College Admission Policies and Pre-College Human Capital Investment: Evidence from China

Chenglei Diao, Chenyuan Liu and **Xiaohan Zhong**

Presented at DSE 2025 HKU
December 10th, 2025

Introduction

Part 1

Introduction: Unbalanced College Expansion and Educational Involution

- China has emerged as one of the world's most prominent cases of rapid higher education expansion over the past three decades:
  - Gross Enrolment Rate (1998-2024): 10%→60% (the World: 18%→ 43%);
  - Admission rates among applicants: 34%→80%.
- New colleges entering into the market, however, are almost entirely located at the low-end (locally administered and privately-owned, with inadequate fundings)
  - # of Higher Education Institutions: Total, 1,200→3,000, centrally administered: 110→130.
- **<u>Chronic scarcity of high-quality colleges:</u>**
  - Admission rates <u>among applicants</u>: four-year universities: staggering at 40%; top 100 universities, 6%, top 30, 2%; among <u>all college students</u>, declined by more than half
- Widened Quality Gap (or stratification) among colleges is also observed, partially due to major central government fundings targeted elite colleges (i.e., Project-985/211).

Introduction: Unbalanced College Expansion and Educational Involution

The widening gap in college admission cutoff scores, 2007-2017

Cutoff
Growth Rate,
2007-2017

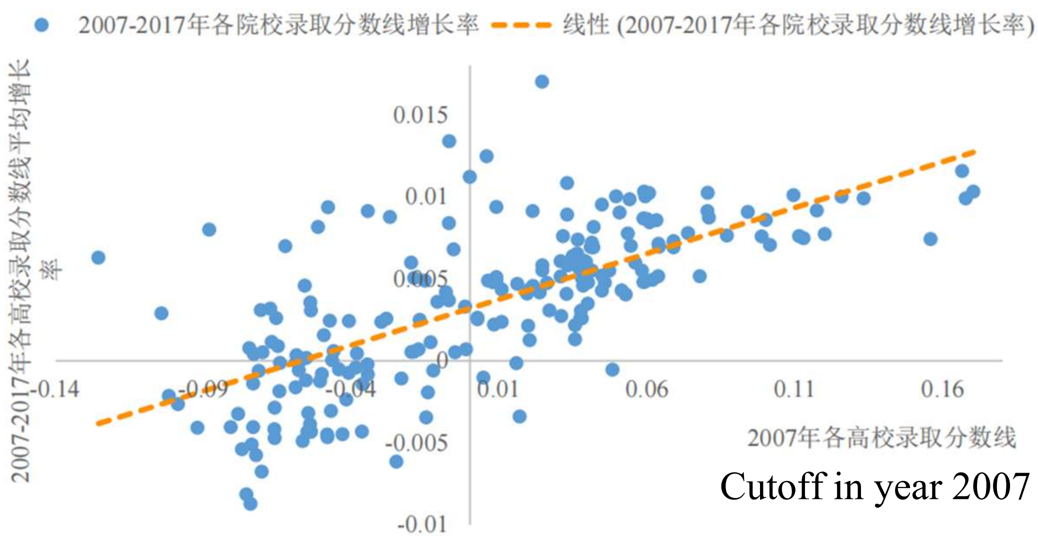


图4.2　2007-2017录取分数线平均增长率与初始录取分数线

来源：李思源（2019） Source: Li(2019)

Introduction: Unbalanced College Expansion and Educational Involution

Fierce competition for admission to high-quality colleges –"educational involution/anxiety" - has become a pressing social concern in China.

- The average daily study time of high school students increased from 6.3 to 9.6 hours during 2008-2018 (NTUS)
- Family education expenditure increased from 7.5% to 9.5% as the proportion of all cash expenditure during 2013-2023 (China Household Survey Yearbook), with current level much higher than the world average (Wei, 2023)
- The unprecedent growth of the extra-curriculum training business, with new firm registrations and total investment more than tripled during 2014-2019, and total sales increasing more than 40%.
- Serious and deteriorating mental health problem of high school students (Hou & Chen, 2021).

Introduction: Research Question

Our research questions:

1. **How does an <u>exam-centric</u> college admissions system affect students' pre-college human capital investment?**
2. **How can we measure the possibly *excess effort* (or overinvestment in human capital) and associated welfare loss?**
3. **What policies from *supply-side* can reduce excess human capital investment and associated welfare loss?**

Our key insights:

- College admission is a **tournament (or contest**), in addition to a matching or resource allocation.
- The incentive for pre-college human capital investment works through **two** channels: Productive channel & **Competitive channel.**
- Not only the aggerate quantity and quality of colleges matters for pre-college HC investment, but also their **quantity-quality composition**.

Introduction: Preview of the Model

In our theoretical model:

- **Students** are endowed with unobservable heterogeneous abilities and make (costly) effort to invest in pre-college human capital and strive for desirable college entrance exam (Gaokao) score.
- **Colleges**, varying in their qualities, admit students according to their exam scores up to their quotas.
- **Wages** after college graduation are jointly determined by student pre-college human capital level and the college quality.
- We solve for a **competitive equilibrium** in which students take the admission cutoff scores as given. The equilibrium is achieved under cutoff scores equating the quota of each college (i.e., supply) and the (expected) number of students who make efforts to meet its cutoff (i.e., demand).

Introduction: Preview of the Results

- We calibrate our model by estimating key parameters in Cobb-Douglas-formed **pre-college human capital production function** and **wage determination equation**, by using data from China's college admissions.

Results (current system):

- High-school students spend 2.3 more hours per day in study compared with under the first best scenario (i.e., 5.9 hours).
- The efficiency loss, measured by the reduction of student surplus from the first-best, is equivalent to 1.5% of the post-college wage level (1/4-1/5 of one-year education return).
- The competitive channel dominates the productive channel at the margin.
- There are large variations among students. Higher-ability students suffer much more.

Introduction: Preview of the Results

Results (counterfactual analysis):

- We explore the possibility of changing the **quality-quantity composition** of colleges in China's college market to **alleviate** the excess efforts and associated efficiency loss.
- Restricted to "resource neutral" policies (i.e., policies not changing the total college quality & quantity in the market), both reducing the quality-gap between colleges and expanding the high-quality college quotas can reduce excess efforts by up to 0.5 hours, deadweight loss up to 0.5 percentage point of wage level, and reduce competitive ratio by up to 13 percentage point.

Literature

Literature often find ambiguous effects of college admission policy on student efforts and academic performance.

- Affirmative action
  - Non-structural (reduced) form: Caldwell, 2010; Antonovics and Backes, 2014; Akhtari et al., 2020; Cotton et al., 2022
  - Structural Form: Bodoh-Creed & Hickman 2017 (BH); Grau, 2018
- "Top 10% Plan"
  - Reduced form: Cortes et al., 2011; Leeds et al., 2017
- China's college admissions
  - Reduced form: Xing, 2013, Luo et al., 2016, Xue, 2022

Other Related Literature

- Stratification in higher education (mostly studied in the U.S.)
  - Sallee et al. (2008), Bound, Hershibein and Long (2009), Hoxby (2009)
- Matching with efforts
  - Moldovanu and Sela (2001); Bodoh-Creed & Hickman (2018); Hafalir et al. (2018)
- Matching in large markets
  - Azevedo and Leshno (2016): cutoff scores as “market prices” and students as “price takers”.

The Model

Part 2

The Model: Students

- **Human capital production function:**

$$S_i = L_i = \alpha_0 A_i^{\alpha_1} E_i^{\alpha_2}, \quad \alpha_0, \alpha_1 > 0, \quad \alpha_2 \in (0,1)$$

Where S, L are the level of pre-college human capital of students, A is student (intrinsic) ability, E is study effort.

- **Cost function:**

$$C(E_i, A_i) = \frac{E_i{}^{\mu}}{A_i}, \mu > 1.$$

- **Wage determination equation:**

$$\mathbb{E}[w_i | P_i, L_i] = \alpha_w P_i{}^{\alpha_P} L_i{}^{\alpha_L}, \quad \alpha_w, \alpha_P > 0, \quad \alpha_L \in (0,1).$$

  - Where w is post-college-graduate wage; P is college quality.

- **Students choose effort level E to maximize utility：**

$$\max U = \mathbb{E}[w_i | P_i, L_i] \text{- } \mathrm{C}(\mathrm{E}_i, \mathrm{A}_i) = \beta P_i{}^{\alpha_P} A_i^{\beta_1} E_i^{\beta_2} - \frac{E_i^{\mu}}{A_i}$$

where $\beta \equiv \alpha_w \alpha_0{}^{\alpha_L}, \beta_1 \equiv \alpha_1 \alpha_L, \beta_2 \equiv \alpha_2 \alpha_L \in (0,1)$.

The Model: Colleges

- There are a large student population but only two (H and L, expanded to a few later) colleges in the market with different qualities: $P_H > P_L$.
- The total quota of the two universities is equal to the number of students: $Q_H + Q_L = N$. Define the proportion of quota as $q_{H,L} \equiv Q_{H,L}/N$
- The colleges admit students according to the college entrance exam (Gaokao) score, S, up to their quota, and give higher priority to students with higher S.

The Model: First-Best Outcome

- **FBO** = Student intrinsic ability is fully observable, and colleges admit students according to their abilities. It must be a *positive assortative matching* (high ability student matched with high quality college).
- **Cutoff ability $A^{\#}$:** is defined as the lowest ability eligible for College H under FBO.
- Students with $A < A^{\#}$ will enroll in College L, with **socially optimal** effort as:

$$E^{o}(A_i, P_L) = \underset{E_i}{argmax}\{\beta P_L^{\alpha_P} A_i^{\beta_1} E_i^{\beta_2} - \frac{E_i^{\mu}}{A_i}\}$$

- Students with $A \geq A^{\#}$ will enroll in College H, with **social optimal** effort as:

$$E^{o}(A_i, P_H) = \underset{E_i}{argmax}\{\beta P_H^{\alpha_P} A_i^{\beta_1} E_i^{\beta_2} - \frac{E_i^{\mu}}{A_i}\}.$$

The Model: Competitive Equilibrium

- We assume student ability is private information and admissions are based on college entrance exam (Gaokao) scores.
- **Competitive Equilibrium** is achieved if both (1) and (2) hold**:**

1. *For any given* cutoff score of College H ($\bar{S}$), students choose their (individually) optimal effort (E), leading to their optimal human capital investment level ($S = L$), as well as their admitted college - depending on $S \geq \bar{S}$ or not.
2. The *equilibrium* cutoff score of College H ($\bar{S}^*$) is determined by： $Q_H = D(\bar{S}^*)$, here $D(\bar{S})$ is the demand function - the number of students choosing an effort level enough to meet the cutoff $\bar{S}$.

The Model: Competitive Equilibrium

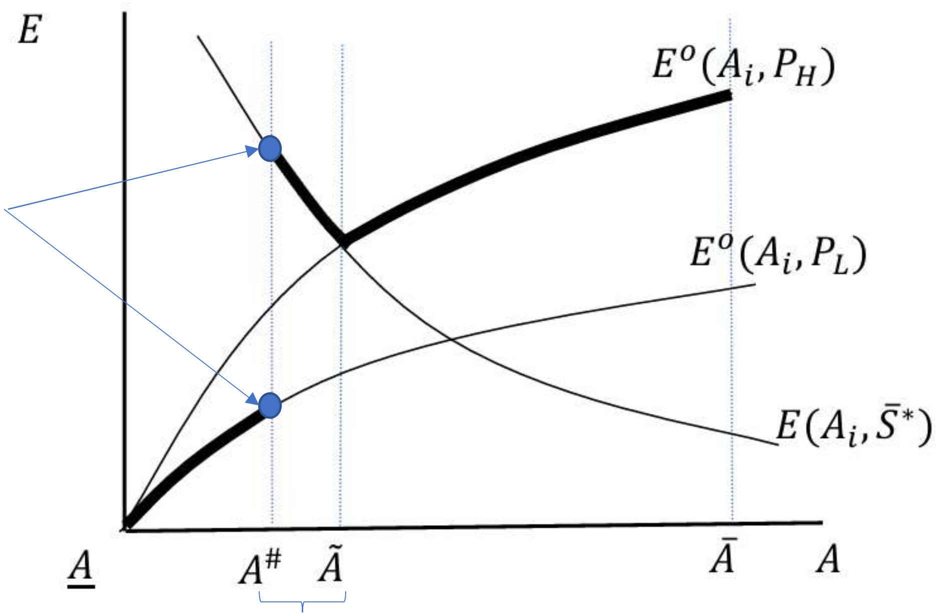
- Denote the effort level just to meet a given cutoff ($\bar{S}$) as:

$$E(A_i, \bar{S}), \text{s.t.}\, \bar{S} = \alpha_0 A_i^{\alpha_1} E_i^{\alpha_2}$$

- **<u>Claim:</u>** The equilibrium cutoff score ($\bar{S}^*$) is achieved if $E\left(A^\#, \bar{S}^*\right) = E^o\left(A^\#, P_L\right)$.
  - That is, the student with $\boldsymbol{A = A^\#}$ is indifferent between making the effort just to meet the cutoff, and the optimal effort to get into the low-quality college for certain.
- Then: students with $\mathrm{A} < \mathrm{A}^\#$ choose $E^o(A_i, P_L)$ and enroll in College L, and students with $\mathrm{A} \geq \mathrm{A}^\#$ choose $Max\ \{E^o(A_i, P_H), E(A_i, \bar{S}^*)\}$ and enroll in College H (by single-crossing property).
- The equilibrium allocation is **the same as FBO** !

The Model: Competitive Equilibrium

Student with ability A#
is indifferent between
these two points.



Marginal students (i.e., students with excess efforts and welfare loss)

Figure 1 Student effort under competitive equilibrium

The Model: Competitive Equilibrium with Random Perturbations

- Modify the human capital production function to:

$$S_i = S(L(A_i, E_i), \varepsilon_i) = \alpha_0 A_i^{\alpha_1} E_i^{\alpha_2} e^{\varepsilon_i}, \varepsilon \sim N(0, \sigma_\varepsilon^2).$$

- S_i is the college entrance exam (Gaokao) scores whereas ε be a disturbance imposed on student true human capital level (L_i).

- The utility maximization problem becomes:

$$max\ U = \mathbb{E}_\varepsilon \left\{ \alpha_w P_i (S_i)^{\alpha_P} L_i{}^{\alpha_L} - \frac{E_i^\mu}{A_i} \right\}.$$

The Model: Competitive Equilibrium with Random Perturbation

The *competitive equilibrium cutoff* ($\bar{S}^*$) *under random perturbations* is solved recursively such that:

- when all the students take $\bar{S}^*$ as given (and deterministic) and make efforts $E_i(\bar{S}^*)$ to maximize its *expected* utility……
- the resulted empirical score distribution $F(\cdot\,|\{E_i(\bar{S}^*)\})$ approaches to $1 - F(\bar{S}^*|\{E_i(\bar{S}^*)\}) = q_H$.

The Model: Productive Channel vs. Competitive Channel

- The **revenue** part of a student's utility function is:

$$R \equiv \mathbb{E}_\varepsilon\left\{\alpha_w P_i^{\alpha_P} {L_i}^{\alpha_L}\right\} = \alpha_w {L_i}^{\alpha_L} \mathbb{E}_\varepsilon\{P_i^{\alpha_P}\}$$

- Then the **marginal revenue** w. r. t. HC investment is:

$$MR_i(L_i) \;\; = \alpha_L \alpha_w L_i^{\alpha_L - 1} \mathbb{E}_\varepsilon\{P_i^{\alpha_P}\} + \alpha_w L_i^{\alpha_L} \frac{\partial \mathbb{E}_\varepsilon\{P_i^{\alpha_P}\}}{\partial L_i}$$

- **Productive Channel** of HC investment:

$$MR_p = \alpha_L \alpha_w L_i^{\alpha_L - 1} \mathbb{E}_\varepsilon\{P_i^{\alpha_P}\}$$

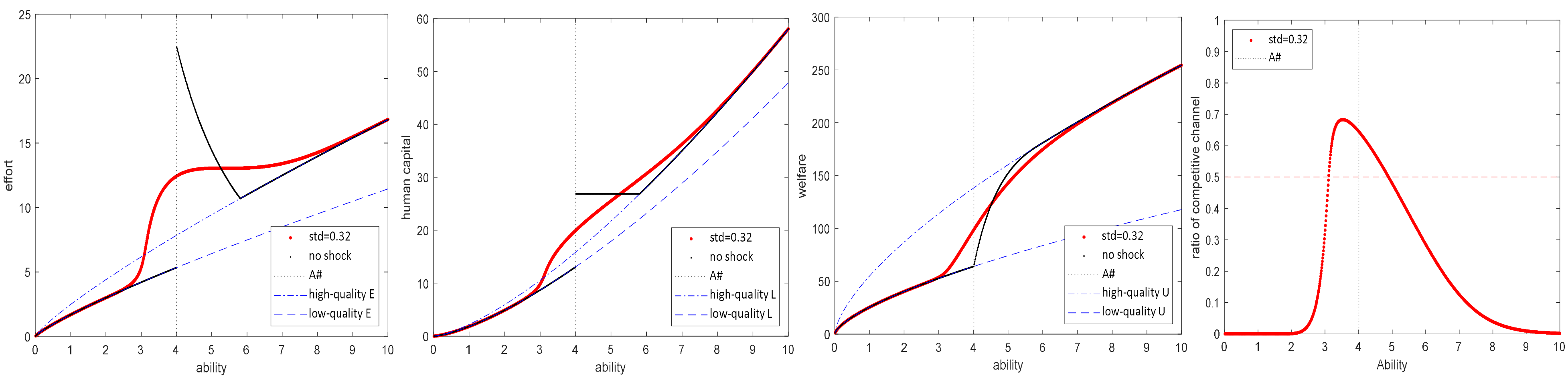
- **Competitive channel** of HC investment:

$$MR_c = \alpha_w L_i^{\alpha_L} \frac{\partial \mathbb{E}_\varepsilon\{P_i^{\alpha_P}\}}{\partial L_i}$$

- The relative strength of the competitive channel is called the **competitive channel ratio**:

$$\boldsymbol{Ratio_c = \frac{MR_c}{MR}}$$

Competitive Equilibrium: With Random Perturbation



(a) effort (b) score(pre-college HC) (c) utility (d) Competitive channel ratio

Figure 3 Competitive equilibrium outcome (without/with random perturbation)

Graph Parameters

Empirical Calibration

Part 3

Data Source

- China Family Panel Survey (CFPS, 2010-2020) (see Table A7)
  - 25 provinces, 14,960 households, six rounds(biennial)
  - Including measures of students' abilities and human capital (by cognitive ability tests, CAT), daily study time, and socioeconomic status
  - In our sample: 1,200-4,700 high school students
- China College Student Survey (CCSS 2013) (see Table A8)
  - Stratified random sampling, 65 colleges of all levels located in 23 provinces. Including measures of college entrance examination(Gaokao) scores, college GPA, post-graduation wages, and socioeconomic characteristics
  - In our sample: 5,929 college students from 30 residential provinces, taking Gaokao in 2009

Model Estimation

1. **Human capital Production Function (from CFPS):**

$$ln(\tilde{S}_i) = ln(\alpha_0) + \alpha_1\, ln(A_i) + \alpha_2\, ln(E_i) + \alpha_{3k}\, ln(Q_{ki}) + \varepsilon_i,$$

where $\tilde{S}_i$ is cognitive ability test (CAT) scores of high school graduates; A_i is an early CAT scores; E_i is study time in high school; Q_{ki} are student SES variables (as controls), ε_i is the residuals.

2. Wage Equation (from CCSS):

$$ln(w_i) = ln(\alpha_w) + \alpha_{P1} College211_i + \alpha_{P2} Tier1_i + \alpha_{P3} Tier23_i + \alpha_L\, ln(L_i) + \alpha_{Qk} \ln(Q_{ki}) + \eta_i.$$

Where w_i is monthly wage offer after college graduation; college quality are categorical variables (with 4 tiers) , and L_i is inter-provincially comparable pre-college HC.

3. Cost Function (μ)**:**

derived from minimizing the sum of squared distance between the predicted equilibrium effort (given μ) and the actual equilibrium effort, i.e., $\mu = argmin \sum_{i=1}^{N} (\hat{E}_i(\mu) - E_i)^2$.

Model Estimation: Human Capital Production Function

Table 4 Estimating human capital production function (CFPS data)

	Dependent variable: Ln(S)				
	All students	The last CAT test in middle school as ability(A)	Years between two CAT tests ≥ 6		
	(1)	(2)	(3)	(4)-IV	(5)
Ln(A)	0.481***	0.559***	0.132*	0.122	0.123*
	(0.11)	(0.10)	(0.07)	(0.0798)	(0.06)
Ln(E)	0.041*	0.034*	0.039***	0.0770	0.042***
	(0.02)	(0.02)	(0.01)	(0.0848)	(0.01)
Ln(Finc)	0.001	0.001	-0.002	-0.00156	-0.005
	(0.01)	(0.01)	(0.01)	(0.00660)	(0.01)
Ln(ParentEdu)	0.035***	0.031***	0.023***	0.0242***	0.019***
	(0.01)	(0.01)	(0.01)	(0.00670)	(0.01)
Ln(FamEduExp)					0.008**
					(0.00)
Male	-0.009*	-0.009*	0.008	0.00661	0.014*
	(0.01)	(0.01)	(0.01)	(0.00930)	(0.01)
Urban	0.014	0.012	0.007	0.00248	0.012
	(0.01)	(0.01)	(0.01)	(0.0135)	(0.01)
Constant	2.121***	1.800***	3.692***	3.654***	3.726***
	(0.44)	(0.39)	(0.31)	(0.224)	(0.26)
N	2,468	2,468	1,306	1268	1,248
R^2	0.41	0.468	0.068	0.056	0.094

Model Estimation: Wage Determination Equation(CCSS)

Table 5 Estimating wage equation (CCSS data)

Dependent variable: Ln(w)		
	Unweighted Sample	Weighted Sample
	(1)	(2)
College211 (vs Vocational)	0.444***	0.468***
	(0.035)	(0.040)
Tier1	0.279***	0.297***
	(0.034)	(0.040)
Tier23	0.162***	0.180***
	(0.031)	(0.038)
Ln(L)	0.499***	0.452***
	(0.058)	(0.090)
Ln(Finc)	0.045***	0.046***
	(0.008)	(0.010)
Ln(ParentEdu)	0.027	0.026
	(0.025)	(0.023)
Male	0.098***	0.102***
	(0.017)	(0.018)
Urban	-0.003	-0.004
	(0.019)	(0.021)
Constant	5.353***	5.526***
	(0.231)	(0.364)
N	3,148	3,145
R^2	0.149	0.152

Model Estimation: Parameter Value

Table 6 Estimated parameter values

	Parameter	Value		Parameter	Value
Human capital production function	α_0	Exp(3.692)	**Wage equation**	α_w	Exp(5.526)
	α_1(ability)	0.132		α_{P1}(Col 211)	0.468
	α_2(effort)	0.039		α_{P2} (Tier 1)	0.297
	α_{31}	-0.002		α_{P3} (Tier 2/3)	0.180
	α_{32}	0.023		α_L (HC)	0.452
	α_{33}	0.008		α_{Q1}	0.046
	α_{34}	0.007		α_{Q2}	0.026
	σ_ε	0.080		α_{Q3}	0.102
	$\hat{b}$	0.536		α_{Q4}	-0.004
Cost function	μ	4.052			

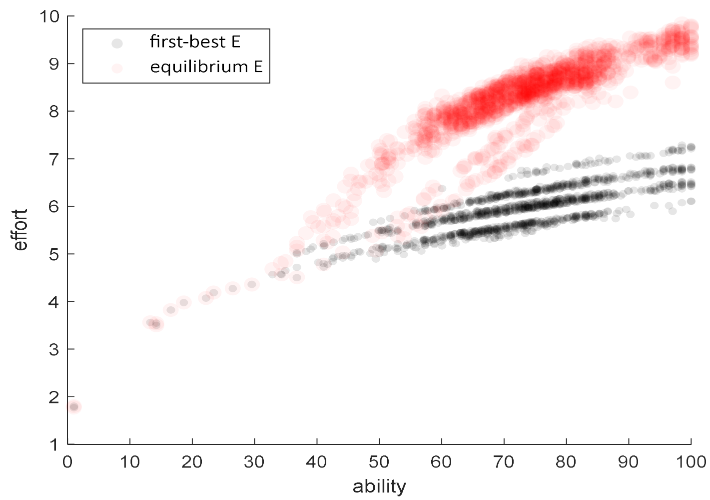
HC production function: $ln(\tilde{S}_i) = ln(\alpha_0) + \alpha_1\, ln(A_i) + \alpha_2\, ln(E_i) + \alpha_{3k}\, ln(Q_{ki}) + \varepsilon_i$.

Wage equation: $ln(w_i) = ln(\alpha_w) + \alpha_{P1} College211_i + \alpha_{P2} Tier1_i + \alpha_{P3} Tier23_i + \alpha_L\, ln(L_i) + \alpha_{Qk} Q_{ki} + \eta_i$

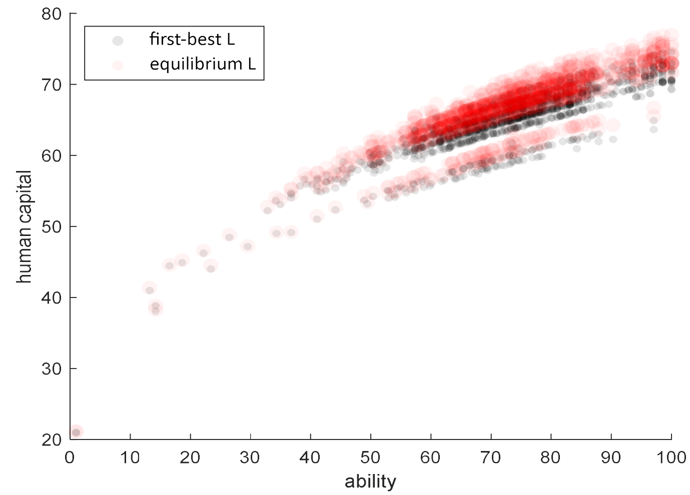
Cost function: $C(E_i, A_i) = \frac{{E_i}^{\mu}}{A_i}$

$\hat{b} = \frac{\mu(\ln(\tilde{S}_i/100))}{\mu(ln(L_i/100))}$ is a parameter to adjust CAT scores to true human capital (more details).

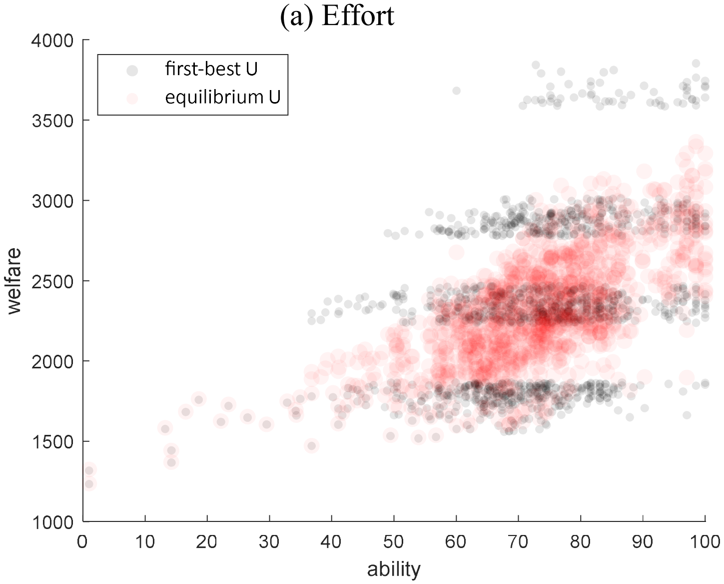
Simulation: Current System



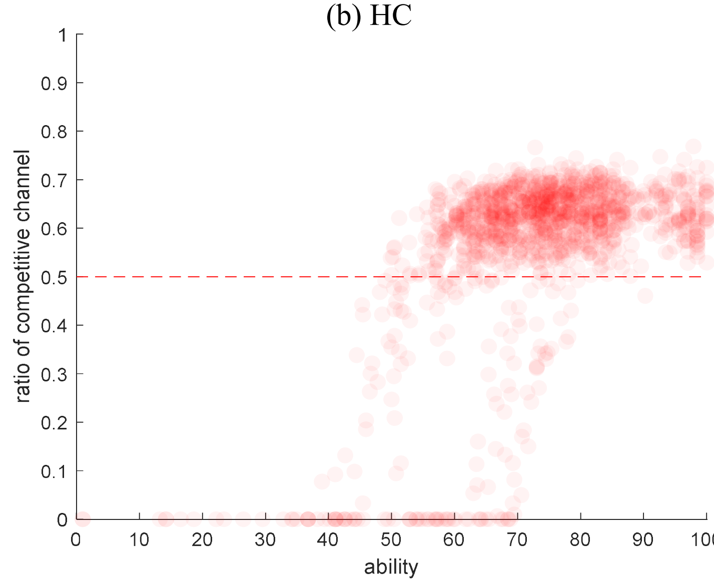
(a) Effort



(b) HC



(c) Welfare



(d) Ratio_c

Figure 10 Simulated equilibrium under the current system (red=eqm; black=FBO)

Simulation: The Benchmark (Current System)

Table 8 Simulated equilibrium under counterfactual policies

Counterfactual policy		**Average effort (hours/day)**	**Average human capital (0-100)**	**Average utility (RMB/ month)**	**First best average utility (RMB/ month**	**Dead-weight loss (%)**	**Competitive channel ratio (0-1)**	**Cutoff scores (0-100)**
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
First best and Benchmark	(0) first best (of benchmark)	5.94	64.61	2351	2351	0	0	71.31, 67.10, 63.17
	(1) Benchmark	8.23	66.59	2315	2351	1.53	0.564	78.39, 70.58, 63.72
Reduce Quality gap	(2) eliminate quality gap between tier 1&2 colleges	8.11	66.52	2316	2348	1.35	0.543	[78.12], 70.48, 63.70
	(3) eliminate quality gap between tier 2&3 colleges	8.00	66.46	2318	2344	1.12	0.524	78.15, [70.38], 63.61
	(4) eliminate quality gap between tier 3&4 colleges	7.74	66.27	2314	2339	1.07	0.428	78.37, 70.37, [63.26]
Increase quota of high-quality colleges	(5) increase quota of tier 1 college by 10%	8.18	66.56	2315	2348	1.41	0.557	77.96, 70.44, 63.61
	(6) increase quota of tier 1&2 colleges by 10%	8.07	66.49	2313	2342	1.26	0.538	77.87, 69.89, 63.10
	(7) increase quota of tier 1-3 colleges by 10%	8.02	66.47	2311	2339	1.21	0.524	77.86, 69.86, 62.28
Add one median-quality college	(8) add one college between tier 1&2 colleges	8.22	66.58	2315	2350	1.48	0.575	79.73, [74.55], 70.57, 63.71
	(9) add one college between tier 2&3 colleges	8.21	66.58	2315	2350	1.47	0.593	78.37, 72.30, [68.19], 63.71
	(10) add one college between tier 3&4 colleges	8.20	66.57	2314	2348	1.46	0.594	78.39, 70.56, 65.98, [61.51]
Change number of students	(11) reduce the total number of students by 10%	8.21	66.55	2359	2394	1.47	0.554	77.93, 69.87, 62.19
	(12) increase the total number of students by 10%	8.23	66.59	2282	2318	1.56	0.574	78.73, 71.11, 64.82

Simulation: The Benchmark (Current System)

- Equilibrium efforts (8.2hrs/day) are much higher than the first best efforts (5.9hrs/day), and the competitive channel dominates the productive channel (0.56).
- The deadweight loss equals to 1.5% reduction of wage level, one fourth to fifth of the return to one-year schooling (5-12%).
- There are also large variations of deadweight loss among students (std=9%), and high-ability students suffer more. (see Table A14)
- The results are largely robust to several alternative specifications (more details.)

Counterfactual Analysis

Part 4

Policy Design

Table 7 Counterfactual Policies

	Counterfactual Policy	**College quality** $(\alpha_{P1}, \alpha_{P2}, \alpha_{P3}, \alpha_{P4})$	**College quota (Σ=1248)** (q_1, q_2, q_3, q_4)
	(1) benchmark	(0.468, 0.297, 0.180, 0)	(59, 303, 489, 397)
reduce quality gap among colleges	(2) eliminate gap between tier 1&2	(0.325, 0.325, 0.180, 0)	(59, 303, 489, 397)
	(3) eliminate gap between tier 2&3	(0.468, 0.2248,0.2248, 0)	(59, 303, 489, 397)
	(4) eliminate gap between tier 3&4	(0.468, 0.297, 0.0994, 0.0994)	(59, 303, 489, 397)
Increase high-quality college quota	(5) tier 1 by 10%, with quality reduced by 10%	(0.4212, 0.297, 0.180, 0)	(65, 303, 489, 391)
	(6) tier 1&2 by 10%, with quality reduced by 10%	(0.4212, 0.2673, 0.180, 0)	(65, 333, 489, 361)
	(7) tier 1-3 by 10%, with quality reduced by 10%	(0.4212, 0.2673, 0.162, 0)	(65, 333, 538, 312)
Add one median-quality college (between tier 1&2, tier 2&3, etc.)	(8) $\alpha_{P5} = (\alpha_{P1} * q_1 + \alpha_{P2} * q_2)/(q_1 + q_2)$, $q_5 = (q_1 + q_2)/3$	(0.468, 0.325=α_{P5}, 0.297, 0.180, 0)	(39, 121=q_5, 202, 489, 397)
	(9) $\alpha_{P5} = (\alpha_{P2} * q_2 + \alpha_{P3} * q_3)/(q_2 + q_3)$, $q_5 = (q_2 + q_3)/3$	(0.468, 0.297, 0.225=α_{P5}, 0.180, 0)	(59, 202, 264=q_5, 326, 397)
	(10) $\alpha_{P5} = (\alpha_{P3} * q_3 + \alpha_{P4} * q_4)/(q_3 + q_4)$, $q_5 = (q_3 + q_4)/3$	(0.468, 0.297, 0.180, 0.099=α_{P5}, 0)	(59, 303, 326, 295=q_5, 265)
Change number of students	(11) reduce the number of students by 10%	(0.468, 0.297, 0.180, 0)	(59, 303, 489, 272)
	(12) increase the number of students by 10%	(0.468, 0.297, 0.180, 0)	(59, 303, 489, 522)

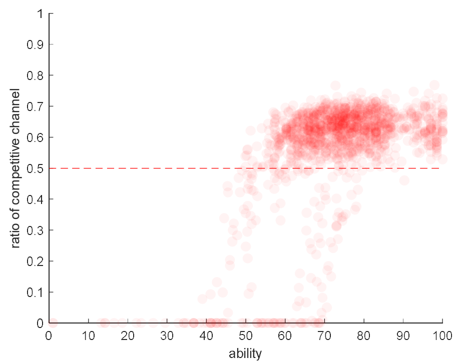
All the designs (except (11)-(12)) keep constant the total college quality in market.

Simulation: Counterfactual Policies

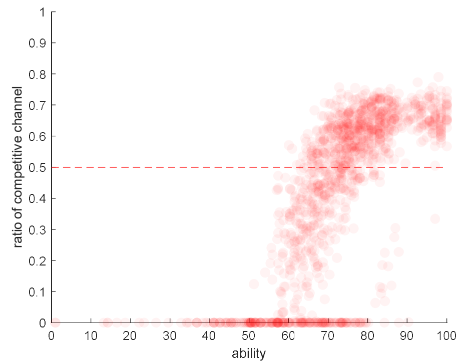
Table 8 Simulated equilibrium under counterfactual policies

Counterfactual policy		**Average effort (hours/day)**	**Average human capital (0-100)**	**Average utility (RMB/ month)**	**First best average utility (RMB/ month**	**Dead-weight loss (%)**	**Competitive channel ratio (0-1)**	**Cutoff scores (0-100)**
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
First best and Benchmark	(0) first best (of benchmark)	5.94	64.61	2351	2351	0	0	71.31, 67.10, 63.17
	(1) benchmark	8.23	66.59	2315	2351	1.53	0.564	78.39, 70.58, 63.72
Reduce Quality gap	(2) eliminate quality gap between tier 1&2 colleges	8.11	66.52	2316	2348	1.35	0.543	[78.12], 70.48, 63.70
	(3) eliminate quality gap between tier 2&3 colleges	8.00	66.46	2318	2344	1.12	0.524	78.15, [70.38], 63.61
	(4) eliminate quality gap between tier 3&4 colleges	**7.74**	**66.27**	**2314**	**2339**	**1.07**	**0.428**	**78.37, 70.37, [63.26]**
Increase quota of high-quality colleges	(5) increase quota of tier 1 college by 10%	8.18	66.56	2315	2348	1.41	0.557	77.96, 70.44, 63.61
	(6) increase quota of tier 1&2 colleges by 10%	8.07	66.49	2313	2342	1.26	0.538	77.87, 69.89, 63.10
	(7) increase quota of tier 1-3 colleges by 10%	**8.02**	**66.47**	**2311**	**2339**	**1.21**	**0.524**	**77.86, 69.86, 62.28**
Add one median-quality college	(8) add one college between tier 1&2 colleges	8.22	66.58	2315	2350	1.48	0.575	79.73, [74.55], 70.57, 63.71
	(9) add one college between tier 2&3 colleges	8.21	66.58	2315	2350	1.47	0.593	78.37, 72.30, [68.19], 63.71
	(10) add one college between tier 3&4 colleges	8.20	66.57	2314	2348	1.46	0.594	78.39, 70.56, 65.98, [61.51]
Change number of students	(11) reduce the total number of students by 10%	8.21	66.55	2359	2394	1.47	0.554	77.93, 69.87, 62.19
	(12) increase the total number of students by 10%	8.23	66.59	2282	2318	1.56	0.574	78.73, 71.11, 64.82

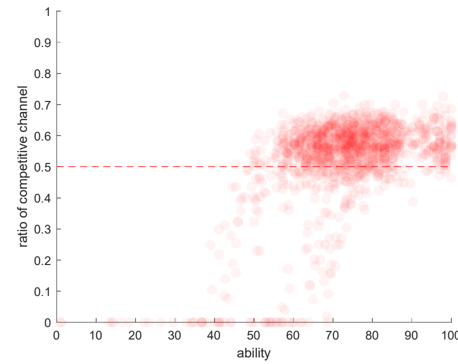
Simulation: Counterfactual Policies



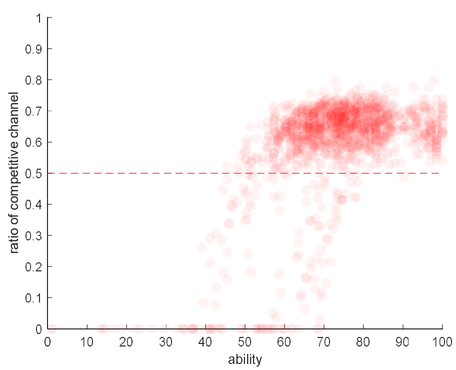
(a) benchmark



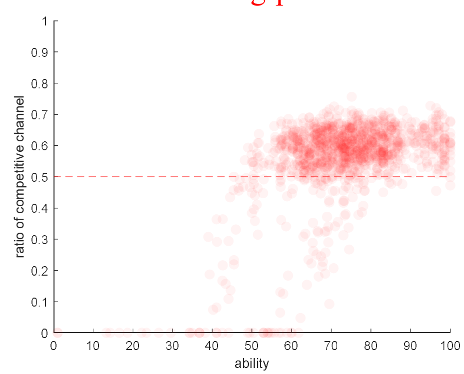
(b) eliminate quality gap between grp 3&4



(c) increase quota of grp1-3 by 10%



(d) add one median-quality college



(e) reduce the number of students by 10%

Figure A2 Competitive channel vs. productive channel under counterfactual policies

Counterfactual Analysis: Simulated Results

- Policies have positive effects on alleviating excessive efforts and reducing deadweight loss, despite of:
  - "constant total college quality/quantity" design
  - Ambiguous (offsetting) effects of some policies
- Reducing quality-gap between *low-quality* colleges and expanding college quota of *all colleges except the lowest-quality group* seem to be more effective in alleviating excess efforts and welfare loss.
- We also compare counterfactual policies under (hypothetical) equilibrium without perturbations. The policy effect in general is larger (more details).

Conclusion

Part 5

Conclusion

- Excessive effort and welfare loss are sizable in the college admissions competition
  - Students tend to generate efforts larger than optimal level and suffer from welfare loss.
  - The competitive channel dominates the productive channel.
  - There can also be large distributive effects among students; those with abilities near the cutoffs of colleges with large quality-gap suffer more.
- Supply side policies by changing quality-quantity composition in the college market can alleviate the excess effort and deadweight loss
  - Narrowing the gap in college quality and expanding the quota of higher-quality ones can reduce excessive efforts and reduce deadweight loss, especially at the low end of college market.

Future Research Questions

- Multidimensional efforts?
- Other Policies:
  - Exam reform: Freely choosing subjects; (hypothetical) restrictions on repetitive college entrance exam taking
  - Affirmative policies (“strong-foundation” plan, “self-strengthening” plan......)
  - Outside option: high school tracking system; oversea higher education market

Thank You!

College Admission Policies and Pre-College Human Capital Investment: Evidence from China

by Chenglei Diao, Chenyuan Liu, and Xiaohan Zhong
Contact the presenter: zhongxh@sem.tsinghua.edu.cn

Man may be either broken in, trained, and mechanically taught, or he may be really enlightened.

-Immanuel Kant

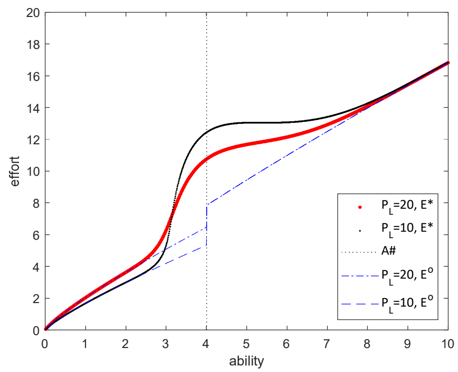
Competitive Equilibrium: Numerical Example

Table 1 Parameter setting in numerical simulation (benchmark)

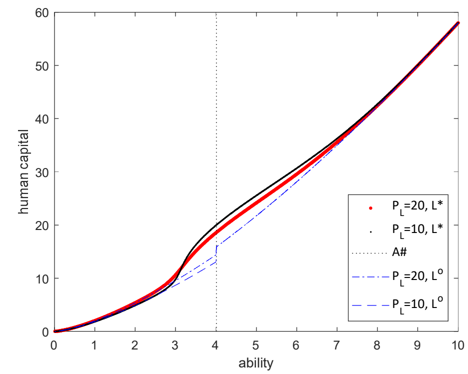
	Meaning	Value
Human capital production function	Constant term	$\alpha_0 = \sqrt{2}$=1.41
	Power of student ability (A)	$\alpha_1 = 1$
	Power of student effort (E)	$\alpha_2 = 0.5$
	Random perturbation	$\varepsilon \sim N(0, \sigma_\varepsilon^2)$, $\sigma_\varepsilon = 0, 0.16, 0.32, 0.64$
Cost function	Power of student effort (E)	$\mu = 2$
Wage determination equation	Constant term	$\alpha_w = 7 * 2^{0.1} = 7.50$
	Power of college quality (P)	$\alpha^P = 0.5$
	Power of human capital (L)	$\alpha^L = 0.4$
Other Parameters	Number of Students	$N = 1000$
	Student ability distribution	$A \sim \text{U}(0,10)$
	College qualities	$P_H = 40;\ P_L = 10$
	Quota proportion of the high-quality college	$q_H = 0.6$

Back

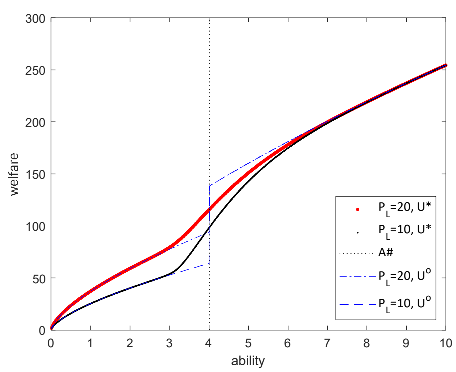
Comparative Statics: Decreasing Quality Gap (by increasing low-quality college quality)



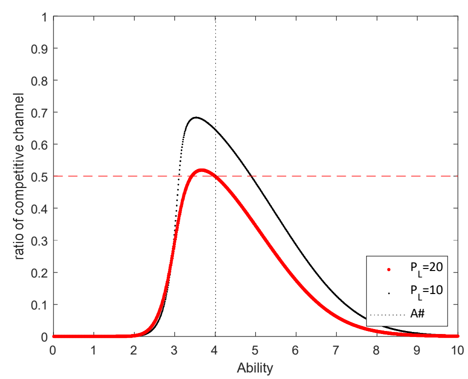
(a) effort



(b) HC



(c) welfare

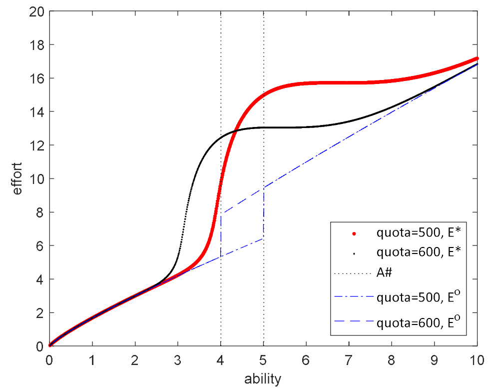


(d) ratio_c

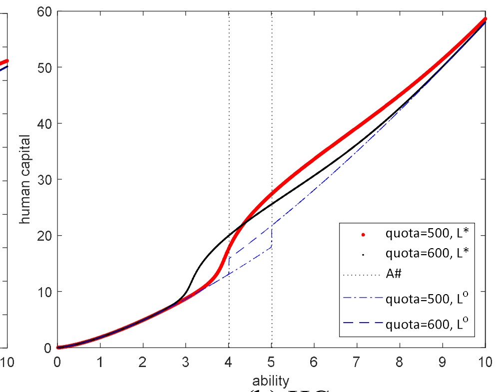
Figure 4 Comparative statics: Decreasing quality gap

P_H=40, P_L=10/20.

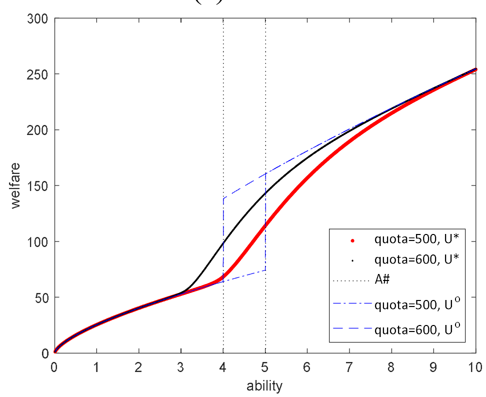
Comparative Statics: Decreasing High-Quality College Quota



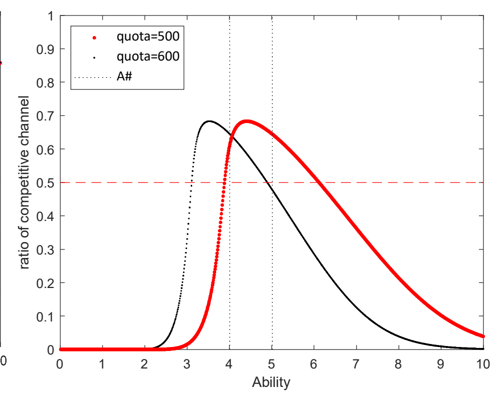
(a) effort



(b) HC



(c) welfare

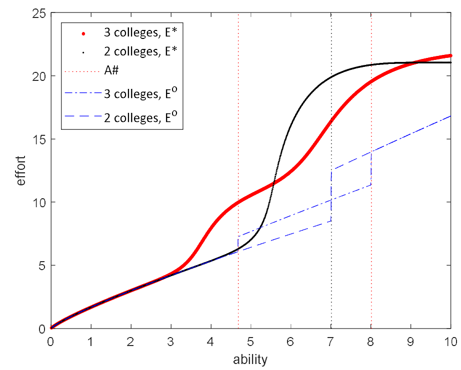


(d) ratio_c

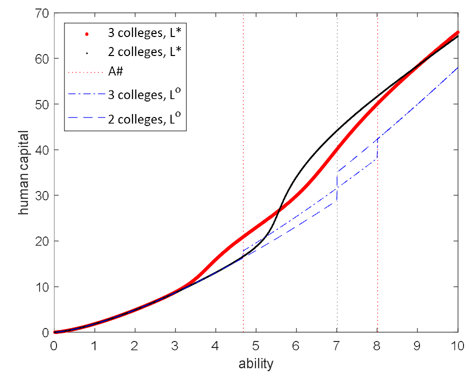
Figure 5 Comparative Statics: Decreasing high-quality college quota

q_H=600/500.

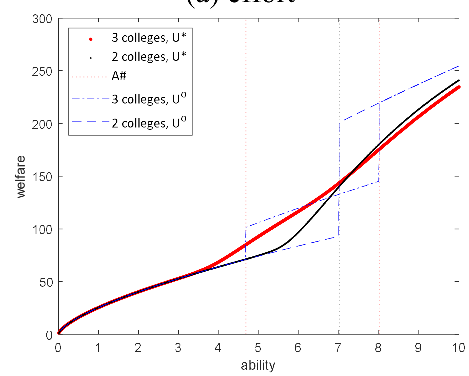
Comparative Statics: Adding One Median-Quality College



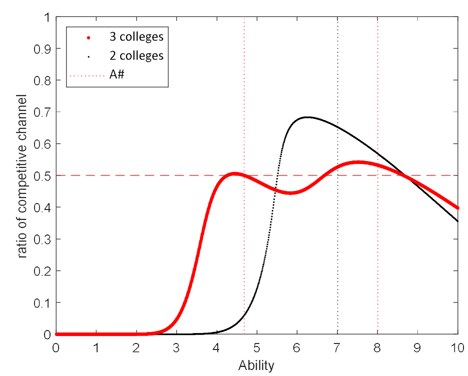
(a) effort



(b) HC



(c) welfare



(d) ratio_c

Figure 6 Comparative statics: Adding one median-quality college

2/3 colleges

Comparative Statics: Some Nonmonotonic Effects (on Deadweight loss)

- With the increase of the random perturbation ($\sigma_\varepsilon \uparrow$), deadweight loss ($\frac{\sum_{i=1}^{N}(U_i^o - U_i^*)}{\sum_{i=1}^{N} U_i^o}$) may change nonmonotonically.
  - $\downarrow$intensive margin (effort level at the cutoff) vs $\uparrow$mismatch(Table A1)

- With the increase of the high-quality college quota, deadweight loss may change nonmonotonically (as well as effort level).
  - $\downarrow$intensive margin vs $\uparrow$extensive margin (# of marginal students)(Table A3)

- With adding median-quality colleges, deadweight loss may change nonmonotonically.
  - $\downarrow$ intensive (depth of quality gaps) vs $\uparrow$extensive margin (# of gaps)(Table A4)

Back

CFPS: Statistic Summary

Table A7 CFPS Variables: Statistic Summary

Variables	**Meanings**	**All regular HS students**	**Sub-sample 1**	**Sub-sample 2**
		(1)	(2)	(3)
Student Ability (A)	Early cognitive ability test (CAT) scores 0-100)		73.64	72.94
Student effort (E)	Average daily study time in high school (hours)	8.92	8.59	8.23
Student pre-college human capital (L)	CAT scores after graduation from high school (0-100)		77.55	75.51
College quality (P)	211 universities and above			0.05
	Batch-1 university, excluding 211			0.24
	Other universities (batch-2&3)			0.39
	non-university tertiary (vocational) colleges			0.32
Wage after graduation (w)	Wage income in the first year after graduation (unit: 1K RMB)			37.94
Gender (male=1)	Gender	0.49	0.49	0.50
Residency (urban=1)	Whether an urban hukou	0.28	0.22	0.17
Finc	Annual household income per capita (unit: 1K RMB)	11.08	7.81	6.91
ParentEdu	Maximum number of years of education for parents	8.85	8.47	8.13
FamEduExp	The education expenses paid by the family for the student last year (unit: 1K RMB)	4.97	3.52	2.87
Sample size		4,675	2,668	1,248

Back

CCSS: Statistic Summary

Table A8 CCSS Variables: Statistic Summary

Variables	Meanings	mean	std
Wage	Highest monthly wage offer after graduation (unit: 1k RMB)	2.997	2.153
L	Expected pre-college human capital($E(\theta^0)$), normalized to 0-100.	66.084	10.304
College211	Whether attending 211 universities, Yes=1，No=1	0.255	0.444
Tier1	Whether attending batch-1 universities (except for 211 colleges), Yes=1, No=0	0.318	0.466
Tier23	Whether attending batch-2/3 universities, Yes=1, No=0.	0.298	0.457
Vocation	Whether attending vocational college, Yes=1, No=0	0.130	0.336
Finc	Annual household income per capita (unit: 1k RMB)	19.536	18.699
ParentEdu	Maximum number of years of education for parents	10.931	3.072
Male	Male=1，Female=0	0.543	0.498
Urban	Urban hukou=1，rural hukou=0	0.425	0.494
Track	Science=1，humanity=0	0.735	0.441

Back

Model Estimation: Adjust CAT to HC

- An issue is that Cognitive Ability Test (CAT) scores after high school graduation ($ln(\tilde{S}_i)$), as a measure of pre-college human capital, may be systematically different from true human capital estimated from CCSS dataset ($ln(L_i)$), although both $\widetilde{S_i}$ and L_i are normalized to 0-100.
- We estimate the following relation between the two:

$$ln(\widetilde{S_i}/100) = b\, ln\left(\hat{L}_i/100\right) + \delta_i$$

with $\delta_i \sim N(0, \sigma_\delta^2)$, through the following procedure: (1) We predict L_i in CCSS by linear regression on covariates of college qualities dummies and individual characteristics (gender, hukou, income, parental education), (2) Replace all the covariate values with its mean value in CFPS in the estimated linear regression to get a predicted mean value of $\mu(ln(L_i))$ for the sample of CFPS, (3) Predict b as:

$$\hat{b} = \frac{\mu(\ln\left(\widetilde{S_i}/100\right))}{\mu(ln\left(\hat{L}_i/100\right))}$$

Where $\mu(\ln\left(\widetilde{S_i}\right))$ is the sample mean in CFPS. Back

Simulation: The Benchmark (Current System)

Table A14 Distributive Effect under the Benchmark: OLS regression

	Excess effort (%)	Deadweight loss (%)	Competitive channel ratio (%)
	(1)	(2)	(3)
Ability	0.406***	0.090**	0.690***
	(0.017)	(0.037)	(0.030)
Ln(Finc)	0.332	-0.862	0.240
	(0.283)	(0.634)	(0.503)
Ln(ParentEdu)	1.141***	0.086	4.072***
	(0.341)	(0.765)	(0.606)
Urban	-1.626**	0.138	-3.075***
	(0.643)	(1.442)	(1.143)
Male	-0.146	-2.058**	0.798
	(0.447)	(1.001)	(0.793)
Dep. Var. Mean	38.316	-1.188	56.384
Std	9.724	17.692	17.306
Min	0.000	-58.586	0.000
Max	57.410	39.230	76.837
N	1,248	1,248	1,248
R^2	0.347	0.010	0.350

Back

Robust Check on Benchmark

Table A15 Robust Check on Benchmark

Scenario	Average effort (hours/ day)	Average human capital (0-100)	Average utility (RMB/ month)	First-best average utility	Dead-weight loss (%)	Competitive channel ratio (0-1)	Cutoff Scores (0-100)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(0) Benchmark: α_1=0.132, α_2=0.039 (with μ=4.052)	8.23	66.59	2315	2351	1.53	0.564	78.39,70.58,63.72
(1) Changing HC function: α_1=0.559, α_2=0.034 (with μ=3.767)	8.22	59.55	2234	2260	1.14	0.503	83.25,66.52,53.88
(2) Changing HC function (IV): α_1=0.122, α_2=0.077 (with μ=4.322)	8.29	66.98	2261	2322	2.62	0.553	79.01,71.14,64.11
(3) Changing perturbation: σ_ε=0.046 (with μ=4.104)	8.29	66.61	2305	2349	1.87	0.580	75.81,69.97,64.55
(4) With sample weights (with μ=4.053)	8.24	66.56	2319	2352	1.42	0.566	78.25,70.40,63.64

Back

Simulated Equilibrium without Perturbations

Table A16 Equilibrium and selected counterfactuals without random perturbations

Scenario	Average effort (hours/ day)	Average human capital (0-100)	Average utility (RMB/ month)	First-best average utility	Deadweight loss (%)	Cutoff Scores (0-100)
	(1)	(2)	(3)	(4)	(5)	(6)
(1') current system	7.85	65.48	2239	2351	4.74%	76.29, 70.80, 67.07
(4') eliminate quality gap between tier 3&4 colleges	7.21	64.60	2280	2339	2.52%	75.80, 71.33, [63.22]
(7') increase quota of tier 1-3 colleges by 10%	7.93	65.79	2261	2339	3.35%	75.96, 70.30, 66.00
(8') add one college between tier 1&2 colleges	8.18	65.98	2274	2350	3.23%	76.13, [71.97], 70.80, 67.07
(11') reduce the total number of students by 10%	8.06	65.84	2302	2394	3.85%	76.07, 70.35, 66.00

Back