Reshuffling University Access

Evidence from Teacher-Based Grading during the COVID-19 Pandemic

Keitaro Ninomiya (Ph.D. Candidate)
Uni. of Illinois at Urbana-Champaign

ASSA

December 31, 2025

UNIVERSITY OF ILLINOIS URBANA-CHAMPAIGN

United States | College admissions

America tries to figure out a fairer way to select students

Sidelining standardised tests won't make college admissions fairer

Harvard Extends Test-Optional Admissions Policy for Four Years

The university joins many others that have eliminated the ACT and SAT requirements, adding fuel to the movement to get rid of standardized test scores.

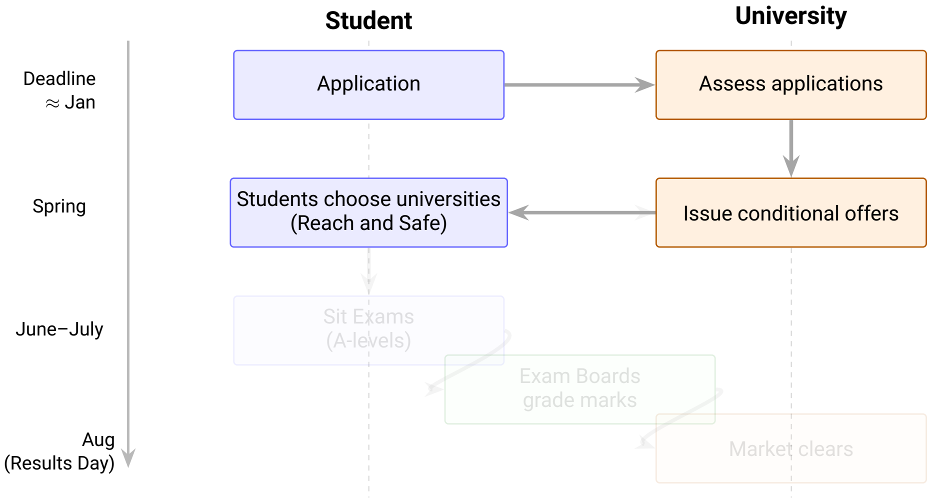
A-level results: is the government doing enough to tackle regional disparities?

How much are the inequality at universities due to the policy makers' choice of grading regimes?

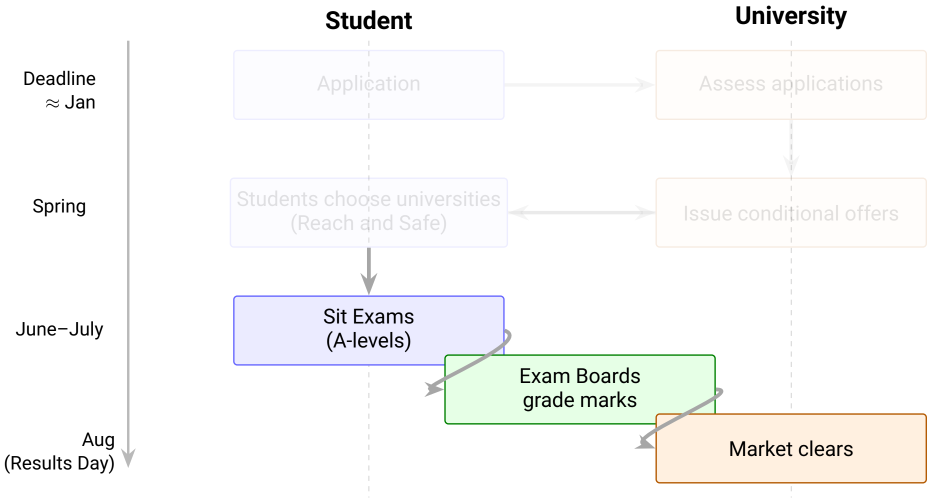
and how do schools' heterogeneous grading policies influence the inequality?

Britains’ replacement of A-levels with teacher assigned grades and the admissions surge

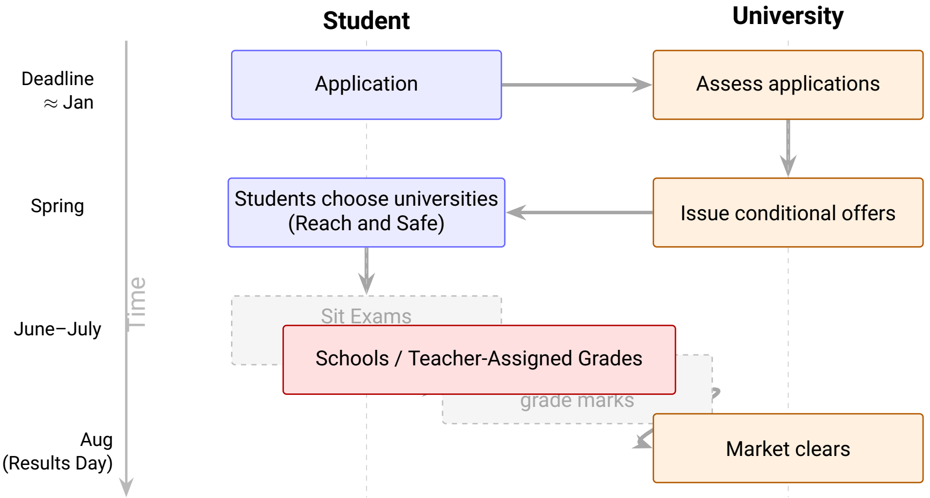
Britains’ replacement of A-levels with teacher assigned grades and the admissions surge



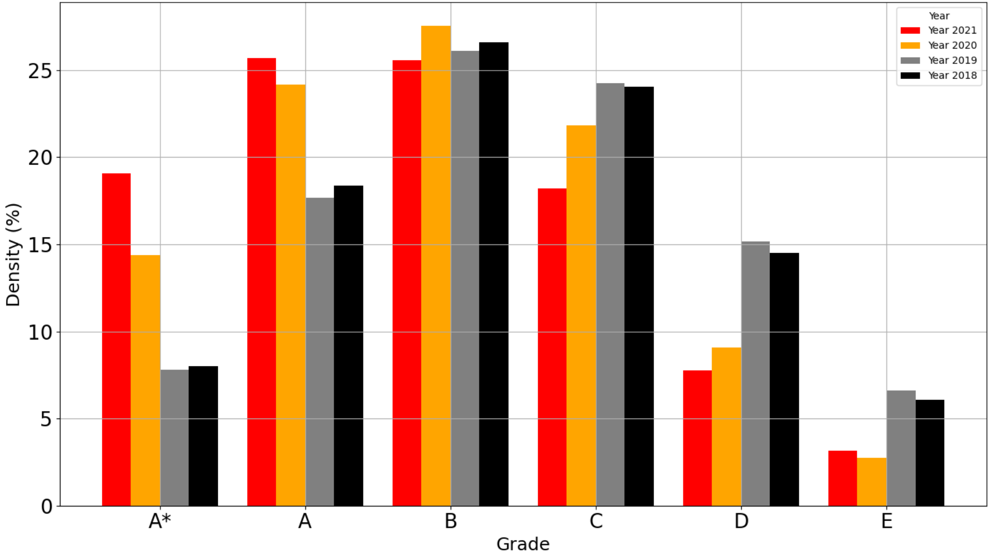
Britains' replacement of A-levels with teacher assigned grades and the admissions surge



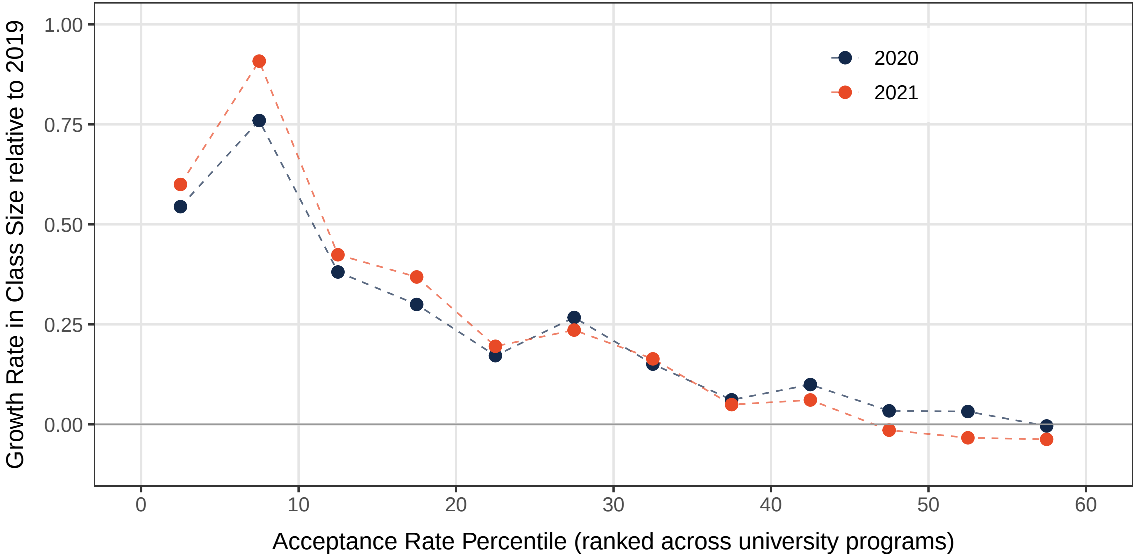
Britains' replacement of A-levels with teacher assigned grades and the admissions surge



Schools issued 200% more top grades.



Subsequent overcrowding at univ.



Research Questions: Grading Policies and Their Consequences

Research Question 1. How did grading policies differ across schools?

Across-school channels

- School **quality** and **type**
- **Adjustments** from 2020 to 2021

Within-school channels

- Student **demographics**
- **Subject** area (STEM vs. non-STEM)

Research Question 2. How did the grading policies change students comp. at unis?

Method: Cross-cohort comparison (COVID vs. non-COVID).
Identification (grading policy): The A-level grade distribution is stationary in the pre-COVID period.
Identification (composition): University–course variation in exposure to applicants' schools' grading policies.

Preview of Results: Heterogeneous Grading Policies and Changes in University Composition

Grade inflation closed attainment gaps across schools.

- Bottom decile schools were 13 p.p more likely to assign A/A* than top decile school.
- Elite-secondary schools had the largest conversion of non-top grade into top grades

Private schools outpaced public via revisions.

Inflation skewed towards non-STEM, female, white, and high-income parent students.

- However, grade inflation levels followed prior academic attainment levels.

Universities admitted lower-achieving but higher-SES students

- A 10% increase in intake reduced average grades by **0.5 s.d.** in Mathematics and **0.4 s.d.** in English.
- The same expansion increased the parental income score by **0.2 s.d.**.

Empirical Model & Identification

1 Motivation

2 Empirical Model & Identification

3 Data

4 Results

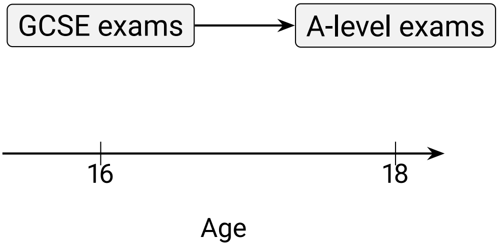
5 Conclusion

Mechanics of how the UK monitors grade distributions

Threat to identification: Confounding time trend and policy effect.

$$E[Y_{i,j,k,t'} | X_{t' \geq 2020}, Standardized] = E[Y_{i,j,k,t} | X_{t<2020}, Standardized]$$

Figure: Timeline for 16–18 education



Comparable Outcomes rule

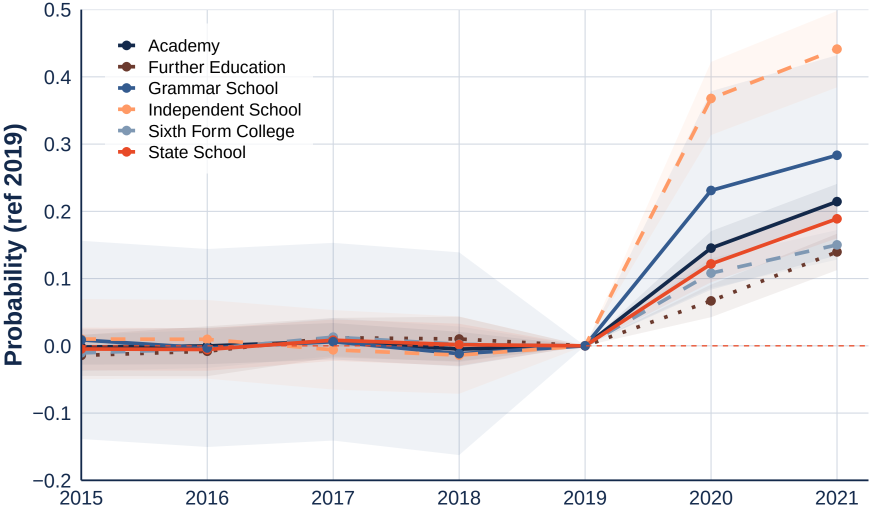
For each subject s, central graders impose:

$$F_{s,t}(Y \mid \text{GCSE}) = F_{s,t-1}(Y \mid \text{GCSE})$$

The **conditional distribution of A-level grades given GCSE scores** is constant across years.

Grade distribution in the UK is stationary

Figure: Conditional probability of achieving an A/A* by school type



Estimating grading policy as responses to grading regime change

$$Y_{i,j,k,t} = \mathbf{1}\{\ell_{i,j,k,t} \geq 0\}, \quad \Pr(Y_{i,j,k,t} = 1 \mid \cdot) = \frac{e^{\ell_{i,j,k,t}}}{1 + e^{\ell_{i,j,k,t}}}$$

$$\ell_{i,j,k,t} = \alpha_j + X_i \eta + \gamma_k + \sum_{\tau \in \{2020,2021\}} D_\tau \times \Big(\underbrace{\Delta \alpha_{j,\tau}}_{\text{Betw.-School Effect}} + X_i \, \Delta \eta_\tau + \Delta \gamma_{k,\tau} \Big)$$

- Y: top-grade indicator; X_i: controls (GCSE, demographics).
- α_j, γ_k: school/subject FE; D_τ: year indicator.
- $\Delta\cdot$: Grading policy.
- Estimation: Maximum likelihood estimation with EB shrinkage.

Data

Universities and Colleges Admissions Service (UCAS) administrative data

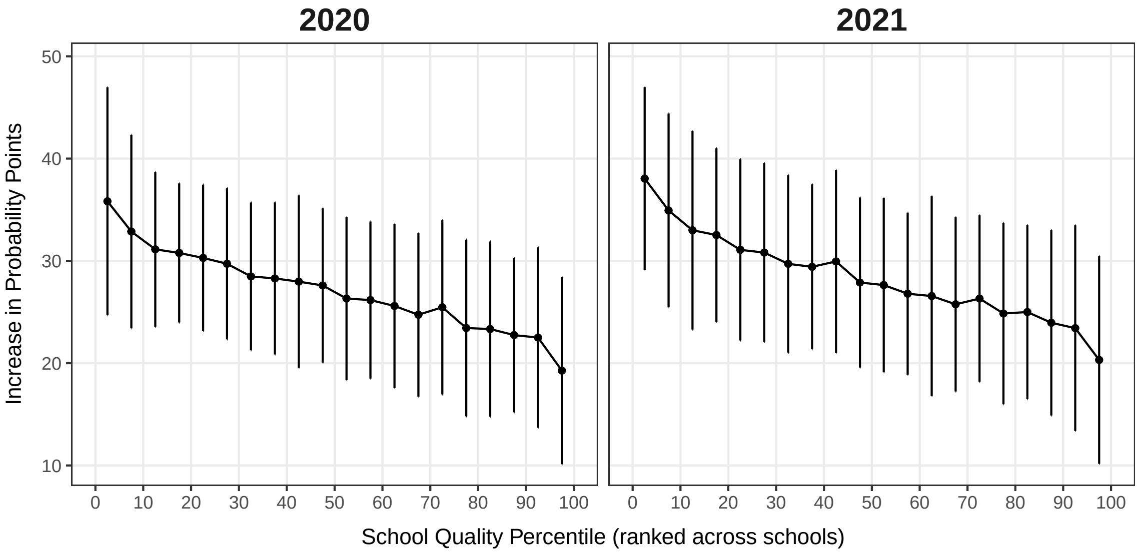
Coverage: (2015-2021).

- Grade outcome for each A-level qualification.
- GCSE qualification scores.
- School of attendance.
- Demographics (e.g. race, gender).
- Income measures.
- Preference rank among university programs.
- Final placement destination.

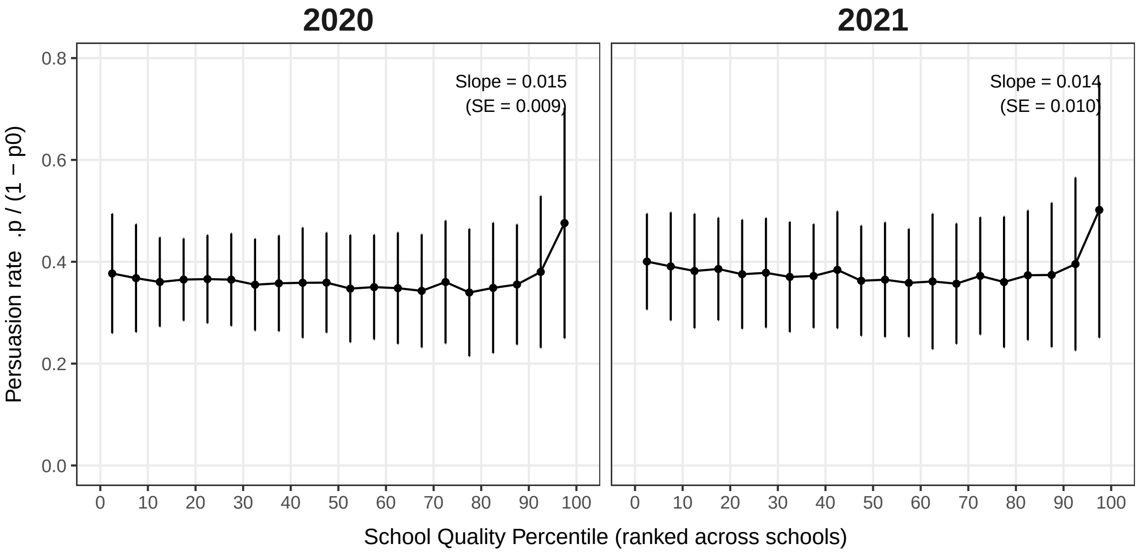
Outline

1. Motivation
2. Empirical Model & Identification
3. Data
4. Results
   - Grading Policy
   - Placement Improvements
   - Composition Change
5. Conclusion

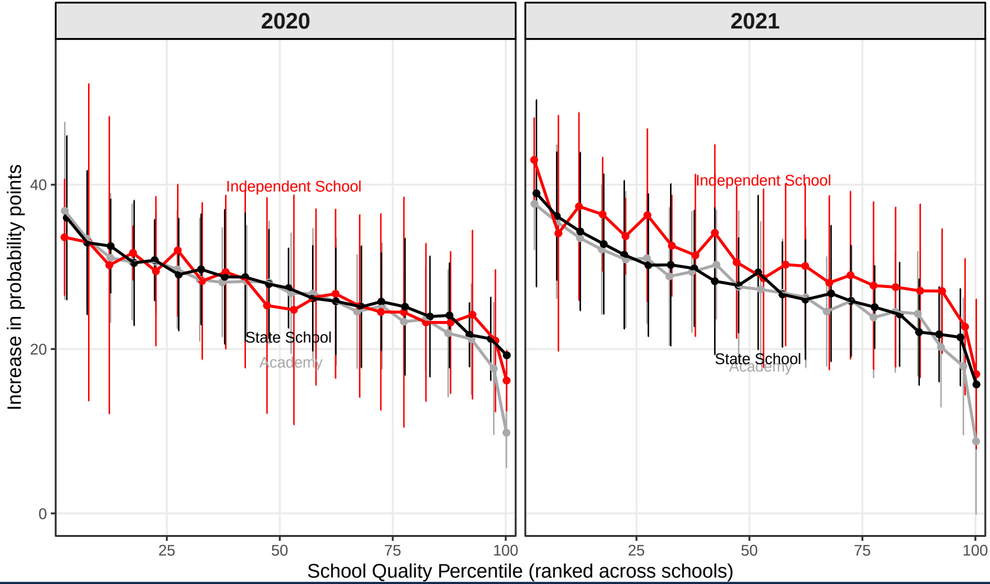
Higher inflation concentrated at lower performing schools



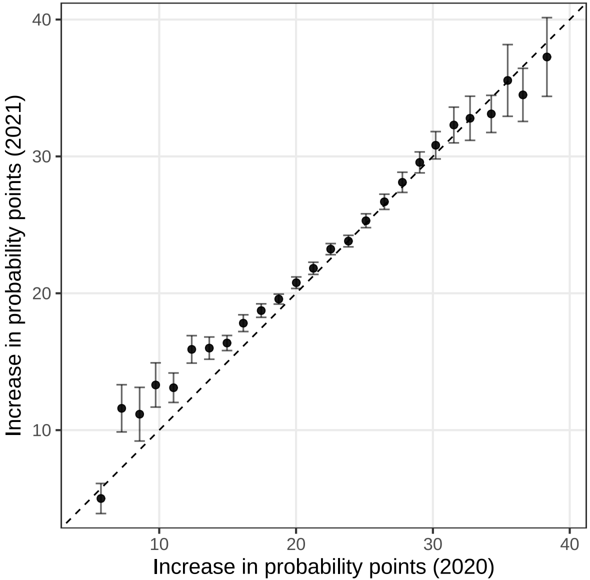
Elite schools converted ill-performing students the most



Private (Independent) schools vs. publicly funded schools



Strictly grading schools loosened standards in second year



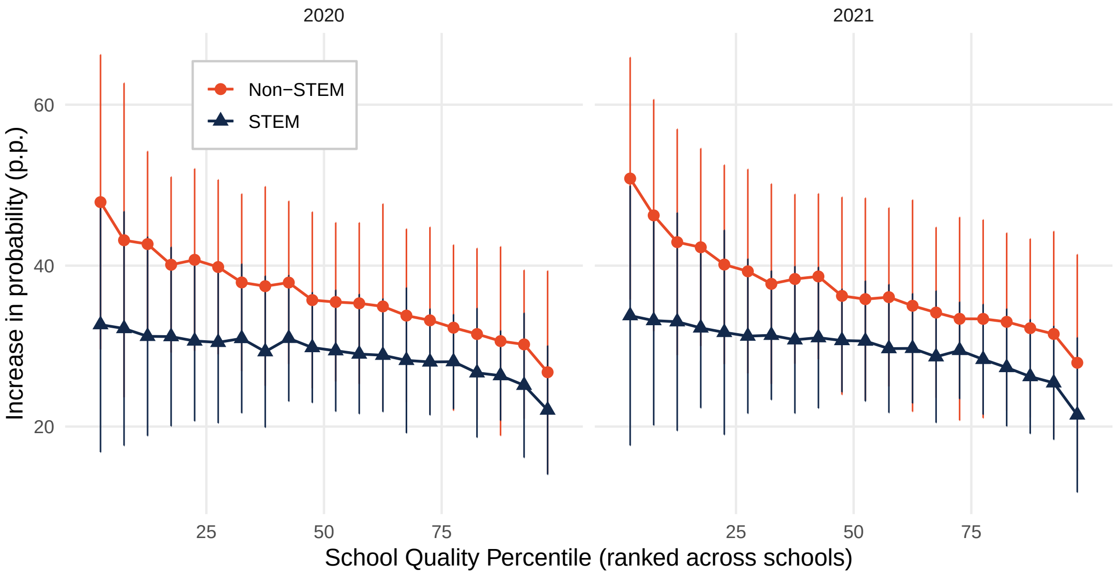
Estimating how teachers assigned grades to different students

$$Y_{i,j,k,t} = \mathbf{1}\{\ell_{i,j,k,t} \geq 0\}, \quad \Pr(Y_{i,j,k,t} = 1 \mid \cdot) = \frac{e^{\ell_{i,j,k,t}}}{1 + e^{\ell_{i,j,k,t}}}$$

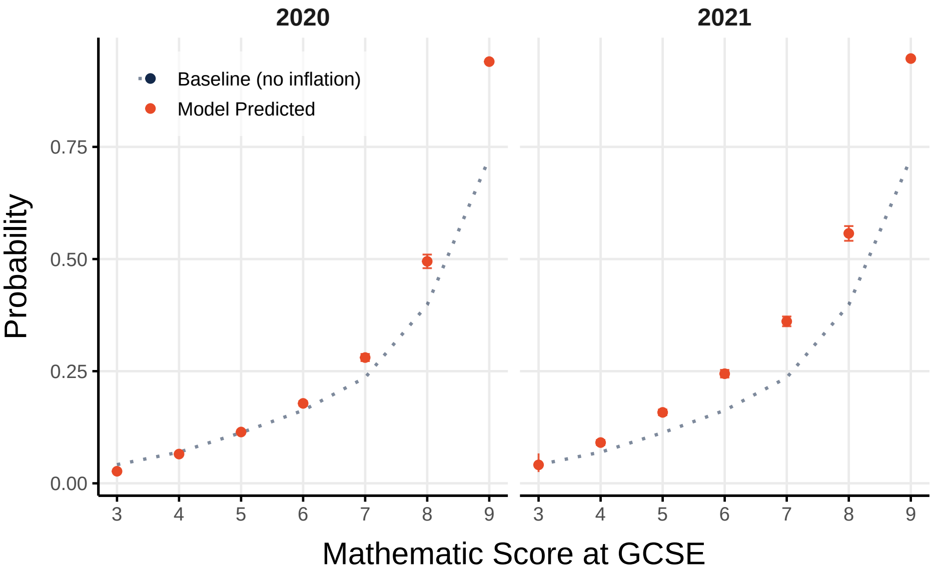
$$\ell_{i,j,k,t} = \alpha_j + X_i\eta + \gamma_k + \sum_{\tau \in \{2020,2021\}} D_\tau \Big(\underbrace{S_k \times \Delta\alpha_{j,\tau}}_{\text{Between-school by subject type}} + \underbrace{X_i \times \Delta\eta_{j,\tau}}_{\text{Within-school}} + \Delta\gamma_{k,\tau} \Big)$$

- S_k: Indicator for non-STEM subjects.
- X_i: Student attributes (e.g. GCSE, demographics)

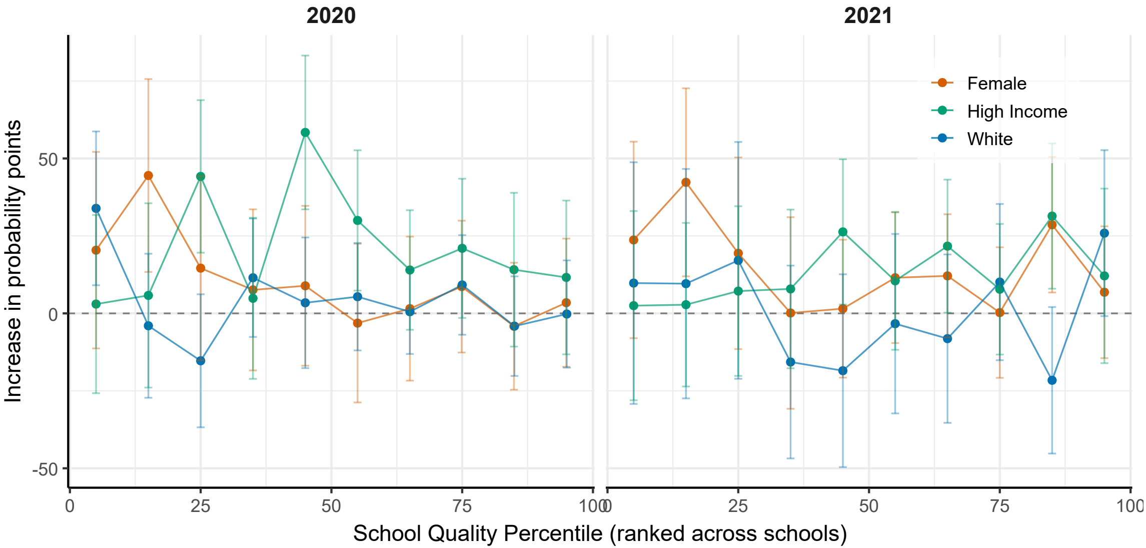
Grading varied more in non-STEM subjects.



Academically well-performing students received higher grades



Female, White, and High Income students receives higher grades



Outline

1 Motivation

2 Empirical Model & Identification

3 Data

4 Results
- Grading Policy
- Placement Improvements
- Composition Change

5 Conclusion

Impact of grade inflation on univ. placements

(How) Did the grading policies improve students' enrollment?

Which schools **translated the grading policy** into placements?

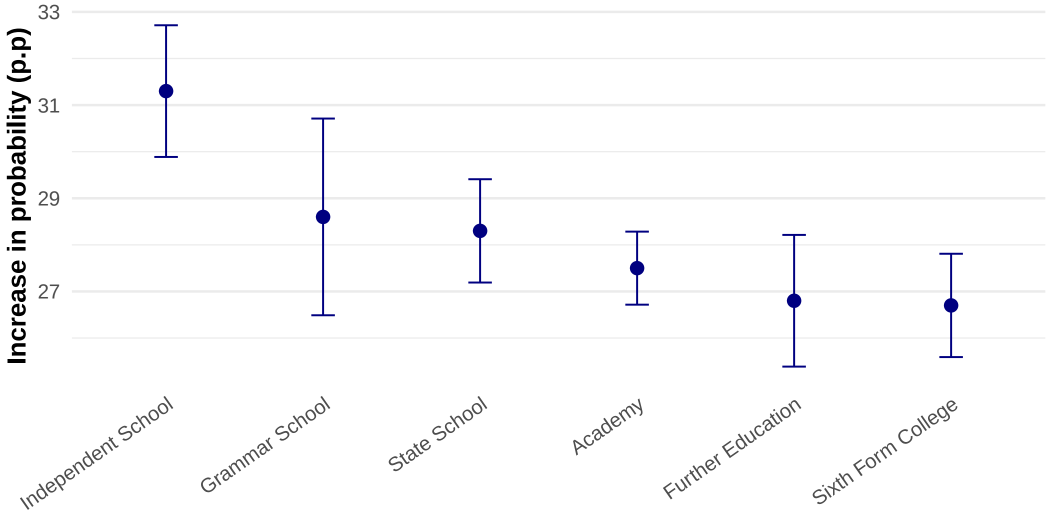
$$Placed_{i,j,t} = \beta \left(D_t \times Inflation_{i,j,t} \right) + \iota_j + \varsigma_{i,j,t}$$

- $Placed_{i,j,t}$: Indicator for student *i* placing at their top choice in university *j* at time *t*.
- ι_j: School fixed effect.
- $Inflation_{i,j,t}$: Degree of grade inflation,

$$Inflation_{i,j,t} \equiv Grade_{i,j,t} - Grade_{i,j,t}^{LASSO}$$

- $Grade_{i,j,t}$: Aggregate A-level score.

Private schools converted inflation into placements the most



Outline

1. Motivation
2. Empirical Model & Identification
3. Data
4. Results
   - Grading Policy
   - Placement Improvements
   - Composition Change
5. Conclusion

Impact of grade inflation on univ. composition

How did grade policies change student composition at universities?

- Did universities enroll students from different parental occupation group?

$$Y_{u,t} = \beta \, \log(N_{u,t}) + \iota_j + \lambda_u + \varsigma_{u,t}$$

$$\log(N_{u,t}) = \delta \cdot \frac{1}{M_{u,t}} \sum_{i \in M_{u,t}} \mathit{Inflation}_{i,u,t}$$

- $Y_{u,t}$: Outcome (e.g., avg. academic score, avg. parental income).
- $N_{u,t}$: Number of accepted students at university u in year t.
- $M_{u,t}$: Number of applicants to university u in year t.
- λ_u: Subject-group fixed effect. ι_j: University fixed effect.

	Academic preparedness Math score (z)		Parental affluence Income score (z)	
	(1)	(2)	(1)	(2)
log N	-0.0351^{***}	-0.0530^{***}	0.0136^{***}	0.0182^{***}
	(0.0070)	(0.0130)	(0.0027)	(0.0047)
Highly selective $\times \log N$	-0.0282^{***}	-0.0879^{*}	0.0108^{***}	0.0312^{*}
	(0.0063)	(0.0357)	(0.0024)	(0.0127)
Mid-selective $\times \log N$	-0.0337^{***}	-0.0470^{**}	0.0131^{***}	0.0161^{**}
	(0.0071)	(0.0159)	(0.0028)	(0.0058)
Less selective $\times \log N$	-0.0389^{***}	-0.0352^{**}	0.0156^{***}	0.0117^{*}
	(0.0081)	(0.0132)	(0.0032)	(0.0049)
University FE		✓		✓
Observations	31,478	31,478	31,592	31,592

Conclusion and Policy Discussion

Who benefits from non-standardized assessments in university admissions?

Proponents	Opponents
"Holistic assessments broaden access."	*"Holistic assessments amplify inequality."*
• Reduces reliance on high-stakes exams • Allows teachers to contextualize performance • Increases diversity in admitted cohorts	• School resources affect credibility of grades • Stronger signalling power for elite schools • Reinforces pre-existing advantage

What the pandemic reveals

- Large upward shift in average grades
- Heterogeneous adjustments across schools and students (e.g. gender, race)
- Grading integrity revised downward in the second year.
- More female and economically affluent students entered selective universities.

Future projects

1. **Effects of grade inflation on academic and labor-market outcomes**
   Did students who gained access to more selective universities perform differently in university courses and subsequent labor markets than conventional students?
   Data: Administrative university transcripts and linked graduate labor-market records.
2. **Behavioral model of grade assignment**
   Teachers preserve within-cohort rank ordering, while grade levels respond to racial and gender biases.
   Data: Administrative records with student ranks within course cohorts.
3. **Student–university matching under strategic responses**
   Estimating equilibrium student allocation when both students and universities optimally respond to grade inflation.

Thank You!

Empirical Challenge 2

- Confounding learning disruptions from COVID with grading effects.

$$E[D_\tau \times (\epsilon_{2020,2021} - \bar{\epsilon})] = 0$$

Solution: I control for the learning disruption during the pandemic by comparing students who took other standardized exams that weren't canceled during the pandemic.

Persistence of grade inflation

$$Y_{i,j,k,t} = \mathbf{1}\{\ell_{i,j,k,t} \geq 0\}, \quad \Pr(Y_{i,j,k,t} = 1 \mid \cdot) = \Lambda(\ell_{i,j,k,t})$$

$$\ell_{i,j,k,t} = \alpha_j + X_i\eta + \gamma_k \qquad 20$$

$$+ \sum_{\tau \in \{2020,2021\}} D_\tau \times \Big(\underbrace{\Delta\alpha_{j,\tau}}_{\text{Betw.-School Effect}} + \underbrace{X_i\,\Delta\eta_\tau}_{\text{Within-School Effect}} + \Delta\gamma_{k,\tau} \Big) \qquad 20$$

I estimate the descriptive estimate of grade inflation within schools.

$$\Delta\alpha_{j,2021} = \rho\,\Delta\alpha_{j,2020} + \varepsilon_j.$$

... and across subject groups.

$$\Delta\alpha_{j,g,2021} = \rho_{\mathrm{g}}\,\Delta\alpha_{j,g,2020} + \rho_{-\mathrm{g}}\,\bar{\Delta\alpha}_{j,-g,2020} + \varepsilon_{j,g},$$

Grading patterns within classrooms

$$Y_{i,j,k,t} = \mathbf{1}\{\ell_{i,j,k,t} \geq 0\}, \quad \Pr(Y_{i,j,k,t} = 1 \mid \cdot) = \Lambda\big(\ell_{i,j,k,t}\big)$$

$$\ell_{i,j,k,t} = \alpha_j + X_i \eta + \gamma_{j,k}$$ 20

$$+ \sum_{\tau \in \{2020,2021\}} D_\tau \times \Big(\underbrace{\Delta\alpha_{j,\tau}}_{\text{Betw.-School Effect}} + \underbrace{X_i \, \Delta\eta_{j,\tau}}_{\text{Within-School Effect}} + \Delta\gamma_{j,k,\tau} \Big)$$ 20

Measuring Impact of grade inflation on univ. placements

- Instrumental variable approach at the school level.

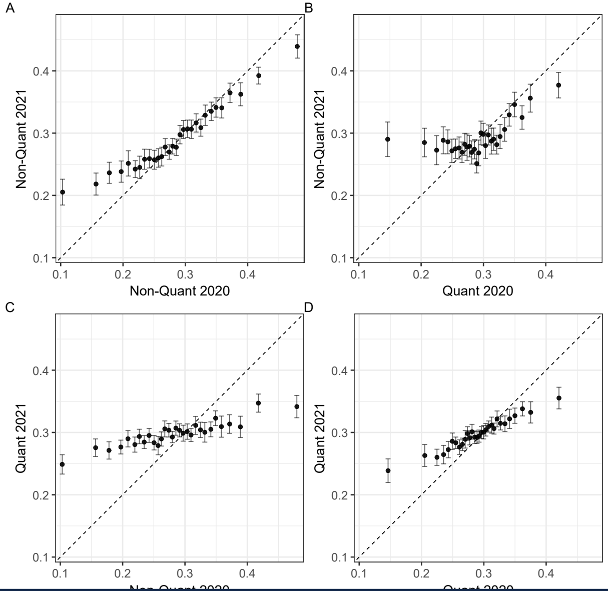
$$\text{Application}_{i,j,t} = \delta Y_{i,j,k,t} + \varsigma_{i,j,t}$$

$$Y_{i,j,k,t} = \mathbf{1}\{\ell_{i,j,k,t} \geq 0\}, \quad \Pr(Y_{i,j,k,t} = 1 \mid \cdot) = \Lambda\big(\ell_{i,j,k,t}\big)$$

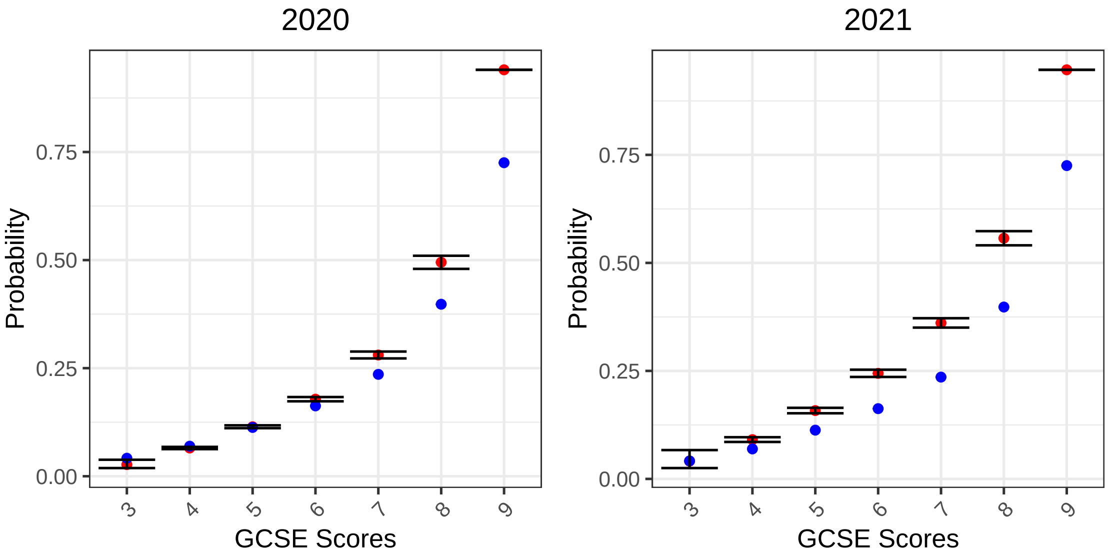
$$\ell_{i,j,k,t} = \alpha_j + X_i \eta + \gamma_k$$

$$+ \sum_{\tau \in \{2020,2021\}} D_\tau \times \Big(\underbrace{\Delta\alpha_{j,\tau}}_{\text{Betw.-School Effect}} + \underbrace{X_i \, \Delta\eta_\tau}_{\text{Within-School Effect}} + \Delta\gamma_{k,\tau} \Big)$$

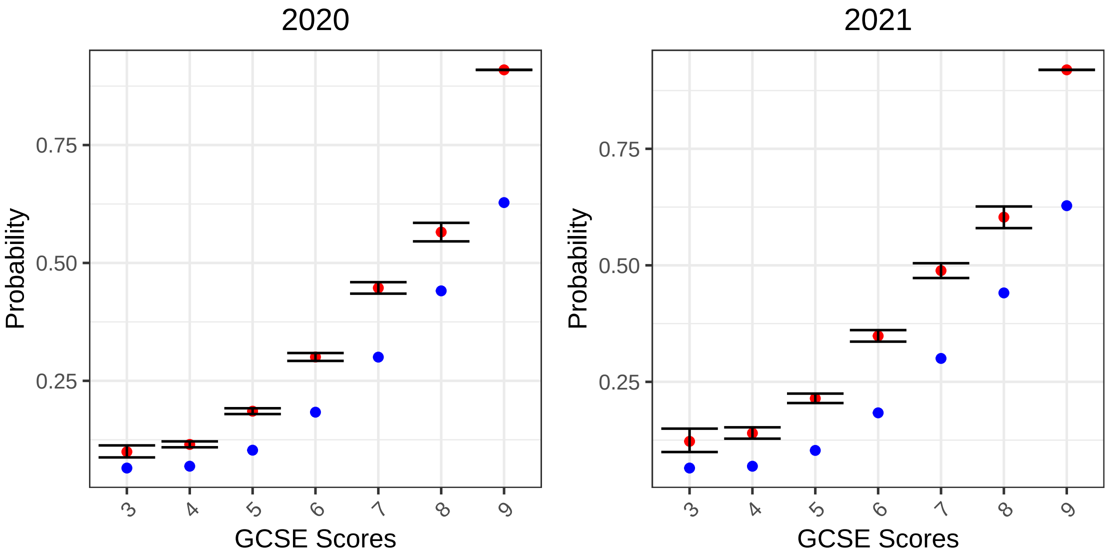
Persistence and adjustment in grading policies



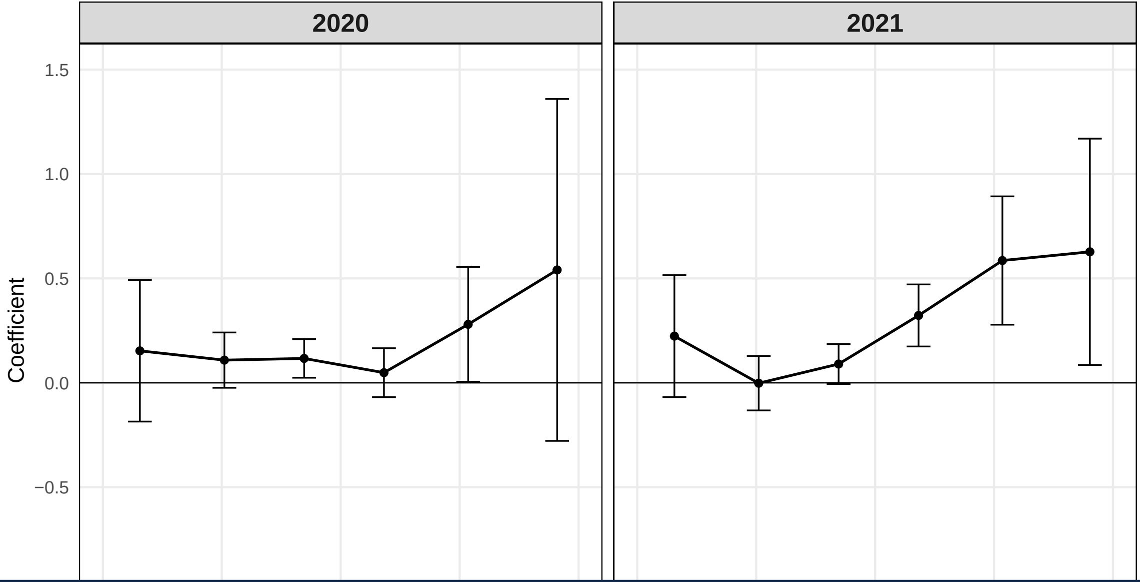
GCSE Math Scores



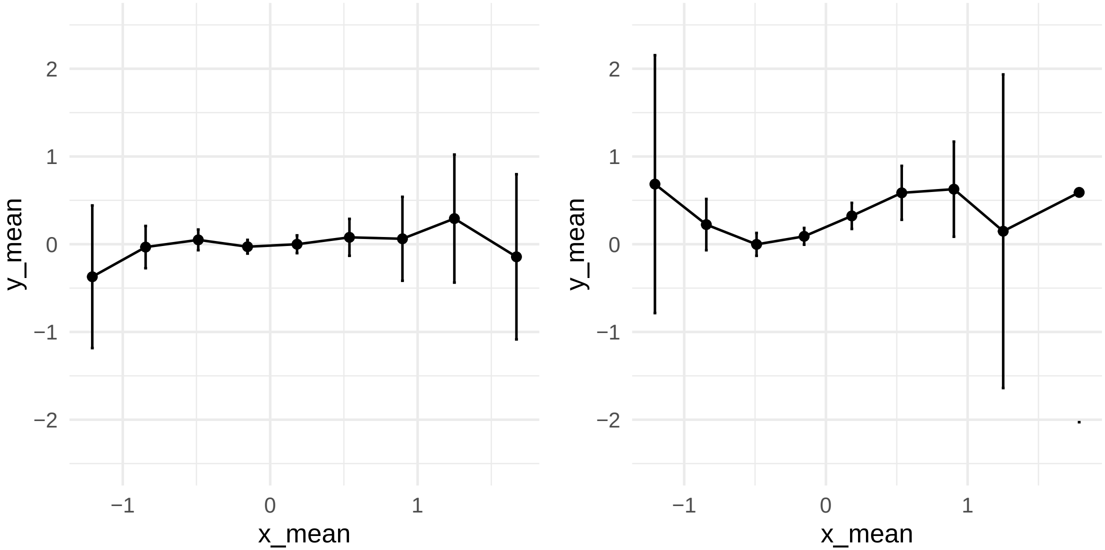
GCSE English Scores



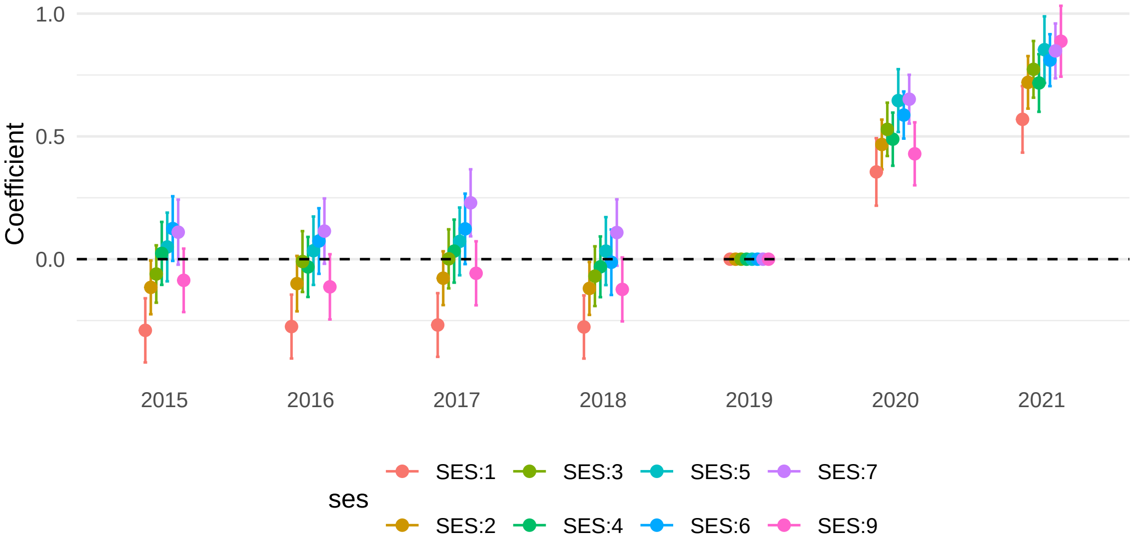
Females students vs. Male students



Ethnic



Conditional on parental occupation class



Changes in university course composition

$$\Delta N_{u,g} = N^t_{u,g} - N^{2019}_{u,g}, \quad N^t_{u,g} = M^t_{u,g} \cdot E_{F_t}[p_t(x)]$$

$M^s_{u,g}$: number of applicants from group g to uni u in year s
F_s: distribution of applicant characteristics in year s
$p_s(x)$: acceptance probability given x under year-s mechanics
$\varphi_t(x)$: Degree of grade inflation.
c: Acceptance threshold.

$$\Delta N_{u,g} = \underbrace{(M^t_{u,g} - M^0_{u,g})\, E_{F_0}[p_0(x)]}_{\text{Application Volume}} + \underbrace{M^t_{u,g}\left(E_{F_t}[p_0(x)] - E_{F_0}[p_0(x)]\right)}_{\text{Composition}} + \underbrace{M^t_{u,g}\left(E_{F_t}[p_t(x)] - E_{F_t}[p_0(x)]\right)}_{\text{Inflation + Threshold}}.$$

$$E_{F_t}[p_t(x)] - E_{F_t}[p_0(x)] = \underbrace{E_{F_t}\big[p(\varphi_t(x), c_0) - p(0, c_0)\big]}_{\text{Inflation effect}} \tag{1}$$

$$+ \underbrace{E_{F_t}\big[p(\varphi_t(x), c_t) - p(\varphi_t(x), c_0)\big]}_{\text{Threshold reaction}}. \tag{2}$$

Estimating p_{if}^{t}

- *Baseline index:* $\eta_i^0 \equiv \mu_i + \phi^0_{s(i),g}$ with $\phi^0 = 0$ by normalization.
- *Inflation-only counterfactual:*

$$\widehat{p}_{if}^{\,\phi\text{-only}} = \widehat{F}\big(\eta_i^0 + \Delta\phi_{s(i),g}\big).$$

- *Threshold remainder (Oaxaca):*

$$\widehat{\Delta p}_{if}^{\,\text{thr}} = \widehat{p}_{if}^{\,1} - \widehat{p}_{if}^{\,\phi\text{-only}}.$$

Firm placement with changing applications

$z_i^t \sim F$; offer at f iff $z_i^t \geq b_{if}^t$, $\quad b_{if}^t = c_f^t - (\mu_i + \phi_{s(i),g}^t)$.

$$p_{if}^t \equiv \Pr(\text{offer at } f \mid i, t) = 1 - F(b_{if}^t).$$

Let $a_{if}^t \equiv \Pr(f(i) = f \mid X_i, t)$ be the prob. student i makes f their firm choice. Total firm inflow for group g:

$$N_{fg}^t = \sum_{i \in \mathcal{R}_g} a_{if}^t \, p_{if}^t, \qquad \Delta N_{fg} = \sum_{i \in \mathcal{R}_g} (a_{if}^1 p_{if}^1 - a_{if}^0 p_{if}^0).$$

$$\Delta N_{fg} = \underbrace{\sum_i (a_{if}^1 - a_{if}^0)\, p_{if}^{(0)}}_{\text{(A) Application mix}} + \underbrace{\sum_i a_{if}^1 \Big[p_{if}^{(\phi^1, c_f^0)} - p_{if}^{(\phi^0, c_f^0)}\Big]}_{\text{(B) Inflation at school}} + \underbrace{\sum_i a_{if}^1 \Big[p_{if}^{(\phi^1, c_f^1)} - p_{if}^{(\phi^1, c_f^0)}\Big]}_{\text{(C) Threshold/tightness}},$$

where $p_{if}^{(0)} \equiv p_{if}^{(\phi^0, c_f^0)}$.

Estimation recipe (no direct c_f^t):

- *Applications:* Estimate $\widehat{a}_{if}^t$ from observed firm choices (e.g., MNL or flexible ML on X_i, school $s(i)$, group g, year t).
- *Acceptance index:* Fit $\widehat{F}$ with baseline index $\eta_i^0 = \mu_i + \phi_{s(i),g}^0$ (normalize $\phi^0 = 0$).
- *Inflation-only counterfactual:*

$$\widehat{p}_{if}^{\,\phi\text{-only}} = \widehat{F}(\eta_i^0 + \Delta\phi_{s(i),g}).$$

- *Assemble the three pieces:*

$$\widehat{(\mathrm{A})} = \sum (\widehat{a}_{if}^1 - \widehat{a}_{if}^0)\, \widehat{p}_{if}^{(0)} \qquad \widehat{(\mathrm{B})} = \sum \widehat{a}_{if}^1 (\widehat{p}_{if}^{\,\phi\text{-only}} - \widehat{p}_{if}^{(0)})$$

Conclusion

- I provide evidence of differential grading policies across schools in the UK during the COVID 19 pandemic.
  - Lower quality schools inflated grade more than high quality schools.
  - Teachers/schools inflated students from their top students wrt. better academic/socio-economic backgrounds.
  - The grade inflation mainly helped students from low quality schools to place into less selective universities.

- Test-optional admissions could help disadvantaged students to obtain **the signals** for moving into selective universities.

  However, such students may not use the grades to enroll themselves into selective universities (Hoxby and Avery [2012]).

  Moreover, test optional policies may backfire by reallocating access to selective universities from academically capable students to students from schools with strong incentive to improve their placement records.

Applies to England

Contents

Grade boundaries change from year to year

The standard of work needed to get each grade remains comparable year on year

Grade boundaries are decided after students take exams and when marking is nearly complete

The number of students achieving each grade can differ between exam boards offering the same qualification

Grade boundaries cannot be adjusted for certain groups of students

The National Reference Test

Grade boundaries change from year to year

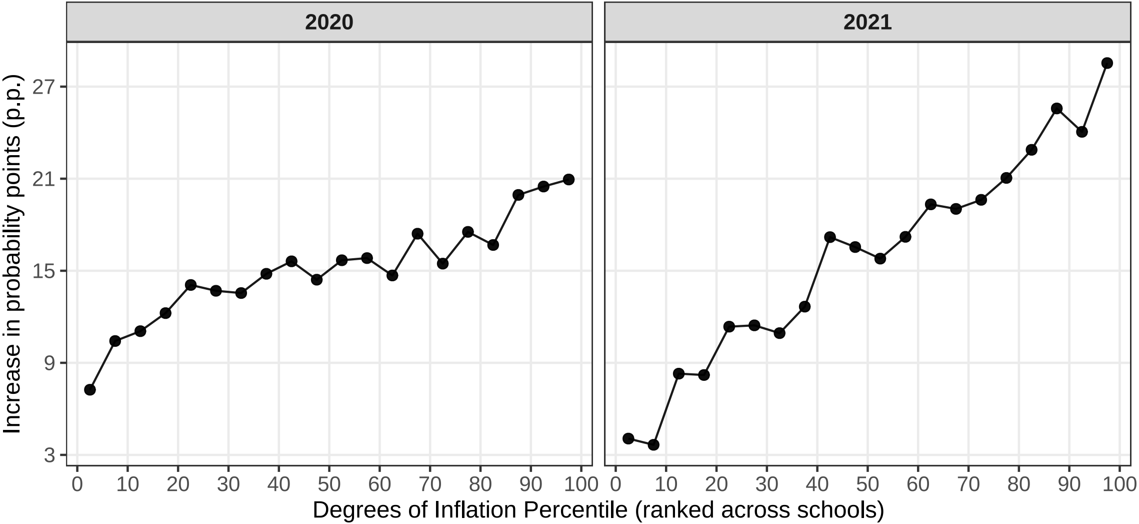
If an exam is easier than in previous years, the grade boundaries for that paper will be higher. If it is harder, the grade boundaries will be lower.

The difficulty of exam questions varies year to year, even though exam boards try to keep the level of demand consistent. That's because it is impossible to determine how difficult students will find a paper until it is taken.

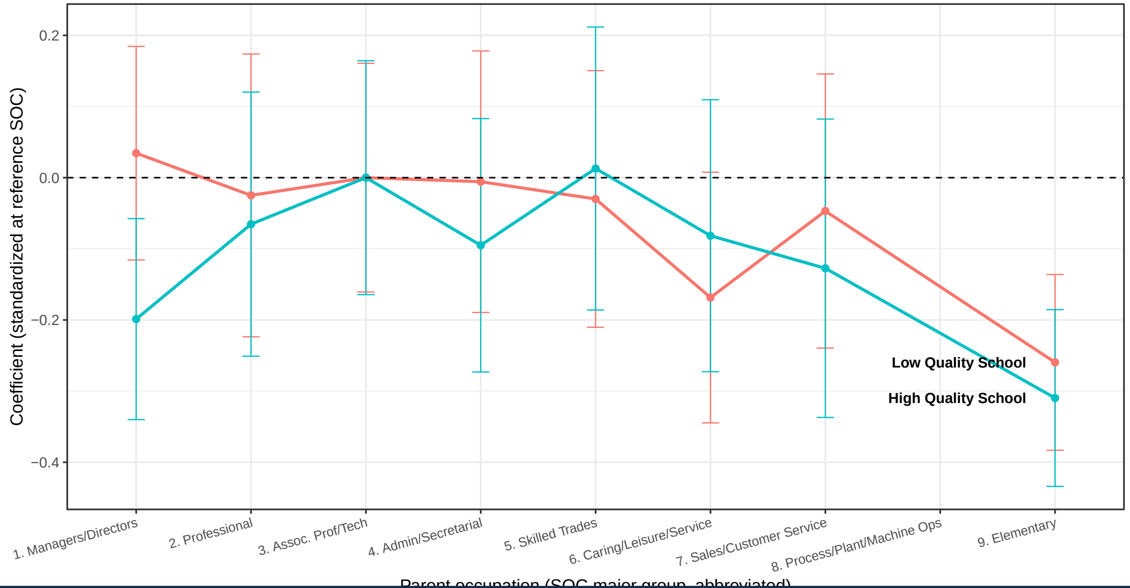
This is why new grade boundaries are set each year – to reflect the difficulty of that particular paper, and to ensure that it is no easier or harder to get a grade in any given year.

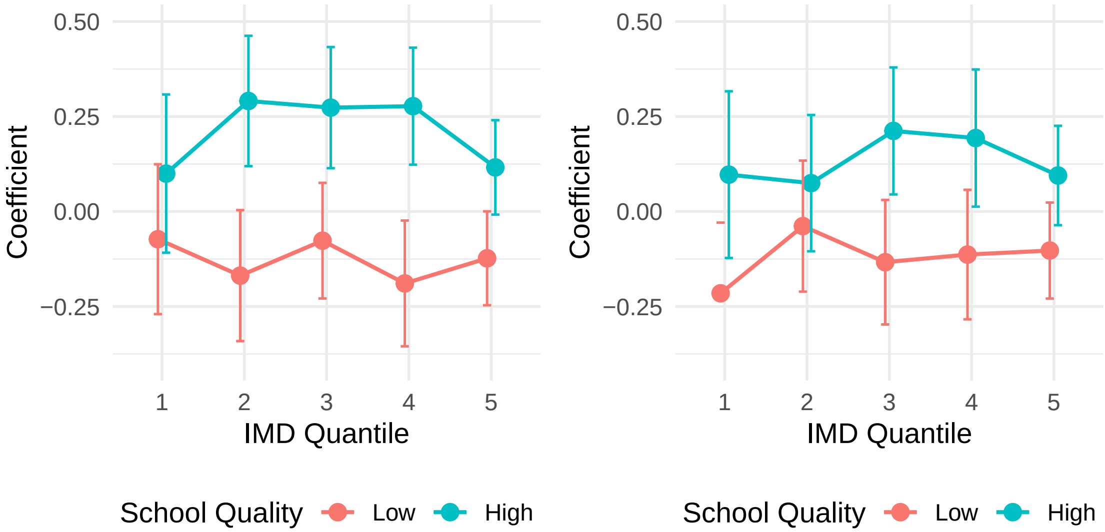
The standard of work needed to get each grade remains comparable year on year

Application success and Grade inflation

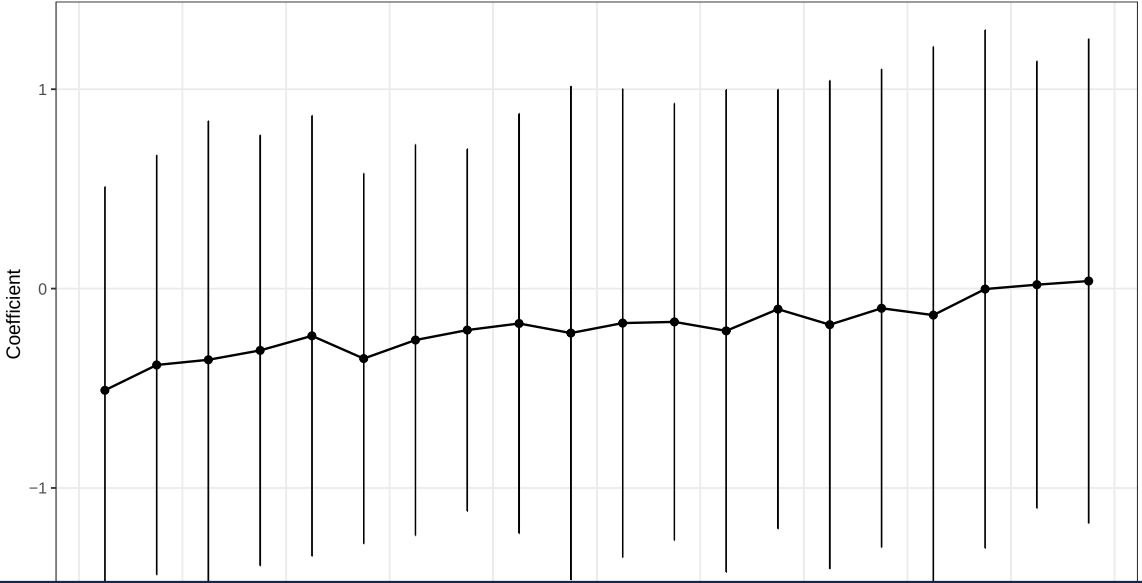


Grading by parental income class in 2020

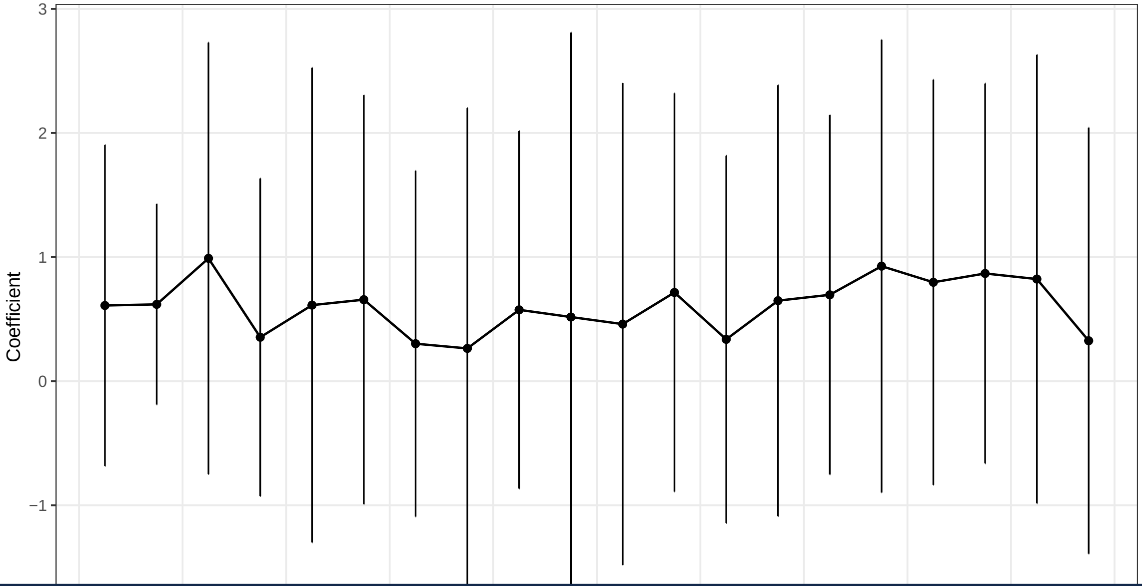



0.50
0.25
0.00
−0.25
Coefficient
1
2
3
4
5
IMD Quantile
School Quality
Low
High
0.50
0.25
0.00
−0.25
Coefficient
1
2
3
4
5
IMD Quantile
School Quality
Low
High

Improvements in application success rate in 2020



Improvements into Selective Univs in 2020



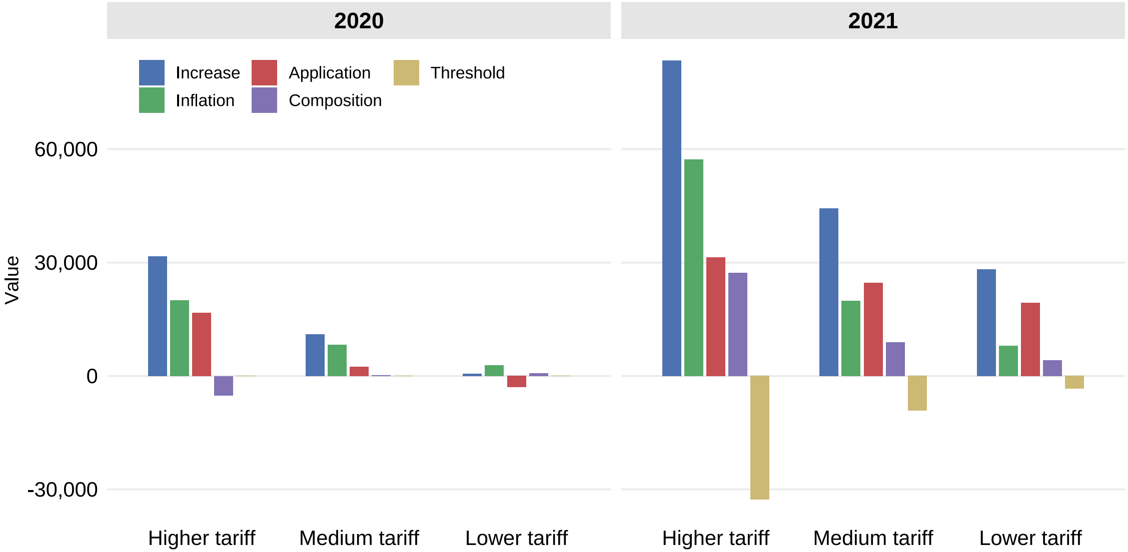
Changes in university course composition

$$\Delta N_{u,g} = N^t_{u,g} - N^{2019}_{u,g}, \quad N^t_{u,g} = M^t_{u,g} \cdot E_{F_t}[p_t(x)]$$

$M^s_{u,g}$: number of applicants from group g to uni u in year s
F_s: distribution of applicant characteristics in year s
$p_s(x)$: acceptance probability given x under year-s mechanics
$\varphi_t(x)$: Degree of grade inflation.
c: Acceptance threshold.

$$\Delta N_{u,g} = \underbrace{(M^t_{u,g} - M^0_{u,g})\, E_{F_0}[p_0(x)]}_{\text{Application Volume}} + \underbrace{M^t_{u,g}\left(E_{F_t}[p_0(x)] - E_{F_0}[p_0(x)]\right)}_{\text{Composition}}$$
$$+ \underbrace{M^t_{u,g}\left(E_{F_t}[p_t(x)] - E_{F_t}[p_0(x)]\right)}_{\text{Inflation + Threshold}}.$$

$$E_{F_t}[p_t(x)] - E_{F_t}[p_0(x)] = \underbrace{E_{F_t}\left[p(\varphi_t(x), c_0) - p(0, c_0)\right]}_{\text{Inflation effect}} \tag{3}$$
$$+ \underbrace{E_{F_t}\left[p(\varphi_t(x), c_t) - p(\varphi_t(x), c_0)\right]}_{\text{Threshold reaction}}. \tag{4}$$


2020
2021
Increase
Application
Threshold
Inflation
Composition
60,000
30,000
0
-30,000
Value
Higher tariff
Medium tariff
Lower tariff
Higher tariff
Medium tariff
Lower tariff

Application success and Grade inflation

	(1)	(2)
Inflation	0.005***	0.008***
	(0.00)	(0.00)
Year	2020	2021
School FE	✓	✓
Control	✓	✓
Observations	1,316,615	1,258,913

Notes: Entries are coefficients; standard errors in parentheses. Stars: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. "Tariff (observed)" is the running variable. School type interactions in col. (2). SEs clustered by school_id. Fixed effects as listed in Panel B.

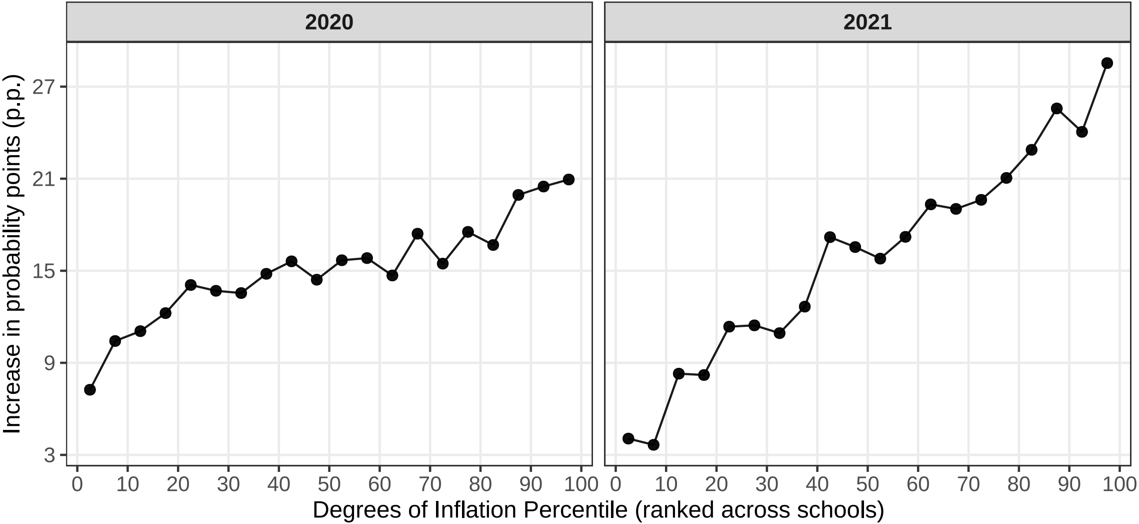
Composition English

	(1)	(2)	(3)	(4)
Baseline				
log N	-3.125^{***}		-4.226^{***}	
	(0.583)		(1.073)	
Tariff group × log N				
Lower tariff × log N		-3.435^{***}		-2.625^{*}
		(0.630)		(1.087)
Higher tariff × log N		-2.457^{***}		-7.207^{*}
		(0.494)		(2.897)
Medium tariff × log N		-2.990^{***}		-3.787^{**}
		(0.554)		(1.322)
Observations	31,495	31,495	31,495	31,495

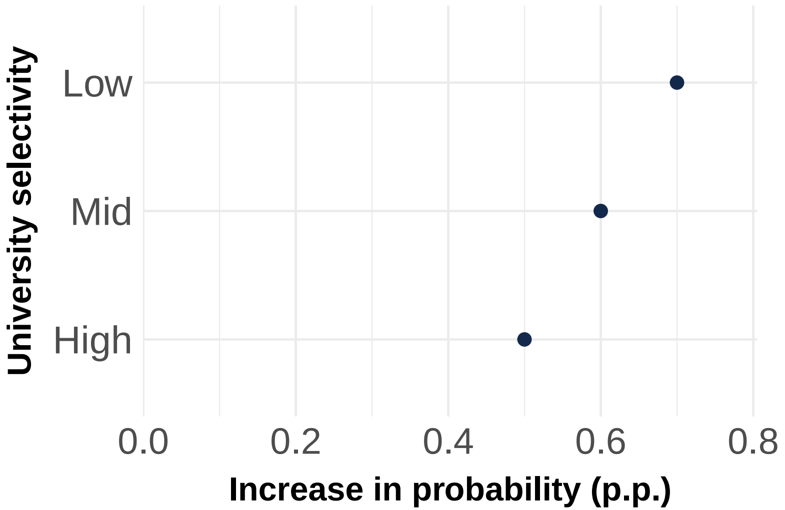
Composition Gender

	(1)	(2)	(3)	(4)
$\log N$	0.357^{***}		0.700^{***}	
	(0.105)		(0.168)	
Tariff group $\times \log N$				
Lower tariff $\times \log N$		0.413^{**}		0.601^{***}
		(0.131)		(0.162)
Higher tariff $\times \log N$		0.324^{**}		0.828^{*}
		(0.102)		(0.333)
Medium tariff $\times \log N$		0.355^{**}		0.707^{***}
		(0.117)		(0.195)
Observations	31,592	31,592	31,592	31,592

Application success and Grade inflation



Applications to less selective universities were the most affected



References I

Caroline M Hoxby and Christopher Avery. The missing one-offs: The hidden supply of high-achieving, low income students. Working Paper 18586, National Bureau of Economic Research, December 2012. URL http://www.nber.org/papers/w18586.