College Affirmative Action and Students' Long-Term Outcomes: A Matter of the Dose

Yi Chen[1] Chao Fu[2] Hongbin Li[3] Teng Li[4] Siyan Tang[4]

[1]ShanghaiTech University

[2]University of Wisconsin-Madison

[3]Stanford University

[4]Sun Yat-sen University

March 12, 2026

Motivation

- Affirmative action in higher education: designed to increase the representation of historically disadvantaged groups or Underrepresented Minority (URM) on campus.
  - Promoting equity by considering group (racial, ethnic, SES) membership alongside academic merit during admissions.
- Affirmative actions are prevalent across the world and vary widely across countries:
  - US, Brazil, South Africa: Race-conscious admissions.
  - India: Constitutionally mandated reservations for Scheduled Castes (SC), Scheduled Tribes (ST), and Other Backward Classes (OBC) in public universities.
  - **China: Bonus points policy in college admissions to support ethnic minority students.**

Motivation

- Putting aside the effect of AA on resource reallocation and focusing on AA-targeted students, debates remain:
  - **Affirmative Side: AA improves opportunities for the disadvantaged.**
    - Increase their college attendance and college quality.
  - **Negative Side: the Mismatch hypothesis**
    - Students may struggle if they are enrolled in programs for which they are under-prepared; they would benefit more from less selective programs.

Motivation

- This debate is hard to settle, because there is truth to both sides: both the quality of the program and match quality can matter.

Motivation

- This debate is hard to settle, because there is truth to both sides: both the quality of the program and match quality can matter.

- Instead of a dichotomous view—AA benefits/hurts the disadvantaged—a more productive discussion should focus on the appropriate *dose* of AA.

Motivation

- This debate is hard to settle, because there is truth to both sides: both the quality of the program and match quality can matter.

- Instead of a dichotomous view—AA benefits/hurts the disadvantaged—a more productive discussion should focus on the appropriate *dose* of AA.

- Although intuitive, this is a hard empirical question because we often do not have transparent variation in *doses*.
  1. In some cases, *variation exists but is murky.* For example, US universities often practice their own AA, but their "holistic admissions" use hard-to-quantify factors, making it hard to know the dose of AA.

  2. In other cases, *AA policy is relatively transparent but has rather limited variation.* Policy adoption or repeal leads to effectively binary variation (e.g., the ban of AA in California).

This Paper

- We take a step forward by answering the question: How does the dose of AA affect short- and long-run outcomes among intended beneficiaries?

- **Research design**:
  - Exploit the combination of centralized college admissions and AA policies in China, where the dose of AA is both transparent and continuously varying.
  - Administrative data of admission outcomes for all college entrance exam (CEE) test takers, 1999-2003.
  - Detailed consumption (and a subset of income) data from UnionPay Merchant Services in 10-23 years after taking the CEE.
  - The Chinese College Student Survey (CCSS) with evidence to bridge the immediate and long-run outcomes.

Preview of Results

- **Effects on college admission**:
  - Increasing the probability of (elite) college admission.
  - Worse ranking relative to peers.
  - Lower-earning majors.

- **Impact on long-run outcomes**:
  - On average, lowering consumption and income.
  - The dose makes the difference:
    - Bottom quintile of bonus intensity: sizable gains, +19% long-run consumption.
    - Middle two quintiles: no detectable long-run effects.
    - Top two quintiles: large losses, 14-28% decline in long-run consumption.
  - The effect is stable across 10-23 years after taking the CEE.
  - Stronger negative effects for students with lower CEE ranking, males, from rural areas, and/or low-quality high schools.

- **Channels**:
  - Less satisfied with GPA, interpersonal relationships, and overall college life.

Literature: Affirmative action

- Extensive literature on the efficacy of affirmative action:
  - **Affirmative action works as intended.**
    - Increased pre-college human capital, college attendance, and college quality (Bagde et al., 2016; Estevan et al., 2019; Bleemer, 2023; Akhtari et al., 2024).
    - Higher early and mid-career earnings (Bleemer, 2024; Machado et al., 2025).
    - Promote socioeconomic mobility (Bleemer, 2022; Black et al., 2023).
  - **Mismatch hypothesis**
    - Quality-fit trade-off (Arcidiacono and Lovenheim, 2016; Dillon and Smith, 2020).
    - Poorer academic performance, along with lower persistence and degree completion due to mismatch (Arcidiacono et al., 2011, 2016; Barrow et al., 2020).
- **Our contributions**:
  - Exploiting annual variation in effective bonus values to evaluate how the intensity of affirmative action affects its intended beneficiaries.
  - Examining long-term economic well-being.

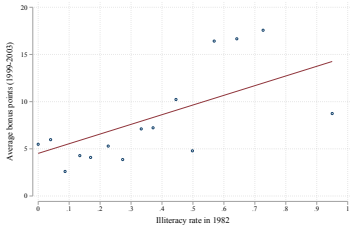
Outline

College Entrance Examination in China

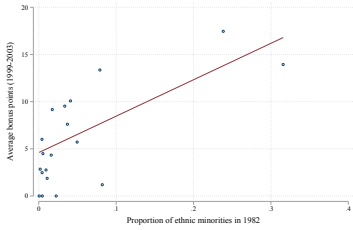
- A comprehensive set of exams assessing high school students' abilities in Math, Chinese, English, and track-specific subjects (e.g., Physics for Science, or History for Humanities).
- Students select either Science or Humanities track in Grade 10, undergo province-track-specific training, and take the CEE at the end of Grade 12.
- Colleges across China allocate province-specific admissions quotas for each major.
- Within each province, admissions are centralized, with applicants ranked solely by CEE (plus bonus) scores.
- Students compete within a province-cohort-track.

Bonus point policy in China

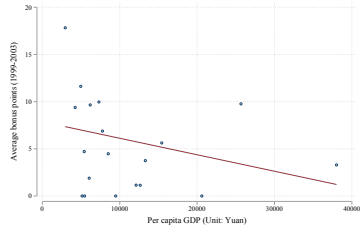
- China's ethnicity-based affirmative action is one of the largest state-sponsored preferential policies for ethnic minorities.
  - Minority students receive bonus points (5-30 points) added directly to their CEE scores.
  - There is a significant correlation between minority status and poverty.
  - Preferential policies are seen as tools for promoting equality.



(a) Illiteracy rate in 1982



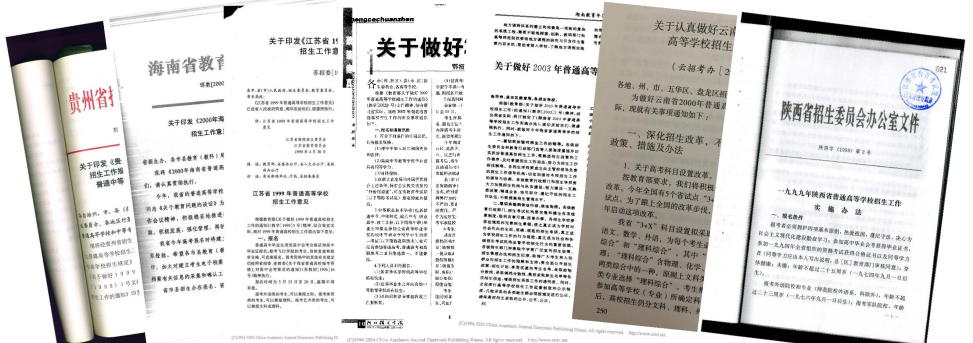
(b) Proportion of minorities in 1982



(c) Per capita GDP

Administrative documents on bonus point policy

- To isolate bonus points related to AA for minorities from other bonus points (e.g., athletes), we collected bonus point policies from different provinces and cleaned the individual bonus points accordingly.
  - Contacting all provincial Departments of Education and searching throughout the E-libraries and university libraries.
  - We collected 68 official documents on bonus point policies and supplemented the remaining information with contemporaneous news sources.
  - In return, a comprehensive dataset of AA policies covering all but four provinces. details



Outline

Administrative data I: All CEE takers, 1999-2003

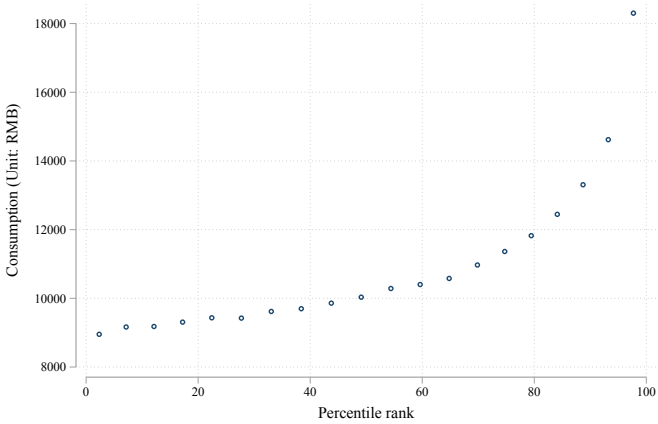
- Registration information (gender, birth year, track, ethnicity, exam county, high school, etc).
  - 5% of all test takers are minorities.
- Test scores.
- Bonus points.
- Admission outcomes (college and major, if admitted).

Administrative data II: Consumption data from China UMS

- UnionPay Merchant Services (UMS，银联商务)
- Detailed consumption data in 2013-2022 (date, category, amount, and bank card type).
- Universe of consumer spending transactions at offline merchants using bank cards and Quick Response (QR) codes (i.e., linked to e-wallets in Alipay and WeChat pay) captured by UMS.
- UnionPay bank card transactions represented about 80% of overall retail consumption in 2018 (Lai et al., 2022), though this share has subsequently fallen.
- The two datasets pre-merged by UnionPay.

Administrative data II: Consumption data from China UMS

- Correlation between CEE rank and consumption.



- A portion of UMS's business involves payroll services for client firms, which provides a subset of annual income data in the database. Among these individuals, the correlation between consumption and income is 0.6.

Administrative data III: Chinese College Student Survey (CCSS)

- National representative survey of graduating college students (2010-2015) across China.
- 90 colleges, selected to represent all regions and tiers (985/211, provincial, vocational).
- Roughly 40,000 respondents, with rich data on CEE scores, ethnicity, college and majors, GPA ranks, college life satisfaction, etc.
- Include occupation and wage offer information at graduation.

Summary statistics

Variable	Observations	Mean	Std. Dev.
Panel A: Demographic characteristics			
Female	9,015,498	0.40	0.49
Age when taking CEE	9,015,498	19.0	1.11
Rural	9,015,498	0.54	0.50
Minority	9,015,498	0.04	0.20
Science track	9,015,498	0.55	0.50
Humanity track	9,015,498	0.27	0.44
Comprehensive track	9,015,498	0.19	0.40
Panel B: CEE performance			
CEE score	9,015,498	426.19	96.07
Rankings	9,015,498	49.76	28.81
Bonus points	9,015,498	0.47	2.70
Bonus points> 0	9,015,498	0.04	0.19
Bonus points (Conditional on Bonus>0)	322,878	13.11	6.18
AA doses (Conditional on Bonus>0)	322,878	0.16	0.08

Summary statistics

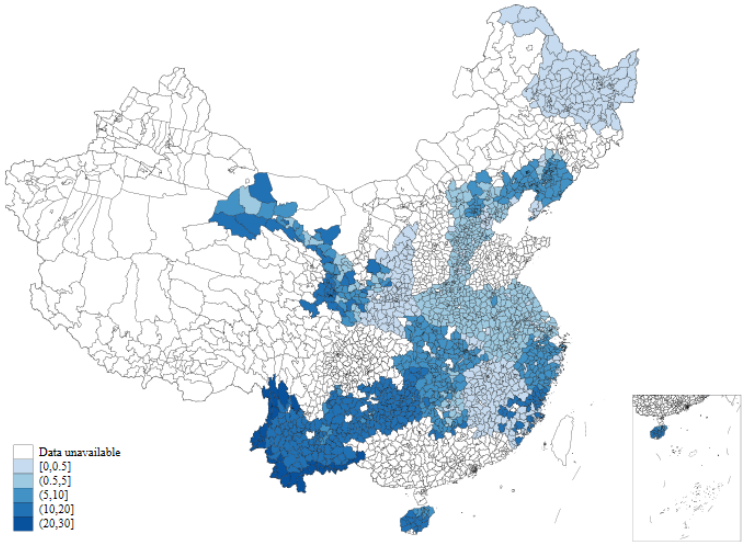
Variable	Observations	Mean	Std. Dev.
Panel C: Consumption & Income (RMB)			
Average annual consumption	9,015,498	10,962.9	34,224.2
Durables	9,015,498	4,329.7	18,415.0
Non-durables	9,015,498	571.8	1,612.7
Leisure	9,015,498	27.5	127.4
Financial Investment	9,015,498	1,740.2	8,160.6
Insurance	9,015,498	76.5	354.8
Bond & stocks	9,015,498	1,369.4	6,816.5
Average annual income	144,172	76,985.9	194,921.3

Empirical strategy - bonus intensity

- Bonus points are stable within region-ethnicity during our sample period.
- But even the same bonus points may have different values across years due to the random realization of CEE score distributions.
  - Relative scale: +30 is 15% of the s.d. in Year A (30/200), but 60% of the s.d. in Year B (30/50).
- We create two measures on bonus intensity:
  - Intensity = (bonus points)/(s.d. of CEE scores).
  - Intensity = (bonus points)/(cutoff score for elite universities - mean CEE scores).
  - The correlations between the two intensity measures is about 0.8.

Empirical strategy - bonus intensity

- Variations in bonus intensity within county.



Empirical strategy - specification

- **Recentering by Borusyak and Hull (2023, EMCA)**:
  - Bonus intensities are always larger for some regions.
  - Exploit temporal variation in bonus intensity within county-ethnicity-track cells, consistent with the approach of controlling for historical mean bonus intensity within each cell.

$$Y_{ijest} = \beta D_{jest} + \gamma Rank_{ijest} + X'_{ijest}\delta + \lambda_{jes} + \theta_{jst} + \epsilon_{ijest}. \quad (1)$$

- Y_{ijest}: a generic outcome, where i, e, j, s, and t denote individual, ethnicity, county, track, and cohort, respectively. (Modified log by Chen and Roth (2024))
- D_{jest}: the bonus intensity for individual i in CEE.
- $Rank_{ijest}$: individual i's CEE rank within cohort-province-track.
- X'_{ijest}: individual characteristics, including gender, age when taking the CEE, rural hukou indicator, and high school quality.
- λ_{jes}: county-ethnicity-track FE.
- θ_{jst}: county-cohort-track FE.
- ϵ_{ijest}: clustered at the county-cohort-track level.

Verifying randomness of bonus intensity

	(1)	(2)	(3)	(4)
	Dependent variable: CEE rank (without bonus)			
AA doses	-23.563***	-4.474	-4.627	1.401
	(1.690)	(3.787)	(3.968)	(2.373)
County-ethnicity-track FE	NO	YES	YES	YES
County-cohort-track FE	NO	NO	YES	YES
Controls	NO	NO	NO	YES
Obs.	9,089,235	9,089,235	9,089,235	9,089,235
R2	0.00	0.04	0.05	0.29

Notes: This table assesses the potential endogeneity of AA doses by regressing students' original CEE ranks on the AA doses they receive. Column (1) reports the unconditional correlation without fixed effects. Column (2) adds county-ethnicity-track fixed effects, and column (3) further includes county-cohort-track fixed effects. Column (4) extends the specification in column (3) by adding a set of individual-level controls, including gender, age at CEE, rural hukou status, and high school quality. * p<0.1, ** p<0.05, *** p<0.01.

Outline

Admission results

	(1) Admission	(2) Admitted by elite colleges	(3) Admitted by top choice	(4) College quality
Panel A: Linear effect				
AA doses	0.310***	0.097***	0.283***	22.104***
	(0.025)	(0.020)	(0.024)	(1.240)
Panel B: Non-linear effect				
1st quintile	0.065***	0.008*	0.036***	2.528***
	(0.006)	(0.004)	(0.005)	(0.289)
2nd quintile	0.031***	-0.001	0.045***	1.394***
	(0.006)	(0.006)	(0.006)	(0.285)
3rd quintile	0.043***	0.015***	0.052***	2.699***
	(0.007)	(0.004)	(0.006)	(0.328)
4th quintile	0.063***	0.026***	0.056***	4.939***
	(0.006)	(0.005)	(0.006)	(0.359)
5th quintile	0.095***	0.025***	0.076***	6.128***
	(0.007)	(0.005)	(0.006)	(0.328)
Controls	YES	YES	YES	YES
County-ethnicity-track FE	YES	YES	YES	YES
County-cohort-track FE	YES	YES	YES	YES
Obs.	9,105,498	9,105,498	9,105,498	5,551,765

Notes: College quality is measured by the leave-one-out average of province-cohort-track percentile rankings of admitted students (ranging from 0 to 100, excluding students with bonus points).

Effect of 10 points over mean of s.d. of CEE scores (=0.1):

- ↑ 3.1% in the Prob. of being admitted to college (5.1% of sample mean)
- ↑ 0.97% in the Prob. of being admitted to an elite college (8.1% of sample mean)
- ↑ 2.83% in the Prob. of being admitted by first choice (6.9% of sample mean)
- ↑ 2.21 points in college quality (3.6% of sample mean)

Relative ranking to peers

	(1) Relative ranking	(2) Relative score
Panel A: Linear effect		
AA doses	-20.344***	-32.357***
	(0.874)	(4.022)
Panel B: Non-linear effect		
1st quintile	-1.283***	-2.673**
	(0.199)	(1.048)
2nd quintile	-1.769***	-4.053***
	(0.200)	(0.972)
3rd quintile	-2.479***	-5.876***
	(0.221)	(0.790)
4th quintile	-4.169***	-8.610***
	(0.247)	(1.133)
5th quintile	-5.468***	-7.049***
	(0.243)	(1.055)
Controls	YES	YES
County-ethnicity-track FE	YES	YES
County-cohort-track FE	YES	YES
Obs.	5,554,036	4,122,707

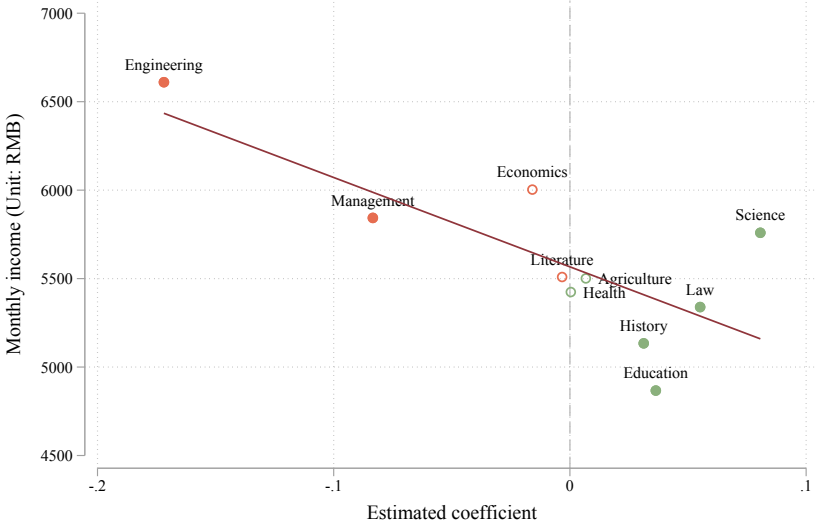
Notes:

- Relative ranking: an individual's percentile rank minus the average percentile rank of peers from the same college and major (excluding the individual).
- Relative score: an individual's CEE score minus the average score of peers from the same college, same major, and using the same format of the exam (excluding the individual). exam formats

Effect of 10 points over mean of s.d. of CEE scores (=0.1):

- Associated with 2.0 percentile points lower than the average rank of peers (13.6% of sample s.d.)
- Associated with 3.2 points lower in exam scores relative to peers (7.3% of sample s.d.)

Major selection



Notes:

- Y-axis: Average income associated with each major, with data sourced from the 2023 Chinese College Graduate Employment Report published by MyCOS.
- X-axis: Estimated coefficients representing the effect of bonus intensity on the prob. of choosing each major.

Major selection

	(1) Top-three majors	(2) Bottom-three majors
Panel A: Linear effect		
AA doses	-0.044	0.142***
	(0.032)	(0.030)
Panel B: Non-linear effect		
1^{st} quintile	-0.008	0.001
	(0.007)	(0.005)
2^{nd} quintile	0.004	-0.002
	(0.006)	(0.005)
3^{rd} quintile	0.005	0.006
	(0.007)	(0.006)
4^{th} quintile	-0.011	0.023***
	(0.009)	(0.008)
5^{th} quintile	-0.011	0.056***
	(0.009)	(0.009)
Controls	YES	YES
County-ethnicity-track FE	YES	YES
County-cohort-track FE	YES	YES
Obs.	5,554,036	5,554,036

Note: "Top-three major" and "Bottom-three major" are defined as a dummy variable equal to 1 if the individual chooses one of the top/bottom three majors with the highest average consumption.

Effect of 10 points over mean of s.d. of CEE scores (=0.1):

- ↓ 0.4% in the Prob. of choosing one of the top-three majors (1.5% of sample mean)
- ↑ 1.4% in the Prob. of choosing one of the bottom-three majors (12.0% of sample mean)

Outline

Long-term outcomes

	(1) Consumption	(2) Durables	(3) Non-durables	(4) Leisure
Panel A: Linear effect				
AA doses	-0.814***	-0.117	-0.535***	-0.159*
	(0.221)	(0.177)	(0.175)	(0.084)
Panel B: Non-linear effect				
1^{st} quintile	0.196***	0.096**	0.129***	0.020
	(0.049)	(0.038)	(0.034)	(0.019)
2^{nd} quintile	0.048	0.077*	0.053	0.006
	(0.048)	(0.044)	(0.042)	(0.018)
3^{rd} quintile	-0.037	0.035	0.031	0.011
	(0.060)	(0.0461)	(0.044)	(0.018)
4^{th} quintile	-0.144***	0.001	-0.078*	0.003
	(0.056)	(0.048)	(0.046)	(0.024)
5^{th} quintile	-0.283***	-0.076	-0.200***	-0.069***
	(0.063)	(0.048)	(0.049)	(0.024)
Controls	YES	YES	YES	YES
County-ethnicity-track FE	YES	YES	YES	YES
County-cohort-track FE	YES	YES	YES	YES
Obs.	9,015,498	9,015,498	9,015,498	9,015,498

Note: To address zero observations, we apply a log transformation with a calibrated extensive margin value, following the approach of Chen and Roth (2024). Specifically, we use $m(Y)$ to transform the level value Y, where $m(Y)$ is defined to equal $log(Y)$ for $Y > 0$ and $-K$ for $Y = 0$. In our baseline specification, we set $K = 0.1$.

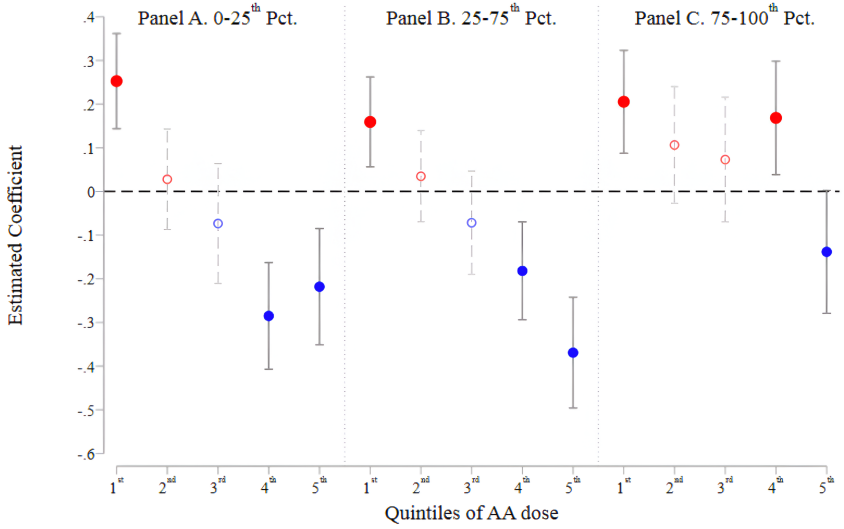
Effect of 10 points over mean of s.d. of CEE scores (=0.1):

- ↓ 7.8% average annual consumption
- Highly nonlinear effect: Positive effect for bottom quintile; Zero effect for middle quintiles; Negative effect for top quintiles

Robustness

- Alternative intensity measures;
- Alternative calibration values;
- Potential selection bias in card usage (Heckman two-step correction);
- Restricting the consumption sample to 2013-2018.

Non-linear AA dose effect across CEE score rank groups



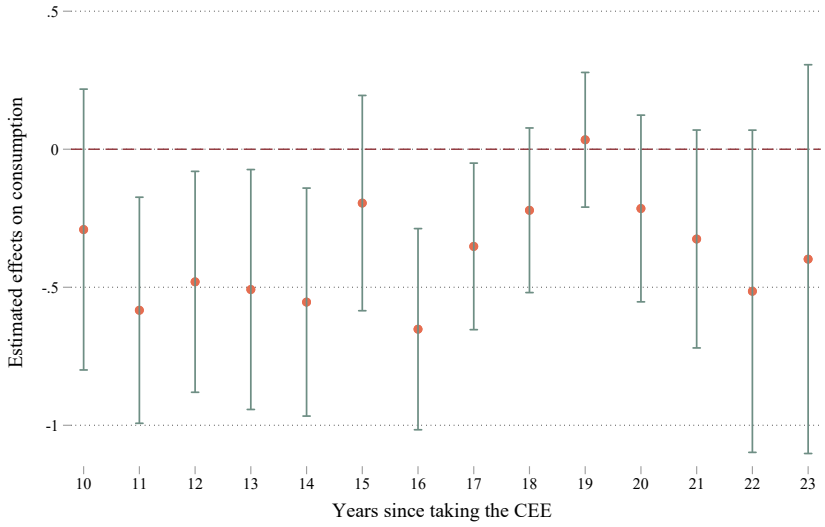
Notes: This figure shows the heterogeneous non-linear effects of AA doses on average annual consumption in later-life across within-province-cohort-track CEE percentile rank groups. Estimates are obtained by interacting five AA-dose quintiles with three CEE rank indicators (0-25$^{\text{th}}$, 25-75$^{\text{th}}$, and 75-100$^{\text{th}}$ percentiles).

Heterogeneity

	(1)	(2)	(3)
		Dependent variable: Average consumption	
Variable (X):	Male	Rural	HS Quality
Panel A: Linear effect			
AA doses × X	-1.559***	-1.233***	0.033***
	(0.124)	(0.129)	(0.005)
AA doses	0.137	0.079	-2.232***
	(0.234)	(0.239)	(0.292)
Panel B: Non-linear effect			
1^{st} quintile × X	-0.175***	-0.038	0.001
	(0.040)	(0.058)	(0.001)
2^{nd} quintile × X	-0.294***	-0.019	0.003*
	(0.039)	(0.052)	(0.002)
3^{rd} quintile × X	-0.146***	-0.225***	0.006***
	(0.043)	(0.049)	(0.002)
4^{th} quintile × X	-0.394***	-0.294***	0.011***
	(0.040)	(0.046)	(0.002)
5^{th} quintile × X	-0.341***	-0.287***	0.006***
	(0.044)	(0.043)	(0.002)
Controls	YES	YES	YES
County-ethnicity-track FE	YES	YES	YES
County-cohort-track FE	YES	YES	YES
Obs.	9,015,498	9,015,498	9,015,498

Notes: Columns (1)–(3) examine heterogeneity by gender (*Male*), household registration status (*Rural*), and high school quality (*High school average CEE ranking*), respectively. High school quality is measured by the average CEE ranking of students from that high school in the first year it appears in the sample.

Dynamics



Note: We estimate the effect of bonus intensity on total consumption for each year since enrollment, with each point representing a separate coefficient estimate.

Impact of AA Doses on financial investment

	(1) Financial investment	(2) Insurance	(3) Bonds & stocks
Panel A: Linear effect			
AA doses	-0.470***	-0.054	-0.478***
	(0.153)	(0.082)	(0.142)
Panel B: Non-linear effect			
1^{st} quintile	0.061	0.077***	0.009
	(0.039)	(0.021)	(0.034)
2^{nd} quintile	-0.013	-0.017	-0.014
	(0.032)	(0.019)	(0.030)
3^{rd} quintile	-0.069*	-0.017	-0.066*
	(0.040)	(0.021)	(0.035)
4^{th} quintile	-0.118***	-0.037*	-0.102***
	(0.042)	(0.022)	(0.038)
5^{th} quintile	-0.150***	-0.013	-0.155***
	(0.043)	(0.022)	(0.040)
Controls	YES	YES	YES
County-ethnicity-track FE	YES	YES	YES
County-cohort-track FE	YES	YES	YES
Obs.	9,015,498	9,015,498	9,015,498

Note: This table reports estimates of the effect of AA doses on later-life expenditure on financial investment. Columns (1)–(3) are average annual expenditure on all financial investments, insurance, and bonds & stocks, respectively.

Academic performance and subjective well-being

	(1) Below median rank	(2) Fail in courses	(3) Sat. with: GPA	(4) Sat. with: Social	(5) Sat. with: Relations	(6) Overall Sat.
Panel A: Linear effect						
AA doses	0.275**	0.225	-0.251*	-0.066	-0.150	-0.185*
	(0.139)	(0.145)	(0.139)	(0.117)	(0.101)	(0.110)
Panel B: Non-linear effect						
1st tercile	0.025	0.003	0.010	0.028	0.014	0.016
	(0.032)	(0.034)	(0.034)	(0.028)	(0.022)	(0.025)
2nd tercile	-0.012	-0.046	0.001	-0.018	-0.003	-0.011
	(0.035)	(0.042)	(0.036)	(0.033)	(0.026)	(0.028)
3rd tercile	0.074*	0.074*	-0.069*	-0.023	-0.047*	-0.051*
	(0.038)	(0.039)	(0.038)	(0.032)	(0.028)	(0.029)
Cell FE	YES	YES	YES	YES	YES	YES
Obs.	3,817	3,776	3,817	3,817	3,817	3,817

Note: The dependent variables in columns (1)–(2) measure academic outcomes: *Below median rank* is a dummy variable equal to 1 if the student's cumulative GPA rank is below the median within their class. *Fail in courses* indicates whether the student failed at least one course during college. Columns (3)–(6) present dummy variables for satisfaction with GPA, social activities, interpersonal relationships, and overall campus life, respectively, constructed from original 5-point scale (1 = "highly dissatisfied"; 5 = "highly satisfied"). Each dummy variable equals 1 if the original score is equal to or greater than 3, and 0 otherwise. To ensure comparability, we restrict the sample to cells that contain at least one individual with positive bonus points.

Conclusion

- We study how the long-term well-being of AA beneficiaries (disadvantaged ethnic minorities) is affected by the *dose* of AA they received.
- On average, higher dose reduces long-tem well-being:
  - Consumption ↓.
  - Durables, non-durables, finance investment ↓.
- The effect is highly non-linear, ranging from sizable gains from small doses to sizable losses from large doses.
- Suggestive evidence on college activities and interpersonal relationships.
- AA debate: It is a matter of the dose.

Thanks and comments welcome!

ltsysu1990@gmail.com