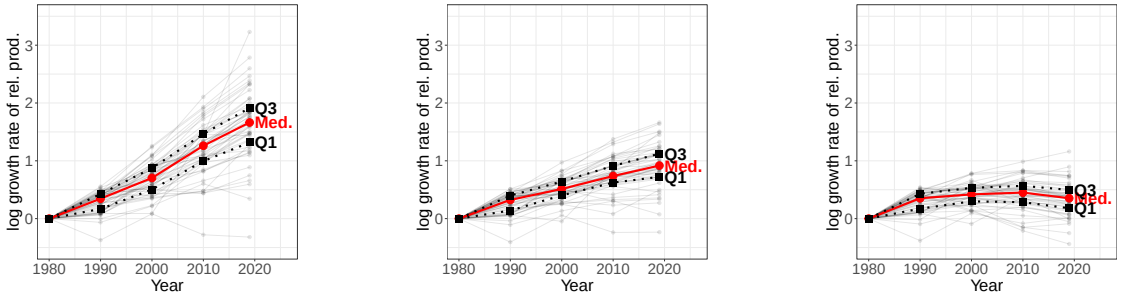
We derive the productivities for all three cases: no, medium, and high agglomeration. Figure 4 illustrates the log growth rates of underlying productivities of the high-skilled relative to the low-skilled, $\log\left(\frac{\bar{A}_{p,t}(e=1,n)}{\bar{A}_{p,t}(e=0,n)}\right) - \log\left(\frac{\bar{A}_{p,1980}(e=1,n)}{\bar{A}_{p,1980}(e=0,n)}\right)$, for the three cases. Regarding the no agglomeration case, the total productivities are equal to the underlying productivities, i.e., $A_{p,t}(e,n) = \bar{A}_{p,t}(e,n)$. The total productivities in the medium and high agglomeration cases are equal to those in the no agglomeration case.

From the total productivities for all three cases (Panel (a)), we observe skill-biased technical changes (SBTC). Productivity growth was faster for high-skilled labor than for low-skilled labor. Furthermore, regional variation in the SBTC is observed, which is in line with the previous literature (e.g., Giannone 2017, and Rubinton 2020). For the underlying reason

[20] A simple illustration of this point is as follows. To economize on notations, let $p \equiv \log\left(\frac{W_t(1,n)}{W_t(0,n)}\right)$, and $q \equiv \log\left(\frac{L_t(1,n)}{L_t(0,n)}\right)$. Let's assume that the demand follows $p = a + bq + \epsilon_d$ and the supply is given by $p = c + dq + \epsilon_s$. In this case, if we run the OLS using the demand equation only, the bias for b is given by $\hat{b} - b = \frac{1}{d-b}\frac{var(\epsilon_d)}{var(q)}$. Therefore, if the supply is sufficiently inelastic so that $d - b > 0$, then the estimate $\hat{b}$ is biased upward.

Figure 4: Log Growth Rates of Relative Productivities Across States and Years

(a) $\frac{\bar{A}_{p,t}(e=1,n)}{\bar{A}_{p,t}(e=1,n)}$ (no aggl.) (b) $\frac{\bar{A}_{p,t}(e=1,n)}{\bar{A}_{p,t}(e=0,n)}$ (med aggl.) (c) $\frac{\bar{A}_{p,t}(e=1,n)}{\bar{A}_{p,t}(e=0,n)}$ (high aggl.)



* Note: A grayed-out line represents observations for a state. Each line represents $\log\left(\frac{\bar{A}_{p,t}(e=1,n)}{\bar{A}_{p,t}(e=0,n)}\right) - \log\left(\frac{\bar{A}_{p,1980}(e=1,n)}{\bar{A}_{p,1980}(e=0,n)}\right)$. (a), (b), and (c) illustrate the statistics for the underlying productivities for the no, medium, and high agglomeration cases.

behind the SBTC and the regional variation, the three cases assign different weights to the part driven by the exogenous productivities and the part endogenously determined by labor force composition. For the no agglomeration case, all the productivity growth is assigned to the exogenous part. For the high agglomeration case (Panel (c)), most of the SBTC is explained by the skill ratio and workforce size, and the exogenous underlying productivity contributes less to the SBTC.

The above observation is important to understand our quantitative results. If productivity is affected by the skill composition of a state, then the higher education policy can have a larger impact on local wages. Furthermore, the degree of agglomeration effects has implications for efficiency as well. Since the individuals do not internalize the agglomeration effects, the equilibrium is inefficient. Therefore, the presence of agglomeration effects implies that there is room for policies to mitigate the distortion.

Lastly, we obtain the housing stock for each region, $H(n)$, which is time-invariant, from the rental market clearing condition in 1980 (Equation (17)) and rents from data.

3.5 Solution Method

To check the fit of the baseline model and to run counterfactual exercises, we need to simulate the model with given primitives. The baseline model is solved (simulated forward in time) as below.

(a) In 1980, given $\{w_{1980}(e,n)\}_{e,n}$ and $\left\{M_{1980}(e^p,n^b,e,n^e,n^w)\right\}_{e^p,n^b,e,n^e,n^w}$, find the equilibrium $\{\tau(1,n)\}_n$, $\psi_{f,1980}(0,1)$, and $\{\psi_{n,1980}(0,n)\}_n$ such that Equations (14) to (16) are satisfied. To evaluate the three equations, all the other prices and allocations need

to be calculated using other equilibrium conditions. Also, note that the relative tax rates are directly from the data, and the levels are identified with $\psi_{f,1980}(0,1)$ and $\{\psi_{n,1980}(0,n)\}_n$.

(b) The solution to the previous step yields the measure of individuals in 1990, $\left\{M_{1990}(e^p, n^b, e, n^e, n^w)\right\}_{e^p,n^b,e,n^e,n^w}$, hence $\{L_{1990}(e,n)\}_{e,n}$. Utilizing Equation (24), derive $W_{1990}(e,n)$ from $A_{p,1990}(e,n)$ and $L_{1990}(e,n)$ for each $n \in \mathcal{N}$.

(c) Repeat Step (a), (b) for following years.

3.6 Model Fit

The model that we laid out does not fit the data perfectly because the individuals' preference parameters were estimated with errors in Equations (19) to (21). For the counterfactual exercises, we allocate the errors to the utility costs to achieve the perfect model fit, following the standard approach of the spatial economics literature (e.g., those surveyed by Redding and Rossi-Hansberg 2017). Let $\hat{m}_{n^w}(e,n,n^w)$ denote the estimated moving costs in Equation (19). We modify the utility costs in our model to vary over time and depend on the birth location so that they account for error terms: $m_{n^w,t}(n^b,e,n^e,n^w) = \hat{m}_{n^w}(e,n^e,n^w) + \hat{\sigma}_{n^w} u_{n^b,e,n^e,n^w,t}$. This modification subsequently changes the value $V_{n^e,t}(n^b,e,n^e)$ given in Equation (4). With the modified $V_{n^e,t}(n^b,e,n^e)$, in Equation (20), we assign the error terms to the utility costs $m_{n^e}(n^b,e,n^e)$ in the same manner. We proceed to the Equation (21) and obtain the perfect model fit.[21]

In Appendix D, we discuss the fit of the under-identified model, where the error terms are not allocated to the moving costs. We demonstrate that the model with no agglomeration effects fits the data well and that the model with agglomeration effects can predict the data moments one period forward well.

The reason for utilizing the exactly identified model instead of the under-identified model is mainly due to the agglomeration effects. With the effects, when we simulate our model forward from 1980, the errors in the earlier periods amplify, having an increasingly large impact on the later period. For example, in the individuals' choice problem in 1980, when the under-identified model predicts that a state has a higher share of high-skilled individuals in 1990 compared to data, it will predict that the state's high-skilled productivity in 1990 is higher than that in data. This implies a higher skill premium for the state in 1990, which subsequently makes the state more high-skill-intensive in the following periods.

[21] For the exact fit, the rental market clearing condition in each period (Equation (17)) should be satisfied by the rents in the data. For this purpose, we derive the rents in each period using the information on the housing stock, $H(n)$, and the total disposable income of a state from data.

4 Counterfactuals

Through the counterfactual exercises, we address the following question: "To what extent have the federal and state higher education policies contributed to the growing cross-state wage inequality since 1980?" We evaluate the impact through the standard deviation of mean wages across US states weighted by the size of the workforce in 1980.[22] We conduct the exercises for three cases: no agglomeration ($\eta_r^1 = \eta_s^1 = 0$), medium agglomeration ($\eta_r^1 = 1.34$, and $\eta_s^1 = 0.13$), and high agglomeration ($\eta_r^1 = 2.67$, and $\eta_s^1 = 0.25$).

While answering the question, we also assess other factors that affect the supply and demand of college education, such as changes in college's market power and production costs. Furthermore, in Appendix E, we utilize our model to answer another question that has attracted attention from both policymakers and researchers: "To what extent did each of the demand and supply factors of college education contribute to the increase in tuition?"

To study how changes in each factor affect the wage inequality, we fix the corresponding model primitives to their initial-year levels. First, we assess the impact of changes in tuition discount rates and changes in colleges' market power. (Recall the Bennett Hypothesis, which asserts the importance of the interaction between demand-side public policies and the market power of colleges.) We fix the tuition discount rates, $\{d_{n,t}(n^b, e, n), d_{f,t}(n^b, e, n)\}_{n^b,e,n,t}$, at their 1980 levels. Second, we fix the colleges' market power terms, $\{\phi_t(e = 1, n)\}_{n,t}$ at the initial-year level. Lastly, we fix both the discount rates and the market power terms to their initial levels. The results are reported in rows two through four in Table 5. Then, we measure the impact of changes in the rate of subsidies to colleges' production from the state governments, $\{s_{n,t}(e = 1, n)\}_{n,t}$, and those from the federal government, $\{s_{f,t}(e = 1, n)\}_{n,t}$, which is reported in row five. The overall impact of changes in both demand-side and supply-side policies is provided in row six. Lastly, we quantify the effects of Baumol's cost disease, i.e., changes in $\{MC(e = 1, n)\}_{n,t}$, on row seven.

The first noticeable pattern is that if there are no agglomeration effects, the counterfactuals have negligible effects on the rising inequality. In the absence of agglomeration effects, high-skilled and low-skilled wages are determined by the measure of high and low-skilled workers and their underlying productivities, $\bar{A}_{p,t}(1, n)$ and $\bar{A}_{p,t}(0, n)$. To illustrate this point, from Equation (12), we can derive two equations that yield the equilibrium wages:

$$\left(\frac{A_{p,t}(0,n)}{W_t(0,n)}\right)^{\rho-1} + \left(\frac{A_{p,t}(1,n)}{W_t(1,n)}\right)^{\rho-1} = 1 \tag{25}$$

[22] The unweighted standard deviation exhibits the same qualitative results, although the magnitude of quantitative results differs (Appendix F).

Table 5: Population-Weighted Standard Deviation of Log Mean Wages in 2019 Across the Baseline and Counterfactuals

	No Aggl. ($\eta_r^1 = 0$, $\eta_s^1 = 0$)		Med. Aggl. ($\eta_r^1 = 1.34$, $\eta_s^1 = 0.13$)		High Aggl. ($\eta_r^1 = 2.67$, $\eta_s^1 = 0.25$)	
	SD	Effect (%)	SD	Effect (%)	SD	Effect (%)
1. Baseline and Data	0.117		0.117		0.117	
* In 1980, stdev. of wages was 0.087.						
2. Discounts	0.118	3	0.119	7	0.122	17
3. Market power	0.117	0	0.114	−10	0.110	−23
4. Discounts + market power	0.117	0	0.114	−10	0.113	−13
5. Subsidies	0.117	0	0.118	3	0.127	33
6. Discounts + subsidies	0.118	3	0.122	17	0.139	73
7. Marginal cost	0.117	0	0.117	0	0.134	57

* Note: SD denotes a weighted standard deviation of log mean wages, with the weights being states' labor force in 1980. Effect calculates the counterfactual scenario's additional contribution to the increase in the standard deviation. For example, the contribution of tuition discounts in the no agglomeration case is calculated by (0.118 - 0.117)/(0.117-0.087) = 3%, meaning that without the changes in tuition discount rates, the standard deviation would have risen 3% more.

$$\frac{W_t(1,n)}{W_t(0,n)} = \left(\frac{L_t(1,n)}{L_t(0,n)}\right)^{-\frac{1}{\rho}} \left(\frac{A_{p,t}(1,n)}{A_{p,t}(0,n)}\right)^{\frac{\rho-1}{\rho}} \tag{26}$$

In the no agglomeration case, underlying productivity and total productivity are equal, i.e., $A_{p,t}(e,n) = \bar{A}_{p,t}(e,n)$. The demand and supply factors of higher education affect the mean wage only by affecting the skill composition of a region, $\{L_t(0,n), L_t(1,n)\}$. For example, factors that affect young individuals' choices in $t-1$ and, in turn, make a state more high-skilled-intensive in t, decrease the high-skilled wage and increase the low-skilled wage in t. The mean wage of the region, $\frac{W_t(0,n)L_t(0,n)+W_t(1,n)L_t(1,n)}{L_t(0,n)+L_t(1,n)}$, can increase or decrease based on which of the two following forces dominates: the larger weights are given to high-skilled wage $W_t(1,n)$ (which tends to be higher than $W_t(0,n)$) versus the decreasing skill premium due to high-skilled becoming more abundant. For the large quantitative effects of the counterfactuals on the spatial inequality, first, one of the two forces should dominate, and second, there should be significant cross-state variation in the changes of the primitives. However, quantitatively this was not the case, as demonstrated by the no agglomeration case in Table 5.

In contrast, in the high agglomeration case, the skill composition of a state can affect the total productivity through the agglomeration effects. Therefore, the skill premium does not necessarily decrease when the factors that make a state more high-skilled-intensive. This

can be seen by plugging Equation (11) into Equations (25) and (26):

$$\left(\frac{\bar{A}_{p,t}(0,n)}{W_t(0,n)}\right)^{\rho-1} + \left\{\frac{\bar{A}_{p,t}(1,n)}{W_t(1,n)}\left(\frac{L_t(1,n)}{L_t(0,n)}\right)^{\eta_r^1}\left(L_t(1,n)+L_t(0,n)\right)^{\eta_r^2}\right\}^{\rho-1} = 1 \tag{27}$$

$$\frac{W_t(1,n)}{W_t(0,n)} = \left(\frac{L_t(1,n)}{L_t(0,n)}\right)^{-\frac{1}{\rho}+\frac{\rho-1}{\rho}\eta_r^1}\left(L_t(0,n)+L_t(1,n)\right)^{\frac{\rho-1}{\rho}\eta_s^1}\left(\frac{\bar{A}_{p,t}(1,n)}{\bar{A}_{p,t}(0,n)}\right)^{\frac{\rho-1}{\rho}}. \tag{28}$$

Now, the factors that make the state more high-skilled-intensive or more populated can allow for higher mean wages. The agglomeration effect works as the force that can counteract the force of the high skill becoming more abundant, decreasing the skill premium. Given this, cross-state variation in the changes in the primitives, such as tuition discount rates or subsidy rates, can have large quantitative effects. This was indeed the case quantitatively. In Table 5, in the high agglomeration case, the counterfactual treatment of the model primitives had large effects. For example, if the tuition discount rates and production subsidy rates had remained at their 1980 level, the population-weighted standard deviation would have increased to 0.139 in 2019 from 0.087 in 1980. In the baseline, the standard deviation was 0.117 in 2019. This implies that the spatial inequality would have risen 73% faster—calculated by $(0.139-0.117)/(0.117-0.087)$—in the counterfactual economy. This means that the higher education policies by the state and federal governments had become more progressive in that they curbed the rising spatial inequality. In the medium agglomeration case, where the parameters η_r^1 and η_s^1 are halves as estimated, the standard deviation would have risen by 17% more in the counterfactual than in the baseline, instead of 73%. This result reiterates the importance of agglomeration effects in understanding the role of higher education policies.

Next, we delve into discovering which features of higher education policies decelerated the rise in inequality. First, we assess the impact of changes in each of the policy-related primitives (tuition discounts by the state governments, tuition discounts by the federal government, production subsidies by the state governments, and production subsidies by the federal government) in isolation. Table 6 demonstrates that the production subsidies provided by the states and the federal tuition discounts are the two main policies that reduced the spatial inequality.

Then, we explore through which mechanism each policy affects states' wage growth and, subsequently, cross-state inequality. We assess how changes in each policy from 1980 to 1990 affect the growth in 2000 wages utilizing our counterfactuals. Note that the higher education policy in 1990 affects wages in 2000 by affecting the stock of high- and low-skilled individuals in 2000. We only evaluate the wages in 2000 because, to study subsequent years, we need to

Table 6: Population-Weighted Standard Deviation of Log Mean Wages in 2019 Across the Baseline and Counterfactuals (Higher Education Policies)

	No Aggl. ($\eta_r^1 = 0$, $\eta_s^1 = 0$)		Med. Aggl. ($\eta_r^1 = 1.34$, $\eta_s^1 = 0.13$)		High Aggl. ($\eta_r^1 = 2.67$, $\eta_s^1 = 0.25$)	
	SD	Effect (%)	SD	Effect (%)	SD	Effect (%)
1. Baseline and Data	0.117		0.117		0.117	
* In 1980, stdev. of wages was 0.087.						
2. Discounts (states)	0.118	3	0.117	0	0.119	7
3. Discounts (federal)	0.117	0	0.118	3	0.120	10
4. Discounts (all)	0.118	3	0.119	7	0.122	17
5. Subsidies (states)	0.117	0	0.119	7	0.128	37
6. Subsidies (federal)	0.117	0	0.116	−3	0.115	−7
7. Subsidies (all)	0.117	0	0.118	3	0.127	33
8. Discounts + subsidies (all)	0.118	3	0.122	17	0.139	73

evaluate the cumulative impact of higher education policy. For example, to study wages in 2010, we need to assess the impact of policies in both 1990 and 2000, which complicates the analysis. Figure 5 illustrates the effect of federal policies. The states where tuition discount rates increased faster or college subsidy rates decreased slower were the ones where average wages increased faster in the baseline. This implies that the higher education policies targeted at promoting higher education, either through demand- or supply-side factors, increase the average wages of the states.[23] This is mainly due to the feature of federal policies that the costs of policies are shared by all states—when we solve the counterfactual economy, the tax rates proportionally scale up or down by the same factor for all types of individuals. Therefore, the states where federal policies promote college education relative to other states have a higher agglomeration of the high-skilled in the next period, and the wage increases faster there.

Unlike federal policies, where costs are shared by all states, state policies promoting higher education do not necessarily lead to faster wage growth (Figure 6). Counterintuitively, for some states, it is optimal to cut back on their higher education policies for wage growth. This is because states need to fund their policies by taxing their workers. Taxation leads to decreased income, which reduces in-migration to the state and increases the out-migration from the state. Furthermore, it can significantly impact the skill premium in the state depending on how progressive tax rates are, which in turn affects college attendance

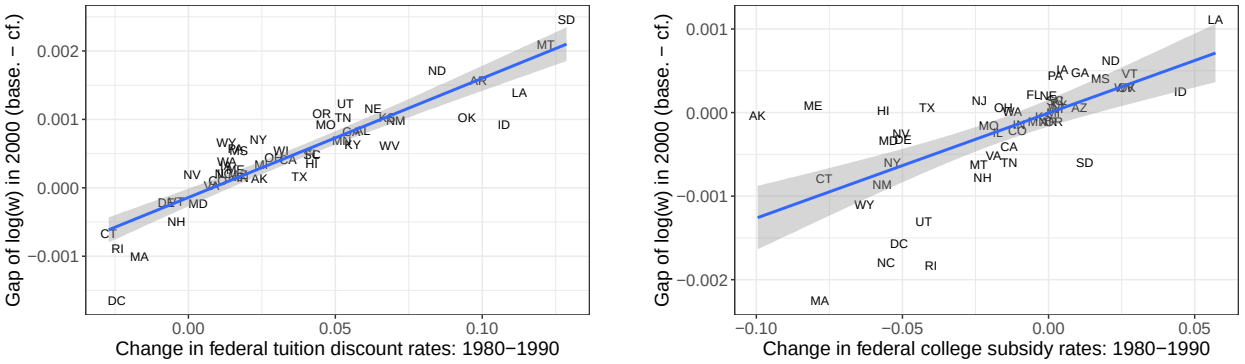
[23]The wage growth is significant for both low- and high-skilled labor. Although only high-skilled productivity is affected by the skill composition of the workforce, low-skilled wages respond through the demand for consumption goods and the substitutability between the skills for the production of goods—empirically, the effects from the increased demand outweighs the substitution towards high-skilled labor, increasing the low-skilled wages when high-skilled productivity rises.

choices. Therefore, the optimal state policy depends on multiple factors idiosyncratic to each state, such as whether it is more prone to the out-migration of high-skilled workers. This characteristic of state policies also reflects the additional value of our model. It is capable of evaluating the costs and benefits of actual and hypothetical policies while accounting for equilibrium effects associated with the policies.

Figure 5: Federal Policies in 1990 and Wages in 2000

(a) Changes in Tuition Discount Rates

(b) Changes in College Subsidy Rates

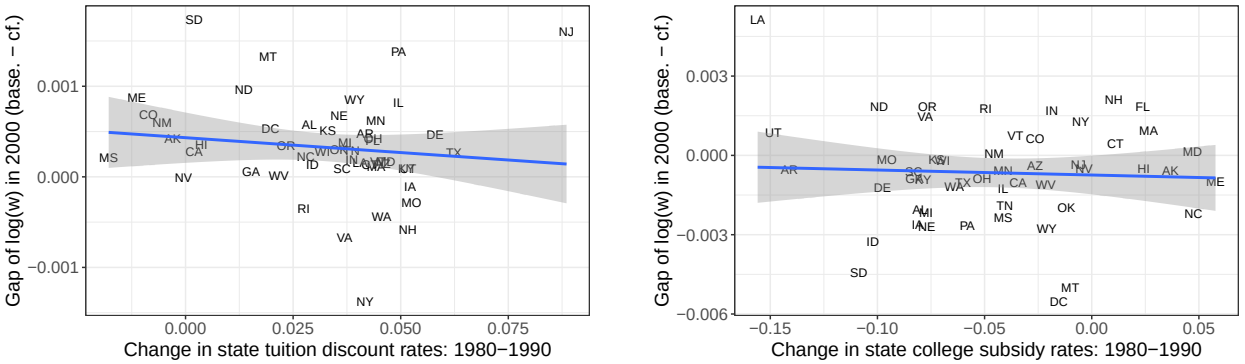


* Note: For (a), the x-axis represents changes in the tuition discount rates from 1980 to 1990, and the y-axis represents the log mean wages of a state in the baseline subtracted by that in the counterfactual when the discount rates in 1990 are set to those in 1980. Analogous explanation holds for (b).

Figure 6: State Policies in 1990 and Wages in 2000

(a) Changes in Tuition Discount Rates

(b) Changes in College Subsidy Rates



5 Conclusion

Higher education policies as a tool to reduce geographical inequality have not received much attention from either policymakers or academia. Industrial policies were mainly proposed

to mitigate the inequality (e.g., Build Back Better Regional Challenge). To explain the rise in spatial inequality, the academic literature has focused on the agglomeration of high-skilled labor due to inter-region migration. In contrast, we discovered that for a state-level analysis, the residents' college attendance is quantitatively more important than the migration of college graduates for the agglomeration of high-skilled workers. Furthermore, higher education market is one of the markets where the impact of public policies is the strongest. Motivated by these observations, we asked and addressed the following question: To what extent have the federal and state higher education policies contributed to the growing cross-state wage inequality since 1980?

By building and estimating a novel quantitative spatial model with endogenous higher education, we found that the higher education policies since 1980 by the federal and state governments significantly contributed to curbing the spatial inequality. If the policies had remained at their 1980 level, the inequality, measured by the population-weighted standard deviation of log mean wages across states, would have increased 73% faster. We also demonstrate that the agglomeration effects played a major role in this quantitative result. If the agglomeration effects were half as strong as the estimated level, the effect of the policies would decrease from 73% to 17%. Lastly, future research includes the robust estimation of agglomeration effects and the derivation of optimal state and federal policies.

References

Acemoglu, Daron. 2007. "Equilibrium bias of technology." *Econometrica*, 75(5): 1371–1409.

Autor, David H. 2014. "Skills, education, and the rise of earnings inequality among the "other 99 percent"." *Science*, 344(6186): 843–851.

Barro, Robert J and Xavier Sala-i Martin. 1992. "Convergence." *Journal of political Economy*, 100(2): 223–251.

Baumol, William J. 1967. "Macroeconomics of unbalanced growth: the anatomy of urban crisis." *The American economic review*, 57(3): 415–426.

Baumol, William J, William G Bowen, et al. 1993. *Performing arts-the economic dilemma: a study of problems common to theater, opera, music and dance.* Gregg Revivals.

Bennett, William J. 1987. "Our greedy colleges." *New York Times*, 18: A27.

Burland, Elizabeth, Susan Dynarski, Katherine Michelmore, Stephanie Owen, and Shwetha Raghuraman. 2023. "The power of certainty: experimental evidence on the effective design of free tuition programs." *American Economic Review: Insights*, 5(3): 293–310.

Cai, Zhifeng and Jonathan Heathcote. 2022. "College tuition and income inequality." *American Economic Review*, 112(1): 81–121.

Davis, Donald R and Jonathan I Dingel. 2020. "The comparative advantage of cities." *Journal of International Economics*, 123: 103291.

Deming, David J and Christopher R Walters. 2018. "The impact of state budget cuts on US postsecondary attainment." *Draft, Harvard University*, 1567–1633.

Desmet, Klaus and Esteban Rossi-Hansberg. 2014. "Spatial development." *American Economic Review*, 104(4): 1211–1243.

Diamond, Rebecca. 2016. "The determinants and welfare implications of US workers' diverging location choices by skill: 1980–2000." *American Economic Review*, 106(3): 479–524.

Donovan, Kevin and Christopher Herrington. 2019. "Factors affecting college attainment and student ability in the us since 1900." *Review of Economic Dynamics*, 31: 224–244.

Eckert, Fabian. 2019. "Growing apart: Tradable services and the fragmentation of the us economy." *mimeograph, Yale University*.

Eeckhout, Jan, Roberto Pinheiro, and Kurt Schmidheiny. 2014. "Spatial sorting." *Journal of Political Economy*, 122(3): 554–620.

Epple, Dennis, Richard Romano, and Holger Sieg. 2006. "Admission, tuition, and financial aid policies in the market for higher education." *Econometrica*, 74(4): 885–928.

Epple, Dennis, Richard Romano, Sinan Sarpça, and Holger Sieg. 2017. "A general equilibrium analysis of state and private colleges and access to higher education in the US." *Journal of Public Economics*, 155: 164–178.

Epple, Dennis, Richard Romano, Sinan Sarpça, Holger Sieg, and Melanie Zaber. 2019. "Market power and price discrimination in the US market for higher education." *The RAND Journal of Economics*, 50(1): 201–225.

Fu, Chao. 2014. "Equilibrium tuition, applications, admissions, and enrollment in the college market." *Journal of Political Economy*, 122(2): 225–281.

Ganong, Peter and Daniel Shoag. 2017. "Why has regional income convergence in the US declined?" *Journal of Urban Economics*, 102: 76–90.

Giannone, Elisa. 2017. "Skilled-biased technical change and regional convergence." *University of Chicago. Unpublished manuscript*.

Gordon, Grey and Aaron Hedlund. 2022. "Accounting for tuition increases across U.S. colleges." *mimeograph*.

Katz, Lawrence F and Kevin M Murphy. 1992. "Changes in relative wages, 1963–1987: supply and demand factors." *The quarterly journal of economics*, 107(1): 35–78.

Kennan, John. 2015. "Spatial variation in higher education financing and the supply of college graduates." National Bureau of Economic Research.

McFadden, Daniel. 1972. "Conditional logit analysis of qualitative choice behavior."

Redding, Stephen J and Esteban Rossi-Hansberg. 2017. "Quantitative spatial economics." *Annual Review of Economics*, 9: 21–58.

Rubinton, Hannah. 2020. "The geography of business dynamism and skill-biased technical change." *FRB St. Louis Working Paper*, , (2020-20).

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek. 2021. "Ipums usa: Version 10.0. Minneapolis, MN: IPUMS, 2020."

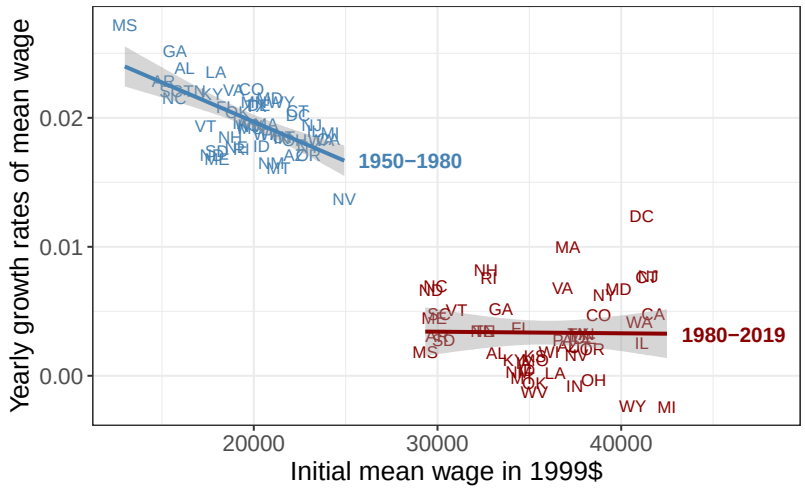
Singleton, John D. 2019. "Incentives and the supply of effective charter schools." *American Economic Review*, 109(7): 2568–2612.

Appendix

A Rise in Spatial Inequality

A.1 Beta Convergence

Figure A.1: Beta Convergence Rates: Pre- vs Post-1980



* Source: Decennial Census and American Community Survey (ACS).
* Note: y-axis represents $\frac{\log(w_{i,t_2})-\log(w_{i,t_1))}{t_2-t_1}$, where i denotes each state, and t_1 and t_2 are the start and end years. Hawaii and Alaska are not included, as they did not utilize the standard census questionnaire in 1950.

A.2 Decomposition of the Rise in Spatial Inequality

This section decomposes the rise in the spatial inequality measured by the variance in log mean wages across states (sigma divergence). We apply Shapley decomposition to the increase of the population-weighted standard deviation since 1980, the right part of Figure 1.[24] The mean wage of state n at time t can be written as the weighted average of low-skilled wages $(W_t(0,n))$ and high-skilled wages $(W_t(1,n))$. Letting $S_t(1,n)$ denote the share of high-skilled in state n, the mean wage is given by

$$\overline{W}_t(n) = W_t(1,n)S_t(1,n) + W_t(0,n)\left(1 - S_t(1,n)\right).$$

Therefore, we can view the log mean wage of the state as a function of the low-skilled and high-skilled wages and the share of high-skilled: $\log\left(\overline{W}_t(n)\right) \equiv F\left(W_t(1,n), W_t(0,n), S_t(1,n)\right)$.

[24] This can also be regarded as a Blinder-Oaxaca decomposition, as well.

Let $Std\left(\{F\left(W_t(1,n), W_t(0,n), S_t(1,n)\right)\}_n\right)$ denote the standard deviation across states at time t.

One way to decompose the change in the standard deviation from time t to $t+k$ is as follows:

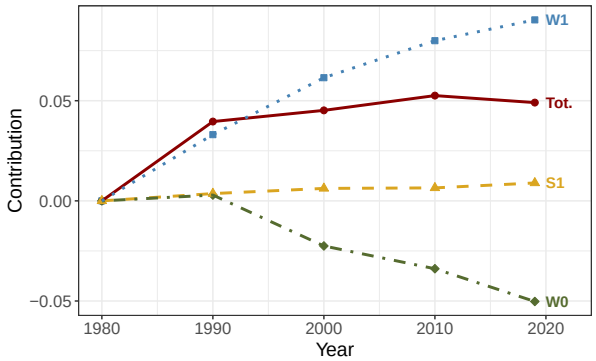
$$
\begin{aligned}
&Std\left(\{F\left(W_{t+k}(1,n), W_{t+k}(0,n), S_{t+k}(1,n)\right)\}_n\right) - Std\left(\{F\left(W_t(1,n), W_t(0,n), S_t(1,n)\right)\}_n\right) \\
&= \underbrace{Std\left(\{F\left(W_{t+k}(1,n), W_{t+k}(0,n), S_{t+k}(1,n)\right)\}_n\right) - Std\left(\{F\left(W_{t+k}(1,n), W_{t+k}(0,n), S_t(1,n)\right)\}_n\right)}_{\text{Contribution of } S_{t+k}(1,n) - S_t(1,n)} \\
&= \underbrace{Std\left(\{F\left(W_{t+k}(1,n), W_{t+k}(0,n), S_t(1,n)\right)\}_n\right) - Std\left(\{F\left(W_{t+k}(1,n), W_t(0,n), S_t(1,n)\right)\}_n\right)}_{\text{Contribution of } W_{t+k}(0,n) - W_t(0,n)} \\
&= \underbrace{Std\left(\{F\left(W_{t+k}(1,n), W_t(0,n), S_t(1,n)\right)\}_n\right) - Std\left(\{F\left(W_t(1,n), W_t(0,n), S_t(1,n)\right)\}_n\right)}_{\text{Contribution of } W_{t+k}(1,n) - W_t(1,n)}.
\end{aligned}
$$

In the above, we assessed the contribution of changes in the high-skilled share, low-skilled wages, and high-skilled wages, by changing each component in order. Since there are six possible paths to conduct this decomposition, we apply the Shapley decopmosition, where we assess all six paths and take averages to measure the contribution of each factor.

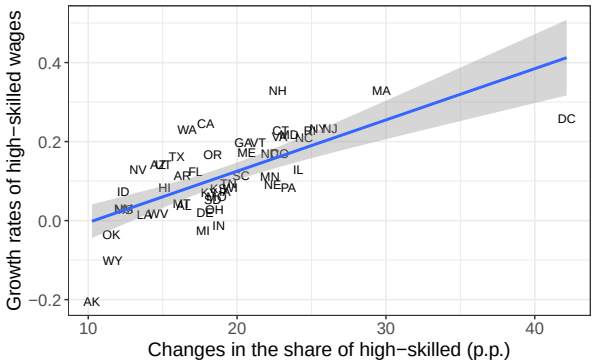
Figure A.2a reports the decomposition results. Changes in the high-skilled wage were the major contributor to the rising inequality, with the changes in the high-skilled share being the next important component. In contrast, the contribution of changes in the low-skilled wages was negative. Furthermore, the two positive components (changes in $\{W_t(1,n)\}_n$ and

Figure A.2: Causes of the Rising Spatial Inequality

(a) Shapley Decomposition



(b) Changes in High-skilled Shares and Wages



* Note on (a): "Tot." denotes the total changes in the population-weighted standard deviation from 1980. "W1", "W0", and "S1" denote the contributions of changes in high-skilled wages, low-skilled wages, and high-skilled shares, respectively.
* Note on (b): y-axis represent the log growth rates of high-skilled wages in 1999$.

$\{S_t(1,n)\}_n$) are positively correlated (Figure A.2b).[25] These two findings suggest that the growing disparity in high-skilled wages and shares are likely the drivers behind the post-1980 rise in the inequality, in line with the existing literature.

B Discrete Choice Problem

This section outlines how to derive the solutions to the discrete choice problem. Let a random variable X follow the Gumbel distribution (type-I extreme value distribution) with a cumulative distribution function, $Q(x) = \exp(-a\exp(-x))$ and associated probability distribution function, $q(x) = a\exp(-x)\exp(-a\exp(-x))$. Then, $\mathbb{E}(x) = \log(a) + \gamma$, where γ is the Euler-Mascheroni constant.

Assume a simple discrete choice problem where a utility-maximizing individual is choosing among $\mathcal{I} = \{1, ..., I\}$. Utility from choosing an option $i \in \mathcal{I}$ is given by $V_i + \sigma\epsilon_i$, where ϵ_i is drawn i.i.d from cumulative distribution $F(x) = \exp(-\exp(-x))$. Let $H(\epsilon_1, ..., \epsilon_I)$ denote the joint cdf of the utility shocks. Let $H_j(\cdot)$ denote the j-th partial derivative of $H(\cdot)$.

Probability of choosing $j \in \mathcal{I}$,

$$
\begin{aligned}
& P_j \\
=& \int_{-\infty}^{\infty} H_j\left(\frac{V_1 - V_j}{\sigma} + \epsilon_j, ..., \frac{V_J - V_j}{\sigma} + \epsilon_j\right) d\epsilon_j \\
=& \int_{-\infty}^{\infty} \exp(-\epsilon_j)\exp\left(-\sum_{i=1}^{I}\exp\left(\frac{V_i - V_j}{\sigma} + \epsilon_j\right)\right) d\epsilon_j \\
=& \int_{-\infty}^{\infty} \exp(-\epsilon_j)\exp\left(-\exp(-\epsilon_j)\sum_{i=1}^{I}\exp\left(\frac{V_i - V_j}{\sigma}\right)\right) d\epsilon_j \\
=& \frac{1}{\sum_{i=1}^{I}\exp\left(\frac{V_i - V_j}{\sigma}\right)} \int_{-\infty}^{\infty} \exp(-\epsilon_j)\sum_{i=1}^{I}\exp\left(\frac{V_i - V_j}{\sigma}\right)\exp\left(-\exp(-\epsilon_j)\sum_{i=1}^{I}\exp\left(\frac{V_i - V_j}{\sigma}\right)\right) d\epsilon_j \\
=& \frac{1}{\sum_{i=1}^{I}\exp\left(\frac{V_i - V_j}{\sigma}\right)} \\
=& \frac{\exp\left(\frac{V_j}{\sigma}\right)}{\sum_{i=1}^{I}\exp\left(\frac{V_i}{\sigma}\right)}.
\end{aligned}
$$

[25] The figure can be also viewed as suggestive evidence behind the agglomeration effects for high-skilled.

Given that j is chosen, the expected utility,

$$
\begin{aligned}
&\mathbb{E}\left[V_j + \sigma\epsilon_j | j = argmax_{i\in\{1,...I\}}(V_i + \sigma\epsilon_i)\right] \\
=&\frac{1}{P_j}\int_{-\infty}^{\infty}(V_j + \sigma\epsilon_j)H_j\left(\frac{V_1 - V_j}{\sigma} + \epsilon_j, ..., \frac{V_J - V_j}{\sigma} + \epsilon_j\right)d\epsilon_j \\
=&\frac{1}{P_j}\int_{-\infty}^{\infty}(V_j + \sigma\epsilon_j)\exp(-\epsilon_j)\exp\left(-\sum_{i=1}^{I}\exp\left(\frac{V_i - V_j}{\sigma} + \epsilon_j\right)\right)d\epsilon_j \\
=&\sigma\int_{-\infty}^{\infty}\left\{\sum_{i=1}^{I}\exp\left(\frac{V_i}{\sigma}\right)\right\}\left(\frac{V_j + \sigma\epsilon_j}{\sigma}\right)\exp\left(-\frac{V_j + \sigma\epsilon_j}{\sigma}\right) \\
&\times\exp\left(-\left\{\sum_{i=1}^{I}\exp\left(\frac{V_i}{\sigma}\right)\right\}\exp\left(-\frac{V_j + \sigma\epsilon_j}{\sigma}\right)\right)d\epsilon_j \\
=&\sigma\int_{-\infty}^{\infty}x\left\{\sum_{i=1}^{I}\exp\left(\frac{V_i}{\sigma}\right)\right\}\exp(-x)\exp\left(-\left\{\sum_{i=1}^{I}\exp\left(\frac{V_i}{\sigma}\right)\right\}\exp(-x)\right)dx \\
=&\sigma\left\{\log\left(\sum_{i=1}^{I}\exp\left(\frac{V_i}{\sigma}\right)\right) + \gamma\right\}.
\end{aligned}
$$

Therefore, the ex-ante utility of the consumer,

$$
\begin{aligned}
&\mathbb{E}\left[\max_{i\in\{1,...I\}}\{V_i + \sigma\epsilon_i\}\right] \\
=&\sum_{j\in\{1,...I\}}P_j\mathbb{E}\left[V_j + \sigma\epsilon_j | j = argmax_{i\in\{1,...I\}}(V_i + \sigma\epsilon_i)\right] \\
=&\sigma\left\{\log\left(\sum_{i=1}^{I}\exp\left(\frac{V_i}{\sigma}\right)\right) + \gamma\right\}.
\end{aligned}
$$

C Data Appendix

This appendix outlines how different variables used in the paper have been constructed from the data.

(i) **Expenditure share on housing:** α

We first construct the housing expenditure as a share of average annual expenditure for each decennial year from the BLS CEX topline means table. The housing expenditure share for different years is obtained as: share in 1990 = 8,703/28,381 = 31%, share in 2000 = 12,319/38,045 = 32%, share in 2010 = 34%, share in 2019 = 20,679/63,036 = 33%. We took the average of the four years, and set $\alpha = 0.325$.

(ii) **Rental rate by state:** $r(n)$

To derive the average rental rate by state, we employ the hedonic price regression,

similar to Eeckhout, Pinheiro, and Schmidheiny (2014). For each year, we run the following regression.

$$(rent)_{is} = (bedroom)_{is} + (units)_{is} + (years)_{is} + F_s + \epsilon_{is},$$

where i represents each census tract in a state s, $(rent)_{is}$ is the median gross rent of the tract, $(bedroom)_{is}$ is the median number of bedrooms across housing units in the tract, $(units)_{is}$ is the median number of units in the structure, and $(years)_{is}$ is the median age of housing units. Each regressor is a categorical variable defined as follows: $(bedroom)_{is} \in \{0, 1, 2, 3, 4, (5 \text{ or more})\}$. $(units)_{is} \in \{(1, \text{detached}), (1, \text{attached}), 2, 3-4, (5 \text{ or more})\}$.
$(years)_{is} \in \{(1, \text{detached}), (5 \text{ or less}), 6-10, 11-20, 21-30, 31-40, (41 \text{ or more})\}$.
We derive the 'average' rent of a state by using the state fixed effect and the median housing characteristics of all 50 states:

$$\widehat{(rent)}_s = \overline{(bedroom)} + \overline{(units)} + \overline{(years)} + \widehat{F}_s,$$

where $\overline{(bedroom)}$ denotes the estimates for the median number of bedrooms of housing units in the US. $\overline{(units)}$ and $\overline{(years)}$ are defined likewise.

For our model parameterization, we are using the rents for a single year. Therefore, we regress the state-year rents, $\widehat{(rent)}_{s,t}$, on state fixed effects and year fixed effects and utilize the state fixed effects as our measure of $r_{1980}(n)$.

(iii) **Housing stock by state:** $H(n)$
The housing stock for a region is derived from the housing market clearing condition and the rental rates and disposable income obtained from data for 1980— specifically,

$$r_{1980}(n)H(n) = \alpha DI_{1980}(n),$$

where $DI_{1980}(n)$ is the total disposable income in a region in 1980 (sum of the right-hand side of the budget constraint across education types).

(iv) **Tax rates on labor income:** $\left\{\psi_t^f(e,n), \psi_t^n(e,n)\right\}_{e,n,t}$
We compute the tax rates on labor income using the NBER-Taxsim model. We compute the federal and state tax rates separately for college- and non-college educated individuals using the mean wages earned by college- and non-college educated individuals in a state at between the working ages of 23–62 years. The tax rates are computed for a representative married individual of age 43 (mean working age in the model).

(v) **Full-time equivalent (FTE) enrollments**
We obtain the total full-time equivalent enrollments in a state as the sum of total full-time enrollments and half the part-time enrollments across four-year degree granting institutions in a state in an academic year.

(vi) **Gross tuition cost:** $\tau_t(e, n)$
We construct gross tuition costs as the revenue from tuition and sales of services per FTE students in a state.

(vii) **Federal discounts:** $d_{f,t}(n^b, e, n)$
We construct federal discounts as the share of total revenue from Pell and other federal grants towards tuition and other services. Note that we assume both in-state and out-of-state are eligible for federal grants.

(viii) **State and local discounts:** $d_{n,t}(n^b, e, n)$
We construct state and local discounts as the share of total revenue from state and local grants towards tuition and other services.

(ix) **Marginal cost:** $MC_t(e, n)$
We construct the marginal cost of operating a college as the expenditure per FTE students.

(x) **Federal subsidies:** $s_{f,t}(e, n)$
We construct federal subsidies as the share of total expenditure financed from federal appropriations.

(xi) **State and local subsidies:** $s_{n,t}(e, n)$
We construct state and local subsidies as the share of total expenditure financed from state and local appropriations.

(xii) **College attendance choice probability:** $Pr_t[e|e^p, n^b]$
We obtain the probability of going to college given state of birth and education of parents from a cross-tabulation matrix constructed from NCES B&B dataset. For the missing values, we apply generalized RAS algorithms to impute them.

(xiii) **College location choice probability:** $Pr_t[n^e|e, n^b]$
We construct the college location choice probabilities for agents who decide to go to college from the share of students studying in a state for different state of birth. We obtain this data from IPEDS.

(xiv) **Work location choice probability:** $Pr_t[n^w|e, n^e]$

We obtain the work location probabilities for non-college-educated individuals from the cross tabulation of state of current residence against state of birth from Decennial Census. We obtain the work location probabilities for college-educated individuals from the cross tabulation of state of college and state of employment from NCES B&B. For the missing values, we apply generalized RAS algorithms to impute them.

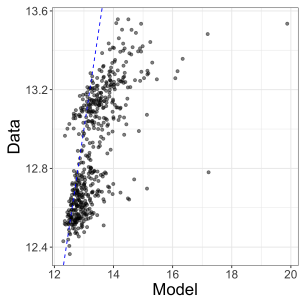
D Fit of the Under-identified Model

In this section, we assess the fit of the under-identified model, where we do not allocate the errors in the preference estimation to the time-varying utility costs. We utilize the five key moments: the three sets of choice probabilities across time, $\{Pr_t[n^w|n^b, e, n^e]\}_{t,n^b,e,n^e,n^w}$, $\{Pr_t[n^e|n^b, e]\}_{t,n^b,e,n^e}$, and $\{Pr_t[e = 1|e^p, n^b]\}_{t,e^p,n^b}$, along with the equilibrium prices of the labor and higher-education markets: $\{w_t(e, n)\}_{t,e,n}$ and $\{\tau_t(e = 1, n)\}_{t,n}$.

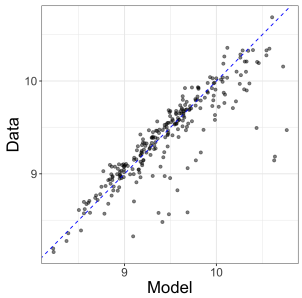
We simulate forward from 1980 to 1990, 2000, 2010, and 2019, for the no-agglomeration case. Figure D.1 describes that the model fits for this case. For the medium-agglomeration

Figure D.1: Model Fit (Non-perfect Fit, No Agglomeration)

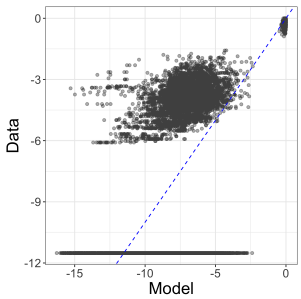
(a) Log Wages: $\{\log(W_t(e, n))\}_{e,n,t}$



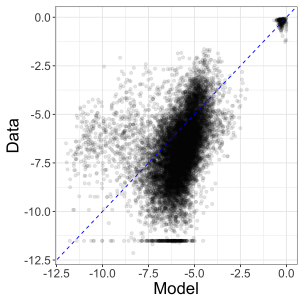
(b) Log Tuition: $\{\log(\tau_t(1, n))\}_{n,t}$



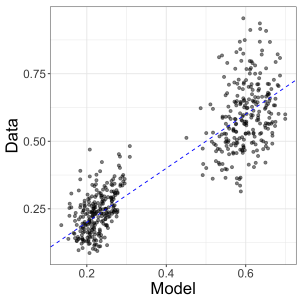
(c) CP for work location: $\{\log(Pr_t[n^w|n^b, e, n^e])\}_{n^b,e,n^e,n^w,t}$



(d) CP for college location: $\{\log(Pr_t[n^e|n^b, e = 1])\}_{n^b,n^e,t}$



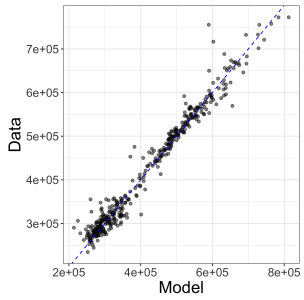
(e) CP for college attendance: $\{\log(Pr_t[e = 1|e^p, n^b])\}_{e^p,n^b,t}$



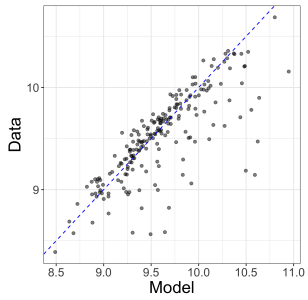
* Note: CP denotes log choice probabilities. Blue dashed lines represent 45-degree lines.

Figure D.2: Model Fit (Non-perfect Fit, Medium Agglomeration)

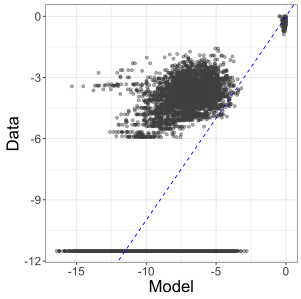
(a) Log Wages: $\{\log(W_t(e,n))\}_{e,n,t}$



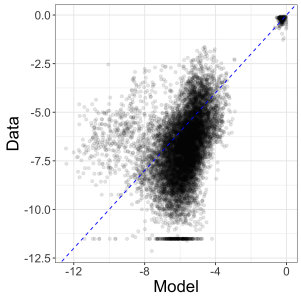
(b) Log Tuition: $\{\log(\tau_t(1,n))\}_{n,t}$



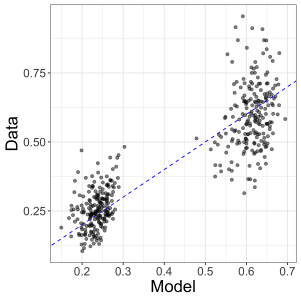
(c) CP for work location: $\{\log\left(Pr_t[n^w|n^b,e,n^e]\right)\}_{n^b,e,n^e,n^w,t}$



(d) CP for college location: $\{\log\left(Pr_t[n^e|n^b,e=1]\right)\}_{n^b,n^e,t}$



(e) CP for college attendance: $\{\log\left(Pr_t[e=1|e^p,n^b]\right)\}_{e^p,n^b,t}$



* Note: CP denotes log choice probabilities. Blue dashed lines represent 45-degree lines.

case, we solve the model for each decade and test the model performance in predicting the next decade's equilibrium prices and allocations. For example, we predict 2000 using the estimated model with the data from 1990 and compare the model fit in 2000. Figure D.2 illustrates the model fit for 1990, 2000, 2010, and 2019. Taken together, we conclude that even without fitting the model perfectly, it gives reasonable predictions.

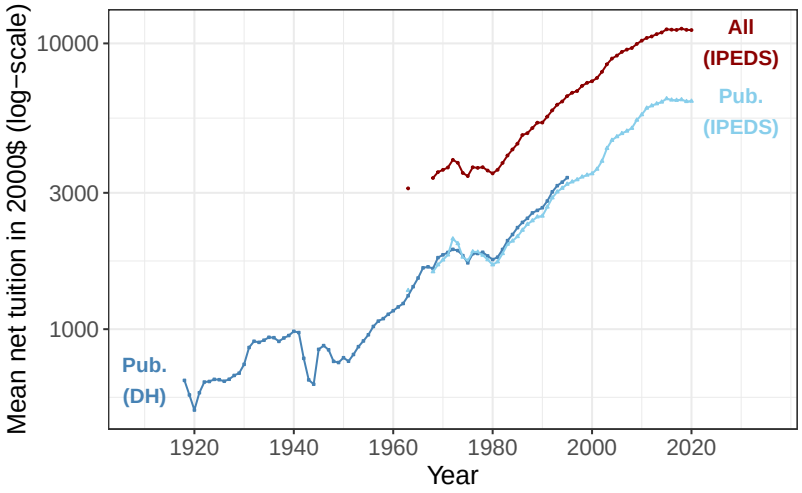
E Rise in College Tuition

In this section, we utilize our model to quantitatively study the factors behind the rise of college tuition (Figure E.1 and Figure 3c). By utilizing the same counterfactuals as in Section 4, we ask and answer: To what extent did each of the demand and supply factors of college education contribute to the increase in tuition?

We report the result in Table 7. For the exposition of the results, we use the medium agglomeration case.[26] In terms of the growth rate of median tuition across 51 states (50 states

[26] Although the quantitative results are different across three cases, the rank orders of each counterfactual scenario's contribution to the tuition hike are similar.

Figure E.1: Average Net Tuition: 1918-2020



* Source: Integrated Postsecondary Education Data System (IPEDS), and Donovan and Herrington (DH, 2019)
* Note on (c): Net tuition represents tuition paid by students after all scholarships and grants. "All" denotes the mean net tuition for all colleges. "Pub." denotes that for public colleges. Nominal tuition is expressed in 2000$ by denominating it with CPI. The reason why the increase in tuition is different from Figure 3c is that for the model parameterization, we adopt the broad notion of tuition, which also includes the sales of services (e.g., room and board).

+ D.C.) from 1980 to 2019, in the baseline and data, the median tuition increased by 36% (Figure 3c). The changes in the tuition discount rates explained 47% $(= (0.36 - 0.19))/0.36)$ of the tuition growth. This is in line with the increase in the discounts provided by the state and federal governments (Figures 3e and 3f). In the existence of colleges' market power, higher discounts make the demand for colleges less elastic, increasing the markups in Equation (14).[27] The overall increase in the colleges' market power terms—illustrated in Table 3—explained 72% of the tuition increase.[28] Changes in the discount rates and market power altogether accounted for 142%. This result echoes the Bennett Hypothesis, under which tuition increased because colleges "exploited" the public programs targeted to reduce tuition with their market power.

Next, changes in the rate of subsidies to colleges' production from the state govern-

[27]By taking the partial derivative of the price elasticity of demand with respect to tuition, one can verify this.

[28]This result should be interpreted with caution. The $\phi_t(e, n)$ is a 'wedge' between a markup from the data and the inverse demand elasticity calculated from preference parameters and data. Therefore, although it is possible that $\phi_t(e, n)$ correctly captures the increasing market power of colleges, it could possibly reflect the degree of model misspecification. Epple et al. (2019) make headway in measuring the degree of market power, but their analysis depends on the cross-sectional data. Extending their analysis to different time horizons could give insights into the extent to which $\phi_t(e, n)$ captures the increasing market power of colleges; however, we leave it for future research.

Table 7: Growth Factor of Median College Tuition across States from 1980 to 2019 in the Baseline and Counterfactuals

	No Aggl. ($\eta_r^1 = 0$, $\eta_s^1 = 0$)		Med. Aggl. ($\eta_r^1 = 1.34$, $\eta_s^1 = 0.13$)		High Aggl. ($\eta_r^1 = 2.67$, $\eta_s^1 = 0.25$)	
	Gr.	% Expl.	Gr.	% Expl.	Gr.	% Expl.
1. Baseline and data	1.36		1.36		1.36	
2. Discounts	1.20	44	1.19	47	1.19	47
3. Market power	1.10	72	1.10	72	1.00	100
4. Discounts + market power	0.85	142	0.85	142	1.00	100
5. Subsidies	1.13	64	1.13	64	1.17	53
6. Discounts + subsidies	1.04	89	0.86	139	1.06	83
7. Marginal cost	1.13	64	1.13	64	1.26	28

* Note: Gr. represents the growth factor of median, i.e., (Median of $\tau_{2019}(e = 1, n)$)/(Median of $\tau_{1980}(e = 1, n)$). % Expl. denotes the percentage of the growth in median tuition explained by the counterfactual scenario. For example, 44% in the no agglomeration case for the discounts is calculated by $(0.36-0.20)/(0.36)$.

ments, $\{s_{n,t}(e = 1, n)\}_{n,t}$, which declined (Figure 3g), and those from the federal government, $\{s_{f,t}(e = 1, n)\}_{n,t}$, which stagnated (Figure 3h), together accounted for 64% of the tuition hike. Notably, the demand-side and supply-side public policies, i.e., tuition discounts and production subsidies by the state and federal governments, accounted for 139% of the increase in tuition. The quantitative importance of public policies mirrors Gordon and Hedlund (2022), who found that 'financial aid expansions and lagging state support [are] salient forces' behind the tuition increase. Lastly, Baumol's cost disease, i.e., the increase in the prodution costs (Figure 3d), explained 64% of the median tuition increase.

F Results with the Unweighted Standard Deviation

This section reports the main results in Table 5 and Table 6 using the unweighted standard deviation.

Table 8: Unweighted Standard Deviation of Log Mean Wages in 2019 Across the Baseline and Counterfactuals

	No Aggl. ($\eta_r^1 = 0$, $\eta_s^1 = 0$)		Med. Aggl. ($\eta_r^1 = 1.34$, $\eta_s^1 = 0.13$)		High Aggl. ($\eta_r^1 = 2.67$, $\eta_s^1 = 0.25$)	
	SD	Effect (%)	SD	Effect (%)	SD	Effect (%)
1. Baseline and data	0.139		0.139		0.139	
* In 1980, stdev. of wages was 0.117.						
2. Discounts	0.139	0	0.140	5	0.142	14
3. Market power	0.139	0	0.136	−14	0.127	−55
4. Discounts + market power	0.139	0	0.137	−9	0.137	−9
5. Subsidies	0.139	0	0.138	−5	0.144	23
6. Discounts + subsidies	0.140	5	0.145	27	0.168	132
7. Marginal cost	0.139	0	0.142	14	0.176	168

Table 9: Unweighted Standard Deviation of Log Mean Wages in 2019 Across the Baseline and Counterfactuals (Higher Education Policies)

	No Aggl. ($\eta_r^1 = 0$, $\eta_s^1 = 0$)		Med. Aggl. ($\eta_r^1 = 1.34$, $\eta_s^1 = 0.13$)		High Aggl. ($\eta_r^1 = 2.67$, $\eta_s^1 = 0.25$)	
	SD	Effect (%)	SD	Effect (%)	SD	Effect (%)
1. Baseline and data	0.139		0.139		0.139	
* In 1980, stdev. of wages was 0.117.						
2. Discounts (states)	0.140	5	0.137	−9	0.135	−18
3. Discounts (federal)	0.139	0	0.141	9	0.148	41
4. Discounts (all)	0.139	0	0.140	5	0.142	14
5. Subsidies (states)	0.140	5	0.141	9	0.152	59
6. Subsidies (federal)	0.139	0	0.138	−5	0.138	−5
7. Subsidies (all)	0.139	0	0.138	−5	0.144	23
8. Discounts + subsidies (all)	0.140	5	0.145	27	0.168	132