The death of exams? Grade inflation and student satisfaction when coursework replaces exams

Petar Stankov, @stankovecon, petar.stankov@rhul.ac.uk

DEE Conference, 09/2023

ROYAL HOLLOWAY UNIVERSITY OF LONDON

What is this paper about?

Rationale for replacing a final exam with coursework

Structure of the assessment reform

Reform outcomes:

    1) Student satisfaction

    2) Grade inflation

Grade simulations: Alternative reforms

Is this reform scalable?

1. Why did coursework need to replace an exam?

Known issues with final exams:

1. Perpetuate **attainment gaps** (Ahlburg & McCall, 2021; Shaw & Tranter, 2021; Sohn, 2012),

2. Reinforce **social inequalities** (Baert & Verhaest, 2021; Feng & Graetz, 2017).

3. Prone to **grade inflation** before COVID:
   - in the UK (Bachan, 2017; Frank et al., 2019),
   - the USA (Jewell et al., 2013) and
   - elsewhere (Karadag, 2021; Müller-Benedict & Gaens, 2020).

4. Online exams during COVID **amplified grade inflation** (Vazquez et al., 2021).

5. Online exams incentivise cheating, collusion, and commissioning.

- **Question**: Is a *superior assessment mix* feasible and desirable?

Superior assessment mix

What is a superior assessment mix?

An alternative assessment mix would be superior if it:

- retains or expands the **learning outcomes**,

- preserves the **grade distribution** or aligns it within professional norms, and

- improves **assessment authenticity**.

Where: mid-sized (55 stud.) advanced UG economics module

Before: a typical pre-pandemic assessment mix

After: the final exam was replaced by a 60% coursework assessment.

Component / Term	Autumn 2021/22	Autumn 2022/23
Assessment Component	Homework 1 (20%)	Homework 1 (20%)
	Homework 2 (20%)	CW1: Presentation (15%)
	Final exam (60%)	CW2: Data work (15%)
		Homework 2 (20%)
		CW3: Policy brief (15%)
		CW4: Video promotion (15%)

More than a mechanical replacement of a final exam:

1. New **learning outcomes**: presentation skills, coding and data-analytical skills, framing results and conclusions for policy circles, and promoting own work for both professional and non-professional audiences on social media.

2. Embed **continuous feedforward and feedback** into the teaching and assessment design

3. Scaffolding the **skills** that came with the new assessments

4. Non-essential material scrapped for more **workshop time**

3. The effects: What worked well?

Positive outcomes:

1. Learning outcome proliferation

2. Continuous engagement due to regular deadlines + feedback/workshop sessions

3. No attainment gaps (AGs): no gender or ethnic disparity between the grades after the reform -> contribution to declining AGs.

=> student satisfaction increased

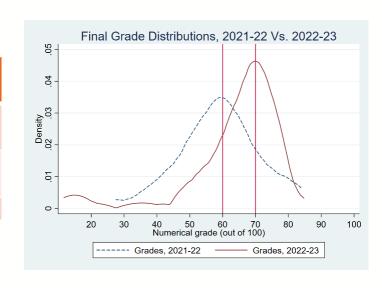# 4. The defects: What did not work well?

Two undesired effects:

1. Workload: the demand for marking, feedback and feedforward soared, standard for work-authentic assessments (Bilgin et al., 2017): Significant resource implications.

2. Grade inflation

Table 2: Summary statistics of Final Grades by academic year

Acad. Yr.	Mean	Median	SD	IQR	% First	% 2.1.	N
2021/22	59.00	58.40	12.71	13.80	19.23	26.92	26
2022/23	63.95	66.95	14.76	12.20	40.00	36.36	55
Total	62.36	65.40	14.24	14.60	-	-	81

Final Grade Distributions, 2021-22 Vs. 2022-23

Grades, 2021-22 — Grades, 2022-23

4.2. Origins of grade inflation (1)

Table 3: Summary Statistics for Homework 1 and Homework 2

Pre-treatment difference in the student abilities over the two cohorts? Not significant.

Acad. Year	Mean	Median	SD	IQR	Min	Max
	Homework 1					
2021/22	70.19	**68.00**	9.02	10.00	55.00	88.00
2022/23	64.30	**65.00**	9.78	10.00	32.00	82.00
	Homework 2					
2021/22	63.17	**65.00**	17.19	13.50	0.00	88.00
2022/23	60.92	**62.00**	8.23	13.00	45.00	85.00
Notes: The table presents the mean, median, standard deviation (SD), inter-quartile range (IQR), minimum and maximum values of homework 1 and homework 2 assignments over the two iterations of the module. Sample size is 26 in 2021/22 and 55 in 2022/23.						

4.2. Origins of grade inflation (2)

Table 4: Coursework grades in 2022/23 and final exam grades in 2021/22

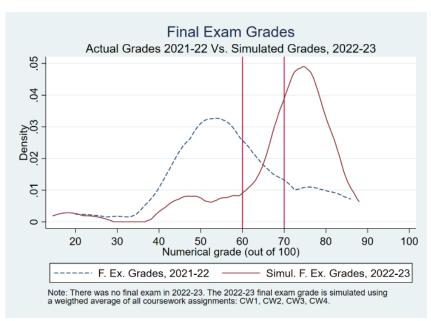| Variable	Mean	Median	P-25	P-75	SD	IQR	Min	Max
CW1	65.18	66.50	58.00	72.00	10.51	14.00	38.00	92.00
CW2	**96.31**	**100.00**	**100.00**	**100.00**	**10.89**	**0.00**	**55.00**	**100.00**
CW3	65.85	66.50	62.00	72.00	8.49	10.00	30.00	78.00
CW4	67.12	65.00	62.00	72.00	9.92	10.00	48.00	92.00
Final Exam 2021/22	57.72	55.00	50.00	66.00	14.65	16.00	20.00	86.00

Notes: The table presents the mean, median, grades at the 25th and the 75th percentile, standard deviation (SD), inter-quartile range (IQR), minimum and maximum values of all four coursework (CW) assignments, as well as the final exam grades in 2021/22. The final exam was replaced by coursework in 2022/23.
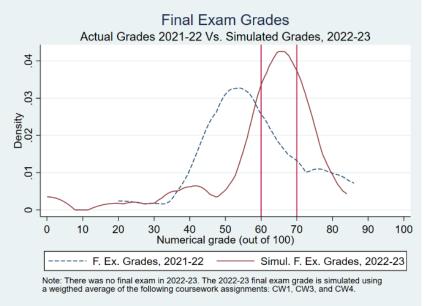
CW2 gives students a real boost.

Counterfactual synthetic final exams:

(a) Final exam with CW2

(b) Final exam w/o CW2
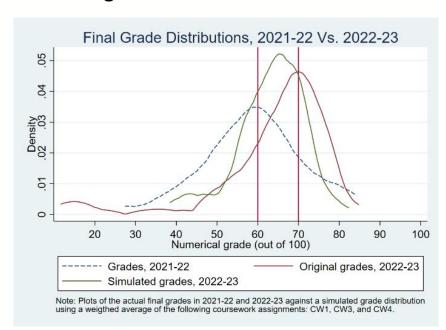
Resulting simulated distribution:

Fig 3. Original and simulated grades

The new, simulated distribution:

- Not too harsh (as 2021/22) nor too generous, as the original 2022/23

- Mean= 63.28, median=65.00, in line w/ expectations.

- First (21.81%) and Upper-second (52.73%),

=> an overall share of good degrees at 75% close to the original 2022/23 (76.36%) and the sector-wide pre-pandemic share (Office for Students, 2022, 2023).

6. Is the reform scalable?

Neither a firm Yes nor a definite No. Several factors at play:

- salience of student satisfaction

- exam authenticity vs. coursework authenticity

- tolerance to grade inflation

- capacity of AI to replace learning: videos, case studies, presentations, coding…

- can AI be employed to reduce admin costs? feedback, emailing and marking

- class size and availability of teaching assistants

More examples on the way.

7. Conclusion, policy implications

- Motivation, structure, and effects of a post-pandemic assessment reform in an advanced undergraduate economics module

- 60% final exam and 40% coursework replaced by continuous coursework =>

    (+) student learning, engagement, and satisfaction

    (-) grade inflation, workload

- Grade simulations offered a way out of grade inflation

=> a superior assessment mix: keep student satisfaction high, preserve grades, stabilise staff workload and proliferate learning outcomes.

=> Before this reform is rolled out at scale, the reform better suited to smaller classes.

The death of exams? Grade inflation and student satisfaction when coursework replaces exams

Petar Stankov, @stankovecon, petar.stankov@rhul.ac.uk

DEE Conference, 09/2023

ROYAL
HOLLOWAY
UNIVERSITY
OF LONDON