Narrowing awarding gaps: the contributory role of policy and assessment type

M. Gabriella Cagliesi, Denise Hawkes & Susan Smith

Published online: 18 May 2023.

Submit your article to this journal ↗

Article views: 609

View related articles ↗

View Crossmark data ↗

Routledge
Taylor & Francis Group

∂ OPEN ACCESS   Check for updates

Narrowing awarding gaps: the contributory role of policy and assessment type

M. Gabriella Cagliesi [a], Denise Hawkes [b] and Susan Smith [c]

[a]Department of Economics, University of Sussex Business School, Brighton, UK; [b]Department of Economics, Anglia Ruskin University, Chelmsford, UK; [c]Department of Finance, Accounting and Economics, UCL School of Management, London, UK

ABSTRACT
Awarding gaps between various groups of students persist across the Higher Education sector, yet the responses designed to address the contributors remain localised. The sudden spread of COVID-19 led to various responses across the University sector creating an unprecedented natural experiment and offering the opportunity to compare outcomes from these measures with prior cohorts. This study seeks to investigate the effects of two COVID-19 interventions on students' performance in the Business and Management discipline at a UK university. The specific COVID-19 measures considered here are the move to online assessments and the new grade policy to ensure the pandemic did not affect students' outcomes adversely. We use a Kernel Propensity Score and a Quantile Difference in Differences models to estimate the treatment effect of the two COVID interventions on the treated group, namely term two students' performances of the academic year 2019/20. Our results indicate that the effects of both COVID interventions supported the outcomes of international students, thereby narrowing the awarding gap. Findings suggest firstly that institutional policies adopted in crises should seek to address potential adverse effects on student outcomes for the period of disruption, indicating that significant care should be taken in their drafting. The policy, in this case, was found to have achieved its aim. Secondly, the move to new modes of assessment combined with detailed briefings from faculty may have served to uncover aspects of the hidden curriculum for this group, contributing to a narrowing of awarding gaps between different groups of students.

Introduction

Recent years have seen an increased policy focus on reducing and eliminating persistent undergraduate awarding gaps across the Higher Education sector. Differences in degree attainment between various ethnic groups were first drawn to the attention of the higher education community in the 1990s (Connor et al., 1996) and have subsequently been confirmed in a range of studies e.g. (Fielding et al. 2008; Richardson 2008). They are now an area of focus for sector policy e.g. Office for Students in the UK, and consequently institutional strategies. Whilst the international awarding gap (differences in

---

degree attainment between students classified as 'home' and those classified as 'international' (non-EU) as per the Higher Education Statistical Agency) has received less attention from policymakers and researchers (with a few notable exceptions e.g. (Iannelli and Huang 2014)) it is also now a focus of higher education institutions as they work to embed more equitable curricula and expand their student intakes. Literature is now starting to draw linkages between the social justice agenda of widening participation and internationalisation activity (Gayton 2020). In short, it is critical to better understand and address the drivers of differential student attainment in higher education.

To date work on awarding gaps has focused on controlling for various characteristics which could affect attainment (Cotton et al. 2016), finding a strong association between entry grades and degree outcomes (Jones et al. 2017; Smith 2016). Studies show that despite controlling for prior attainment, age, gender, and discipline most ethnic minorities experience awarding gaps and that females outperform males (Cotton et al. 2016). Qualitative work associates student outcomes with feelings of belonging, and successful transition to higher education (Jones 2018).

Whilst existing studies allude to assessment practices' contribution to the observed gaps (Richardson 2015) empirical evidence is lacking. This research, therefore, seeks to contribute to the debate by illuminating the extent to which mode of assessment is a factor as a response to Richardson's (2015) observation that 'we do not know what aspects of teaching and assessment practices are responsible for variations in the attainment gap.' (278)

The research focuses on students studying Business and Management subjects, which was selected because more qualifications were awarded in Business and Management than any other subject in 2019/20 (HESA 2021), with non-EU students comprising 27% of the 2019/20 cohort. Whilst the study is conducted in one discipline at one UK institution the findings are of interest more broadly as awarding gaps are coming into focus in other countries both within their 'home' cohorts (Palacios and Alvarez 2016) and also for 'international' cohorts (Crawford and Wang 2015). The advantage of adopting a case study approach is that it enables control of key variables, with students who have entered the discipline based on similar entry criteria, and who have had a comparable experience of teaching and assessment during their studies (Jones 2018).

This observational study seeks to uncover the extent to which the mode of assessment contributes to the reported awarding gaps by undertaking a statistical analysis at a modular level comparing the pre-COVID assessment regime to the wholesale change in the assessment regime in response to the COVID-19 pandemic. In the UK, many universities adopt a degree classification algorithm that includes student performance in the second and final year of the degree, albeit with variations in how they are weighted. The study offers new insights into the influence of assessment type on student outcomes at second year (level 5 (L5)) and final year (level 6 (L6)), including first attempt, deferred or referred assessments (resits), and the use of the exceptional COVID-19 policy (the'No-Detriment policy' (ND)) in progression and award decisions. The ND approach was used across UK universities to mitigate the potential adverse effects of the rapid adoption of online assessment during the first wave of lockdowns and the general pandemic disruption. However, institutional policy details varied significantly (Chan 2022). The changes implemented in response to the pandemic emergency led to focus on the following two research questions (RQ):

RQ1: To what extent did the institutional policy approach adopted in response to COVID-19 create a backstop for student grades?

RQ2: To what extent did the change in assessment modes contribute to narrowing (or otherwise) of observed awarding gaps?

Awarding gaps

The continued persistence of awarding gaps of various degrees of magnitude across different institutions and subjects that cannot be explained by entry qualification (Richardson, Mittelmeier, and Rienties 2020) implies that there is some degree of influence arising from the structures of teaching

and assessment (Richardson 2015). As a result, both the institutional context and the cohort composition are likely to be important considerations.

The literature related to awarding gaps is increasingly nuanced, taking into account the complexities of intersectionality, and structural barriers among other factors (Codiroli Mcmaster 2021). It now spans a range of studies focusing on the student experience and its influence on attainment, often finding multiple determinants ranging from the curricula and learning environment; staff-student relationships; social, cultural, and economic capitals; and psychosocial and identity factors (HEFCE 2015). To investigate these determinants, further studies focus on each aspect and its impact on retention and attainment. For example, finding that Chinese students' final degree mark was not influenced by gender, prior academic performance, prior academic qualifications or degree programme (Crawford and Wang 2015). The majority of studies adopt qualitative research methods, are typically of small scale, and are conducted at one institution. Recent studies have also started to engage with questions of intersectionality (Richardson, Mittelmeier, and Rienties 2020). To date, most studies tend to focus on how the student can change themselves rather than the university structures changing to accommodate a broader conception of students (Koutsouris, Mountford-Zimdars, and Dingwall 2021).

Prior studies indicate that international students can misinterpret the requirements of UK higher education and therefore overestimate their likelihood of being awarded a 'good' degree reflecting a lack of understanding of the system (Cotton et al. 2016). This misunderstanding may be partially explained by the concept of the hidden curriculum, the difference between the curriculum as designed by the academic, and how it is experienced by the learner (Barnett and Coate 2005) which can form a helpful construct to explain differential student outcomes. From origins in secondary education, the hidden curriculum offers explanatory potential in higher education where 'less densely codified curricula provide a landscape of potential for multiple hidden curricula to exist' (Cotton, Winter, and Bailey 2013, 193). The hidden curriculum exists in sharp contrast with the explicitly technical nature of contemporary higher education with clear learning outcomes, assessment briefings, etc., and assumed transparency of the curriculum (Orón Semper and Blasco 2018). Four primary meanings have been attributed to the concept of the hidden curriculum: implicit expectations relayed by educators; unintended learning outcomes; implicit messages conveyed by the structures of higher education, and students' interpretations of how to achieve reward and success (Portelli 1993). The hidden curriculum is experienced variably by students as 'Students are not simply responding to the given subject—they carry with them the totality of their experiences of learning and being assessed' (Boud 2013, 39).

The hidden curriculum has been disaggregated into various components both from the perspective of the educator and the student. From an educator's perspective three primary influences have been highlighted, firstly what to include or exclude from the curriculum; secondly the approach to teaching and the messages this conveys; and finally the physical or virtual learning environment (Cotton, Bailey, and Tosdevin 2020). Despite efforts to create transparency in assessment, the hidden curriculum remains likely to operate below the surface (Wicking 2020). Sambell and McDowell (1998) observe that innovative assessments can disrupt the hidden curriculum thereby reducing the gap between the formal curriculum and the hidden. As such changing the assessment dialogue, prompted by the pandemic may have served to uncover aspects of the hidden curriculum. This comprised both a change in classroom behaviour as educators worked to prepare students for revised assessment modes along with the practical implementation of the revised assessments. To investigate the effect of assessment mode changes on student outcomes our work evaluates data at a modular level before the application of the institutional no-detriment policy.

Data

Our data set is drawn from administrative data held by the institution in question, a mid-sized research-intensive UK University. It covers all UK-domiciled and overseas (non-EU) students in the

Table 1. Descriptive statistics: performance by year. Panel Award Board (*).

HESA Business and Management subjects			UK domiciled		International Non-UK		International Chinese		Direct Entry		Direct Entry International
2019/20	No.	Male	No.	Male	No.	Male	No.	Male	Intern.	China	China %
L6	755	57.8%	289	62.3%	446	54.9%	311	51.1%	305	251	82.3%
L5	881	59.8%	348	66.1%	553	55.7%	253	48.9%	288	185	64.2%
2018/19											
L6	678	53.8%	221	58.4%	457	51.6%	310	50.0%	308	242	78.6%
L5	762	57.2%	288	62.2%	474	54.2%	317	50.1%	317	259	81.7%
2017/18											
L6	663	53.5%	184	51.1%	479	54.5%	325	49.5%	337	267	79.2%
L5	740	55.8%	262	61.8%	478	52.5%	318	50.9%	321	252	78.5%

* Progression Award Board (PAB) results on first attempt (January/May exams)

HESA Business and Management subject areas. This study includes three years of data, 2017/18, 2018/19, and 2019/20 on students' biographical characteristics and the grades achieved on all individual modules. These characteristics are summarised in Table 1.

The academic data include information about their journey through the university including: which academic year they joined the university, learning needs, and other adjustments such as extenuating circumstances claims, as well as information on the degree outcomes and grades awarded. Throughout we use the terminology of the Framework for Higher Education Qualification that uses the 'level' codes L3, L4, L5, L6 to indicate different stages of student progression, namely Foundation Year (L3), Year 1 (L4), Year 2 (L5), and final year (L6) of a degree.

The information on module grades is very detailed. The data record each student's first attempt grade received in the module's exam session (as recorded at the Progression Award Board (PAB)) as well as the resit (referral) or sit grades (deferral) received in the August exam session if these occur. These two sets of grades are referred to as 'First attempt' and 'Final' (after resit) grades, respectively. They coincide when students either succeed at a first attempt or are not granted any re-sits or sits.

A third grade is available for the modules delivered in the second term (T2) of 2019/20. This grade reflects the application of the ND or safety net policy designed to respond to the COVID-19 emergency, and it is referred to as 'Final after ND' grade. The key feature of the ND considered in this study is to use the 2019/20 student's T1 average grade (i.e. the average of modules taught under non-pandemic circumstances) to identify whether their T2 grades (the COVID term) seemed unusually low and to uplift the T2 module results if this were the case. Whenever the ND policy is not binding or not applicable, the 'Final after ND' grades coincide with the 'Final' (after resit) ones.

In addition to the ND policy, the University responded to the physical restrictions imposed by the pandemic emergency by moving all on-Campus activities and all assessments online. Traditional assessments were replaced by alternative instruments such as Take-Home Paper (TAP) or online quizzes (Multiple Choice Questions (MCQs)). The granularity of the students' sets of grades for pre-COVID and COVID teaching terms enables us to exploit the natural experiment nature of the pandemic emergency and investigate the effects of both measures - the one-off effect of the ND policy and the change in assessment mode still in place. The findings concerning changes in modes of assessment will help inform future assessment policy.

Methodology

The data available can be explored either as a pooled set of repeated cross-sections or in a longitudinal format. We started with the repeated cross-section set and used a Difference-in-Differences (DiD) treatment effect approach to examine the average differences in grades between the cohort affected by COVID and the two cohorts graduating in the two years before. We explored further the differential impact of interventions by looking at different points of the cohorts' grade distributions, using quartile DiD estimations. The longitudinal data set was used to estimate a panel fixed effect

model as a robustness check to explore if the within-individual effect rather than the within-cohort effect would confirm the findings of the original DiD estimates.

The DiD approach

All three years share similar conditions regarding fee levels and admission from an institutional policy perspective. To explore whether COVID interventions used in T2 2019/20 produced any treatment effects on award gaps, we apply a DiD treatment approach and use the 2019/20 cohort as the treated group, and their T2 performance as the treatment effect. We then compare their T1 and T2 outcomes with those of the control cohorts (2018/19) not affected by COVID interventions. We do this comparison for L6 students as well as for L5 students.

The repeated cross-section analysis exploits the difference in the cohorts of the students and requires some additional assumptions around the comparability of the overall student cohorts. Therefore, to check the robustness of our results, we also produce DiD estimates of the two cohorts not affected by COVID measures, comparing T1 and T2 outcomes of the 2017/18 cohort with those of the 2018/19 cohort.

The DiD models were estimated using robust standard errors. We took account of covariates and used the Kernel matching option to line up comparison individuals according to sufficient observable factors and to remove systematic differences in the evaluation outcome between treated and non-treated groups. The balancing properties of the DiD models were all tested.

We also apply the same DiD approach to investigate treatment effects at different quartiles of the distributions of cohorts' performances.

Longitudinal approach

While the DiD approach allows us to compare the performances of different cohorts of students at the same level of progression, the longitudinal analysis enables us to control for individual differences, as the same person's performance is compared through time. International students from China represent around 70% of all international students in each cohort (see Table 1), they have an unusually balanced gender distribution and almost all of them (80%) join the university in the second year of the degree study. We focus therefore only on their T1 and T2 performances at L5 and L6 so that and the same cohort of students in the panel are sampled four times.

To investigate how COVID interventions could have affected awarding gaps, we apply the longitudinal fixed effect model to two cohorts of overseas students from China, those who graduated the year before COVID and those who graduated in the COVID year. The presence of modules' repeaters, students in placement, or studying abroad creates attrition in data that renders the panel unbalanced. The evolution of the cohort is set out in table 2.

Table 2. Evolution of the Longitudinal Cohort. Evidence of an unbalanced cohort (International students from China).

	HESA Business and Management subjects							
	Pre-COVID cohort: unbalanced panel (students sampled 4 times)				COVID cohort: unbalanced panel (students sampled 4 times)			
	L5 in 2017/18		L6 in 2018/19		L5 in 2018/19		L6 in 2019/20	
Overseas from China	T1	T2	T1	T2	T1	T2	T1	T2(*)
Female	167	190	158	149	159	170	155	144
Male	173	205	155	149	159	167	151	154
Total	340	395	313	298	318	337	316	298
% of overseas	67.6%	69.8%	68.9%	68.7%	67.2%	64.9%	66.8%	66.4%

(*) = COVID term

To exploit the information on students' experience over the two years of the degree, we will assume some degree of comparability between levels of the degree and across cohorts. We will discuss these conditions in the results sections.

Results

We present three main sets of results: the repeated cross-sectional DiD, the repeated cross-sectional DiD quantile regression, and the fixed effect panel regression.

Repeated cross-sectional DiD

For each of the DiD estimates presented below we estimated a full DiD model. An example of the repeated cross-sectional DiD for L6 is presented in Appendix (Table A1). Table 3 below presents the DiD estimates for the different sets of grades using the repeated cross-sectional structure.

The first two rows of Table 3 highlight the effects of the ND policy and the switch to online assessments on UK domestic and international students. The changes in assessments did not produce any statistically significant impact on the UK domestic students' grades, either in T2 or in the resit stage. However, their grades were affected by applying the ND policy, which produced similar statistically

Table 3. Summary DiD Results 2018/19 and 2019/20.

	Comparing differences in T2 with differences in T1 across cohorts					
	L6 Years 2018/19 and 2019/20			L5 Years: 2018/19 and 2019/20		
Average effects	First attempt January/May	Final after August Resits	Final After ND	First attempt January/May	Final after August Resits	Final After ND
All students Effects on: Average Class of grade	1.529* (0.889) ↑; same (>60)	1.549* (1.88) ↑; same (>60)	3.964*** (0.830) ↑; same (>60)	2.908*** (0.901) ↑; same (>60)	2.711*** (0.849) ↑; same (>60)	4.714*** (0.853) ↑; same (>60)
UK domiciled Effects on: Average Class of grade	0.073 (1.366) no effect same (>60)	−0.237 (1.245) no effect same (>60)	2.744** (1.247) ↑: same (>60)	0.371 (1.362) no effect same (>60)	0.253 (1.302) no effect same (>60)	2.684** (1.306) ↑: same (>60)
International all Effects on: Average Class of grade	2.333** (1.035) ↑; same (<60)	2.531*** (0.952) ↑: same (<60)	4.595*** (0.955) ↑; **change (>60)**	4.980*** (1.100) ↑; same (<60)	4.835*** (1.033) ↑: same (<60)	6.535*** (1.038) ↑; **change (>60)**
Inter. (excl. China) Effects on: Average Class of grade	1.114 (2.059) no effect same (<60)	1.938 (1.803) no effect same (<60)	4.299** (1.807) ↑; **change (>60)**	4.946*** (1.877) ↑; same (<60)	4.876*** (1.751) ↑; **change (>60)**	6.783*** (1.757) ↑; **change (>60)**
Inter. (only China) Effects on: Average Class of grade	3.038*** (1.140) ↑; same (<60)	2.974*** (1.081) ↑; same (<60)	4.890*** (1.085) ↑; **change (>60)**	5.097*** (1.360) ↑; same (<60)	4.908*** (1.280) ↑; same (<60)	6.388*** (1.288) ↑; same (<60)
Female & Int. China Effects on: Average Class of grade	3.326** (1.399) ↑; **change (>60)**	3.177** (1.319) ↑; **change (>60)**	5.448*** (1.327) ↑; **change (>60)**	4.690*** (1.745) ↑; same (<60)	4.532*** (1.644) ↑; same (<60)	6.178*** (1.657) ↑; **change (>60)**
Male & Int. China Effects on: Average Class of grade	2.767 (1.743) no effect same (<60)	2.780* (1.663) ↑; same (<60)	4.363*** (1.662) ↑; same (<60)	5.479*** (2.062) ↑; same (<60)	5.265*** (1.938) ↑; same (<60)	6.566*** (1.947) ↑; same (<60)

Standard Error in parentheses. Inference: *** = 1%; ** = 5%; * = 10*. (Bold indicates a change of grade classification: averages change from 50s (2:2 degree classification) into 60s (2:1 degree classification))

significant gains at L5 and L6. However, these gains did not change the grade classification, and the overall group average remained in the 60s interval.

The international students gained from the COVID interventions; as a result, their average final grade crossed the 60% threshold, increasing their probability of being awarded a higher degree classification. In addition, the grade increase at L5 was substantially bigger than at L6. The substantial proportion of international students entering the UK HE system directly at L5 may partly explain this result. The COVID measures have helped these Direct Entry (DE) students, mitigating their transition into the system.

The third and fourth rows of Table 3 show the differential impact of the COVID interventions between the international students from China and the rest of the international students. The results show that, at L6, this latter group of students behaves similarly to the UK domestic group. For them, only the ND intervention produced a significant effect on grades, big enough to induce a change in class grades. However, at L5, the same category of students shows a different story benefitting from both COVID interventions (and crossing the 60% threshold in their final after resit grades). By contrast, the international students from China benefitted from both COVID interventions- the change in assessments in T2 and in resits, and the ND policy - and at any level of progression. However, only those at L6 crossed the 60% threshold with the application of the ND policy.

Given that the international Chinese students represent a significant proportion of overseas students, we focus on this group and report their results in the lower part of Table 3, splitting them by gender. Our findings suggest that female students from China benefitted from both COVID interventions, while male did not.

To summarise, the combination of changes in assessments, and ND policy has increased the average of all groups of students and enabled the international students, except the male Chinese group, to reach a GPA group average that crossed the threshold of the 60% (an upper second classification), and catch up with the domestic students. However, for most groups, except Chinese females, the gains are mainly to be attributed to the ND policy itself rather than to the change in assessments or the possibility of resits. Overall, our findings present a variety of heterogeneous responses and effects of COVID policies on students' gender, ethnicity, and country of domicile.

To check for robustness, we applied the same model to compare the previous two pre-COVID cohorts, who were not affected by the emergency interventions (i.e, 2017/18 and 2018/19).

These results show that when COVID interventions were not in place, there were no differential effects between T1 and T2 grades across different cohorts at L6. The absence of any statistically significant difference between 2018/19 and 2017/18 performances indicates that in the absence of intervention that any gap in performances was constant between term 1 and term 2 (not closing or increasing) for pre-COVID cohorts at L6 (this finding is summarised in the first two columns of Table 4). As a result, the presence of statistically significant differences between the 2018/19 and 2019/20 cohorts capture the treatment effect of COVID policies.

Level 5 shows some statistically significant differences between the T2 and T1 also in the pre-COVID periods. As explained earlier, these are related to the effect of the direct entry (DE) cohorts and their T1 impact. Even taking account of this DE effect in the previous results, the female international students from China have a different response to the COVID changes than other groups of international and home students, benefitting from COVID measures.

The repeated cross-sectional DiD quantile regression

The remaining results seek to understand better these differential responses by looking at quartile distribution instead of averages. Table 5 presents the DiD quantile regression results for L6 and L5 international students from China, by gender.

We omit to present the pre-COVID results at L6 because, like before, we did not find any significant differences in grades. Although this is not a formal test of the common trend, which is not

Table 4. Summary DiD results 2017/2018 and 2018/19.

| | Comparing differences in T2 with differences in T1 across cohorts | | | |
| | L6 Years 2017/18 and 2018/19 | | L5 Years: 2017/18 and 2018/19 | |
Average effects	First attempt January/May	Final after August Resits	First attempt January/May	Final after August Resits
All students Effects on: Average Class of grade	−0.898	−1.423	−1.729*	−1.706*
	(0.918)	(0.887)	(0.899)	(0.880)
	no effect	no effect	↓;	↓;
	same (> 60)	same (> 60)	same (<60)	same (<60)
UK domiciled Effects on: Average Class of grade	−1.314	−2.201	1.423	1.329
	(1.938)	(1.718)	(1.433)	(1.403)
	no effect	no effect	no effect	no effect
	same (>60)	same (>60)	same (>60)	same (>60)
International Effects on: Average Class of grade	−0.503	−0.845	−3.393***	−3.277***
	(0.935)	(0.940)	(1.032)	(1.012)
	no effect	no effect	↓;	↓;
	same (<60)	same (<60)	same (<60)	same (<60)
Intern. (excl. China) Effects on: Average Class of grade	−1.176	−1.433	−3.033	−3.205*
	(1.888)	(1.850)	(1.930)	(1.863)
	no effect	no effect	no effect	↓;
	same (<60)	same (<60)	same (<60)	same (<60)
Intern. (only China) Effects on: Average Class of grade	−0.350	−0.704	−3.355***	−3.125***
	(1.031)	(1.053)	(1.185)	(1.175)
	no effect	no effect	↓;	↓;
	same (<60)	same (<60)	same (<60)	same (<60)
Female & Inter. China Effects on: Average Class of grade	−1.237	−1.500	−4.285**	−3.909**
	(1.345)	(1.389)	(1.678)	(1.613)
	no effect	no effect	↓;	↓;
	same (<60)	same (<60)	same (<60)	same (<60)
Male & Inter, China Effects on: Average Class of grade	0.644	0.197	−2.546	−2.479
	(1.533)	(1.551)	(1.641)	(1.670)
	no effect	no effect	no effect	no effect
	same (<60)	same (<60)	same (<60)	same (<60)

Standard Error in parentheses. Inference: *** = 1%; ** = 5%; * = 10*. Yellow indicates a negative effect.

Table 5. DiD Quartile regressions. International Students from China ONLY. COVID cohorts.

| | Comparing differences in T2 with differences in T1 across cohorts | | | | | |
| | L6 Years 2018/19 and 2019/20 | | | L5 Years: 2018/19 and 2019/20 | | |
All	First attempt January/May	Final August Resits	Final After ND	First attempt January/May	Final August Resits	Final After ND
25th	**3.667*****	**4.750*****	**6.625*****	**7.167*****	**7.083*****	**8.833*****
	(1.368) (CL)	**(1.369) (CL)**	**(1.373) (CL)**	**(1.225) (CL)**	**(2.238) (CL)**	**(1.222) (CL)**
50th	**4.750*****	**4.750*****	**6.250*****	6.583***	**3.333***	**7.500*****
	(1.052) (CU)	**(1.315) (CU)**	**(1.312) (CU)**	(1.237)	**(1.852) (CU)**	**(1.239) (CU)**
75th	3.250**	3.250**	4.500 ***	4.167***	1.833	5.500***
	(1.357)	(1.360)	(1.135)	(1.525)	(1.915)	(1.454)
Female						
25th	4.250**	4.500***	6.750***	**6.667*****	**7.333*****	**9.333*****
	(2.102)	(1.623)	(1.631)	**(2.408) (CL)**	**(2.329) (CL)**	**(1.851) (CL)**
50th	3.500	**3.750***	**6.500*****	**5.000*****	**5.000*****	**6.250*****
	(2.157)	**(2.166) (CU)**	**(1.891) (CU)**	**(1.716) (CU)**	**(1.805) (CU)**	**(1.810) (CU)**
75th	3.250*	3.000	5.000*	4.333**	4.667***	6.500***
	(1.874)	(1.874)	(2.570)	(1.867)	(1.791)	(1.557)
Male						
25th	2.000	**4.750****	**5.750****	**10.000*****	**7.667*****	**8.417*****
	(2.812)	**(2.326) (CL)**	**(2.254) (CL)**	**(1.810) (CL)**	**(1.635) (CL)**	**(1.797) (CL)**
50th	4.750***	5.500***	**6.750*****	7.250***	7.500***	7.917***
	(1.617)	(1.616)	**(1.344) (CU)**	(2.549)	(2.002)	(1.814)
75th	**4.000***	**4.500****	**6.000*****	3.333	3.000	5.250***
	(2.102) (CU)	**(2.109) (CU)**	**(1.861) (CU)**	(2.184)	(1.959)	(1.887)

CL = change of grade classification, from 40s into 50s; **CU** = change of grade classification, from 50s into 60s

required in the Kernel type DiD, it does suggest in the absence of any intervention the gap in performance between the cohorts is constant from term 1 to term 2. From Table 5, we can observe three things. Firstly, the change to online assessments enabled students in the lower quartile to substantially increase their grades, moving from the 40s interval into the 50s range. This event is reported in Table 5 using the bold acronym CL, which denotes an average grade change from a third into a lower second class. Finally, the application of ND reinforced this move. These results hold for both L5 and L6 international Chinese students, but with some differences between genders (mainly because female students in the lower quartile were already in the 50s range).

Secondly, finalist students (L6) in the second quartile also benefitted from the online assessment change, and their average grade moved from the 50s into the 60s range. The bold acronym CU denotes an average grade change from a lower to an upper-second grade classification, a move reinforced by the ND policy. Interestingly, for L5 students (most of whom are DE), the grade crossing classification occurred only after the resit exam period and after the ND policy. Gender differences are also present in this second quartile. For example, the average female grade crossed the 60s grade boundaries earlier than the males because female students start with a higher average than males.

Thirdly, the COVID policies did not create additional first-class degrees for the upper quartile. But both assessment changes and ND policy helped L6 male students from China move up to an average of 60, joining the female students already positioned in that grade range.

The DiD quantile regression across three different sets of grades (first attempt, final after resits, and final after ND policy) reveals the importance of stepping away from the average effects for all students and considering the distributional impact of a change in policies across student demographics and levels of student attainment.

The fixed effect panel regression

The final set of results is based on comparing performances of the same student as they progress from L5 into L6 instead of comparing different cohorts. We consider each student's T1 and T2 grades at L5 and L6 and estimate a fixed effect model. We then compare the results of the students who graduated under the first COVID lockdown with those who graduated the year before, focusing only on the international students from China. The results of the estimates are reported in Table 6 and Appendix (Table A3).

Looking at the students in the pre-COVID cohort and tracking them through L5 and L6 stages, we noticed an improvement from T1 to T2 for L5 students (most of them DE), as they adjust to university academic life, but a noticeable drop at L6 in their T1 and T2 grades, perhaps due to more challenging modules.

In the COVID cohort, we can also find a significant drop in grades in the first term of L6 (T1). This result is relevant because the ND was based on taking that term as a minimum grade. However, their T2 average of L6 does not remain low, but COVID measures lift it, with some differential effects. For example, students who did not invoke the ND policy showed a substantial increase in grades relative to those who used it. This fact would suggest that the change in the assessment modes was very

Table 6. Fixed effect model. International students from China.

Average grades	COVID cohort			Pre–COVID cohort	
	First attempt	Final after resits	Final after ND	First attempt	Final after resits
L5-T1	56.88	57.55	57.55	54.31	54.34
L5-T2	56.71	57.29	57.29	57.44	58.52
L6-T1	54.24	54.99	55.09	54.31	54.34
L6-T2 (average all students)	57.92	58.69	60.66	54.31	54.34
L6-T2 students without ND	64.64	64.47	64.38		
L6-T2 student with ND	55.16	56.24	59.09		

effective for some students (those who did not need the ND policy) but less effective or ineffective for those who needed to invoke the ND policy.

The combination of the two COVID measures has thus helped move students closer to or above the 60% threshold, and the ND (by using T1 grades as a minimum grade) has offered a proper lower bound to the students it was meant to help. We believe that the ND policy was judiciously designed. It did not inflate grades for subjects in the HESA Business and Management categorisation (it would have been more generous had it included the average of level 5 grades). In addition, it helped students who invoked it keep the trajectory of T1 grades, and it remained neutral to those who did not use it, whose grades improved due to changes in assessments.

Conclusions

The study highlighted the contribution of both changes to degree awarding policies adopted at the time of the onset of the COVID-19 pandemic and the changes to modes of assessment that resulted to accommodate education continuity. It offers a rich insight into the detailed work that institutions can undertake to further understand the effect of assessment change on awarding gaps in the student body. Such changes combined with the briefings offered by faculty related to revised assessment modes may have served to uncover the previously implicit aspects of assessment for international students supporting their outcomes and leading to a reduction in the awarding gap. In addition, it offers a commentary on the effect of institutional policy adopted at the time and may inform future construction of force majeure regulations should they be required in the future.

Limitations include the generalisability from one institutional context and academic framework. Future research could undertake a comparative analysis of institutional force majeure policies and their effects on the awarding gaps at the time of the pandemic. This may help to inform sector understanding of the relationship between assessment type, institutional policy, and grade inflation.

Acknowledgments

Disclosure statement

ORCID

M. Gabriella Cagliesi 🄳 http://orcid.org/0000-0002-5208-730X
Denise Hawkes 🄳 http://orcid.org/0000-0002-7305-3846
Susan Smith 🄳 http://orcid.org/0000-0003-0411-9327

References

Barnett, Ronald, and Kelly Coate. 2005. *Engaging the Curriculum in Higher Education*. Maidenhead: Open University Press.
Boud, David. 2013. *Enhancing Learning Through Self-Assessment*. Sydney: Routledge.
Chan, C. 2022. "A review of the changes in higher education assessment and grading policy during covid-19." *Assessment and Evaluation in Higher Education*. doi:10.1080/02602938.2022.2140780.
Codiroli Mcmaster, Natasha. 2021. *Ethnicity Awarding Gaps in UK Higher Education in 2019/20*. York: Advance HE. https://www.advance-he.ac.uk/knowledge-hub/ethnicity-awarding-gaps-uk-higher-education-201920.
Connor, Helen, Ivana La Valle, Nii Djan Tackey, and Sarah Perryman 1996. *Ethnic Minority Graduates: Differences by Degrees. Report 309*. Brighton: Institute for Employment Studies.

Cotton, Debby, John Bailey, and Matthew Tosdevin. 2020. "Higher Education and the Climate Emergency: Exploring the Hidden Curriculum of the Campus." In *The Hidden Curriculum of Higher Education*, edited by Tim Hinchcliffe, 29–40. York: Advance HE.

Cotton, D. R. E., M. Joyner, R. George, and P. A. Cotton. 2016. "Understanding the Gender and Ethnicity Attainment Gap in UK Higher Education." *Innovations in Education and Teaching International* 53 (5): 475–86. doi:10.1080/14703297.2015.1013145.

Cotton, Debby, Jennie Winter, and Ian Bailey. 2013. "Researching the Hidden Curriculum: Intentional and Unintended Messages." *Journal of Geography in Higher Education* 37 (2): 192–203. doi:10.1080/03098265.2012.733684.

Crawford, Ian, and Zhiqi Wang. 2015. "The Impact of Individual Factors on the Academic Attainment of Chinese and UK Students in Higher Education." *Studies in Higher Education* 40 (5): 902–20. doi:10.1080/03075079.2013.851182.

Fielding, A., C. M. J. Charlton, D. Z. Kounali, and G. B. Leckie. 2008. Degree Attainment, Ethnicity and Gender: Interactions and the Modification of Effects-A Quantitative Analysis'.

Gayton, Angela M. 2020. "Exploring the Widening Participation-Internationalisation Nexus: Evidence from Current Theory and Practice." *Journal of Further and Higher Education* 44 (9): 1275–88. doi:10.1080/0309877X.2019.1678014.

HEFCE. 2015. Causes of Differences in Student Outcomes', 141.

HESA. 2021. 'Where Students Come from and Go to Study (2019/20) (SB258)'. 27 January 2021. https://www.hesa.ac.uk/news/27-01-2021/sb258-higher-education-student-statistics/location.

Iannelli, Cristina, and Jun Huang. 2014. "Trends in Participation and Attainment of Chinese Students in UK Higher Education." *Studies in Higher Education* 39 (5): 805–22. doi:10.1080/03075079.2012.754863.

Jones, Steven. 2018. "Expectation vs Experience: Might Transition Gaps Predict Undergraduate Students' Outcome Gaps?" *Journal of Further and Higher Education* 42 (7): 908–21. doi:10.1080/0309877X.2017.1323195.

Jones, Steven, Maria Pampaka, Daniel Swain, and Julian Skyrme. 2017. "Contextualising Degree-Level Achievement: An Exploration of Interactions Between Gender, Ethnicity, Socio-Economic Status and School Type at One Large UK University." *Research in Post-Compulsory Education* 22 (4): 455–76. doi:10.1080/13596748.2017.1381287.

Koutsouris, George, Anna Mountford-Zimdars, and Kristi Dingwall. 2021. "The 'Ideal' Higher Education Student: Understanding the Hidden Curriculum to Enable Institutional Change." *Research in Post-Compulsory Education* 26 (2): 131–47. doi:10.1080/13596748.2021.1909921.

Palacios, Angélica M. G., and Rafael D. Alvarez. 2016. "An Analysis of Nonfirst-Generation Community College Men of Color: Comparing GPA, Noncognitive, and Campus Ethos Differences Across Race." *Community College Journal of Research and Practice* 40 (3): 180–7. doi:10.1080/10668926.2015.1112319.

Portelli, John P. 1993. "Exposing the Hidden Curriculum." *Journal of Curriculum Studies* 25 (4): 343–58. doi:10.1080/0022027930250404.

Richardson, John T.E. 2008. "The Attainment of Ethnic Minority Students in UK Higher Education." *Studies in Higher Education* 33 (1): 33–48. doi:10.1080/03075070701794783.

Richardson, John T.E. 2015. "The Under-Attainment of Ethnic Minority Students in UK Higher Education: What We Know and What We Don't Know." *Journal of Further and Higher Education* 39 (2): 278–91. doi:10.1080/0309877X.2013.858680.

Richardson, John T. E., Jenna Mittelmeier, and Rienties. Bart. 2020. "The Role of Gender, Social Class and Ethnicity in Participation and Academic Attainment in UK Higher Education: An Update." *Oxford Review of Education* 46 (3): 346–62. doi:10.1080/03054985.2019.1702012.

Sambell, K., and L. Mcdowell. 1998. "The Construction of the Hidden Curriculum: messages and meanings in the assessment of student learning." *Assessment and Evaluation in Higher Education* 23 (4): 391–402. doi:10.1080/0260293980230406.

Semper, Orón, José Víctor, and Maribel Blasco. 2018. "Revealing the Hidden Curriculum in Higher Education." *Studies in Philosophy and Education* 37 (5): 481–98. doi:10.1007/s11217-018-9608-5.

Smith, Emma. 2016. "Can Higher Education Compensate for Society? Modelling the Determinants of Academic Success at University." *British Journal of Sociology of Education* 37 (7): 970–92. doi:10.1080/01425692.2014.987728.

Wicking, Paul. 2020. "Formative Assessment of Students from a Confucian Heritage Culture: Insights from Japan." *Assessment & Evaluation in Higher Education* 45 (2): 180–92. doi:10.1080/02602938.2019.1616672.

Appendix

Table A1. DiD Estimates. Comparing COVID and Pre-COVID cohorts of L6 students.

HESA Business and Management subjects				
Control group: in T1 = Co1; in T2 = Co2 Treated group: in T1 = Tr1; in T2 = Tr2	Co1 = 647; Co2 = 576 total Co = 1223 Tr1 = 671; Tr2 = 636 total Tr = 1307 1318 1212 N.obs = 2530		Co1 = 671; Co2 = 636 total Co = 1307 Tr1 = 765; Tr2 = 726 total Tr = 1491 1436 1362 N.obs = 2798	
	Academic Years	Final Grade	Academic Years	Final Grade after ND
Before treatment (T1)				
	$Co1_{2017.18}$	59.146	$Co1_{2018.19}$	60.317
	$Tr1_{2018.19}$	59.302	$Tr1_{2019.20}$	59.874
Diff. before treatment.	$Tr1_{2018.19} -$	0.156	$Tr1_{2019.20} -$	−0.444
	$Co1_{2017.18} = A$	(0.629)	$Co1_{2018.19} = A$	(0.591)
After treatment (T2)				
	$Co2_{2017.18}$	61.646	$Co2_{2018.19}$	61.231
	$Tr2_{2018.19}$	60.379	$Tr2_{2019.20}$	64.751 (COVID)
Diff. after treatment	$Tr2_{2018.19} -$	−1.267**	$Tr2_{2019.20} -$	3.52***
	$Co2_{2017.18} = B$	(0.626)	$Co2_{2018.19} = B$	(0.583)
Diff-in-Diffs (treatment effect)	B − A	−1.423	B − A	**3.964***
		(0.887)		(0.830)

Kernel propensity score matching; repeated cross section; robust Std. Error
R-square: 0.01; Inference: *** < 0.01; ** < 0.05; * < 0.1
Covariate: Race(White); UK; sex (Female)
Balancing two sample t-test of the difference in the means of the covariates between the control and treated groups in T1.

	2017-2018; N = 1318				2018-2019; N = 1436			
T1	Mean Co1	Mean Tr1	Diff.	Pr ($|T|>|t|$)	Mean Co1	Mean Tr1	Diff.	Pr ($|T|>|t|$)
Ave. Fin. Grade	59.146	59.302	0.156	0.803	60.317	59.874	−0.444	0.452
Race (White)	0.258	0.267	0.01	0.682	0.365	0.369	0.004	0.883
UK (dom.)	0.290	0.300	0.01	0.711	0.374	0.379	0.004	0.865
Sex (Female)	0.462	0.462	0	0.996	0.441	0.410	−0.030	0.246

Results: we cannot reject the null hypothesis that the covariates are balanced.

Table A2. DiD Estimates. Comparing COVID and Pre-COVID cohorts of L6 International students ONLY.

HESA Business and Management subjects.				
Co1 = 471; Co2 = 406 total Co = 877 Tr1 = 454; Tr2 = 434 total Tr = 888 925 840 N.obs. = 1765			Co1 = 454; Co2 = 434 total Co = 888 Tr1 = 473; Tr2 = 449 total Tr = 992 927 883 N.obs. = 1810		
	Year	Final Grade	Year	Final Grade after ND
Before treatment (T1)				
	$Co1_{2017.18}$	57.311	$Co1_{2018.19}$	56.753
	$Tr1_{2018.19}$	56.669	$Tr1_{2019.20}$	56.436
Diff before treat.	$Tr1_{2018.19} -$	−0.642	$Tr1_{2019.20} -$	−0.316
	$Co1_{2017.18} = A$	(0.660)	$Co1_{2018.19} = A$	(0.663)
After treatment (T2)				
	$Co2_{2017.18}$	58.925	$Co2_{2018.19}$	57.725
	$Tr2_{2018.19}$	57.438	$Tr2_{2019.20}$	62.004
Diff after treat.	$Tr2_{2018.19} -$	−1.488**	$Tr2_{2019.20} -$	4.280***
	$Co2_{2017.18} = B$	(0.670)	$Co2_{2018.19} = B$	(0.688)
Diff-in-Diff	B − A	−0.845	B − A	**4.595***
		(0.940)		(0.955)

Kernel propensity score matching; repeated cross section; robust Std. Error
R-square: 0.01; Inference: *** < 0.01; ** < 0.05; * < 0.1
Covariate: Race(White); sex (Female)
Balancing two sample t-test of the difference in the means of the covariates between the control and treated groups in T1.

	Term 1 2017-2018; N = 925				Term 1 2018-2019; N = 927			
	Mean C1	Mean T1	Diff.	Pr ($|T|>|t|$)	Mean C1	Mean T1	Diff.	Pr ($|T|>|t|$)
Ave. Fin. Grade	57.311	56.669	−0.642	0.331	56.752	56.436	−0.316	0.634
Race (White)	0.101	0.103	0.002	0.926	0.120	0.123	0.003	0.885
Sex (Female)	0.453	0.480	0.027	0.409	0.465	0.438	−0.027	0.415

Results: we cannot reject the null hypothesis that the covariates are balanced.

Table A3. Fixed Effect Model.

Average grades	COVID cohort			Non–COVID cohort	
	First attempt	Final after resit	Final after ND	First attempt	Final after resits
Constant (L6-T1 is the ref. categ.)	54.236***	54.988***	55.087***	54.305***	54.336***
	(0.311)	(0.282)	(0.262)	(0.260)	(0.264)
L5-T1	2.646***	2.555***	2.446***	0.700	0.586
	(0.523)	(0.469)	(0.467)	(0.519)	(0.506)
L5-T2	2.473***	2.306***	2.200***	3.178***	3.184***
	(0.506)	(0.475)	(0.472)	(0.433)	(0.451)
L6-T2 (average all students)	3.687***	3.652***	5.567***	0.619	0.604
	(0.608)	(0.550)	(0.466)	(0.428)	(0.416)
L6-T2 students without ND	10.400***	9.483***	9.350***		
	(0.952)	(0.911)	(0.904)		
L6-T2 student with ND	−9.481***	−8.235***	−5.343***		
	(1.111)	(1.043)	(1.013)		
Observations	1,269	1,269	1,269	1,346	1,346
Number of ID (clusters)	348	348	348	409	409
R-squared **within**	0.046	0.052	0.118	0.056	0.051
R-squared **within (separating ND)**	0.129	0.126	0.150		

Robust standard errors in parentheses; statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$