From Classroom to Prosperity: Fostering Development Through Higher Education

Alvaro Cox

Yale University

April 16, 2024

Introduction

- Is education an engine or a consequence of development?
  - If reducing access costs boosts development, what are the mechanisms?
- Previous research:
  - Struggles with causality, focuses on elementary education.
  - Ignores the role of the firm, GE, and spillovers from large-scale policies.
- **This paper:** education reform from Brazil, novel mechanism and structural model:
  - New evidence on the causal effect of rising college education on development.
  - New mechanism: reducing access costs to college helps firms to expand.
  - Structural model to quantify the benefits of reducing access costs at the region level.

Roadmap

1. Reform and rich data to assess the development consequences of rising college attainment.
   - Exploit heterogeneous effects of reform on college attainment across labor markets.
   - Development: increase in income and white-collar occupations, decline in agricultural employment.
   - New mechanism: increase in the prevalence of large firms.

Roadmap

1. Reform and rich data to assess the development consequences of rising college attainment.
   - Exploit heterogeneous effects of reform on college attainment across labor markets.
   - Development: increase in income and white-collar occupations, decline in agricultural employment.
   - New mechanism: increase in the prevalence of large firms.
2. New structural model.
   - Connects college attainment to aggregate productivity gains through firms' expansion.
   - College attainment, firm size distribution, and development jointly determined.

Roadmap

1. Reform and rich data to assess the development consequences of rising college attainment.
   - Exploit heterogeneous effects of reform on college attainment across labor markets.
   - Development: increase in income and white-collar occupations, decline in agricultural employment.
   - New mechanism: increase in the prevalence of large firms.
2. New structural model.
   - Connects college attainment to aggregate productivity gains through firms' expansion.
   - College attainment, firm size distribution, and development jointly determined.
3. Quantify the effect of increasing college attainment on Brazil's 2000-2010 growth experience.
   - Structural estimation of the model and validation using empirical evidence from the reform.
   - Reducing education costs drives 18% of GDPpc growth and 7% of decline in agricultural labor.

Related Literature and Contribution

- The Role of Human Capital in Economic Growth
  - Nelson and Phelps (1966), Lucas Jr (1988), Bils and Klenow (2000), Caselli and Coleman II (2001), Galor and Moav (2004), Aghion et al. (2009), Kaboski (2009), Jones (2014), Manuelli and Seshadri (2014), Buera and Kaboski (2012), Buera, Kaboski, et al. (2022), Porzio, Rossi, and Santangelo (2022). **Within-country:** Duflo (2001), Hsiao (2022), Akresh, Halim, and Kleemans (2023), Khanna (2023), Nimier-David (2022).
  - Use large-scale education policy to obtain causal effects of higher education on local development.
  - Novel model connecting college attainment to aggregate productivity born from firm-level expansion.
  - Model validation using empirical analysis as non-targeted moments.
- Human Capital and Firms.
  - Chandler (1977), Bloom et al. (2013), Bloom et al. (2014), Hjort and Poulsen (2019), Akcigit et. al. (2021), Hjort et al. (2022).
  - Evidence on firm expansion from increasing the supply of workers with a college degree.

Outline

Outline

Institutional Context: Higher Education in Brazil

Institutional Context: Higher Education in Brazil

Private Colleges

Federal Council of Education

*Professional Organizations

Non-For-Profit Colleges

Opening of Colleges & Courses

Institutional Context: Higher Education in Brazil

Private Colleges

Federal Council of Education

*Professional Organizations

Non-For-Profit Colleges

Opening of Colleges & Courses

INEP

For-Profit Colleges

Flexible Entry

ENADE

5yr-Accreditation Cycle

Institutional Context: Higher Education in Brazil

Private Colleges

Federal Council of Education

*Professional Organizations

Non-For-Profit Colleges

Opening of Colleges & Courses

INEP

For-Profit Colleges

Flexible Entry

ENADE

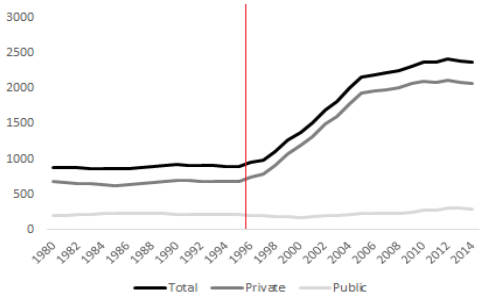
5yr-Accreditation Cycle

Public Colleges

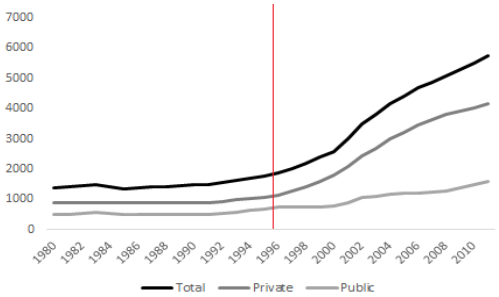
Elite Institutions with limited slots

Expansion in Private Colleges and Enrollment

(a) Number of Colleges

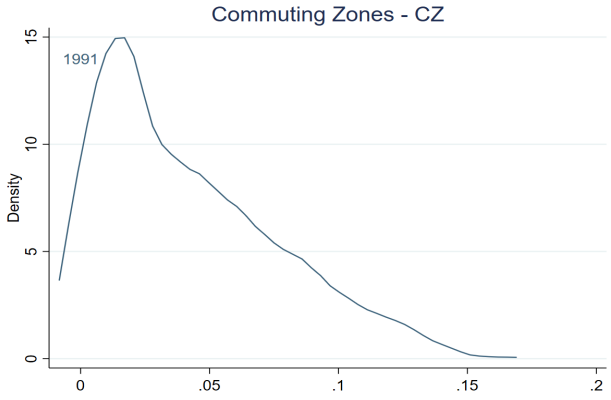


(b) Enrollment



▸ Other HE Policies

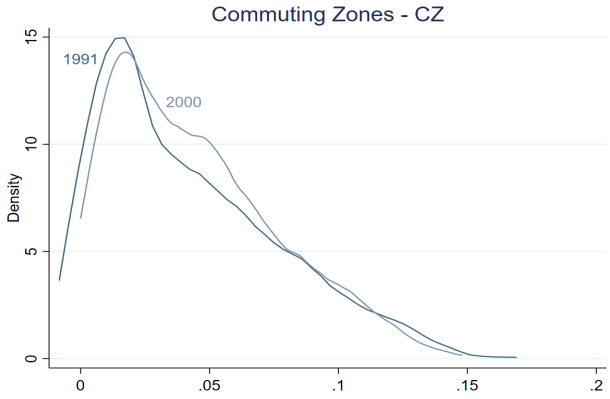
Expansion in College Attainment of the Youth (25-34 years old)



Share of Youth with College Degree

▸ Municipalities ▸ Reasons Expansion

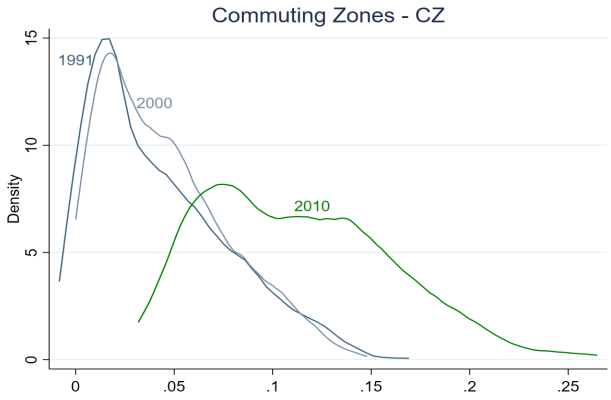
Expansion in College Attainment of the Youth (25-34 years old)



Share of Youth with College Degree

▸ Municipalities ▸ Reasons Expansion

Expansion in College Attainment of the Youth (25-34 years old)



Share of Youth with College Degree

▸ Municipalities ▸ Reasons Expansion

Characteristics of New College Graduates

- **Migration to Enrollment**
  - 95% of college students from private (public) colleges travel at most 85km (221km) to enroll in college.
- **Majors**
  - Business Administration, Accounting, Law, Medicine, and Education.
- **Quality of new colleges**
  - Distribution of value-added in for-profit (new) and non-for-profit colleges (old) is similar.
- **Employment Profile**
  - Sectors: Services. Occupations: Administrative and Professionals. Job Status: 75% Formal Jobs.

Sector—Occupation Allocation | Major composition | Formality by major | Quality provided | Migration to Enrollment

Outline

The Effect of College Attainment on Local Development, 2000 - 2010

- Empirical Specification:

$$\Delta Y_r = \beta \Delta h_r + X_r \gamma + \theta_r^s + \epsilon_r$$

- Rise in college attainment (Δh_r) and outcomes (ΔY_r) include:
  - Change in average income growth, change in the employment share of agriculture and white collar occupations, growth in the number of firms.

- **Identification concerns:** Higher education investment in regions with better growth prospects.

- Identification Strategy:

  1. Control for observable growth prospect: 1991 Development and human capital, and pre-trends 1991-2000.
  2. Instrumental variable and Diff-in-Diff:

▸ Data used ▸ Controls

Empirical Strategy: Instrumental Variable Approach

- Instrument: Relative scarcity of Colleges **before the reform:**

$$Z_r = \frac{\#Colleges_{r,1991}}{\#\text{20-29yo with High School but no College}_{r,1991}}$$

- After the reform, private colleges entered those regions where colleges were relatively more scarce.
- Exclusion Restriction:
  - Conditional on X_r, post-reform, locations where colleges were relatively scarce (lower Z_r) had better development outcomes only because they expanded more college attainment.

College attainment, income, and structural transformation

Table: Effect of rising college attainment on development

Dep. Variable	Income Growth (logs)			Change Share Working in Agriculture			Change Share Working in White-Collar Occ.		
Panel A: All Cohorts									
Δ h	0.43**	0.62***		-0.20	-0.54***		0.35***	0.51***	
	(0.18)	(0.19)		(0.15)	(0.04)		(0.05)	(0.02)	
Panel B: Young Cohorts									
Δ h	0.99**	0.93***		-0.23	-0.53***		0.80***	0.88***	
	(0.39)	(0.26)		(0.15)	(0.04)		(0.03)	(0.01)	
Obs.	486	486		486	486		486	486	
Controls	No	Yes		No	Yes		No	Yes	
Specification	OLS	OLS	IV	OLS	OLS	IV	OLS	OLS	IV

Controls 1991 initial development: average income, population, agricultural and high-skill service employment share, high school attainment, college attainment, and had college dummy. Pre-trend 1991-2000: income, population growth, change in agricultural and high-skill employment. Uses population weights.

▸ First Stage ▸ Population weights

College attainment, income, and structural transformation

Table: Effect of rising college attainment on development

Dep. Variable	**Income Growth (logs)**			**Change Share Working in Agriculture**			**Change Share Working in White-Collar Occ.**		
Panel A: All Cohorts									
Δ h	0.43**	0.62***	0.58	-0.20	-0.54***	-0.61**	0.35***	0.51***	0.57***
	(0.18)	(0.19)	(0.38)	(0.15)	(0.04)	(0.28)	(0.05)	(0.02)	(0.11)
Panel B: Young Cohorts									
Δ h	0.99**	0.93***	0.86*	-0.23	-0.53***	-0.62**	0.80***	0.88***	0.88***
	(0.39)	(0.26)	(0.46)	(0.15)	(0.04)	(0.29)	(0.03)	(0.01)	(0.08)
Obs.	486	486	486	486	486	486	486	486	486
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Specification	OLS	OLS	IV	OLS	OLS	IV	OLS	OLS	IV

Controls 1991 initial development: average income, population, agricultural and high-skill service employment share, high school attainment, college attainment, and had college dummy. Pre-trend 1991-2000: income, population growth, change in agricultural and high-skill employment. Uses population weights.

▸ First Stage ▸ Population weights

College attainment increases the prevalence of medium size firms

$$\Delta y_{ir}^{b} = \beta^{b} \Delta h_r + X_r \Gamma + \theta_{ir}^{s} + \theta_i + u_{ir}^{b}$$

Table: Effect of College Attainment on Firms

Dep. Variable	Growth Number of Firms (1996-2010)					
Sample	All Firms		Small Firms L 1-5		Large Firms L>5	
Specification	(1)	(2)	(3)	(4)	(5)	(6)
Δ h	0.29***	0.53	0.22***	0.25	0.26***	0.72***
	(0.06)	(0.36)	(0.05)	(0.32)	(0.04)	(0.23)
R^2	0.701	0.700	0.695	0.695	0.736	0.734
Obs.	486	486	486	486	486	486
Method	OLS	IV	OLS	IV	OLS	IV

Notes: Includes controls for initial and pre-trend development and Industry Fixed Effects.

▸ World Bank Survey ▸ Composition RAIS ▸ College occupations RAIS

College attainment increases the prevalence of medium size firms

$$\Delta y_{ir}^{b} = \beta^{b} \Delta h_r + X_r \Gamma + \theta_{ir}^{s} + \theta_i + u_{ir}^{b}$$

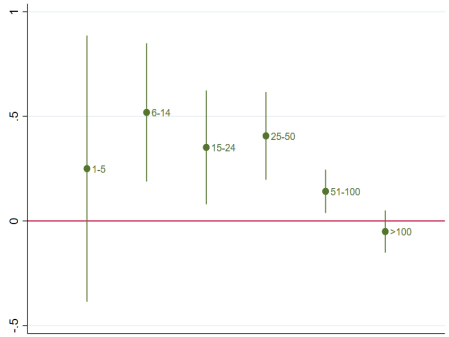
Table: Effect of College Attainment on Firms

Dep. Variable	Growth Number of Firms (1996-2010)					
Sample	All Firms		Small Firms L 1-5		Large Firms L>5	
Specification	(1)	(2)	(3)	(4)	(5)	(6)
Δ h	0.29***	0.53	0.22***	0.25	0.26***	0.72***
	(0.06)	(0.36)	(0.05)	(0.32)	(0.04)	(0.23)
R^2	0.701	0.700	0.695	0.695	0.736	0.734
Obs.	486	486	486	486	486	486
Method	OLS	IV	OLS	IV	OLS	IV

Notes: Includes controls for initial and pre-trend development and Industry Fixed Effects.

▸ World Bank Survey ▸ Composition RAIS ▸ College occupations RAIS

Figure: Growth in Number of Firms by Bin-Size



Notes: Beta coefficients of the IV regression for each bin size.

Exclusion restrictions and taking stock

- DinD analysis shows the effect of the instrument is only significant and relevant for the post-reform period, not before.
- **Taking stock:** rising college attainment increases income, white collar occupations and prevalence of larger firms.
- From relative effects to quantifying aggregate effects.
  - Empirical evidence used as *non-target* moments to validate the model.

Robustness | Diff-in-Diff | Annual RAIS | Ex-post migration | Returns to College

Outline

A Theory of Higher Education and Local Development

Assumptions:

- Static model of R locations that trade among each other.
- Two types of agents, youth (ψ) and old ($(1-\psi)$). Only youth invest in education.
- Two types of workers, skilled (H) and unskilled (L), are immobile across space.
- Preferences are homothetic and defined over two sectoral aggregates of regional varieties.

A Theory of Higher Education and Local Development

Assumptions:

- Static model of R locations that trade among each other.
- Two types of agents, youth (ψ) and old ($(1-\psi)$). Only youth invest in education.
- Two types of workers, skilled (H) and unskilled (L), are immobile across space.
- Preferences are homothetic and defined over two sectoral aggregates of regional varieties.

Main Components of the Model:

I. **College Choice**

  - Decision based on region-specific returns to college and access costs.

A Theory of Higher Education and Local Development

Assumptions:

- Static model of R locations that trade among each other.
- Two types of agents, youth (ψ) and old ($(1-\psi)$). Only youth invest in education.
- Two types of workers, skilled (H) and unskilled (L), are immobile across space.
- Preferences are homothetic and defined over two sectoral aggregates of regional varieties.

Main Components of the Model:

I. **College Choice**

- Decision based on region-specific returns to college and access costs.

II. **Production**

- Two tradable sectors: Agriculture (A) and Non-Agriculture (NA).
- The NA sector is endogenously determined by firms heterogeneous in productivity.
- Skilled worker intensity increases with firm size (Scarcity of skill induces scarcity of larger firms).

College Choice

- Individuals heterogeneous in ability $\epsilon_i \sim F(.)$ choose to get a degree based on prospective income.

$$e^i = \begin{cases} w_r^h \epsilon_i - w_r^h \mathbf{C_r}, & \text{if College Graduate} \\ w_r^l, & \text{otherwise} \end{cases}$$

College Choice

- Individuals heterogeneous in ability $\epsilon_i \sim F(.)$ choose to get a degree based on prospective income.

$$e^i = \begin{cases} w_r^h \epsilon_i - w_r^h \mathbf{C_r}, & \text{if College Graduate} \\ w_r^l, & \text{otherwise} \end{cases}$$

- Regional college costs $\mathbf{C_r}$, college skill premium $\bar{\omega} = \frac{w^h}{w^l}$, and college attainment of the youth $h(\bar{\omega}_r, \mathbf{C_r}) = 1 - F\left(\frac{1}{\bar{\omega}_r} + \mathbf{C_r}\right)$. Supply of college skill in efficiency units:

$$H_r^s = \left(\underbrace{\psi_r E\left[\epsilon_i | \epsilon_i \geq \frac{1}{\bar{\omega}_r} + \mathbf{C_r}\right] h(\bar{\omega}_r, \mathbf{C_r})}_{\text{Youth College Skill}} + \underbrace{(1-\psi_r) h_r^o}_{\text{Old College Skill}} - \underbrace{\psi_r \mathbf{C_r} h(\bar{\omega}_r, \mathbf{C_r})}_{\text{Local Costs}} \right) N_r$$

- Back

Production

- NA Sector: Heterogeneous firms in $z_i \sim \mathbf{G_r(z)} = \mathbf{1} - \left(\frac{\mathbf{B_r}}{\mathbf{z}}\right)^{\lambda}$, use H, and L_m to produce.

$$y_i = z_i^{1-\alpha-\beta} L_{mi}^{\alpha} \left(\mathbf{\Gamma_r} H_i + \kappa\right)^{\beta}$$

- Where $\alpha + \beta < 1$ captures the notion of limited span of control, and κ helps to determine differences in demand for college workers by firm size.

$$\frac{H_i}{L_{mi}} = \theta_0(\underbrace{\bar{\omega}_r}_{-}) - \kappa\theta_1(\underbrace{\Gamma_r}_{-}, \underbrace{\bar{\omega}_r}_{+})\frac{z_i^{\theta_2}}{\bar{y}^{\theta_3}}$$

Production

- NA Sector: Heterogeneous firms in $z_i \sim \mathbf{G_r(z)} = \mathbf{1} - \left(\frac{\mathbf{B_r}}{\mathbf{z}}\right)^{\lambda}$, use H, and L_m to produce.

$$y_i = z_i^{1-\alpha-\beta} L_{mi}^{\alpha} \left(\mathbf{\Gamma_r} H_i + \kappa\right)^{\beta}$$

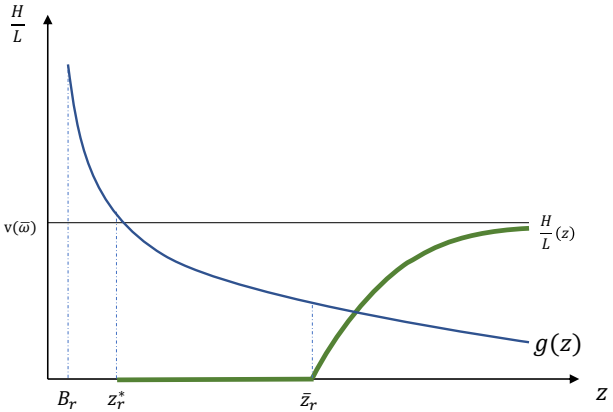
- Where $\alpha + \beta < 1$ captures the notion of limited span of control, and κ helps to determine differences in demand for college workers by firm size.

$$\frac{H_i}{L_{mi}} = \theta_0(\underbrace{\bar{\omega}_r}_{-}) - \kappa\theta_1(\underbrace{\Gamma_r}_{-}, \underbrace{\bar{\omega}_r}_{+})\frac{z_i^{\theta_2}}{\bar{y}^{\theta_3}}$$

- Define $z_r^*(\bar{\omega}_r)$ and $\bar{z}_r(\bar{\omega}_r)$, productivity thresholds for active firms and firms hiring H, respectively.

$$y(z_i) = \begin{cases} \theta_r \frac{w_r^l}{P_r^{na}} \left(\frac{z_i}{\bar{z}_r}\right)^{\frac{1-\alpha-\beta}{1-\alpha}} & \text{for } z_i \in [z_r^*(\bar{\omega}_r), \bar{z}_r(\bar{\omega}_r)) \\ \theta_r \frac{w_r^l}{P_r^{na}} \frac{z_i}{\bar{z}_r} & \text{for } z_i \geq \bar{z}_r(\bar{\omega}_r) \end{cases}$$

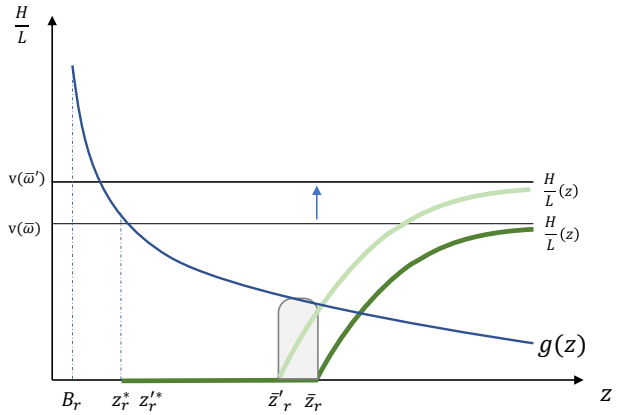
Optimal composition of labor as a function of productivity



▸ Proposition

Optimal composition of labor as a function of productivity

- A reduction in C_r increases the number of firms hiring College Workers.



▸ Proposition

Outline

Structural Estimation - Roadmap

Structural parameters:

$$\Omega^1 = \{\gamma_1, \gamma_2, \alpha, \beta, \lambda, \kappa, f_e, f_o, F(.)\}$$

Region specific college costs and productivities:

$$\Omega^2 = \{C_{r,t}, A^a_{r,t}, B_{r,t}, \Gamma_{r,t}\}_{r,t}$$

- Recover $\{\Omega^1, \Omega^2\}$ using rich administrative data, equilibrium development accounting, matching moments from the data to the model and calibration.

Structural Parameters Ω^1 | Estimation of C_r | Plot of $\hat{C}_r$ | Model Validation

Structural Estimation - Roadmap

Structural parameters:

$$\Omega^1 = \{\gamma_1, \gamma_2, \alpha, \beta, \lambda, \kappa, f_e, f_o, F(.)\}$$

Region specific college costs and productivities:

$$\Omega^2 = \{C_{r,t}, A^a_{r,t}, B_{r,t}, \Gamma_{r,t}\}_{r,t}$$

- Recover $\{\Omega^1, \Omega^2\}$ using rich administrative data, equilibrium development accounting, matching moments from the data to the model and calibration.
- Given $\hat{\Omega}^1, \hat{\Omega}^2$, random shock that reduces 2000 $\mathbf{C}_r$, to validate the model.

Structural Parameters Ω^1 | Estimation of $\mathbf{C}_r$ | Plot of $\hat{\mathbf{C}}_r$ | Model Validation

Structural Estimation - Roadmap

Structural parameters:

$$\Omega^1 = \{\gamma_1, \gamma_2, \alpha, \beta, \lambda, \kappa, f_e, f_o, F(.)\}$$

Region specific college costs and productivities:

$$\Omega^2 = \{C_{r,t}, A^a_{r,t}, B_{r,t}, \Gamma_{r,t}\}_{r,t}$$

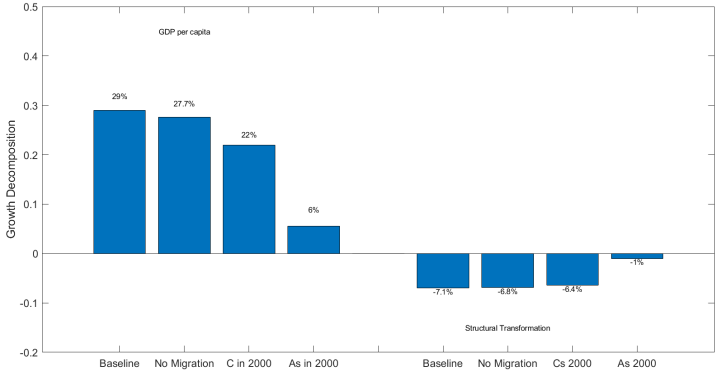
- Recover $\{\Omega^1, \Omega^2\}$ using rich administrative data, equilibrium development accounting, matching moments from the data to the model and calibration.
- Given $\hat{\Omega}^1, \hat{\Omega}^2$, random shock that reduces 2000 $\mathbf{C}_r$, to validate the model.
- Growth accounting, recompute the aggregate growth keeping $\mathbf{C}_r$ at 2000 levels.

▸ Structural Parameters Ω^1 ▸ Estimation of $\mathbf{C}_r$ ▸ Plot of $\hat{\mathbf{C}}_r$ ▸ Model Validation

Outline

Counterfactual Brazil with high access costs to college

Figure: Counterfactual Growth



- Evidence of heterogeneity in the gains from reducing college costs.

▸ Taking migration into account ▸ Heterogeneity across regions plot ▸ Heterogeneity across regions reg

Conclusions

- New evidence: higher education causes income growth, and labor reallocation outside agriculture.
- College expansion increases the prevalence of medium-size firms.
- New theory that connects college education to aggregate productivity gains through the expansion of firms.
- The expansion in college attainment accounts for around 18% of the Brazilian development process between 2000 and 2010.
- Important complementarities in education investments: Regions with high skill-biased technical change benefited more from college expansion.

Other Important Policies

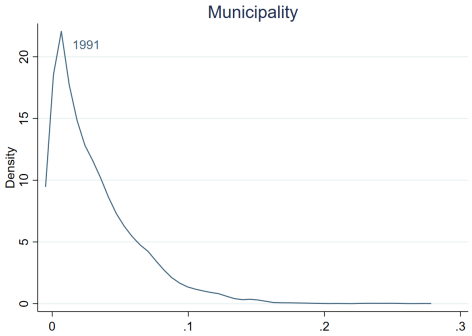
- SISU - Centralized Online System of Vacancies offered by Public Universities established in 2010.
  - Selection based only on ENEM and free of charge.

- PROUNI - Scholarship program covering 50/100% cost of tuition.
  - Requires minimum average score at ENEM, and household income per person must be less than 1.5 or 3 MW.
  - Participating colleges have the following benefits and conditions:
    - Provide one full scholarship for each of 11 regular students, and use partial scholarships up to 8.5% of gross income.
    - Scholarships must allocate the same share across majors.
    - Fiscal benefits and expanded demand.

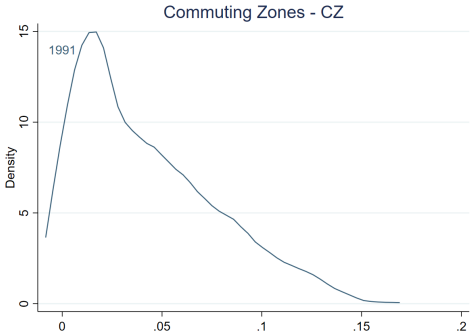
- FIES - College loans to students.
  - Established in 2001. It was relatively small before 2010.
  - Condition: requires a minimum average score at ENEM and household income per person under 3 MW.

Back

Expansion in College Attainment of the Youth (25-34 years old)



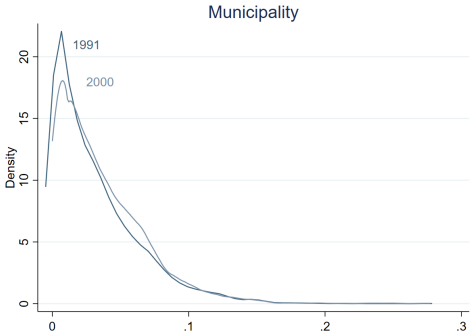
Share of Youth with College Degree



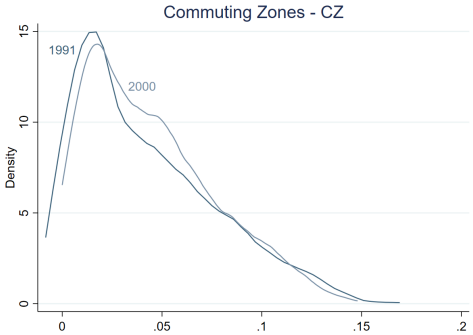
Share of Youth with College Degree

Back

Expansion in College Attainment of the Youth (25-34 years old)



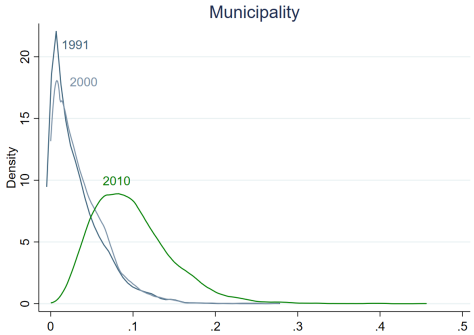
Share of Youth with College Degree



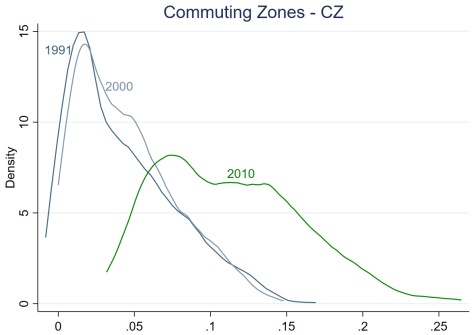
Share of Youth with College Degree

Back

Expansion in College Attainment of the Youth (25-34 years old)



Share of Youth with College Degree



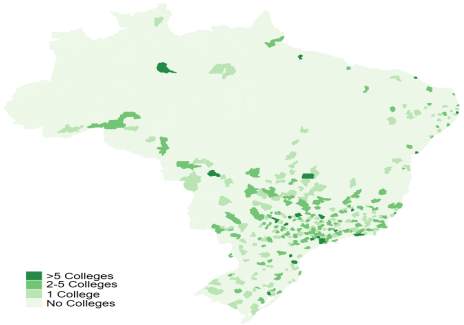
Share of Youth with College Degree

Back

Colleges Before The Reform

(a) Occupancy Rate

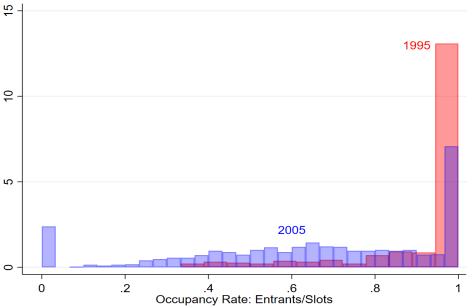
(b) Spatial Distribution



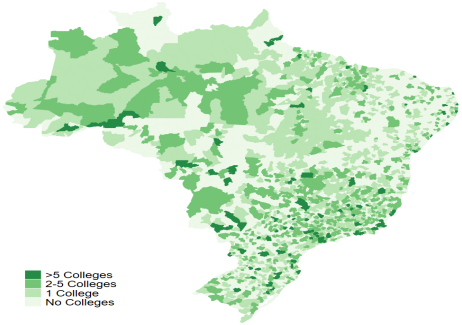
Back

Colleges After The Reform

(a) Occupancy Rate



(b) Spatial Distribution



Back

Characteristics of New College Graduates

- **Migration to Enrollment**

Table: Distance Traveled in Km to Enroll in College, by Institution Type

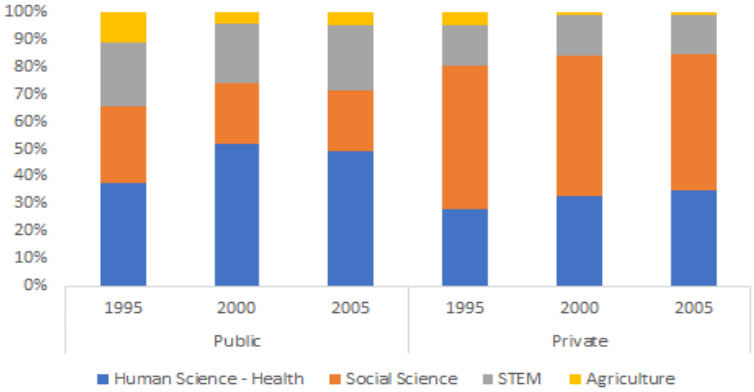
Institution Type	Mean	Std. Dev.	P10	P25	P50	P75	P90
Private NFP	47	195	0	0	0	28	77
Private FP	47	191	0	0	0	27	85
Public	92	274	0	0	0	59	221

Notes: Distance measured in Km. Source:Higher Education Census and Socioeconomic Data from ENEM.

Back

Most Demanded Majors

* Business Administration and Law had the highest enrollment in Private Universities.

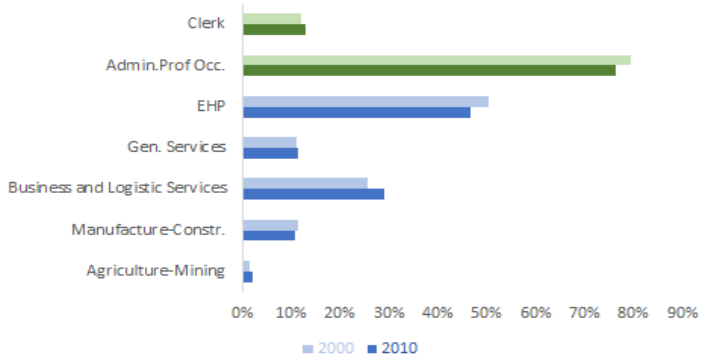


Back

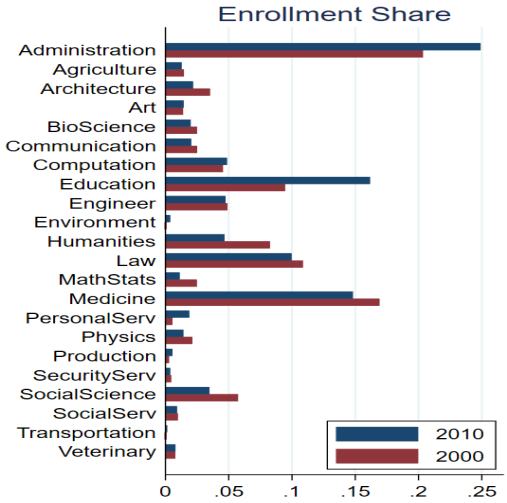
Same Occupation Sector Profiles for **New** College Graduates

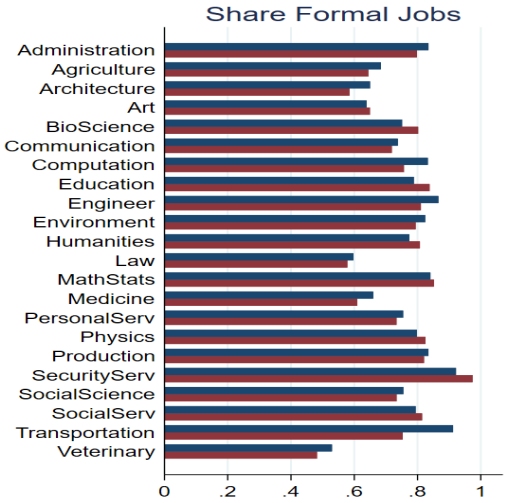
- Compare college graduates between 23-34yo pre- and post-reform. Similar job profiles: formal workers, white-collar occupations, in Business and EHP services sectors.

Figure: Job Profile of College Graduates 23-34yo



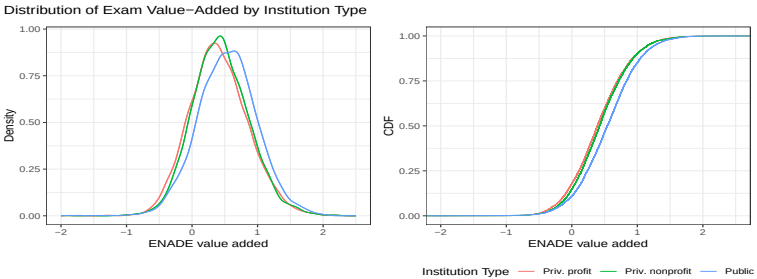
Same Levels of Formality for **New** College Graduates





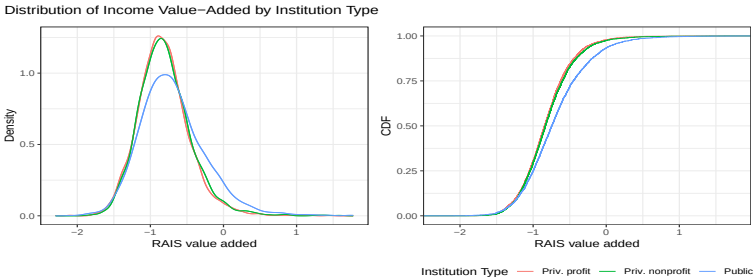
Quality provided

- Measuring quality differences across colleges:
  - Link pre-college test scores (ENEM), college affiliation, and post-college outcomes (RAIS and ENADE).



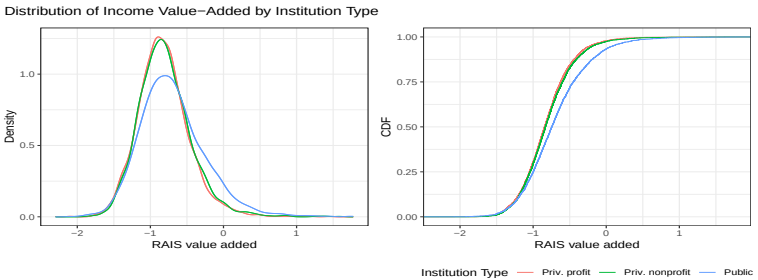
Quality provided

- Measuring quality differences across colleges:
  - Link pre-college test scores (ENEM), college affiliation, and post-college outcomes (RAIS and ENADE).



Quality provided

- Measuring quality differences across colleges:
  - Link pre-college test scores (ENEM), college affiliation, and post-college outcomes (RAIS and ENADE).



- Implications of quality deterioration:
  - Deterioration of average quality provided creates attenuation bias relative to a case with no quality downgrade.

Back

Data

- National population censuses 1980, 1991, 2000 and 2010.
- Higher education census 1995, 2005, 2009-2012.
- Matched employer-employee database 1991 - 2014.
- Microdata of ENEM scores 2009 - 2012.

Back

Controls Used

Table: Baseline Controls

List of Variables	N	mean	sd	min	max
% Adult Population with at least HS, 1991	486	0.09	0.04	0.01	0.32
% Adult Population with College Degree, 1991	486	0.04	0.03	0.001	0.16
% Working Informal Sector, 1991	486	0.60	0.12	0.32	0.85
% Working in High Skill Services, 1991	486	0.18	0.06	0.05	0.45
Δ % Working in High Skill Services, 1991-00	486	0.02	0.03	-0.26	0.13
% Working in Agriculture, 1991	486	0.38	0.19	0.01	0.79
Δ % Working in Agriculture, 1991-00	486	-0.07	0.05	-0.31	0.11
Average Income, 2000	486	2.36	0.39	1.43	3.41
Average Income, 1991	486	8.43	0.45	7.18	10.52
Log Population, 1991	486	11.99	0.97	7.43	16.27
Δ Log Population, 1991-00	486	0.13	0.13	-0.28	1.13
Had HEIs in 1995	486	0.44	0.49	0	1
#HEIs/# Youth with High School (No College), 1991	486	0.02	0.03	0	.18

Notes: The geographic unit is the CZ. These variables will be used as controls in the empirical analysis.

Back

World Bank Survey

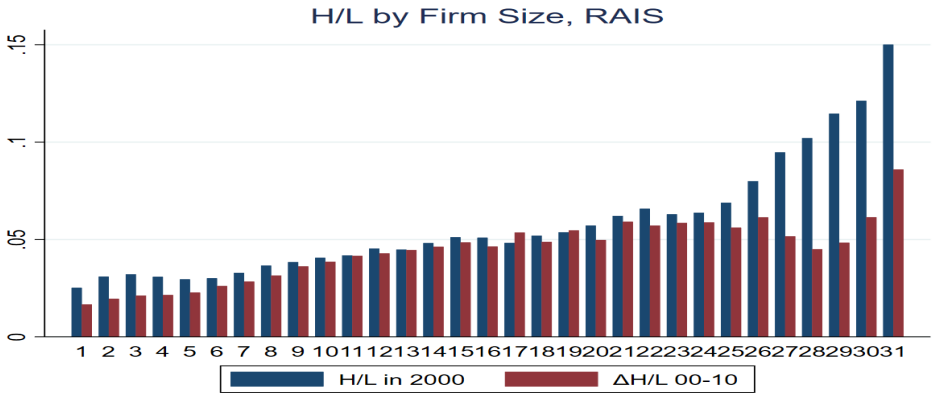
Variable	Low Education of Workforce					
	Biggest Constraint			One Constraint		
Region	Brazil	LAC	World	Brazil	LAC	World
	(1)	(2)	(3)	(4)	(5)	(6)
All	12.6%	7.8%	9.8%	74.9%	28.6%	20.8%
Small Firms 6 to 19 Workers	13.4%	6.7%	8.6%	71%	28.1%	19.3%
Medium Firms 20 to 99 Workers	12.4%	9.8%	11.5%	86.3%	27.6%	22.6%
Large Firms 100+ Workers	4.6%	11.9%	14.2%	69.9%	36.1%	25%

Notes: The table provides information on the survey results conducted on a representative sample of 1,803 firms in Brazil. The survey explores the main constraints firms face to grow.

Back

College to no college ratio in the RAIS

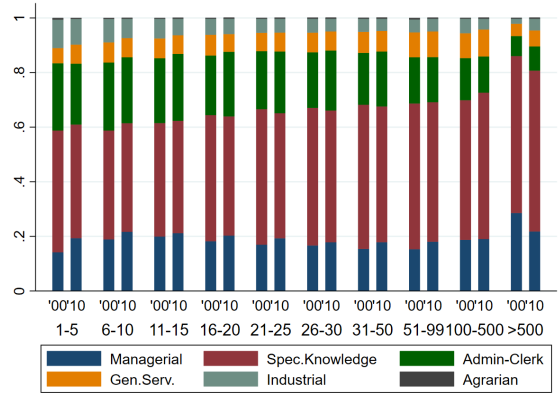
Figure: College/No-College Workforce Composition



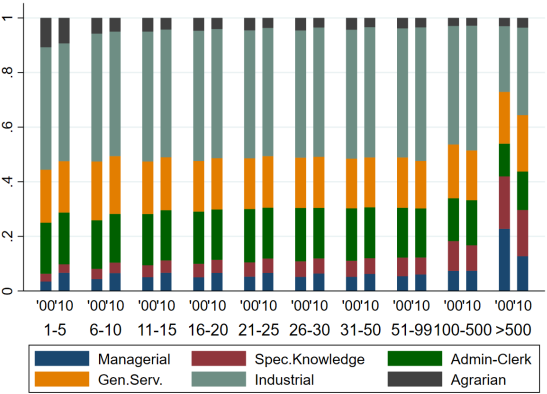
Back

Occupation of college graduates in RAIS

Figure: Occupation Differences Within Formal Firms



(a) 25-34yo College Workers



(b) No-College Workers

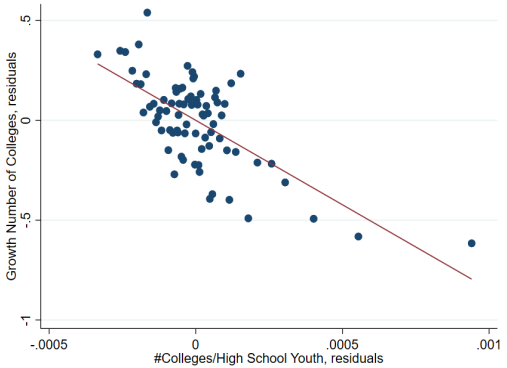
Back

First stage: instrument, entry, college attainment

- Private colleges entered under-served locations.

$$\Delta log(\#Colleges) = \alpha Z_r^{1991} + X_r\gamma + \theta_s + u_r$$

Dep. Variable	Growth #Colleges	
Specification	(1)	(2)
Z_r	-120.4***	-84.56***
	(10.69)	(11.21)
F-Test		
R^2	0.60	0.67
Obs.	486	486
Initial Development	No	Yes
Pre-Trends	No	Yes

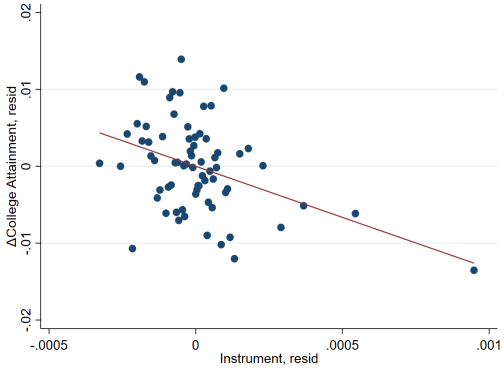


First stage: instrument, entry, college attainment

- Faster increase in college attainment.

$$\Delta h_r^{2000-10} = \alpha Z_r^{1991} + X_r\gamma + \theta_s + u_r$$

Dep. Variable	Growth	#Colleges	$\Delta h^{2000-10}$ (25-34)	
Specification	(1)	(2)	(1)	(2)
Z_r	-120.4***	-84.56***	-0.18***	-0.14***
	(10.69)	(11.21)	(0.04)	(0.04)
F-Test			21.69	13.01
			0.00	0.00
R^2	0.60	0.67	0.56	0.61
Obs.	486	486	486	486
Initial Development	No	Yes	No	Yes
Pre-Trends	No	Yes	No	Yes



Use of population weights

Table: Demographics Across the Instrument Variation

Sample	Observations	1991 Pop (1000s)
With Colleges	215	514
With Colleges (Excluding RJ,SP)	213	419
No Colleges	269	133

Back

Robustness

- Analysis is robust to controlling for other development drivers studied in Brazil:
  - 1991 Trade liberalization (Dix-Carneiro and Kovak 2017).
  - Minimum wage increase (Engbom and Moser 2022).
  - Introduction of new agricultural technologies (Bustos, Caprettini, and Ponticelli 2016).

Back

Discussion on Exclusion Restriction: Diff-in-Diff Analysis

- I use the following specification to assess the effect of the instrument in a diff-in-diff setting:

$$y_{rt} = \delta_r + \delta_t + \beta_{2000} 1_{t=2000} * Z_r + \beta_{2010} 1_{t=2010} * Z_r + X_r^{t-20} \Gamma + \theta_r^s * t + u_{rt}$$

Discussion on Exclusion Restriction: Diff-in-Diff Analysis

- I use the following specification to assess the effect of the instrument in a diff-in-diff setting:

$$y_{rt} = \delta_r + \delta_t + \beta_{2000} 1_{t=2000} * Z_r + \beta_{2010} 1_{t=2010} * Z_r + X_r^{t-20} \Gamma + \theta_r^s * t + u_{rt}$$

Table: Diff-in-Diff Analysis

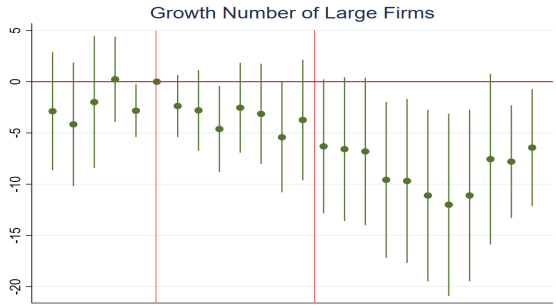
Dep. Variable	College Attainment (25-34)		Share of People in Agriculture		Share of People in White Collar Occ.	
Specification	(1)	(2)	(3)	(4)	(5)	(6)
Z_r * 1{t=2000}	4.26		-2.93		2.83*	
	(3.51)		(3.84)		(1.55)	
Z_r * 1{t=2010}	-16.81***	-17.50***	12.33***	12.83***	-13.05***	-13.53***
	(3.61)	(3.95)	(3.95)	(3.90)	(1.60)	(3.92)
R^2	0.856	0.856	0.699	0.699	0.372	0.371
Obs.	958	958	958	958	958	958

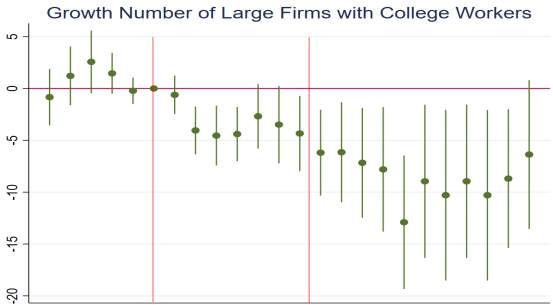
Notes: Columns (2), (4) and (6) estimate the specification excluding $1_{t=2000} * Z_r$. All specification control for lagged initial levels of human capital and Development, including the average local income, the population level, the share of people working in agriculture and high-skill services, the percentage of people who completed high school and college, and a dummy of whether the location had college pre-reform.

Discussion on Exclusion Restriction: Evidence From Firms

- Effect should transpire only when College graduates enter the labor market after 2003.

$$\Delta y_{ir}^{t-1996} = \delta + \sum_{t=1991}^{2014} \beta^t 1_t * Z_r + X_r^{t-20}\Gamma + \theta_r^s + \theta_r^i + \Delta u_{ir}^{t-1996}$$





Back

Discussion on Exclusion Restriction: returns to college

- Strategy aims to control for demand forces → isolate expansion in the supply of skills.

Table: Returns to College

Outcome	Change in Returns to College, 2000 - 10		
Specification	Reduced form	OLS	IV
	(1)	(2)	(3)
Z_r	102.75**		
	(42.35)		
Δh_r		0.91**	-4.33**
		(0.46)	(2.07)
Obs.	486	486	486
R^2	0.141	0.081	0.107

Controls 1991 initial development: average income, population, agricultural and high-skill service employment share, high school attainment, college attainment, and had college dummy. Pre-trend 1991-2000: income, population growth, change in agricultural and high-skill employment.

- A 6% increase in college attainment → 25% reduction in college returns.

Back

Is Migration After College a Concern?

- Education might complement migration choice (Hsiao 2022).
  - Use permanent migration patterns from the Census.

Is Migration After College a Concern?

- Education might complement migration choice (Hsiao 2022).
  - Use permanent migration patterns from the Census.

Sample	All Education Groups		College Graduates	
Dep. Variable	Share Immigrants			
Z_r	3.84	-2.31	5.82	4.01
	(8.17)	(6.90)	(12.48)	(11.48)
R^2	0.392	0.668	0.358	0.584
Dep. Variable	Share Emigrants			
Z_r	23.26***	6.29	38.41***	18.56
	(5.19)	(4.94)	(10.85)	(11.73)
R^2	0.376	0.569	0.255	0.335
Dep. Variable	Net Migration			
Z_r	-19.82*	-8.80	-33.15	-14.94
	(11.61)	(9.24)	(20.96)	(19.25)
R^2	0.270	0.647	0.268	0.528
Obs.	486	486	486	486
Controls	No	Yes	No	Yes

* Net migration of College Grads:

$$NM_r^e = \frac{Immigrant_r^e - Emigrant_r^e}{Locals_r^e}$$

- $Corr(Z_r, NM_r^e) > 0 \rightarrow$ Attenuation Bias.
- $Corr(Z_r, NM_r^e) < 0 \rightarrow$ Upward Bias.

Back

Equilibrium

- Given $F(.)$ and $\{\psi_r, N_r, \mathbf{C_r}\}$, an equilibrium is an allocation $\{L_{ar}, \mathbf{H}_r^{na}, H_r\}_r$ and a price system $\{w_r^l, \bar{\omega}_r, P_r^a, P_r^{na}\}_r$ that satisfies equilibrium conditions $\forall r$:

$$w_r^l L_{ar} = \left(\frac{\left(w_r^l (A_r^a)^{-1}\right)^{1-\gamma_2}}{\sum_j \left(w_j^l (A_j^a)^{-1}\right)^{1-\gamma_2}} \right) \left(\frac{\left(\sum_j \left(w_j^l (A_j^a)^{-1}\right)^{1-\gamma_2}\right)^{\frac{1-\gamma_1}{1-\gamma_2}}}{\sum_{s'} \left(\sum_j \left(w_j^l (A_j^{s'})^{-1}\right)^{1-\gamma_2}\right)^{\frac{1-\gamma_1}{1-\gamma_2}}} \right) \sum_m \Upsilon_m$$

$$w_r^l \mathbf{H_r^{na}} = \left(\frac{\left(w_r^l (A_r^{na})^{-1}\right)^{1-\gamma_2}}{\sum_j \left(w_j^l (A_j^{na})^{-1}\right)^{1-\gamma_2}} \right) \left(\frac{\left(\sum_j \left(w_j^l (A_j^{na})^{-1}\right)^{1-\gamma_2}\right)^{\frac{1-\gamma_1}{1-\gamma_2}}}{\sum_{s'} \left(\sum_j \left(w_j^l (A_j^{s'})^{-1}\right)^{1-\gamma_2}\right)^{\frac{1-\gamma_1}{1-\gamma_2}}} \right) \sum_m \Upsilon_m$$

$$\left(\psi_r E\left[\epsilon_i | \epsilon_i \geq \frac{1}{\bar{\omega}_r} + \mathbf{C_r}\right] h(\bar{\omega}_r, \mathbf{C_r}) + (1-\psi_r) h_r^o - \psi_r \mathbf{C_r} h(\bar{\omega}_r, \mathbf{C_r})\right) N_r = H_r^d$$

$$\left(\psi_r F\left(\frac{1}{\bar{\omega}_r} + \mathbf{C_r}\right) + (1-\psi_r)(1-h_r^o)\right) N_r = L_r^d$$

- and is consistent with optimal education choice, optimal consumption, and profit maximization in

Demand system

- Preferences follow a two-layer CES structure which delivers the following demands:

$$C_r^s = \left(\frac{\bar{P}_r^s}{\bar{P}_r}\right)^{-\gamma_1} \frac{1}{\bar{P}_r} \Upsilon_r \qquad C_{rm}^s = \left(\frac{P_m^s}{\bar{P}_r^s}\right)^{-\gamma_2} C_r^s$$

- γ_1 and γ_2 are the sector and variety elasticity of substitution, respectively. Aggregate income in each r:

$$\Upsilon_r = \left(\psi E_r[e^i] + (1-\psi) w_r^l (\bar{\omega}_r h_r^o + (1 - h_r^o))\right) N_r$$

$$E_r[e^i] = w_r^l \left[F\left(\frac{1}{\bar{\omega}_r} + \mathbf{C_r}\right) + \bar{\omega}_r h(\bar{\omega}_r, \mathbf{C_r}) \left(E\left[\epsilon_i | \epsilon_i \geq \frac{1}{\bar{\omega}_r} + \mathbf{C_r}\right] - \mathbf{C_r} \right) \right]$$

Back

The effects of falling access costs to college $C_r \downarrow$

Proposition

When $\frac{\kappa}{f_o}\frac{\bar{\omega}_r}{\Gamma_r} > \frac{\beta}{1-\alpha}$. *A reduction in* C_r *has the following theoretical implications:*

1. *Increases the supply of young college workers.*
2. *Reduces the equilibrium level of college skill premium* $\bar{\omega}_r$.
3. *The effect on the number of firms hiring college workers, larger in size, increases.*
4. *Employment in NA sector, both college and no-college, increases.*
5. *Average income increases.*

▸ Equilibrium Definition ◂ Back

Structural Estimation Ω^1: Estimating λ

- Estimation of λ using firm-level administrative data.

$$L_{mi} + H_i = \underbrace{\left(\frac{\alpha+\beta}{\beta}\right)\frac{f_r^o \zeta_r}{\bar{z}_r}}_{\Phi_r} z_i - f_r^o \zeta_r$$

- 1-to-1 relationship between employment and productivity z_i. As $z_i \sim G_r(.)$ Pareto:

$$G_r\left(\Phi_r z_i - f_r^o \zeta_r \leq l\right) = \bar{G}_r(l)$$

$$1 - \bar{G}_r(l) = \left(\frac{B_r \Phi_r}{l + f_r^o \zeta_r}\right)^{\lambda}$$

- In logs:

$$ln(1 - \bar{G}_r(l)) = \lambda\theta_r - \lambda ln\left(l + f_r^o \zeta_r\right)$$

- Specify l_k for different size bins and run a regression to recover $\lambda = 1.3$.

Back

Recovering A_r^a, B_r

- Using equilibrium conditions, and data on $\{w_r^l, L_r^a\}_r$ and P^a, the relative price of A to NA. We can obtain an expression for A_r^a, A_r^{na} :

$$A_r^a = \frac{1}{P^a}\left(\frac{w_r^{\gamma_2} L_{ar}}{\sum_j w_j L_{aj}}\right)^{\frac{1}{\gamma_2 - 1}}$$

$$A_r^{na} = \left(\frac{w_r^{\gamma_2} \mathbf{H_r^{na}}}{\sum_j w_j \mathbf{H_r^{na}}}\right)^{\frac{1}{\gamma_2 - 1}}$$

- Finally, using the expression we derive for A_r^{na} we can obtain B_r:

$$A_r^{na} = \alpha \kappa^{\beta}\left(\frac{1-\alpha}{\alpha f_r^o}\right)^{1-\alpha}(z_r^*(B_r))^{1-\alpha-\beta}$$

Back

Structural Estimation Ω^2: Recovering $\mathbf{C}$

Estimated ability distribution, demand, and supply for college deliver three moments:

$$\bar{\omega}\left(\hat{E}\left[\epsilon_i|\epsilon_i \geq \frac{1}{\bar{\omega}_r} + \mathbf{C}_r\right] - \mathbf{C}_r\right)\frac{H_r}{L_r^{na}} = \frac{P_1}{\tilde{\zeta}(\Gamma_r, \bar{\omega}) - P_2} \tag{1}$$

$$h_r = 1 - \hat{F}\left(\frac{1}{\bar{\omega}_r} + \mathbf{C}_r\right) \tag{2}$$

$$\hat{\theta}_{1r} = \bar{\omega}_r \hat{E}\left[\epsilon_i|\epsilon_i \geq \frac{1}{\bar{\omega}_r} + \mathbf{C}_r\right] \tag{3}$$

Proposition

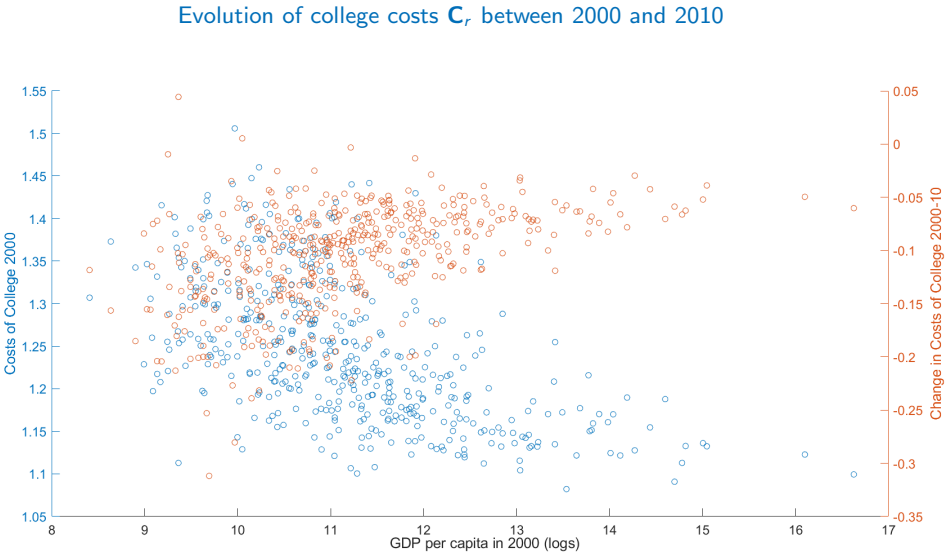
There is a unique solution to the system of equations (1) - (3) given by:

$$\mathbf{C}_r = \frac{\hat{\theta}_{1r}F^{-1}(1-h_r) - \hat{E}\left[\epsilon_i|\epsilon_i \geq F^{-1}(1-h_r)\right]}{\hat{\theta}_{1r}}$$

Solution for $\bar{\omega}_r$ and Γ_r in the paper.

Recovering A_r and B_r Back

Estimates of C_r



Model Validation

- Replicate regressions using the model-generated data.
  - Baseline 2000, counterfactual: $\{\Gamma_{r,2000}, A^a_{r,2000}, B_{r,2000}\}_{r,t}$ and $\{\mathbf{C}^{shock}_r\}_r$.

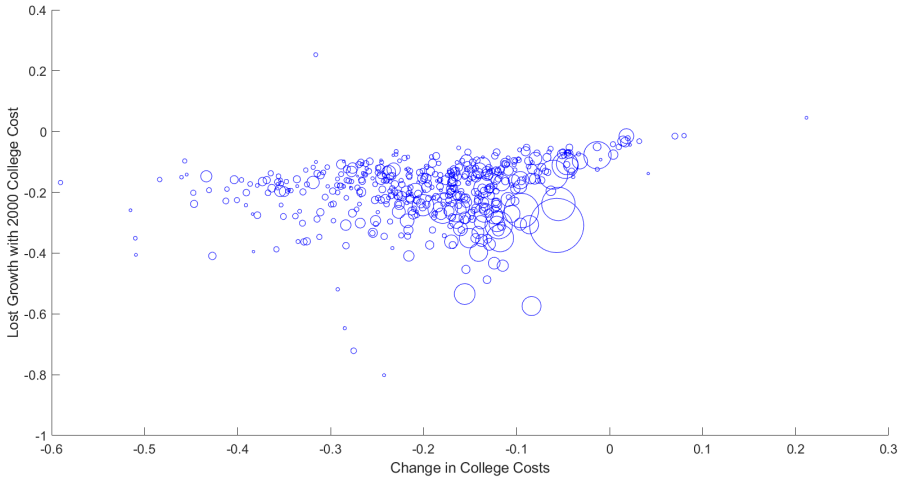
$$\Delta y_r = \beta \Delta h_r + X_r \gamma + v_r$$

Table: Comparison between model Regression and Empirical Regression

Outcome Specification	Data (OLS) (1)	Data (IV) (2)	**Model** **(3)**
Δ Average Income	0.59	0.52	**0.42**
	(0.08)	(0.40)	**(0.01)**
Δ % Working Agriculture	-0.53	-0.61	**−0.41**
	(0.04)	(0.28)	**(0.02)**
Δ % Working White-Collar	0.51	0.57	**0.44**
	(0.02)	(0.11)	**(0.01)**
Δ log#Firms	0.29	0.53	**−0.35**
	(0.06)	(0.36)	**(0.01)**
Δ log#Large Firms	0.26	0.72	**0.50**
	(0.04)	(0.23)	**(0.01)**

◂ Recovering $\{C_r\}_r$ ◂ Recovering $\{A^a_r, B_r\}_r$ ◂ Back

Spatial heterogeneity: college access and income growth

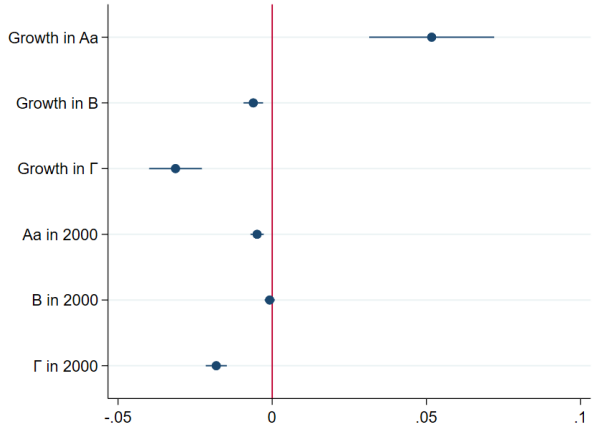


Back

Complementarity between education costs and technology growth

$$\text{lost growth}_r = \theta_0 + \text{fundamentals}_r^{2010}\rho_1 + \Delta\text{fundamentals}_r^{2000-10}\rho_2 + \epsilon_r$$

Assessing heterogeneity in the reduction in GDP per capita growth



Structural Estimation Ω^1

Table: Structural Parameters

Parameters	Target	Value
	Ability Distribution	
$F(.)$	Math Score Distribution ENEM 2009-2012	Lognormal(0,0.16^2)
	Preference Parameters	
γ_1	Sector Allocation of Labor	2.02
γ_2	From Pelegrina (2022)	4.8
	Production Function Parameters	
α	Share of Manufacturing Workers	0.11
β	Employment Share of College Workers - US	0.7
λ	Tail of employment distribution in RAIS	1.3
f_r^e	Normalization	0.1
f_r^o	Normalization	1
κ	Normalization	1

Example λ Back

Structural Estimation Ω^2: Recovering $\mathbf{C}_r$

Step 1: Approximate F(.) using distribution of math scores in ENEM exam.

- Connect three administrative datasets: (i) ENEM scores, (ii) Higher Education Census, (iii) RAIS.

$$ln(y_i) = \rho T_i + X_i\gamma + e_i$$

- Where $T_i \sim N(0,1)$ are standardized test results. I use $\epsilon_i = exp(\rho T_i) \sim LogNorm(0, \rho^2)$.

College Choice Back

Structural Estimation Ω^2: Recovering $\mathbf{C}_r$

Step 1: Approximate F(.) using distribution of math scores in ENEM exam.

- Connect three administrative datasets: (i) ENEM scores, (ii) Higher Education Census, (iii) RAIS.

$$ln(y_i) = \rho T_i + X_i\gamma + e_i$$

- Where $T_i \sim N(0,1)$ are standardized test results. I use $\epsilon_i = exp(\rho T_i) \sim LogNorm(0, \rho^2)$.

Step 2: Estimate Income differences between H and L, for each r.

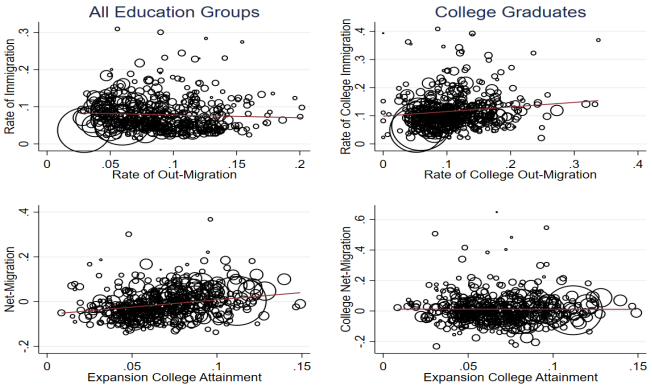
$$y_{ir} = \theta_{0r} + \theta_{1r} D_{ir}^{HE} + X_{ir}\Gamma_r + \epsilon_{ir}$$

where $D_{ir}^{HE} = 1$ if individual i has a college degree.

College Choice | Back

Correction for migration

- Heterogeneity in the prevalence of migration among college graduates.



- Net migration of college graduates creates biased $\mathbf{C}_r$ estimates.
  - Use mobility patterns from the Census to "bring back" individuals to their origin to recover correct $\mathbf{C}_r$.

Back