

# How do assessment patterns affect course marks at LSE?

Ellen Austin, ESE Analyst, December 2021

## Overview

Assessment at LSE is an integral part of the learning experience, as well as a way of measuring students' attainment on a course or programme. It should be developmental, helping students to deepen their understanding of the material and confidently demonstrate the learning that they experience across different courses as they progress through their programme.

This analysis shows, at a high level, that different types of assessment can influence course outcomes for students, and that certain groups of students may experience systematically different outcomes from their peers with a particular form of assessment.

There are some clear patterns in the assessment outcomes for students. While some of the effects involved are relatively small, over all the courses taken in an academic year they could add up to create a big difference in outcomes for students. This is particularly important given the diversity of routes that students take through their programme at LSE – over 80% of third year undergraduates take a unique combination of courses.

The analysis presented here is indicative and cannot make claims about causality. With this observational data we have no comparison group, and the findings can only show associations. Nonetheless, the findings provide insight into factors that may differentiate student experience of assessment and student assessment outcomes.

Data about assessments at LSE is held across multiple systems, and while we can reconstruct high-level likely assessment patterns for most LSE students in recent (pre-COVID) years, there are some difficult edge cases. More detail on assessment data can be found in Appendix A.

## Key recommendations

- In a subset of undergraduate and taught postgraduate quantitative courses,<sup>1</sup> pilot alternative methods of assessment and review the impact for students from racialised ethnic backgrounds and/or with mental health difficulties.
- Explore, with qualitative and quantitative research, why some students, especially in the taught postgraduate cohort, appear to benefit from higher assessment counts on courses.
- With LSE Life, develop and run targeted exam preparation for students from non-UK European and North American countries, who experience worse outcomes when assessed via in-person exams. Evaluate this intervention, ideally through a randomised treatment group or, if that is not possible/ethical, through a quasi-experimental study.
- Undertake further developmental work to the underlying data so that:
  - Additional assessment data about e.g. formative assessments, assessment types, wordcounts can be added into the dataset to understand more about assessment patterns.
  - Data about optionality on programmes can be integrated with assessment data to understand how programme design may affect assessment pathways for students.

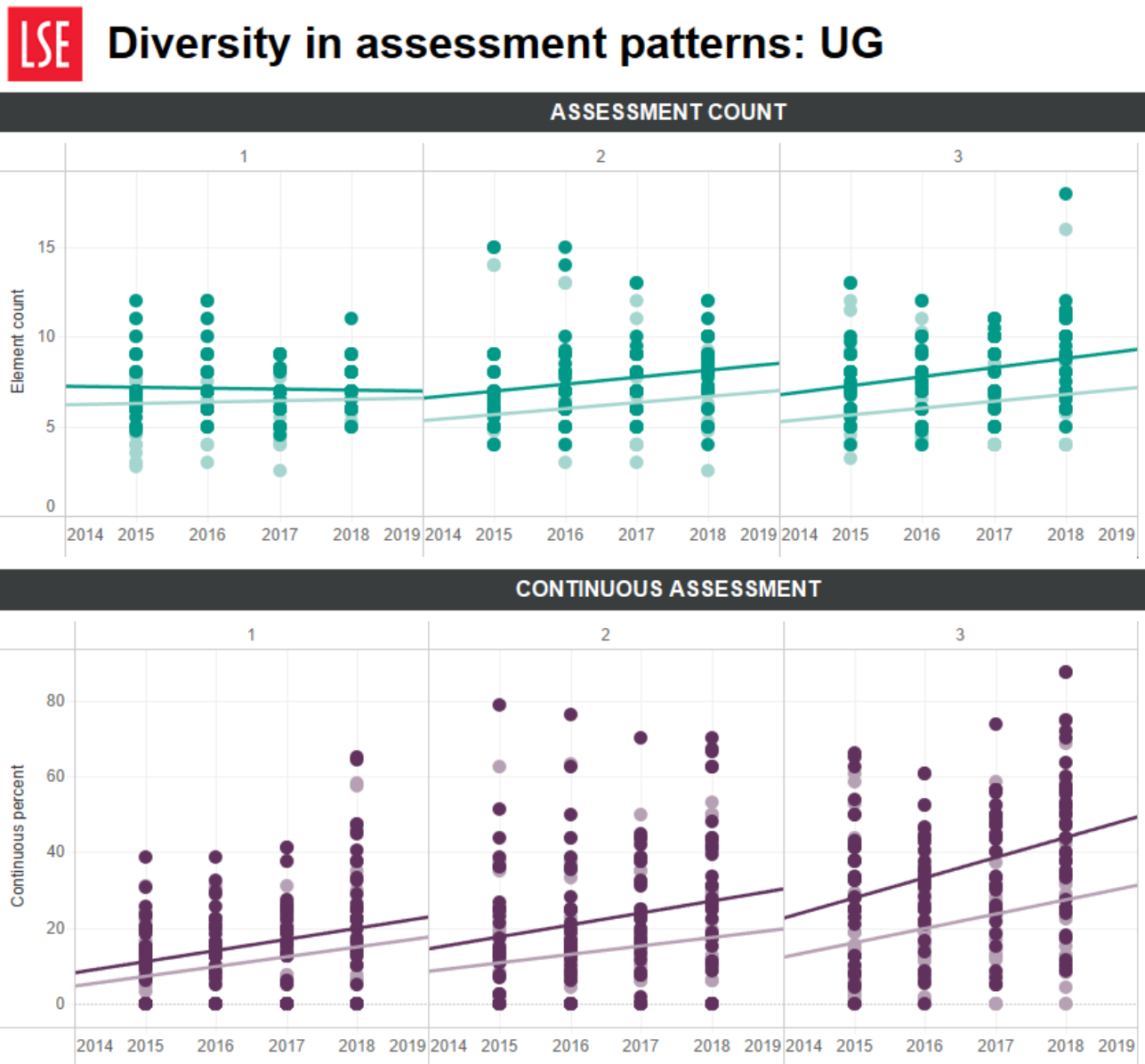
---

<sup>1</sup> Courses were split into quantitative and qualitative based on their department; Annex C provides a list of quantitative and qualitative departments.

## Changes to assessment on programmes over time

Looking at our best possible reconstruction of assessments taken by students over time, some trends emerge. We have not considered data from 2019/20 or 2020/21 given that student assessments in those years were so atypical.

Figure 1: Undergraduate assessment on programmes over time



At undergraduate level students are taking more assessment elements during a year, and the proportion of the year's credit assessed by coursework methods is also increasing. This is shown in the upward slope of the lines in Figure 1 – green lines representing the in-year count of assessment elements and purple representing the proportion of credit assessed via coursework assessment.

We can also see that, when it comes to assessment, the diversity of student experience on programmes is increasing over time, especially in Year 2 and Year 3 of study. This can be seen in the growing gap between the fitted lines for the lower quartile (light green/purple) and the upper quartile (dark green/purple) in Figure 1. As the gap gets wider, the experience of students on a single programme tends to become more diverse.

Figure 2: Taught postgraduate assessment on programmes over time



## Diversity in assessment patterns: PGT

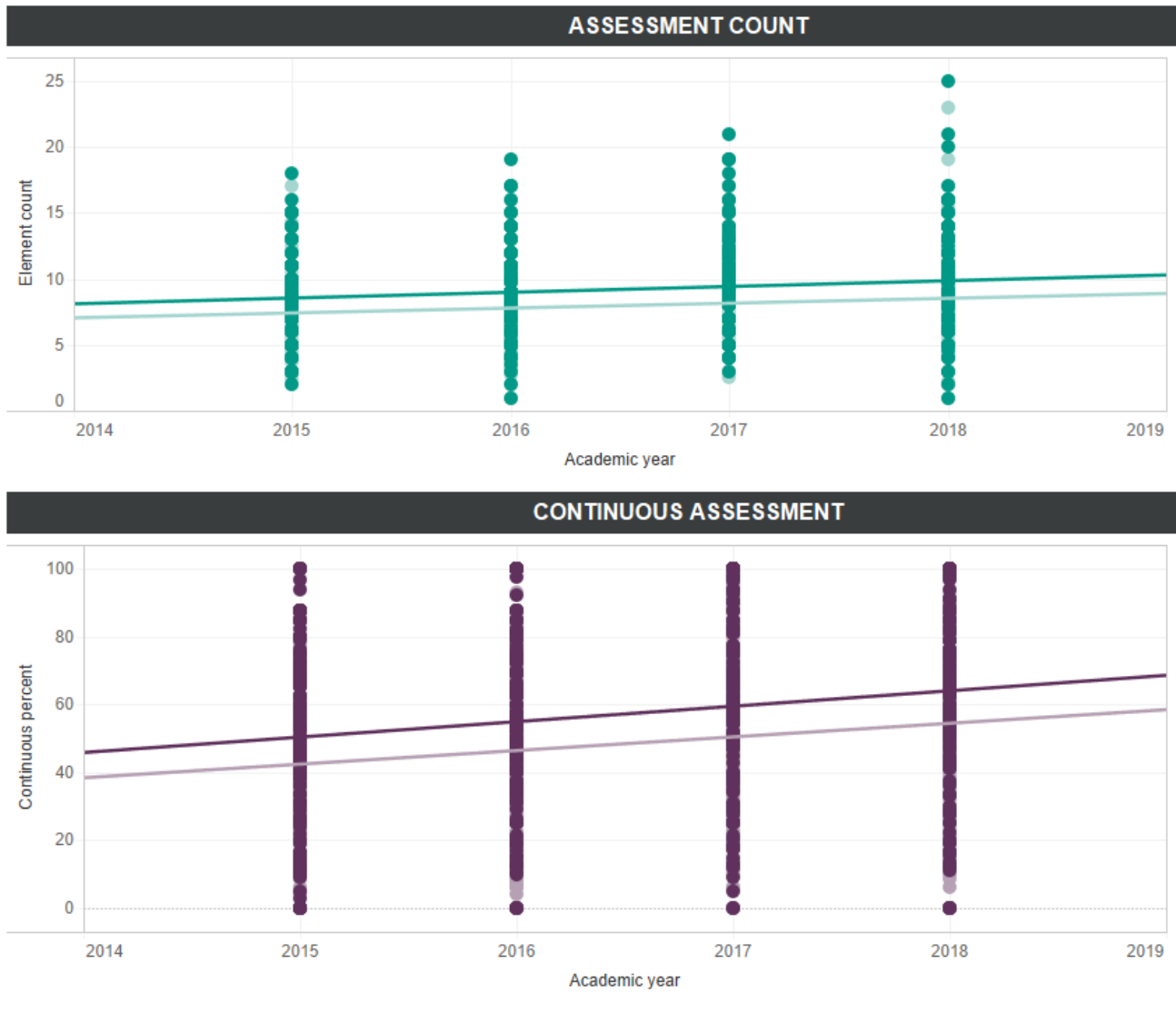


Figure 2 shows the same data for taught postgraduate programmes. As with undergraduates, the element count and percentage of coursework assessment seems to be increasing over time, but student experience on the programmes appears to remain more consistent – there is not the same growing gap between the dark and light lines on the graph.

These overall trends may mask other patterns within the data where, for example, the level of optionality on a programme affects assessment patterns (which seems likely).

## Effect of assessment type on course results

We ran four models, each looking at a different group of courses, to understand possible relationships between assessment patterns and student results. Figure 3 shows the splits and counts of students considered in each grouping.

**Figure 3: Course results included in analysis by programme level and disciplinary group**

	FD		PGT	
	Qualitative	Quantitative	Qualitative	Quantitative
2015	8,078	11,426	17,239	10,761
2016	8,688	11,419	19,359	12,614
2017	8,682	11,860	19,459	13,640
2018	9,133	13,437	19,042	13,923

Figure 4 shows the relationship between the assessment variables included in the model and student marks on their course. The four assessment variables were as follows:

- Percentage of course assessed by in-person exam
- Percentage of course assessed by take-home exam
- Number of assessment elements on the course
- Number of assessment elements taken by the student in the academic year

**Figure 4: Relationship between assessment variables and student marks on courses**

Study level	Department	Variable	2015	2016	2017	2018
FD	All qual	10pp increase: in-person exam	-0.10	-0.07	-0.08	-0.06
		10pp increase: take-home exam	-0.17	-0.19	-0.10	-0.11
		1 additional assessment element on course	-0.15	-0.23	0.14	-0.07
		1 additional assessment element in year	0.12	0.06	0.13	0.01
	All quant	10pp increase: in-person exam	-0.11	-0.24	-0.24	-0.15
		10pp increase: take-home exam	-0.33	-0.22	-0.30	-0.15
		1 additional assessment element on course	-1.10	-0.31	-2.45	-2.35
		1 additional assessment element in year	0.23	0.07	0.00	0.10
PGT	All qual	10pp increase: in-person exam	-0.13	-0.12	-0.08	-0.09
		10pp increase: take-home exam	0.01	0.06	0.03	0.00
		1 additional assessment element on course	-0.17	0.44	0.39	0.03
		1 additional assessment element in year	0.05	0.05	0.13	0.15
	All quant	10pp increase: in-person exam	-0.03	0.09	0.03	-0.09
		10pp increase: take-home exam	-0.22	-0.11	-0.04	-0.09
		1 additional assessment element on course	0.63	0.48	0.65	0.44
		1 additional assessment element in year	0.13	0.14	0.11	0.11

### Assessment count on courses affects marks for some, not all, students

Although we may worry about assessment overload, the number of assessment elements on a course do not show a consistently negative relationship with student outcomes. When controlling for other characteristics, more assessment elements on a course typically correlate with lower student marks on undergraduate quantitative courses. But on taught postgraduate quantitative courses, more assessment elements seem linked to better marks. Similarly, there is no clear relationship between the number of assessment elements in a year of study and a student's outcome on individual courses – and any effects observed appear small.

### Non-coursework assessment usually lowers marks

There is a clearer relationship between non-coursework assessment and student marks. At undergraduate level, both take-home and in-person exams relate to worse student marks – the more weight attached to non-coursework assessments on a course, the lower the marks tend to be. For taught postgraduates on qualitative courses, we see this same relationship for in-person exams (but not take-home); for taught postgraduates on quantitative courses we see the relationship for take-home exams (but not in-person).

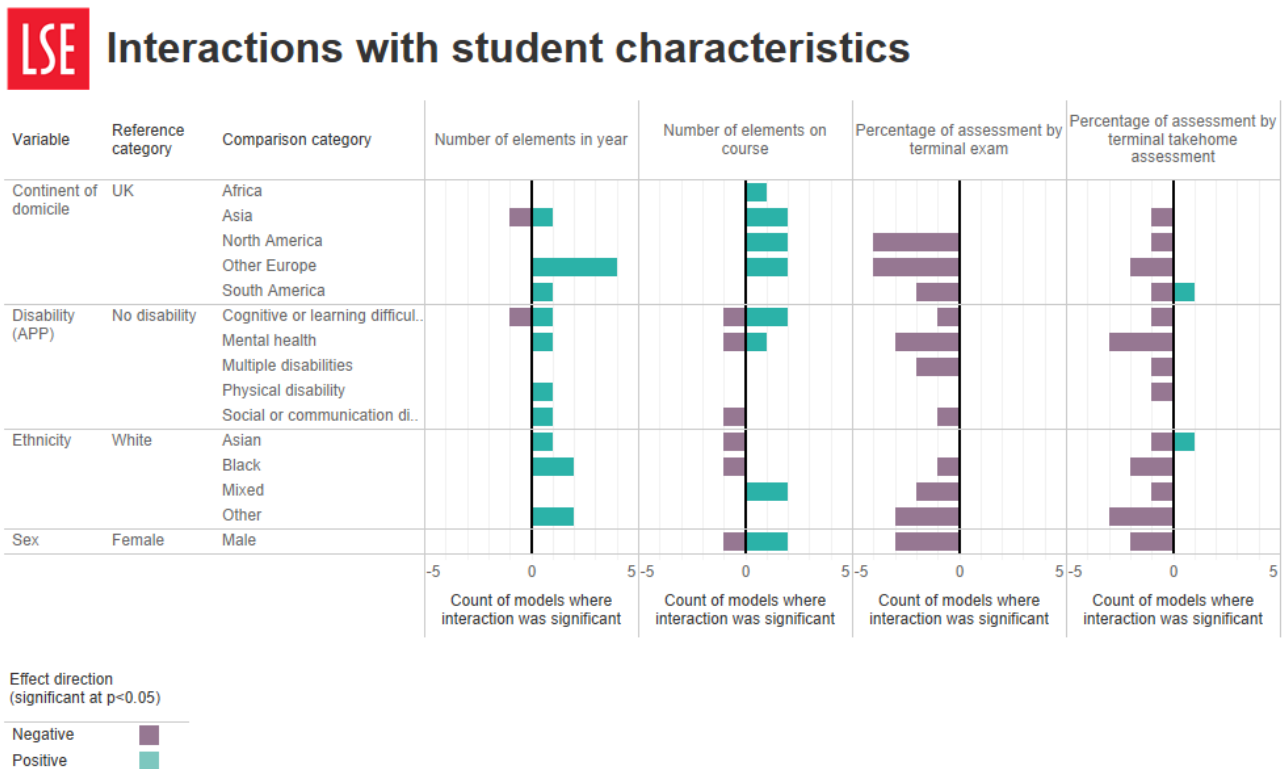
The effects are not necessarily large – a 10 percentage point change in non-coursework assessment weight can be correlated with an average decrease of between 0.06 and 0.3 marks. But when aggregated up this effect could be much larger. For example, while most students in the second year of BSc Economics have zero coursework assessment on the programme, some outliers may have as much as 60% of their overall grade in the year determined by coursework assessment. This could mean a big difference in overall marks for the year.

### Impact of demographic characteristics

We then moved on to look at interactions between assessment types and student characteristics. We wanted to understand whether some groups of students are more likely than others to experience negative (or positive) outcomes based on assessment methods.

Figure 5 shows the number of the four models where the interaction was significant, and the direction of the effect. The bigger the bar, the more models showed a significant interaction between the demographic characteristic and the assessment variable. A fuller model output is available in Annex B.

**Figure 5: Interactions between student characteristics and assessment methods**



### **Students from North America and non-UK EU countries consistently perform better with coursework assessment**

Of the four demographic characteristics we tested, a student's country of domicile showed the most consistent pattern across departments and study levels. Controlling for other characteristics, students from North America and non-UK European countries experience worse outcomes than UK-domiciled peers when assessed by exam rather than coursework. This applies on quantitative and qualitative courses, and at undergraduate and postgraduate level. A 10 percentage point increase in exam-based assessment can be associated with a drop in between 0.1 and 0.5 marks for students domiciled in these areas. Taught postgraduates from South America experience the same disadvantage.

At the same time, taught postgraduates from Asia, North America and non-UK Europe experience an uplift in marks with each additional assessment element on the course. Adding one assessment element is associated with an increase of 0.35-1 marks.

### **Students with a mental health condition experience disadvantage from in-person exams**

With the exception of taught postgraduates on qualitative courses, students with declared mental health difficulties experienced a 0.1-0.5 mark penalty with each additional 10 percentage points of their course that was assessed by in-person exam. In quantitative courses this same disadvantage appeared for students with multiple declared disabilities.

### **Effects of ethnicity are more apparent on quantitative courses**

When compared to their white peers, almost all students of colour on quantitative undergraduate courses experience a disadvantage from final exams (both in-person and take-home). A 10 percentage point increase in exam-based assessment could be correlated with a drop of 0.2-0.5 marks for students on these courses from Black, Mixed or Other ethnic backgrounds – Asian students experienced a disadvantage for take-home exams but not in-person exams.

### **Other effects may exist but be difficult to detect due to student numbers**

Where the number of students with a certain characteristic on a particular type of course is small, it may be difficult for the model to identify statistically significant effects. There may be patterns of outcomes that are not picked up by this model, and any further qualitative research should bear this in mind.

## Conclusions and recommendations

The most consistent finding relates to the relative underperformance (compared to UK-domiciled students) of students from non-UK European and North American countries in courses where assessment is heavily weighted towards exams. Any changes to assessment methods must recognise this, while also considering that some students perform better in exams and would potentially be disadvantaged by a move away from them. This pattern might reasonably be explained by the education systems (either Level 3 or undergraduate) that these students have experienced before coming to LSE, where coursework assessment is more common than terminal. One option would be to deliberately target exam preparation activities towards sub-groups who we might expect to derive more benefit from them. This will need to be carefully framed so as not to appear to place responsibility with the students. This intervention would be an excellent candidate for evaluation under the School's developing evaluation framework.

Students on undergraduate quantitative courses from Black, Mixed or Other ethnic backgrounds also appear to experience worse outcomes from terminal assessments compared to White students. This is harder to explain than the difference based on country of domicile, and a possible intervention might be to trial more diverse assessment methods on a subset of quantitative courses to review the impact on all students. Again, this could be a good candidate for evaluation as part of LSE's Access and Participation Plan.

Students with a mental health difficulty, again in quantitative programmes, also experience a disadvantage from in-person examinations. The intervention identified above could also be used to understand the impact on their performance.

The number of assessment elements on a course appears to have a positive impact for many (though not all) students, especially at taught postgraduate level. It is not entirely clear what may be driving this. Perhaps students have benefitted from exposure to a wider range of ways of demonstrating their learning. Or maybe the assessment load has been distributed more evenly across the year, meaning that their performance is less affected by stress than if everything rested on a handful of terminal assessments – a single 'bad day' has less impact. Further qualitative research – or, if we invest in cleansing and organising assessment data, quantitative research may be possible – will help us to understand the causal mechanisms, and how they operate for different groups of students, to ensure that changes to assessment practice retain some of these benefits.

As outlined in Annex A, data on courses, programmes and assessments are held across multiple systems and can be very difficult to reconcile. Cleansing and reconciliation of this data would allow us to review how individual students experience assessment throughout their programme, as well as allowing more detailed analysis on, for example, the relationship between formative and summative assessment.

The key recommendations are therefore as follows:

- In a subset of undergraduate and taught postgraduate quantitative courses, pilot alternative methods of assessment and review the impact for students from racialised ethnic backgrounds and/or with mental health difficulties.
- Explore, with further qualitative and quantitative research, the possible explanations for the positive impact of assessment count for certain sub-groups of students, especially in the taught postgraduate cohort.
- Develop and run targeted exam preparation (building on existing resources in LSE Life) for students from non-UK European and North American countries, who experience worse outcomes when assessed via in-person exams. Evaluate this intervention, ideally through a randomised treatment group or, if that is not possible/ethical, through a quasi-experimental study.

- Undertake further developmental work to the underlying data so that:
  - o Additional assessment data about e.g. formative assessments, assessment types, wordcounts can be added into the dataset to understand more about assessment patterns.

Data about optionality on programmes can be integrated with assessment data to understand how programme design may affect assessment pathways for students.



## Annex A: data on assessments, courses and programmes

Data about student assessments on their LSE programmes is held in two main systems:

- CAPIS – LSE’s course management system contains relatively detailed data about formative and summative assessments. This data is entered by departments to produce the course guides and has not, therefore, been structured for onward analysis. Much of the data – especially about formative assessments – is held in free text boxes and would need to be coded to be used for onward visualisation and analysis. Courses may have different assessment patterns depending on the student’s programme of study.
- SITS – LSE’s student management system contains data about student module takings and results. In recent years results data is available at the individual assessment level but this has not been available historically. Where students have deferred assessments, or retaken them the following year without repeat teaching, the assessment pattern may differ from the one held in CAPIS.

Further data, necessary to explore assessment patterns in more detail, is held in REG, LSE’s programme regulations management system. This system holds data on which courses are available on which programmes and would allow us to identify, for example, whether students are obliged to sit an assessment (because the course is compulsory on their programme) or whether they may have been able to choose between different assessment types when selecting the course they take to meet the requirements of a paper on their programme. We might expect students on more structured programmes (with fewer optional courses, or a smaller range of options on courses) to have a more purposefully designed assessment experience.

## Annex B – interaction coefficients

This table shows the interactions between student characteristics (variables) and assessment variables. Coefficients that are significant at  $p < 0.05$  are highlighted in purple (negative relationship) and green (positive relationship).

Variable	Comparison category	Study level	Department	Assessment variable							
				1 additional element in year		1 additional element on course		10pp increase - in-person exam		10pp increase - take-home exam	
				Effect size	p	Effect size	p	Effect size	p	Effect size	p
Continent of domicile	Africa	FD	Qualitative	0.19	0.35	-0.40	0.48	-0.02	0.81	-0.06	0.60
			Quantitative	-0.60	0.08	-1.39	0.15	0.05	0.78	-0.13	0.52
		PGT	Qualitative	0.06	0.64	0.27	0.46	-0.08	0.10	0.12	0.25
			Quantitative	0.14	0.27	0.95	0.03	-0.18	0.07	-0.34	0.06
	Asia	FD	Qualitative	0.10	0.25	0.08	0.72	-0.04	0.37	-0.13	0.01
			Quantitative	-0.29	0.03	-0.72	0.05	-0.01	0.81	0.07	0.30
		PGT	Qualitative	0.16	0.01	0.50	0.01	-0.05	0.05	0.03	0.59
			Quantitative	0.11	0.08	0.78	0.00	0.02	0.74	-0.14	0.18
	Australasia	FD	Qualitative	-0.24	0.58	1.06	0.27	-0.16	0.42	-0.18	0.41
			Quantitative	-0.67	0.41	2.35	0.13	-0.38	0.11	0.21	0.45
		PGT	Qualitative	-0.09	0.48	-0.27	0.47	-0.03	0.54	-0.20	0.06
			Quantitative	-0.01	0.98	-0.04	0.95	-0.12	0.34	-0.10	0.70
North America	FD	Qualitative	-0.36	0.09	0.54	0.24	-0.23	0.00	-0.20	0.06	
		Quantitative	-0.23	0.59	-0.15	0.88	-0.47	0.01	-0.20	0.28	
	PGT	Qualitative	0.06	0.26	0.37	0.01	-0.09	0.00	0.04	0.30	
		Quantitative	0.10	0.13	0.76	0.00	-0.12	0.01	-0.27	0.01	
Other Europe	FD	Qualitative	0.14	0.03	0.08	0.65	-0.13	0.00	-0.17	0.00	
		Quantitative	0.23	0.04	-0.17	0.57	-0.18	0.00	-0.09	0.13	
	PGT	Qualitative	0.10	0.01	0.45	0.00	-0.10	0.00	0.01	0.73	
		Quantitative	0.10	0.02	0.63	0.00	-0.06	0.02	-0.19	0.00	
South America	FD	Qualitative	0.15	0.61	0.67	0.39	-0.03	0.84	-0.23	0.23	
		Quantitative	0.09	0.94	1.48	0.55	-0.29	0.40	-0.58	0.15	
	PGT	Qualitative	0.26	0.00	-0.43	0.06	-0.18	0.00	0.14	0.03	
		Quantitative	0.20	0.06	0.28	0.43	-0.29	0.00	-0.53	0.00	
Disability (APP)	Cognitive or learning difficulty	FD	Qualitative	0.08	0.41	0.22	0.40	-0.12	0.02	-0.18	0.01
			Quantitative	0.82	0.00	-0.78	0.01	-0.24	0.06	-0.14	0.29
		PGT	Qualitative	-0.29	0.02	0.82	0.01	-0.01	0.81	0.25	0.05
			Quantitative	0.13	0.32	0.45	0.01	-0.19	0.07	-0.24	0.06
	Mental health	FD	Qualitative	0.23	0.03	-0.14	0.59	-0.10	0.03	-0.22	0.00
			Quantitative	0.33	0.14	-0.80	0.01	-0.40	0.00	-0.30	0.01
		PGT	Qualitative	-0.03	0.80	0.31	0.26	-0.03	0.38	0.11	0.30
			Quantitative	0.19	0.13	0.43	0.01	-0.38	0.00	-0.43	0.00
	Multiple disabilities	FD	Qualitative	0.02	0.87	-0.23	0.59	0.00	0.98	-0.26	0.00
			Quantitative	0.63	0.14	-0.65	0.63	-0.48	0.03	-0.51	0.06
		PGT	Qualitative	-0.23	0.21	0.53	0.30	0.02	0.67	-0.05	0.75
			Quantitative	0.48	0.06	0.55	0.53	-0.37	0.04	-0.57	0.32
Physical disability	FD	Qualitative	0.31	0.04	-0.44	0.28	-0.08	0.28	-0.18	0.04	
		Quantitative	0.40	0.31	0.22	0.80	-0.09	0.54	-0.03	0.86	

		PGT	Qualitative	-0.18	0.20	0.46	0.24	-0.06	0.22	0.01	0.95
			Quantitative	0.17	0.31	0.26	0.65	-0.11	0.34	-0.06	0.79
	Social or communicati on disability	FD	Qualitative	0.31	0.41	-2.45	0.02	-0.32	0.08	-0.15	0.52
			Quantitative	0.59	0.93	16.77	0.38	14.32	0.26	18.57	0.23
		PGT	Qualitative	-1.43	0.20	-1.13	0.34	-0.04	0.85	-0.03	0.97
			Quantitative	0.09	0.05	10.44	0.60	-0.14	0.00	0.48	0.86
Ethnicity	Asian	FD	Qualitative	0.05	0.40	-0.01	0.96	0.01	0.78	-0.03	0.42
				Quantitative	0.42	0.00	-2.18	0.00	-0.07	0.11	-0.22
		PGT	Qualitative	0.03	0.67	0.29	0.13	-0.01	0.55	0.10	0.04
				Quantitative	0.10	0.08	0.21	0.30	0.00	0.98	0.04
	Black	FD	Qualitative	0.04	0.75	-0.16	0.59	-0.08	0.15	-0.16	0.02
				Quantitative	0.27	0.21	-2.65	0.00	-0.47	0.00	-0.48
		PGT	Qualitative	0.35	0.00	-0.14	0.65	0.07	0.09	0.14	0.14
				Quantitative	0.28	0.02	0.34	0.39	-0.12	0.18	-0.27
	Mixed	FD	Qualitative	-0.06	0.49	0.42	0.05	-0.02	0.58	-0.05	0.29
				Quantitative	0.15	0.42	-1.05	0.06	-0.21	0.02	-0.22
		PGT	Qualitative	-0.06	0.40	0.31	0.12	0.00	0.91	0.07	0.26
				Quantitative	0.14	0.10	0.61	0.03	-0.17	0.00	-0.01
Other	FD	Qualitative	0.23	0.11	-0.54	0.18	-0.13	0.06	-0.20	0.02	
			Quantitative	0.54	0.03	-1.07	0.19	-0.50	0.00	-0.31	0.04
	PGT	Qualitative	-0.04	0.59	0.33	0.17	-0.11	0.00	0.04	0.58	
			Quantitative	0.35	0.00	0.12	0.71	-0.27	0.00	-0.47	0.00
Sex	Male	FD	Qualitative	0.07	0.15	-0.01	0.91	-0.02	0.46	-0.07	0.03
				Quantitative	0.03	0.78	-0.98	0.00	-0.27	0.00	-0.13
		PGT	Qualitative	-0.06	0.18	0.40	0.00	-0.09	0.00	0.01	0.90
				Quantitative	0.07	0.10	0.44	0.01	-0.11	0.00	-0.14

## Annex C – models

We ran two regressions analyses to generate the findings in this paper. Full detail on the variables and excluded datapoints are in Table 1 below.

The first analysis was a mixed effects linear regression to model student outcomes by discipline and study level. We included a student-level random effect to account for differences between students that could not be captured by variables included in the model, and an interaction term for the influence of each fixed variable within an academic year.

This analysis confirmed that, for each subset of students, the effect of each variable of interest (the assessment variables) was similar across academic years. We were therefore able to aggregate data for the academic years 2016/7-2018/9 in order to examine the effect of student characteristics on the variables of interest. We performed this aggregation to ameliorate the effect of relatively small numbers of students with certain characteristics within the subsets of students.

We then ran the second analysis, a mixed effects linear regression with a student-level random effect and interaction terms for the four demographic characteristics of ethnicity, disability, sex and continent of domicile. This generated the findings outlined in the second section of the report.

**Table 1: Variables included in the analysis models**

Variable	Description	Values (reference value in bold)	Exclusions
Student module mark	Mark achieved by student on module	Percentage (1-100)	0 or Null
Academic year	Academic year of study	<b>2015/6</b> 2016/7 2017/8 2018/9	
Module department	Department of module	<b>Accounting (quant)</b> Economics (quant) Finance (quant) Mathematics (quant) Management (quant) Methodology (quant) Psychological & Behavioural Science (quant) Statistics (quant) Anthropology (qual) Economic History (qual) European Institute (qual) Gender Institute (qual) Geography & Environment (qual) <b>Government (qual)</b> Health Policy (qual) International Relations (qual) Law (qual) Language Centre (qual) Media & Communications (qual) Philosophy, Logic and Scientific Method (qual) Public Policy (qual) Sociology (qual) Social Policy (qual)	

Module discipline	Was the module in the student's home department, their own discipline or the other discipline?	<b>Home department</b> Quant module on quant programme Quant module on qual programme Qual module on qual programme Qual module on quant programme	
Disability	Student's declared disability (based on APP categories)	<b>No disability</b> Cognitive or learning difficulty Mental health Multiple disabilities Physical disability Social or communication disability	
Ethnicity	Student's ethnicity (based on HESA categories)	Asian Black Mixed Other <b>White</b>	Prefer not to say Unknown
Sex	Student sex	<b>Female</b> Male	Other
Domicile continent	Student continent of domicile	Africa Asia Australasia North America Other Europe South America <b>UK</b>	Unknown
Year of study	Student year of study	<b>1</b> 2 3 4	
Terminal in-person exam	Percentage of module assessment determined by in-person exam	Percentage (0-100)	
Terminal take-home exam	Percentage of module assessment determined by take-home exam	Percentage (0-100)	
Element count on course	Count of individual assessment elements within the module	Count	
Element count in year	Count of individual assessment elements across all modules taken by the student in the academic year	Count	