



Assessing Assessment: The Role of Institutional Data in Understanding Awarding Gaps

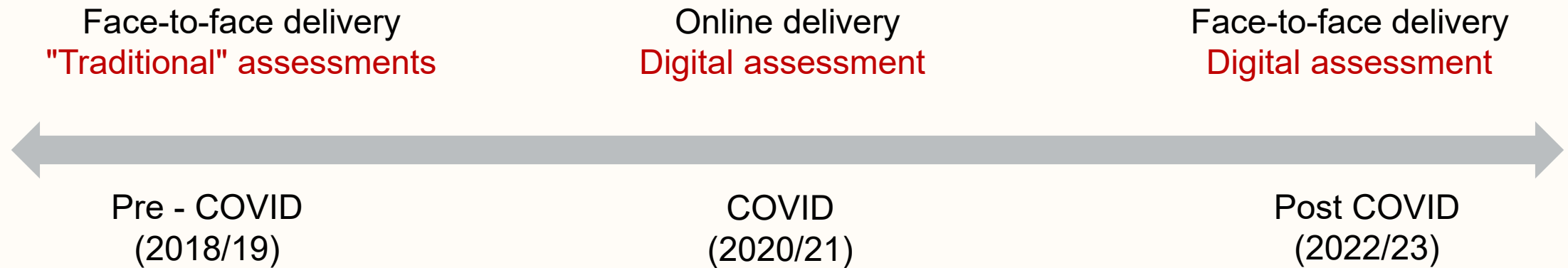
Gabriella Cagliesi and Valeria Terrones

Development in Economics Education Conference 2025
September 2025

Outline

1. Motivation and research question
2. Data
3. Descriptive statistics
4. Methodology
5. Research design
6. Main results
7. Policy implications

1. Motivation and research question



How have different **student factors** and **assessment formats** shaped awarding gaps across the pre-, during-, and post-pandemic periods?

2. Data

Source: Large institutional administrative dataset spanning for three years.

Focus on:

- Department of Economics, undergraduate single degrees (BSc Economics and BA Economics)
- Domestic UK students only (student-level data)
- Three non-consecutive academic years selected to capture distinct educational contexts

Pre-pandemic (2018/19): Face-to-face teaching + Traditional assessments

Pandemic (2020/21): Online teaching + Digital assessments

Post-pandemic (2022/23): Face-to-face teaching + Digital assessments

Data Structure:

- Performance of all grades by type and time of assessment in each academic year taken by all students of the Economics Department
- Clustering: Multiple assessments per student per year

2. Data

Dependent Variable:

- Assessment grade: Performance on student summative assessments

Independent Variables:

- Ethnicity: White, Black, Asian, Mixed, Other
- Gender
- Socioeconomic status: (using Free School Meals Eligibility criteria)
- Disability status: declared disability
- Module characteristics: academic level, type of course (quantitative or qualitative)
- Assessment formats (various categories)
- Evaluation period (in term, out of term, high stake)
- Placement
- Progression levels (first, second and third year)
- Academic year: 2018/19, 2020/21, 2022/23 (interacted with ethnicity, gender and FSM)

3. Descriptive statistics

Variable	Academic year			Avg.
	2018/2019	2020/2021	2022/2023	
Ethnicity				
White	67.5	64.5	63.8	65.2
Black	10.7	8.3	8.6	9.1
Asian	10.1	14.1	13.3	12.6
Mixed	9.3	9.1	10.1	9.5
Other	2.4	4.0	4.2	3.6
Gender				
Female	24.8	26.3	23.1	24.7
Male	75.3	73.7	76.9	75.3
FSM				
Received	7.5	10.5	8.5	8.8
Not received	92.5	89.5	91.5	91.2
Disability				
None	82.7	83.3	78.3	81.7
Cognitive	4.1	5.3	6.1	5.2
Mental Health	7.5	5.5	7.8	6.8
Multile	2.7	2.2	2.7	2.5
Phyiscial	2.7	3.3	3.6	3.2
Social	0.3	0.3	1.4	0.7

Variable	Academic year			Avg.
	2018/2019	2020/2021	2022/2023	
Assessment type				
Unseen Exam (UEX)	39.2	0.0	0.2	12.2
MCQ	14.9	32.1	45.2	30.7
Essay	16.5	13.0	17.3	15.4
Project	8.5	6.9	8.4	7.8
Report	0.2	25.7	18.6	15.7
Take Away Paper	6.4	6.9	5.2	6.3
Presentaton	6.4	6.9	5.2	6.3
Engagement	6.7	9.5	1.7	6.3
Evaluation period				
In-term	49.6	56.6	54.0	53.7
Out-of-term	40.0	37.2	39.4	38.7
High stake	10.4	6.2	6.7	7.7
N° of observations	4 679	5 982	4 473	(t=15 134)
N° of students	285	339	285	(t=909)

4. Methodology: Mixed-Effects Model

Why we use a Mixed-Effect Model approach?

Fixed and random effect's structure:

1. **Fixed effects:** group differences (gender, ethnicity, socioeconomic status) to control for unobserved heterogeneity
2. **Random effects:** account for unobserved heterogeneity at the **student** level.
 - Random intercepts: individual baseline ability differences
 - Random slopes: module type (quantitative or qualitative), student progression level, assessment timing, and time trends (academic year), allowing individual variation in response to module characteristics.

Clustering: adjusts standard errors for within-student correlation across multiple observations.

Dependent variable = fixed effects | random effects(slope): intercepts, vce (cluster)

*Total mark = gender, age, socioeconomic status, (...) | student_id: i.dummy_out i.module_level
i.year, vce(cluster student_id)*

5. Research design

How we use a Mixed-Effect Model approach?

1. Mixed effect model for **each year** (pre, during and post pandemic) .
2. Mixed effect model using a **pooled data** to be able to test the hypothesis across years.

Four main specifications:

Model 1: general model

- Students' performance in all types of summative assessment distinguishing only ...out/in high stake

Model 2: quantile model

- Similar as model 1 but using quantile analysis

Model 3: assessment format analysis (coursework)

- Students' performance only using in-term assessments to distinguish different assessment formats (timed coursework, essays and others, written reports, technical reports, interactive and engagement) during teaching period.

Model 4: timed and untimed assessment analysis (final exams)

- Students' performance in final exams (timed and untimed) during the evaluation period.

6. Main results: summary

Model 1: general model

- **Ethnicity:** ethnic awarding gaps remain entrenched, with only a temporary narrowing during the emergency online period of 2020/21.
- **Female:** outperform male (4.22 points) ($p=0.001$)
- **Socioeconomic status:** socioeconomic disadvantage, (FSM), interact differently with gender. Pre-COVID—often posing greater challenges for females—the post-COVID period witnessed a marked reversal. FSM males emerged as the most educationally disadvantaged group, experiencing a ten-point decline in performance.
- **Disabilities conditions:** students with mental health conditions face persistent penalties, which intensified in 2022.

Model 2: quantile analysis (stress the differences in quartile)

- **Asian** students achieved complete parity across all quantiles, reversing severe pre-pandemic penalties.
- **Black** students retained persistent 4–5 point gaps, most acute among low achievers (**Q25**).
- **Socioeconomic** disadvantage intensified, with a dramatic reversal for FSM males—turning a 7.6-point advantage at **Q25** pre-COVID into an 11-point penalty—representing the largest intersectional shift observed.

6. Model 1: persistent ethnic inequalities

- All groups –except the Asian students- experienced lower grades in 2022 relative to 2018. (graph 1)
- Some minority ethnic groups show significant and stable penalties relative to White British students (–4.6 to –6.0 points), indicating that core ethnic awarding gaps remain embedded despite the temporary narrowing observed in 2020.

Graph 1: Ethnic Performance Gaps

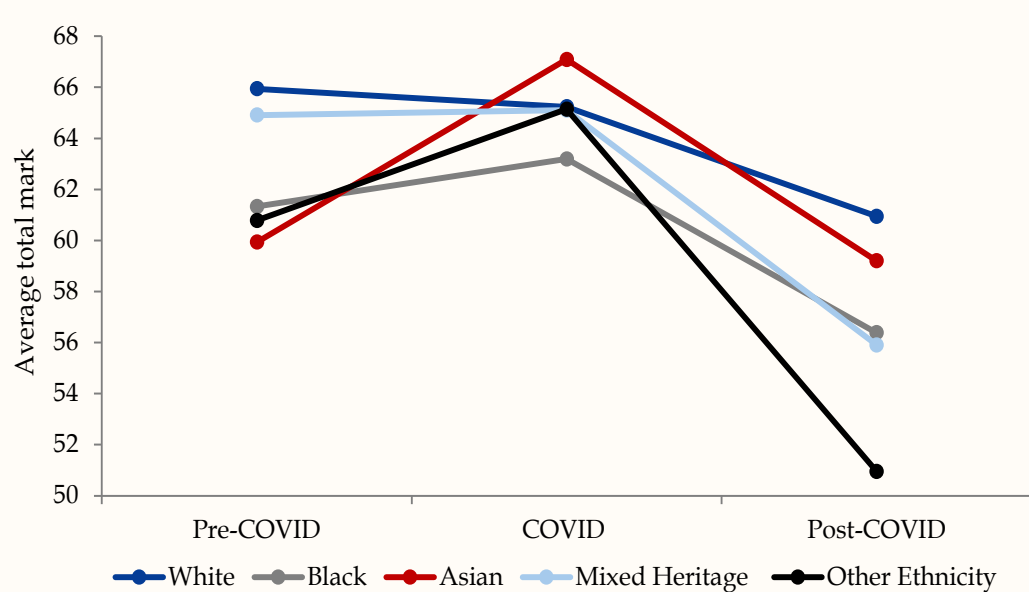
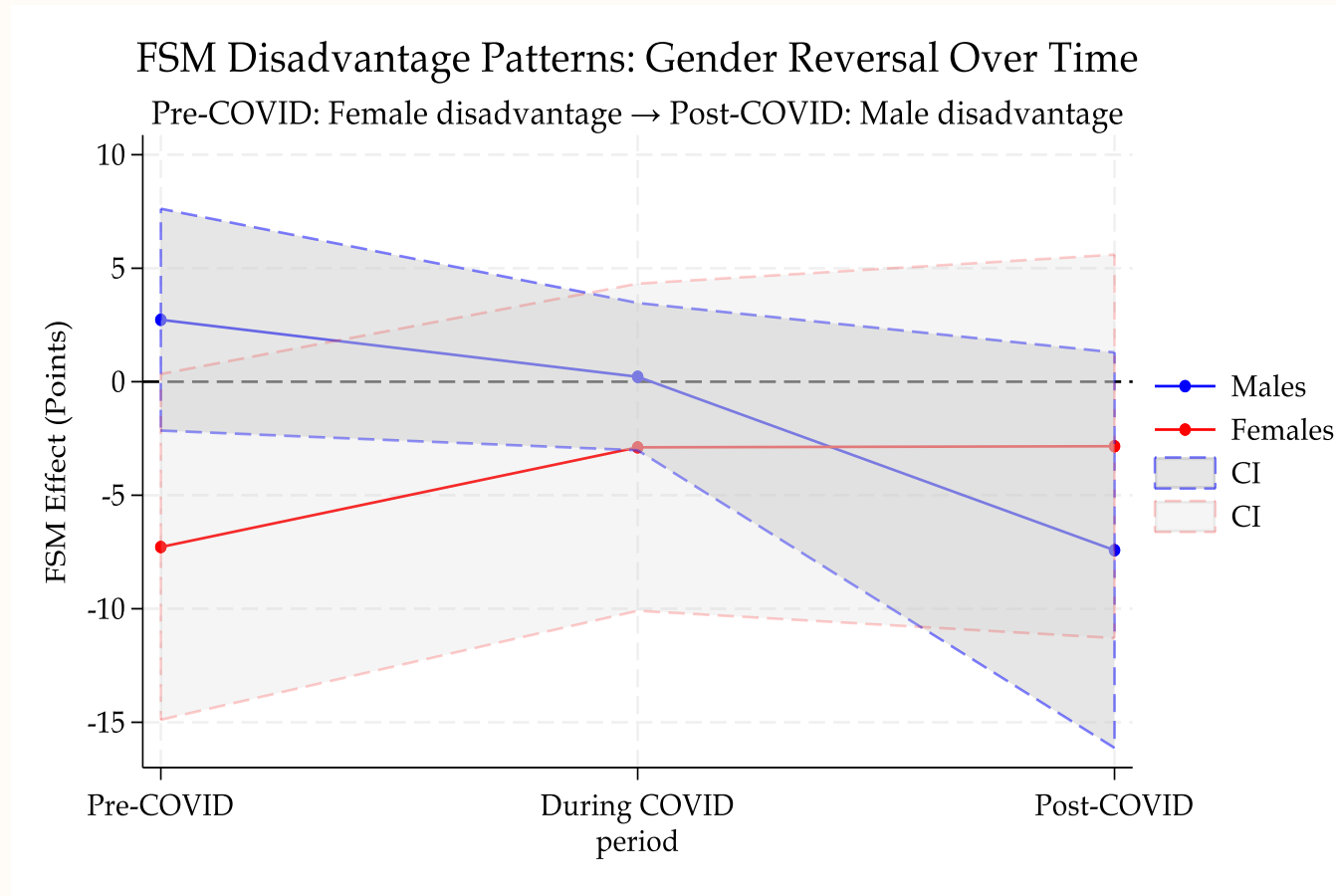


Table 1: Ethnic Performance Gaps

Ethnic Group	Pre-COVID Gap (2018/19)	COVID Gap (2020/21)	Post-COVID Gap (2022/23)	COVID Status	Post-COVID Status
Black	-4.61** (p=0.014)	-2.04 (p=0.364)	-4.56* (p=0.076)	Gap Reduced	Gap Persists
Asian	-6.01*** (p=0.002)	1.85 (p=0.244)	-1.75 (p=0.403)	Gap Eliminated	Major Improvement
Mixed Heritage	-1.03 (p=0.561)	-0.13 (p=0.946)	-5.04* (p=0.064)	No Baseline Gap	Gap Emerged
Other Ethnicity	-5.16 (p=0.189)	-0.10 (p=0.974)	-9.99** (p=0.050)	No Baseline Gap	Significant Gap

6. Model 1: gender and FSM

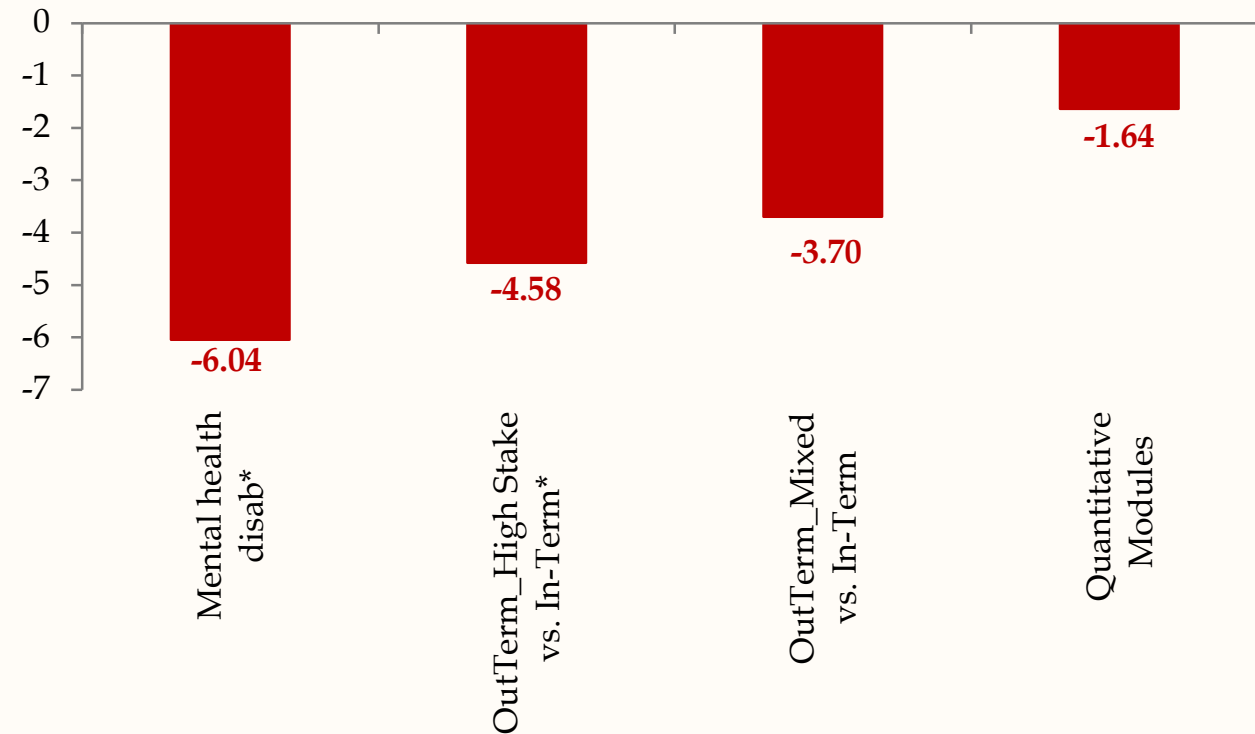
- Gender advantages provided some protection for female FSM students, making working-class males the most vulnerable group in the post-pandemic landscape.
- Female students outperformed their male peers by more than four points ($p=0.001$).
- Socioeconomic disadvantage, (FSM) interact differently with gender: pre-COVID-greater challenges for females; almost no effect in Covid; post-COVID reversal. FSM males emerged as the most educationally disadvantaged group, experiencing a ten-point decline in performance ($p=0.046$). This represents a statistically significant intersectional shift ($p=0.050$), where the combined identity of being male and from a low-income background transitioned from relative resilience to acute vulnerability.



6. Model 1: other variables

- Students with specific disabilities (mental health conditions) faced a disadvantage during the period of study.
- Quantitative subjects also, imposed a consistent performance penalties across all demographic groups;
- Out of term and particularly high stake exams show a negative effect

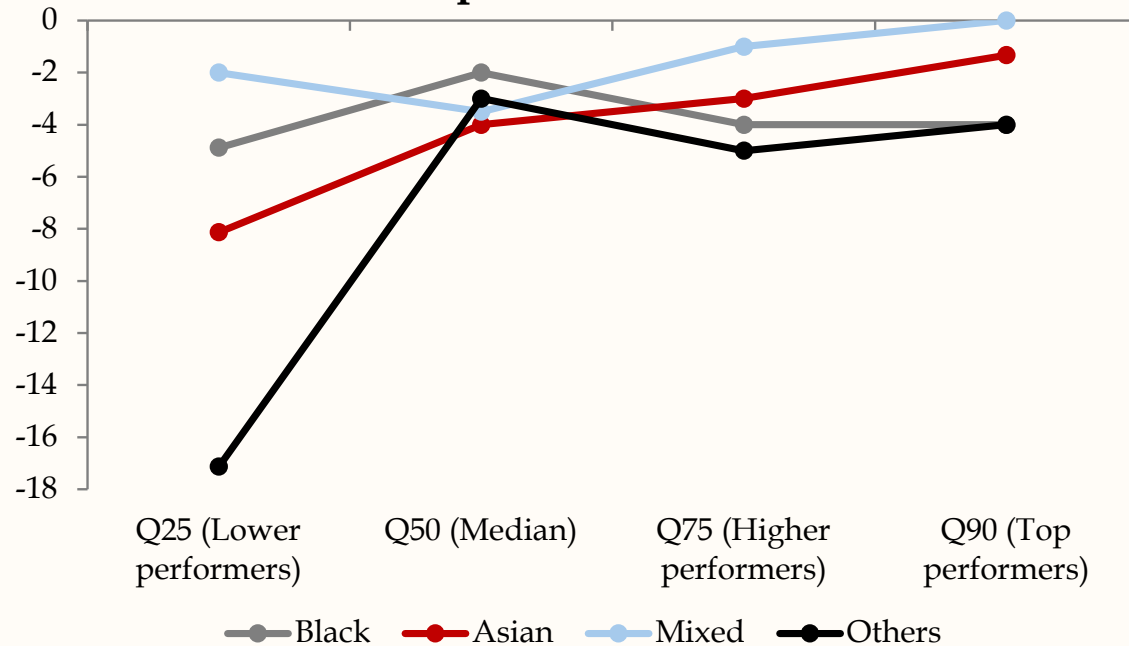
Graph 3: Effects on other relevant variables
between 2022 and 2018



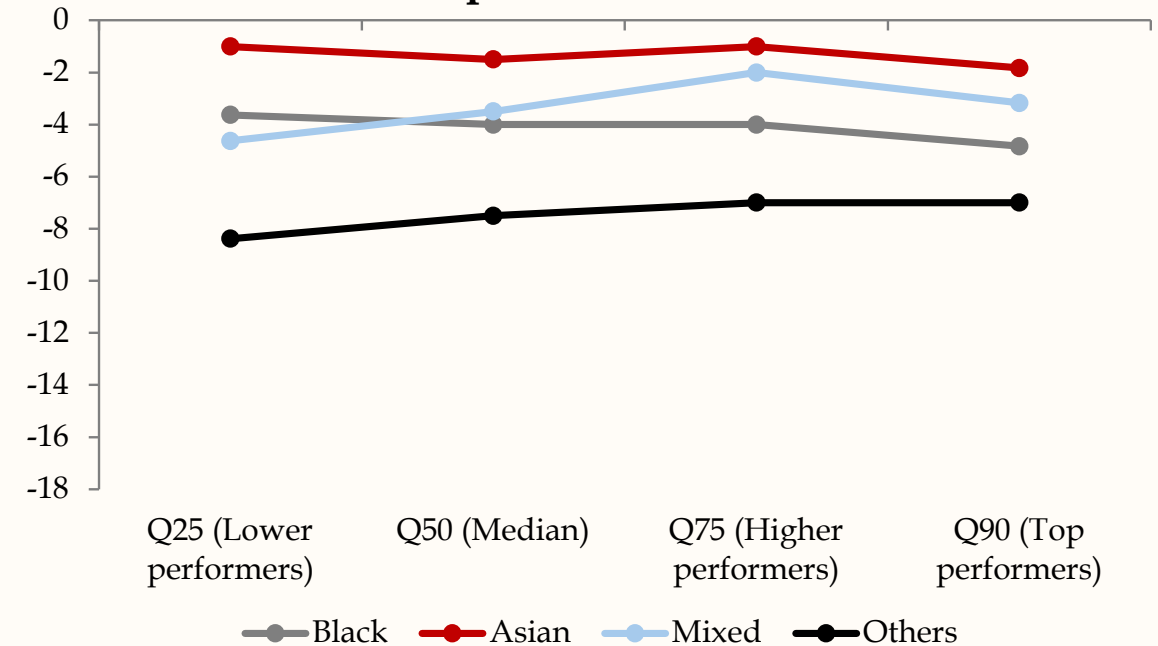
6. Model 2: quantile analysis

- **Distribution of ethnic inequalities:** Asian students achieved equity across the entire distribution, and “other ethnicity” and “black” students show the most severe and persistent disadvantages.

Graph 4: Distribution of the Ethnic Performance Gaps, pre-COVID



Graph 5: Distribution of the Ethnic Performance Gaps, post-COVID



6. Model 2: quantile analysis

Table 2: FSM Intersectionality across performance distribution

Quantile	Pre-COVID	During COVID	Post-COVID
Q25 (Lower performers)	+4.63* (p=0.005)**	-0.19 (p=0.886)	-11.13* (p<0.001)**
Q50 (Median)	+2.00 (p=0.145)	0.00 (p=1.000)	-3.50 (p=0.021)**
Q75 (Higher performers)	+2.00 (p=0.114)	+1.00 (p=0.322)	-5.00* (p<0.001)**
Q90 (Top performers)	+2.33 (p=0.115)	0.00 (p=1.000)	-5.50* (p=0.001)**

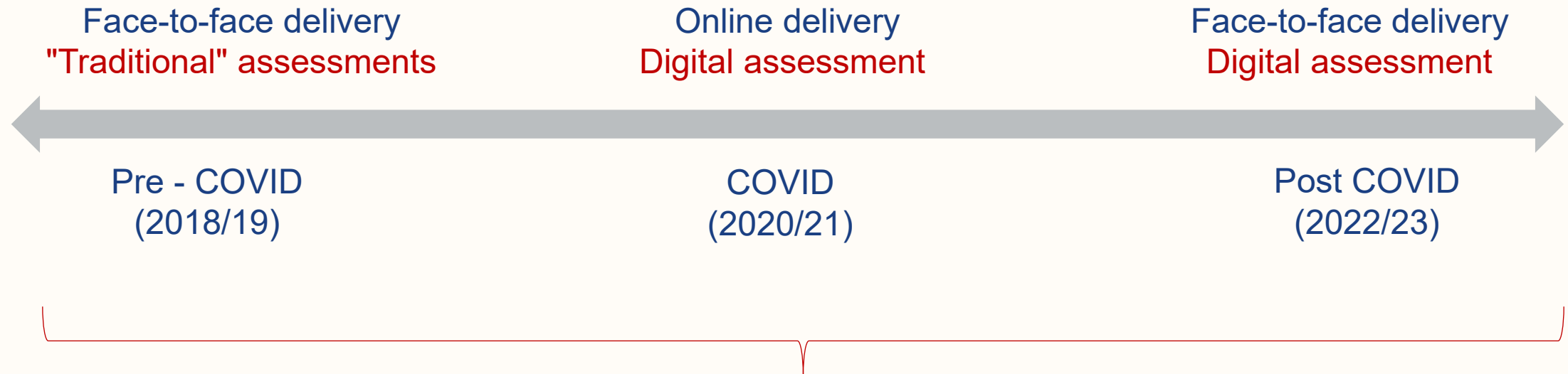
- **Socioeconomic disadvantage intensified with a dramatic reversal for FSM males** —turning a 7.6-point advantage at Q25 pre-COVID into an 11-point penalty—representing the largest intersectional shift observed.
- These findings show that post-pandemic inequality is dynamic, multiplicative, and most severe among lower-achieving students, requiring targeted, intersectional interventions.

*** p<0.01, ** p<0.05, * p<0.1

Robust standard errors in parentheses.

Source: Source: institutional data from QAA-funded Collaborative Enhancement Project "Using Institutional Data to Address the Causes of Awarding Gaps".

6. Model 3: assessment format analysis (coursework)

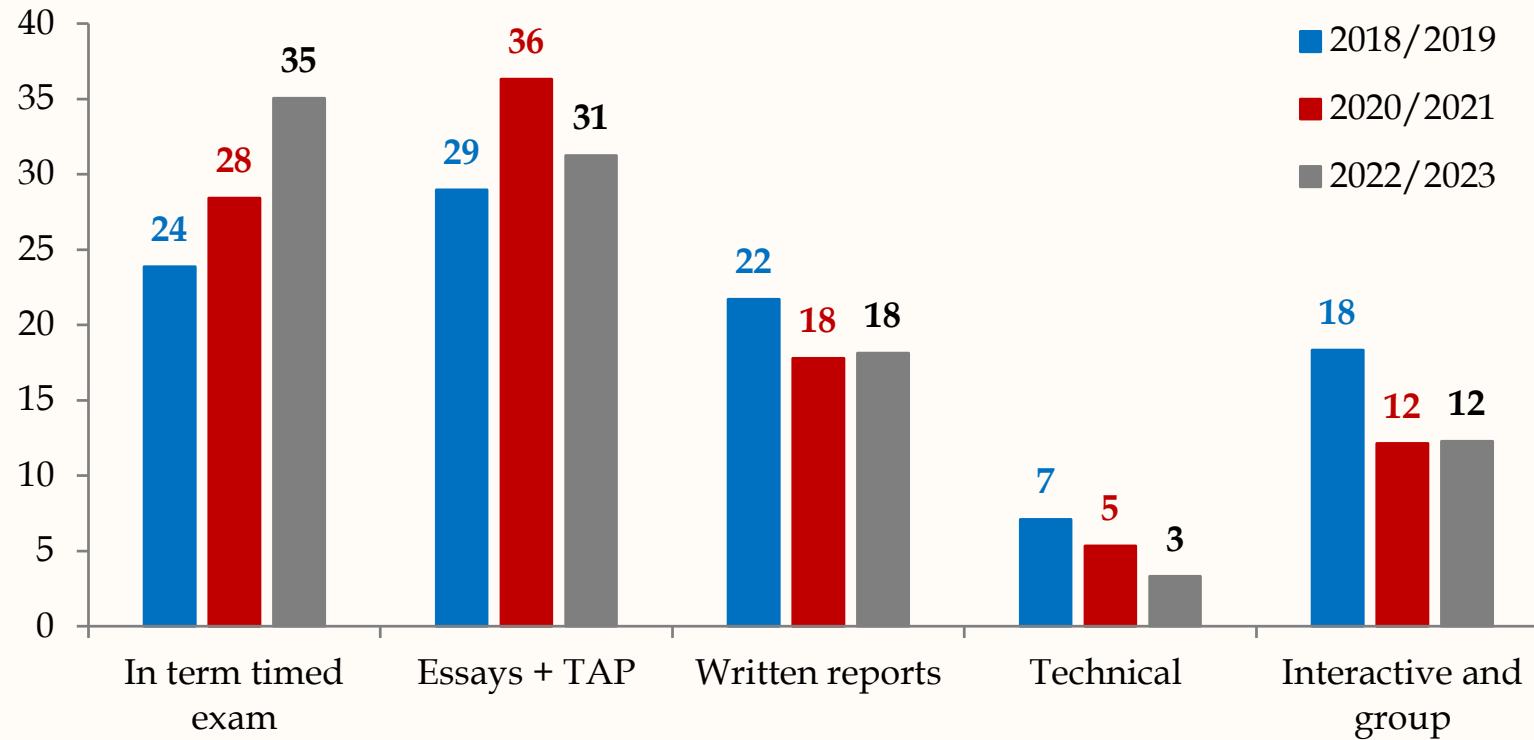


Type of assessments:

- Timed assessments (e.g. test (2018, 2020, 2022); CEX (2020 and 2022); MCQ (2020))
- Essays + TAP
- Written reports (e.g. reports, project, portfolio)
- Technical (e.g. problem sets, media, software exercises)
- Interactive and engagement (e.g. group presentation, oral, group written submission)

6. Model 3: Assessment format analysis (coursework)

Graph 6: Assessment format analysis (% of total)



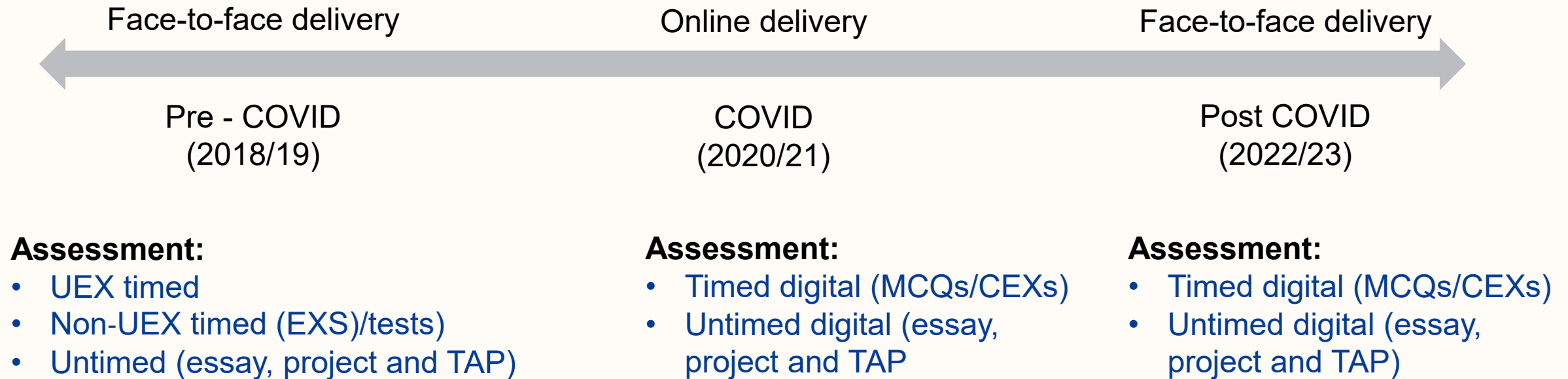
6. Model 3: assessment format analysis (coursework)

Table 3: GAPS post covid considering different assessment formats

Ethnic Group	Timed Exams	Essays & Other	Written Reports	Technical Reports	Interactive
Black Students	-1.05	-5.19	-12.44**	-0.24	-2.29
Asian Students	+1.19	-4.83	-1.34	-3.10	-8.33
Mixed Heritage	+4.50	-11.05**	-13.57*	-9.64	-7.50
Other Ethnicity	+2.37	-11.03*	-12.45	-10.68*	-20.40(**)

- Post-COVID, Black students/mixed heritage have a significant disadvantage in written reports (in pre-covid they has an advantage); Asian have eliminated the disadvantage in the same type of assessment.
- Maybe resource access issues (technology, internet, study spaces); digital divide impacts on portfolio/project submission

6. Model 4: timed and untimed assessment analysis



(untimed final exams exclude: problem sets group work, presentation, practical assessments)

Assesment (%)	2018/2019	2020/2021	2022/2023
UEX (Traditional)	79.28	0	0
Non-UEX (Timed)	5.32	33.92	50.34
Non-UEX (Untimed)	15.4	66.08	49.66

6. Model 4: timed and untimed assessment analysis

Table 4: Awarding gaps considering timed and untimed assessments outside teaching period (point changes, p-values)

Ethnic Group	2018 UEX	2018 Timed exams	2018 Untimed exams	2020 Timed exams	2020 Untimed exams	2022 Timed exams	2022 Untimed exams
Black	-5.000** (0.027)	-3.959 (0.440)	-1.084 (0.618)	0.035 (0.990)	-2.836 (0.316)	-2.369 (0.500)	-6.479** (0.034)
Asian	-4.147 (0.138)	-5.437 (0.363)	-14.858*** (0.004)	0.508 (0.817)	2.094 (0.295)	0.443 (0.815)	-1.687 (0.481)
Mixed	-0.192 (0.922)	9.220*** (0.007)	-5.136 (0.223)	-2.113 (0.351)	3.916** (0.029)	-1.947 (0.3035)	-4.411 (0.115)
Other	-7.874 (0.109)	-3.307 (0.839)	-17.702 (0.167)	-2.786 (0.449)	-0.031 (0.992)	-13.621** (0.019)	-10.273*** (0.002)
Gaps group	Black Disadvan.	Mixed Advant.	Asian Disadvan.	No gaps	Mixed Disadvant.	Other G. Disadvan.	Black and Other Disadv.

- Pre pandemic: UEX for black and Untimed written assessments (essays, take-home reports) for Asian were linked to large disparities.
- Timed and untimed assessments, during covid (2020) eliminated gaps; gaps reappeared in 2022 for some groups (untimed exam: so, it is not just a matter of online exams).

*** p<0.01, ** p<0.05, * p<0.1 P-values in parenthesis. White students as a reference group. Positive values indicate improvements.

7. Policy implications

This two-step approach demonstrates how initial broad-brush findings can mask important pedagogical mechanisms, highlighting the importance of granular institutional data analysis for evidence-based assessment policy.

Equity-oriented assessment reform must extend beyond diversifying formats to:

- Recognise assessment format as a structural determinant of awarding gaps;
- Balance inclusivity and integrity when considering online unproctored assessments in an era of AI;
- Target interventions for groups facing compounded disadvantages, particularly working-class males and students with mental health conditions;
- Retain accessibility innovations that benefited physically disabled students during the pandemic.



Thank you!

Gabriella Cagliesi and Valeria Terrones

Development in Economics Education Conference 2025
September 2025