On the Spatial Distribution of Colleges*

Jacob Wright[†]

University of Minnesota

December 8, 2025

CLICK HERE FOR MOST RECENT VERSION

Abstract

Given the decentralized nature of American higher education, there is substantial cross-state heterogeneity in: (1) in-/out-of-state tuition, (2) spending per student, and (3) in-/out-of-state capacity. Empirically, I show that larger and wealthier states charge higher in-/out-of-state tuition, spend more per student, and offer fewer seats to out-of-state students. I develop a novel model that embeds a heterogeneous-agent lifecycle structure within a quantitative spatial framework, featuring an endogenous distribution of college quality and firm activity across locations, as well as migration-based sorting of students and workers. I estimate the model to replicate cross-state dynamics of migration and education choices, wages, and college characteristics. The model rationalizes key empirical observations and provides a clear economic understanding of why college policies differ across states. For a fixed level of college expenditures, optimal federal policy increases aggregate welfare by 3.3%. The federal solution guides a policy analysis, which shows that spatial policies are markedly more effective at increasing welfare than standard proposals.

JEL Codes: H52, H75, I23, I28, R12, R13

*A special thanks to Ellen McGrattan, Todd Schoellman, and Michael Waugh for their invaluable guidance. I also thank Manuel Amador, Anmol Bhandari, Gorkem Bostanci, Elizabeth Caucutt, Varadarajan Chari, Giovanni Gallipoli, Juliana Gamboa-Arbelaez, Thomas Hasenzagl, Kyle Herkenhoff, Thomas Holmes, Torsten Jaccard, Oksana Leukhina, Lance Lochner, Joseph Mullins, Anusha Nath, Sergio Ocampo-Díaz, Andrii Parkhomenko, Jesse Perla, Amil Petrin, Pau Pujolas, Gajendran Raveendranathan, Kjetil Storesletten, Michal Szkup, Abigail Wozniak, Bipul Verma, Angela Zheng, and participants at the University of Iowa, University of British Columbia, Duke University, Western University, McMaster University, Queen's University, Minneapolis Federal Reserve, and University of Toronto for many helpful comments and discussions. I am grateful for support from the Hutcheson-Lilly Dissertation Fellowship for completion of this research.

†email: wrig0974@umn.edu. website: jacob-wright.me.

1 Introduction

There is considerable public debate over how to design and finance a system of higher education. The United States is an ideal setting to study these questions, given the highly decentralized nature of American higher education, which has led to dramatic heterogeneity in policy choices across states. This paper studies three main aspects in which systems of higher education differ: (1) in-/out-of-state tuition, (2) spending per student, and (3) in-/out-of-state seat availability, and asks: "What is the optimal federal-level distribution of college tuition, spending, and capacity across U.S. states?"

In answering this normative question, I also address the natural positive question: "Why do college systems differ substantially across states?" I begin by documenting a set of motivating facts in the data to (1) investigate which state characteristics are associated with particular college policies and (2) motivate the quantitative framework. States with larger populations and higher wages tend to charge higher tuition, allocate more resources per student, and offer fewer seats to out-of-state students. These states also exhibit lower rates of alumni outmigration and greater inflows of college-educated workers from other states.

To rationalize these relationships and answer my main research questions, I develop a novel quantitative framework. The model endogenizes the relationships between migration flows, firm activity, college quality, and higher education policies across states. It thereby captures how heterogeneity in location fundamentals across space determine patterns of college quality and enrollment, firm activity, and government policy.

The model consists of two parts, which together provide a natural representation of the U.S. economy and system of higher education. The first is a demand curve for college, where the key decision-makers are households, firms, and colleges that take as given in-/out-of-state tuition, spending per student, and in-/out-of-state capacity. The second part is a state optimization problem over these college policies, taking as given this demand for college.

Modeling the college demand curve features a lifecycle model of human capital accumulation and migration embedded in a general equilibrium spatial framework. High school graduates with differing ability, human capital, and parental income choose whether to attend college and, if so, in which state. Each state features a college of differentiated quality and an admission cutoff for in- and out-of-state students. Consistent with observed data, college quality affects the trajectory of lifecycle wage growth. Given wages and housing prices, college graduates then decide where to work. In each state, a representative firm with differentiated productivity combines human capital from high school and college graduates. A state government operates in each location, and a federal government operates across the economy as a whole. Both levels of government choose a labor tax to finance exogenous government expenditures. College qualities and admission cutoffs, wages, and house prices, are determined endogenously in equilibrium to clear the market for college seats, skilled and unskilled labor, and houses. Taxes are endogenously determined to balance the government budget.

Given this college demand curve, each state then individually optimizes domestic utilitarian welfare, taking as given the actions of all other states and the equilibrium choices of households, firms, and colleges. The solution to this problem is a Nash equilibrium in which no state has a profitable deviation.

This paper's main results focus on using the model to understand why college policies differ across states and on conducting a counterfactual exercise analyzing the optimal federal-level distribution of college policies across the United States. For these results, it is important that the model fits both the current level of cross-state migration and the elasticity of cross-state migration with respect to key policy variables, as well as the optimal college policy choices of state governments.

First, this includes a rich set of cross-sectional moments on conditional migration and education choices, which ensure that the model accurately predicts which types of agents study and work in which locations. A second key elasticity for policy choices and human capital accumulation is the change in future earnings with respect to college expenditures per student. I use two instruments to causally estimate this elasticity. The first is a shift-share style instrument based on shocks to state government budgets and lagged state shares of college financing, and the second is exogenous changes to tuition freezes or caps. I provide a novel estimate of this moment and find that a 1% increase in college expenditures increases wages ten years later by 0.18%. I then calibrate the model to match.

The model is also externally validated against two additional sets of key moments. First, in addition to matching the current distribution of migration flows, the model performs well in predicting changes in migration as college policies change. It replicates the elasticities of in-state college attendance and in-state work with respect to local wages, college tuition, and college spending per student. Second, the model's Nash equilibrium performs well in replicating the main motivating facts and correlations established in the data, as well as levels and heterogeneity in cross-state college policies. This second external validation exercise is crucial for using the model as a lens to understand why college policies differ drastically across states and to answer the main positive research question.

Using the calibrated model, I show that the empirical correlations between state size and college policies are primarily driven by supply-side factors. Expanding capacity in California is more costly than in North Dakota, leading to higher tuition and lower out-of-state capacity. By contrast, the correlation between state wages and college policies is primarily determined by demand-side factors. High-wage states experience stronger demand for their labor markets and, correspondingly, for their colleges. As tuition rises, the elasticity of residual college demand is lower for colleges in states with strong labor markets. Hence, these states optimally set higher in-/out-of-state tuition.

The negative correlation between out-of-state capacity and state wages is driven by migration flows and the marginal return to an out-of-state student. Consider the high-wage state of California versus the low-wage state of North Dakota. Graduates of California's colleges are more likely to remain and work in-state, increasing demand for California's colleges among out-of-state students. Moreover, California attracts a larger inflow of college graduates from other states, regardless of its policies toward out-of-state students. Given this strong external demand for its colleges and labor markets, California optimally devotes tax revenues to in-state students. By contrast, North Dakota attracts out-of-state students by setting tuition low and capacity high, with some of these students remaining in-state after graduation.

To answer this paper's normative research question, I perform an optimal policy analysis at the federal level, where a counterfactual policymaker simultaneously chooses all policies to maximize national welfare. For a fixed aggregate level of tax revenues allocated to colleges, optimally setting college policies in a centralized system increases aggregate welfare by 3.3% relative to the decentralized system. The federal solution improves welfare for each state individually, indicating substantial welfare losses from strategic interaction in the state-level Nash optimum.

There are two main efficiency reasons why the federal solution differs from the state solution. First, the strategic interactions (and monopoly power) present in the decentralized equilibrium are no longer present under the federal solution. Second, the federal government internalizes the entire spatial distribution of college qualities and firm productivities relative to population, agent abilities, and other state primitives.

These differences translate into several important dissimilarities between the state and federal policy solutions. First, the federal government eliminates the premium on out-of-state tuition and removes capacity constraints specific to out-of-state students. While it is natural for a state-level policymaker to subsidize in-state students by charging high premiums and rationing availability to out-of-state students, this a clear inefficiency from the perspective of the federal policymaker, which seeks to promote migration and ensure that the highest-ability students are matched with the highest-quality colleges, and most productive labor markets. Second, while the decentralized college system features substantial variance across states, heterogeneity actually increases further under the federal solution. Third, all motivating correlations between population, wages, and college policies reverse sign.

Together, the second and third differences imply that in the centralized system, colleges in large and high-wage states become larger, less competitive, cheaper, and lower quality, while colleges in small and low-wage states become smaller, more competitive, costlier, and higher quality. This policy reflects both the efficiency and redistribution motives of the federal policymaker. It is efficient to expand capacity and lower tuition barriers to high-quality colleges with highly productive labor markets, in states such as California. This also serves as redistribution: rising college capacity increases the supply of college-educated workers and pushes down skill prices. Moreover, spending per student and average ability decline at colleges in states like California, lowering college quality and further reducing earnings. The opposite occurs in states like North Dakota, leading to income convergence across locations.

There are three main takeaways from the federal solution that are relevant to policy setting: (1) elimination of the out-of-state tuition premium, (2) an increase in the percentage of out-of-state students, and (3) an expansion of college capacity in high-productivity states. I apply these dynamics to policy design and find that implementing spatial policies is significantly more effective than standard proposals by the literature or public officials. Relative to the benchmark equilibrium, a federal policy that prohibits states from setting an out-of-state tuition premium or reserving college seats for in-state students increases aggregate welfare by 0.94%, nearly one-third of the increase under the fully optimal federal policy. The welfare gain from eliminating cross-state tuition and capacity frictions is significantly larger than a national free-tuition policy. Targeting capacity expansions to high-productivity colleges, rather than to all states, raises welfare by 1.5 times more.

Given positive welfare gains from offering free tuition or expanding capacity, a standard model might suggest that these are advisable policies. However, these welfare gains are far from optimal. Without solving the full federal government optimality problem within the quantitative spatial model of higher education developed in this paper, evaluating these alternative policies would not be possible.

The remainder of this paper is structured as follows. First, I review the relevant literature in *Section 1.1*. In *Section 2*, I contrast higher education systems and establish a large degree of heterogeneity across the United States. I then empirically motivate correlations between state characteristics and college policies. Given the empirical facts established, the quantitative theory is developed in *Section 3*. *Section 4* calibrates the quantitative framework. In *Section 5*, I solve for optimal-state level policies in a decentralized college system. I show that the solution rationalizes key empirical facts. *Section 6* conducts a welfare analysis for a counterfactual centralized education system and contrasts the solution of the federal and state-level optimal policy problems. *Section 7* applies what is learned from the optimal federal solution to a spatial policy analysis. *Section 8* concludes.

1.1 Related Literature

I bridge several strands of the quantitative literature that have so far remained separate. In doing so, I provide a novel theory of the U.S. higher education sector and its relation to firm activity and worker migration decisions across space.

Structural Models of Higher Education – The modeling of the higher education sector relates to a relatively small literature that develops structural models generating an endogenous distribution of university quality. Those most similar to my paper are Hendricks et al. (2021) and Epple et al. (2017). However, both papers study substantially different questions: the former investigates the rise in college attendance following World War II, while the latter examines the interaction between financial aid at public and private colleges. More closely related to the present question are Brotherhood et al. (2023) and Wright and Zheng (2025), who, as in my work, study the reallocation of college seats across different levels of college quality. However, these papers primarily focus on income-based affirmative action policies and their effects on intergenerational mobility.

The primary methodological difference between my paper and this strand of literature is that I model a competitive university sector embedded within a general equilibrium spatial framework, featuring a production sector and housing markets. This allows me to address a broader set of macroeconomic questions.

The model developed here also relates to two contemporaneous papers that both study the evolution of skill-biased technical change and cross-state wage inequality since the 1980s, in a spatial setting: Lee and Verma (2025) and Capelle et al. (2025). This paper differs in its research question and richer modeling of higher education and lifecycle human capital accumulation. In particular, I model endogenous and differentiated college qualities, as well as endogenous in-/out-of-state tuition, in-/out-of-state capacity, and spending per student, along with the dynamic interaction between college quality and human capital accumulation over the lifecycle.

Higher Education Financing Across Regions – An extensive empirical literature studies the effects of state-level differences in college financing and tuition levels and is reviewed in Kane (2006). On the quantitative side, this work relates to several papers that examine differences in college tuition subsidies, financing, and quality, however, at the cross-country level rather than the cross-state level. Herrington (2015), Holter (2015), and Wright (2023) find that a significant portion of cross-country differences can be explained by their respective systems of higher education.

Optimal Policy and Higher Education – This paper also relates to a broader quantitative literature that studies optimal education taxation and subsidy policies. For example, Krueger and Ludwig (2016) and Abbott et al. (2019) compute optimal tuition subsidies financed through progressive labor taxes. While this paper also examines optimal tuition subsidization, I do so within a spatial framework, focusing primarily on evaluating the efficiency implications of geographic variation in college tuition, quality, and seat availability.

Migration and Human Capital Accumulation – This paper is closely related to an industrial organization (structural micro) literature studying migration and human capital accumulation in response to spatial variation in college costs and availability. Fu et al. (2022) studies the welfare consequences of "education deserts" and shows that student college demand is elastic with respect to the proximity of colleges. Ishimaru (2023) investigates the link between birth location, migration, and economic outcomes.

Most similar to this paper in both research question and approach to modeling migration is Kennan (2015), who examines how tuition differences across states explain state-level variation in the proportion of college graduates in the local labor force. He finds that tuition differences have large effects on college enrollment that are not eliminated through migration.

I differ from these works along several lines. First, I model the macroeconomy in a general equilibrium setting. Second, I consider an endogenous college quality distribution across locations, in addition to differences in college financing and availability. Given the macroeconomic structure of my model, I am able to undertake an aggregate welfare analysis and evaluate optimal government policy.

Quantitative Spatial Economics – Finally, this paper relates to a large, quantitative spatial literature studying place-based policies, as reviewed in Redding and Rossi-Hansberg (2017). Spatial models and optimal policy typically focus on firms or local infrastructure investments, as in Kline and Moretti (2014) and Fajgelbaum and Gaubert (2020). While my model incorporates spatial firm heterogeneity, it primarily focuses on the higher education sector when evaluating place-based and optimal government policies. My model also differs from this line of work in that most papers in this literature depend on agglomeration economies as the main source of inefficiency. In contrast, my framework does not rely on agglomeration externalities for any of the results.

An important avenue for future work, which could be pursued through the lens of my model, is to consider firm-side policies in the presence of an endogenous college sector and a spatial distribution of skilled and unskilled human capital.

2 The Geography of Colleges and Government Policy

This section documents several empirical facts to motivate the model and subsequent quantitative analysis. First, I describe the large degree of heterogeneity in state college systems across the United States. I then present motivating relationships between state characteristics and education policies. The quantitative spatial model developed in *Section 3* provides an explicit theory of these empirical observations. This allows me to better analyze heterogeneity across states and evaluate the aggregate efficiency of alternative policy regimes.

For all data analysis, I focus on public, four-year degree-granting institutions with at least 250 students enrolled. All data are for the 2017–2018 academic year. This amounts to 572 colleges across 50 U.S. states (the District of Columbia is excluded). I choose this time period in order to avoid large changes to the higher education sector during the COVID-19 pandemic. Details of the empirical calculations and data sources are in *Appendix A.1* and further results exploring colleges and higher education policy across states are reported in *Appendix A.2*.

2.1 Cross-State Heterogeneity

The state government controls three main aspects of the public college system: (1) financing (2) in-/out-of-state tuition, and (3) in-/out-of-state capacity. In this section, I document the significant cross-state heterogeneity along these three dimensions.[1,2]

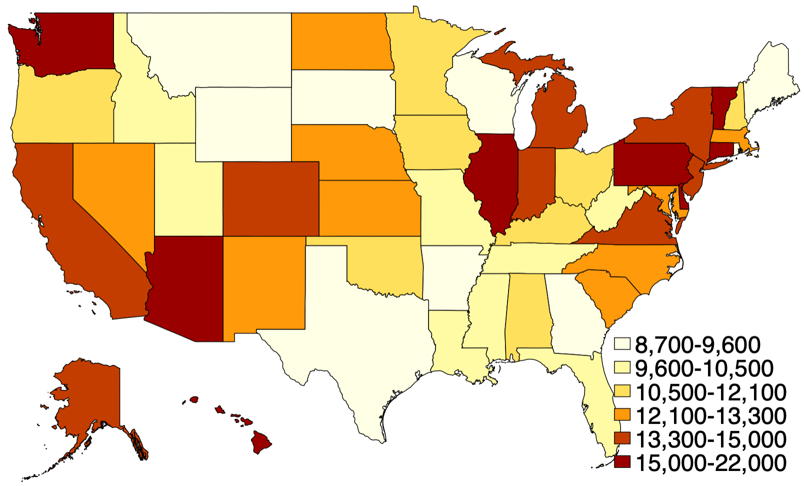
Financing – State financing of colleges vary along several dimensions. First, the amount spent per student. *Figure 1* plots instructional expenditures per full-time equivalent (FTE) student across states. There is substantial variation across the United States, with Pennsylvania spending nearly twice as much as California and three times as much as Texas.

It is not immediately clear which states "value" education most based on their spending patterns. *Figure E.11* shows that Texas spends a small amount per student but allocates a large portion of total government expenditures to education, whereas California spends a large amount per student but allocates a smaller portion of its total expenditures to education. On the other hand, *Figure E.12* shows that Florida spends a relatively small amount

[1]While it is true that variation within states can also be substantial, I focus on cross-state variation here.

[2]In practice, the public college budgetary process involves a complicated set of negotiations with the state government and college board. This bureaucratic process also differs significantly across states. However, in the majority of states, the state government has the legal authority to set tuition, capacity, and college budget, which ultimately determines spending per students. See Parmley et al. (2009) for an in-depth discussion of higher education budget practices across the United States.

Figure 1: **Instructional Spend per FTE Across States**



Notes: This figure plots total instructional spending per full-time equivalent student across the 50 U.S. states. Amounts are reported in 2017 real dollars. Data are taken from the Integrated Postsecondary Education Data System (IPEDS) for the 2017-2018 academic year.

per student, yet more than half of the college budget is financed through state tax revenues.[3] This implies that other forms of revenue, such as tuition or donations, are much more important to college budgets in some states than in others.

Tuition – In-state and out-of-state tuition both vary significantly across states. *Figure 2* plots per-semester, in-state, sticker tuition prices. In-state tuition in Pennsylvania is more than double that of neighboring New York. All states also set out-of-state tuition levels higher than in-state tuition.[4] *Figure 3* plots the ratio of out-of-state to in-state tuition. The figure shows similarly large disparities across states but with a pattern entirely different from that in *Figure 2*. At one end of the spectrum, Minnesota and Illinois charge roughly 75% more to out-of-state students, whereas North Carolina and Florida charge over 400% more.

Capacity – I present several summary statistics showing capacity varies widely across states. First, in-state admission rates (in-state admission offers over in-state applications) range from below 50% in California to above 95% in Iowa. *Figure 4* plots total admission rates.[5]

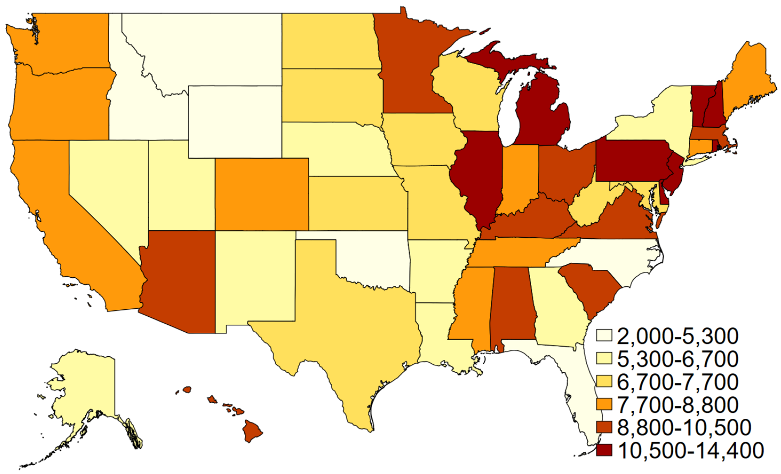
A more direct measure of immediately available public college seats is total college enrollment per high school student. While this measure does not account for demand or the

[3]Federal appropriations represent a small portion of college budgets in all states. The vast majority of federal college funding comes in the form of Title IV aid.

[4]Reciprocity agreements exist for certain regional groupings of states (e.g., Minnesota and Wisconsin). I will account for these arrangements when taking the quantitative model to the data in *Section 4*.

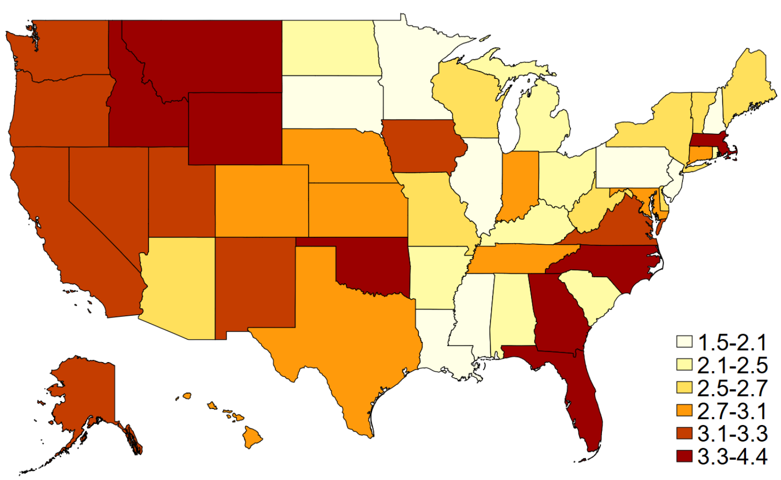
[5]Due to data limitations, I am unable to plot in-state admission rates for all public colleges and states.

Figure 2: **In-state Sticker Tuition**



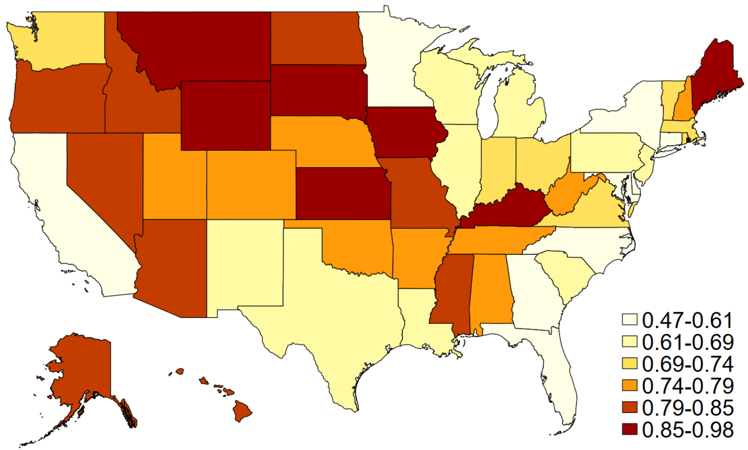
Notes: This figure plots in-state sticker tuition across the 50 U.S. states. Amounts are reported in 2017 real dollars. Data are taken from the Integrated Postsecondary Education Data System (IPEDS) for the 2017-2018 academic year.

Figure 3: **Out-of-state to In-state Tuition Ratio**



Notes: This figure plots the ratio of out-of-state to in-state sticker tuition across the 50 U.S. states. Data are taken from the Integrated Postsecondary Education Data System (IPEDS) for the 2017-2018 academic year.

Figure 4: **Admission Rate**



Notes: This figure plots the admission rate across the 50 U.S. states. Data are taken from the Integrated Postsecondary Education Data System (IPEDS) for the 2017-2018 academic year.

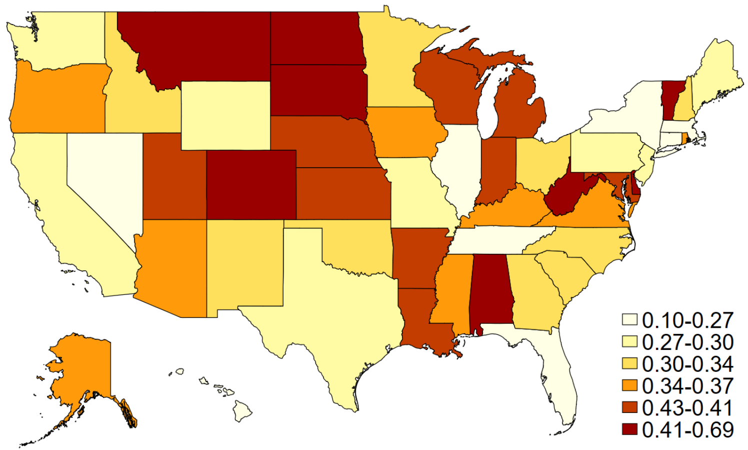
availability of private options it serves as a good proxy. Given differences in tuition levels, a college will strictly prefer out-of-state students. Hence, the state government must set a cap on out-of-state students or a minimum on in-state students. Assuming these caps bind, in-state capacity can be inferred as in-state enrollment to high school graduates.

Figure 5 plots this statistic and shows that, states with low admission rates also tend to have a relatively low number of available college seats per high school student. Again, California and New York are among the lowest in the country, with fewer than 0.3 spots per high school student. At the other end of the spectrum are the Dakotas, with nearly 0.7 spots per high school student.

While these trends could be driven by the relative availability of public versus private colleges, *Figure E.1* in *Appendix E* suggests this is not the case. Although there is variation across states, public college enrollment as a percentage of total enrollment is high on average at 79.1%. I discuss the role of private colleges further when introducing the quantitative model in *Section 3*.

Finally, states also vary significantly in terms of out-of-state capacity, which I represent as the percentage of the student body which is out-of-state. *Figure E.13* shows that for states like New York, California, or Texas, out-of-state students represent less than 10% of the student body. On the other hand, in states like North Dakota or Vermont nearly 70% of students originate from out-of-state.

Figure 5: **College Enrollment per High School Student**



Notes: This figure plots the number of full-time equivalent college students per high school student across the 50 U.S. states. Data are taken from the Integrated Postsecondary Education Data System (IPEDS) for the 2017-2018 academic year.

2.2 Motivating Facts

Given this significant heterogeneity, it is natural to ask: "Why are college systems so different across states?" While I ultimately address this question using a quantitative theory, I first present several relationships observed in the data to motivate this quantitative analysis and an understanding of these cross-state differences.

I find that the two quantitatively important correlates of state higher education policy are population size and college graduate wages. *Tables 1* and *2* show the ordinary least squares (OLS) relationships between these variables and the main college policy parameters set by state governments: (1) out-of-state tuition, (2) in-state tuition, (3) out-of-state capacity, and (4) spending per student.

Regression results in columns (1) and (2) show that larger states charge higher in-state and out-of-state tuition, on average. Similarly, states with higher college wages charge higher tuition for both in- and out-of-state students. Column (3) from both tables shows that wealthier and more populous states enroll relatively fewer out-of-state students. Finally, column (4) shows there is no clear relationship between spending per student and state size. However, there is a strong positive correlation between spending per student and average college wages in a state, as shown in column (4) of *Table 2*.[6]

[6]While the correlations are only intended to serve as motivation for the quantitative analysis, I perform robustness analysis to ensure the correct model design is motivated. In particular, I control for the racial composition, democratic vote share, private college percentage, and the percentage of students studying at a land grant college. These results are reported in *Tables E.1* and *E.2* and all correlations are quite similar to those reported here.

Table 1: **Motivating OLS – State Policies and Population**

	(1)	(2)	(3)	(4)
	$\ln(T^{out})$	$\ln(T^{in})$	ln(% in-state)	ln(spend per)
ln(*population*)	0.101	0.150	0.130	-0.018
	(0.028)	(0.020)	(0.000)	(0.599)
constant	8.39	6.68	-2.34	11.06
	(0.000)	(0.000)	(0.000)	(0.000)
R^2	0.11	0.12	0.48	0.01
Observations	50	50	50	50

Notes: This table reports the correlation between state population and college policies. P-values shown in brackets. Data are taken from the Integrated Postsecondary Education Data System (IPEDS) for the 2017-2018 academic year and the American Community Survey (ACS).

A large portion of the variation in out-of-state capacity across states is explained by population size, with column (3) reporting an R^2 of 0.48. For instance, states like California educate very few out-of-state students, whereas states like Vermont educate significantly more out-of-state than in-state students. While state size explains little of the variation in spending per student, a significant portion is explained by state college wages, with an R^2 of 0.32.

I present two additional statistics to provide further insight into the potential causes underlying the observed correlations. First, in column (1) of *Table 3*, I regress the two key correlates discussed above, college earnings and population, on the percentage of a state's college alumni who leave the state to work elsewhere. The regression also includes controls for the unemployment rate and the percentage of students who are in-state. Column (2) conducts the same regression using the net inflow of college graduates as the dependent variable. On average, larger and higher wage states experience lower outflows of their own alumni and higher net inflows of college graduates.

Given this, *Figure 6* shows that expenditures per student, as presented in *Figure 1*, do not tell the full story. A more relevant measure is the cost of net college graduates, where cost is defined to be total state expenditures per net college graduates. That is, the cost of producing a college graduate who remains in-state or of "importing" a graduate who earned a degree elsewhere. *Figure 6* plots this cost. For example, Michigan spends a large amount per college student but relatively little to produce a net college graduate, whereas West Virginia spends less per student but considerably more to produce a net graduate.

Table 2: **Motivating OLS – State Policy and College Wages**

	(1) ln(T^{out})	(2) ln(T^{in})	(3) ln(% in-state)	(4) ln(spend per)
ln(*college wages*)	0.661	0.891	0.706	1.441
	(0.041)	(0.106)	(0.008)	(0.000)
constant	2.54	-0.97	-8.10	-4.80
	(0.473)	(0.872)	(0.007)	(0.160)
R^2	0.09	0.06	0.14	0.32
Observations	50	50	50	50

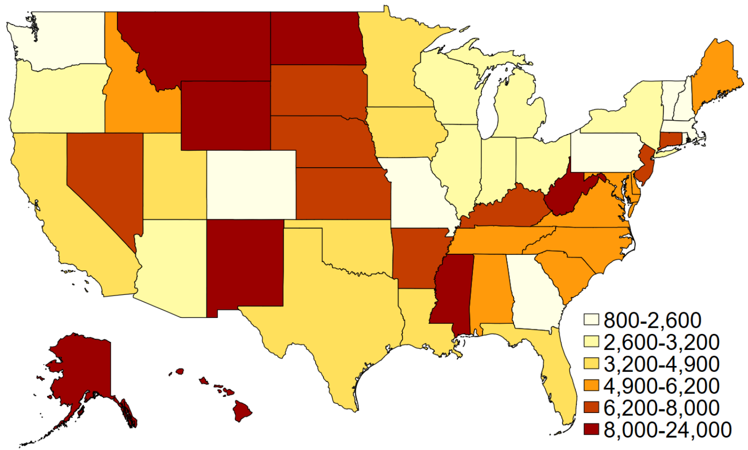
Notes: This table reports the correlation between state college wages and college policies. P-values shown in brackets. Data are taken from the Integrated Postsecondary Education Data System (IPEDS) for the 2017-2018 academic year and the American Community Survey (ACS).

Table 3: **Alumni Migration Patterns and State Characteristics**

	(1) ln(% alum leave)	(2) ln(net import)
ln(*college earn*)	-0.735	2.631
	(0.069)	(0.056)
ln(*population*)	-0.134	0.277
	(0.014)	(0.004)
ln(*unemployment*)	0.160	-0.315
	(0.353)	(0.0376)
ln($\%in - state$)	-3.36	
	(0.000)	
constant	8.14	
	(0.048)	
R^2	0.92	0.56
Observations	50	50

Notes: This table reports the regression of college characteristics on the percentage of a state's college alumni that leave the state for work, and the net import of college graduates. P-values shown in brackets. Data are taken from the Integrated Postsecondary Education Data System (IPEDS) for the 2017-2018 academic year, the American Community Survey (ACS), and Conzelmann et al. (2023).

Figure 6: **Cost of Retained and Imported College Graduates**



Notes: This figure plots state expenditures per graduate student retained (i.e., remaining in-state after graduation) and imported (i.e., moving to the state after graduation). Data are from open access sources provided by Conzelmann et al. (2023).

2.3 Discussion

Taking stock, I have presented two sets of motivating observations from the data. First, there is substantial heterogeneity across states in their systems of higher education. In particular, states differ significantly in (1) in-/out-of-state tuition, (2) in-/out-of-state capacity, and (3) spending per student. The second set of empirical observations shows that larger and higher-wage states tend to charge higher tuition, admit fewer out-of-state students, and spend more per student. These states also experience lower alumni outflows and larger inflows of college graduates from other states.

3 Model

The objective of the model developed in this section is to analyze optimal policies at the federal level (normative research question) and to understand why state governments choose their respective college policies (positive research question). The model features two main sections, which together serve as the benchmark U.S. economy. The first is a rich college demand curve, where households are the key decision-makers taking college policies as given. The second is an optimal state-level college policy problem over in-/out-of-state tuition, spending per student, and capacity, taking college demand as given.

3.1 College Demand

3.1.1 Overview

Time is discrete and has an infinite horizon. One model period represents five biological years and is denoted by j. The economy is populated by measure one of J overlapping generations with a uniform demographic structure.

The economy contains a discrete number of locations indexed by $\ell \in \mathfrak{L}$. Each location is populated with an initial mass $\mathcal{M}_\ell$ of high school graduates with heterogeneous family backgrounds, ability, and human capital. Agents first choose whether to attend the local college, attend college elsewhere, or immediately enter the workforce. Agents then choose which location to work in. In order to move across locations, an individual must pay a moving cost. Agents consume and purchase housing services, which are in fixed supply for each location. While working over their lifecycle, individuals are endowed with one unit of time which they supply to the labor markets inelastically and accumulate human capital on-the-job via learning by doing.

A college of differentiated quality operates in each location. The college sets a separate admissions policy for in- and out-of-state students that specifies the ability needed for attendance. In each location, a representative firm aggregates human capital from high school or college graduates in order to produce consumption goods. A state government operates in each location and a federal government operates across the economy as a whole. Both governments tax labor market earnings proportionally, in order to finance exogenous government expenditures.

In this model, the primitives and sources of ex-ante heterogeneity that define a state, are: (1) location, (2) population, (3) available land, (4) firm productivity for college and non-college labor, (5) college quality productivity, and (6) state and federal appropriations.

3.1.2 Agents

In this section, I describe an agent's lifecycle. I first lay out preferences and the static optimization problem. I then describe the lifecycle optimization problem. Lastly, I present the sequence of discrete choices over education and migration.

Static Problem – An agent who chose education type e, working location ℓ_w, and has a current human capital level h, solves the following static optimization problem to allocate labor market earnings over consumption c, and housing d,

$$
\begin{aligned}
\mathcal{U}(e, \ell_w, h) &= \max_{c,d} \left\{ \ln \left(\left(\frac{c}{1-\xi} \right)^{1-\xi} \left(\frac{d}{\xi} \right)^{\xi} \right) \right\} \\
\text{subject to,} & \qquad (1) \\
c + p(\ell_w) \cdot d &= \big(1 - \tau(\ell_w)\big) \cdot \omega(e, \ell_w) \cdot h
\end{aligned}
$$

where, $p(\ell_w)$ is the price of housing in location ℓ_w, $\tau(\ell_w)$ is the total tax liability in location ℓ_w, and $\omega(e, \ell_w)$ is the wage paid to an individual of education type e in location ℓ_w. Given ln utility, this problem can be analytically solved, with the solution given by,

$$
\mathcal{U}(e, \ell_w, h) = \ln \Big(\big(1 - \tau(\ell_w)\big) \cdot \omega(e, \ell_w) \cdot h \Big) - \xi \ln \Big(p(\ell_w) \Big) \qquad (2)
$$

The first term gives the natural logarithm of disposable income. I denote disposable income as $\Omega(e, \ell_w, h)$. The second term is the natural logarithm of the rental rate on houses multiplied by its income share ξ. Together, this gives the utility of an agent for period j, with the constant term being omitted.

Lifecycle Problem – Each agent inelastically supplies one unit of time each period, and accumulates human capital via learning-by-doing. Given the solution to the problem defined by equation (1), agents receive lifetime utility according to,

$$
\begin{aligned}
&\max \sum_{j=0}^{J} \beta^j \cdot \mathcal{U}(e, \ell_w, h_j) \\
&\text{subject to,} \qquad (3) \\
&h_{j+1} = \big(a \cdot Q(\ell_c)\big) h_j^{\nu} + (1 - \delta) h_j
\end{aligned}
$$

Clearly this is a trivial maximum problem as static optimality holds each period, and with inelastic labor supply, human capital stock is given. β is the time discounting rate, and $\delta \in [0, 1]$ is the period depreciation rate of human capital. Let the value of this problem for an agent who did not attend and attended college be denoted as, $W^{nc}(a, h, \ell_w)$ and $W^{c}(a, h, \ell_c, \ell_w)$, respectively. College quality Q_ℓ from location ℓ_c and ability a, determine the efficiency with which human capital is produced. $Q(\ell_c)$ is normalized to one for those who do not attend college.

A reduced form literature finds that (controlling for observable characteristics) earnings levels and growth rates vary significantly across college qualities.[7] Motivated by this finding, I employ this model structure where college alters the growth of human capital over the agent's lifecycle. That is, for a monetary and opportunity cost, the agent may purchase a more productive technology for producing human capital. This modeling structure and its empirical relevance will be discussed further in *Section 4* when I take the model to the data. It is important to note that I do not impose $Q(\ell_c) > 1$ and instead allow for the possibility that college quality plays no role in lifetime earnings dynamics.

Discrete Choices – All discrete choices in the model are subject to idiosyncratic preference shocks. These shocks are distributed according to the Type I Extreme Value distribution with shape parameter equal to zero and scale parameter of one. Given this, all discrete choice problems can be solved analytically following McFadden (1972).[8]

At age $j = 1$ an agent's initial state is characterized by ability a, human capital h, parental resources p, and birth location ℓ_b. The initial endowments are drawn from the distribution $F(a, h, p)$, which is constant across locations. The probability of being born in a given location is $Pr(\ell_b)$.

Ability a affects how much an individual learns and benefits from college, as well as the efficiency with which they accumulate human capital over the lifecycle. Human capital h determines labor market earnings. Parental resources p provide financing for consumption and tuition if the agent chooses to attend college.

Given the initial state, agents make an irrevocable decision to attend college or work as a high school graduate. This determines their education type e_c for college and e_{nc} for non-college. That is, high school graduates first solve,

$$V(a, h, \ell_b, p) = \mathbb{E}_{\epsilon_e}\Big[\max_{c,nc}\Big\{V^c(a, h, \ell_b, p) + \sigma_e\epsilon_e, V^{nc}(a, h, \ell_b) + \sigma_e\epsilon_e\Big\}\Big] \tag{4}$$

where, $V^c(a, h, \ell_b, p)$ is the optimal continuation value of attending college, and $V^{nc}(a, h, \ell_b)$ is the optimal continuation value of entering the workforce immediately as a high school graduate. ϵ_e is drawn from the Type I Extreme Value distribution and the parameter σ_e controls the dispersion of these shocks. The solution to this problem is given by the choice probabilities and corresponding value function,

[7] See for example, Chetty et al. (2020), Cellini and Turner (2019), or Kirkeboen et al. (2016).

[8] Throughout I omit the Euler-Mascheroni constant in the solution to each discrete choice problem.

$$Pr[e = \bar{e}|a, h, \ell_b, p] = \frac{\exp\left(\frac{V^{\bar{e}}(a,h,\ell_b,p)}{\sigma_e}\right)}{\sum_{\hat{e}} \exp\left(\frac{V^{\hat{e}}(a,h,\ell_b,p)}{\sigma_e}\right)} \tag{5}$$

$$V(a, h, \ell_b, p) = \sigma_e \ln\left(\sum_{\hat{e}} \exp\left(\frac{V^{\hat{e}}(a, h, \ell_b, p)}{\sigma_e}\right)\right) \tag{6}$$

For those who choose to attend college, they spend one period studying and graduate with certainty. While in college they cannot work and must consume using parental resources p. The value of attending college is given by,

$$Z(a, h, \ell_b, p, \ell_c) = \ln\left(p - T^i(\ell_c) - T^o(\ell_c)\mathbb{1}_{\ell_b \neq \ell_c}\right) - \delta(\ell_b, \ell_c) + \beta V^{cw}(a, h, \ell_c) \tag{7}$$

where, ℓ_c is the location of college attendance, $T^i(\ell_c)$ is in-state tuition, and $T^o(\ell_c)$ is the additional tuition premium paid by out-of-state students.[9] Moving costs from the location of birth to the location of college attendance are given by the function $\delta(\ell_b, \ell_c)$. The optimal continuation value of working with a college degree is given by $V^{cw}(a, h, \ell_c)$. The agent chooses in which location to attend college by solving,

$$V^c(a, h, \ell_b, p) = \mathbb{E}_{\epsilon_{cq}}\left[\max_{\ell_c : a \geq \bar{z}_l} \left\{Z(a, h, \ell_b, p, \ell_c) + \sigma_{cq}\epsilon_{cq}\right\}\right] \tag{8}$$

where $\bar{z}_\ell$ is the admission cutoff to college in location ℓ, which will be discussed further in *Section 3.1.3*. This gives choice probabilities and value function,

$$Pr[\ell_c = \bar{\ell}_c|a, h, \ell_b, p] = \frac{\exp\left(\frac{Z(a,h,\ell_b,p,\bar{\ell}_c)}{\sigma_{cq}}\right)}{\sum_{\hat{\ell}_c} \exp\left(\frac{Z(a,h,\ell_b,p,\hat{\ell}_c)}{\sigma_{cq}}\right)} \tag{9}$$

$$V^c(a, h, \ell_b, p) = \sigma_{cq} \ln\left(\sum_{\hat{\ell}_c} \exp\left(\frac{Z(a, h, \ell_b, p, \hat{\ell}_c)}{\sigma_{cq}}\right)\right) \tag{10}$$

Given the solution to equation (3), choice problems over work locations for college and high school graduates are given by,

$$V^{cw}(a, h, \ell_c) = \mathbb{E}_{\epsilon_{cw}}\left[\max_{\ell_w} \left\{W^c(a, h, \ell_c, \ell_w) - \delta(\ell_c, \ell_w) + \sigma_{cw}\epsilon_{cw}\right\}\right] \tag{11}$$

$$V^{nc}(a, h, \ell_b) = \mathbb{E}_{\epsilon_{ncw}}\left[\max_{\ell_w} \left\{W^{nc}(a, h, \ell_w) - \delta(\ell_b, \ell_w) + \sigma_{ncw}\epsilon_{ncw}\right\}\right] \tag{12}$$

[9] Let $Z(a, h, \ell_b, p, \ell_c) = -\infty$ for any consumption values less than or equal to zero.

The resulting choice probabilities and value functions are,

$$Pr[\ell_w = \bar{\ell}_w | a, h, \ell_c] = \frac{\exp\left(\frac{W^c(a,h,\ell_c,\bar{\ell}_w) - \delta(\ell_c,\bar{\ell}_w)}{\sigma_{cw}}\right)}{\sum_{\hat{\ell}_w} \exp\left(\frac{W^c(a,h,\ell_c,\hat{\ell}_w) - \delta(\ell_c,\hat{\ell}_w)}{\sigma_{cw}}\right)} \tag{13}$$

$$V^{cw}(a, h, \ell_c) = \sigma_{cw} \ln\left(\sum_{\hat{\ell}_w} \exp\left(\frac{W^c(a, h, \ell_c, \hat{\ell}_w) - \delta(\ell_c, \hat{\ell}_w)}{\sigma_{cw}}\right)\right) \tag{14}$$

$$Pr[\ell_w = \bar{\ell}_w | a, h, \ell_b] = \frac{\exp\left(\frac{W^{nc}(a,h,\bar{\ell}_w) - \delta(\ell_b,\bar{\ell}_w)}{\sigma_{ncw}}\right)}{\sum_{\hat{\ell}_w} \exp\left(\frac{W^{nc}(a,h,\hat{\ell}_w) - \delta(\ell_b,\hat{\ell}_w)}{\sigma_{ncw}}\right)} \tag{15}$$

$$V^{nc}(a, h, \ell_b) = \sigma_{ncw} \ln\left(\sum_{\hat{\ell}_w} \exp\left(\frac{W^{nc}(a, h, \hat{\ell}_w) - \delta(\ell_b, \hat{\ell}_w)}{\sigma_{ncw}}\right)\right) \tag{16}$$

Given the initial distribution and choices probabilities, one can derive the mass of individuals working in the economy. The mass of individuals of education type e_{nc} and e_c, ability a, human capital h, education location ℓ_c, and working location ℓ_w is,

$$\mu(e_{nc}, a, h, \ell_w) = \sum_{\ell_b} \int_p Pr(\ell_b) \cdot F(a, h, p) \cdot Pr[e = nc | a, h, \ell_b, p] \cdot Pr[\ell_w = \bar{\ell}_w | a, h, \ell_b] \, dp \tag{17}$$

$$\begin{aligned} \mu(e_c, a, h, \ell_c, \ell_w) = &\sum_{\ell_b} \int_p Pr(\ell_b) \cdot F(a, h, p) \cdot Pr[e = c | a, h, \ell_b, p] \\ &\cdot Pr[\ell_c = \bar{\ell}_c | a, h, \ell_b, p] \cdot Pr[\ell_w = \bar{\ell}_w | a, h, \ell_c] \, dp \end{aligned} \tag{18}$$

In a similar manner, I will use $\mu(\cdot)$ to denote the mass of individuals with some set of characteristics throughout this paper, even when the mass is not given by equations (18) or (18). When it does not cause confusion, and with a slight abuse of notation, I will use $\mu(e, a, h, \ell_c, \ell_w)$ to represent the mass of either college or non-college educated workers, with a redundant state variable $\ell_c = \ell_w$ for non-college workers. Finally, since moves are not permitted across the lifecycle and human capital accumulates deterministically, I suppress j as a state in the mass of agents.

3.1.3 Colleges

Each location has one public college.[10] The quality Q_ℓ of each college is given by,

$$Q_\ell = \bar{q}_\ell \big(m_\ell\big)^\theta \big(\bar{a}_\ell\big)^{1-\theta} \tag{19}$$

where, $\bar{a}_\ell$ represents the average ability of the student body and is commonly referred to as "peer effects". Spending per student is given by m_ℓ. θ controls the elasticity of college quality with respect to spending per student. Lastly, $\bar{q}_\ell$ is a fixed college quality term representing intangible capital such as reputation or job placement networks.

Colleges set an admissions criterion that specifies a minimum student ability for acceptance as $\bar{z}$. The objective function of a college is lexicographic. Its first priority is to maximize enrollment until it hits capacity. Once capacity is met, the college's objective is to maximize quality, which implies setting the highest value of $\bar{z}$ that maintains full enrollment. Given separate capacity for in and out-of-state students, colleges will set separate admission cutoffs $\bar{z}^i_\ell$ and $\bar{z}^o_\ell$.

The remaining objects are taken as given by the college, and the optimal policy problem generating these policies is discussed in *Section 3.2*. Net tuition for in-state and out-of-state students is fixed at T^i_ℓ and $T^i_\ell + T^o_\ell$, respectively. Likewise, capacity is set by the policymaker at $\mathcal{C}^i_\ell$ and $\mathcal{C}^o_\ell$. Finally, spending per student, m_ℓ, is also set by the state-level policymaker. Note, I have defined T^o_ℓ to be the premium on out-of-state tuition in addition to in-state tuition. However, for simplicity and with abuse of notation, I may refer to T^o_ℓ simply as out-of-state tuition.

3.1.4 Firms

Production is performed in each location by a representative firm that combines labor inputs from workers with a college degree (c) and those without a college degree (nc). The production technology is given by the CES function,

$$Y_\ell = \Big(\bar{A}(e_{nc}, \ell) Y(e_{nc}, \ell)^\eta + A(e_c, \ell) Y(e_c, \ell)^\eta\Big)^{\frac{1}{\eta}} \tag{20}$$

[10] I abstract here from private colleges. The majority of students attending private colleges are out-of-state. As seen in *Figure E.3* average out-of-state tuition exhibits significant overlap with private tuition. Hence, one can also interpret migrating out-of-state for college as attending a private college. Moreover, 79% of students attend a public college, and hence I capture the large majority of college bound students with this assumption.

The elasticity of substitution between the two types of labor is given by $\frac{1}{1-\eta}$. The exogenous skill-augmenting productivity for non-college labor is given by, $\bar{A}(e_{nc}, \ell)$. The skill-augmenting productivity for college labor is a function of college quality in location ℓ and a fixed component,

$$A(e_c, \ell) = \bar{A}(e_c, \ell)\big(Q_\ell\big)^\lambda \tag{21}$$

Here, λ governs the strength of spillovers from the state college to firms. The benchmark model will set $\lambda = 0$, and $\lambda > 0$ is considered as an extension.

Labor is aggregated according to,

$$Y(e_{nc}, \ell_w) = \sum_{j=1}^{J} \iint \mu(e_{nc}, a, h, \ell_w) h(e_{nc}, j, a, \ell_w)\, da\, dh \tag{22}$$

$$Y(e_c, \ell_w) = \sum_{j=2}^{J} \sum_{\ell_c} \iint \mu(e_c, a, h, \ell_c, \ell_w) h(e_c, j, a, \ell_c, \ell_w)\, da\, dh \tag{23}$$

3.1.5 Governments

The economy contains two types of governments. Each location features a state government and a federal government operates at the economy-wide level. State and federal governments must finance exogenous expenditures S_ℓ and F using proportional labor taxes. In equilibrium, state and federal governments set taxes τ^s_ℓ and τ^f, respectively, such that budgets are balanced,

$$\tau^s_\ell \Big(\sum_{\ell_c} \sum_{e} \int_a \int_h \mu(e, a, h, \ell_c, \ell_w = \ell) \big(\omega(e, \ell_w) h\big)\, da\, dh \Big) = S_\ell \tag{24}$$

$$\tau^f \Big(\sum_{\ell_w} \sum_{\ell_c} \sum_{e} \int_a \int_h \mu(e, a, h, \ell_c, \ell_w) \big(\omega(e, \ell_w) h\big)\, da\, dh \Big) = F \tag{25}$$

At the federal level, government expenditures consist of transfers to state governments. The federal government provides a fraction $\pi_\ell \in [0, 1]$ of F to each locality, such that $\sum_\ell \pi_\ell = 1$. At the state level, government expenditures consist of three components related to colleges: (1) capacity, $\mathcal{C}^i_\ell$ and $\mathcal{C}^o_\ell$; (2) subsidized tuition, T^i_ℓ and T^o_ℓ; and (3) spending per student, m_ℓ. Here, these policies are taken as given. I will discuss the optimal choice over these policies in *Section 3.2*.

3.1.6 Housing

Location ℓ_w has a fixed supply of housing $D(\ell_w)$. It is assumed that housing is owned by an outside landlord. Housing prices are derived from the solution to the household's static optimality problem, equation (2). That is, total spending on housing services is equal to the preference for housing services ξ, times total disposable income in a location,

$$p(\ell_w)D(\ell_w) = \xi \sum_e \sum_{\ell_c} \int_a \int_h \mu(e,a,h,\ell_c,\ell_w)\Omega\big(e,\ell_w,h\big)\, da\, dh \tag{26}$$

3.1.7 College Demand Equilibrium

Given model parameters, primitives, and college policies $\{T^i_\ell, T^o_\ell, m_\ell, C^i_\ell, C^o_\ell\}_{\ell\in\mathfrak{L}}$, a college demand equilibrium consists of prices and allocations,

- College qualities: $\{Q_\ell\}_\ell$
- Admissions cutoffs: $\{\bar{z}^i_\ell, \bar{z}^o_\ell\}_\ell$
- Wages: $\{\omega(e,\ell)\}_{e,\ell}$
- Firm labor demands: $\{Y(e,\ell)\}_{e,\ell}$
- Firm productivities: $\{A(e_c,\ell)\}_\ell$
- House prices: $\{p_\ell\}_\ell$
- State government tax rates: $\{\tau^s_\ell\}_\ell$
- Federal government tax rate: τ^f
- Value functions: $V(a,h,\ell_b,p), V^c(a,h,\ell_b,p), V^{nc}(a,h,\ell_b), V^{cw}(a,h,\ell_c),$ $W^{nc}(a,h,\ell_w), W^c(a,h,\ell_c,\ell_w)$
- Choice probabilities: $Pr[e=\bar{e}|a,h,\ell_b,p], Pr[\ell_c=\bar{\ell}_c|a,h,\ell_b,p],$ $Pr[\ell_w=\bar{\ell}_w|a,h,\ell_c], Pr[\ell_w=\bar{\ell}_w|a,h,\ell_b]$

such that,

1. Given college qualities, admissions cutoffs, wages, and house prices, the value functions, and choice probabilities are consistent with agent optimization.
2. Labor demands, wages, and outputs are consistent with firm optimization problem.
3. Given choice probabilities, college qualities are consistent with college optimization.
4. State and federal government budgets balance.
5. $\forall \ell \in \mathfrak{L}$, markets clear,
    (a) Consumption goods markets.
    (b) Labor markets.
    (c) Admissions markets.
    (d) Housing markets.

3.2 State Government Problem

In this section I lay out the state government problem for optimal college policies, taking as given, the college demand equilibrium as defined in *Section 3.1.7*. First note, the cost of capacity is given by the function,

$$K\big(\mathcal{C}_\ell, m_\ell, \ell\big) = \bar{f} + \mathcal{V}(\mathcal{C}_\ell, \ell) + m_\ell \mathcal{C}_\ell \tag{27}$$

where $\bar{f}$ represents fixed costs, and $\mathcal{V}(\mathcal{C}_\ell, \ell)$ denotes variable costs. Variable costs depend on a college's location through land/housing prices. Fixed and variables costs are independent of college quality and I will refer to them as "custodial costs". The last term represents the sum of spending per student m_ℓ.

The state governments play a one-shot simultaneous move Nash game, which is defined in *Section 3.2.1*. That is, each state solves its optimal policy problem taking as given the choices of all other states.[11] A state government l, chooses in-state capacity $\mathcal{C}_l^i$, out-of-state capacity $\mathcal{C}_l^o$, spending per student m_l, in-state tuition T_l^i, and out-of-state tuition T_l^o, in order to maximize aggregate utilitarian welfare of agents born in-state l, by solving,

[11]The formulation in *Section 3.1.3* where colleges choose admission cutoffs given state government policies is without loss of generality and is implicit from the state governments decisions here.

$$
\max_{\{\mathcal{C}_l^i,\mathcal{C}_l^o,m_l,T_l^i,T_l^o\}} \int_a \int_h \int_p \mu(a,h,p,\ell_b = l) V(a,h,p,\ell_b = l)\, da\, dh\, dp
$$

subject to, (28)

$$
K\big(\mathcal{C}_l^i + \mathcal{C}_l^o, m_l, l\big) = T_l^i \int_a \int_h \int_p \mu(e_c, a, h, \ell_b = l, p, \ell_c = l)\, da\, dh\, dp
$$

$$
+ \big(T_l^i + T_l^o\big) \sum_{\ell_b \neq l} \int_a \int_h \int_p \mu(e_c, a, h, \ell_b, p, \ell_c = l)\, da\, dh\, dp + S_l + \pi_l \cdot F
$$

$$
Q_l = \bar{q}_l \big(m_l\big)^{\theta} \big(\bar{a}_l\big)^{1-\theta}
$$

$$
\int_a \int_h \int_p \mu(e_c, a, h, \ell_b = l, p, \ell_c = l)\, da\, dh\, dp \leq \mathcal{C}_l^i
$$

$$
\sum_{\ell_b \neq l} \int_a \int_h \int_p \mu(e_c, a, h, \ell_b, p, \ell_c = l)\, da\, dh\, dp \leq \mathcal{C}_l^o
$$

+ college demand equilibrium conditions

Total state expenditures are cost of capacity and spending per student $K(\mathcal{C}_l, m_l, l)$. Revenues consist of in-state and out-of-state tuition, a fixed state spending S_l, which is generated by equilibrium labor taxes τ_l^s, and some fraction π_l of fixed federal expenditures F.

The second constraint is the college quality production function and the third and fourth constraints are the capacity constraints for in- and out-of-state students, respectively. The government's problem must also be consistent with all conditions which characterize a competitive equilibrium, as defined in *Section 3.1.7*. That is, this problem can be thought of as a static Ramsey problem, and is not a planner problem of choosing allocations.

An alternative to maximizing welfare of those born in-state l, is to maximize welfare of those living in-state l. I have chosen the current specification as this alternative produces counterfactual choices by the state government wherein they set $T_l^i + T_l^o < T_l^i$ in order to attract high-ability out-of-state applicants. In principle, this could also be the case with the chosen specification given sufficiently strong peer-effects; however I find it is not observed for the calibrated model.

3.2.1 Equilibrium

For each state $\ell \in \mathfrak{L}$, a strategy is the vector of college policies $s_\ell = (\mathcal{C}^i_\ell, \mathcal{C}^o_\ell, m_\ell, T^i_\ell, T^o_\ell)$ where, $s_\ell \in \mathcal{S}_\ell \equiv \mathbb{R}^3_+ \times \mathbb{R}^2$. The joint strategy space is defined as $\mathcal{S} \equiv \prod_{\ell \in \mathfrak{L}} \mathcal{S}_\ell$. Assume the existence of a unique equilibrium for the college demand curve (as defined and discussed in *Section 3.1.7*) for each strategy s. Then, for any $s \in \mathcal{S}$, let $x(s)$ denote this residual-demand equilibrium induced by s. The payoff for state ℓ is the utilitarian welfare of agents born in ℓ, evaluated at $x(s)$, and denoted as $W_\ell(s)$. Then the simultaneous-move normal-form game is defined as,

$$\mathcal{G} \;=\; \big\langle\, \mathfrak{L},\;\, (\mathcal{S}_\ell)_{\ell \in \mathfrak{L}},\;\, (W_\ell)_{\ell \in \mathfrak{L}} \,\big\rangle$$

Given all other policies $s_{-\ell}$, the best-response correspondence of state ℓ is given by,

$$\mathrm{BR}_\ell(s_{-\ell}) \equiv \arg\max_{s_\ell \in \mathcal{S}_\ell} \; W_\ell(s_\ell, s_{-\ell})$$

subject to,

at $x(s_\ell, s_{-\ell})$, state budget constraint and capacity constraints defined in problem (28) hold.

Then, a policy profile $\hat{s} \in \mathcal{S}$ is a (pure-strategy) Nash equilibrium if and only if,

$$\forall\, \ell \in \mathfrak{L}, \quad \hat{s}_\ell \in \mathrm{BR}_\ell\big(\hat{s}_{-\ell}\big).$$

That is, evaluated at the residual-demand equilibrium $x(\hat{s})$, no state ℓ can increase its objective $W_\ell(s_\ell, s_{-\ell})$ by a one-shot unilateral and feasible deviation in $(\mathcal{C}^i_\ell, \mathcal{C}^o_\ell, m_\ell, T^i_\ell, T^o_\ell)$.

3.3 Equilibrium Selection

Given the presence of peer effects in the college demand equilibrium (defined in *Section 3.1.7*) multiple equilibria may arise. Generally speaking, quantitative spatial models often guarantee a unique equilibrium by requiring that congestion forces are "stronger" than agglomeration forces. I do not include firm-level productivity agglomeration forces in the benchmark model, and the only explicit source of agglomeration is the peer effects. Standard uniqueness proofs do not apply in the current setting with a dynamic lifecycle problem. However, it can be shown, for a simplified version of the lifecycle problem, that for the calibrated equilibrium, the congestive forces of house prices dominate the agglomeration forces of peer effects, and a unique equilibrium exists.

For the equilibrium to the Nash game, (defined in *Section 3.2.1*), I computationally, show that a stable and unique equilibrium exists in the parameter space within a neighborhood of the calibrated equilibrium. Moreover, when optimizing, I consider a large number of other candidate points to ensure that the correct unique equilibrium was found. *Section C* contains extensive details of the computational algorithm.

Finally, equilibrium uniqueness in the counterfactual exercise in *Section 6.1* where a federal government simultaneously chooses all 250 college policies, follows from the discussion of the college demand equilibrium.

4 Calibration

The calibration involves choosing parameters so that the model replicates a steady-state equilibrium of the U.S. economy. I calibrate the model to the 50 U.S. states for the 2015-2019 period. The model is calibrated to match household college decisions by high school graduates across locations and parental income levels, as well as subsequent migration and earnings dynamics by college attendance and ability. Ensuring that the "correct" agents move to the "correct" locations is key to optimal policy setting in the counterfactual exercises. Importantly, for the household and policymaker dynamics, I also ensure that the model matches the elasticity of earnings with respect to college expenditures. After calibration, I externally validate the model using out-of-sample predictions for key elasticities central to the policymakers' optimal problem. I also externally validate the solution to the decentralized Nash optimum, comparing the motivating correlations between the model and the data, as well as levels and variance in college policies across states.

Table 4 summarizes model parameters that are externally calibrated and *Table 5* summarizes the internally calibrated parameters and parameters derived from model inversion. Each section below describes a block of the economy in which I discuss both externally and internally calibrated parameters. Details are discussed throughout this section and additional information is reported in *Appendix B*. All moments and parameters estimated for each individual state are omitted here and reported in *Appendix E*.

The calibration follows a two-step procedure. The first step is conducted in partial equilibrium, with skill prices, housing rental rates, and taxes taken from the data. Agent preference parameters, the initial distribution, and college parameters are chosen to match choice probabilities and human capital accumulation dynamics. In the second step, I recover the remaining firm, and housing parameters by inverting structural equations from the model.

Table 4: **Externally Calibrated Parameters**

Parameter	Value	Description	Source
Agents			
β	0.98	Time discounting factor	Annual interest rate, 0.02
J	8	Number of model periods	5-year periods, age 18-58
ξ	0.33	Housing share of total expenditures	BLS CEX
$\mu(\ell_b)$	–	Mass of high school students in state ℓ	ACS
μ_h	1.0	Human capital distribution, mean	Normalization
σ_h	0.56	Human capital distribution, std.	ACS
ν	0.50	Human capital production	Huggett et al. (2011)
δ	0.02	Human capital depreciation	Huggett et al. (2011)
Colleges			
$\bar{f}$	2.67×10^7	College custodial fixed costs	IPEDS and NHGIS
κ_1	11,817	College custodial FTE linear term	IPEDS and NHGIS
κ_2	0.219	College custodial FTE quadratic term	IPEDS and NHGIS
κ_3	12.37	College custodial FTE×rent interaction term	IPEDS and NHGIS
Firms			
η	0.31	College to non-college CES elasticity	Katz and Murphy (1992)
Government			
S_ℓ	–	State college expenditures in state ℓ	IPEDS
$\pi_\ell F$	–	Federal college expenditures in state ℓ	IPEDS

Notes: This table reports model parameters directly observed and calculated from the data. Values that are not calculated by state are reported here. Values calculated for each state individually are reported in *Appendix* E.

Table 5: **Internally Calibrated Parameters**

Parameter	Value	Description	Source
Preferences			
γ_1^e	0.0012	Moving cost, education	Choice probabilities
γ_2^e	-2.1179	Home bias, education	Choice probabilities
γ_1^w	0.0004	Moving cost, working	Choice probabilities
γ_2^w	9.5459	Home bias, working	Choice probabilities
γ_3^w	4.9769	Unskilled labor, working	Choice probabilities
σ_e	0.1502	Dispersion, education	Choice probabilities
σ_{cq}	0.8910	Dispersion, college location	Choice probabilities
σ_{cw}	1.7335	Dispersion, college work	Choice probabilities
σ_{ncw}	2.5625	Dispersion, non-college work	Choice probabilities
Distributions			
μ_a	0.3176	Ability, mean	Average earnings profile
σ_a	0.5751	Ability, std.	Average earnings variance
μ_p	0.9341	Parent resources, mean	College by parent income
σ_p	0.4387	Parent resources, std.	College by parent income
Colleges			
θ	0.6112	Elasticity of earnings w.r.t. spending	IPEDS, Grapevine, and Deming and Walters (2017)
$\bar{q}_\ell$	–	College fixed quality college	Earnings profiles by quality, Chetty et al. (2020)
Firms			
$\bar{A}(e, \ell)$	–	Firm fixed productivity	Firm prod. function, skill prices.
Housing			
H_ℓ	–	Housing stock	Housing market clearing, rent price

Notes: This table reports internally calibrated model parameters, their value, a brief description, and the moment target. Values that are not calculated by state are reported here. Values calculated for each state individually are reported in *Appendix E*. Preference parameters, initial distribution parameters, and the elasticity of earnings with respect to college spending per student are estimated by GMM. The remaining sets of parameters are estimated by model inversion.

4.1 Agents

The set of parameters here are the moving costs and dispersion parameters. Agent preference parameters are estimated so that the choice probabilities over education, college location, and work location in the model are closest to those in the data. Given that the sum of probabilities equals one, the log-odds ratio is used for estimation. From equations (5), (9), (13), and (15), the odds ratios can be derived as follows,

$$\ln\left(\frac{Pr[e=c|a,h,\ell_b,p]}{Pr[e=nc|a,h,\ell_b,p]}\right)=\frac{1}{\sigma_e}\Bigg(V^c(a,h,\ell_b,p)-V^{nc}(a,h,\ell_b)\Bigg) \tag{29}$$

$$\begin{aligned}&\ln\left(\frac{Pr[\ell_c=\bar{\ell}_c|a,h,\ell_b,p]}{Pr[\ell_c=1|a,h,\ell_b,p]}\right)\\ &=\frac{1}{\sigma_{cq}}\Bigg(\ln\left(p-T^i(\ell_c)-T^o(\ell_c)\mathbb{1}_{\ell_b\neq\ell_c}\right)-\delta(\ell_b,\ell_c)+\beta V^{cw}(a,h,\ell_c)\\ &\quad-\Big(\ln\left(p-T^i(1)-T^o(1)\mathbb{1}_{\ell_b\neq 1}\right)-\delta(\ell_b,1)+\beta V^{cw}(a,h,1)\Big)\Bigg)\end{aligned} \tag{30}$$

$$\begin{aligned}&\ln\left(\frac{Pr[\ell_w=\bar{\ell}_w|a,h,\ell_c]}{Pr[\ell_w=1|a,h,\ell_c]}\right)\\ &=\frac{1}{\sigma_{cw}}\Bigg(W^c(a,h,\ell_c,\bar{\ell}_w)-\delta(\ell_c,\bar{\ell}_w)-\Big(W^c(a,h,\ell_c,1)-\delta(\ell_c,1)\Big)\Bigg)\end{aligned} \tag{31}$$

$$\begin{aligned}&\ln\left(\frac{Pr[\ell_w=\bar{\ell}_w|a,h,\ell_b]}{Pr[\ell_w=1|a,h,\ell_b]}\right)\\ &=\frac{1}{\sigma_{ncw}}\Bigg(W^{nc}(a,h,\bar{\ell}_w)-\delta(\ell_b,\bar{\ell}_w)-\Big(W^{nc}(a,h,1)-\delta(\ell_b,1)\Big)\Bigg)\end{aligned} \tag{32}$$

Typically, quantitative spatial models feature location- (and possibly education-) specific amenity values that enter the preferences of agents, allowing the model to achieve an exact fit to the data. In the quantitative model developed in the previous section, I omit location amenity values, and so the model is over-identified. Matching the choice probabilities instead works entirely through earnings (human capital accumulation), initial distributions, moving costs, and preference shocks.

The data identifying choice probabilities originate from several different sources. The education decision, equation (5), is taken from the geocoded National Longitudinal Survey of Youth 1997 (NLSY97). Ability is defined using Armed Services Vocational Aptitude Battery (ASVAB) scores. Human capital is defined as in Wright and Wurdinger (2025), using the GPA-weighted stock of high school credits.

The choice of university location, equation (9), combines data from the NLSY97 and IPEDS. Due to sample size limitations, I cannot directly estimate this choice probability using NLSY97 data. Instead, I first calculate transition probabilities from state of birth to state of college degree using IPEDS. Next, I assign each college a quality in the NLSY97 using average later-life earnings of those attending the given college. This is the natural definition which closely aligns with model college quality. I group colleges into quartiles of quality and obtain transition probabilities to these college qualities, contingent on ability, human capital, and parental resources. Finally, I note that for the full state-by-state transition matrix obtained using IPEDS data, any difference in transition probabilities for states with the same quality college must be due to moving costs or idiosyncratic preference shocks. Hence, I take differences within quality quartiles and add them to the transition probabilities obtained from the NLSY97.

For the transition probabilities from college location to working location for each state, I use the Baccalaureate and Beyond (B&B) 2016/2020 survey. I then use the same procedure in the NLSY97 as just described, but group states based on the skill price. Finally, for the transition probabilities from birth location to working location for non–college educated workers, I use the American Community Survey (ACS) and again use the NLSY97, with groupings based on state skill prices. I discuss calculation of skill prices using the Panel Study of Income Dynamics (PSID) in *Section 4.3*.

4.1.1 Demographics

The demographic structure of agents is uniform across the lifecycle. It would be a simple extension of the model to allow for a different mass of agents at different ages. However, demographic variation is not central to the main research question addressed in this paper. The initial mass of agents in each state is given by the high school population, taken from the 2015-2019 5-year ACS sample.

4.1.2 Preferences

The preference parameters affecting household decisions over college and working locations are the preference shocks for each discrete decision: σ_e, σ_{cq}, σ_{cw}, and σ_{ncw}, and the moving costs. Moving costs for the education migration decision are parameterized as,

$$\delta^c(\ell_b, \ell_c) = \gamma_1^e \cdot d(\ell_b, \ell_c) + \gamma_2^e \cdot \mathbb{1}(\ell_b \neq \ell_c) \tag{33}$$

where $d(\cdot, \cdot)$ is the Euclidean metric measuring the distance between birth location and education location. γ_1^e captures moving costs, and γ_2^e captures home bias, or the extra utility received by an individual studying in their home state. The moving costs associated with the work migration decision are given by,

$$\delta^w(\ell, \ell_w, e) = \gamma_1^w \cdot d(\ell, \ell_w) + \gamma_2^w \cdot \mathbb{1}(\ell \neq \ell_w) + \gamma_3^w \cdot \mathbb{1}(e = nc) \tag{34}$$

where $\ell = \ell_b$ for a non–college educated worker ($e = nc$), and $\ell = \ell_c$ for a college educated worker ($e = c$). γ_3^w captures a set of unobservable factors affecting the mobility of non-college educated workers, given the well established observation that college educated workers have a much higher degree of geographic mobility.

Birth and work locations are defined as the population-weighted centroid of a given state, and education locations as the student weighted centroid of the state. The former is calculated using population-weighted centroids for each block group from the U.S. Census. The latter is calculated using geographical coordinates provided for each college in IPEDS.

4.1.3 Initial Distributions

There are six parameters governing the variance and mean of initial distributions for ability, human capital, and parental resources: σ_a, μ_a, σ_h, μ_h, σ_p, and μ_p. The mean of the initial distribution of human capital is normalized to unity and the variance taken from initial earnings when graduating high school and entering the labor force. The initial distribution of ability is internally calibrated to match average earnings growth rates, and average earnings variance over the lifecycle. Finally, the initial distribution of parental resources targets college attendance rates by parental income quintiles.

4.1.4 Human Capital Production

Period depreciation and the curvature of human capital production are taken to be standard values from the literature, with $\delta = 0.02$ and $\nu = 0.50$. I discuss the estimation of college quality Q and college production function below.

4.2 Colleges

For colleges, I focus on public, primarily four-year, degree-granting institutions with at least 250 students enrolled. All IPEDS data are for the 2017-2018 academic year. This amounts to 572 colleges across the 50 U.S. states. I choose this time period in order to avoid large changes to the higher education sector during the COVID-19 pandemic. Data details are reported in *Appendix A*.

4.2.1 Tuition

I construct a measure of net tuition price for in-state and out-of-state students. IPEDS includes an estimate of net price for each college; however, it is not reported separately for in-state and out-of-state students. Instead, I use gross sticker tuition and subtract average federal and state tuition grants per student. At the federal level, this includes Pell Grants, which I assume are awarded uniformly to both in-state and out-of-state students. Federal Title IV funding is not included in the college/state budget constraint; however, it is included in the calculation of net tuition price. That is, while Title IV funds affect the net tuition price, these funds are not readily available for colleges to spend on other purposes, as is the case with direct federal funding. I account for several tuition reciprocity schemes which are discussed in detail in *Appendix A*.

4.2.2 Spending per Student

For spending per student, I follow Epple et al. (2006) and consider m_ℓ to consist of "quality-enhancing" expenditures. Inputs to expenditures per student include instructional spending in the benchmark calibration. I perform robustness analysis using broader or differing definitions, which are reported in *Appendix B.1.2*, and show that no main results in the paper are sensitive to this definition.

4.2.3 Capacity

For capacity, I use the total count of full-time equivalent students enrolled, disaggregated by in-state and out-of-state status. I do not count international students as out-of-state and instead drop them from the analysis. International student enrollment has risen dramatically over the last 20 years, however, they still represent a small share of the total student body, at less than 4% on average.

4.2.4 Cost Function

From equation (27) that the state government's budget constraint can be summarized as,

$$\bar{f} + \mathcal{V}(\mathcal{C}_\ell, \ell) + m_\ell \mathcal{C}_\ell = T_\ell + S_\ell + \pi_\ell \cdot F \tag{35}$$

where m_ℓ is spending per student, as discussed in the previous section. "Custodial costs" are given by a fixed cost $\bar{f}$ and a variable cost function $\mathcal{V}(\mathcal{C}_\ell, \ell)$. Revenues are then a function of total tuition revenue T_ℓ and total state and federal appropriations.[12] Given observable data, the budget constraint then provides a direct expression for custodial costs, which I denote as K^c_ℓ.[13] Custodial costs are estimated following a similar approach to Epple et al. (2006). Assume they can be parameterized as,

$$K^c_{i,\ell} = \bar{f} + \kappa_1 \mathcal{C}_{i,\ell} + \kappa_2 \mathcal{C}^2_{i,\ell} + \kappa_3 (p_\ell \cdot \mathcal{C}_{i,\ell}) + \epsilon_{i,\ell} \tag{36}$$

The identification assumption necessary to estimate each parameter using ordinary least squares (OLS) is that measurement error in custodial costs, $\epsilon_{i,\ell}$ (at college i in state ℓ), is uncorrelated with college size. The industrial organization literature takes a different view and interprets $\epsilon_{i,\ell}$ as an unobserved shock to the cost function (Berry, 1994). Under this interpretation, it is likely that $\mathbb{E}[\epsilon_{i,\ell} \mid \mathcal{C}_{i,\ell}] \neq 0$. To overcome this issue, I employ an instrumental variable (IV) strategy, where college size is instrumented with endowment level.

College capacity (size), $\mathcal{C}_{i,\ell}$, is taken to be the total full-time equivalent (FTE) undergraduate enrollment. *Section 4.4* estimates p_ℓ as the average rent price across each state, as is common in this class of models. While rent prices are the relevant measure for scaling student residential buildings, land prices would be a better measure for other types of physical expansion. However, rent prices are highly correlated with land prices and serve as a reasonable proxy in equation (36) to capture cross-state differences in costs. It is through these rental rates that the cost function depends on the location. However, I do not estimate a separate cost function for each state and assume that there are no unobserved cost differences across states.

Results from estimating equation (36) are reported in *Table 6*. Under all specifications, the cost function is convex in the number of students. Column (1) reports the estimation

[12] While colleges do receive revenues from gifts, endowment income, or other ventures, an investigation of heterogeneity across states along these categories is beyond the scope of the present paper. Moreover, *Figure E.7* in *Appendix E* shows that tuition and government revenues represent the vast majority of the college budget constraint across all states.

[13] In the benchmark calibrated U.S. economy, this equation is simply an accounting identity. It becomes a relevant constraint in the context of problem (28) when the state government chooses an optimal level of tuition, spending per student, and capacity.

Table 6: **Custodial Cost Function Estimation**

	(1) Custodial Costs	(2) Custodial Costs	(3) Custodial Costs
κ_1	10,084	8,582	11,817
	(0.000)	(0.000)	(0.271)
κ_2	0.314	0.311	0.219
	(0.000)	(0.000)	(0.210)
κ_3		1.34	12.37
		(0.264)	(0.000)
$\bar{f}$	1.44×10^7	1.52×10^7	2.67×10^7
	(0.236)	(0.212)	(0.520)
Observations	543	543	505
IV	No	No	Yes

Notes: This table reports the results of the custodial cost function estimation. Column (3) instruments for full-time equivalent students using endowment levels. P-values are reported in brackets. Data are taken from IPEDS for college variables and from NHGIS for local rent prices.

omitting the interaction between college size and rent prices. Column (2) adds this interaction. Finally, Column (3) reports the results of the IV regression.

Throughout, fixed costs (the regression constant) are large but not precisely estimated. The linear and quadratic terms, κ_1 and κ_2 on college size are stable across specifications but lose power under the IV regression. Finally, the effect of rent prices, κ_3 on college costs is large and an important predictor of custodial costs.

4.2.5 Production

College quality fixed component $\bar{q}$ is chosen to match earnings growth profiles over the ages 23-35, with data taken from Chetty et al. (2020). I then choose $\bar{q}$ to match this profile for each college.

The remaining parameter to calibrate for colleges is θ, the elasticity of college quality with respect to spending per student. θ is internally calibrated to match the elasticity of median graduate earnings with respect to college spending per student. θ is calibrated after the the first-step of the procedure, so that the counterfactual economy when increasing spending per student matches changes in graduate earnings. This introduces a fixed point in the calibration process, I first guess θ, and then update after the first-step of the calibration procedure is completed. I then re-calibrate internal parameters, and update θ until convergence is achieved. See *Appendix C* for details of the formal procedure.

In the data, the estimate of interest for college i at time t is,

$$y_{i,t} = \beta \ln \left(spend\ per_{i,t} \right) + \gamma X_{i,t} + \delta_t + \lambda_i + \epsilon_{i,t} \tag{37}$$

where $y_{i,t}$ is ln-median income of cohort t, eight to ten years post graduation, $X_{i,t}$ is a vector of controls, δ_t are time fixed effects, and λ_i are college fixed effects. I use IPEDS data for each variable, with earnings data available for 1997-2008. To be consistent with the college cost function estimation, I use total instructional spending per FTE as spending per student.

Estimating equation (37) through OLS would likely result in a biased estimate of the coefficient of interest, β, due to unobserved individual-level heterogeneity in ability. Instead, I employ an instrumental variables (IV) strategy similar to Deming and Walters (2017) and instrument for ln-spending per student, $\ln \left(spend\ per_{i,t} \right)$.

I consider two instruments. The first is a shift-share (i.e., Bartik, 1991) style instrument based on shocks to state budgets. The idea is that state budget cuts made "across the board" will have a larger impact on colleges where state appropriations represent a greater share of the total budget.[14] I construct the instrument with initial 1997 shares of total college revenues accounted for by state appropriations, using IPEDS data. I then interact this with yearly state appropriations allocated to higher education, per college-aged (19-23) person. Yearly state appropriations for higher education are taken from Grapevine, and the population of college-aged individuals is taken from the Current Population Survey (CPS). That is, the instrument is given by,

$$Z^b_{i,t} = \left(\frac{approp_{i,1997}}{rev_{i,1997}} \right) \cdot \left(\frac{state\ approp_{s,t}}{pop_{s,t}} \right) \tag{38}$$

where it is implicit in the notation that state s is the state in which college i is located.

The second instrument uses variation in tuition increase caps or freezes across states. I use data compiled by Deming and Walters (2017), again for the period 1997-2008. Over this period, seventeen states enacted legislation capping or freezing tuition increases at all public colleges. I define the instrument $Z^f_{i,t}$ to equal one if the state instituted a tuition freeze or cap between periods $t-1$ and t, and zero otherwise.

The results of each instrumental variables regression are reported in *Table 7*. Column (1) reports results using the shift-share instrument, and Column (2) reports results using

[14] State budget cuts are typically (though not always) applied across all colleges. See Deming and Walters (2017) for a full discussion.

Table 7: **Elasticity of Graduate Earnings to Spending per Student**

	(1) ln(median earnings)	(2) ln(median earnings)
ln(*spend per*)	0.173	0.182
	(0.000)	(0.000)
Instrument	Bartik	Cap/Freeze
Observations	2,826	2,865
F-stat	267.1	220.3
Controls	Yes	Yes
Time Effects	Yes	Yes
Fixed Effects	Yes	Yes

Notes: This table reports the results of the IV regressions estimating the elasticity of college graduates' median earnings with respect to total instructional spending per student. Earnings are measured eight to ten years after graduation. The first column reports results using the shift-share instrument, and the second column reports results using the tuition caps instrument. Data are taken from IPEDS (1997-2008), Grapevine, the CPS, and Deming and Walters (2017).

the tuition freeze/cap instrument. Both coefficients are of similar magnitudes and are highly significant. Together, they indicate that a 1% increase in spending per student translates into an approximately 0.18% increase in median earnings of graduates.

4.3 Firms

Consider first the benchmark model with no spillovers ($\lambda = 0$). The elasticity of substitution between college- and non-college educated labor, η, is externally calibrated to $0.31 = 1 - 1/1.44$, following Katz and Murphy (1992) and Heckman et al. (1998). Firm fixed productivities, $\bar{A}(e_{nc}, \ell)$ and $\bar{A}(e_c, \ell)$, are then obtained by inverting the production function.

Defining total output in location ℓ as Y_ℓ, firm's profit maximization problem gives,

$$\begin{aligned} \ln\big(A(e,\ell)\big) = & \ln\big(\omega(e,\ell)\big) + \big(1-\eta\big)\ln\big(Y(e,\ell)\big) \\ & - \big(1-\eta\big)\ln\big(\omega(e_{nc},\ell)Y(e_{nc},\ell) + \omega(e_c,\ell)Y(e_c,\ell)\big) \end{aligned} \tag{39}$$

Given data for skill prices, $\omega(e,\ell)$, and labor supplies, $Y(e,\ell)$ (implied by calibrated choice probabilities), fixed productivities $\bar{A}(e,\ell)$ can be recovered.

Skill prices paid by firms in the model, $\omega(e,\ell_w)$, are not directly observable in the data.

Earnings observable in the data are given by $\omega(e, \ell_w) \cdot h$ in the model. However, I use yearly migration decisions of college- and non-college educated workers in the Panel Study of Income Dynamics (PSID) to infer skill prices. That is, assuming small yearly changes in human capital accumulation (which is empirically verifiable), any difference in income post-migration must be due to differences in skill prices. Hence, I regress log-differences in yearly earning, $y_{i,s,t}$, for those who migrated across states, on time effects δ_t, state fixed-effects λ_s, and a vector of demographic controls $X_{i,s,t}$,

$$y_{i,s,t} = \beta X_{i,s,t} + \delta_t + \lambda_s + \epsilon_{i,s,t} \tag{40}$$

where i is individual, s is state, and t is time. The state fixed-effects then capture difference in skill-prices.

I set California to be the omitted dummy, so all skill prices are measured relative to California. On average, relative differences across states are larger for college-educated workers than for non-college-educated workers. For non-college educated workers (conditional on age and sex), skill prices in Alabama are 62% lower than in California, whereas they are 7% higher in Washington. For college-educated workers, skill prices in West Virginia are 81% lower than in California, whereas they are only 12% lower in New York.

4.3.1 College Quality Spillovers

I extend the benchmark model to allow for a broad set of positive college externalities by calibrating $\lambda > 0$. I do so because the federal government may have incentives to spend in a given location for reasons beyond those captured in the benchmark economy.

To estimate the college spillover parameter λ, I use publicly available data from a series of papers: Andrews (2023), Russell et al. (2024), and Andrews and Russell (2025). This dataset compiles natural experiments in which college location assignment was nearly random. These papers study the effects of college openings on local attainment rates, invention, economic mobility, and inequality. I use the same methods to estimate the effect of college openings on long-run firm productivity.

Given this empirical estimate, I then run a model experiment in which a college is opened in a new location. I choose fixed productivity parameter λ to replicate the observed productivity effect in the data. I perform robustness on λ and show that the main results do not change (are in fact strengthened) with the presence of productivity spillovers.

4.4 Housing

4.4.1 Expenditure Share

The housing expenditure share, ξ, is estimated from the BLS CEX Top Means Tables. It is calculated as the ratio of total housing expenditures to average annual expenditures. An average value is calculated for the years 2015-2019. This gives $^{19{,}590}/_{59{,}526} = 33\%$.

4.4.2 Rent Prices

To calculate average rental prices for each state I follow a method similar to Eeckhout et al. (2014) and run a hedonic regression for census tract i and state s,

$$rent_{i,s} = bedrooms_{i,s} + units_{i,s} + age_{i,s} + \theta_s + \epsilon_{i,s} \tag{41}$$

where $rent_{i,s}$ is median rent, $bedrooms_{i,s}$ is the median number of bedrooms, $units_{i,s}$ is the median number of units in the structure, $age_{i,s}$ is the median age, and θ_s are state fixed effects. Using the fixed effects from this regression, average rent in a state is then assumed to be given by,

$$\overline{rent}_s = bedroom + units + age + \hat{\theta}_s \tag{42}$$

where *bedrooms*, *units*, and *age* are now national median levels. Data are taken from the 5-year ACS sample for the years 2015-2019, and is accessible through IPUMS National Historical GIS. Average rent is highest in California and lowest in South Dakota, with a ratio of 2.33 between the two states.

4.4.3 Housing Stock

Housing stocks $D(\ell_w)$ can be derived from the market clearing condition given by equation (26) and summarized as,

$$p(\ell_w)D(\ell_w) = \xi \cdot \Big(\text{Disposable Earnings}(\ell_w)\Big) \tag{43}$$

Observable data for housing rental rates and total disposable earnings implied by choice probabilities is then used to infer housing stocks by states.

4.5 Government

State and federal college expenditures are taken from IPEDS for the sample of colleges defined above in *Section 4.2*. Total expenditures are defined as the sum of university revenues from state and federal appropriations, grants, and contracts. Recall from *Section 4.2* that I allow federal Title IV aid to affect net tuition, but not state revenues. Data on disposable income are taken from the National Income and Production Accounts (NIPA) tables. Given these two data points, equations (24) and (25) then imply state and federal taxes, and the government budget constraint will balance by construction in the calibrated equilibrium.

4.6 Model Fit

Prior to conducting the counterfactual exercises and policy experiments, I verify that the model replicates key features of the U.S. economy. The model is able to match targeted moments well and predict key elasticities out of sample. I compare the solution to the optimal state policy problem and data in *Section 5.1*.

4.6.1 Targeted Moments

The internally calibrated parameters not estimated via model inversion are estimated using the generalized method of moments (GMM). These parameters include the preference parameters $[\gamma_1^e, \gamma_2^e, \gamma_1^w, \gamma_2^w, \gamma_3^w, \sigma_e, \sigma_{cq}, \sigma_{cw}, \sigma_{ncw}]$, initial distribution parameters $[\mu_a, \sigma_a, \mu_p, \sigma_p]$, college fixed productivity for each state $\bar{q}_\ell$, and the elasticity of college quality with respect to spending per student, θ. See *Appendix B* for formal details and the computational procedure.

Figure E.8 plots the model fit for each of the four discrete choices: the college versus no-college education decision, the college location choice, and the work location choices for both college and non-college graduates. The overidentified model performs well in generating accurate choice probabilities, with *Table 8* reporting an OLS slope near unity and a high R^2. The discrete choice that fits least well is the work location decision of non-college workers; however, these agents are not central to the state or federal government's decisions over college policies.

Figure E.9 compares wage growth rates by college quality in the model and the data. The model nearly exactly replicates these growth rates. *Figure E.9* also compares the qualities implied by $\bar{q}$ in the first stage of calibration to those generated in the full general equilibrium, again showing a near-exact fit.

Table 9 reports the remaining moments matched during the first stage of the calibration

Table 8: **Model Fit – Discrete Choices**

Discrete Choice	Slope	R^2
Education	1.02	0.97
College location	0.99	0.98
Work location, non-college	0.83	0.92
Work location, college	0.94	0.97

Notes: This table reports the fit of the choice probabilities between the model and data. The first column reports the slope of an OLS regression of model output on empirical data, and the second column the resulting R^2.

Table 9: **Model Fit – Targeted Moments**

Moment	**Model**	**Data**	**Source**
Average earnings growth	1.79	1.79	ACS
Average earnings std.	1.01	1.01	ACS
College attendance, parent income Q1	0.30	0.31	CPS
College attendance, parent income Q2	0.38	0.36	CPS
College attendance, parent income Q3	0.43	0.41	CPS
College attendance, parent income Q4	0.48	0.48	CPS
College attendance, parent income Q5	0.53	0.53	CPS
Earnings to college spending elasticity	0.18	0.18	See *Section 4.2.5*

Notes: This table compares the model generated moments to the data counterparts for selected targeted moments.

procedure. The model precisely matches average earnings growth rates and variance over the lifecycle. College attendance is accurately fit for parental income quintiles as well. The model exactly matches the elasticity of earnings with respect to college expenditures per student.

Figure E.10 shows that skill prices and housing rental prices obtained through model inversion exhibit a near-exact fit. State taxes are less well matched; however, this is due to the fact that no parameter directly affects taxes. I am simply comparing model-generated data to the empirical counterpart, and hence the fit is quite reasonable.

4.6.2 External Validation

I now verify several additional elasticities that are central to the state governments' optimal policy problems but were not targeted when calibrating the model. First, I compare the elasticity of in-state college attendance with respect to spending per student, in-/out-of-state tuition, and college-educated wages, both in the model and in the data. Second, I compare the elasticity of alumni leaving the state to work and the net import of college graduates with respect to college-educated wages. I externally validate the untargeted solution to the Nash optimum in the decentralized college system in *Section 5*.

To compute the in-state college attendance elasticities in the data, I estimate the following panel regression for state s, in year t, via OLS,

$$\ln(L_{s,t}) = \beta_1 \ln(T^{in}_{s,t}) + \beta_2 \ln(T^{out}_{s,t-5}) + \beta_3 \ln(M_{s,t-5}) + \beta_4 \ln(E_{s,t}) + \gamma X_{s,t} + \lambda_s + \epsilon_{s,t} \tag{44}$$

$L_{s,t}$ is the percentage of a given state's college students studying outside their home state. $T^{in}_{s,t}$ and $T^{out}_{s,t-5}$ are in- and out-of-state tuition, respectively. $M_{s,t-5}$ is total instructional spending per student, and $E_{s,t}$ is the wage of college graduates. $X_{s,t}$ is a vector of controls: admission rate, average SAT score, percentage of the student body from out of state, and the unemployment rate. Finally, λ_s is a state fixed effect. Out-of-state tuition and spending per student are lagged by five years to account for delayed student expectations over changes in college quality. The coefficients of interest, which will be compared to the model counterparts, are β_1–β_4. *Appendix B* reports the full regression results and additional specifications. *Appendix B* also includes a robustness check where the elasticities for spending per student and wages are estimated using an IV regression strategy. The coefficients are of similar magnitude across all robustness exercises.

To compute the elasticity of alumni leaving the state (β_5) and net import of college grads (β_6) with respect to college-educated wages, I estimate the following cross-sectional regression for state s, by OLS,

$$\ln(L_s) = \beta_{5/6} \ln(E_s) + \gamma X_s + \epsilon_s \tag{45}$$

L_s is the percentage of a given state's alumni who move outside of the state for work after graduation, or is the percent net import of college graduates. E_s is the wage of college graduates. X_s is a vector of controls: unemployment rate and the percentage of student body from out of state. The two coefficient of interest, which will be compared to the model counterpart, are β_5 and β_6. Due to data limitations, I am unable to estimate this regression using panel data, as was done for the first set of elasticities. This cross-sectional regression covers the time period 2011–2017. *Appendix B* reports the full regression results.

The first four elasticities $\beta_1 - \beta_4$ use within-state, cross-time variation for identification. The model is static, and so I employ the following method to most closely match these moments. I exogenously shock each state by changing in-/out-of-state tuition, high-skilled productivity, and spending per student. I shock in-/out-of-state tuition and spending per student by one percentage point, and I shock high-skilled productivity by an amount that induces a one percentage point increase in college earnings. I then average the responses across states and compare them to the data. This requires an additional 4×50 solves of the model. For this reason, these moments are used for external validation, as it would be computationally infeasible to solve the model an additional 200 times at each step of the calibration routine. For the two cross-sectional elasticities β_5 and β_6 I simply regress wages on the two outcomes variables of interest in the calibrated model.

Table 10 compares the data estimates β_1–β_6 to the model-generated moments. The model performs well with out-of-sample predictions for the effects of wages on in-state college attendance, in-state alumni employment, and net import of college graduates. In both cases, higher college wages cause fewer high school graduates and alumni to leave the state. While in-state tuition raises college expenditures (and hence quality), the increased financial burden dominates, and in both the model and the data, the percentage of high school graduates leaving the state rises. In contrast, out-of-state tuition and instructional spending more directly raise college quality, and in both cases, the percentage of students leaving the state falls. The model does quite well at predicting the elasticity for instructional spending, although it slightly overestimates the elasticity for out-of-state tuition compared to the data. Overall, these elasticities are central to the state government's optimal policy decisions, and the model's performance lends additional confidence to the results of the counterfactual exercises.

5 Decentralized College System

I begin in *Section 5.1* by externally validating the model by comparing the model Nash optimum to the data. I then provide economic intuition for the data through the lens of the optimal policy problem and address the main positive question in *Section 5.2*.

5.1 Nash Optimum versus Data

I begin by solving the Nash equilibrium for problem (28). *Table 11* reports average college policies in the Nash equilibrium compared to the data. Given that these moments are

Table 10: **Model Fit − External Validation**

Moment	**Model**	**Data**	**Description**
β_1	0.059	0.024	leaving state studying to in-state tuition
β_2	-0.13	-0.031	leaving state studying to out-of-state tuition
β_3	-0.21	-0.18	leaving state studying to spending per student
β_4	-0.17	-0.22	leaving state studying to college wages
β_5	-1.94	-1.47	leaving state working to college wages
β_6	3.67	3.31	college grad net import to college wages

Notes: This table compares the model generated moments to the data counterparts for out-of-sample moments. Data for β_1-β_4 is taken from IPEDS and the CPS for the years 1990-2017. Data for β_5 and β_6 is taken from IPEDS, the CPS, and Conzelmann et al. (2023).

untargeted and the model is well overidentified, it performs remarkably well at replicating both the mean and variance of college policies across states. The one area where the model struggles is the level of in-state tuition. There are two main reasons for this. First is the utilitarian welfare function. In general, there exists a set of Pareto weights in problem (28) that exactly match college policies. I have abstracted from that here, as a key objective of the paper is to explain cross-state differences in college policies through underlying differences in estimable state primitives that determine economic forces.[15] Second, the model features one tuition price for all income levels. If the state government optimally chose a full tuition schedule, tuition would be lower for low-income agents and higher for high-income agents, likely increasing mean tuition.[16] *Table E.6* reports the standard deviation in policy choices between the Nash optimum and the data, where the model also performs well in generating similar levels of heterogeneity.

Next, I compare the motivating correlations reported in *Tables 1* and *2* to those generated by the Nash optimum by running the same eight regressions, as reported in *Table 12*. The Nash optimum produces correlations that are quite similar to those observed in the data. Correlations with population are closer to the data in the Nash optimum than those with wages, with the latter being weaker in the model than in the data. The main reason is that the model distribution of parental college resources does not vary across states, whereas in the data, parents in California can, on average, spend significantly more on college than parents

[15] *Figure E.18* plots in-state tuition for a representative state as a higher welfare weight is placed on high-income agents. Here, optimal in-state tuition increases, given that high-income agents prefer higher-quality colleges, even at the cost of higher tuition. This has an intuitive interpretation, as states likely place higher welfare weights on high-income agents, who are typically the households with political power and who vote most often. A thorough analysis of these political economy considerations is beyond the scope of this paper.

[16] Choosing the optimal tuition schedule is computationally infeasible.

Table 11: **Mean State Policies in Nash Optimum versus Data**

	Nash Optimum	**Data**
T^{in}	\$2,669	\$7,700
T^{out}	\$27,563	\$20,780
% in-state	61.9%	72.9%
spend per	\$11,844	\$12,616
total % col	48.9%	40.6%

Notes: This table reports average college policies in the model solution compared to data counterparts.

Table 12: **Motivating OLS – Nash Optimum versus Data**

	$\ln(pop)$		$\ln(col\ wage)$	
	Nash Optimum	Data	Nash Optimum	Data
$\ln(\%\ in - state)$	0.19	0.13	0.67	0.71
$\ln(T^{out})$	0.45	0.10	0.37	0.66
$\ln(T^{in})$	0.21	0.15	0.09	0.89
$\ln(spend\ per)$	-0.01	-0.02	1.26	1.44

Notes: This table reports the motivating OLS regressions coefficients from *Section 2* compared to model counterparts.

in North Dakota. Additionally, the model performs worse at generating the correlation between in-state tuition and wages than between out-of-state tuition and wages. This is mainly because some states set in-state tuition to zero, providing less variation between state characteristics and in-state tuition in the model than in the data.

5.2 Understanding Cross-State College Policy Differences

Given that the model's optimal policy solution produces correlations quite similar to those in the data, I am able to use this optimal policy problem to provide economic intuition for the forces driving these empirical correlations. In doing so, I also answer the paper's main positive question: "Why do college policies differ substantially across states?"

In the model, all differences across states arise from variation in state primitives: (1) geographic location, (2) initial population, (3) housing stock, (4) firm productivity for college

Table 13: **Shapley Decomposition of State College Policies and Primitives**

	Shapley	% of R^2
Population	0.24	0.35
Firm Productivity	0.16	0.24
College Productivity	0.07	0.10
Housing Stock	0.06	0.09
Appropriations	0.15	0.22

Notes: This table reports the Shapley decomposition from regressing each state primitive on state policies. The contribution of appropriations and firm productivities are summed and reported combined.

and non-college labor, (5) college quality productivity, and (6) state and federal appropriations. These primitives determine both the relative cost of supplying college seats and the residual demand for college seats faced by state policymakers.

Tables E.7 and *13* present two decompositions of state policies into the contributions of each primitive. The first conducts a counterfactual exercise in which the state of Pennsylvania is "changed" to Wyoming by sequentially adding primitives of Wyoming to Pennsylvania. For example, the first row shows that when Pennsylvania's population is replaced with that of Wyoming, it explains 30.1% of the difference in out-of-state tuition (column two). However, this decomposition is computationally infeasible to perform for all combinations of primitives across all states.

The second decomposition table reports a Shapley decomposition, in which I regress state primitives on state policies and then consider all possible permutations of adding primitives. For each permutation, I record the marginal R^2 contribution from the primitive, and then average these contributions across permutations. From column one *Table 13* the decomposition shows that location, population, and firm productivities are the three largest contributors to heterogeneity in optimal college policies across states, accounting for 32%, 24%, and 16% of the differences, respectively, where location is the obtained as the residual of the decomposition.

However, simply decomposing heterogeneity with respect to primitives does not provide economic intuition as to *why* certain policies are optimal for certain types of states. Through the lens of the model, I illustrate the economics underlying these results, focusing on the three key primitives: location, population, and firm productivity. In particular, I explain the economic mechanisms underlying each of the eight correlations in *Table 12*. This answers the main positive research question: "Why do college systems differ substantially across states?"

To begin, I consider the determinants of the correlations between population size and college policies (column one of *Table 12*). These correlations are mainly driven by supply-side factors. Given a convex cost function, as population increases it becomes more costly to expand capacity. In addition, more populous states tend to have higher rents. Hence, for example, the marginal cost of an additional student in Pennsylvania is more than double that in Wyoming. This translates into relatively smaller capacity, higher tuition, and less spending per student. By contrast, higher state appropriations increase the supply of seats, lower tuition, and raise spending per student, as the state has more funds available to devote to higher education.

Turning to the correlations between wages and college policies (column two of *Table 12*), these relationships are mainly driven by demand-side factors. To understand how geography and other state primitives shape college demand, I begin at a high level by simply plotting normalized residual college demand across states in *Figure 7*, before turning to a specific discussion of how college demand maps into differing optimal policies.[17]

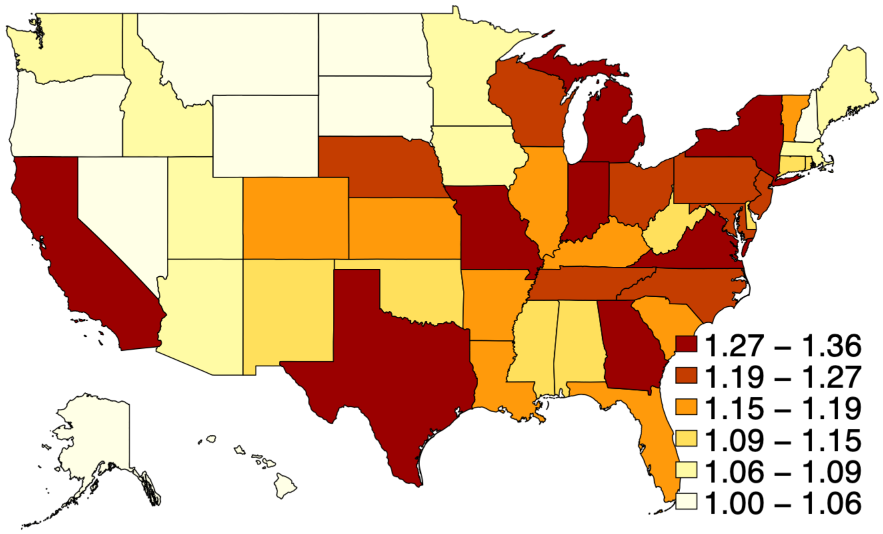
First, consider the example of Pennsylvania and Wyoming. The three largest determinants of college demand are location, population, and wages. Pennsylvania is large, high wage, and located in an urban region of the country, and consequently experiences high demand for its colleges. By contrast, Wyoming is small, low wage, rural, and faces low demand for its colleges.

Now consider population explicitly. Florida and Vermont both have mid-levels of demand, despite the latter being much less populous than the former. Vermont is near the large-population states along the eastern United States, with many students unable to attend competitive colleges in New York or Ohio. As a result, Vermont faces relatively large demand for its size. On the other hand, with the exception of Georgia, Florida is located near smaller states. Moreover, Florida charges very high tuition in both the data and the Nash optimum relative to a state like Georgia, which then attracts most of the demand in the region.

Next consider firm productivity for college labor $\bar{A}_c$, which manifests as differences in college wages across states. On average, states with higher college wages tend to exhibit higher residual demand for colleges and the unconditional correlation between $\bar{A}_c$ and residual demand is 0.46. For example, California has very high college wages and correspondingly high college demand. By contrast, Washington also has high college wages but low demand. There are two key reasons for this difference. First, Washington does not border any highly populated states, and so the pool of potential students is relatively small. Second, most

[17] To isolate differences in residual demand that stem from state primitives rather than state policies, the plot is calculated for a counterfactual economy in which each state sets in-/out-of-state tuition and capacity at average levels. Residual demand is calculated as the sum of in- and out-of-state demand.

Figure 7: **College Residual Demand (normalized) Across States**



Notes: This figure plots college residual demand across states. Residual demand is defined for mean tuition and capacities for all states. Residual demand is calculated as the sum of in- and out-of-state and normalized to one for the lowest residual demand state.

states near Washington are also near California, which offers higher wages and higher college quality, thereby drawing the bulk of residual demand away from Washington. More generally, it is state primitives *relative* to the population- and distance-weighted primitives of other states in the economy that determine college demand.[18]

Next, consider the state primitive of college quality productivity, $\bar{q}$. Take for example, Wisconsin which has a high $\bar{q}$, with its public colleges among the best in the country. Consequently, Wisconsin experiences high demand for its colleges. Arizona also has a high $\bar{q}$; however, geographically it lies between Texas and California, which dominate college demand in this region. I find that for out-of-state demand, $\bar{q}$ is relatively unimportant, as colleges tend to compete for out-of-state students primarily through tuition and capacity. By contrast, for in-state demand and the margin of attending college versus not attending, $\bar{q}$ is significantly more important.

Finally, the two smaller contributors to college demand in a given state are relative housing stocks and state appropriations. When relative housing stocks are small, rental rates are high and demand is low. For state appropriations, the effect on demand is ambiguous. As discussed above, the supply-side effect of higher appropriations is to lower tuition and raise college quality, both of which increase demand. However, higher appropriations also

[18]Higher non-college labor productivity $\bar{A}_{nc}$ is associated with lower demand on average, particularly for in-state students who substitute toward non-college work.

come at the cost of higher state taxes. On average, I find that this second effect dominates and higher appropriations slightly reduce college demand.

Now, consider the specific mechanisms determining the correlations between wages and college policies in *Table 12*. First, the positive correlation between percent in-state students and wages (column two, row one). There are two main benefits to admitting an additional out-of-state student. One, they typically face higher tuition and have higher average ability, increasing college quality for domestic students. Two, many students remain in-state to work after graduating, which lowers taxes. I show that migration flows are the main driver of explaining this negative correlation between wages and relative out-of-state capacity.

Figure 8 plots the model relationship between college skill prices and the percentage of college workers originating from out of state. As skill prices rise, more workers migrate to the state after graduating from colleges elsewhere. Moreover, *Figure E.31* shows that higher skill-price states experience fewer college graduates leaving to work elsewhere after graduation. Combined, this implies that in a high-wage state such as Pennsylvania, many college graduates enter and few leave. Hence, Pennsylvania can set out-of-state capacity low while still maintaining a large number of college workers in the economy. By contrast, to attract college workers, a low-wage state such as Wyoming must offer relatively high out-of-state capacity.

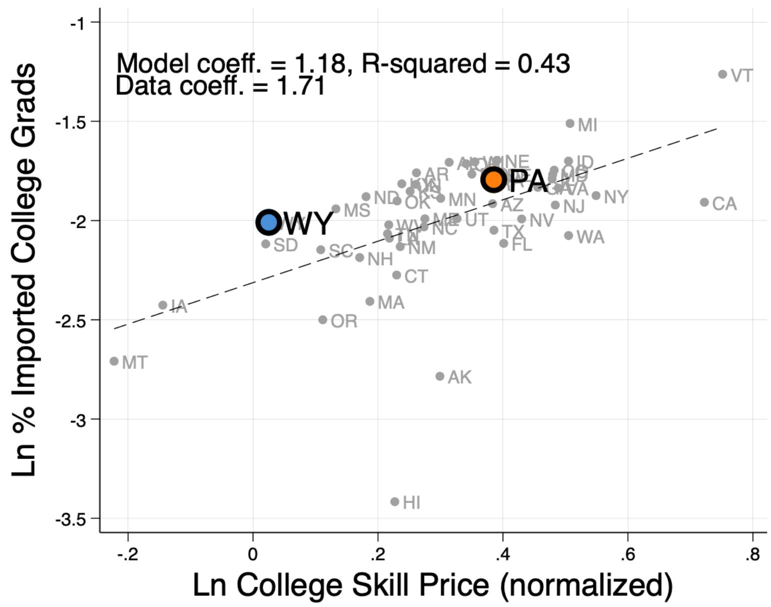
Figure 9 illustrates this logic by plotting the percent change in college workers against the percent change in out-of-state capacity. This elasticity is much steeper for Wyoming than for Pennsylvania, implying that the marginal return to an additional out-of-state college student is higher in Wyoming. As a result, it is optimal for Wyoming to set relative out-of-state capacity higher than Pennsylvania.

Now consider the positive correlation between out-of-state tuition and wages (column two, row two of *Table 12*). The benefit of raising out-of-state tuition is that spending per student rises, which increases college quality for in-state students. The costs of raising out-of-state tuition are twofold. First, average ability at the college falls as some students are priced out and demand declines among others, lowering college quality. Second, the number of students willing to migrate to the state for college falls, reducing the percentage of out-of-state students. This results in fewer college workers in the state, reducing domestic welfare. Thus, even if the first two effects combined raise college quality, welfare may still fall if the reduction in the number of college workers is sufficiently large.

Mechanically, the increase in spending per student is constant across state primitives.[19]

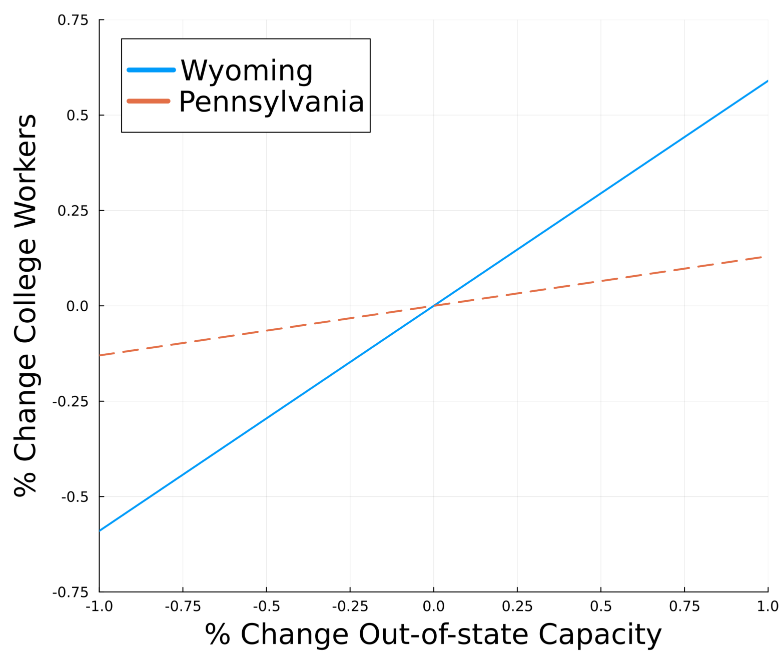
[19]The marginal return to welfare is not. For example, states with higher $\bar{q}$ experience higher marginal returns, all else equal. I abstract from this and other reasons the marginal return on welfare differs for simplicity of exposition.

Figure 8: **Elasticity of Imported College Graduates to College Skill Prices**



Notes: This figure plots the scatter of the natural log of % of college workers who graduated college in another state, to local college skill prices. The dashed black line reports the linear line of best fit, and is compared to the model equivalent.

Figure 9: **Change in College Workers to Change in Out-of-state Capacity**



Notes: This figure plots the change in number of college workers living in a state, to changes in out-of-state college capacity in the same state.

By contrast, the negative effects of increasing out-of-state tuition are not of the same magnitude across varying levels of firm productivity for college labor. *Figure 10* plots a heat map of average ability as out-of-state tuition and firm college labor productivity vary.[20] Average ability declines more slowly in high-productivity (high-wage) states. More specifically, moving vertically in *Figure 10*, the reduction in average ability (heat portion) is smaller for states with higher labor productivity (moving horizontally).

Figure 11 plots the percentage of the student body from out-of-state as a function of out-of-state tuition and firm college labor productivity. Again, the elasticity of out-of-state demand with respect to tuition is smaller for highly productive states. Hence, because the negative effects of raising out-of-state tuition are smaller for high-wage states, it is optimal for these states to set higher out-of-state tuition.

Returning to the example of Wyoming and Pennsylvania, they are marked as the black dots in both *Figures 10* and *11*, with Pennsylvania having high out-of-state tuition and high labor productivity, and Wyoming having low out-of-state tuition and low labor productivity. In both figures, Pennsylvania exhibits similar average ability and percentage of out-of-state students as Wyoming, even though its out-of-state tuition is higher. This reflects the fact that labor productivity (wage) is higher in Pennsylvania than in Wyoming.

The dynamics for in-state tuition and college wages (column two, row three of *Table 12*) are similar to those for out-of-state tuition, except that the relevant migration margin concerns students leaving the state to study elsewhere. When in-state tuition rises, many of these students leaving do not return to work in the state. *Figures E.29* and *E.30* plot the relevant figures for understanding in-state tuition.

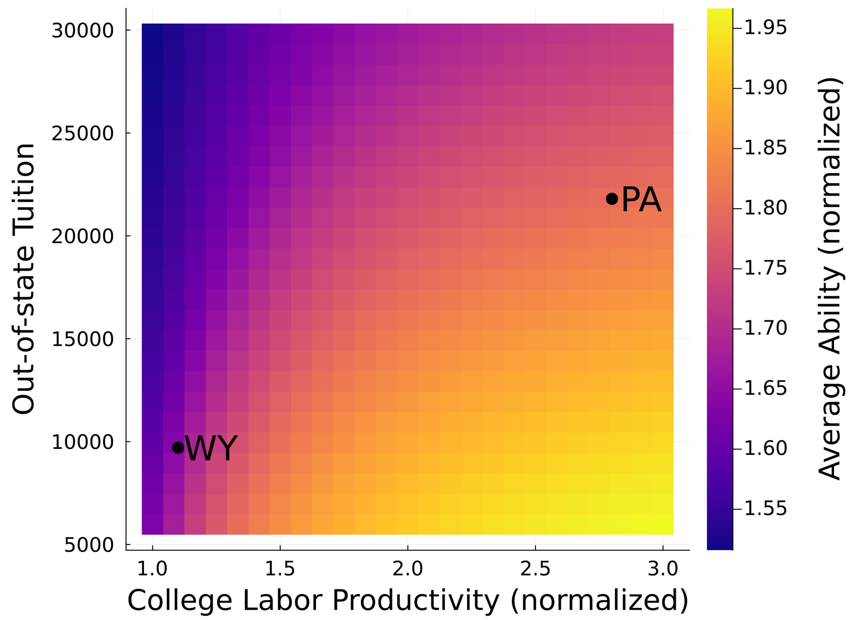
Finally, the positive correlation between wages and spending per student (column two, row four) can be understood as the residual from the above three correlations. Higher wages are associated with higher in- and out-of-state tuition and relatively lower out-of-state capacity. By the budget constraint, this translates into higher spending per student.

6 Centralized College System

A federal college system was proposed early in U.S. history by several founding fathers. The strongest proponent was James Madison, who addressed the issue in each State of the Union speech and put forth a bill that passed the Senate but ultimately failed in the House of Representatives. Today in the United States, the federal government's explicit involvement in higher education is limited to a handful of military colleges. However, fully centralized

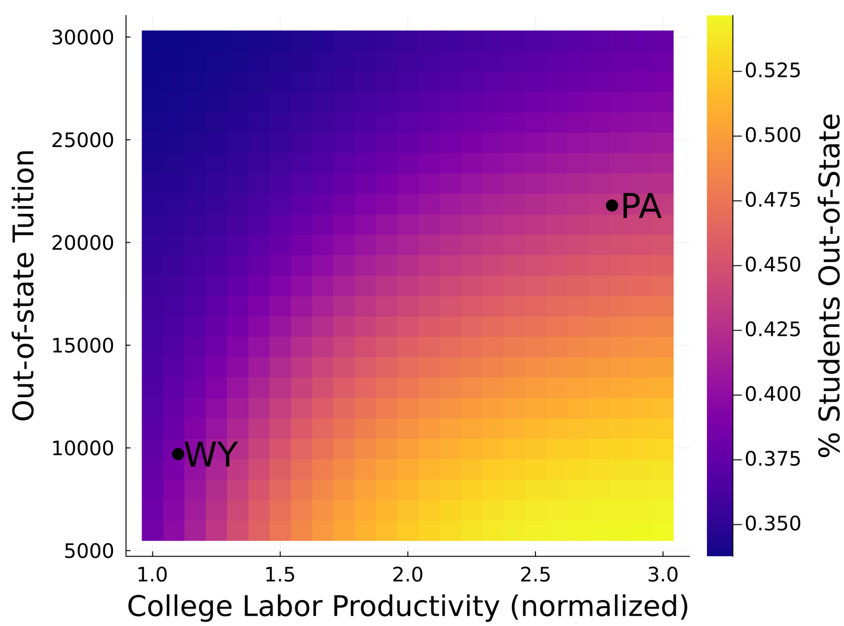
[20] To illustrate the mechanism, the figure is plotted by varying primitives for a generic state, assuming capacity does not bind.

Figure 10: **Mean College Ability to Out-of-state Tuition and Firm Labor Productivity**



Notes: This figure plots changes in college average ability (heat colors) to changes in out-of-state tuition and firm college labor productivity.

Figure 11: **% Students Out-of-state to Out-of-state Tuition and Firm Labor Productivity**



Notes: This figure plots changes in the percentage of student body out-of-state (heat colors) to changes in out-of-state tuition and firm college labor productivity.

higher education systems do exist in other countries, such as Germany or Brazil. Suppose there were a federal system of higher education in the United States today. Would one observe different patterns than presented in *Sections 2* and *5*, where each state acts individually?

I begin in *Section 6.1* by laying out the federal policy problem. In *Section 6.2.1*, I discuss the model and data objects that suggest why the federal policymaker may choose different optimal policies than each state acting individually. In *Section 6.2.2*, I compare the solutions to the state and federal Ramsey problems and provide economic intuition for the differences. Finally, *Section 6.3* concludes with a welfare analysis at the federal level.

6.1 Federal Government Problem

The federal government's problem is the natural extension of equation (28) where now the vector of capacities, spending per student, and tuition are chosen jointly to solve,

$$\max_{\{\mathcal{C}_\ell, m_\ell, T^i_\ell, T^o_\ell\}_{\ell\in\mathfrak{L}}} \sum_{\ell_b} \int_a \int_h \int_p \mu(a,h,p,\ell_b) V(a,h,p,\ell_b)\, da\, dh\, dp$$

subject to, (46)

$$\sum_{\ell} K\big(\mathcal{C}_\ell, m_\ell, \ell\big) = \sum_{\ell_c=\ell_b} \Big(T^i(\ell_c)\int_a\int_h\int_p \mu(e_c,a,h,\ell_b,p,\ell_c)\, da\, dh\, dp\Big)$$

$$\sum_{\ell_b\neq\ell_c} \Big(\big(T^i(\ell_c)+T^o(\ell_c)\big)\int_a\int_h\int_p \mu(e_c,a,h,\ell_b,p,\ell_c)\, da\, dh\, dp\Big) + \sum_{\ell} S_\ell + F$$

$$Q_\ell = \bar{q}_\ell\big(m_\ell\big)^\theta\big(\bar{a}_\ell\big)^{1-\theta} \qquad \forall \ell \in \mathfrak{L}$$

$$\sum_{\ell_b}\int_a\int_h\int_p \mu(e_c,a,h,\ell_b,p,\ell_c)\, da\, dh\, dp \leq \mathcal{C}(\ell_c) \qquad \forall \ell_c \in \mathfrak{L}$$

+ college demand equilibrium conditions

Note, the federal government is appropriating the sum of state revenues generated by labor taxes τ^s_ℓ. I formulate the federal problem in this way to isolate welfare changes exclusively due to the system of higher education and not any changes in tax distortions. Also note, under the solution to this counterfactual exercise, aggregate appropriations spent on college which are generated by labor taxation does not change. However, total revenues from in-state and out-of-state tuition (and hence total college expenditures) may change.

This problem does not produce the social planner allocation. The government in this problem is not able to choose all allocations as a planner would. In particular, the government is unable to dictate, which agents attend college, where each agent attends college, and where each agent is employed. Given peer-effects (and possible college spillovers), the solution to the federal government problem may see welfare improvements over the decentralized equilibrium, but will not be the welfare maximizing allocation of a social planner.

6.2 Federal versus State College Systems

In this section, I first discuss the economic forces that lead to different dynamics between the federal and state Ramsey policy problems. I then compare the solutions to both problems, with a focus on understanding differences in the motivating correlations from *Section 2*.

6.2.1 Federal versus State Problems

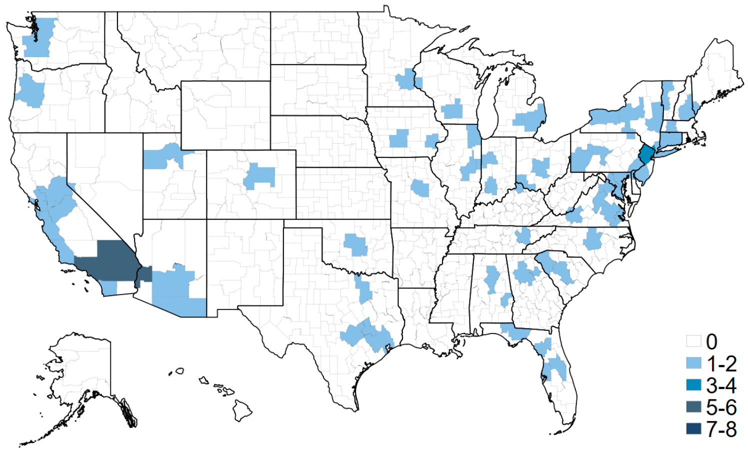
The state Ramsey problem chooses a best response given the actions of all other states. These strategic interactions are absent in the federal Ramsey problem. A second consequence of maximizing national social welfare is that the federal policymaker has motives for *redistribution* across states. While the spatial concentration of firm productivity and household wealth is well documented (Moretti, 2008; Chetty et al., 2014; Hsieh and Moretti, 2019), I show that college availability and quality are also highly concentrated across space in the United States.

Figure 12 plots the number of public colleges ranked among the top 150 nationally, by commuting zone. There is clear clustering along the coasts, with most central states lacking a high-quality public college. Overall, 411 of 741 commuting zones (18% of the total population) lack a public college of *any* quality. The federal Ramsey problem may therefore find it optimal to redistribute from states with many high-quality colleges (e.g., California or New York) to those with none (e.g., Wyoming or Nebraska).

The second key difference between the federal and state Ramsey problems is that the federal policymaker can internalize the entire distribution of college quality productivities relative to population, firm productivities, and other state primitives. In doing so, the federal policymaker also has motives towards *efficiency*.

Returning to *Figure 12*, the national distribution of college qualities relative to population is relevant to the optimal federal solution. An empirical literature studying "education deserts" finds that student demand for college is elastic to distance (Fu et al., 2022; Hillman, 2017; Turley, 2009), with this mechanism operating in the model through moving costs. Hence, the current college system, with high spatial concentration, may "miss" high-ability students due to the lack of colleges in certain states or regions.

Figure 12: **Top 150 Ranked Public Colleges by Commuting Zone**



Notes: This figure plots the count of top 150 nationally ranked public colleges (out of all colleges) located within each commuting zone. Rankings are taken from U.S. News (2020). Commuting zones are defined as in Autor and Dorn (2013).

The countrywide distribution of college qualities relative to firm productivities is also key to the federal Ramsey solution. Unlike firms, whose average age is five years, 73% of students attend a public college whose location was determined more than 100 years ago. Consider, for example, Penn State, founded in 1862 as the Agricultural College of Pennsylvania. Its centralized rural location had a clear policy rationale at the time, to provide agricultural expertise to local farmers in a centralized location.

Table 14 shows that, on average, colleges built over 100 years ago are located in (relatively) less densely populated areas today than when they were founded. Each county is nationally ranked by population density in 1920 and 2020, and then assigned a percentile. I take the difference between the percentile rank in 2020 and 1920 for the county in which a college is located. Hence, if a college is located in a county that became relatively less dense, this difference will be negative. This difference is then regressed on a dummy variable indicating whether the college was built over 100 years ago, and controls for size, land-grant status, and initial density. I find that colleges built over 100 years ago are associated with nearly a 10-percentile decrease in relative county population density. In the model, this fixed initial distribution of quality is represented by $\bar{q}$.

While the federal policymaker can reallocate spending and capacity toward more urban locations, the stock of intangible capital that determines college quality, such as name recognition, buildings, and placement networks is fixed and cannot be readily relocated. This is the fixed portion of college quality assumed to have been determined by historical college dynamics. To see the model mechanisms of this force, consider an example with two states

Table 14: **Changes in Relative Population Density by College Founding**

	(1) Density Difference
$i.100\ years\ old$	-9.31
	(0.000)
FTE	0.001
	(0.000)
$i.land\ grant$	0.55
	(0.776)
$dense\ rank$ 1920	-0.51
	(0.000)
$constant$	43.93
	(0.000)
R^2	0.68
observations	570

Notes: The dependent variable is the difference between the county population density percentile ranking (where the college is located) in 2020 and 1920. P-values are shown in brackets. Founding data are from the author's own collection. Historical county-level population density is taken from the U.S. Census Bureau. Control variables are obtained from the Integrated Postsecondary Education Data System (IPEDS).

that are ex-ante identical except for one with a high and the other with a low, historic level of college quality (i.e., $\bar{q}$). For simplicity of exposition assume that there is no separate out-of-state tuition.

Figure 13 plots this example economy, where both states have equal firm productivity. Each panel plots optimal policy as the federal government alters the fraction of total revenue allocated to the state (π_ℓ). The federal government finds it optimal to provide the majority of revenues to the high-$\bar{q}$ state. Spending per student between the two states is similar, and tuition is set to zero in the low-$\bar{q}$, but not in the high-$\bar{q}$ state. The government spends the additional revenue at the high-quality college, significantly expanding capacity. *Figure E.28* shows that this same pattern is even starker when the high-productivity firms are also located near the high-quality colleges.

However, *Figure 14* shows that when the high-quality colleges are located near the (relatively) low-productivity firms (e.g., Wisconsin), the federal policymaker now finds it optimal to spend roughly half of revenues on each location and set capacity close to equal. That is, any mismatch between historical college locations and present firm activity significantly alters decisions by a federal policy maker. In this situation, the federal government would "like to" move the location of the high-$\bar{q}$ colleges, however they are constrained in doing so.

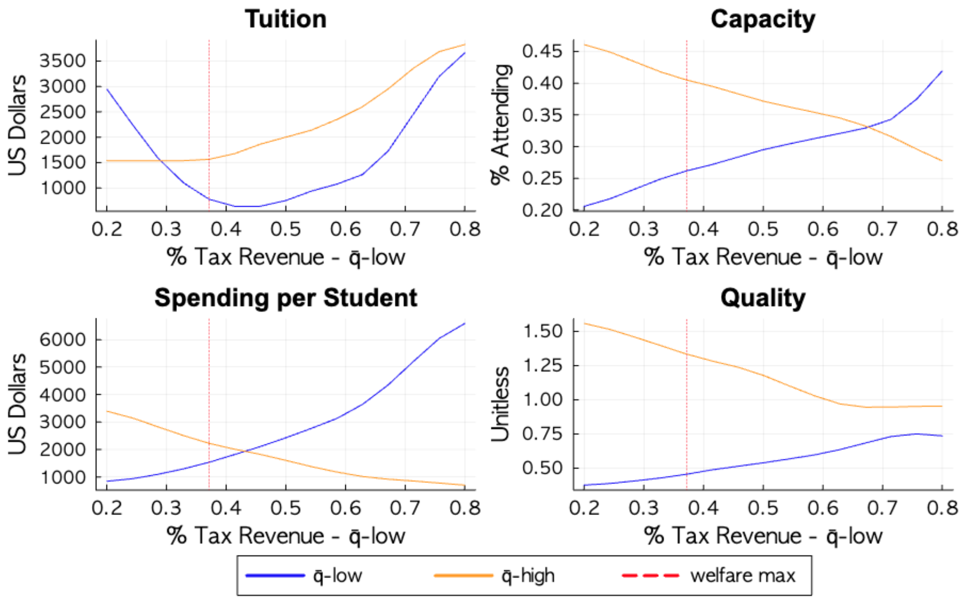
The final key difference between the state and federal Ramsey problems is that the college production function features externalities through peer effects. Individual agents do not internalize the impact of their own ability on college quality. While the federal Ramsey cannot fully internalize these externalities, a larger portion can be internalized at the national level by adjusting in-/out-of-state tuition and capacities, which in turn affect student and worker migration.

6.2.2 Federal versus State Solutions

Table 15 reports differences in the averages and heterogeneity of college policies between the federal and state solutions. First, while low in-state tuition was the optimal policy for most states under the decentralized system, it is not the optimal policy under the centralized regime. Second, rows one and two show that the federal policymaker eliminates the premium on out-of-state tuition. On average, out-of-state tuition is lower than in-state tuition, which acts as a moving subsidy for out-of-state students. While states maximizing domestic welfare have a natural incentive to subsidize in-state students by setting out-of-state tuition high, cross-state pricing frictions are inefficient from the perspective of the federal Ramsey problem.

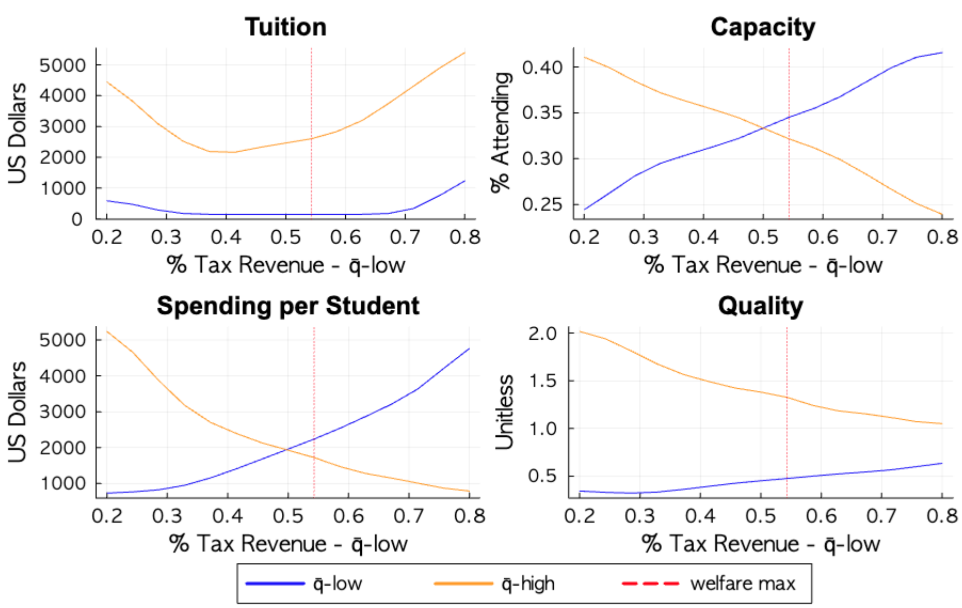
Under the federal Ramsey solution, the percentage of out-of-state students rises dramatically from 38.1% to 81.1%. In effect, the federal policymaker eliminates out-of-state capacity

Figure 13: **Federal Government Solution, Same Firm Productivity**



Notes: This figure plots the solution to the federal governments problem for two different states with the same productivity.

Figure 14: **Federal Government Solution, Unmatched Firm-College Productivity**



Notes: This figure plots the solution to the federal governments problem for two different states with different productivity. Firm productivity is "unmatched" with colleges in the sense that high-productivity firms are not located in the same location as high fixed $\bar{q}$ colleges.

Table 15: **Federal Optimal Policies Compared to State**

	Federal	**State**
mean T^{in}	\$11,505	\$2,669
mean T^{out}	\$10,629	\$27,563
mean % in-state	18.9%	61.9%
mean spend	\$16,837	\$11,844
std(T)	\$3,840	\$602.3
std(% in-state)	23.8%	15.4%
std(spend)	\$23,619	\$3,429

Notes: This table reports the mean and standard deviation in optimal federal college policies compared to optimal state college policies.

constraints. The intuition is straightforward. Given heterogeneity in initial conditions for ability, human capital, and parental income, the probability that an agent is born in the "right" state to attend college is low. Hence, the federal policymaker removes constraints on out-of-state students attending their most preferred college. By contrast, in the decentralized college system, states have an incentive to ration out-of-state seats to maintain high residual demand and high out-of-state tuition, thereby increasing quality and reducing tuition for in-state students. Given expanded out-of-state capacity and more students attending their most preferred college, total college attendance rises significantly, by 21.5 percentage points.

Although substantial heterogeneity in college policies exists across states and in the data, rows five to eight show that the standard deviation of college policies increases under the optimal centralized regime. Before turning to an understanding of the economics of this result, I first revisit (again) the motivating OLS regressions from *Section 2*.

Table 16 reports the same regressions as those empirically documented in *Section 2*, compared between the federal and state solutions. Under the federal solution, the correlations between population, wages, and college policies are of the opposite sign compared to the state solution or data. Larger and wealthier states are associated with lower in-/out-of-state tuition, higher relative out-of-state capacity, and lower spending per student. *Table E.9* reports federal college policies as a function of conditional correlations with state primitives.

Together with the increase in the variance of college policies, these correlations imply that colleges in large and high-wage states become larger, less competitive, cheaper, and lower quality, while colleges in small and low-wage states become smaller, more competitive, costlier, and higher quality. This federal policy reflects both the *efficiency* and *redistribution* motives of the federal policymaker.

Table 16: **Motivating OLS – Federal and State Solution**

	$\ln(pop)$		$\ln(col\ wage)$	
	Federal	State	Federal	State
$\ln(T^{out})$	-0.03	0.19	-0.06	0.67
$\ln(T^{in})$	-0.04	0.45	-0.08	0.37
$\ln(\%\ in-state)$	-0.14	0.21	-0.44	0.09
$\ln(spend\ per)$	-0.93	-0.02	-0.57	1.26

Notes: This table reports the motivating OLS regressions coefficients from *Section 2* compared between the federal and state solutions.

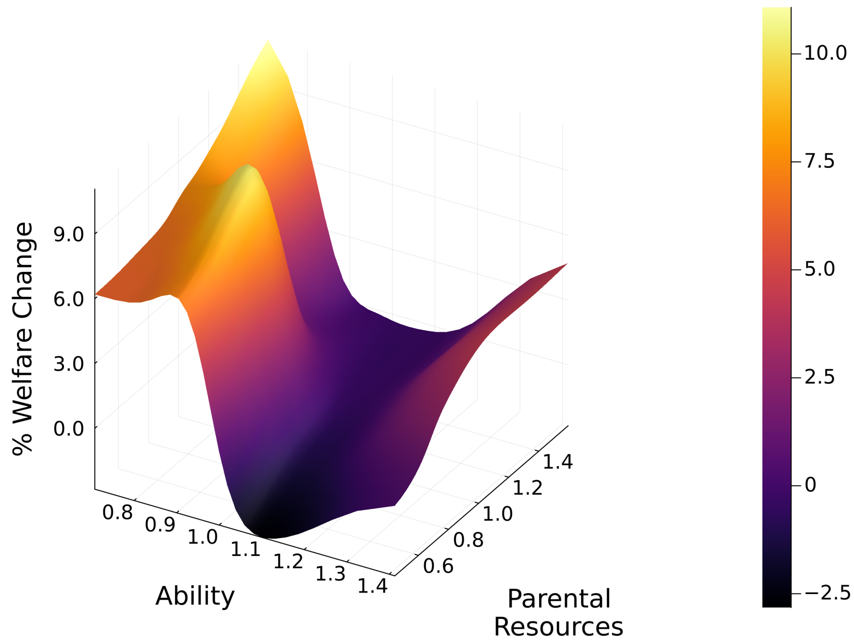
It is efficient to expand capacity and lower tuition barriers to high-quality colleges in highly productive labor markets. That is, given highly productive firms and high-quality colleges in California, it is optimal for the federal Ramsey to have more agents studying and living in California. *Figure E.32* illustrates this dynamic by plotting lifetime earnings as a function of the state primitives for college quality, productivity, and firm-college labor productivity. Agents experience higher lifetime earnings (and hence higher utility) in highly productive states. Therefore, it is optimal to allow many agents to attend college and work in these states. Moreover, *Figure E.14* shows that the number of alumni leaving the state is decreasing in firm wages. Hence, if the federal Ramsey planner were to expand college capacity in low-wage states, most agents would simply leave after graduating.

This policy also serves as redistribution, and the variance in skill prices and earnings falls by 8.1% and 19.7%, respectively. Rising college capacity in highly productive states increases the supply of college-educated workers and pushes down skill prices. Additionally, in high-productivity states with high-quality colleges, rising capacity decreases average ability and spending per student, which reduces college quality. The opposite occurs in low-productivity states, which raises skill prices and college quality. This leads to income convergence across locations.

6.3 Welfare Analysis

I now consider welfare changes when a federal government simultaneously chooses all 250 college policies across states to maximize aggregate welfare. This allocation does not coincide with each individual state's preferred policy choice, and each state would find it profitable to deviate from the federal solution when maximizing the welfare of its own residents. Following the logic of the preceding section, a state like Pennsylvania would find it profitable to increase

Figure 15: **Welfare Changes from Centralized System by Ability and Parental Resources**



Notes: This figure plots welfare changes associated with the federal system of higher education (averaged across states) by agent ability and parental resources. Ability and parental resources are reported relative to the mean.

out-of-state tuition and reduce out-of-state capacity, whereas Wyoming would increase out-of-state capacity. *Table E.12* reports details of the profitable deviation from the federal solution for Pennsylvania and Wyoming.

There are large welfare gains of 3.3% when the federal government chooses all college policy parameters optimally. This welfare increase occurs *without* any aggregate increase in college appropriations. Appropriations are simply allocated more efficiently and with the intent of redistribution.

However, welfare effects are heterogeneous, and not all types of agents are better off under the federalized system. *Figure 15* plots the heterogeneous welfare changes (averaged across states) by ability and parental income. The agents experiencing the largest welfare gains are primarily low-ability and low-income individuals who were unable to meet admission cutoffs or afford college in the decentralized system. In fact, *Figure 15* shows a sharp change in welfare gains around the income or ability threshold at which agents transition from being unable to attend college to being able to attend college. High-ability agents also experience large welfare gains, as many can now attend a high-quality college of their choosing at a more affordable tuition rate and benefit from ability–quality complementarities over the lifecycle.

Medium-ability agents fare the worst and experience welfare losses under the centralized college system. This largely arises because these agents, in high-quality and high-productivity states, are no longer able to attend the in-state institution and are crowded out by higher-ability out-of-state students. This generates a U-shaped relationship between ability and welfare gains. The relationship between welfare and parental resources, by contrast, is increasing for most of the state space. The primary reason is that many low-income agents who would have attended college in-state see tuition rise under the centralized system.

Welfare changes also vary significantly at the average state level. However, every state achieves higher welfare for its own domestic residents than under state-level optimization. An important implication of this result is that there are substantial welfare losses from strategic interaction in the decentralized Nash equilibrium. Competition and monopoly power between states create spatial frictions that make all states worse off than under a "cooperative" equilibrium in the federal solution.

Large states such as California see the smallest increase, at 0.7%, whereas small states such as Connecticut experience the largest increase, at 5.1%. On average, *Table E.9* shows that states with high-productivity firms and high-quality colleges (e.g., Pennsylvania) see the smallest welfare gains, whereas states with low-productivity firms and low-quality colleges (e.g., Wyoming) see the largest welfare gains. This result is intuitive given the discussion in the preceding sections, where it was shown that, on net, agents from Wyoming prefer to migrate to Pennsylvania for college.

7 Spatial Policies and Higher Education

In this section, I take insights from the full federal Ramsey solution presented in *Section 6* and apply it to national level policy setting for the U.S. economy. Three main takeaways from the federal Ramsey problem which are most relevant for national policy setting are: (1) eliminate out-of-state tuition premium, (2) increase percentage of out-of-state students, and (3) expand college capacity in high-productivity states. *Table 17* reports federal policy exercises and the associated welfare changes when solving a constrained Nash optimum problem for the decentralized college system.

First, column one considers a federal policy eliminating the out-of-state tuition premium. That is, I solve for the constrained Nash optimum in which states cannot choose an out-of-state tuition rate separate from in-state tuition. Otherwise, the problem is identical to that solved in equation (28). Here, welfare falls by 0.21%. When states can no longer charge higher tuition for out-of-state students, the incentives to admit them are significantly reduced. Most states now simply do not admit any out-of-state students, and welfare falls.

Table 17: **Welfare Changes of Federal Policy Experiments**

(1) No OS Tuit.	(2) No OS Cap.	(3) No OS Tuit.Cap.	(4) ↑ Cap. High Prod.	(5) Free Tuit.	(6) ↑ Cap.
-0.21%	0.67%	0.94%	0.49%	-0.29%	0.33%

Notes: This table plots the changes in welfare for selected policy experiments. Numbers are percentage point changes relative to benchmark calibrated U.S. economy.

Optimal in-state tuition is now higher as well, given that it can no longer be subsidized by out-of-state students. Consequently, in-state capacity and average ability also fall.

Second, column two eliminates specific capacity constraints for out-of-state students, mimicking the increase in the percentage of out-of-state students under the federal Ramsey solution. That is, I solve for the constrained Nash optimum in which states cannot choose separate in- and out-of-state capacities, and as a result there is now a single admissions cutoff. Welfare rises by 0.67%. Most states increase total capacity relative to the benchmark solution, in order to protect in-state students from being heavily crowded out of college by out-of-state students. The increase in welfare stems from a higher number of students attending their most preferred colleges nationally. Given the rich initial heterogeneity in agent characteristics, the probability that an individual's most preferred college is their home-state college is low. Eliminating out-of-state–specific capacity effectively increases available out-of-state capacity and lowers the out-of-state admissions cutoff.

However, greater demand and competition from out-of-state applicants raise the in-state admission criteria. The new single cutoff lies strictly between the in-state and out-of-state cutoffs in the benchmark economy. As a result, welfare rises the most for out-of-state students just below the admission cutoff in the benchmark economy, who are now able to attend their most preferred college. Conversely, welfare falls the most for low-income, medium-ability in-state students who are no longer admitted to their in-state college due to higher admission standards *and* who cannot afford to attend college out-of-state because of the higher out-of-state tuition premium.

Third, column three combines these two policies, with welfare now rising by 0.94%, nearly one-third of the gain under the full federal solution. Eliminating cross-state tuition frictions now has positive effects on aggregate welfare, as states can no longer simply react by not offering out-of-state capacity. Moreover, many low-income in-state students who were crowded out by out-of-state students under capacity-only reforms are now able to adjust their choices and more easily attend a preferred out-of-state college. However, this policy does not increase welfare for each state individually. A state like California (high out-of-state college demand) experiences welfare losses. Given the large residual demand for Californian colleges, most in-state students are crowded out of high-quality colleges and must move out-of-state to attend college. Additionally, spending per student falls due to lower out-of-state revenues, and hence even those domestic students not crowded out attend lower-quality colleges. The opposite is true for a state like Wyoming (low out-of-state college demand), where welfare rises significantly.

Fourth, column four increases in- and out-of-state capacity in the 25 most productive states by 10 percentage points, raising welfare by 0.49%. The main source of this welfare gain is the increase in the number of agents attending college and realizing higher lifetime earnings. As discussed in the federal Ramsey solution, this policy also serves as a mechanism for income convergence across states. Spending per student falls in the states where capacity expands, lowering college quality and reducing lifetime human capital accumulation for those attending these colleges. Additionally, the supply of college-educated workers rises in these states, which pushes down skill prices and decreases the lifetime earnings of those working in the state.

Columns five and six compare these spatial policies to standard policies proposed in the existing literature (Krueger et al., 2025; Abbott et al., 2019; Kennan, 2015; Epple et al., 2006) or by public officials. The first is a financing system for American higher education that mirrors that of Europe, where payments fall primarily on taxpayers rather than on individual households. This is implemented through a free national tuition policy and results in a 0.29% decrease in aggregate welfare.[21] The second reports the commonly proposed policy of expanding capacity at public state colleges. In- and out-of-state capacity are expanded by 6.4 percentage points to generate the same aggregate increase as the spatial policy targeted at high-productivity states. This policy raises welfare by 0.33%.

Eliminating out-of-state tuition and capacity frictions yields a large and positive welfare gain, whereas free tuition reduces welfare. Targeting capacity increases to high-productivity colleges, rather than to all states, raises welfare by 1.5 times more. Given that offering free tuition or expanding capacity may generate welfare gains, a standard model or analysis would suggest that these are advisable policies. However, the welfare gains associated with such policies are far from optimal. Without the full solution to the federal Ramsey problem and the novel quantitative spatial model of higher education developed in this paper, evaluating these policies would not be possible.

[21] If state appropriations also rise, there exists a free-tuition college policy that raises welfare; however, this is not comparable to the other policies considered throughout the paper, where aggregate appropriations are held fixed.

8 Conclusion

This paper studies the spatial distribution of higher education systems. I focus on the United States, where a uniquely decentralized college system has led to substantial heterogeneity in tuition, spending, and capacity across space. I ask the normative question: "Is the current American spatial distribution of college tuition, spending, and capacity optimal?" I also answer the natural positive question: "Why do college systems differ substantially across states?"

I begin by documenting several empirical facts: state size and wealth are positively correlated with in-/out-of-state tuition and spending per student, and negatively correlated with the relative capacity for out-of-state students. I then develop a novel quantitative spatial model that incorporates endogenous migration decisions across locations, distribution of college qualities, lifecycle human capital accumulation, and differentiated firms and wages across locations.

To answer the paper's main positive research question, I first solve for optimal college policies in a decentralized college system and show that the model replicates well the motivating facts and correlations documented in the data. The state-level optimal solution is a Nash equilibrium in which each state maximizes the welfare of its residents, taking as given the actions of all other states. Given that the state-level optimal policy problem rationalizes the observed differences in higher education across the United States, I use it to understand the economic forces underlying differences in college policies across states.

The correlation between state size and college policies is primarily driven by supply-side factors. Expanding capacity in California is more costly than in North Dakota; hence tuition is set higher and relative out-of-state capacity lower. By contrast, the correlation between state wages and college policies is mainly determined by demand-side factors. Higher-wage states experience greater demand for their labor markets, and consequently for their colleges, which leads them to set higher in-/out-of-state tuition. The negative correlation between out-of-state capacity and state wages is driven by migration flows and the marginal return to an out-of-state student. California attracts a large influx of college graduates from out of state, and few in-state graduates leave to work elsewhere. Hence, California sustains a high number of college workers with a relatively small capacity for out-of-state students. North Dakota, by contrast, must offer relatively high out-of-state capacity in order to increase the number of college workers in the state.

To answer the paper's normative question, I conduct a welfare analysis of optimal policy at the federal level. In the federal problem, the government simultaneously chooses in-/out-of-state capacity, tuition, and spending per student. For a fixed level of government appropriations, I find that optimally choosing college policies allows the federal government to improve aggregate welfare by 3.3%. The federal solution increases welfare relative to the state problem for every state individually, indicating substantial welfare losses from strategic interaction in the decentralized Nash equilibrium.

Finally, I apply the lessons from the full federal solution to policy design. The three main takeaways from the federal problem are: (1) eliminate the out-of-state tuition premium, (2) increase the percentage of out-of-state students, and (3) expand college capacity in high-productivity states. I show that implementing these spatial policies is significantly more effective than standard proposals in the literature or by policy officials. A federal policy eliminating the out-of-state tuition premium and out-of-state-specific admissions increases aggregate welfare by 0.94%, nearly one-third of the gain under the fully optimal federal policy. By contrast, a free national tuition policy decreases welfare by 0.29%. Only by solving the full federal Ramsey problem within the novel quantitative spatial model developed in this paper is it possible to correctly evaluate and design these spatial policies.

References

Abbott, Brant, Giovanni Gallipoli, Costas Meghir, and Giovanni L Violante, "Education Policy and Intergenerational Transfers in Equilibrium", *Journal of Political Economy* 127 (2019), 2569–2624.

Andrews, Michael J, "How do Institutions of Higher Education Affect Local Invention? Evidence from the Establishment of US Colleges", *American Economic Journal: Economic Policy* 15 (2023), 1–41.

Andrews, Michael J and Lauren Russell, "Historical Place-Based Investments and Contemporary Economic Mobility and Inequality: Impacts of University Establishment", *Working Paper* (2025).

Autor, David H and David Dorn, "The Growth of Low-skill Service Jobs and the Polarization of the U.S. Labor Market", *American Economic Review* 103 (2013), 1553–1597.

Bartik, Timothy J, "Who benefits from state and local economic development policies?" (1991).

Berry, Steven T, "Estimating Discrete-choice Models of Product Differentiation", *The RAND Journal of Economics* (1994), 242–262.

Brotherhood, Luiz, Bernard Herskovic, and Joao Ramos, "Income-based Affirmative Action in College Admissions", *The Economic Journal* 133 (2023), 1810–1845.

Capelle, Damien, Charly Porcher, and Hannah Rubinton, "The Geography of Higher Education and the Spatial Propagation of Skill-Biased Technological Change" (2025).

Cellini, Stephanie Riegg and Nicholas Turner, "Gainfully Employed?: Assessing the Employment and Earnings of For-profit College Students Using Administrative Data", *Journal of Human Resources* 54 (2019), 342–370.

Chetty, Raj, John N Friedman, Emmanuel Saez, Nicholas Turner, and Danny Yagan, "Income Segregation and Intergenerational Mobility Across Colleges in the United States", *The Quarterly Journal of Economics* 135 (2020), 1567–1633.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez, "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States", *The quarterly journal of economics* 129 (2014), 1553–1623.

Conzelmann, Johnathan G, Steven W Hemelt, Brad J Hershbein, Shawn Martin, Andrew Simon, and Kevin M Stange, "Grads On the Go: Measuring College-specific Labor Markets for Graduates", *Journal of Policy Analysis and Management* (2023).

Deming, David J and Christopher R Walters, "The Impact of State Budget Cuts on US Postsecondary Attainment", *Working Paper 23736* (2017).

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

Epple, Dennis, Richard Romano, Sinan Sarpça, and Holger Sieg, "A General Equilibrium Analysis of State and Private Colleges and Access to Higher Education in the US", *Journal of Public Economics* 155 (2017), 164–178.

Epple, Dennis, Richard Romano, and Holger Sieg, "Admission, Tuition, and Financial Aid Policies in the Market for Higher Education", *Econometrica* 74 (2006), 885–928.

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

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

Guvenen, Fatih and Serdar Ozkan, *TikTak: A Multistart Global Optimization Algorithm*, 2021.

Heckman, James J, Lance Lochner, and Christopher Taber, "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents", *Review of Economic Dynamics* 1 (1998), 1–58.

Hendricks, Lutz, Christopher Herrington, and Todd Schoellman, "College Quality and Attendance Patterns: A Long-run View", *American Economic Journal: Macroeconomics* 13 (2021), 184–215.

Herrington, Christopher, "Public Education Financing, Earnings Inequality, and Intergenerational Mobility", *Review of Economic Dynamics* 18 (2015), 822–842.

Hillman, Nicholas W, "Geospatial Analysis in Higher Education Research", *Higher Education: Handbook of Theory and Research* (2017), 529–575.

Holter, Hans A, "Accounting for Cross-country Differences in Intergenerational Earnings Persistence: The Impact of Taxation and Public Education Expenditure", *Quantitative Economics* 6 (2015), 385–428.

Hsieh, Chang-Tai and Enrico Moretti, "Housing Constraints and Spatial Misallocation", *American economic journal: macroeconomics* 11 (2019), 1–39.

Huggett, Mark, Gustavo Ventura, and Amir Yaron, "Sources of Lifetime Inequality", *American Economic Review* 101 (2011), 2923–2954.

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

Kane, Thomas J, "Public Intervention in Post-secondary Education", *Handbook of the Economics of Education* 2 (2006), 1369–1401.

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

Kennan, John, "Spatial Variation in Higher Education Financing and the Supply of College Graduates", *National Bureau of Economic Research Working Paper* (2015).

Kirkeboen, Lars J, Edwin Leuven, and Magne Mogstad, "Field of Study, Earnings, and Self-selection", *The Quarterly Journal of Economics* 131 (2016), 1057–1111.

Kline, Patrick and Enrico Moretti, "People, Places, and Public Policy: Some Simple Welfare Economics of Local Economic Development Programs", *Annual Review of Economics* 6 (2014), 629–662.

Krueger, Dirk and Alexander Ludwig, "On the Optimal Provision of Social Insurance: Progressive Taxation Versus Education Subsidies in General Equilibrium", *Journal of Monetary Economics* 77 (2016), 72–98.

Krueger, Dirk, Alexander Ludwig, and Irina Popova, "Shaping Inequality and Intergenerational Persistence of Poverty: Free College or Better Schools?", *Journal of Monetary Economics* 150 (2025), 103694.

Lee, Sang Min and Bipul Verma, "Higher Education Policies and Spatial Inequality Across US States" (2025).

McFadden, Daniel, "Conditional Logit Analysis of Qualitative Choice Behavior" (1972).

Moretti, Enrico, "Real Wage Inequality" (2008).

Parmley, Kelli, Allison Bell, Hans L'Orange, and Paul Lingenfelter, "State Budgeting for Higher Education in the United States", *Report, State Higher Education Executive Officers* (2009).

Redding, Stephen J and Esteban Rossi-Hansberg, "Quantitative Spatial Economics", *Annual Review of Economics* 9 (2017), 21–58.

Russell, Lauren C, Lei Yu, and Michael J Andrews, "Higher Education and Local Educational Attainment: Evidence from the Establishment of US Colleges", *Review of Economics and Statistics* 106 (2024), 1146–1156.

Turley, Ruth N López, "College Proximity: Mapping Access to Opportunity", *Sociology of Education* 82 (2009), 126–146.

Wright, Jacob, "Quality, Inequality, and Mobility: American vs. Canadian Systems of Higher Education", *Available at SSRN: https://ssrn.com/abstract=4953531* (2023).

Wright, Jacob and Alex Wurdinger, "Declining Teen Employment: Causes and Consequences", *Available at SSRN: https://ssrn.com/abstract=4953600* (2025).

Wright, Jacob and Angela Zheng, "From Kindergarten to College: The Impact of Education Policies over the Lifecycle", *IZA Discussion Papers, No. 17301, Institute of Labor Economics* (2025).

Appendix

"On the Spatial Distribution of Colleges"

Jacob Wright

Contents

A Data

A.1 Data Sources

In this Appendix, I report details of each dataset used in *Section 2* for motivating statistics and *Section 4* for model parameters and moments.

Integrated Postsecondary Education Data System – The Integrated Postsecondary Education Data System is built from interrelated surveys conducted annually by the U.S. Department of Education's National Center for Education Statistics. The survey is one of the key datasets used for much of the analysis throughout the paper. The survey covers a large array of variables on admissions, enrollment, graduation rates, and college finances. Most variables cover the 1990 to 2022 time period.

College Scorecard – The College Scorecard dataset is maintained by the Department of Education. The key variable used from the College Scorecard is earnings by former students. This data are reported for 6, 7, 8, 9, and 10 years after a given pooled cohort entered college. Median earnings are given for individuals who are working and not enrolled in a college at the point of measurement. Earnings are defined as wages and deferred compensation from the IRS W-2 form.

National Association of College and University Business Officers – The National Association of College and University Business Officers is a membership organization representing more than 1,900 colleges and universities across the country. Among other activities, it conducts an annual study of college endowments in conjunction with the Teachers Insurance and Annuity Association of America. The total endowment variable is the one used here.

Grapevine – Grapevine is an annual data collection effort that reports on state funding for higher education, jointly produced by the State Higher Education Executive Officers (SHEEO) Association and the Center for the Study of Education Policy at Illinois State University. Consistent with the approach in Deming and Walters (2017), I use Grapevine's appropriation figures instead of those from IPEDS due to concerns about double-counting state support across different campuses within the same institution and potential inaccuracies in administrator responses to the IPEDS survey.

Baccalaureate and Beyond – The Baccalaureate and Beyond (B&B) is a nationally representative longitudinal study that tracks individuals who earn a bachelor's degree in a specific academic year, following their outcomes one, four, and ten years after graduation. This allows me to obtain migration rates from the state of college attendance to the subsequent state of employment. I use the B&B as opposed to the NLSY, as the sample size for college graduates is roughly 10 times larger, and I am able to use the 2016–2018 time period as opposed to the late 1990s or early 2000s as in the NLSY. B&B microdata is restricted access; however, the National Center for Education Statistics (NCES) Datalab provides the creation of custom, publicly available tables.

College Founding Dates – To the best of my knowledge, there does not exist a publicly available dataset covering college founding or establishment dates. I build this dataset by scraping Wikipedia. The overwhelming majority of college Wikipedia pages contain a box along the right-hand side of the page with a field labeled "Established" or "Founded," as seen in *Figure E.17*. This makes scraping simple and accurate using a list of college names from IPEDS. A few remaining colleges were then manually added. In total, I gather 1,610 founding dates for private and public colleges. I have made the dataset containing these compiled founding dates publicly available at https://doi.org/10.3886/E233781V1. College founding dates vary in accuracy in the sense that, as colleges grow, merge, split, etc. over time, there is not a standardized way to determine how the founding date evolves. However, for the simple correlation motivating statistics, the officially reported college dates serve as a reasonable proxy on average.

Deming and Walters (2017) – For data on tuition freezes and caps by state, I take data from Deming and Walters (2017). This paper compiles the data for the years 1990–2013 using a combination of official sources and Lexis-Nexis searches of state newspapers. All caps or freezes were coded only under the condition that they could be independently verified.

Chetty et al. (2020) – Data for wage profiles for the ages 23–35 is taken from Chetty et al. (2020). Publicly available statistics are published using microdata taken directly from IRS tax records.

Conzelmann et al. (2023) – Conzelmann et al. (2023) construct a dataset on college alumni by scraping publicly available LinkedIn (LI) profiles from institutional webpages. The resulting data offer aggregated information on the geographic distribution of graduates from nearly all public and private nonprofit higher education institutions in the U.S. Through a series of validation checks, the authors show that the LinkedIn-based data closely align with official government sources in both graduate counts and location patterns. I use data from Conzelmann et al. (2023) as opposed to the B&B or other sources when I require alumni migration patterns at the individual college level instead of the aggregate state level.

Panel Study of Income Dynamics – The Panel Study of Income Dynamics (PSID) launched in 1968 with a nationally representative cohort of over 18,000 individuals from 5,000 U.S. families. It has since tracked these individuals and their descendants over time, collecting detailed information on income, mobility, and other socioeconomic outcomes. I use the PSID to calculate skill prices, as it is the largest dataset that allows me to do so. While the ACS contains information on migration, it does not report previous years' earnings.

National Longitudinal Survey of Youths – The NLSY97 is a nationally representative longitudinal dataset administered by the U.S. Bureau of Labor Statistics, tracking a cohort of American youth born between 1980 and 1984. Participants were between 12 and 17 years old when they were first surveyed in 1997. The survey provides detailed information on a wide range of topics, including education, labor market outcomes, fertility, public assistance, health, family characteristics, and individual beliefs. For our analysis, we use data from the NLSY97 on student achievement and family background to discipline parameters related to intergenerational processes.

American Community Survey – The American Community Survey (ACS) is an ongoing, nationally representative survey conducted by the U.S. Census Bureau. It provides annual data on demographic, economic, housing, and social characteristics of the U.S. population. I use ACS microdata to measure cross-state migration patterns, population sizes, educational attainment, and earnings. The ACS is available annually from 2005 onward.

Current Population Survey – The Current Population Survey (CPS) is a monthly and annual, nationally representative survey conducted by the U.S. Census Bureau and the Bureau of Labor Statistics. It provides detailed information on earnings and educational attainment. I use the CPS over the ACS when I require annual data prior to 2005, and the significantly larger sample size of the ACS is not required.

United States Census – The U.S. Census, conducted every ten years by the U.S. Census Bureau, provides detailed demographic, social, housing, and economic data on the entire U.S. population. The decennial census offers comprehensive population counts and characteristics at various geographic levels, including states, counties, and census tracts.

I use U.S. Census data for several purposes. First, to calculate rent prices at the state level. I use the National Historical Geographic Information System (NHGIS), which is a comprehensive source of historical U.S. Census data and geographic boundary files, covering the period from 1790 onward. It delivers harmonized summary statistics on demographics, housing, agriculture, and economic indicators across multiple geographic levels—ranging from states to census blocks. The database is particularly valuable for conducting longitudinal and spatial research using consistent census geography. Second, I use Census data to calculate population-weighted centroids for each state. I do so using the population and coordinates reported by Block Groups from the U.S. Census. Lastly, I use the U.S. Census for historically comparable county population estimates dating back to 1900.

National Income and Product Accounts – The National Income and Product Accounts (NIPA), maintained by the U.S. Bureau of Economic Analysis (BEA), provide comprehensive data on the economic activity of the United States. These accounts include measures such as GDP, personal income, and government expenditures, offering a consistent framework to track the flow of income and production across sectors. In this paper, I use NIPA for data on disposable income after taxes by state. The above datasets all report pre-tax earnings or income.

Consumer Expenditure Survey – The Consumer Expenditure Survey (CEX), produced by the U.S. Bureau of Labor Statistics, provides detailed information on household spending across a wide range of categories. I use the published "Top Line Mean" tables, which report average annual expenditures by major category, to compute the share of total housing-related expenses—such as rent, mortgage payments, utilities, and maintenance—relative to average total annual household expenditures.

A.2 Higher Education Across States

In this appendix, I report additional details beyond those discussed in *Section 2* on cross-state differences in higher education. All figures are shown in *Appendix E*.

Financing – In addition to spending per student, state financing of colleges differs in several other important ways. *Figure E.11* shows that higher education accounts for a much larger share of total government outlays in Texas than in California or Pennsylvania, with the former spending three times as much as the latter. Another major dimension of variation is the proportion of total college expenditures financed by state tax revenues. *Figure E.12* plots the average share of a college's budget funded by state appropriations. In California and New York, nearly half of the college budget comes from state appropriations—a share that is even higher when including state grants and contracts. At the other end of the spectrum, colleges in Pennsylvania and Michigan fund less than 10% of expenditures through state appropriations. This implies that alternative revenue sources, such as tuition or donations, are much more important for college budgets in some states than in others.

Migration – Now consider student migration across states. *Figure E.13* plots the percentage of a state's college student body that originates from out of state. For example, New York and Texas have relatively few out-of-state students, with less than 10% of their student populations coming from elsewhere. In contrast, in Vermont, the majority of college students are from outside the state. Alternatively, consider student outflows. *Figure E.14* plots the percentage of students who attend college outside their home state. Nearly all students who graduate from high school in California remain in-state for college, whereas the majority of students from New Jersey leave to attend college elsewhere.

Measures of Ability – I now contrast different measures of student ability at public colleges across states. First, *Figure E.15* plots median ACT composite scores by state.[22] Average ACT composite scores fall around the 86^{th} percentile of nationwide scores in states like New York and Michigan, whereas states such as South Dakota and Idaho have average scores near the 66^{th} percentile.

A second common proxy for the academic composition of college students is the 6-year (or 150%) graduation rate, shown in *Figure E.16*. Once again, New York and Michigan have high rates, exceeding 65%, while Idaho and South Dakota have graduation rates below 50%. These graduation rates are important for understanding the cost of producing a graduate, as opposed to spending per enrolled student. A state may spend relatively little per student but have a high dropout rate, which results in a high cost per graduate. An interesting extension to this paper would be to endogenize graduation rates in this context.

[22] In the data, I observe the 25^{th} and 75^{th} percentiles and define the median as the midpoint of these two scores.

Tuition and Spending per Student – The dependent variables in the motivating regressions presented in *Section 2* are also significantly correlated with each other. *Table E.3* reports the relationship between spending per student and in-/out-of-state tuition. Given a state and college budget constraint, we see that states with higher in- and out-of-state tuition spend significantly more per student on higher education.

B Calibration

This section reports additional details and robustness exercises for the model calibration presented in *Section 4*. All figures and tables are shown in *Appendix E*.

B.1 Colleges

B.1.1 Reciprocity Agreements

Full Reciprocity – Reciprocity agreements are not automatic, and students must apply and prove they qualify. The only full statewide reciprocity agreements are between Minnesota and Wisconsin, and Minnesota and North Dakota, for all public colleges. Minnesota and South Dakota had a past agreement, but South Dakota exited the agreement in Fall 2024, and Minnesota will exit in Fall 2025.[23]

Regional Compacts – There are four large regional compacts: (1) Western Undergraduate Exchange, (2) Midwest Student Exchange Program, (3) Academic Common Market (Southern region), and (4) New England Regional Student Program.

These compacts differ from full state-to-state reciprocity, as they offer discounts for in-state tuition at a specific set of schools and for a specific set of majors. For instance, the Academic Common Market only offers tuition discounts for majors not offered in the home state, and the New England Regional Student Program provides a 25%–75% discount, which varies by campus at participating colleges. I exclude these four agreements from the analysis.

Other – There also exist reciprocity agreements exclusively for residents of specific border counties, campuses, or professional schools. I exclude these from the analysis. An example of a larger county-based agreement is that between Indiana and Ohio, which is active from 2023 to 2025. However, this agreement mostly includes community colleges, which are excluded from the sample of colleges in this analysis. In general, most county-specific agreements are geared toward community colleges.

B.1.2 Spending per Student

Here I perform robustness analysis for the definition of spending per student. As discussed in *Section 4.2.2* I follow Epple et al. (2006) and consider m_ℓ to consist of total instructional spending. I consider three additional definitions of spending per student. (1) instructional

[23] Minnesota also has full tuition reciprocity with the Canadian province of Manitoba.

salaries, (2) total instructional expenditures + total research expenditures, (3) total instructional expenditures + total student services expenditures.

The key place in which this definition affects the counterfactual exercise is in the estimation of the cost function and college production function. The correlation between the benchmark definition of spending per student and the above three is 0.87, 0.99, and 0.99, respectively. While the definition will affect the calibrated level of $\bar{q}$, given the high correlation across definitions, the relative $\bar{q}$ does not change significantly. For the estimation of the cost function, the results are quite similar across all specifications and do not affect the main counterfactual results.

B.1.3 Quality Function

The main issue with estimation of the quality function and human capital function is that θ is cannot be separately identified from any elasticity between ability and quality in the production of human capital. A solution would be to estimate the elasticity of later life earnings to college peer-effects however in the empirical literature this is notoriously not well estimated. The alternative that I have employed is to assume equal elasticity between college quality and ability, and perform two robustness checks.

The effect of spending on college quality (subsequent wages) is well identified. The residual is large, and can be either explained by peer-effects or a productivity term $\bar{q}$. I rerun the analysis of this paper assuming that all differences are either solely attributable to $\bar{a}$ or $\bar{q}$. I find that no key results in *Table 12* or the welfare analysis are substantially changed.

B.2 External Validation

Here I report regression details and robustness specifications for the elasticities reported for external validation in *Section 4.6.2*. *Table E.4* in *Appendix E* reports all results. The preferred specification used in *Section 4.6.2* is reported in column (1). The remaining columns show varying robustness and instrumental variable specifications.

The instrumental variable for spending per student is constructed as in Deming and Walters (2017) and described in *Section 4.6.2*. The instrumental variable for wages is a standard Bartik-style instrument, where lagged industrial shares by state are interacted with national wage changes by industry. The coefficients of interest are quite stable across all specifications.

C Computation

C.1 College Demand Equilibrium

The solution algorithm for solving the college demand equilibrium proceeds as follows,

1. Provide an initial guess for aggregate labor demand in each location, $Y(c, \ell)$ and $Y(nc, \ell)$, and for college qualities Q_ℓ.

   1.1 Solve for all other endogenous objects using the model's equilibrium equations.

2. Solve the household's lifecycle problem backwards and obtain values $W^{nc}(a, h, \ell_w)$ and $W^c(a, h, \ell_c, \ell_w)$ for non-college and college workers, respectively.
3. Solve discrete choice problems sequentially, working backwards.
4. Solve for admissions cutoffs that clear in-/out-of-state capacity at each college. This requires a modified solution to the sequence of discrete choice problems.
5. Update guesses using $d \in (0, 1)$ weighting on the newly computed objects.
6. Repeat steps 2 through 5 until convergence is achieved in the sup-norm.

C.2 Calibration

The parameters to be internally estimated consist of the preference parameters, $[\gamma_1^e, \gamma_2^e, \gamma_1^w, \gamma_2^w, \gamma_3^w, \sigma_e, \sigma_{cq}, \sigma_{cw}, \sigma_{ncw}]$, initial distribution parameters $[\mu_a, \sigma_a, \mu_p, \sigma_p]$, college fixed productivity for each state $\bar{q}_\ell$, and the elasticity of college quality with respect to spending per student, θ. Call this vector of parameters Θ. Define the loss function to minimize as $\mathcal{L}(\Theta) \equiv \big(m(\Theta) - \hat{m}\big)' \mathbf{W} \big(m(\Theta) - \hat{m}\big)$, where $m(\Theta)$ is the vector of model-simulated moments, $\hat{m}$ is the vector of data-equivalent moments, and $\mathbf{W}$ is a weighting matrix. I set the weighting matrix to be diagonal and inversely proportional to the standard deviation of the estimated values of $\hat{m}$.

The internal calibration procedure then involves a nested fixed point problem and proceeds as follows,

1. Guess θ and the vector $\bar{q}_\ell$.
2. Solve for the set of parameters $[\gamma_1^e, \gamma_2^e, \gamma_1^w, \gamma_2^w, \gamma_3^w, \sigma_e, \sigma_{cq}, \sigma_{cw}, \sigma_{ncw}, \mu_a, \sigma_a, \mu_p, \sigma_p]$ to minimize the distance between the model and data.[24]

[24] I use the TikTak global optimization method from Guvenen and Ozkan (2021) with 500 starting points and L-BFGS as the local optimizer to solve for this set of parameters.

3. Solve for $\bar{q}_\ell$ to match the earnings growth rate in the model and data using a bisection root-finding routine.
4. Repeat steps 2 and 3 until convergence on $\bar{q}_\ell$ is achieved.
5. Shock the model with an exogenous 1 percentage point increase in spending at each college. Compare the change in earnings to the elasticity estimated in the data. Using a bisection routine, find θ that matches the model elasticity to the data.
6. Repeat steps 2-5 until convergence in θ is achieved.

C.3 State Government

The state problem is solved using a modified L-BFGS quasi-Newton algorithm to allow for non-negatively constraints on all choice variables. Gradients are calculated using the central finite differences method. The state's problem is akin to a Ramsey problem, and so the full general equilibrium (see *Appendix C.1*) must be solved at each point in the optimization routine. Solving the state's problem then proceeds as follows,

1. Guess an set of initial vectors for each state $T_\ell^i, T_\ell^o, \mathcal{C}_\ell^i, \mathcal{C}_\ell^o$, and m_ℓ.
2. Solve each state problem, holding fixed all other state policies at initial guess.[25]
3. Update policy vectors using a weight $d \in (0, 1)$ on solution from previous iteration.
4. Repeat steps 2-3 until convergence is achieved on $T_\ell^i, T_\ell^o, \mathcal{C}_\ell^i, \mathcal{C}_\ell^o$, and m_ℓ.

C.4 Federal Government

The federal government's problem is a high-dimensional optimization problem consisting of 249 choice variables. As in the state government's problem, the full general equilibrium (see *Appendix C.1*) must be solved at each point in the optimization routine. Given this dimensionality, the problem is best suited to an L-BFGS quasi-Newton optimizer. However, two main issues arise.

First, the problem features multiple local maxima and minima. To address this, I employ a multi-start L-BFGS procedure à la TikTak, with 100 starting points. I confirm that the global solution is reached by comparing results to those obtained from a global derivative-free black-box optimizer based on a Natural Evolution Strategy algorithm.

Second, finite-difference derivatives would require $2 \times 249 = 498$ solves of the model and are therefore infeasible. Instead, I employ reverse-mode automatic differentiation. To avoid

[25] This step of the solution algorithm is easily parallelized over 50 cores.

back propagation through every iteration of the fixed-point solver, I recast the problem as an implicit system and apply the Implicit Function Theorem to differentiate the equilibrium directly.

D Model

D.1 Additional Results

D.1.1 Allocating an Additional Dollar

The state government can allocate revenues toward: (1) reducing in-state tuition, (2) increasing in-state capacity, (3) reducing out-of-state tuition, (4) increasing out-of-state capacity, and (5) spending to directly increase college quality. In this section, I provide the economic intuition behind the optimal choice over these five objects.

For simplicity and clarity of exposition, consider a stylized five-state model with two states of interest: one with high, and the other with low, out-of-state college demand. That is, one with a high admissions cutoff for out-of-state students and the other with near open admissions. Call these high- and low-demand states, respectively. Again, the quantitative dynamics are similar, however considerably more complicated with 50 states.

I choose these state types as they best highlight differences across states. *Figures E.19* and *E.20* plot the optimal decisions for the state government as funding through state labor tax appropriations is increased. That is, as an extra dollar becomes available, how does the state chooses to optimally allocate it.

For low levels of state appropriations, the low-demand state sets a relatively high in-state tuition and a small out-of-state capacity. Spending per student is also quite low, and quality is relatively low as well. Conversely, the high-demand state is able to set out-of-state tuition and capacity much higher. This allows for an almost full subsidization of in-state tuition, and spending per student that is significantly higher than in the low-demand state. Mean ability is lower in the high-demand state than in the low-demand state, as the higher capacity at the high-demand state results in a lower average ability of the student body. However, overall, the higher spending per student dominates, and quality is higher in the high-demand state.

As state appropriations increase, the low-demand state steadily decreases in-state tuition, and increases in-state capacity and spending per student. Out-of-state tuition remains relatively flat, and out-of-state capacity rises slightly. As spending per student and quality increase, so does demand, and hence, admissions criteria and average ability also increase.

For the high-demand state, as appropriations rise, tuition is first reduced to zero, and in-state capacity flattens out at "full capacity". That is, the level at which all students who wish to attend college are able to do so. Out-of-state tuition and out-of-state capacity remain relatively flat. Where behavior differs from the low-demand state is when state appropriations become sufficiently high, it is more beneficial to restrict out-of-state capacity and lower out-of-state tuition. This causes a large increase in admissions criteria and hence peer effects, which dominate the decrease in spending per student, and so quality continues to rise. Out-of-state tuition and capacity then remain at this new level, and all further increases in state appropriations are directed toward increasing spending per student.

For this calibrated model, both types of state governments find it optimal to simultaneously increase spending per student and capacity, and decrease tuition as revenues increase. However, this is not true in all regions of the parameter space. For example, under a different parameterization, it may be optimal for a policymaker to hold spending per student constant until tuition has been brought to zero. I now discuss in more detail the key trade-offs and parameters involved in allocating an additional dollar of college funding. This determines the relative rate at which these objects change. I largely ignore a discussion of general equilibrium changes in wages and housing rental rates, as these significantly complicate the analysis.

Consider directing a marginal amount of additional revenue toward spending per student, m_ℓ. The first-order effect is to increase quality. Increased college quality has several effects. First, it increases earnings for those who attend the given college. Increased earnings directly affect the welfare of those attending the college. It also increases output (and reduces taxes) in the given state, improving welfare for those not attending the college as well. The key model parameter here is θ, the elasticity of earnings with respect to spending per student. A higher θ increases the relative benefit of directing funds toward spending per student.

A second effect of increased spending per student is increased demand for the college. This raises admissions cutoffs and, hence, average ability, which further increases quality. Importantly, increased demand also occurs for out-of-state students. Given complementarities between ability and college quality in the human capital production function, this attracts higher-ability students from out-of-state to attend the given college. A portion of these high-ability students then remain in-state to work after graduating. *Figure E.25* shows that the average ability of college-educated laborers in a given state increases monotonically as spending per student is exogenously increased. Importantly, however, *Figure E.25* reflects an exogenous increase in spending per student, without respecting the state budget constraint. When spending per student increases, it must come at the cost of either higher in-/out-of-state tuition or lower in-/out-of-state capacity.

Suppose that the higher spending per student comes at the cost of a *ceteris paribus* decrease in in-state capacity. The first-order effect is simple, "one" fewer student is educated, but all other students accumulate more human capital over their lifecycles. *Figure E.26* shows that decreasing capacity also raises the average ability of the student body, due to increased admissions cutoffs.

The key parameters here are those that control the convexity of the college cost function (κ_2) relative to the concavity of the human capital production function (ν). For high (low) capacity and low (high) spending per student, both the cost function and the human capital production function are steep (flat). Hence, (in partial equilibrium) it is more beneficial to direct existing resources toward spending per student. *Figure E.21* shows welfare as a function of in-state capacity (bottom x-axis) and the spending per student (top x-axis). Welfare increases monotonically until it reaches a peak, at which point the state policymaker is indifferent between increasing spending or increasing capacity. The welfare function is plotted for the model in general equilibrium, but the same pattern is also observed in partial equilibrium.

Suppose instead that the higher spending per student comes at the cost of a *ceteris paribus* increase of in-state tuition. This causes demand to fall for all types of students, and some students to be "priced out" of college. In particular, there will be high-ability but low-income students who are no longer able to attend college. *Figure E.27* plots average ability at a college as in-state tuition prices are exogenously increased. For sufficiently low demand (i.e., open admissions) and high tuition prices, the mass of those attending college also begins to fall.

Important parameters here are the mean and variance of the distribution of parental resources, μ_p and σ_p. These parameters determine the rate at which students are no longer able to afford college as tuition increases. *Figure E.22* plots welfare as a function of in-state tuition (bottom x-axis) and the spending per student (top x-axis). Again, the function increases monotonically until an indifference point is reached at the peak.

Finally, the policy maker could instead, reduce out-of-state capacity or increase out-of-state tuition. The welfare increases of an out-of-state student is twofold. First, higher average ability of the student body and higher tuition increase college quality. Second, a portion of the out-of-state student remain in-state to work after graduating. As shown in *Section 5.1*, for a high-productivity state, this second benefit is small. For this type of state, they choose tuition and average ability (capacity) for out-of-state students such that the marginal effect on college quality is equal. Moreover, as just discussed, it must be that this marginal effect on quality is equal to the marginal effect of increasing in-state capacity or reducing in-state tuition. For a state with low-productivity there is the additional benefit of a relatively large increase of in-state college workers. This type of state will then set the marginal benefit of an out-of-state student on quality below the marginal benefit of increasing in-state capacity of reducing in-state tuition.

D.1.2 Constrained Federal Solution

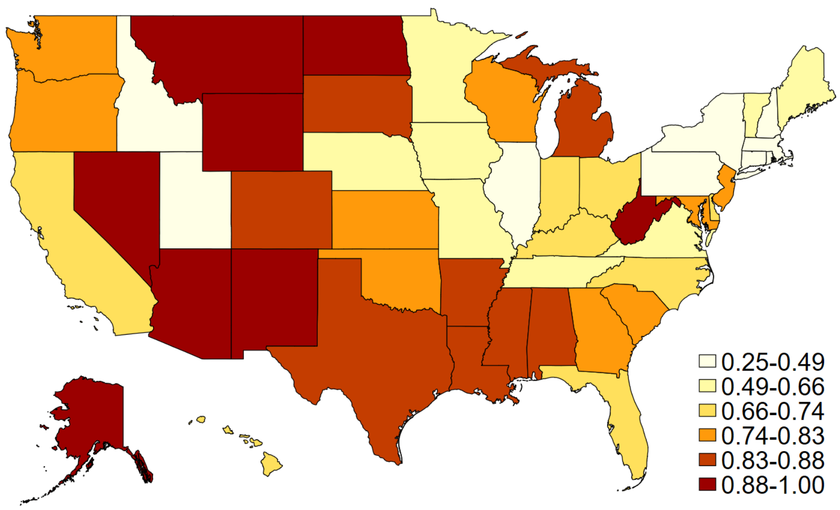
Table E.10 reports welfare changes when the federal problem is constrained to choosing a subset of college policy parameters.[26] *Table E.10* shows that increasing capacity by itself is the most effective, raising welfare by 1.9%, followed by spending per student at 0.61% and tuition at 0.21%.

Tuition is lower on average when chosen in isolation, since it is not accompanied by an increase in capacity. Capacity is also lower when spending per student cannot be reallocated across states. When only spending per student is chosen, it becomes strongly positively correlated with firm productivity. In this case, the efficiency motive dominates redistribution, in contrast to the full counterfactual. Finally, rows four and five show that while tuition alone generates the smallest welfare gain, it interacts strongly with spending and capacity to increase welfare.

[26] For simplicity I group the choice of in-/out-of-state tuition and capacity. When tuition or capacity changes, spending per student also rises or falls. I keep proportional spending per student fixed at the calibrated benchmark equilibrium levels.

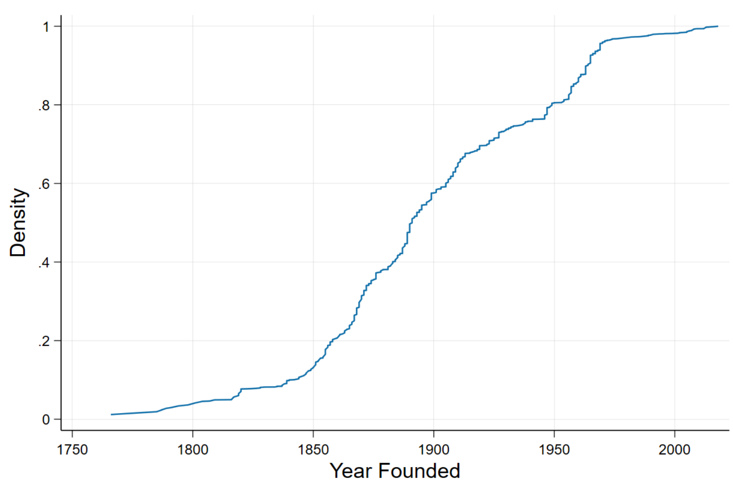
E Additional Figures and Tables

Figure E.1: **Percentage of Total Students Attending Public College**



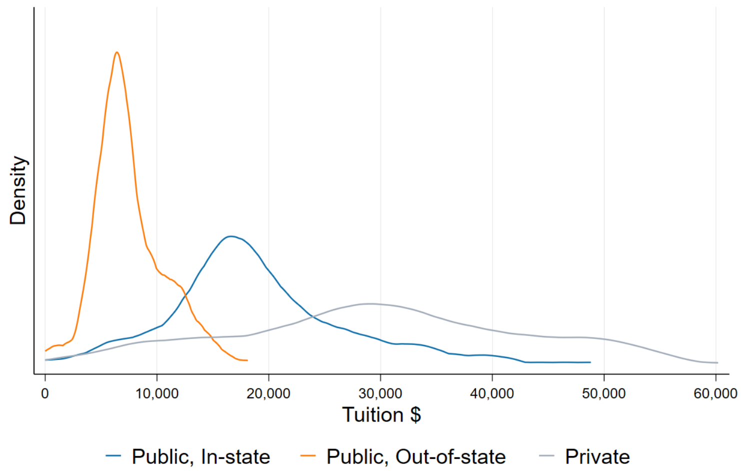
Notes: This figure plots the percentage of total college students enrolled at a public college across states. Data is taken from the Integrated Postsecondary Education Data System (IPEDS) for the 2017-2018 academic year.

Figure E.2: **Cumulative Density Function of College Founding Dates**



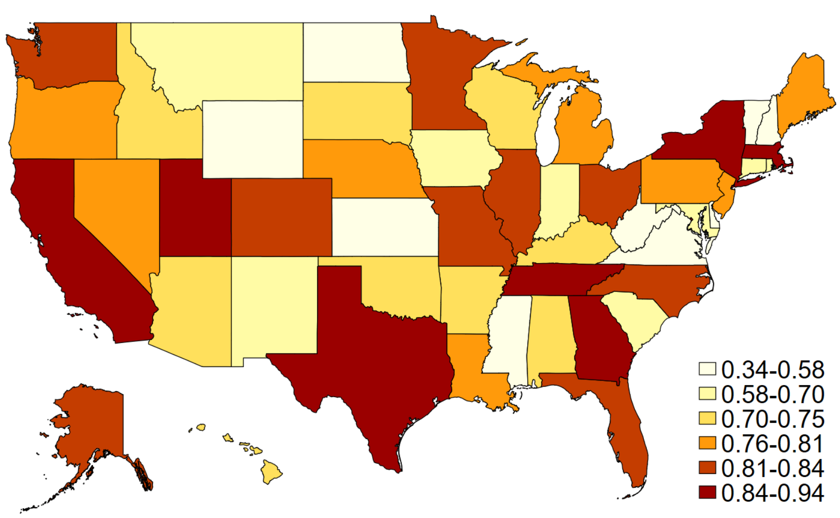
Notes: This figure cumulative percentage of college attendance by founding data. Data are taken from the Integrated Postsecondary Education Data System (IPEDS) for the 2017-2018 academic year for enrolled students and own source for founding dates.

Figure E.3: **Distribution of Tuition by Type**



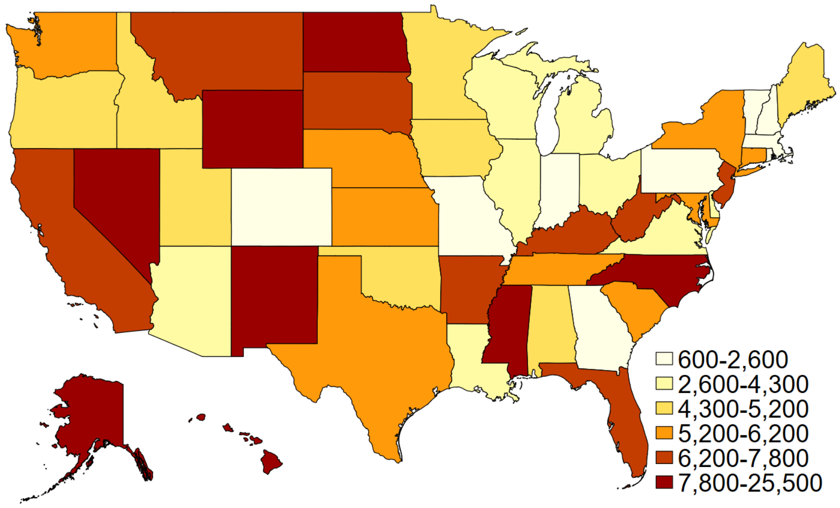
Notes: This figure plots the distribution of tuition by in-state public, out-of-state public, and private. Data are taken from the Integrated Postsecondary Education Data System (IPEDS) for the 2017-2018 academic year.

Figure E.4: **Percentage of Alum Living in State**



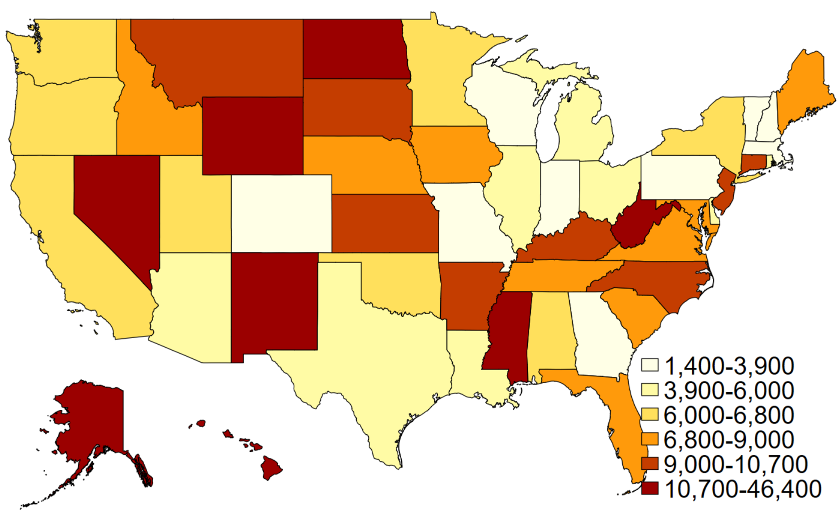
Notes: This figure plots the percentage of college grads working in the same state in which they attended college. Data are from Baccalaureate and Beyond.

Figure E.5: **Cost of Produced College Graduates**



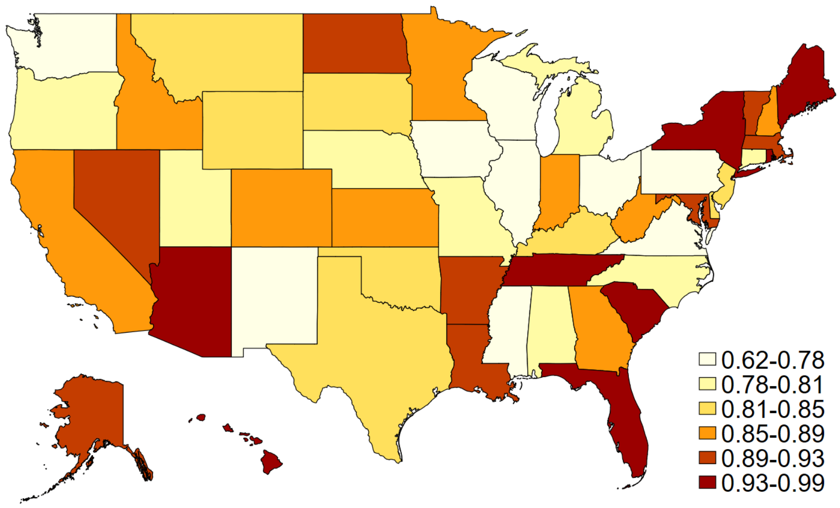
Notes: This figure plots the cost of producing a college graduate. Data open access from Conzelmann et al. (2023).

Figure E.6: **Cost of Retained College Graduates**



Notes: This figure plots the cost of retaining a college graduate. That is, the cost of educating a student that then remains to work in state once graduating. Data are open access from Conzelmann et al. (2023).

Figure E.7: **Percentage of Total Revenues from Tuition or Government Sources**



Notes: This figure plots the percentage of total college revenues received from tuition, or government sources (excluding tuition subsidies and grants). Data are taken from the Integrated Postsecondary Education Data System (IPEDS) for the 2017-2018 academic year.

Table E.1: **Motivating OLS with Controls – State Policies and Population**

	(1) $\ln(T^{out})$	(2) $\ln(T^{in})$	(3) ln(% in-state)	(4) ln(spend per)
ln(*population*)	0.113	0.160	0.132	0.066
	(0.045)	(0.020)	(0.028)	(0.230)
constant	7.92	5.48	-2.08	10.35
	(0.000)	(0.000)	(0.000)	(0.000)
R^2	0.20	0.26	0.54	0.41
Observations	50	50	50	50

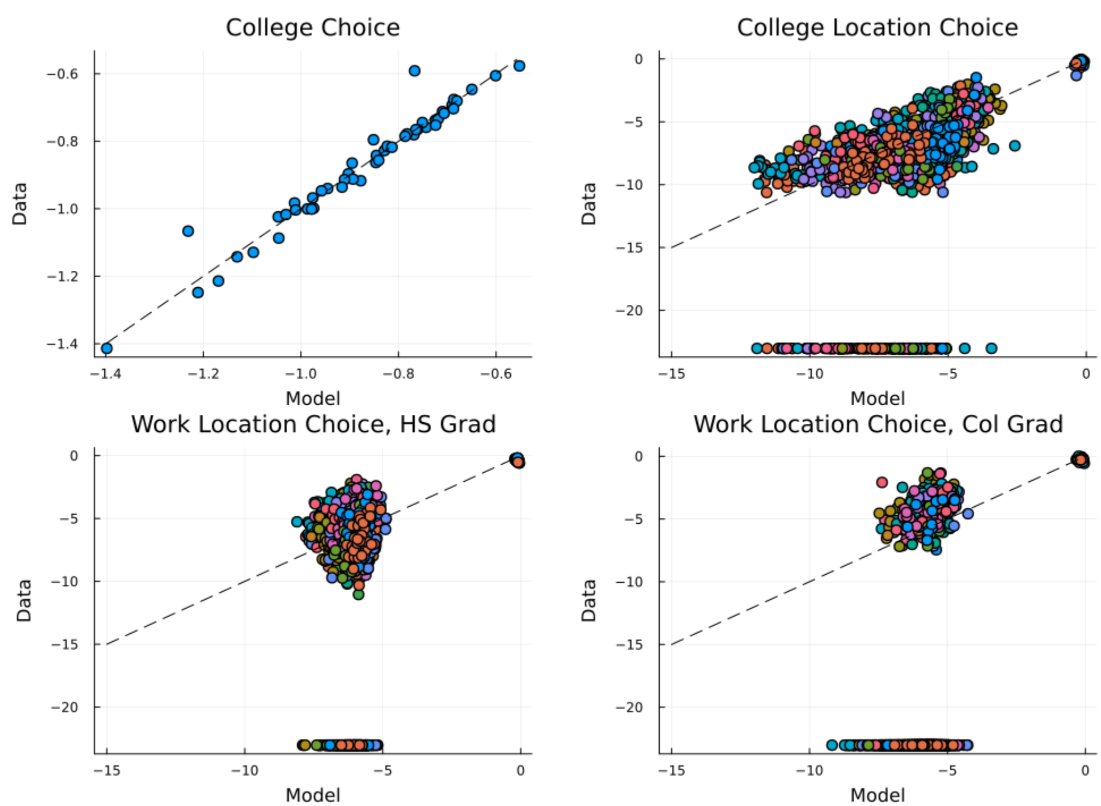
Notes: This table reports the correlation between state population and college policies, while including controls for race, democratic vote share, private college share, and land grant school shares. P-values shown in brackets. Data are taken from the Integrated Postsecondary Education Data System (IPEDS) for the 2017-2018 academic year and the American Community Survey (ACS).

Table E.2: **Motivating OLS with Controls – State Policy and College Wages**

	(1) $\ln(T^{out})$	(2) $\ln(T^{in})$	(3) ln(% in-state)	(4) ln(spend per)
ln(*college wages*)	0.680	0.933	0.517	1.34
	(0.098)	(0.091)	(0.125)	(0.012)
constant	1.70	-2.69	-4.44	-3.58
	(0.773)	(0.731)	(0.251)	(0.531)
R^2	0.22	0.20	0.19	0.36
Observations	50	50	50	50

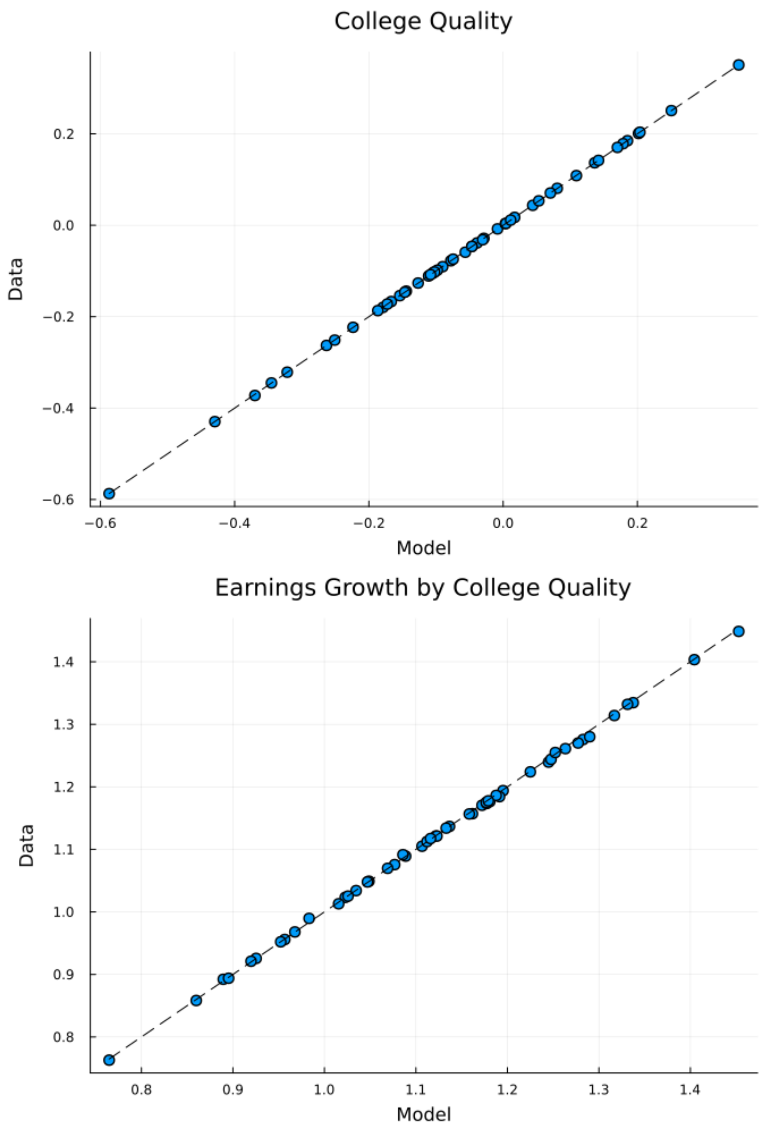
Notes: This table reports the correlation between state college wages and college policies, while including controls for race, democratic vote share, private college share, and land grant school shares. P-values shown in brackets. Data are taken from the Integrated Postsecondary Education Data System (IPEDS) for the 2017-2018 academic year and the American Community Survey (ACS).

Figure E.8: **Model Fit – Choice Probabilities**



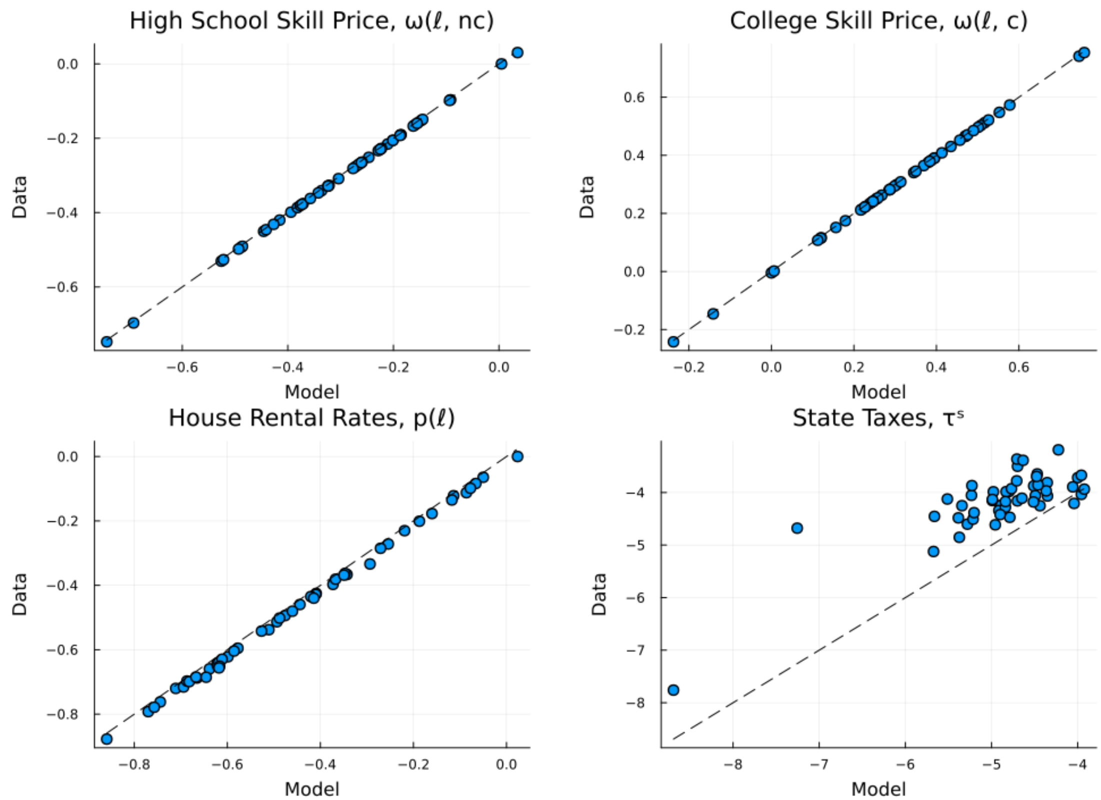
Notes: The model with data choice probabilities for each discrete choice. The 45-degree line represents a perfect fit between the model and data.

Figure E.9: **Model Fit – Earnings Growth and College Qualities**



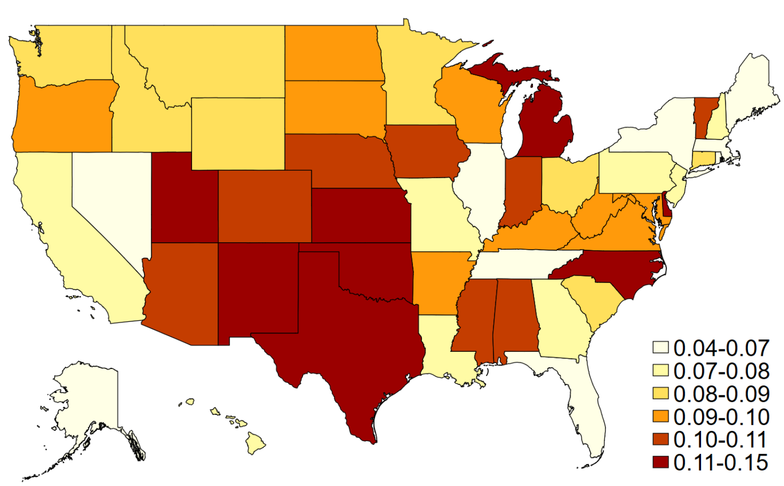
Notes: The model with data choice probabilities for each discrete choice. The 45-degree line represents a perfect fit between the model and data.

Figure E.10: **Model Fit – Skill Prices, Rental Prices, and Tax Rates**



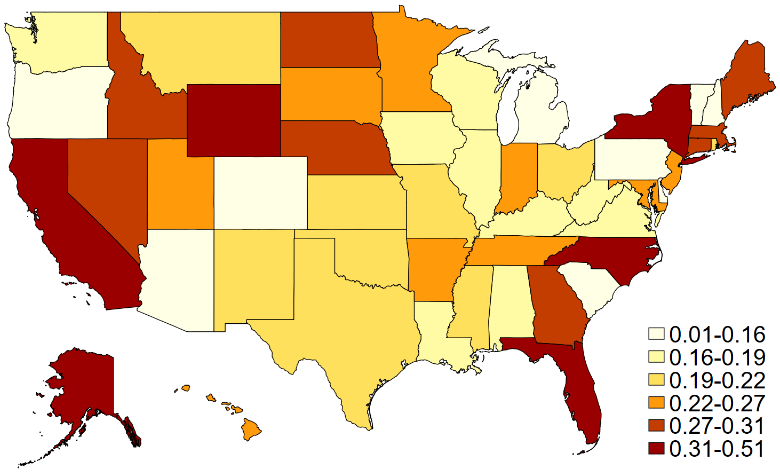
Notes: The model with data choice probabilities for each discrete choice. The 45-degree line represents a perfect fit between the model and data.

Figure E.11: **Higher Education Expenditures as Proportion of Total State Outlays**



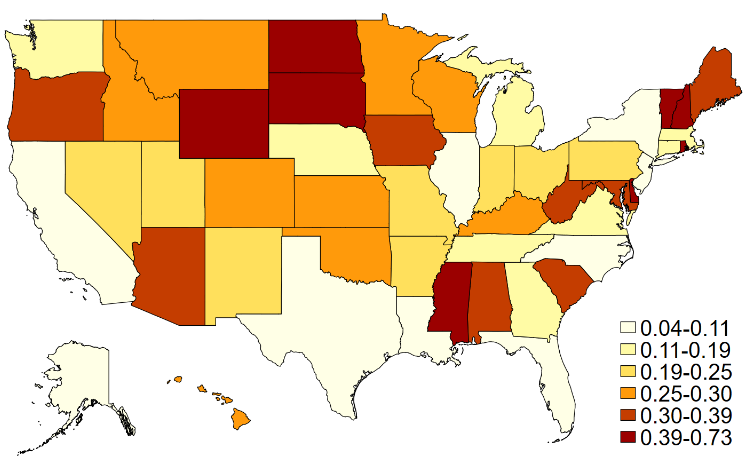
Notes: This figure plots the proportion of total state outlays allocated to higher education across the 50 U.S. states. Data are taken from the U.S. Census Bureau's State and Local Government Finances. Average values are reported for fiscal years 2016-2018 to account for short-run fluctuations.

Figure E.12: **Proportion of College Budget Financed by State Appropriations**



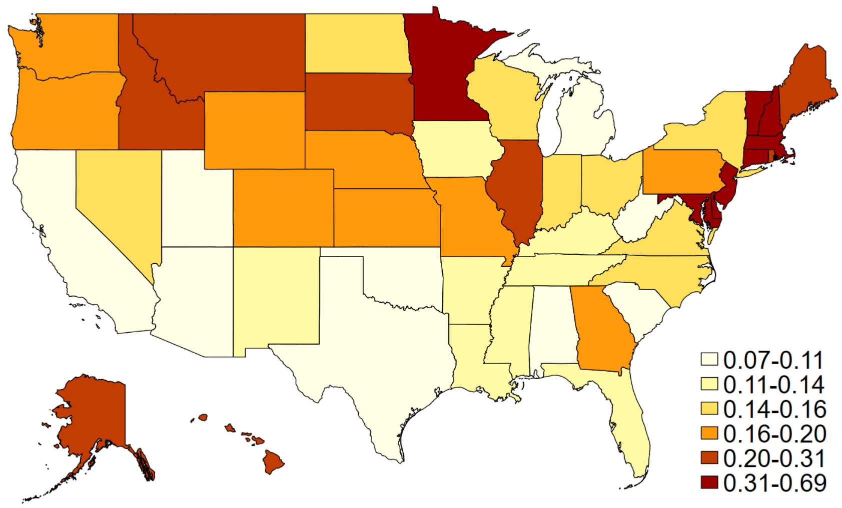
Notes: This figure plots the proportion of total college expenditures financed by state government revenues across the 50 U.S. states. Data are taken from the Integrated Postsecondary Education Data System (IPEDS) for the 2017-2018 academic year.

Figure E.13: **Percentage of Student Body From Out-of-State**



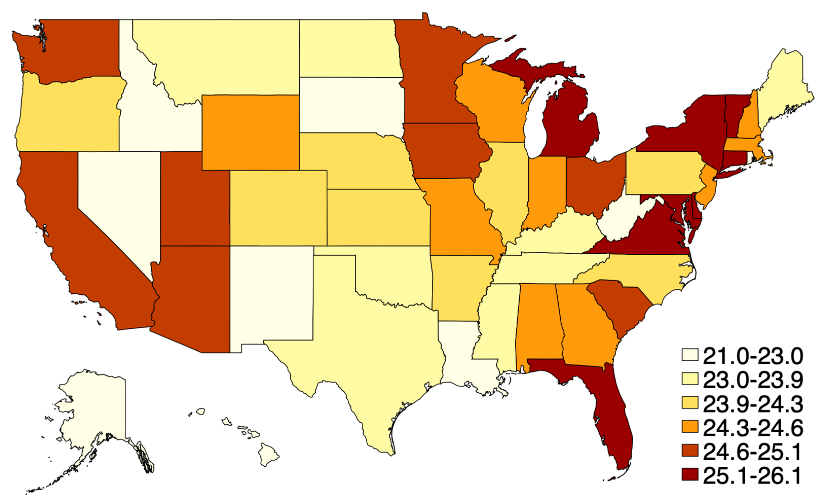
Notes: This figure plots the percentage of out-of-state students across the 50 U.S. states. Data are taken from the Integrated Postsecondary Education Data System (IPEDS) for the 2017-2018 academic year.

Figure E.14: **Percentage of Students Leaving Home State for College**



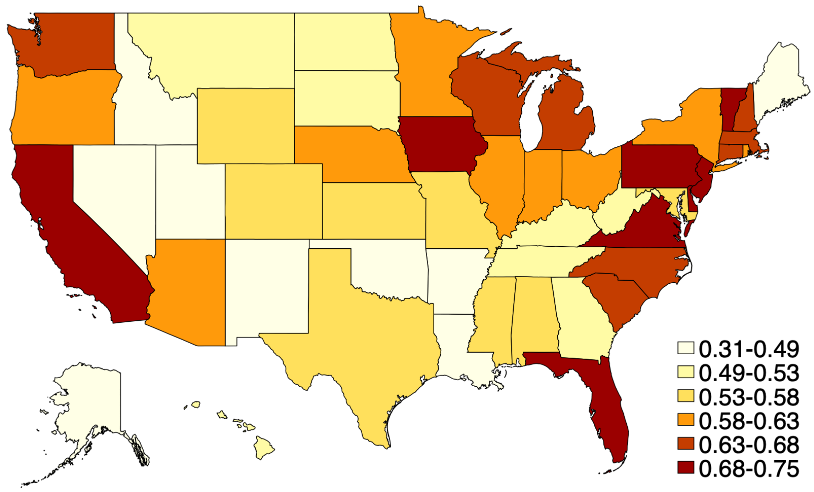
Notes: This figure plots the percentage of high students that study outside state of high school graduation, across states. Data are taken from the Integrated Postsecondary Education Data System (IPEDS) for the 2017-2018 academic year.

Figure E.15: **Median ACT Composite Scores**



Notes: This figure plots the median American College Testing (ACT) composite scores at public colleges across the 50 U.S. states. Composite scores are an average of English, Mathematics, Reading, and Science scores. Data are taken from the Integrated Postsecondary Education Data System (IPEDS) for the 2017-2018 academic year.

Figure E.16: **6-year Graduation Rate**



Notes: This figure plots the 6-year graduation rate at public colleges across the 50 U.S. states. Data are taken from the Integrated Postsecondary Education Data System (IPEDS) for the 2017-2018 academic year.

Figure E.17: **University of Minnesota–Twin Cities Wikipedia Page**

Notes: This figure shows the Wikipedia page of the University of Minnesota–Twin Cities used for scraping founding dates.

Table E.3: **Motivating OLS – Policy Variable Correlations**

	(1) ln(spend per)	(2) ln(spend per)
$\ln(T^{in})$	0.195	
	(0.084)	
$\ln(\%in - state)$		0.381
		(0.043)
constant	9.52	11.35
	(0.000)	(0.000)
R^2	0.06	0.08
observations	50	50

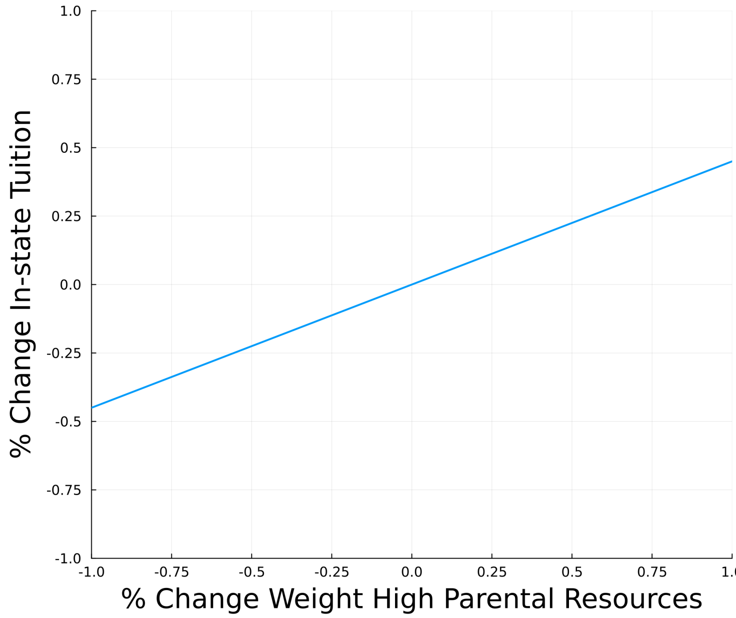
Notes: P-values show in brackets. Data are taken from the Integrated Postsecondary Education Data System (IPEDS) for the 2017-2018 academic year and the American Community Survey (ACS).

Table E.4: **External Validation Estimation and Robustness**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	ln(leave grad)	ln(leave grad)	ln(leave grad)	ln(leave grad)	ln(leave grad)	ln(leave alum)	net import
$\ln(T^{in})$	0.059	0.061	0.056				
	(0.007)	(0.077)	(0.038)				
$ln(col\ earn)$	-0.164	-0.243			-0.178	-1.47	3.31
	(0.146)		(0.016)		(0.004)	(0.002)	(0.002)
$L5.\ln(spend)$	-0.216	-0.211	-0.244	-0.181			
	(0.141)	(0.031)	(0.021)	(0.023)			
$L5.\ln(T^{out})$	-0.134	-0.133					
	(0.150)	(0.053)					
$\%\ out-state$	2.01	2.01	2.08	2.10	2.09	-2.83	-2.83
	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$admit\ rate$	-0.493	-0.503	-0.480	-0.525	-0.545		
	(0.095)	(0.022)	(0.102)	(0.005)	(0.003)		
$med\ act$	0.008	0.008	0.012	0.017	0.015		
	(0.642)	(0.567)	(0.497)	(0.154)	(0.171)		
$unemp$	0.026	0.025	0.484	0.622	0.436	-0.057	-0.334
	(0.967)	(0.958)	(0.507)	(0.218)	(0.355)	(0.810)	(0.522)
$constant$	1.74	-0.146	1.45	-1.45	-0.461	13.76	-36.44
	(0.260)	(0.786)	(0.318)	(0.013)	(0.566)	(0.010)	(0.002)
IV	No	No	No	Yes	Yes	No	No
Observations	692	692	826	829	829	50	50

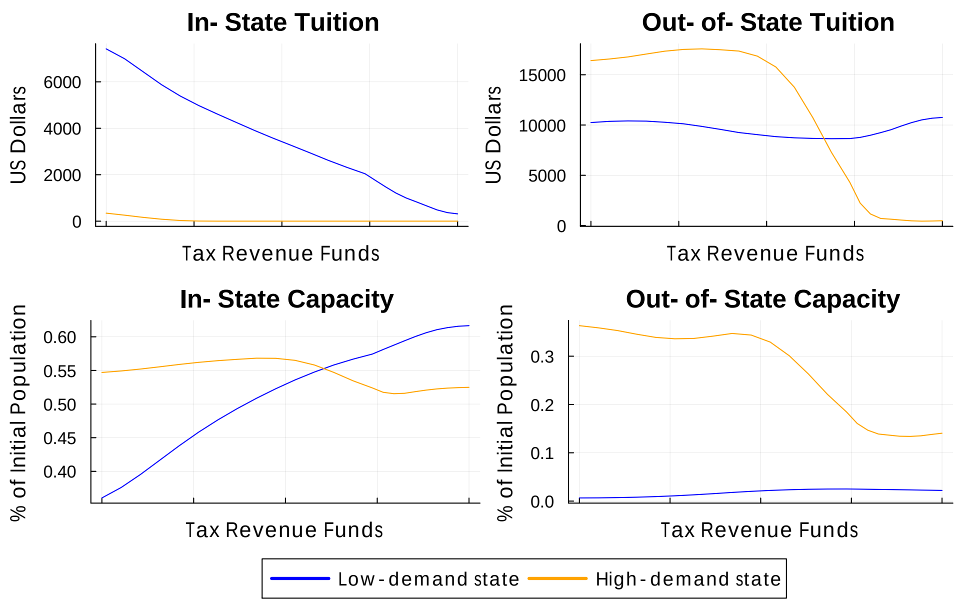
Table E.5: *Notes:* P-values show in brackets. Data are taken from the Current Population Survey, Integrated Postsecondary Education Data System, and Conzelmann et al. (2023). Specifications (1) to (5) feature clustered standard errors and fixed effects.

Figure E.18: **Pareto Weighting and Optimal In-state Tuition**



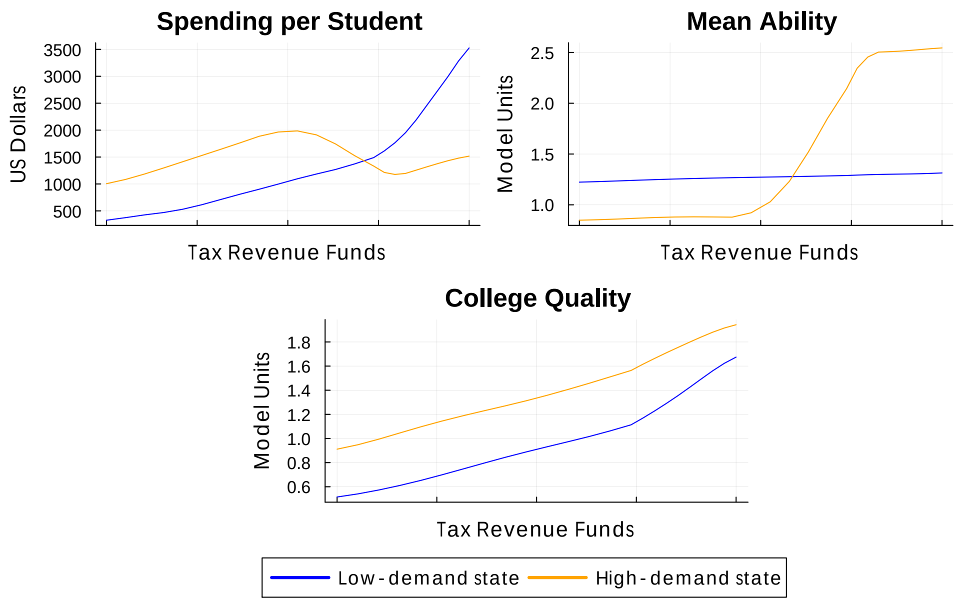
Notes: This figure plots the change in optimal in-state tuition, as the Pareto weights on high-income agents are increased.

Figure E.19: **State Government Solution – Tuition and Capacity**



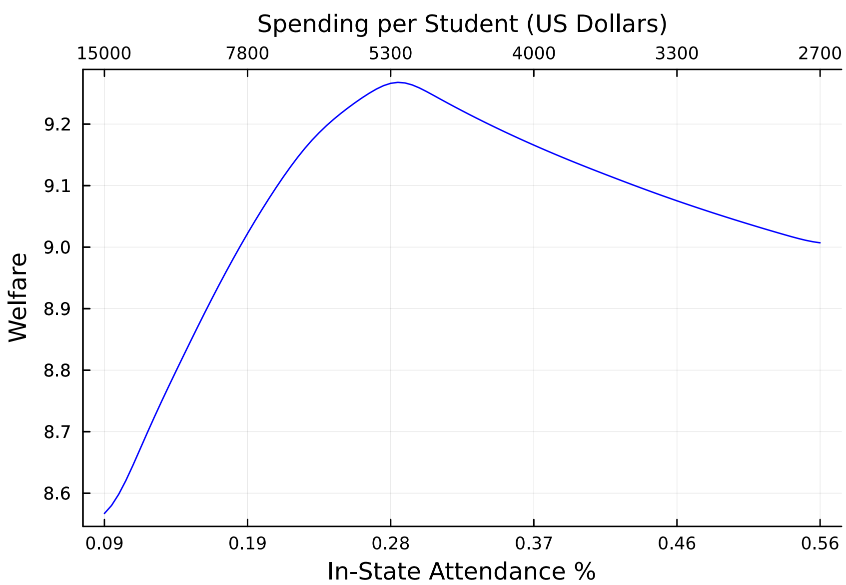
Notes: This figure plots the solution for the optimal policy problem of two example states when college revenues financed through labor taxes are increased.

Figure E.20: **State Government Solution – Spending, Peer effects, and Quality**



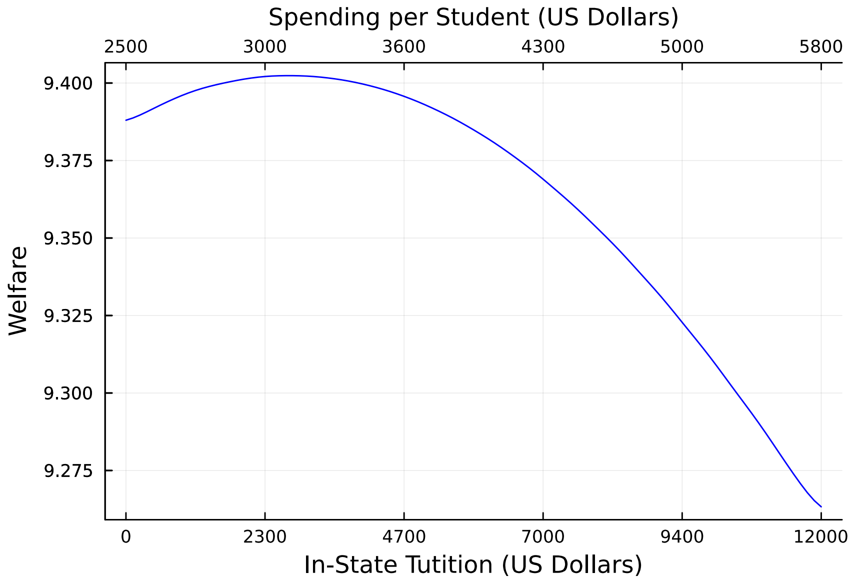
Notes: This figure plots the solution for the optimal policy problem of two example states when college revenues financed through labor taxes are increased.

Figure E.21: **State Welfare – In-State Capacity and Spending per Student**



Notes: This figure plots state welfare as in-state capacity is increased. From the budget constraint this immediately implies a level of spending per student. All else is held fixed. The bottom x-axis shows increases in capacity, and the top x-axis shows corresponding decreases in spending per student.

Figure E.22: **State Welfare – In-State Tuition and Spending per Student**



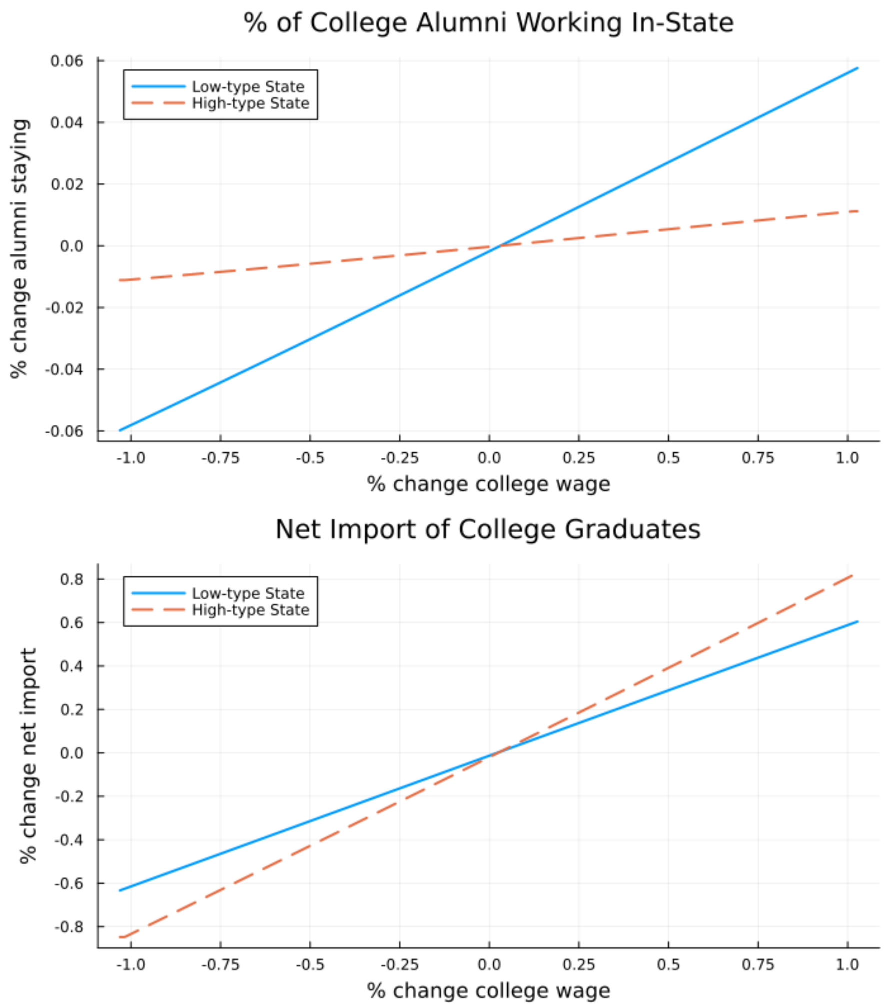
Notes: This figure plots state welfare as in-state tuition is increased. From the budget constraint this immediately implies a level of spending per student. All else is held fixed. The bottom x-axis shows increases in tuition, and the top x-axis shows corresponding increases in spending per student.

Table E.6: **College Policy Heterogeneity in Model and Data**

	Model	Data
In-state tuition	1.52	1.02
Out-of-state tuition	0.24	0.23
Spending per student	0.36	0.32
In-state attendance per HS student	0.22	0.46
Out-of-state attendance per HS student	0.92	0.53

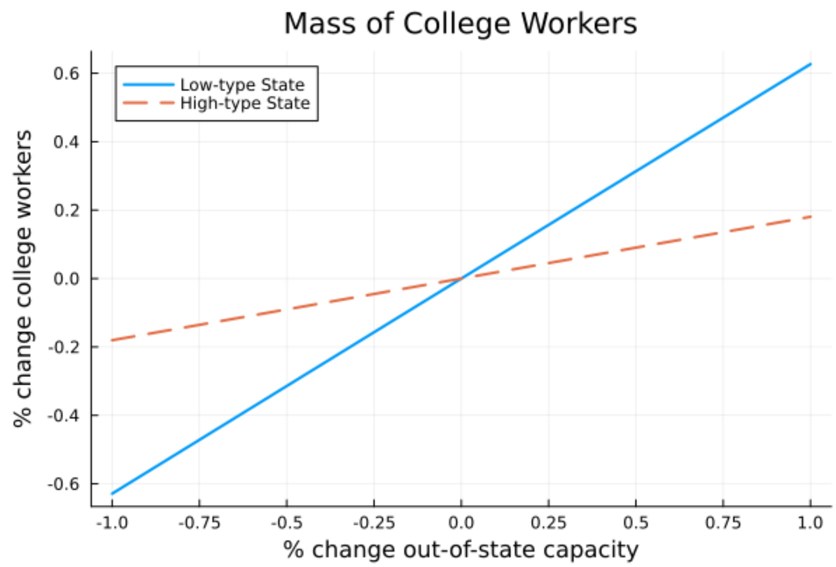
Notes: This table compares heterogeneity in college policies between the model and data. All moments are untargeted.

Figure E.23: **Changes in In-State Alumni and Net College Grad Imports**



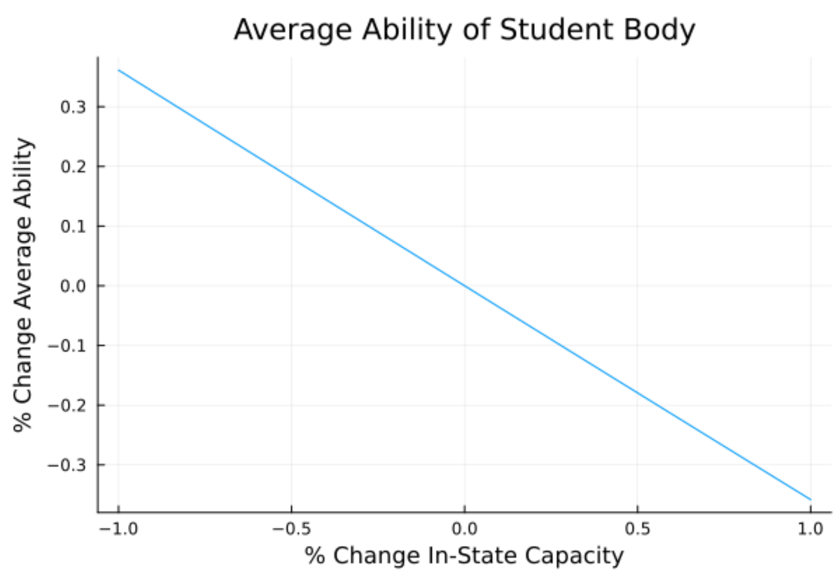
Notes: This figure plots the percentage point change in the percentage of alumni who stay to work in-state (top panel) and the net import of college graduates (bottom panel) by percentage point change in college skill prices.

Figure E.24: **Changes in College Workers and Out-of-State Capacity**



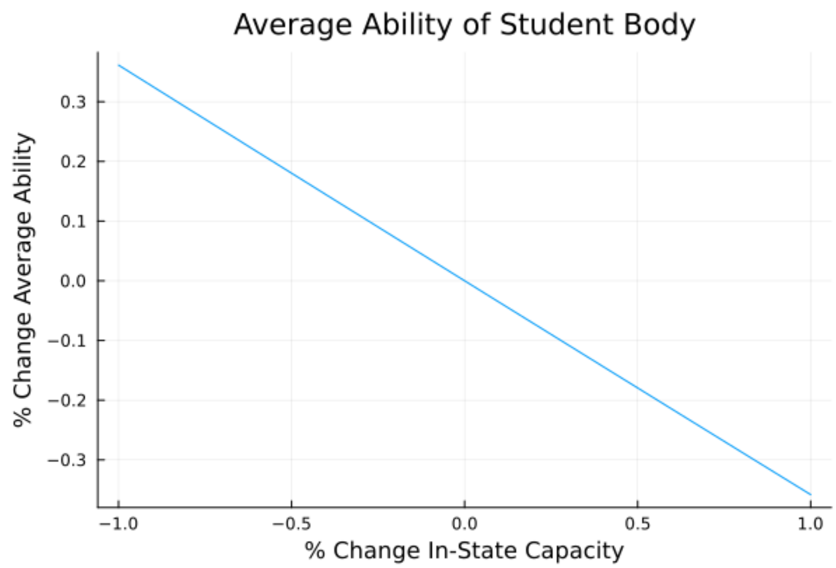
Notes: This figure plots the percentage point change in the percentage of college workers in a state by percentage point change in out-of-state capacity.

Figure E.25: **Changes in Peer Effects and Spending per Student**



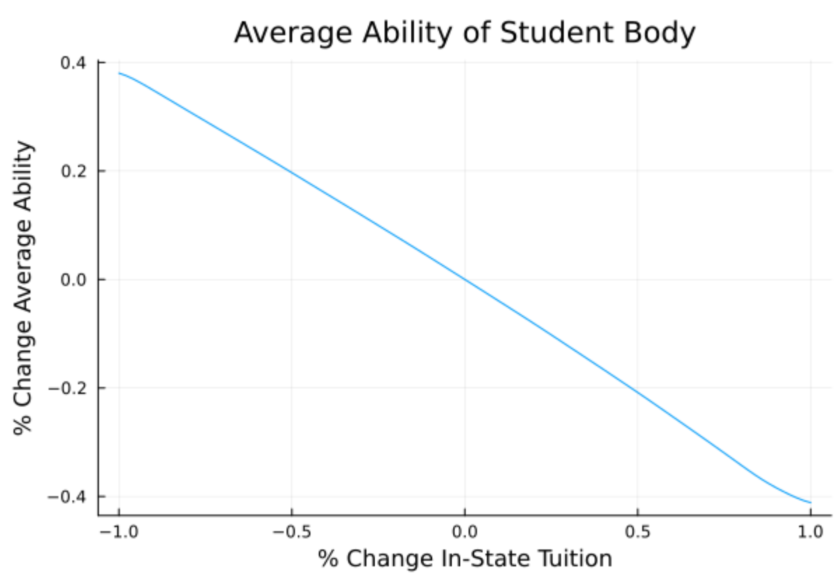
Notes: This figure plots the percentage point change in average ability of college body by percentage point change of spending per student. Spending per student is change exogenously and the government budget constraint in general does not hold here.

Figure E.26: **Changes in Peer Effects and In-State Capacity**



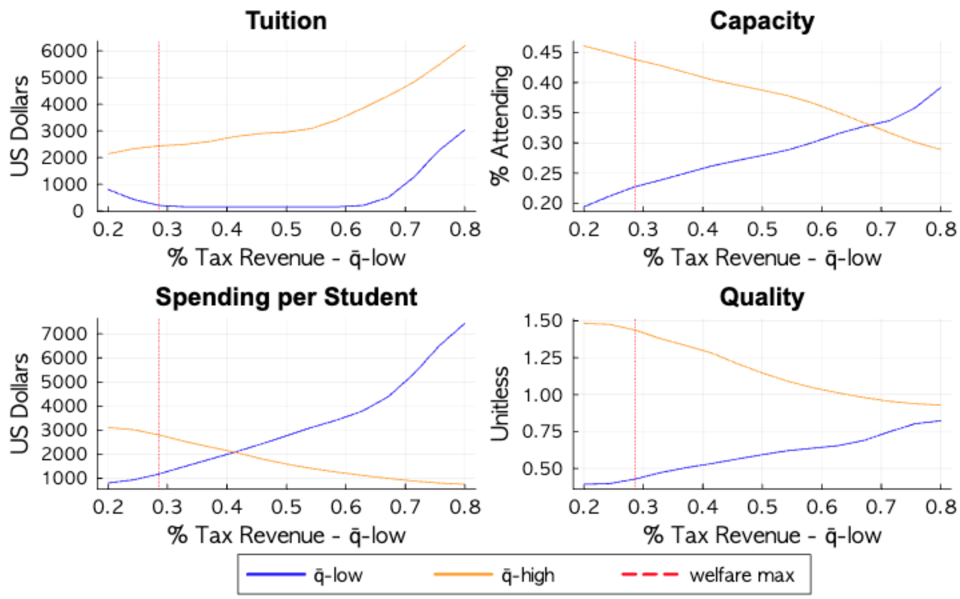
Notes: This figure plots the percentage point change in average ability of college body by percentage point change of in-state capacity.

Figure E.27: **Changes in Peer Effects and In-State Tuition**



Notes: This figure plots the percentage point change in average ability of college body by percentage point change of in-state tuition.

Figure E.28: **Federal Government Solution, Matched Firm-College Productivity**



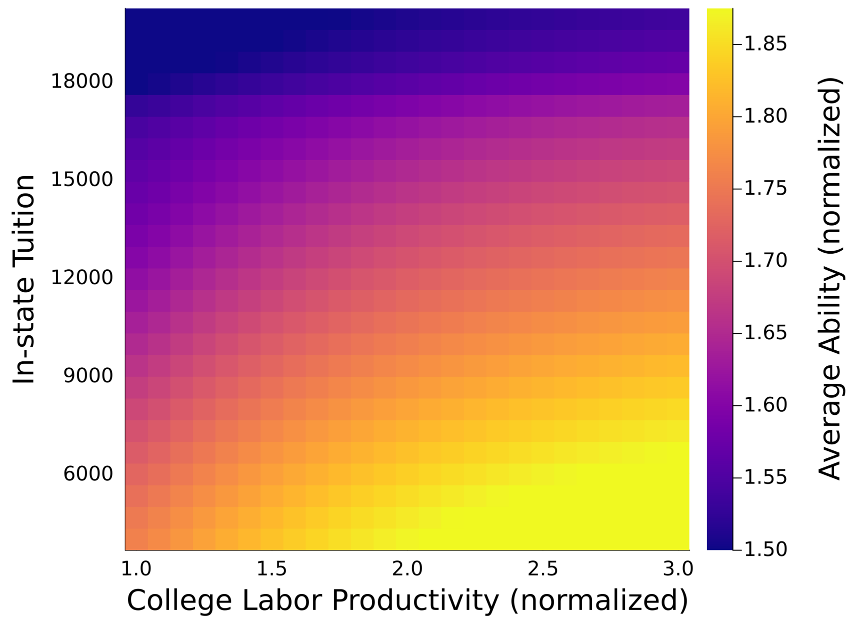
Notes: This figure plots the solution to the federal governments problem for two different states with different productivity. Firm productivity is "matched" with colleges in the sense that high productivity firms are located in the same location as high fixed $\bar{q}$ colleges.

Table E.7: **Pennsylvania Changed to Wyoming**

	T^{in}	T^{out}	% In-state	Spend per
Initial Population	39.3%	30.1%	50.7%	52.0%
Housing stock	0.1%	0.0%	0.1%	0.2%
Firm productivities	2.5%	1.5%	15.2%	3.9%
College productivity	20.1%	25.2%	9.5%	10.5%
Tax revenues	10.2%	1.3%	3.3%	15.2%
Residual (≈ location)	27.8%	41.9%	21.2%	18.2%

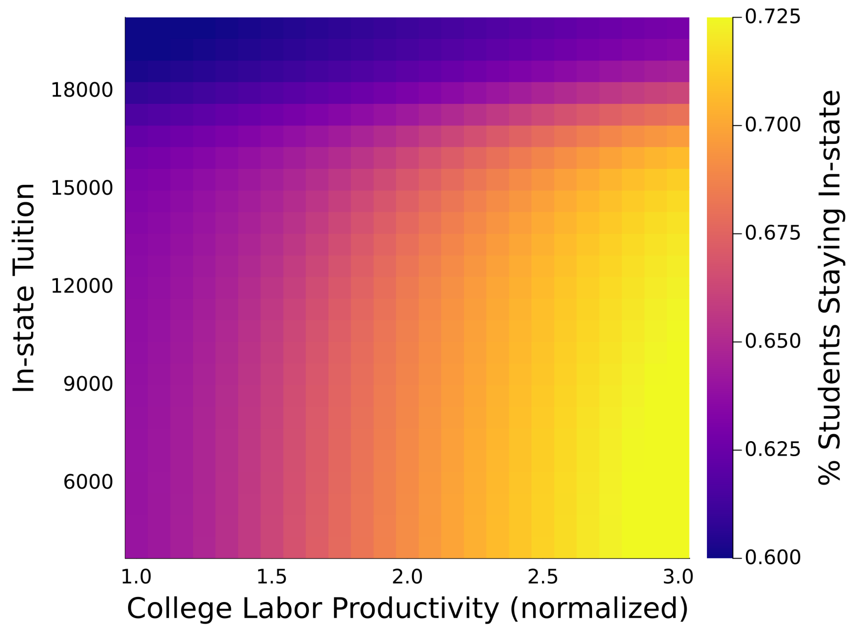
Notes: This table reports a decomposition of college policies between Pennsylvania and Wyoming. Each row changes one additional primitive from Pennsylvania to Wyoming. For instance, the first row changes Pennsylvania's population to that of Wyoming, and the second row changes Pennsylvania's population and housing stock to Wyoming's.

Figure E.29: **Mean College Ability to In-state Tuition and Firm Labor Productivity**



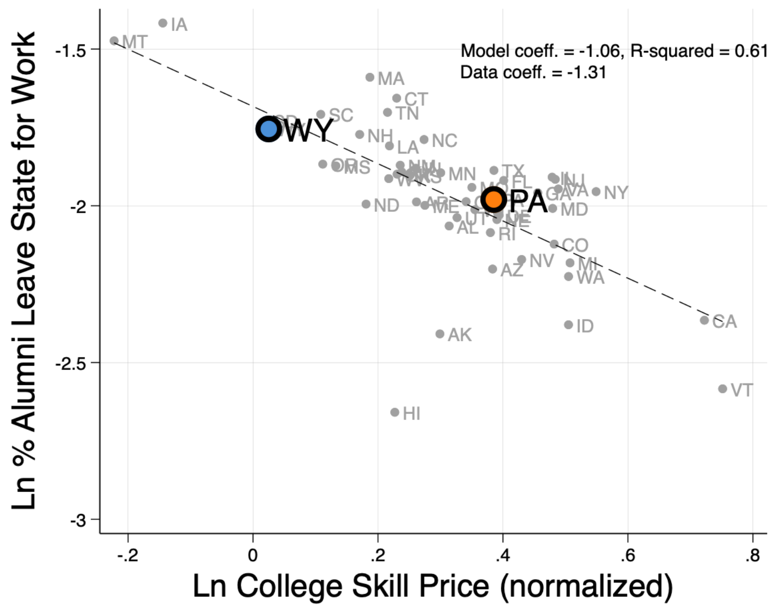
Notes: This figure plots changes in college average ability (heat colors) to changes of in-state tuition and firm college labor productivity.

Figure E.30: **% Students In-state to Out-of-state Tuition and Firm Labor Productivity**



Notes: This figure plots changes in the percentage of student body out-of-state (heat colors) to changes of in-state tuition and firm college labor productivity.

Figure E.31: **Elasticity of Leaving College Graduates to College Skill Prices**



Notes: This figure plots the scatter of the natural log of % of college who leave the state to study elsewhere, to local college skill prices. The dotted black line reports the linear line of best fit, and is compared to the model equivalent.

Table E.8: **Changes in College Policies Between Data and Nash Optimum**

	(1) *norm*	(2) $ln(\Delta\ \%\ in)$	(3) $ln(\Delta\ T^{in})$	(4) $ln(\Delta\ T^{out})$	(5) $ln(\Delta\ spend)$
$\ln(pop)$	-1.09	0.1041	0.800	0.593	-0.255
	(0.008)	(0.000)	(0.001)	(0.076)	(0.000)
$\ln(\bar{A}_c/\bar{A}_{nc})$	-0.16	0.003	0.043	0.017	-0.074
	(0.038)	(0.106)	(0.129)	(0.002)	(0.037)
$\ln(\bar{q})$	-1.35	-0.201	-0.951	-0.960	0.988
	(0.000)	(0.106)	(0.000)	(0.031)	(0.000)
$\ln(appropriations)$	1.19	0.496	-0.618	-0.520	-0.737
	(0.000)	(0.106)	(0.000)	(0.087)	(0.000)
R^2	0.93	0.81	0.95	0.73	0.66

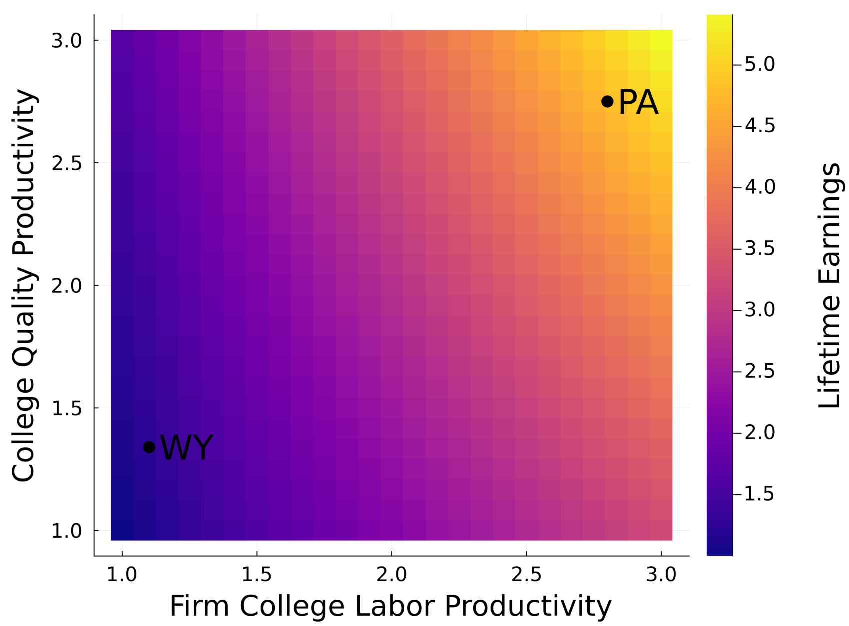
Notes: This table reports the natural log of changes in college policies between the data and decentralized Nash equilibrium. Column (1) reports the distance between the Nash optimum and data, as measure by the $L2 - norm$. Column (2) reports changes in % of students which are in-state, column (3) in-state tuition, column (4) out-of-state tuition, and column (5) spending per student. P-values shown in brackets.

Table E.9: **State Policies and Primitives in Federal Solution**

	T^{in}	T^{out}	% Col.	Spend Rel. Pop.	Spend Rel. Rev.
pop	-347	-318	0.007	-0.003	-0.001
	(0.085)	(0.079)	(0.000)	(0.009)	(0.000)
$\bar{q}$	-11,437	-10,484	0.109	-0.162	0.018
	(0.027)	(0.033)	(0.002)	(0.003)	(0.000)
$\bar{A}_c$	-16,279	-14,922	0.105	-0.158	0.049
	(0.011)	(0.009)	(0.002)	(0.003)	(0.000)
$\bar{A}_c \cdot \bar{q}$	16,660	15,271	-0.135	0.148	-0.028
	(0.099)	(0.087)	(0.007)	(0.004)	(0.000)
constant	22,174	20,326	-0.030	0.195	-0.033
	(0.035)	(0.0341)	(0.045)	(0.000)	(0.000)

Notes: This table reports state primitives regressed against optimal federal solutions for each state. Population in 100,000. $\bar{q}$ and $\bar{A}_c$ demeaned. P-values shown in brackets.

Figure E.32: **Lifetime Earnings and State Primitives**



Notes: This figure plots lifetime earnings for changes in-state college quality productivity and firm college labor productivity. All variables are normalized.

Table E.10: **Welfare Changes for Individual College Policies**

	Δ Welfare
Tuition only	0.21%
Capacity only	1.9%
Spending only	0.61%
Spending and tuition	0.92%
Spending and capacity	2.4%

Notes: This table reports aggregate welfare changes when only an individual college policy can be chosen. Changes in spending per student are changed in fixed proportions to the full optimal federal solution.

Table E.11: **Welfare Changes from Centralized College System by State Primitives**

	(1) $\Delta\, Welfare$	(2) $\Delta\, Welfare$
A_u	-0.003	0.0101
	(0.090)	(0.077)
$\bar{q}$	-0.042	0.082
	(0.091)	(0.097)
$A_u \times \bar{q}$		-0.018
		(0.075)
pop	0.030	0.043
	(0.069)	(0.016)
A_s	-0.033	-0.022
	(0.056)	(0.042)
$constant$	1.10	1.0
	(0.000)	(0.000)

Notes: This table reports average welfare changes by state primitives. Variables are normalized relative to the mean.

Table E.12: **One-shot Profitable Deviation from Centralized Solution**

	ΔT^{in}	ΔT^{out}	ΔC^{in}	ΔC^{out}	Δq
Pennsylvania					
Iteration 1	-21%	+112%	+61%	-78%	+26%
Wyoming					
Iteration 1	-29%	-18%	+9%	+44%	-19%

Notes: This table reports changes in college policies for Wyoming and Pennsylvania from a profitable one-shot deviation over the federal solution.