A Longitudinal Analysis of Module Grade Classification at LSE

Yuhong Wu, Chengchun Shi, and Ellen Austin
London School of Economics and Political Science

Abstract

In July 2019, the Office for Students (OfS) published a report (OfS 2019.28) on changes in the proportion of first- and upper second-class degrees awarded by UK universities from 2010-11 to 2017-18. The report investigates the extent to which the changes in degree attainment can be statistically accounted for by changes in individual characteristics. The increase in proportion that is unexplained by changes in the graduate population over the years is interpreted as evidence of 'grade inflation' in the UK higher education sector.

A follow-up study of OfS 2019.28, which investigates the degree classification in LSE, identified similar trends in the change of proportions of first- and upper second-class degrees awarded by LSE. The study builds on the analysis of OfS 2019.28 by investigating grade inflation at the departmental level and reveals that the extent of grade inflation is heterogeneous among departments. The study also points out that although there is 6.2 percentage points unexplained increase in the proportion of first- and upper second-class degrees attained at LSE, it is substantially lower than the sector-level, which is 13.3 percentage points.

The aim of this report is to apply similar methods to a different variable of interest, which is the individual module grades for students. This report seeks to extend the analysis in the LSE grade inflation study by performing statistical modelling at the module level and expanding the time series considered to include 2009-10 to 2019-20 inclusive. A hypothetical group is introduced in the report to close the attainment gaps between individuals, providing insights into the effects of individual characteristics on grade attainment. This report also includes the comparison between the grade inflation trends of quantitative and qualitative modules to show the effect of module level characteristics.

This report finds out that in 2019-20, 18.0 percentage points of first- and upper second-class grade attainment change, and 18.1 percentage points of first-class grade attainment change since 2009-10 are unexplained by changes in the student population. The unexplained changes would be lower when the attainment gaps are hypothetically 'removed', which become 12.6 percentage points and 13.4 percentage points respectively. Whilst the extent of grade inflation is heterogeneous across different modules and departments, most modules

and departments experienced a huge unexplained increase in first- and upper second-class in 2019-20. This report also observes that quantitative modules tend to experience higher grade inflation, while qualitative modules exhibit a more discernible trend in first-class grade inflation.

Introduction

From 2018 to 2020, the Office for Students (OfS) published a series of reports (OfS 2018.54, OfS 2019.28, and OfS 2020.52) on changes in degree classifications over time across UK universities, focusing on the problem of grade inflation. Grade inflation can be defined as an upward shift in student grades over an extended period without a corresponding increase in student achievement. Using statistical modelling at the individual student level, the reports aimed to quantify changes in proportions of first- and upper second-class degrees that are unexplained by changes in the graduate population over the years 2010-11 to 2018-19. These unexplained changes were then used to determine the amount of grade inflation over the years.

In LSE, a follow-up study of OfS 2019.28 investigated the changes in the proportion of first- and upper second-class degrees awarded by LSE from 2010-11 to 2017-18. The study built on OfS 2019.28 by examining the grade inflation more closely at department level in LSE. The study revealed that grade inflation in LSE becomes noticeable starting in academic year 2014-2015. It estimated a 6.2 percentage point unexplained increase in first- and upper second-class degrees awarded by LSE in academic year 2017-18 relative to 2010-11. There is 7.9 percentage points increases in first-class degrees unexplained by the change in graduating population. The results for first- and upper second-class degrees combined are shown in Table 1 and the results for first-class degrees alone are shown in Table 2.

Table 1: Summary of observed and unexplained changes in upper degree attainment

Academic Year	Observed proportion (%)	Observed change (pp)	Unexplained change (pp)	S.E. of Unexplained change
2010-11 (ref.)	84.5	-	-	-
2011-12	84.5	0.0	0.8	1.4
2012-13	82.2	-2.3	-2.1	1.4
2013-14	84.7	0.2	1.4	1.4
2014-15	87.2	2.7	3.4	1.4
2015-16	90.0	5.5	5.8	1.4
2016-17	91.5	7.0	5.7	1.3
2017-18	91.7	7.2	6.2	1.4

Table 2: Summary of observed and unexplained changes in first-class degree attainment

Academic Year	Observed proportion (%)	Observed change (pp)	Unexplained change (pp)	S.E. of Unexplained change
2010-11 (ref.)	18.6	-	-	-
2011-12	18.2	-0.4	-0.8	1.5
2012-13	20.4	1.8	1.4	1.5
2013-14	20.9	2.3	2.0	1.5
2014-15	25.3	6.7	6.3	1.5
2015-16	25.3	6.7	5.9	1.5
2016-17	26.1	7.0	6.2	1.6
2017-18	27.8	7.2	7.9	1.5

The LSE grade inflation report considers only UK-domiciled, full-time students on Home (UK) fee status who graduated in the academic years from 2010-11 to 2017-18, amounting to a total of 5,015 students. Both fixed- and mixed-effects logistic regression models are used to predict expected degree classification attainment, taking the following factors into account:

- Department
- Year of graduation
- Qualifications on entry

- Age
- Disability
- Ethnicity
- Gender
- Participation of Local Areas (POLAR4) quintile

To investigate grade inflation in LSE, fixed-effects and mixed-effects logistic regression models are fitted to predict the probability of an individual obtaining 1) a first- or upper second-class degree combined ("upper degree") or 2) a first-class degree only, using the above factors as covariates. The academic year of degree attainment is set to be 2010-11 to predict the probability of obtaining first- or upper second-class degrees if the students were graduating in 2010-11.

The predicted proportion of first- and upper second-class degrees are then compared with the observed proportion to obtain the unexplained change and inflation rate for each academic year. The inflation rates at the department level, aggregating the proportions from 2011-12 to 2017-18, are used to rank the departments for consistently high inflation rates. The top six departments with highest aggregate inflation rate are shown in Table 3. The study finds out that the extent of grade inflation is heterogeneous among departments, with high within-department variation over time.

Table 3: Six departments with highest aggregate inflation rate

Upper degrees	First-class degrees
Accounting	Social Policy
Mathematics	Accounting
Statistics	Government
Economics	Philosophy, Logic and Scientific Method
Social Policy	Geography & Environment
Sociology	Sociology

Methodology

The current analysis adopts a similar method as the LSE grade inflation study. It considers module grades of 9177 students who took some of the 690 modules in 31 departments in the academic years from 2009-10 to 2019-20. However, unlike in the OfS analysis and the

LSE grade inflation study, the variable of interest is individual module grade and factors at the module level are also included as predictor variables.

Both fixed- and mixed-effects logistic regression models are used to predict expected module grade, taking the factors in Table 4 into account. A detailed explanation of the variables is presented in Annex C.

Table 4: Predictor variables considered in the models

Factors at student level	Factors at module level
• Year of attainment • Qualifications on entry • Ethnicity • Disability • Sex • Age group • Attainment of Undergraduate Bursary • POLAR4 quintile • IMD quintile • Tundra quintile • ACORN group	• Module code • Module credit • Module level • Number of students

Model Comparison

Following the method used in the previous LSE grade inflation study, a fixed-effects regression model (Simplified model) is fitted to predict the probability of a single individual obtaining 1) a first- or upper second-class grade combined ("Upper-class Grade") or 2) a first-class grade only, using the above factors at student level. The model is then extended by including factors at module level (Full model), which significantly lowers the deviance, Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) of the original model. The model fit statistics are provided in Annex C. The regression equation and corresponding coefficient estimates are given in Annex D.

As some modules are introduced later than the others, this report takes the reference year to be the year when the module was established. The unexplained change in module grade

relative the reference year is calculated by comparing the observed proportion of good grades or first-class grades awarded to the proportion that is predicted by the model. To understand the extent of grade inflation better, the grade inflation rate is calculated as the ratio of the observed proportion to the predicted proportion for each academic year.

Next, four mixed-effects models are fitted with 1) a random intercept for each module, 2) a random intercept for each module, and for academic years that are nested within modules, 3) a random intercept for each department, and for modules that are nested within departments and 4) a random module intercept and a random year coefficient for each module. The regression equation and corresponding coefficient estimates are provided in Annex E. Comparing model fit statistics, mixed-effects model 4 has the lowest deviance, AIC and BIC. However, mixed-effects model 4 leads to non-convergence issue in predicting first-class grades. Furthermore, mixed-effects models generate similar results.

After computing the predicted proportion for upper class grades and first-class grades using the full fixed-effects model and mixed effects model 4 respectively, Figure 1 shows that the full fixed-effects model and mixed-effects model 2 give similar predictions of the aggregate trend. For individual modules, full fixed-effects model predicts the trend more accurately, while mixed-effects models tend to predict a constant trend.

Fig.1: Trend for observed and predicted proportion every year

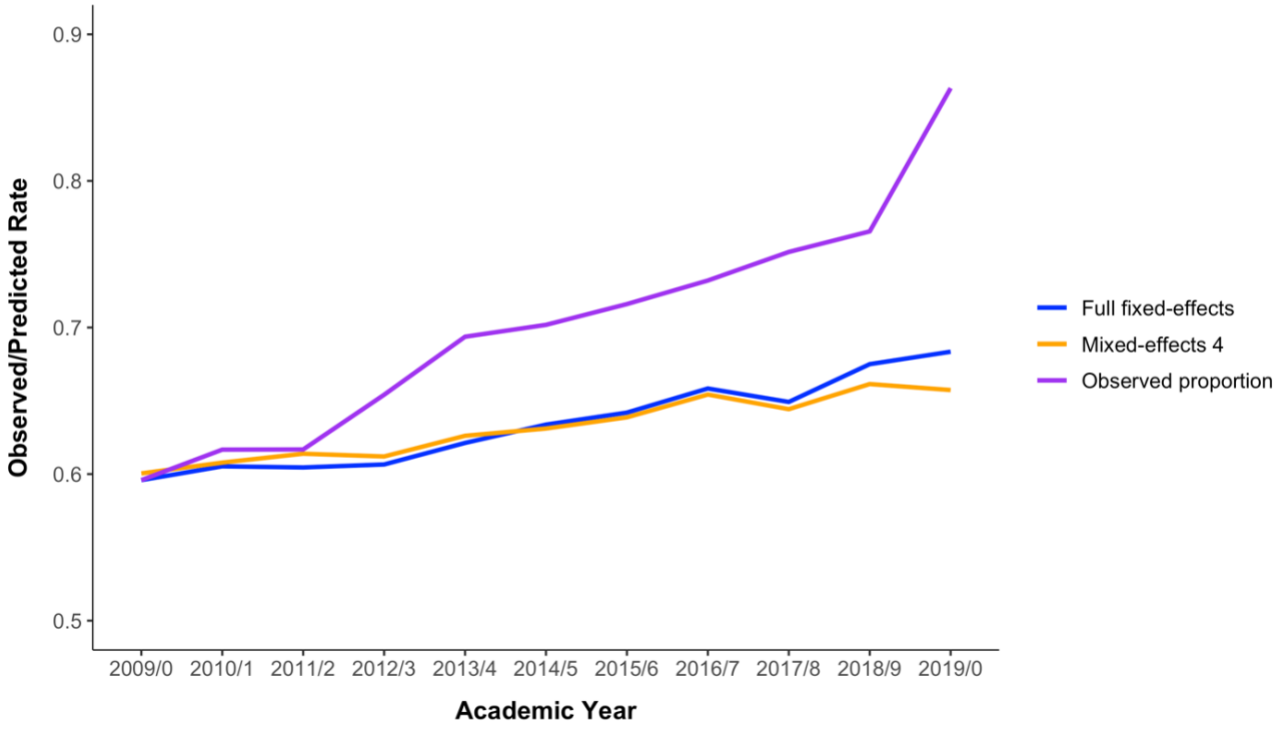
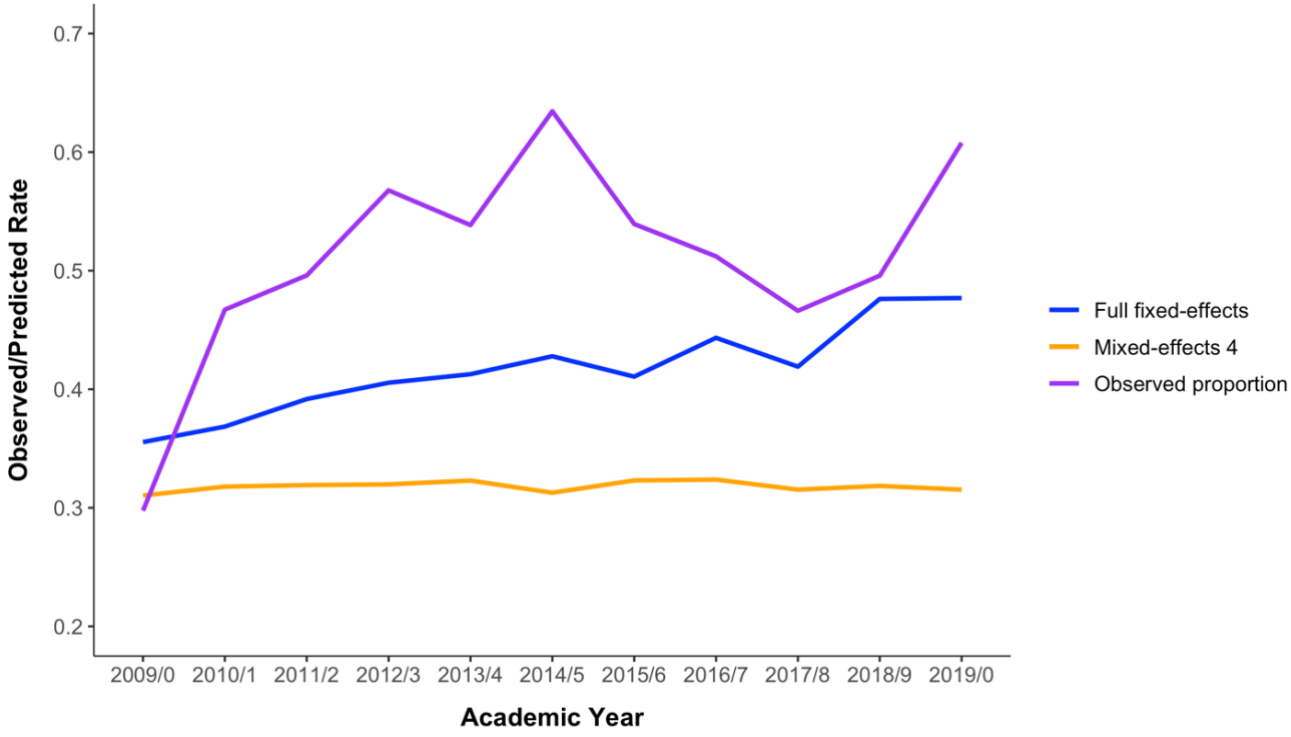


Figure 2 shows the predicted trends for MA100. Across the six fixed-effects and mixed-effects models, the full fixed-effects model has the lowest deviance, AIC and BIC. The model fit statistics are provided in Annex C. As such, results based on the full fixed-effects model are presented in the main text.

Fig.2: Observed and predicted proportion of Module MA100



To understand the overall trend in module grades, results at the departmental level and aggregate results for each academic year are also reported. This is achieved by comparing the observed proportion of upper-class grades or first-class grades awarded in each department for each academic year, with the predicted proportion for each department using the full fixed-effects model. The aggregate results for each year are obtained by comparing the observed and predicted proportion of upper-class grades or first-class grades awarded for each academic year.

To compare grade inflation rates across modules, three different approaches are taken. The first approach is to compute the aggregate grade inflation rates over the years 2010-11 to 2019-20 by calculating the observed and predicted proportion of students obtaining upper-class grades over 10 years, taking 2009-10 as the reference year. The ranking only considers the modules that cover all the academic years in range. The second approach is to identify the frequency that the module grade inflation rate exceeds the aggregate inflation

rate for each department. The modules with inflation rates exceeding the aggregate departmental inflation rate most frequently are taken to have consistently high inflation rate compared to departmental level. The third approach is to identify the frequency that the module grade inflation rate exceeds the aggregate inflation rate for each academic year. The modules with inflation rates exceeding the aggregate inflation rate most frequently are taken to have consistently high inflation rate.

Using the similar method employed by OfS 2020.52, a hypothetical group is created to artificially close the gap between individuals of differing characteristics. The hypothetical group is taken to possess the highest attaining individual characteristics, which are "considered to be those showing the greatest (most positive) regression coefficient estimates" in the full fixed-effects model. The unexplained changes and inflation rates in module grades are then calculated for the hypothetical group.

Results

Descriptive analysis of module grade from 2009-10 to 2019-20

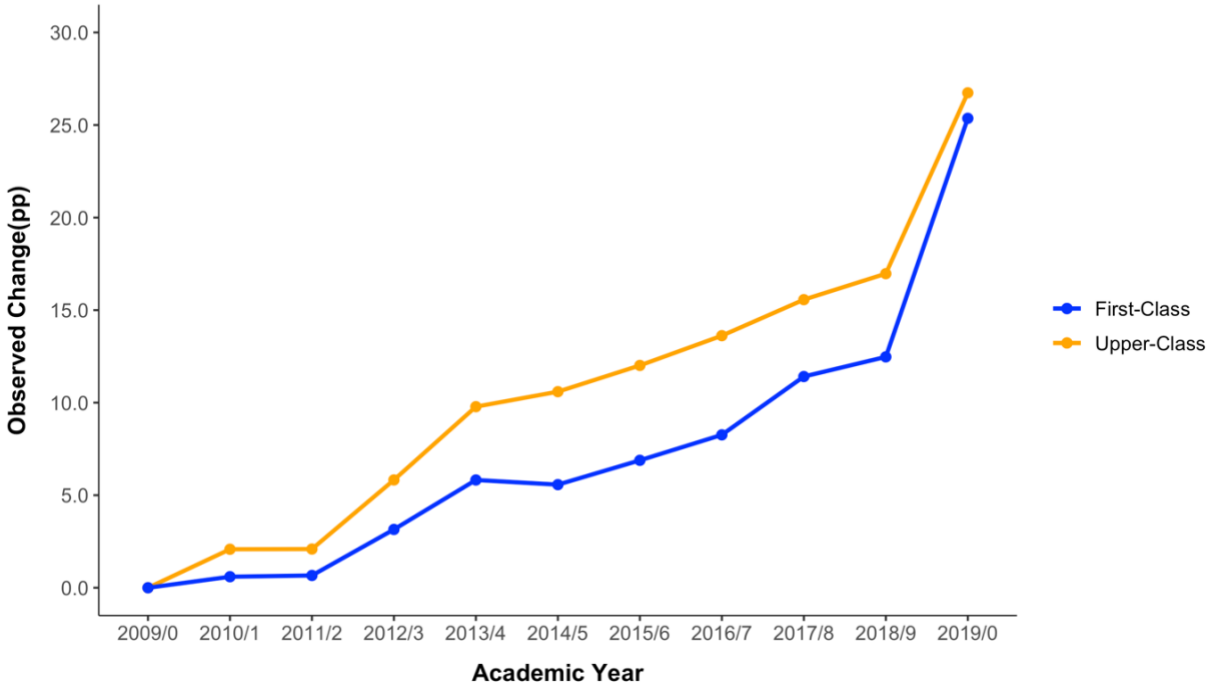
Table 5 below provides a summary of the percentage of students in the sample obtaining classified grades, including passes and fails, in the academic years 2009-10 and 2019-20. The proportion of students in the sample obtaining a first-class grade has risen by approximately 12.4 percentage points between 2009-10 and 2018-19 and experienced a larger increase of 12.9 percentage points in 2019-20 alone. The proportion of students obtaining upper second-class grades has remained relatively constant at around 45%, whilst the proportion of those who obtained lower grades has fallen by approximately 26.7 percentage points between 2009-10 and 2019-20.

Table 5: Summary of module grade classifications for academic years 2009-10 and 2019-20

	2009-10		2018-19		2019-20	
Module Grade	**Number**	**Percentage**	**Number**	**Percentage**	**Number**	**Percentage**
First	1463	17.2%	2988	29.6%	4450	42.5%
Upper Second	3618	42.4%	4732	46.9%	4585	43.8%
Other	3446	40.4%	2364	23.5%	1431	13.7%
Total	8527	100%	10084	100%	10466	100%

Next, the observed change in the proportion of students obtaining an upper-class grade or a first-class grade in each year relative to academic year 2009-10 is plotted in Figure 3 below. The plot indicates that the observed proportion of students obtaining an upper-class grade, or a first-class grade has been increasing since 2009-10, with a huge jump of more than 10 percentage points in 2019-20. By 2019-20, the observed change in the proportion of upper-class grade has risen to around 26.7% and the observed change for first-class grade has risen to around 25.4%.

Fig.3: Trend for observed change every year

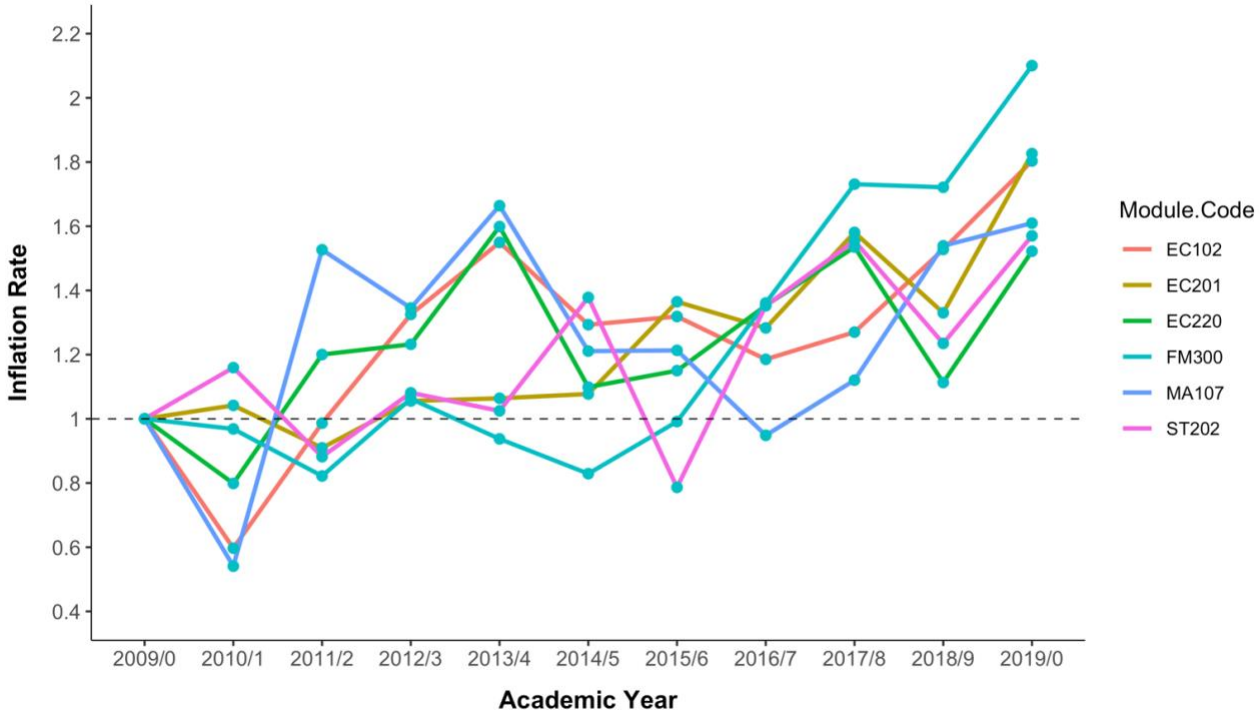


Results from statistical modelling

Grade Inflation in Modules

Figure 3 below present the inflation rate from 2009-10 to 2019-20 for six modules identified to have the highest grade inflation, ranked by their aggregate grade inflation rates across the academic years. For upper-class grades, the top six modules (EC102, EC201, EC220, FM300, MA107 and ST202) are shown respectively. The grade inflation rates for more modules are included in Annex B.

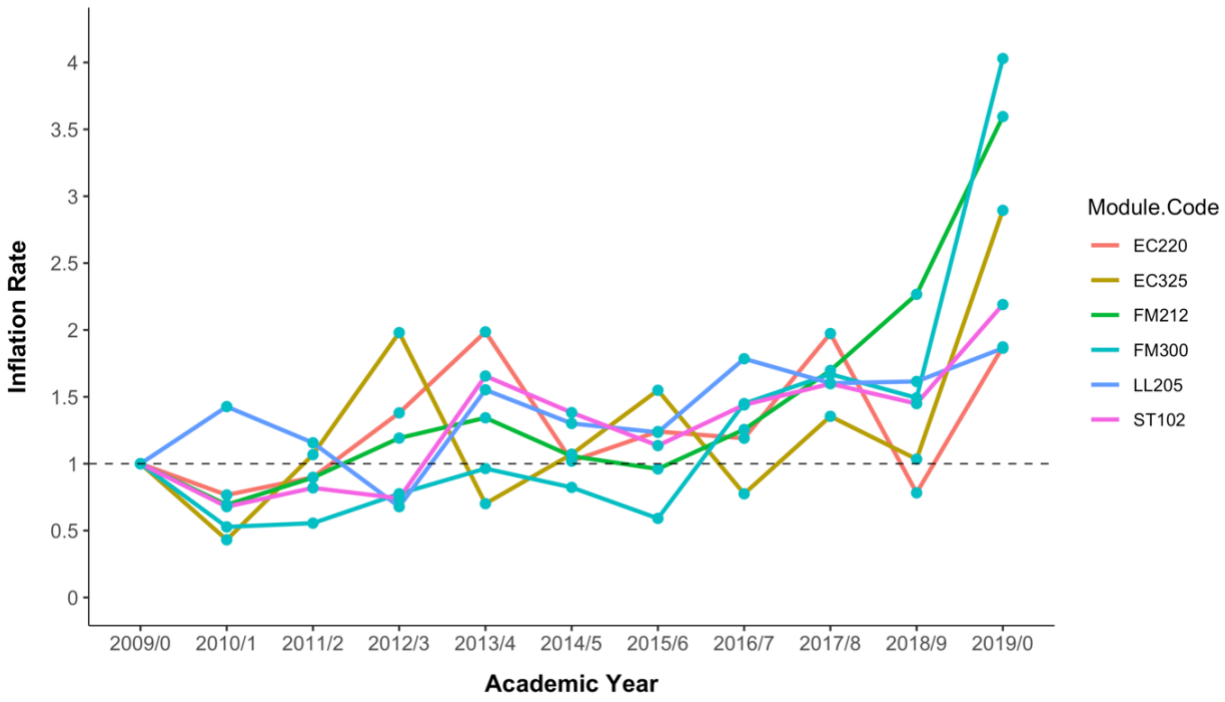
Fig.4: Six modules with high aggregate inflation rate



There are large fluctuations in grade inflation in academic year 2010-11, 2013-14 and 2017-2018, some modules experienced huge increases in grade inflation, whilst the others experienced drops. Although the extent of grade inflation is heterogenous in different academic years, there is generally an upward trend in grade inflation rate for the modules that exhibited the highest grade inflation. Most modules exhibit grade inflation after 2012-13. All six modules experienced a large increase in grade inflation in academic year 2019-20.

Figure 5 below shows the first-class inflation rate from 2009-10 to 2019-20 for six modules identified to exhibit the highest grade inflation (EC220, EC325, FM212, FM300, LL205, ST102). The first-class grade inflation rates for more modules are included in Annex B.

Fig.5: Modules with high aggregate inflation rate for First Class

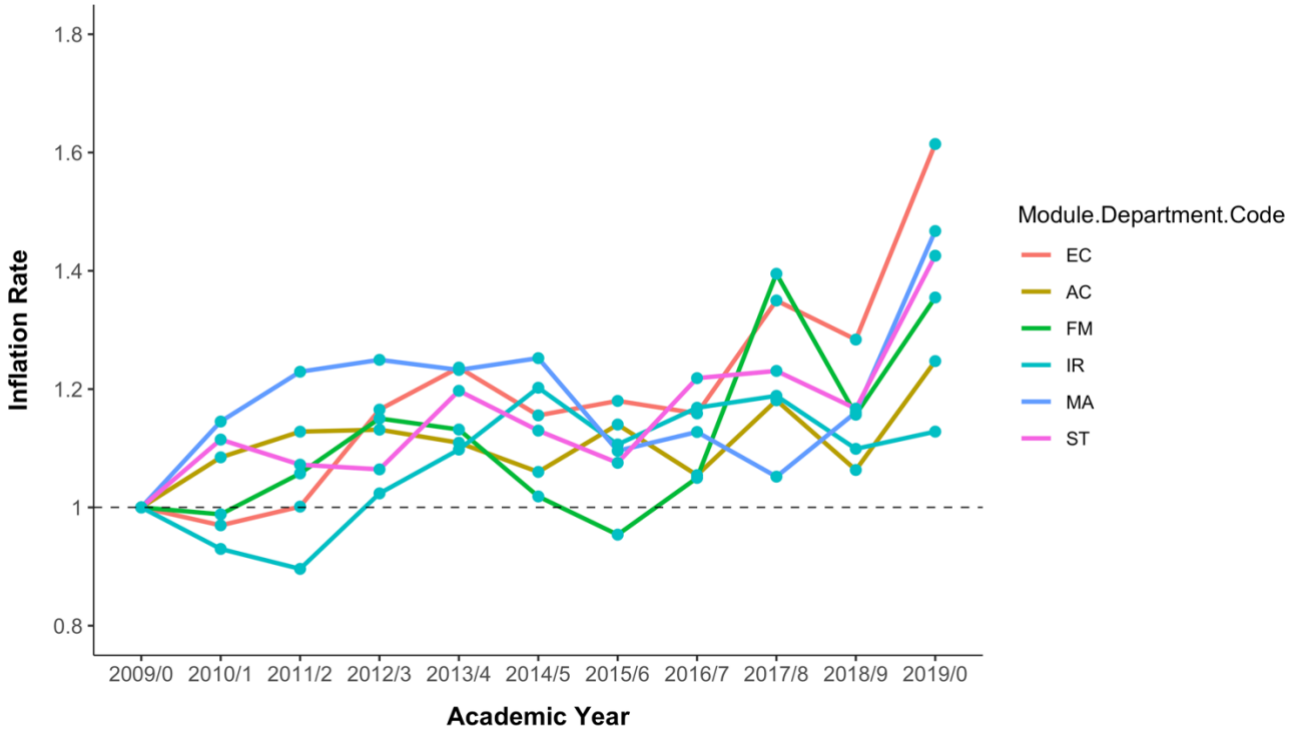


The first-class grade inflation rates are much higher than the upper-class grade inflation rates. There are fluctuations in the first-class grade inflation rates in academic years 2013-14, 2017-18 and 2018-19. Despite the fluctuations, there is generally upward trend in the inflation rates of the modules identified. All six modules experienced a huge increase in grade inflation rate in 2019-20, especially FM212 and FM300, which have increases in inflation rate from approximately 0.7 in 2010-11 to 3.6 in 2019-20 and from 0.5 to 4.0 respectively.

Grade Inflation in Departments

Figure 6 below present upper-class grade inflation rates for the departments identified to have the highest grade inflation from 2009-10 to 2019-20, ranked by their aggregate upper-class inflation rates across the academic years. The observed and predicted proportions at the departmental level are calculated by aggregating the upper-class or first-class grades awarded in each department. For upper-class grades, the top six departments (Accounting, Economics, Finance, International Relations, Mathematic, and Statistics) are shown respectively. The grade inflation rates for more departments are included in Annex B.

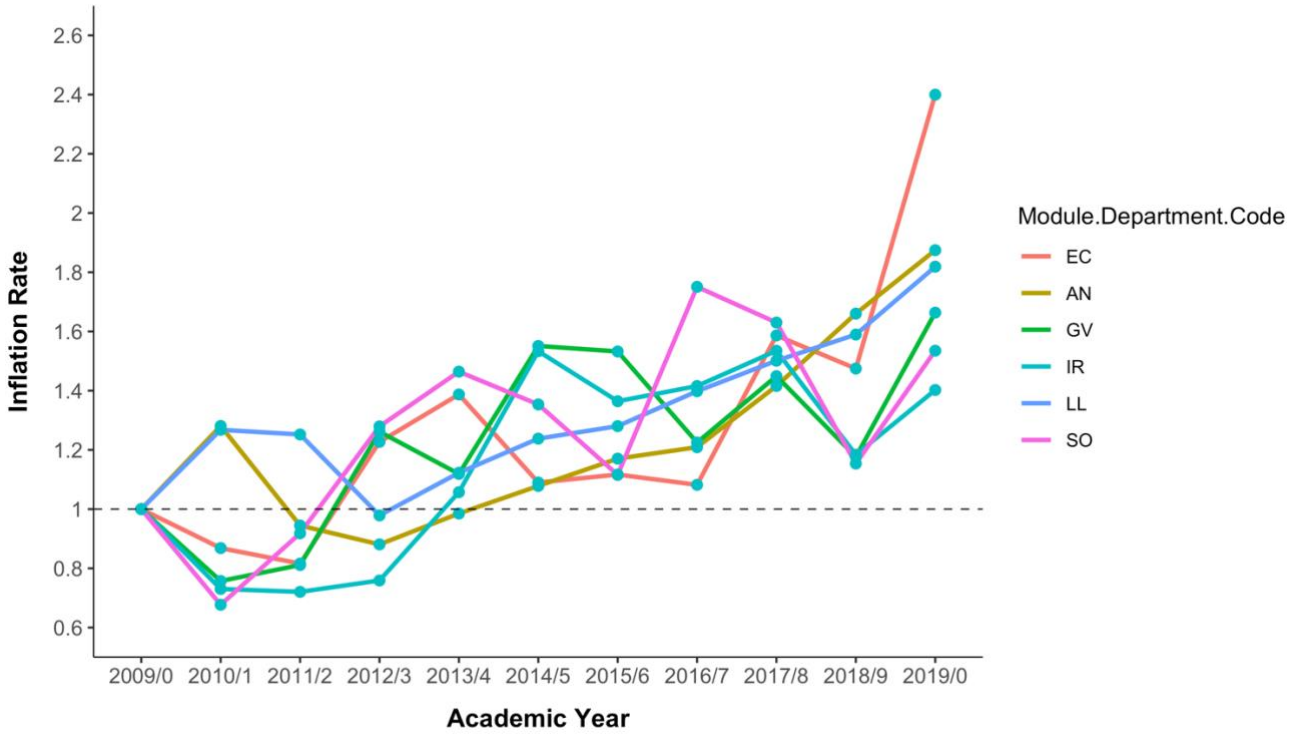
Fig.6: Departments with high aggregate inflation rate



It can be observed that for upper-class grade attainment, there is generally an upward trend for the departments that exhibited the highest grade inflation. There are large fluctuations in inflation rates experienced in academic year 2017-18, most departments experienced increases in inflation rates whilst the Mathematics department experienced a drop in inflation rate. All six departments experienced a huge rise in inflation rate in academic year 2019-20.

Figure 7 below shows first-class grade inflation rates for the departments identified to have the highest grade inflation from 2009-10 to 2019-20, ranked by their aggregate first-class inflation rates across the academic years. The top six departments identified (Anthropology, Economics, Government, International Relations, Law and Sociology) differ from those identified to have highest upper-class grade inflation rates. The first-class grade inflation rates for more departments are included in Annex B. From Figure 6 and 7, quantitative departments have higher upper-class inflation rates and qualitative departments have higher first-class inflation rates.

Fig.7: Departments with high aggregate inflation rate for First-class



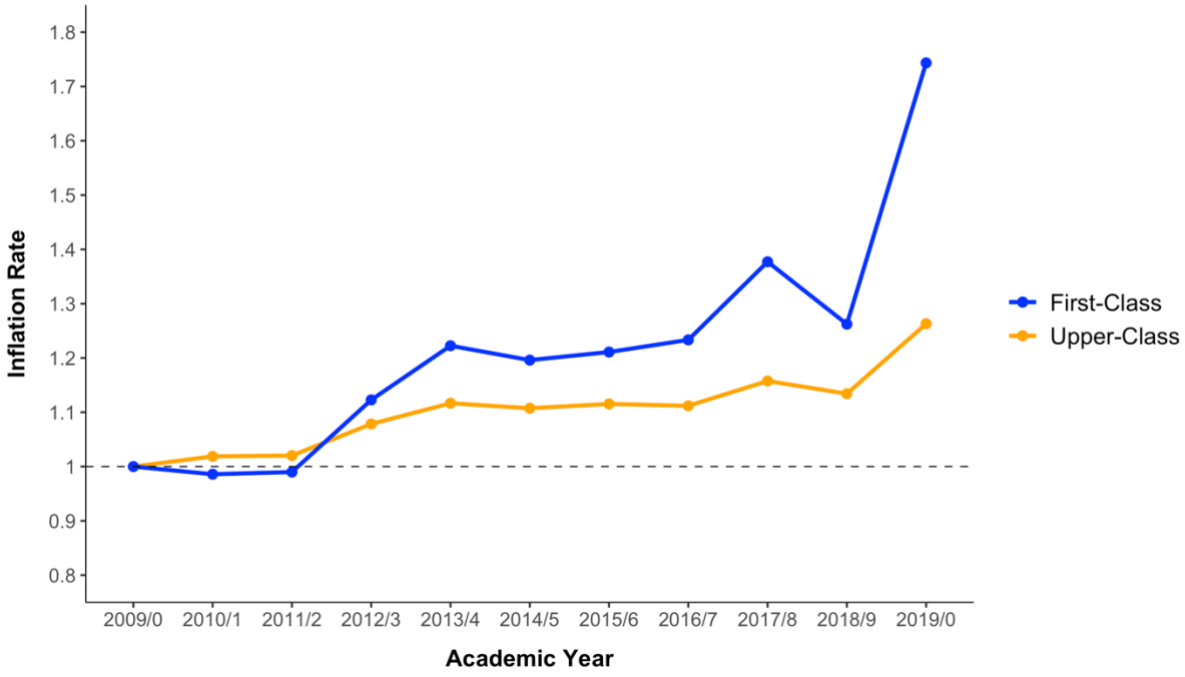
The first-class grade inflation rates at departmental level are generally higher than the upper-class grade inflation rates. Despite the fluctuations in some academic years, there is a much more discernible upward trend in the first-class grade inflation rates. In academic years 2017-18 and 2019-20, all six departments experienced a huge increase in grade inflation rates.

The results at the department level echo the findings in the previous LSE grade inflation study. Departments like Accounting, Economics, Government, Mathematics, Statistics and Sociology are found to experience consistently high degree inflation in the previous study. The high module grade inflation in these departments could potentially be the driving factor of the degree inflation.

Grade Inflation of the sample population

Figure 8 below presents the aggregate upper- and first-class grade inflation rate for each academic year. This is obtained by comparing the observed and predicted proportion of all students obtaining an upper-class or first-class grade for each academic year in the sample.

Fig.8: Aggregate Inflation Rate for first- and upper-classes

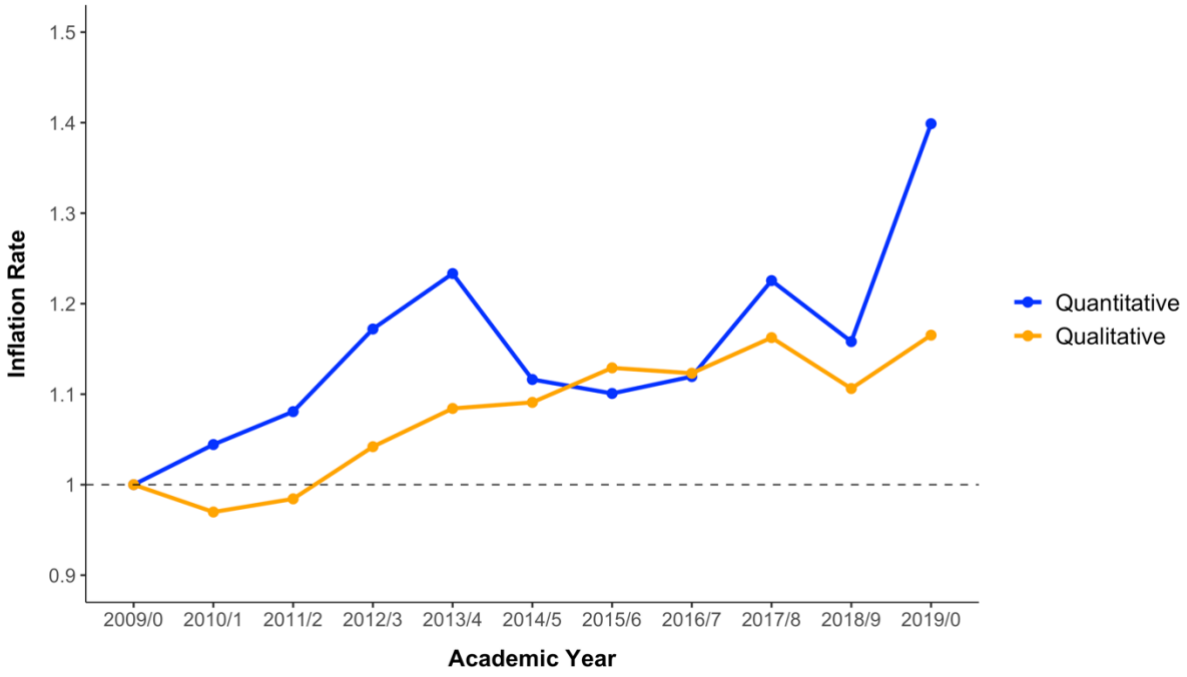


While there are small dips in grade inflation rates in academic year 2014-15 and 2018-19, there is generally an upward trend in both upper-class and first-class grade inflation rates across the years. The first-class inflation rates are generally higher than the upper-class inflation rates. In 2019-20, the aggregate grade inflation rate experienced a huge increase, especially for first-class grades.

Grade Inflation for Quantitative and Qualitative modules

Figure 9 below presents the aggregate upper-class grade inflation rates from 2009-10 to 2019-20 for quantitative and qualitative modules. There are 147 modules classified as quantitative modules and 323 modules classified as qualitative modules in the sample. The grade inflation rates are computed by comparing the observed and predicted proportion of students attaining upper-class or first-class grades in quantitative or qualitative modules.

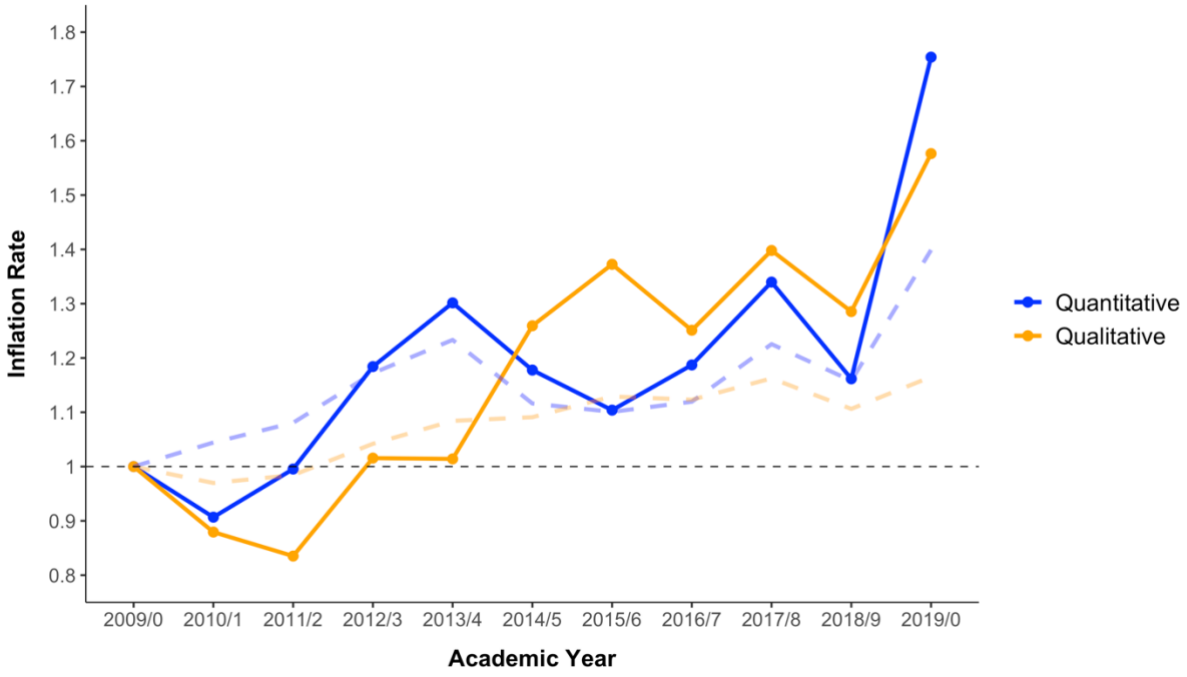
Fig.9: Aggregate Inflation Rate for Quantitative and Qualitative modules



There is generally an upward trend in grade inflation rate for both quantitative and qualitative modules. Although there are dips in grade inflation rate in 2015-16 and 2018-2019, quantitative modules generally experienced higher upper-class grade inflation than qualitative modules. In 2019-20, the grade inflation rate for quantitative modules experienced a large increase from approximately 1.2 to 1.4.

Figure 10 below shows the trends of first-class grade inflation rates for quantitative and qualitative modules. The trends of upper-class grade inflation rates are superimposed as dashed lines.

Fig.10: Aggregate First-Class Inflation Rate for Quantitative and Qualitative modules

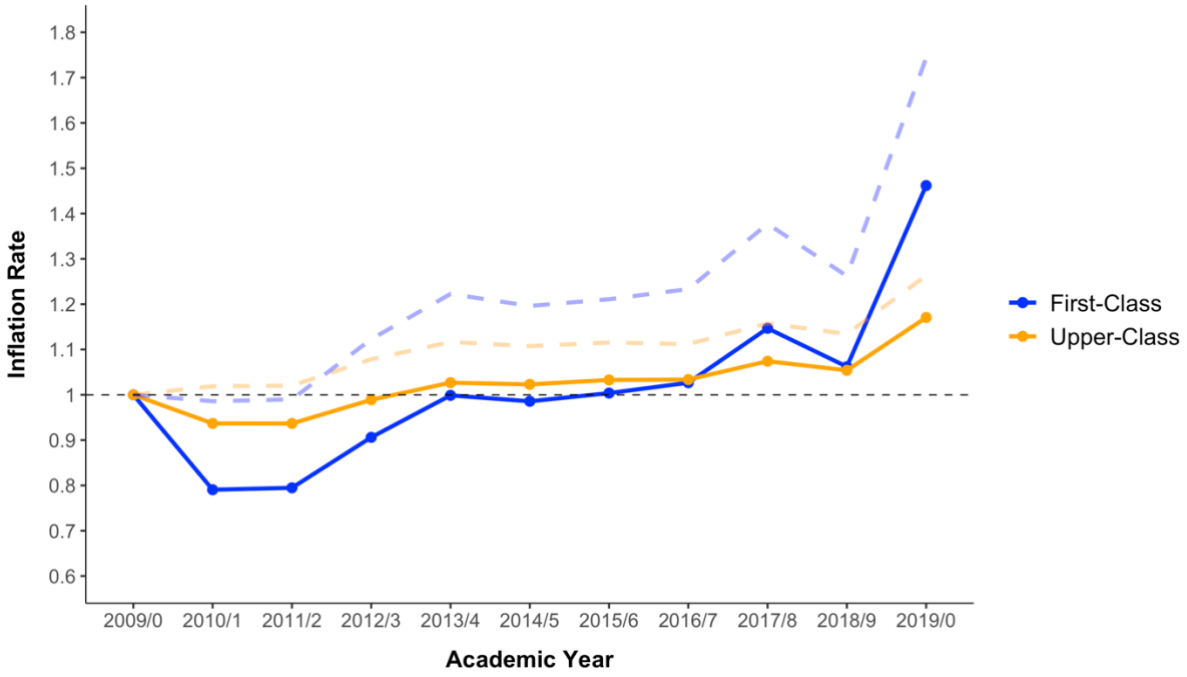


Both quantitative and qualitative modules experience high first-class inflation rates. The grade inflation rate of qualitative modules exceeds that of quantitative modules from 2014/5 to 2018-9. In 2019-20, both quantitative and qualitative modules experienced huge increase in grade inflation rates. More importantly, comparing the upper-class and first-class inflation rates, the grade inflation for qualitative modules could be largely attributed to inflation in first-class grades. On the other hand, the grade inflation for quantitative modules could be driven by the inflation in upper second-class grades.

Grade inflation for the hypothetical group

Figure 11 below shows the grade inflation rates from 2009-10 to 2019-20 for the hypothetical group. The hypothetical group is formed by closing the attainment gaps between individual students. This is achieved by setting the individual contextual characteristics to be the highest attaining categories, as indicated by the highest (most positive) coefficient estimates in the full fixed-effects model. For first- and upper second-class grades combined, the highest attaining students for the additional individual contextual variable groups are white, non-disabled males from POLAR4 quintile 2 areas. The trends for aggregate upper- and first-class grade inflation rate are superimposed as dashed lines.

Fig.11: Aggregate Inflation Rate for the hypothetical group

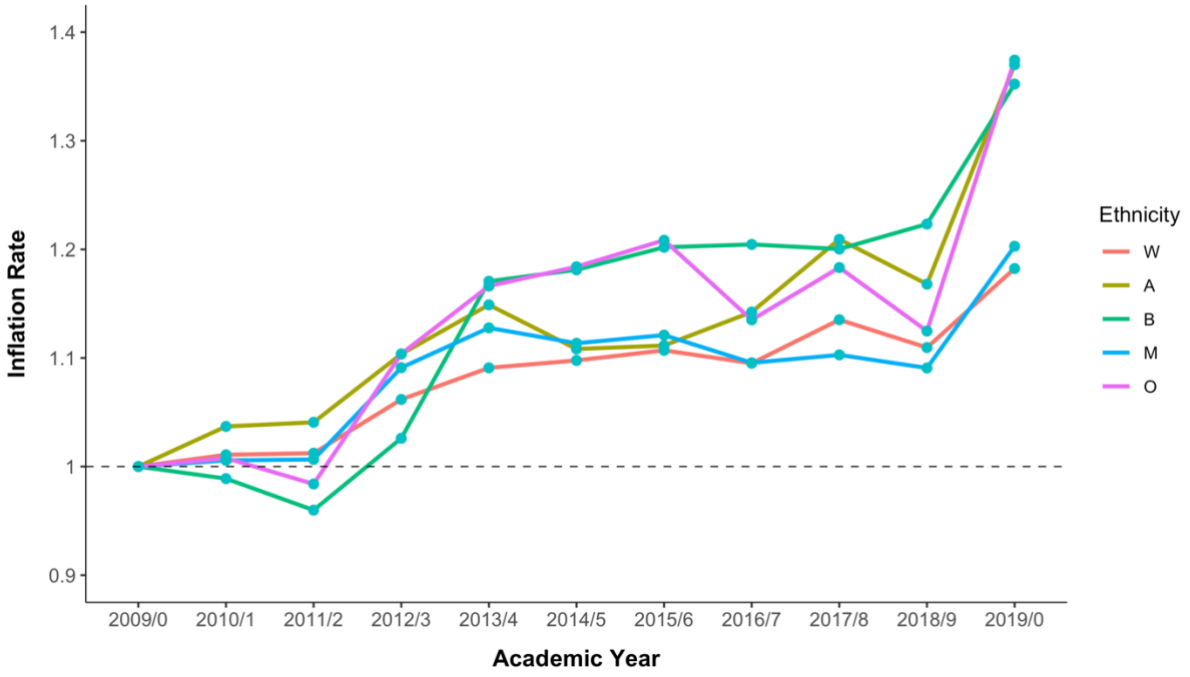


The trends in grade inflation rates for the hypothetical group are similar to the trend of the original sample. However, when we removed attainment gaps for individual contextual characteristics and assumed all students attained at the highest levels predicted by the model, the grade inflation rates are generally lower than the grade inflation rates for the original sample by about 0.2 units. Between 2009-10 and 2013-14, removing the attainment gap eliminates grade inflation effect. This suggests that to some extent, part of the grade inflation could be due to changes in student population. However, in year 2019-20, there are still huge unexplained increases in first- and upper second-class grade attainment of the hypothetical group.

Grade inflation for students from different background

The analysis of hypothetical group suggests that there might be alternative explanations to grade inflation. This prompts us to observe grade inflation rates for different groups of students. Figure 12 below shows the upper-class grade inflation rates for different ethnicity groups. The ethnicity groups are categorized as *"A" (Asian), "B" (Black), "M" (Mixed), "W" (White) and "O" (Other)* in the dataset.

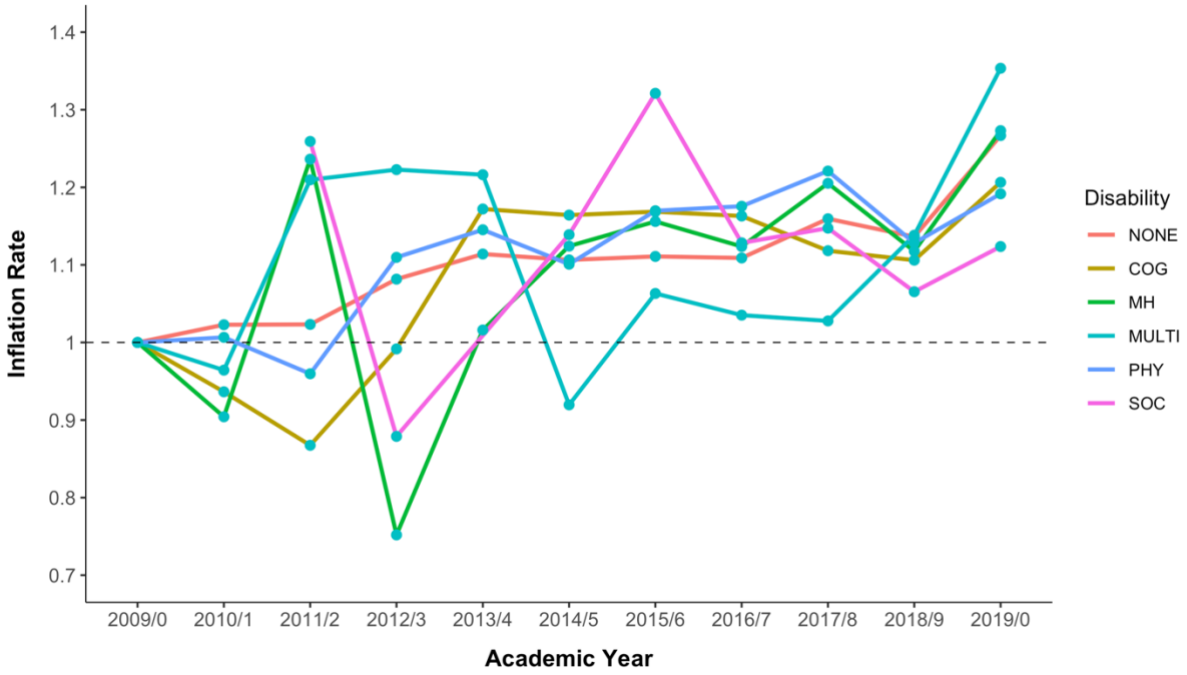
Fig.12: Inflation rate for different ethnicity groups



It can be observed that the black ethnicity group had the lowest upper-class inflation rate before 2013-14, after which they experienced approximately the highest upper-class inflation rate. This might suggest that LSE has been working on becoming a more inclusive community since 2013-14. There might be more support and resources provided for the ethnic minorities. As such, the ethnic minorities might attain a higher grade than their counterparts in 2009-10.

Similar observations can be made for the graph of inflation rates for students with disability. Figure 13 below illustrates the heterogeneous extent of grade inflation for students with disability. The disability statuses are categorized as *"NONE" (Not disabled), "COG" (Cognitive or learning difficulties), "MH" (Mental health condition), "PHY" (Sensory, medical or physical impairments), "SOC" (Social or communication impairment) and "MULTI" (Other or multiple impairments)* in the dataset.

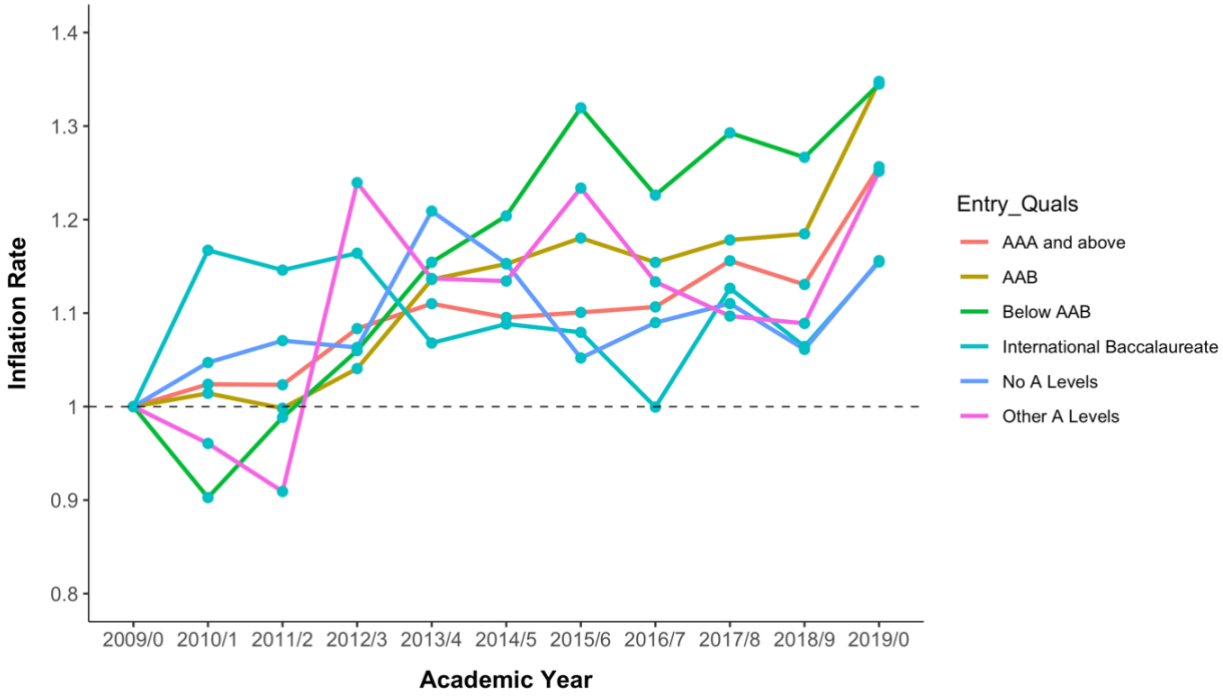
Fig.13: Inflation rate for different disability groups



There are more fluctuations in the trends as the data sample for students with disability is small. In general, we can observe that after 2014-15, the groups identified as “MH” and “PHY” have higher inflation rates than the group with no disability. This might suggest that LSE has developed better teaching and learning environments for the groups with disabilities. It is also observed that all groups experience a higher grade inflation rate in 2019-20, which might suggest that the online examination format is more accessible for groups with disabilities. However, as the data sample is small, there tends to be large standard errors in our estimation.

Figure 14 below shows the trends of upper-class inflation rates for students with different entry qualifications.

Fig.14: Inflation rate for different entry qualifications



From the graph, it can be observed that since 2013-14, the students who entered with A-Level grades “AAB” or “Below AAB” experienced higher grade inflation rates than students who entered with grades “AAA and above”. This might suggest that LSE has improved its quality of teaching such that students who were disadvantaged could perform better than predicted.

Graphs of first-class inflation rates for different groups of students are not included due to length constraint. Nonetheless, the graphs of first-class inflation rates exhibit similar trends. The analysis above suggests that students from less advantaged backgrounds are performing better than they are expected to in the baseline year. Hence, to a certain extent, the grade inflation rates might be attributed to the improvements in teaching and learning environments in the LSE.

Discussion

Evaluating grade inflation at LSE

The results presented in the previous section suggest that a certain level of grade inflation exists at LSE for both upper-class and first-class grade attainment. The grade inflation for

first-class grades is much higher than that for upper-class grade attainment at both module level and departmental level.

Results at both module level and departmental level show that grade inflation becomes drastic in academic year 2019-20. This is mostly likely due to the impact of Covid-19, which caused a complete change in course delivery and examination format. This led to several consequences that are potentially related to grade inflation. Firstly, students might perform better in open book examinations as they could refer to the materials and demonstrate a better understanding of the content. Secondly, to accommodate the different time zones that students were in, there was a 24-hour window to attempt and submit the answer. This might alleviate the effects of anxiety for students, which could influence the grades obtained in examinations. Lastly, as this was the first year experiencing such circumstances and using online lecture structure, the teaching quality and quality of examination contents might be compromised. It would be natural to hypothesise a relationship exists between Covid-19 and grade attainment. Further research is necessary to test this hypothesis and find out ways to improve quality control in exceptional circumstances.

The results also indicate that there are large fluctuations in both upper-class and first-class grade attainment across the academic years. For some modules, especially the qualitative modules, this could be due to the small pool of student population in each module. As such, any increase in actual number of grades attained would lead to a high increase in observed proportion. For other modules with a large student population, this implies greater quality control needs to be exercised within the modules.

From the previous LSE grade inflation study, some departments identified to have high degree inflation rates are identified to exhibit high grade inflation rates in this report as well. These departments include Accounting, Economics, Government, Mathematics, Sociology and Statistics. Although degree inflation is associated with Graduate degree, which aggregates the module grades over three or four years, the observation suggests that greater quality control needs to be exercised within these departments.

However, from the hypothetical group analysis and the inflation rates for different student subpopulations, we could observe that differences in abilities among students might explain parts of the grade inflation rates. In more recent years, there might be additional support and resources channelled to students from less advantaged backgrounds to help them achieve better results. From this point of view, the grade inflation rate maintained at a reasonable range might not be a worrying trend for the school. Having said that, constant monitoring of

grade inflation rates is necessary to prevent any sudden fluctuations or unwarranted increases in upper-class grade attainment.

Caveats in the method used to analyse grade inflation

As the data used in this report does not consist of the whole population in LSE, there might exist certain sampling bias when computing the observed proportion of upper-class grades awarded. In the qualitative modules, the sample size of the student population is generally small, leading to huge fluctuations in the unexplained changes. Subsequently, grade inflation might seem more severe in these modules than they experienced. It is difficult to distinguish between grade inflation and rise in upper-class grade attainment due to unrealistic observed proportion from a small student population. The results presented in the previous section are obtained by removing those modules with small student samples.

There is known to be a change in course structure in academic year 2017-18 and a change of examination format in 2019-20. These could have a sizable impact on grade attainment. However, the effects of these changes are hard to quantify in logistic regression models. It would also be difficult to compare results in these academic years with other years at the same level.

When aggregating the module grades at the departmental level and for each academic year, it should be cautioned that the size of the student population in each module could result in varying effects on the grade inflation at the department level. To illustrate, the grade inflation of a module with 500 students would have a larger impact than that of a module with 50 students on the departmental grade inflation.

There are several possible extensions of this study. Firstly, more module level factors could be included, such as the form of assessment (formative or summative). The results based on these module level factors could provide insights about ways to improve quality control in the modules. Secondly, the actual observed proportion of first- and upper second-class grades in each module and department should be included, as the proportions calculated from a sample of the student population can be biased. Moreover, the upper- and first-class grades could be further segregated into high upper-class grades and borderline upper-class grades according to examination marks. Regression models on module marks can also be introduced. Observing the change in grade attainment at a more granular level would allow the school to distinguish slight improvement in marks, which can lead to increases in borderline upper-class grades, from large unexplained increases in examination marks.

Lastly, the results from 2020-21 can be included, which would provide valuable insights into the effects of Covid-19 and online examination format on grade attainment. Using academic year 2018-19 as the baseline year to predict the Covid-19 grade trajectories would be helpful in understanding the potential factors of grade inflation.

Conclusion

The OfS reports and the follow-up study in LSE have revealed the trends in degree attainment in the sector and in LSE specifically. This report has taken their results a step further by examining the inflation in module grades, which is the key driver of inflation in degree attainment.

The results show that there exists some inflation in both upper-class grade and first-class grade attainment since 2009-10, especially for first-class grades. When the attainment gaps are removed, it seems that some, but not all of the grade inflation could be explained by the change in individual student characteristics. The problem of grade inflation has become more severe in academic year 2019-20, possibly due to Covid-19. Additional analysis about the Covid-19 school arrangements could shed light on how to improve quality control at exceptional circumstances.

This report also reveals that quantitative modules, generally experienced a higher grade inflation than qualitative modules. However, for first-class grades, qualitative modules have higher grade inflation rates in most academic years. This could be due to different structures of examination and coursework, which is an important area of improvement in quality control.

There are several shortcomings of the method used to investigate grade inflation in this report, including the specification of the model and the omission of important factors that might affect module grade attainment. Moving ahead, improvements on these aspects could shed light on the possible causes of grade inflation at the module level.

Annex A: Grade Classifications Summary

1. Table A1 presents the breakdown of the grade classifications from 2009-10 to 2019-20 for the sampling population (9177 students) considered in the model.

2. The table headings are as follows:
   Number (N): The number of grades attained by students in the sample.
   Percentage (%): The proportion of grades in the sample.
   Others: Including Lower Second-class, Third-class and Fail grades.

Table A1: Grade classification summary for academic years 2009-10 to 2019-20

Academic Year	First (N)	First (%)	Upper Second (N)	Upper Second (%)	Others (N)	Others (%)	Total (N)
2009-10	1463	17.2%	3618	42.4%	3446	40.4%	8527
2010-11	1584	17.8%	3918	43.9%	3420	38.3%	8922
2011-12	1590	17.8%	3913	43.9%	3419	38.3%	8922
2012-13	1792	20.3%	3980	45.1%	3052	34.6%	8824
2013-14	1980	23.0%	3997	46.4%	2639	30.6%	8616
2014-15	2074	22.7%	4330	47.5%	2721	29.8%	9125
2015-16	2267	24.0%	4484	47.6%	2678	28.4%	9429
2016-17	2402	25.4%	4516	47.8%	2532	26.8%	9450
2017-18	2748	28.6%	4481	46.6%	2390	24.8%	9619
2018-19	2988	29.6%	4732	46.9%	2364	23.4%	10084
2019-20	4450	42.5%	4585	43.8%	1431	13.7%	10466

Annex B: Results from statistical modelling

1. Table B1 presents the aggregate modelling results of Upper-Class grades for the student population considered in the model from 2009-10 to 2019-20. Table B2 presents the aggregate modelling results of First-Class grades for the student population considered in the model from 2009-10 to 2019-20.

2. The table headings are as follows:

   Observed proportion: The proportion of Upper-Class or First-Class grades attained in the sample.
   Observed change (pp): The change in proportion of Upper-Class or First-Class grades attained relative to 2009-10.
   Unexplained change (pp): The unexplained change in grade attainment relative to 2009-10 (calculated using the full fixed-effects model). A negative number of percentage points (pp) indicates the attainment in the academic year is beneath that of the aggregate proportion in 2009-10 with the effect of explanatory variables accounted for.
   Inflation rate: The indicator of grade inflation, calculated by comparing the observed proportion over the predicted proportion

Table B1: Aggregate modelling results for Upper-Class grades from 2009-10 to 2019-20

Academic Year	Observed proportion (%)	Observed change (pp)	Unexplained change (pp)	Inflation rate
2009-10 (ref.)	59.6	-	-	-
2010-11	61.7	2.1	1.1	1.019 (0.011)
2011-12	61.7	2.1	1.2	1.020 (0.011)
2012-13	65.4	5.8	4.8	1.078 (0.011)
2013-14	69.4	9.8	7.2	1.117 (0.011)
2014-15	70.2	10.6	6.8	1.107 (0.010)
2015-16	71.6	12.0	7.4	1.115 (0.010)
2016-17	73.2	13.6	7.4	1.112 (0.009)
2017-18	75.2	15.6	10.2	1.158 (0.009)
2018-19	76.6	17.0	9.1	1.134 (0.008)
2019-20	86.3	26.7	18.0	1.263 (0.009)

Table B2: Aggregate modelling results for First-Class grades from 2009-10 to 2019-20

Academic Year	Observed proportion	Observed change (pp)	Unexplained change (pp)	Inflation rate
2009-10 (ref.)	17.2	-	-	-
2010-11	17.8	0.6	-0.3	0.986 (0.029)
2011-12	17.8	0.7	-0.2	0.990 (0.031)
2012-13	20.3	3.2	2.2	1.123 (0.032)
2013-14	23.0	5.8	4.2	1.223 (0.031)
2014-15	22.7	5.6	3.7	1.196 (0.031)
2015-16	24.0	6.9	4.2	1.211 (0.029)
2016-17	25.4	8.3	4.8	1.233 (0.029)
2017-18	28.6	11.4	7.8	1.377 (0.033)
2018-19	29.6	12.5	6.2	1.262 (0.024)
2019-20	42.5	25.4	18.1	1.743 (0.030)

3. Table B3 and B4 present the results for the hypothetical group. The hypothetical group is formed by setting the individual characteristics to the highest attaining categories, which are identified as the most positive coefficients in the full fixed-effects model. The coefficient estimates are shown in Annex D.

4. The following characteristics are transformed:
   Sex: Set to “Male”
   Disability: Set to “No disability”
   Ethnicity: Set to “White”
   POLAR Quintile: Set to “Quintile 2”

Table B3: Aggregate modelling results for Upper-Class grades attained by hypothetical group

Academic Year	Observed proportion	Observed change (pp)	Unexplained change (pp)	Inflation rate
2009/0	59.6	-	-	-
2010/1	61.7	2.1	-4.2	0.937
2011/2	61.7	2.1	-4.2	0.937
2012/3	65.4	5.8	-0.7	0.989
2013/4	69.4	9.8	1.8	1.027
2014/5	70.2	10.6	1.6	1.023
2015/6	71.6	12.0	2.3	1.033
2016/7	73.2	13.6	2.4	1.034
2017/8	75.2	15.6	5.2	1.074
2018/9	76.6	17.0	3.9	1.054
2019/0	86.3	26.7	12.6	1.171

Table B4: Aggregate modelling results for First-Class grades attained by hypothetical group

Academic Year	Observed proportion	Observed change (pp)	Unexplained change (pp)	Inflation rate
2009/0	17.2	-	-	-
2010/1	17.8	0.6	-4.7	0.790
2011/2	17.8	0.7	-4.6	0.795
2012/3	20.3	3.2	-2.1	0.906
2013/4	23.0	5.8	0.0	0.999
2014/5	22.7	5.6	-0.3	0.986
2015/6	24.0	6.9	0.1	1.004
2016/7	25.4	8.3	0.7	1.027
2017/8	28.6	11.4	3.7	1.147
2018/9	29.6	12.5	1.7	1.062
2019/0	42.5	25.4	13.4	1.462

5. Table B5 and B7 present the modules with consistently high Upper-Class or First-Class inflation rate, ranked by three approaches: 1) aggregate inflation rate from 2010-11 to 2019-20, 2) frequency of exceeding the module departments' inflation rate, and 3) frequency exceeding the sampling population's aggregate inflation rate. Table B6 and B8 present the inflation rates in each year for the top 10 modules with highest aggregate inflation rates.

Table B5: Top 10 modules with consistently high Upper-Class inflation rate

Highest aggregate IR	Exceeding Dep IR most frequently	Exceeding Overall IR most frequently
FM300 (1.304)	FM212 (9)	GY100 (9)
EC201 (1.295)	GV100 (9)	LL104 (9)
MA107 (1.289)	GY100 (9)	LL106 (9)
EC102 (1.283)	LL106 (9)	EC100 (8)
EC220 (1.270)	AC100 (8)	EC102 (8)
ST202 (1.262)	AN101 (8)	EC210 (8)
MA100 (1.258)	HY116 (8)	GV100 (8)
EC100 (1.250)	LL104 (8)	MA100 (8)
ST102 (1.238)	MA100 (8)	MA103 (8)
MA203 (1.237)	PH101 (8)	MA203 (8)

Table B6: Grade inflation rate for Upper-Class grades for 10 modules (standard errors in parentheses)

Module	2010/1	2011/2	2012/3	2013/4	2014/5	2015/6	2016/7	2017/8	2018/9	2019/0	Mean
FM300	0.968	0.822	1.062	0.937	0.829	0.991	1.361	1.731	1.721	2.100	1.187
	(0.17)	(0.135)	(0.135)	(0.13)	(0.116)	(0.11)	(0.121)	(0.11)	(0.097)	(0.098)	
EC201	1.042	0.910	1.055	1.064	1.077	1.365	1.283	1.581	1.330	1.826	1.227
	(0.08)	(0.08)	(0.089)	(0.09)	(0.089)	(0.09)	(0.081)	(0.07)	(0.065)	(0.071)	
MA107	0.541	1.527	1.345	1.664	1.211	1.213	0.949	1.120	1.539	1.610	1.217
	(0.08)	(0.1)	(0.105)	(0.1)	(0.098)	(0.09)	(0.102)	(0.09)	(0.088)	(0.082)	
EC102	0.597	0.987	1.325	1.550	1.293	1.319	1.185	1.270	1.528	1.803	1.239
	(0.06)	(0.064)	(0.059)	(0.06)	(0.061)	(0.06)	(0.065)	(0.06)	(0.065)	(0.065)	
EC220	0.799	1.200	1.232	1.599	1.099	1.150	1.353	1.535	1.113	1.522	1.237
	(0.1)	(0.12)	(0.094)	(0.09)	(0.102)	(0.1)	(0.094)	(0.08)	(0.093)	(0.084)	
ST202	1.160	0.882	1.081	1.025	1.378	0.787	1.352	1.554	1.235	1.570	1.175
	(0.18)	(0.134)	(0.147)	(0.16)	(0.128)	(0.15)	(0.115)	(0.09)	(0.107)	(0.097)	
MA100	1.268	1.266	1.400	1.305	1.483	1.313	1.155	1.112	1.041	1.274	1.256
	(0.07)	(0.068)	(0.064)	(0.07)	(0.069)	(0.07)	(0.067)	(0.07)	(0.063)	(0.072)	
EC100	1.020	0.617	1.153	1.371	1.094	1.225	1.063	1.445	1.471	1.494	1.162
	(0.1)	(0.108)	(0.109)	(0.13)	(0.098)	(0.06)	(0.078)	(0.07)	(0.072)	(0.066)	
ST102	1.093	1.048	0.901	1.348	1.290	1.084	1.243	1.334	1.219	1.797	1.216
	(0.07)	(0.067)	(0.063)	(0.07)	(0.063)	(0.06)	(0.061)	(0.05)	(0.058)	(0.054)	
MA203	1.058	1.340	1.246	1.503	1.186	0.711	1.006	0.885	1.207	1.933	1.166
	(0.15)	(0.147)	(0.159)	(0.14)	(0.174)	(0.14)	(0.17)	(0.16)	(0.157)	(0.122)	

Table B7: Top 10 modules with consistently high First-Class inflation rate

Highest aggregate IR	Exceeding Dep IR most frequently	Exceeding Overall IR most frequently
EC325 (1.400)	FM212 (10)	LL205 (9)
LL205 (1.395)	HY113 (8)	EC315 (6)
EC220 (1.337)	SO110 (8)	FM212 (6)
ST102 (1.315)	AC211 (7)	MA100 (6)
ST202 (1.311)	AN100 (7)	MA103 (6)

MA100 (1.306)	EH240 (7)	MA203 (6)
ST107 (1.302)	GV100 (7)	ST102 (6)
MA209 (1.286)	LL104 (7)	ST202 (6)
MA203 (1.261)	LL106 (7)	AC330 (5)
FM212 (1.250)	LL205 (7)	EC220 (5)

Table B8: Grade inflation rate for First-Class grades for 10 modules (standard errors in parentheses)

Module	2010/1	2011/2	2012/3	2013/4	2014/5	2015/6	2016/7	2017/8	2018/9	2019/0	Mean
EC325	0.431	1.068	1.980	0.701	1.071	1.549	0.775	1.355	1.034	2.894	1.128
	(0.19)	(0.239)	(0.235)	(0.19)	(0.233)	(0.22)	(0.139)	(0.19)	(0.172)	(0.201)	
LL205	1.427	1.156	0.679	1.553	1.301	1.235	1.785	1.602	1.615	1.863	1.374
	(0.41)	(0.341)	(0.274)	(0.35)	(0.374)	(0.39)	(0.472)	(0.47)	(0.462)	(0.482)	
EC220	0.767	0.898	1.380	1.986	1.020	1.241	1.191	1.973	0.782	1.873	1.235
	(0.15)	(0.199)	(0.179)	(0.21)	(0.175)	(0.17)	(0.177)	(0.19)	(0.143)	(0.198)	
ST102	0.679	0.819	0.741	1.656	1.382	1.136	1.440	1.598	1.449	2.190	1.228
	(0.09)	(0.094)	(0.086)	(0.12)	(0.109)	(0.11)	(0.099)	(0.11)	(0.109)	(0.115)	
ST202	1.132	0.720	1.045	0.797	1.536	0.709	1.493	1.858	1.285	1.644	1.158
	(0.25)	(0.196)	(0.222)	(0.25)	(0.232)	(0.2)	(0.184)	(0.18)	(0.183)	(0.156)	
MA100	1.281	1.085	1.475	1.230	1.923	1.338	1.210	1.306	1.090	1.121	1.288
	(0.11)	(0.109)	(0.117)	(0.12)	(0.119)	(0.12)	(0.112)	(0.11)	(0.106)	(0.112)	
ST107	0.897	1.242	1.209	1.749	1.168	1.477	1.026	1.179	1.342	1.536	1.261
	(0.12)	(0.138)	(0.148)	(0.14)	(0.12)	(0.14)	(0.128)	(0.12)	(0.128)	(0.116)	
MA209	0.900	1.068	1.950	0.680	0.893	0.583	1.015	1.471	1.641	2.903	1.168
	(0.2)	(0.226)	(0.208)	(0.15)	(0.234)	(0.18)	(0.238)	(0.25)	(0.268)	(0.224)	
MA203	0.605	1.202	1.308	1.390	1.057	0.756	0.795	0.752	1.373	2.757	1.091
	(0.18)	(0.23)	(0.238)	(0.22)	(0.273)	(0.19)	(0.224)	(0.23)	(0.251)	(0.228)	
FM212	0.693	0.896	1.192	1.343	1.057	0.961	1.256	1.697	2.267	3.595	1.332
	(0.1)	(0.123)	(0.124)	(0.14)	(0.115)	(0.13)	(0.13)	(0.16)	(0.309)	(0.381)	

6. Table B9 and B11 present the departments with consistently high Upper-Class or First-Class inflation rate, ranked by two approaches: 1) aggregate inflation rate from 2010-11 to 2019-20, and 2) frequency exceeding the sampling population's aggregate inflation

rate. Table B10 and B12 present the inflation rates in each year for the top 10 departments with highest aggregate inflation rates.

Table B9: Top 10 departments with consistently high Upper-Class inflation rate

Top 10 Departments with consistently high Inflation rate (IR)	
Highest aggregate IR	**Exceeding Overall IR most frequently**
Economics (1.233)	Economics (9)
Mathematics (1.231)	Mathematics (9)
Finance (1.204)	Statistics (8)
Statistics (1.196)	Finance (7)
Accounting (1.148)	Law (6)
International Relations (1.144)	Accounting (5)
Sociology (1.132)	Psychology (5)
Social Policy (1.123)	Social Policy (5)
Philosophy (1.119)	Economic History (4)
Economic History (1.119)	International Development (4)

Table B10: Grade inflation rate for Upper-Class grades for 10 departments

Dep	2010/1	2011/2	2012/3	2013/4	2014/5	2015/6	2016/7	2017/8	2018/9	2019/0	Mean
EC	0.970	1.001	1.165	1.236	1.156	1.180	1.159	1.350	1.284	1.614	1.200
MA	1.145	1.229	1.249	1.232	1.252	1.096	1.128	1.052	1.161	1.467	1.196
FM	0.988	1.057	1.150	1.132	1.019	0.954	1.050	1.395	1.157	1.355	1.117
ST	1.115	1.072	1.064	1.197	1.130	1.075	1.219	1.231	1.167	1.426	1.165
AC	1.085	1.128	1.131	1.109	1.060	1.140	1.054	1.181	1.063	1.247	1.119
IR	0.930	0.896	1.024	1.098	1.202	1.106	1.168	1.189	1.099	1.128	1.079
SO	0.932	0.978	1.077	1.108	1.100	1.129	1.136	1.113	1.119	1.218	1.088
SP	0.942	0.935	0.982	1.153	1.102	1.297	1.257	1.251	1.086	0.994	1.092
PH	1.017	0.929	1.003	1.045	1.084	1.126	1.082	1.064	1.175	1.171	1.067
EH	1.008	0.981	1.103	1.008	1.101	1.116	1.028	1.165	1.164	1.202	1.085

Table B11: Top 10 departments with consistently high Upper-Class inflation rate

Top 10 Departments with consistently high First-Class Inflation rate	
Highest aggregate IR	**Exceeding Overall IR most frequently**
Government (1.448)	Economics (9)
Sociology (1.433)	Mathematics (9)
International Relations (1.369)	Statistics (8)
Anthropology (1.367)	Finance (7)
Law (1.360)	Law (6)
Economics (1.352)	Accounting (5)
Economic History (1.331)	Psychology (5)
Social Policy (1.324)	Social Policy (5)
Geography (1.320)	Economic History (4)
International History (1.317)	International Development (4)

Table B12: Grade inflation rate for First-Class grades for 10 departments

Dep	2010/1	2011/2	2012/3	2013/4	2014/5	2015/6	2016/7	2017/8	2018/9	2019/0	Mean
GV	0.757	0.811	1.261	1.119	1.551	1.532	1.225	1.449	1.177	1.664	1.218
SO	0.677	0.918	1.279	1.464	1.354	1.116	1.751	1.630	1.154	1.535	1.244
IR	0.730	0.721	0.759	1.057	1.533	1.364	1.416	1.535	1.184	1.402	1.123
AN	1.281	0.945	0.881	0.985	1.078	1.170	1.209	1.416	1.660	1.875	1.216
LL	1.268	1.252	0.979	1.123	1.238	1.280	1.398	1.501	1.590	1.819	1.326
EC	0.868	0.816	1.227	1.387	1.090	1.117	1.082	1.586	1.475	2.400	1.245
EH	0.927	0.860	1.509	1.213	1.317	1.318	1.209	1.278	1.458	1.842	1.265
SP	0.893	0.724	0.949	1.019	1.015	1.665	1.596	1.841	1.241	1.008	1.145
GY	0.786	0.940	1.150	1.349	1.330	1.510	1.180	1.392	1.185	1.583	1.216
HY	0.921	0.901	0.932	1.288	1.136	1.192	1.336	1.351	1.365	1.578	1.180

Annex C: Description of variables used in the models and model fit statistics

1. This annex provides a summary of the variables used in the model, and the categories within those variables.

2. The simplified variable names used in the model equations are listed in parentheses.

Table C1: Variables used in the fixed- and mixed- effects logistic regression models in this report

Model Variables	Description
Year of attainment (Year)	Academic year of grade attainment: *2009-10 (ref.)* *2010-11* *2011-12* *2012-13* *2013-14* *2014-15* *2015-16* *2016-17* *2017-18* *2018-19* *2019-20*
Qualifications on entry (Entry_qual)	Entry qualification of student: AAA and above (ref.) AAB Below AAB International Baccalaureate No A Levels Other A Levels
Ethnicity (Ethn)	Ethnicity of student: *White(ref.)*

	Asian *Black* *Mixed* *Other* *Prefer not to say* *Unknown*
Disability (Disb)	Declared disability status of student: *No disability (ref.)* *Cognitive or learning difficulties* *Mental health condition* *Social or communication impairment* *Sensory, medical, or physical impairment* *Other or multiple impairments*
Gender (Gend)	Gender of student: *Male (ref.)* *Female* *Other*
Age (Age)	Age on entry into university: *Under 21 (Young) (ref.)* *Over 21 (Mature)*
Participation of Local Areas (POLAR4) quintile (POLAR)[1]	Young participation quintile of student: *Quintile 1* *Quintile 2* *Quintile 3* *Quintile 4* *Quintile 5 (ref.)* *Unknown*
Tracking Underrepresentation by Area (TUNDRA) quintile (TUNDRA)[2]	Young participation quintile of student: *Quintile 1* *Quintile 2* *Quintile 3*

[1] Check POLAR 4 classification standards at *https://www.officeforstudents.org.uk/data-and-analysis/young-participation-by-area/*

[2] Check TUNDRA classification standards at *https://www.officeforstudents.org.uk/data-and-analysis/young-participation-by-area/about-tundra/*

	Quintile 4 *Quintile 5 (ref.)* *Unknown*
Index of multiple deprivation (IMD) quintile (IMD)[3]	Deprived quintile of student: *Quintile 1* *Quintile 2* *Quintile 3* *Quintile 4* *Quintile 5 (ref.)* *Unknown*
ACORN group (ACORN)[4]	Consumption status classification of student: *Group A Lavish lifestyle (ref.)* *Group B Executive wealth* *Group C Mature money* *Group D City sophisticates* *Group E Career climbers* *Group F Countryside communities* *Group G Successful suburbs* *Group H Steady neighbourhoods* *Group I Comfortable seniors* *Group J Starting out* *Group K Student life* *Group L Modest means* *Group M Striving families* *Group N Poorer pensioners* *Group O Young hardship* *Group P Struggling estates* *Group Q Difficult circumstances*
Module Level (Mod_lvl)	Module level: *Level I (ref.)* *Level II*

[3] Check IMD classification standards at *https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019*

[4] Check ACORN classification standards at *https://www.caci.co.uk/sites/default/files/resources/Acorn%20User%20Guide%202020.pdf*

	Level III *Level IV*
Module Credit (Mod_cred)	Module credit: *0 unit* *0.5 unit* *1 unit (ref.)*
Undergraduate bursary (Burs)	Attainment of undergraduate bursary: *None* *In year of module grade attainment* *In other year*
Module Student Count (Stu_count)	Number of students in the module (Continuous variable)
Programme department (Prog_dpt)	Programme department of student: *Omitted (18 departments in total)*
Module Code (Code)	Module Code: *Omitted (690 modules in total)*

3. Model fit statistics for 2 fixed-effects models and 3 mixed-effects models are shown for Upper-Class grades in Table C2, and First-Class grades in Table C3.

Table C2: Model fit statistics for the models for Upper-Class grades attainment

Statistic	-2logLikelihood	Akaike Information Criterion	Baysian Information Criterion
Full fixed-effects model (Equation D1)	100305	101856.7	109254
Simplified fixed-effects (Equation D2)	106647	108092.6	114984.7
Mixed-effects model 1 (Equation E1)	108527	108597	108931
Mixed-effects model 2 (Equation E2)	107564	107616	107864
Mixed-effects model 3 (Equation E3)	108366	108438	108781
Mixed-effects model 4 (Equation E4)	107016.1	107196.1	108054.1

Table C2: Model fit statistics for the models for Upper-Class grades attainment

Statistic	-2logLikelihood	Akaike Information Criterion	Baysian Information Criterion
Full fixed-effects model (Equation D1)	98748	100299.6	107696.9
Simplified fixed-effects (Equation D2)	101793	103238.8	110130.8
Mixed-effects model 1 (Equation E1)	103656	103726	104059
Mixed-effects model 2 (Equation E2)	102767	102819	103067
Mixed-effects model 3 (Equation E3)	103555	103627	103970
Mixed-effects model 4 (Equation E4)	102334.4	102514.4	103372.4

Annex D: Fixed-effects logistic regression model used in this report

1. This annex details the full and simplified logistic regression models used to predict the probability of student *i* attaining an upper-class grade or a first-class grade, after accounting for the effect of the explanatory variables in Annex C above.

2. There are 2250 students taking one of the 106 modules classified as *"Unknown"* type in the variable *Module_in_home_department* (indicating whether the students are taking a module in their home department), which leads to multi-collinearity with the specific module codes. The variable is excluded from the models.

3. The specifications of the 'full' and 'simplified' models are displayed in Equations D1 and D2 respectively.

Equation D1: Full fixed-effects logistic regression model

$$\textit{Attained an upper class grade OR first class grade} \sim \textit{Binomial}\ (n_i, \pi_i)$$

$$\begin{aligned} logit(\pi_i) = \beta_0 &+ \sum_{Year=1}^{11} \beta_{Year} X_{Year,i} + \sum_{Entry\_qual=8}^{22} \beta_{Entry\_qual} X_{Entry_{qual},i} \\ &+ \sum_{Ethn=23}^{29} \beta_{Ethn} X_{Ethn,i} \\ &+ \sum_{Disb=30}^{35} \beta_{Disb} X_{Disb,i} \\ &+ \sum_{Gend=36}^{38} \beta_{Gend} X_{Gend,i} + \sum_{Age=39}^{40} \beta_{Age} X_{Age,i} + \sum_{POLAR4=41}^{46} \beta_{POLAR4} X_{POLAR4,i} \\ &+ \sum_{TUNDRA=47}^{52} \beta_{TUNDRA} X_{TUNDRA,i} + \sum_{IMD=53}^{58} \beta_{IMD} X_{IMD,i} \\ &+ \sum_{ACORN=59}^{75} \beta_{ACORN} X_{ACORN,i} \\ &+ \sum_{Mod\_lvl=76}^{79} \beta_{Mod\_lvl} X_{Mod_{lvl},i} \\ &+ \sum_{Mod\_cred=80}^{82} \beta_{Mod\_cred} X_{Mod_{cred},i} \\ &+ \sum_{Burs=83}^{85} \beta_{Burs} X_{Burs,i} + \sum_{Prog\_dpt=86}^{103} \beta_{Prog\_dpt} X_{Prog\_dpt,i} \\ &+ \sum_{Code=104}^{793} \beta_{Code} X_{Code,i} + \beta_{Stu\_count} Stu\_count \end{aligned}$$

Equation D2: Simplified fixed-effects logistic regression model

$$\textit{Attained an Upper-class grade OR First class grade} \sim \textit{Binomial}\ (n_i, \pi_i)$$

$$logit(\pi_i) = \beta_0 + \sum_{Year=1}^{11} \beta_{Year} X_{Year,i} + \sum_{Entry\_qual=8}^{22} \beta_{Entry\_qual} X_{Entry_{qual},i} \sum_{Ethn=23}^{29} \beta_{Ethn} X_{Ethn,i} + \sum_{Disb=30}^{35} \beta_{Disb} X_{Disb,i} + \sum_{Gend=36}^{38} \beta_{Gend} X_{Gend,i} + \sum_{Age=39}^{40} \beta_{Age} X_{Age,i} + \sum_{POLAR4=41}^{46} \beta_{POLAR4} X_{POLAR4,i} + \sum_{Code=47}^{736} \beta_{Code} X_{Code,i}$$

4. The βs represent the fixed-effects coefficients for the categorical variables in the model, and the Xs (0 or 1) are binary indicators representing whether an individual possess the categories of the characteristics (variable names and categories are listed in Annex C).
5. Estimates of the variable coefficients, their standard errors (SE) and p-values for the full model are shown for Upper-Class grades in Table D1, and for First-Class grades in Table D2. The estimates for the simplified model are omitted due to space constraint.

Table D1: Variable Coefficient estimates of Model Equation D1 for Upper-class grade attainment

Effects		Full model estimate	Full model SE	Full model p-value
Intercept	Intercept	0.815	0.095	<0.0001
Year	*2009-10 (ref.)*	-	-	-
	2010-11	0.063	0.036	0.077
	2011-12	0.068	0.037	0.070
	2012-13	0.266	0.036	<0.0001
	2013-14	0.427	0.037	<0.0001
	2014-15	0.423	0.037	<0.0001
	2015-16	0.475	0.037	<0.0001
	2016-17	0.495	0.038	<0.0001
	2017-18	0.687	0.038	<0.0001
	2018-19	0.661	0.038	<0.0001
	2019-20	1.390	0.042	<0.0001
Entry_qual	*AAA and above (ref.)*	-	-	-
	AAB	-0.606	0.023	<0.0001
	Below AAB	-0.950	0.042	<0.0001

	International Baccalaureate	0.080	0.063	0.202
	No A Levels	0.006	0.048	0.907
	Other A Levels	-0.312	0.068	<0.0001
Ethn	*White (ref.)*	-	-	-
	Asian	-0.337	0.020	<0.0001
	Black	-0.497	0.038	<0.0001
	Mixed	-0.103	0.033	0.002
	Other	-0.372	0.051	<0.0001
	Prefer not to say	-0.185	0.068	0.006
Disb	*No disability (ref.)*	-	-	-
	Cognitive or learning difficulties	-0.104	0.043	0.015
	Mental health condition	-0.258	0.067	0.000
	Social or communication impairment	-0.044	0.216	0.839
	Sensory, medical, or physical impairment	-0.050	0.065	0.439
	Other or multiple impairments	-0.173	0.073	0.018
Gend	*Male (ref.)*	-	-	-
	Female	-0.006	0.017	0.713
	Other	-0.501	0.221	0.023
Age	*Young (ref.)*	-	-	-
	Mature	-0.295	0.062	<0.0001
Mod_cred	*1 unit (ref.)*	-	-	-
	0.5 unit	-0.019	0.896	0.983
	0 unit	-6.886	0.521	<0.0001
Mod_lvl	*Level I (ref.)*	-	-	-
	Level II	0.304	0.053	<0.0001
	Level III	0.617	0.057	<0.0001
	Level IV	1.273	0.536	0.018
Burs	*None (ref.)*	-	-	-
	In year	-0.068	0.020	0.001
	In other year	-0.096	0.031	0.002
POLAR	*Quintile 5 (ref.)*	-	-	-
	Quintile 1	0.084	0.053	0.111
	Quintile 2	0.136	0.037	0.000
	Quintile 3	0.075	0.029	0.010
	Quintile 4	0.050	0.022	0.027
	Unknown	0.028	0.098	0.774
TUNDRA	*Quintile 5 (ref.)*	-	-	-
	Quintile 1	0.151	0.056	0.007
	Quintile 2	-0.068	0.037	0.065

	Quintile 3	-0.026	0.029	0.369
	Quintile 4	-0.042	0.023	0.065
	Unknown	-0.131	0.028	<0.0001
IMD	*Quintile 5 (ref.)*	-	-	-
	Quintile 1	-0.136	0.041	0.001
	Quintile 2	-0.088	0.030	0.003
	Quintile 3	-0.026	0.025	0.296
	Quintile 4	-0.031	0.022	0.166
	Unknown	0.318	0.239	0.184
ACORN	*Group A*	-	-	-
	Group B	0.049	0.029	0.088
	Group C	0.066	0.040	0.101
	Group D	-0.025	0.039	0.533
	Group E	0.048	0.042	0.258
	Group F	0.162	0.056	0.004
	Group G	-0.003	0.038	0.928
	Group H	-0.041	0.040	0.310
	Group I	0.075	0.125	0.547
	Group J	-0.011	0.060	0.854
	Group K	0.217	0.082	0.009
	Group L	-0.097	0.052	0.062
	Group M	-0.027	0.062	0.663
	Group N	-0.159	0.111	0.149
	Group O	-0.136	0.079	0.084
	Group P	-0.100	0.047	0.035
	Group Q	-0.087	0.093	0.349
	Group R	-0.059	0.107	0.584
	Unknown	-0.281	0.212	0.184
Stu_count		2.95E-04	0.0002	0.138
Prog_dpt	*EC (ref.)*	-	-	-
(Five departments listed for reference)	*AC*	-0.859	0.043	<0.0001
	IR	-0.578	0.077	<0.0001
	LL	-0.310	0.087	<0.001
	MA	-0.317	0.036	<0.0001
	ST	-0.553	0.044	<0.0001
Code (Omitted)*				

* Fixed effect coefficients estimate for Module Codes are omitted due to space constraint.

Table D2: Variable Coefficient estimates of Model Equation D1 for First-class grade attainment

Effects		Full model estimate	Full model SE	Full model p-value
Intercept	Intercept	-0.737	0.102	<0.0001
Year	*2009-10 (ref.)*	-	-	-
	2010-11	-0.020	0.043	0.639
	2011-12	-0.015	0.044	0.741
	2012-13	0.166	0.042	<0.0001
	2013-14	0.304	0.042	<0.0001
	2014-15	0.288	0.042	<0.0001
	2015-16	0.322	0.041	<0.0001
	2016-17	0.367	0.041	<0.0001
	2017-18	0.558	0.041	<0.0001
	2018-19	0.494	0.041	<0.0001
	2019-20	1.154	0.041	<0.0001
Entry_qual	*AAA and above (ref.)*	-	-	-
	AAB	-0.644	0.027	<0.0001
	Below AAB	-1.035	0.063	<0.0001
	International Baccalaureate	0.131	0.054	0.016
	No A Levels	0.021	0.045	0.633
	Other A Levels	-0.266	0.066	<0.0001
Ethn	*White (ref.)*	-	-	-
	Asian	-0.321	0.020	<0.0001
	Black	-0.562	0.047	<0.0001
	Mixed	-0.064	0.033	0.051
	Other	-0.292	0.056	<0.0001
	Prefer not to say	-0.314	0.069	<0.0001
Disb	*No disability (ref.)*	-	-	-
	Cognitive or learning difficulties	0.074	0.042	0.078
	Mental health condition	-0.041	0.067	0.545
	Social or communication impairment	-0.511	0.224	0.022
	Sensory, medical, or physical impairment	0.155	0.062	0.013
	Other or multiple impairments	-0.086	0.077	0.263
Gend	*Male (ref.)*	-	-	-
	Female	-0.094	0.017	<0.0001

	Other	-0.398	0.249	0.111
Age	*Young (ref.)*	-	-	-
	Mature	0.026	0.064	0.691
Mod_cred	*1 unit (ref.)*	-	-	-
	0.5 unit	-0.330	0.966	0.732
	0 unit	-6.163	0.695	<0.0001
Mod_lvl	*Level I (ref.)*	-	-	-
	Level II	0.358	0.050	<0.0001
	Level III	0.617	0.053	<0.0001
	Level IV	1.446	0.279	<0.0001
Burs	*None (ref.)*	-	-	-
	In year	-0.088	0.020	<0.0001
	In other year	-0.078	0.033	0.018
POLAR	*Quintile 5 (ref.)*	-	-	-
	Quintile 1	0.091	0.054	0.092
	Quintile 2	0.118	0.037	0.001
	Quintile 3	0.086	0.029	0.003
	Quintile 4	0.045	0.023	0.046
	Unknown	-0.303	0.098	0.002
TUNDRA	*Quintile 5 (ref.)*	-	-	-
	Quintile 1	0.046	0.056	0.406
	Quintile 2	-0.094	0.038	0.013
	Quintile 3	0.038	0.029	0.199
	Quintile 4	-0.036	0.023	0.121
	Unknown	0.005	0.028	0.844
IMD	*Quintile 5 (ref.)*	-	-	-
	Quintile 1	-0.064	0.044	0.149
	Quintile 2	-0.052	0.031	0.090
	Quintile 3	0.025	0.025	0.305
	Quintile 4	0.014	0.022	0.516
	Unknown	0.525	0.305	0.085
ACORN	*Group A*	-	-	-
	Group B	0.020	0.028	0.490
	Group C	0.110	0.038	0.004
	Group D	0.013	0.039	0.738
	Group E	0.001	0.043	0.974
	Group F	0.092	0.053	0.082
	Group G	-0.060	0.038	0.117
	Group H	-0.137	0.041	0.001
	Group I	-0.207	0.116	0.075
	Group J	0.073	0.060	0.224
	Group K	0.039	0.080	0.625

	Group L	-0.135	0.056	0.015
	Group M	-0.024	0.064	0.706
	Group N	-0.068	0.119	0.570
	Group O	-0.056	0.085	0.510
	Group P	-0.138	0.051	0.007
	Group Q	-0.113	0.105	0.284
	Group R	-0.012	0.110	0.917
	Unknown	-0.073	0.283	0.798
Stu_count		-3.81E-05	0.0002	0.852
Prog_dpt *(Five departments listed for reference)*	*EC (ref.)*	-	-	-
	AC	-0.767	0.045	<0.0001
	IR	-0.544	0.075	<0.0001
	LL	-0.390	0.091	<0.0001
	MA	-0.283	0.037	<0.0001
	ST	-0.519	0.044	<0.0001
Code (Omitted)*				

* Fixed effect coefficients estimate for Module Codes are omitted due to space constraint.

Annex E: Mixed-effects logistic regression model used in this report

1. This annex details the mixed-effects logistic regression models used to predict the probability of student i attaining an upper-class grade or a first-class grade, after accounting for the effect of the explanatory variables in Annex C.

2. The specification of the model with a random intercept for each module is displayed in Equation E1. The model with a random intercept for each module, and for academic years that are nested within modules is presented in Equation E2. The model with a random intercept for each department, and for modules that are nested within departments is presented in Equation E3. The model with a random module intercept and a random year coefficient for each module is presented in Equation E4.

Equation E1: Mixed-effects logistic regression model 1

$$\textit{Attained an Upper-class grade OR First-class grade} \sim \textit{Binomial}\ (n_i, \pi_i)$$

$$\begin{aligned} logit(\pi_{ij}) = {} & \beta_{0j} + u_{0j} \\ & + \sum_{Year=1}^{11} \beta_{Year} X_{Year,ij} + \sum_{Entry\_qual=8}^{22} \beta_{Entry\_qual} X_{Entry_{qual},ij} \sum_{Ethn=23}^{29} \beta_{Ethn} X_{Ethn,ij} \\ & + \sum_{Disb=30}^{35} \beta_{Disb} X_{Disb,ij} \\ & + \sum_{Gend=36}^{38} \beta_{Gend} X_{Gend,ij} + \sum_{Age=39}^{40} \beta_{Age} X_{Age,ij} + \sum_{POLAR4=41}^{46} \beta_{POLAR4} X_{POLAR4,ij} \end{aligned}$$

Equation E2: Mixed-effects logistic regression model 2

$$\textit{Attained an Upper-class grade OR First-class grade} \sim \textit{Binomial}\ (n_i, \pi_i)$$

$$\begin{aligned} logit(\pi_{ij}) = {} & \beta_{0j} + u_{0j} + u_{Yj} \\ & + \sum_{Year=1}^{11} \beta_{Year} X_{Year,ij} + \sum_{Entry\_qual=8}^{22} \beta_{Entry\_qual} X_{Entry_{qual},ij} \sum_{Ethn=23}^{29} \beta_{Ethn} X_{Ethn,ij} \\ & + \sum_{Disb=30}^{35} \beta_{Disb} X_{Disb,ij} \\ & + \sum_{Gend=36}^{38} \beta_{Gend} X_{Gend,ij} + \sum_{Age=39}^{40} \beta_{Age} X_{Age,ij} + \sum_{POLAR4=41}^{46} \beta_{POLAR4} X_{POLAR4,ij} \end{aligned}$$

Equation E3: Mixed-effects logistic regression model 3

Attained an Upper-class grade OR First-class grade ~ Binomial (n_i, π_i)

$$logit(\pi_{ij}) = \beta_{0j} + u_{0j} + u_{Dj} + \sum_{Year=1}^{11} \beta_{Year} X_{Year,ij} + \sum_{Entry\_qual=8}^{22} \beta_{Entry\_qual} X_{Entry_{qual},ij} \sum_{Ethn=23}^{29} \beta_{Ethn} X_{Ethn,ij} + \sum_{Disb=30}^{35} \beta_{Disb} X_{Disb,ij} + \sum_{Gend=36}^{38} \beta_{Gend} X_{Gend,ij} + \sum_{Age=39}^{40} \beta_{Age} X_{Age,ij} + \sum_{POLAR4=41}^{46} \beta_{POLAR4} X_{POLAR4,ij}$$

Equation E4: Mixed-effects logistic regression model 4

Attained an Upper-class grade OR First-class grade ~ Binomial (n_i, π_i)

$$logit(\pi_{ij}) = \beta_{0j} + u_{0j} + \sum_{Year=1}^{11} (\beta_{Year} + u_{Year,j}) X_{Year,ij} + \sum_{Entry\_qual=8}^{22} \beta_{Entry\_qual} X_{Entry_{qual},ij} \sum_{Ethn=23}^{29} \beta_{Ethn} X_{Ethn,ij} + \sum_{Disb=30}^{35} \beta_{Disb} X_{Disb,ij} + \sum_{Gend=36}^{38} \beta_{Gend} X_{Gend,ij} + \sum_{Age=39}^{40} \beta_{Age} X_{Age,ij} + \sum_{POLAR4=41}^{46} \beta_{POLAR4} X_{POLAR4,ij}$$

3. The βs represent the fixed-effects coefficients which are common to individuals across all modules and years, and the Xs (0 or 1) are binary indicators representing whether an individual possess the categories of the characteristics (variable names and categories are listed in Annex C).

   u_{0j} is the random intercept for Module j, with $u_{0j} \sim N(0, \sigma^2_{u_0})$,

   u_{Yj} is the random intercept for Module j in academic year Y with $u_{Yj} \sim N(0, \sigma^2_{u_Y})$,

   u_{Dj} is the random intercept for Module j in Department D with $u_{Dj} \sim N(0, \sigma^2_{u_D})$, and

   $u_{Year,j}$ is the random coefficient for Module j in academic year $Year$ with

   $u_{Year,j} \sim N(0, \sigma^2_{u_{Year}})$

4. Estimates of the fixed effects coefficients, their standard errors (SE) and p-values for mixed-effects model 2 are shown for Upper-Class grades in Table E1, and for First-Class grades in Table E2. Mixed-effects model 4 coefficients are not included due to space constraint.

Table E1: Fixed-effects coefficient estimates of Model Equation E2 for Upper-class grade attainment

Effects		Fixed-effects estimate	Fixed-effects SE	Fixed-effects p-value
Intercept	Intercept	2.157	0.045	<0.0001
Entry_qual	*AAA and above (ref.)*	-	-	-
	AAB	-0.802	0.022	<0.0001
	Below AAB	-1.233	0.040	<0.0001
	International Baccalaureate	0.125	0.062	0.045
	No A Levels	-0.085	0.046	0.064
	Other A Levels	-0.394	0.066	<0.0001
Ethn	*White (ref.)*	-	-	-
	Asian	-0.391	0.018	<0.0001
	Black	-0.601	0.036	<0.0001
	Mixed	-0.162	0.032	<0.0001
	Other	-0.449	0.050	<0.0001
	Prefer not to say	-0.264	0.065	<0.0001
Disb	*No disability (ref.)*	-	-	-
	Cognitive or learning difficulties	-0.153	0.042	<0.001
	Mental health condition	-0.264	0.066	<0.0001
	Social or communication impairment	-0.071	0.215	0.740
	Sensory, medical, or physical impairment	-0.041	0.063	0.516
	Other or multiple impairments	-0.259	0.071	<0.001
Gend	*Male (ref.)*	-	-	-
	Female	-0.033	0.016	0.044
	Other	-0.487	0.215	0.024
Age	*Young (ref.)*	-	-	-
	Mature	-0.374	0.059	<0.0001
POLAR	*Quintile 5 (ref.)*	-	-	-
	Quintile 1	0.012	0.041	0.772
	Quintile 2	0.069	0.030	0.022
	Quintile 3	-0.014	0.025	0.571
	Quintile 4	-0.013	0.020	0.508

	Unknown	-0.052	0.053	0.327

Table E2: Fixed-effects coefficient estimates of Model Equation E2 for First-class grade attainment

Effects		Fixed-effects estimate	Fixed-effects SE	Fixed-effects p-value
Intercept	Intercept	-0.649	0.036	<0.0001
Entry_qual	*AAA and above (ref.)*	-	-	-
	AAB	-0.821	0.026	<0.0001
	Below AAB	-1.307	0.061	<0.0001
	International Baccalaureate	0.166	0.054	0.002
	No A Levels	-0.041	0.043	0.337
	Other A Levels	-0.282	0.066	<0.0001
Ethn	*White (ref.)*	-	-	-
	Asian	-0.391	0.018	<0.0001
	Black	-0.601	0.036	<0.0001
	Mixed	-0.162	0.032	<0.0001
	Other	-0.449	0.050	<0.0001
	Prefer not to say	-0.264	0.065	<0.0001
Disb	*No disability (ref.)*	-	-	-
	Cognitive or learning difficulties	-0.153	0.042	<0.001
	Mental health condition	-0.264	0.066	<0.0001
	Social or communication impairment	-0.071	0.215	0.740
	Sensory, medical, or physical impairment	-0.041	0.063	0.516
	Other or multiple impairments	-0.259	0.071	<0.001
Gend	*Male (ref.)*	-	-	-
	Female	-0.033	0.016	0.044
	Other	-0.487	0.215	0.024
Age	*Young (ref.)*	-	-	-
	Mature	-0.374	0.059	<0.0001
POLAR	*Quintile 5 (ref.)*	-	-	-
	Quintile 1	-0.024	0.043	0.574
	Quintile 2	0.039	0.031	0.207
	Quintile 3	0.002	0.026	0.932
	Quintile 4	-0.019	0.020	0.362
	Unknown	-0.077	0.056	0.173

5. Estimates of the variance components and their standard errors for the random intercepts and random year coefficients in mixed-effects model 2 are shown for Upper-Class in Table E3, and First-Class grades in Table E4.

Table E3: Variance component estimates for Model Equation E2 for Upper-Class grade attainment

	Random Effect	glmer2 estimate	glmer2 SE
Module Code (Intercept)	$\sigma^2_{u_0}$	0.681	0.825
Academic Year (Intercept)	$\sigma^2_{u_Y}$	0.432	0.657

Table E4: Variance component estimates for Model Equation E2 for First-Class grade attainment

	Random Effect	glmer2 estimate	glmer2 SE
Module Code (Intercept)	$\sigma^2_{u_0}$	0.443	0.666
Academic Year (Intercept)	$\sigma^2_{u_Y}$	0.325	0.570