CESifo WORKING PAPERS

8135
2020

February 2020

Class Size Effects in Higher Education: Differences across STEM and Non-STEM Fields

Elif Kara, Mirco Tonin, Michael Vlassopoulos

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 8135*

Class Size Effects in Higher Education: Differences across STEM and Non-STEM Fields

Abstract

In recent years, many countries have experienced a significant expansion of higher education enrolment. There is a particular interest among policy makers for further growth in STEM subjects, which could lead to larger classes in these fields. This study estimates the effect of class size on academic performance of university students, distinguishing between STEM and non-STEM fields. Using administrative data from a large UK higher education institution, we consider a sample of 25,000 students and a total of more than 190,000 observations, spanning six cohorts of first-year undergraduate students across all disciplines. Our identification of the class size effects rests on within student-across course variation. Overall, we find that larger classes are associated with significantly lower grades (effect size of -0.04) and the effect varies across academic fields, with no effect in non-STEM fields, and a large effect in STEM fields (-0.08). We further explore the heterogeneity of the effect along the dimensions of students' socio-economic status, ability, and gender, finding that in STEM disciplines smaller classes appear to be particularly beneficial for students from a low socio-economic background, with higher attainment in A-levels and to male students.

JEL-Codes: I210, I230, I280.

Keywords: class size, higher education, student academic performance, STEM.

Elif Kara
Bursa Uludag University
Turkey - 16059 Nilüfer/BURSA
elifkara@uludag.edu.tr

Mirco Tonin
Free University of Bozen-Bolzano
Bozen-Bolzano / Italy
Mirco.Tonin@unibz.it

Michael Vlassopoulos
University of Southampton
United Kingdom – Southampton SO17 1BJ
M.Vlassopoulos@soton.ac.uk

February 18, 2020
This project has received IRB approval from the University of Southampton, UK.

1 Introduction

There is an ongoing discussion in the economics of education literature on the impact of class size on student outcomes. The majority of studies have focused on student outcomes at the primary and secondary education level (Angrist and Lavy, 1999; Hoxby, 2000; Hanushek, 2003; Krueger, 1999, 2003). However, larger classrooms may affect various student outcomes in higher education as well, through different channels (Cuseo, 2007).[1] Moreover, in recent years, increasing university enrolment rates and financial pressure faced by higher education institutions have resulted in larger class sizes, and this is drawing more attention to the issue in higher education as well.[2] The interest is not least due to the fact that student to staff ratios appear to influence student satisfaction (Lenton, 2015), as well as being used as metrics to construct university league tables, so, in the increasingly competitive marketplace for higher education students, policy makers and university administrators are forced to pay attention to the issue. Also, class size affecting academic attainment in higher education can have far-reaching implications for students, as there is evidence that graduating with a higher degree classification improves labour market outcomes (Walker and Zhu, 2011; Naylor et al., 2015; Feng and Graetz, 2017).

Knowing what is the impact of larger class sizes on educational outcomes is becoming particularly relevant for science, technology, engineering, and mathematics (STEM) fields. There is a recognition in many countries that there is a shortage of STEM workforce, with implications for technological innovation and economic growth that depend crucially on workers with these skills. Indeed, the UK government is aiming to address the shortfall in STEM skills by investing over 400 million in STEM education (Department for Business and Strategy, 2017). It is very timely then to examine if the expected increase in post-secondary education enrolment in STEM fields and the possible associated increase in class size will have any implications for student learning and educational attainment in these subjects. This will inform the appropriate allocation of

[1] For instance, students might not be paying close attention in larger classes. Hence, there may be less involvement by students, but they may spend more time studying outside of the (large) classes. On the teaching side, lecturers might find it more challenging to identify the right level to pitch the material in larger classes and might have less time to devote to answering students' questions during class and office hours or reply to their emails. Furthermore, the quantity and type of course assignments and assessments and the level of feedback provided might change by class size. Overall, there might be less faculty-student and peer interaction about course material in larger classes.

[2] In the UK, first year enrolment rates have increased between 2007/08 and 2016/17 by 19.6% for first undergraduate degree and 24.2% for postgraduate studies (https://www.hesa.ac.uk/data-and-analysis/students). According to the analysis in Huxley et al. (2018) using data collected in 2013, class size at UK universities has increased considerably since the 1963 Robbins Report.

funding between policies to encourage enrolment (e.g. student bursaries) and investment in teaching resources needed to cater for the needs of larger cohorts of STEM students.

To this end, this paper examines to what extent does class size affect student academic performance - distinguishing between STEM and non-STEM fields - using student-level administrative data from a large UK higher education institution. We use six cohorts (2007/08 through 2013/14) of first-year undergraduate students across all disciplines for a total sample of 25,000 students and more than 190,000 observations. Our identification rests on within student - across course variation in class size that allows us to estimate class size effects accounting for time-invariant student characteristics, such as ability. That is, in our main specification we estimate student fixed-effects regressions, where we also control for various dimensions of the peer group composition. Moreover, we explore non-linearities in the effect, as well as heterogeneity by students' socio-economic background, A-level grades, and gender. Understanding how class size effects vary across subject groups and along the distribution of class size is very important for policy makers and practitioners (e.g. university administrators) as it informs the decision of where to allocate scarce resources. Studying which categories of students are mostly affected is useful to further our understanding of the distributional impact of policies affecting class size. The dimensions we explore are of course not exhaustive. This exercise, however, is informative in that it highlights groups that may be particularly affected by class size.

The overall results, when we consider all the students together, show that larger classes are associated with significantly lower grades (effect size of -0.04). We then divide students into two broad subject groups: STEM fields that include areas such as engineering, mathematics and natural sciences; and non-STEM subjects that encompass areas such as humanities, social sciences and management. The analysis performed at the subject group level indicates that the overall class size effect is indeed not uniform across subject areas. In particular, we find a negative and significant effect for STEM fields (effect size of -0.08), but no effect for non-STEM fields. In STEM fields, we also find evidence of heterogeneity along important dimensions. In particular, when considering non-linearities, we find a benefit from reducing class size when moving from mid-sized to smaller classes, and from very large to large classes. We also explore heterogeneity along the dimensions of socio-economic status, prior academic attainment, and gender. The overall results indicate that the negative effect in STEM fields is more severe for students from a disadvantaged background, for students with previous higher academic achievement (as measured by A-level grades), and for males.

Differences in student and instructor practices among disciplines could explain the heterogeneous class size effects we find across STEM and non-STEM fields. Indeed, a growing literature on undergraduate teaching and learning argues that there are disciplinary differences not only in curriculum content and assessment practices, but also in faculty teaching beliefs and activities, types of teaching method, and student learning requirements (see Neumann, 2001; Neumann et al., 2002; Jones, 2012). For instance, content in the so-called hard sciences (such as chemistry, engineering) is generally fixed and cumulative; the teaching and learning activities tend to be more concentrated and instructive. On the contrary, content in the so-called soft disciplines (such as philosophy, management) is likely to be not fixed; teaching and learning activities are broadly constructive and interpretative. In addition, in hard disciplines, teaching and learning activities are more likely to be faculty-focused whereas in soft disciplines those activities are more student-focused (see Lindblom-Ylanne et al., 2006; Neumann et al., 2002).

This paper fills an important gap in the literature examining class size effects in higher education. It is the first study, to the best of our knowledge, to show how STEM and non-STEM fields differ and how there are negative and non-linear class size effects in STEM fields. We do so using a very similar identification approach as that in Bandiera et al. (2010), which ensures that our results are comparable to this earlier study. Beyond this, the rich administrative dataset allows us to explore heterogeneity of class size effects across other important dimensions, such as, ability, socioeconomic background and gender.

Our findings have important implications. First of all, they suggest that the drive to expand enrolment in STEM subjects should be accompanied by investment in teaching resources to avoid a deterioration in students' achievement. This might seem obvious, but it is not the conclusion we would reach in non-STEM subjects. Moreover, as not all students are affected in the same way, the impact of such policies is not only in the level of achievement, but also in its distribution, therefore affecting the equity of the educational system. Our finding that in STEM fields students with lower socioeconomic background are particularly affected by larger classes suggests, for instance, that developments that would lead to higher student to staff ratios in these fields would disproportionately impact a group that is already disadvantaged in its access to tertiary education.

In the next section, we review the relevant literature. We then describe the institutional background, the data and some descriptive statistics. The following section presents the empirical method. Section 5 presents the results, while the last section concludes.

2 Related Literature

Related literature on class size effects in higher education can be grouped into two categories, by student outcome measures used, as follows: studies assessing the class size effects on student evaluations and papers investigating the class size effects on measured student achievement.

There is nearly a consensus within the first group of studies that there is a negative and significant class size effect on student evaluations regarding their learning, instructor performance and course assessments. Bedard and Kuhn (2008), Mandel and Sussmuth (2011) and Sapelli and Illanes (2016) analyse the impact of class size on student evaluations of instructor performance for a U.S., a German and a Chilean university, respectively. All of these studies find a negative impact of class size on student evaluations of instructor effectiveness. Cheng (2011) and Monks and Schmidt (2011) both evaluate the class size effects on student learning, instructor and course assessments at higher education institutions in the United States. Cheng (2011) finds that increasing enrolment has negative and significant effects on student satisfaction in some disciplines such as Sociology, Political Science, Computer Science and Engineering, and Mechanical and Aerospace Engineering, but has no effect on others. Monks and Schmidt (2011), in turn, show that class size has a negative impact on the student-rated outcomes of amount learned, the instructor rating, and course rating.

The second strand of related literature looks at class size effects on measured academic performance. Kokkelenberg et al. (2008) assess how class size affects the grade higher education students earn at a US public university. They find that mean grades decline as class size increases, up to class sizes of twenty, and more gradually but monotonically through larger class sizes. Bandiera et al. (2010) investigate the causal impact of class size on the academic achievement of postgraduate students in a UK university. They show that the class size effect is negative and significant only for the smallest and largest ranges of class sizes and zero in intermediate class sizes. In addition, high-ability students are more affected by changes in class size, particularly when class sizes are very large. De Giorgi et al. (2012) examine how class size and class composition influence the labour market as well as the academic performance of college students at Bocconi University in Italy. They find that a one standard deviation increase in class size results in a 0.1 standard deviation deterioration of the average grade. Also, the effect is heterogeneous as it is bigger for males and lower-income students. De Paola et al. (2013) analyse the class size effects on college students mathematics and language test performance at a medium-sized Italian university. They show that there is a significant and sizeable negative impact on student performance only in mathematics. In contrast to other evidence shown above, they find that the negative effect is significantly bigger for low-ability and smaller for high-ability students. Diette and Raghav (2015) also assess the relationship between class size and student achievement using data from a selective liberal arts college. They show that on average grades of students decrease as class size increases. In particular, first-year students and students with low SAT scores experience on average larger negative effects from increases in class sizes. Gaggero and Haile (2019) find a negative effect of class size on postgraduate grades in a UK university, using a regression discontinuity approach that exploits a policy whereby in courses where enrolment size reached a certain level, students were split into two groups. Finally, Bettinger et al. (2017) investigate class-size effects on student success in online courses and on student persistence at DeVry University in the United States, even if their sample variation in class size is rather limited. They find little evidence of class size effects on average or for a range of specific course types. This suggests that small class size changes have little effect in online educational settings.

Several studies in this literature consider only a limited range of subjects. The study by De Giorgi et al. (2012), for instance, is on an institution focused on economics and management, while De Paola et al. (2013) consider only students enrolled in economics and pharmacy and nutritional sciences. Other studies, like Bandiera et al. (2010) and Kokkelenberg et al. (2008), cover a wider variety of subjects, but there is no systematic comparison of STEM vs non-STEM subjects, a distinction that is particularly relevant given the current policy discussion.

3 Institutional Background and Data

This paper uses student-level administrative data from a UK higher education institution. The university is a large, research-intensive, Russell Group university[3] with 31 academic units grouped into eight faculties, offering over 200 undergraduate as well as

[3] https://russellgroup.ac.uk/

postgraduate taught and research degree programmes. We focus on first-year full-time students enrolled in three-year B.Sc., B.A., B.Eng., LL.B. or four-year integrated M.Eng. and M.Sci. degree programmes. Students take compulsory and optional courses throughout their degree. The common objective of the first year of each degree programme is to provide a solid foundation in the degree core subjects through compulsory courses and to broaden the field of study by the choice of optional courses.

Our sample covers 25,442 first-year undergraduate students enrolled full-time in one of the 193 different degree programmes offered by 26 different academic units.[4] The sample spans academic years 2007/08 through 2013/14. There are a total of 190,231 student-course level observations; of these 45% we classify as compulsory and 55% as optional courses.[5] One limitation of the data that are available to us is that they concern first-year students only. However, first-year students can be more affected by larger classes (see, e.g., Diette and Raghav, 2015), for instance due to their transition to a new academic environment. Moreover, as several papers in this literature use first-year students (see, e.g., De Paola et al., 2013; Ho and Kelman, 2014), this makes our results more comparable.

Our key dependent and explanatory variables are final grades and class size, respectively. Final grades range from 0 to 100. It should be noted that first year grades do not count toward the final degree classification, though students need to pass the first year courses in order to progress (the pass mark for undergraduate students is 40). Class sizes are calculated as the sum of students who are formally enrolled in a particular course.

In the econometric analysis, we divide students into two broad groups based on field of study: one includes degrees in science, technology, engineering and mathematics (STEM fields) from 12 academic units (e.g. mechanical engineering, physics and astronomy, chemistry - see table 14 in the appendix for details) amounting to 89,662 student-course level observations and a total of 10,011 students.[6] Non-STEM fields, in turn, include

[4] Students who are enrolled in units within Health Sciences and Medicine are excluded from the sample due to the use of a non-numerical grading system in those subjects. We also exclude the few students who take less than 4 or more than 20 modules, and the few modules with less than 5 students.

[5] In the administrative dataset, whether a course is compulsory or optional for a student was not readily available. In order to extract this information, we first calculated how many students within a programme-year take each course. If a course is taken by 100% of the students within the same programme-year, we designate it as a compulsory course, otherwise we classify it as optional. However, this approach could lead to misclassification as particularly popular optional courses may be classified as compulsory, while courses that are compulsory only for a subset of the programme cohort, e.g. those without A-level in a given subject, may be classified as optional. We introduce this variable to improve comparability with Bandiera et al. (2010), where they control for whether a course is core or elective.

[6] In defining STEM fields we follow a common classification that excludes social and behavioral sciences from STEM fields (see for example Kokkelenberg and Sinha, 2010, and Wakeham, 2016).

degrees from 14 academic units (e.g. modern languages, law, social sciences) and comprise 100,569 student-course level observations and a total of 15,431 students. Overall, in terms of number of observations, we note that the grouping of subjects gives rise to two disciplinary categories that are rather similar.[7]

3.1 Descriptive statistics

In this section, we will present some descriptive statistics concerning the main variables of interest. Given the aim of the paper, we also disaggregate by field and by class size quintiles.

First of all, in table 1 we show descriptive statistics on our outcome variable, final grades (Figure 3 in the Appendix provides the distribution). Overall, the average final grade is 60 (o.s.d. 14), with STEM having a slightly higher figure than non-STEM. In all cases, the between and within students standard deviations are of comparable magnitude. In terms of number of students, non-STEM has more students than STEM. We note also that there is a difference in terms of average number of courses per student, ranging from 6.5 for non-STEM to 9 for STEM.

Table 1: Descriptive Statistics on Grades

	All	**STEM**	**Non-STEM**
Mean	60.0	61.4	58.8
Overall Standard Deviation	14.10	16.13	11.87
Standard Deviation Between Students	10.27	12.28	8.58
Standard Deviation Within Student	9.54	10.81	8.26
Min	0	0	0
Max	100	100	100
Number of Observations	190,231	89,662	100,569
Number of Students	25,442	10,011	15,431
Average # Courses per student	7.5	9.0	6.5

Notes: Grades denote the final exam results and range from 0 to 100. The between and within standard deviations account for the unbalanced panel data.

Next, in table 2 we look at our main explanatory variable, class size (Figure 4 in the Appendix provides the distribution). Overall, class sizes range between 5 and 389, with an average of 148 (o.s.d. 80). Across disciplines, non-STEM display a slightly higher average, at 149 (o.s.d. 81) than STEM, at 146 (o.s.d. 79). The between students standard deviation is higher than the within one for both fields.

[7] Table 14 in the Appendix shows descriptive statistics for each individual academic unit that compose the two disciplines.

Table 2: Descriptive Statistics on Class Sizes

	All	STEM	Non-STEM
Mean	148.0	146.5	149.38
Overall Standard Deviation	80.45	79.45	81.30
Standard Deviation Between Students	62.83	61.96	63.38
Standard Deviation Within Student	54.43	53.74	55.04
Min	5	5	5
Max	389	389	369
Number of Observations	190,231	89,662	100,569
Number of Students	25,442	10,011	15,431
Average # Courses per student	7.5	9.0	6.5

Notes: Class sizes are calculated as the sum of students who are formally enrolled in a particular course and took the final exam. The between and within standard deviations account for the unbalanced panel data.

As we will explore non-linearities in class size effects, in table 3 we present descriptive statistics by class size quintiles. Beyond average grades, we also report the average share of female students within each quintile, as well as the average share of British students, which are both peer group characteristics that we use in the empirical analysis. We also construct indices of ethnic and academic fragmentation.[8] The former accounts for the diversity of students (in a course-year) in terms of ethnicity and the latter measures heterogeneity of students (in a course-year) in terms of the academic unit to which they belong.

What emerges is that across quintiles average grades are similar, with only a slight decrease at the fifth quintile, for class sizes over 146. The ratio of female students is above 40% and that of British students is around 80%. The index of ethnic fragmentation increases above the third quintile, while academic fragmentation is U-shaped. This means that students in small and large classes tend to be more heterogeneous in terms of the academic unit they belong to compared to students in mid-sized classes.

[8] $D = 1 - \sum_{i=1}^{k} P_i^2$ where D is the index of diversity, P is the proportion of observations in the i^{th} category, k is the number of categories. The index is the probability that two randomly selected individuals from the population are in different categories. The categories we have in ethnicity are white, black, Asian, Chinese and other, whereas the different categories in academic units are Archaeology, Art, Biological Sciences, Chemistry, Electronics and Computer Science, Education, English, Faculty of Engineering and Environment (& Sciences), Film, Geography, History, Law, Management, Mathematical Sciences, Modern Languages, Music, Ocean and Earth Science, Philosophy, Physics and Astronomy, Psychology and Social Sciences.

Table 3: Descriptive Statistics by Class Size Quintiles

Class Sizes	Mean Grades	Female Ratio	Ethnic Frag.	Academic Frag.	Brit. Ratio
<24	**60.04** (15.11)	**0.49** (0.28)	**0.30** (0.21)	**0.26** (0.28)	**0.79** (0.24)
24-44	**60.91** (13.54)	**0.42** (0.25)	**0.30** (0.20)	**0.20** (0.24)	**0.82** (0.19)
45-85	**60.85** (13.89)	**0.43** (0.22)	**0.29** (0.18)	**0.12** (0.16)	**0.84** (0.16)
86-146	**60.74** (14.70)	**0.41** (0.26)	**0.34** (0.17)	**0.12** (0.16)	**0.81** (0.16)
147-389	**59.28** (13.81)	**0.48** (0.21)	**0.39** (0.19)	**0.26** (0.23)	**0.79** (0.16)

Notes: Class sizes are ordered from the first to the fifth quintile. Standard deviations are in parentheses. All the ratios are first calculated by course-year; we then take the averages by class size percentiles (and academic disciplines). The female ratio is the share of female students; the ethnic fragmentation takes the value of 0 if all students in a course-year belong to same ethnic group and 1 if none of the students in a course-year belong to the same ethnic group; the academic fragmentation takes the value of 0 if all students in a course-year share the same academic unit and 1 if all students in a course-year are from different academic units; the British ratio is the share of British students.

We repeat the same exercise for the two fields separately, using quintiles defined on the basis of the own class size distribution of each field. Starting with STEM in table 4, we see how average grades increase slightly after the first quintile, but drop again at the last one. STEM is characterized by a low female presence, as well as a relatively higher incidence of non-British students. Ethnic fragmentation is relatively high, while academic fragmentation varies a lot, being very low in the middle quintile, but high for very large classes in the last quintile.

Table 4: Descriptive Statistics by Class Size Quintiles: STEM

Class Sizes	Mean Grades	Female Ratio	Ethnic Frag.	Academic Frag.	Brit. Ratio
<31	**60.4** (17.19)	**0.25** (0.19)	**0.38** (0.18)	**0.17** (0.28)	**0.72** (0.18)
31-67	**62.9** (15.87)	**0.30** (0.21)	**0.35**(0.19)	**0.18** (0.27)	**0.75** (0.20)
68-103	**62.6** (16.24)	**0.22** (0.15)	**0.38** (0.16)	**0.05** (0.10)	**0.75** (0.15)
104-162	**62.0** (15.92)	**0.30** (0.19)	**0.35** (0.17)	**0.13** (0.17)	**0.80** (0.16)
163-389	**60.2** (16.05)	**0.34** (0.19)	**0.44** (0.16)	**0.34** (0.27)	**0.76** (0.12)

Notes: See note to table 3

Concerning non-STEM, in table 5 we see that the average grade slightly drops as class size increases, from 59.8 in the first quintile to 58.4 in the last. The share of female and British students is higher than in STEM. In addition, ethnic diversity is also lower and it generally increases by class size. Academic diversity, in turn, is U-shaped.

Table 5: Descriptive Statistics by Class Size Quintiles: Non-STEM

Class Sizes	Mean Grades	Female Ratio	Ethnic Frag.	Academic Frag.	Brit. Ratio
<23	**59.8** (13.53)	**0.60** (0.23)	**0.27** (0.22)	**0.28** (0.27)	**0.80** (0.26)
23-45	**59.8** (10.85)	**0.56** (0.21)	**0.21** (0.15)	**0.19** (0.19)	**0.92** (0.11)
46-90	**59.1** (11.90)	**0.55** (0.17)	**0.22** (0.15)	**0.15** (0.16)	**0.90** (0.11)
91-163	**58.8** (12.05)	**0.61** (0.19)	**0.32** (0.18)	**0.16** (0.17)	**0.86** (0.14)
164-369	**58.4** (11.80)	**0.59** (0.16)	**0.37** (0.21)	**0.22** (0.20)	**0.80** (0.19)

Notes: See note to table 3

4 Econometric Model

We assess the effect of class size on students' final exam performance relying on within student variation in class sizes exploiting the panel nature of the data. In particular, following Bandiera et al. (2010), we employ a panel data specification of the following form:

$$y_{i,k} = \alpha_i + \beta CS_k + \lambda X_{ik} + \gamma P_k + u_{ik}, \quad (1)$$

where $y_{i,k}$ is the final grade of student i on course k, in which by course, we mean a course given in a specific year, so that Econ101 in 2009 is different from Econ101 in 2010. α_i is a fixed effect for student i, which captures the student's innate ability, motivation, etc., CS_k measures the class size on course k, which is the number of students who enrolled on course k. X_{ik} is a dummy variable that shows whether a course is compulsory or elective for student i. P_k controls for peer group composition (heterogeneity) including the share of female students, the ethnic and academic diversity of students, and the share of British students in the course. Finally, u_{ik}, is a residual random term, clustered by course to capture common unobservable shocks to students' grades. Compared to Bandiera et al. (2010), we lack information on who is teaching a specific course in a given year, and therefore cannot control for faculty. However, this turns out not to be very relevant in the results by Bandiera et al. (2010) and, in any case, we will compare our results to their specification without faculty controls.

Besides the class size effect represented by $\hat{\beta}$, we report, in line with the literature, the implied effect size, defined as $\hat{\beta}[\text{sd}(CS_k\ )/\text{sd}(y_k)]$ where the standard deviations are calculated within students whenever we include a student fixed effect.

We then examine whether there are any heterogeneous class size effects using a similar panel data specification of the following form:

$$y_{i,k} = \alpha_i + \sum_{q=2}^{5} \beta_q Q_{qk} + \lambda X_{ik} + \gamma P_k + u_{ik}, \tag{2}$$

where Q_{qk} is equal to one if class size is in q-th quintile of class size distribution, and zero otherwise. All other controls are defined as above and u_{ik} is still clustered by course.

5 Results

5.1 Overall Results

In this section, we first estimate the overall class size effect, first for the whole dataset, then separately for the two fields. We also investigate heterogeneity by looking at the effect by quintiles of class size. Finally, we further delve into heterogeneous effects by splitting the sample according to socio-economic status, ability, and gender.

Table 6 presents overall class size effects.[9] In column 1, without controlling for any other factors on student performance, the implied effect of class size is -0.0453 and significantly different from zero. This shows that a one standard deviation increase in class size reduces students' final grades by 0.0453 standard deviations of the overall distribution of final grades, i.e. larger classes are associated with significantly lower grades. In column 2, we add student fixed effects as well as a control for whether the course is compulsory or elective for the student. The estimated class size effect is still negative and significant and slightly bigger than in column 1, at -0.0517. This indicates that a one standard deviation increase in class size (increase of about 54 students), would lead to a reduction of the average grade of 0.0517 of the within student standard deviation. In column 3, the peer composition of the class is also controlled for. The estimated class size effect is still negative and significant but slightly smaller than in column 2, at -0.0431. As a term of comparison, Bandiera et al. (2010) find larger effects, as they estimate an unconditional implied effect size of -0.074, further dropping to -0.082 within student and to -0.093 when also controlling for class composition.

[9] The full regression results are presented in Table 15.

Table 6: Class Size Effects

	Unconditional (1)	Within Student (2)	Class Composition (3)
Class Size	-0.00794***	-0.00906***	-0.00755***
	(0.00170)	(0.00190)	(0.00193)
Implied Effect Size	-0.0453***	-0.0517***	-0.0431***
	(0.00972)	(0.0109)	(0.0110)
Student Fixed Effect	No	Yes	Yes
Adjusted R-squared	0.002	0.474	0.476
Observations	190,231	190,231	190,231
Number of Students	25,442	25,442	25,442

Notes: Dependent variable is final exam scores. Course-year clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. In columns 2 and 3 we control for whether the course is a compulsory or elective course for a student. In column 3 we also control for the peer group characteristics – the share of women, the ethnic fragmentation among students, the fragmentation of students by academic unit, and the share of British students. The implied effect size is defined as the effect on mean final grade of a one standard deviation increase in class size divided by the standard deviation of grades. Only in column 1, are these standard deviations calculated over all students; in the other columns the standard deviations refer to the within student values.

5.2 Results by Field

We now perform the same exercise for the two fields separately. Table 7 presents class size effects on STEM students' performance, while table 8 on non-STEM.

Table 7: Class Size Effects: STEM

	Unconditional (1)	Within Student (2)	Class Composition (3)
Class Size	-0.0116***	-0.0207***	-0.0166***
	(0.00310)	(0.00303)	(0.00323)
Implied Effect Size	-0.0573***	-0.103***	-0.0825***
	(0.0153)	(0.0151)	(0.0161)
Student Fixed Effect	No	Yes	Yes
Adjusted R-squared	0.003	0.501	0.504
Observations	89,662	89,662	89,662
Number of Students	10,011	10,011	10,011

Notes: See note to table 6.

What emerges is that there are considerable differences across the two fields. There is indeed a negative and significant class size effect for STEM students, with an implied effect size of -0.083 when we control for student fixed effects and for class composition (column 3). On the opposite, as soon as we add student fixed effects, the class size effect for students in non-STEM is one order of magnitude smaller and statistically not different from zero.[10]

[10] We formally test the difference of the class size effects across the two fields by running a fully interacted model and confirm that the impact of class size is indeed different across STEM and non-STEM.

Table 8: Class Size Effects: Non-STEM fields

	Unconditional	Within Student	Class Composition
	(1)	(2)	(3)
Class Size	-0.00427***	0.00111	0.00117
	(0.00158)	(0.00204)	(0.00206)
Implied Effect Size	-0.0292***	0.00737	0.00781
	(0.0109)	(0.0136)	(0.0137)
Student Fixed Effect	No	Yes	Yes
Adjusted R-squared	0.011	0.433	0.434
Observations	100,569	100,569	100,569
Number of Students	15,431	15,431	15,431

Notes: See note to table 6.

Comparing column (1) to column (2) in table 7, we notice how in the case of STEM the coefficient becomes more negative when including student fixed effects. This implies positive ability sorting, with more high-ability students in bigger classes. The opposite is true for non-STEM. There is a negative and significant class size effect in column (1) in table 8, while when we include student fixed effects, the effect disappears.

5.3 Non-linear Class Size Effects

After having established whether or not a class size effect is present across the two fields, we now explore non-linearities. A negative overall effect could be due to a uniform negative effect over the whole class size distribution, or a combination of no effect over a sizeable part of the distribution (e.g. for class sizes below the median) and a strong negative effect over the rest of the distribution (e.g. for very large classes). The policy implications of these two scenarios are of course very different. Moreover, the lack of an overall effect could hide significant effects over some portions of the class size distribution.

To explore these aspects, in table 9 we look at heterogeneous class size effects estimating equation 2, where we use dummies for the quintiles of class size distribution (see section 3 for descriptive statistics by quintile). In column 1, we look at the whole sample of students. The results show that indeed the class size effects on student performance are heterogeneous. To begin with, there is no significant class size effect comparing the omitted category – the first quintile (corresponding to class sizes 5-23) – to the second quintile (class sizes 24-44) or the third quintile (class sizes 45-85). However, there is a sizeable and statistically significant negative class size effect when comparing the first quintile to the fourth quintile (class sizes 86-146) or to the fifth quintile (class sizes 147-389). Furthermore, we test whether there are any class size effects moving between consecutive quintiles. The results, also in table 9, show that from the second quintile onwards there is a significant negative class size effect between consecutive quintiles, albeit the significance moving from the fourth to the fifth quintile is only at the 10% level (p-value: 0.065). It thus appears that, when considering all the disciplines together, there is a negative effect of larger classes with the exception of the lower section of the class size distribution, and the negative effects are quantitatively strong once we consider very large classes. As a term of comparison, Bandiera et al. (2010) show a negative effect moving from the first to the second quintile, as well as from the second to the third and from the fourth to the fifth.

Table 9: Non-Linear Class Size Effects

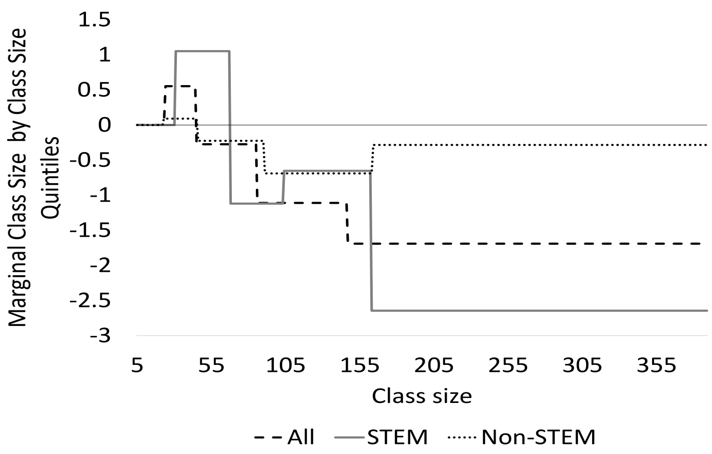
	All (1)	STEM (2)	Non-STEM (3)
Class Size: Quintile 2	0.554	1.053	0.0917
	(0.461)	(0.775)	(0.430)
Class Size: Quintile 3	-0.275	-1.118	-0.226
	(0.457)	(0.738)	(0.460)
Class Size: Quintile 4	-1.108**	-0.653	-0.689
	(0.447)	(0.772)	(0.486)
Class Size: Quintile 5	-1.690***	-2.645***	-0.284
	(0.452)	(0.806)	(0.495)
Student Fixed Effect	Yes	Yes	Yes
Test:Quintile 2=Quintile 3 (p-value)	0.035	0.001	0.405
Test:Quintile 3=Quintile 4 (p-value)	0.028	0.439	0.268
Test:Quintile 4=Quintile 5 (p-value)	0.065	0.000	0.264
Adjusted R-squared	0.477	0.484	0.434
Observations	190,231	89,662	100,569
Number of Students	25,442	10,011	15,431

Notes: Dependent variable is final exam scores. Course-year clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. In all columns, we control for whether the course is a compulsory or elective course for the student, the peer group characteristics–share of women, the ethnic fragmentation among students, the fragmentation of students by academic unit and share of British students. For All, the class quintiles are characterised by class sizes 5-23, 24-44, 45-85, 86-146 and 147-389 from the first to the fifth quintiles respectively. In STEM, the class quintiles are characterised by class sizes 5-30, 31-67, 68-103, 104-162 and 163-389. In non-STEM, the class quintiles are characterised by class sizes 5-22, 23-45, 46-90, 91-163 and 164-369.

When we repeat the same exercise only for students in STEM (column 2), we find that the negative effect of class size is present when comparing the first quintile (class sizes 5-30) to the last one (class sizes 163-389), as well as the second to third and the fourth to fifth. The differences between the first and the second, as well as the third and fourth quintiles are instead insignificant. Therefore, there are beneficial effects when moving both from mid-sized to smaller classes and from very large to large classes. Column 3, in turn, considers only students in non-STEM and the results confirm the lack of a significant class size effect displayed in the main analysis also when considering different portions of the class size distribution. Figure 1 plots the different estimates for the sample as a whole, as well as for the two fields, clearly displaying the difference between

STEM and non-STEM.

Figure 1: Non-Linear Class Size Effects



Notes: The class quintiles are characterised by class sizes 5-23, 24-44, 45-85, 86-146 and 147-389 from the first to the fifth quintiles respectively. In STEM, the class quintiles are characterised by class sizes 5-30, 31-67, 68-103, 104-162 and 163-389. In non-STEM, the class quintiles are characterised by class sizes 5-22, 23-45, 46-90, 91-163 and 164-369.

This analysis highlights where the negative class size effect we detected in the main analysis for STEM arises from. As mentioned, understanding where in the distribution a negative effect appears is essential from a policy perspective, as scarce resources can then be directed towards reducing the size of classes in the relevant range rather than across the board.

5.4 Heterogeneous Class Size Effects by Socio-Economic Status

Here we explore heterogeneity of the class size effect in term of students' socio-economic status (SES), defined on the basis of parental occupation, an information we have for almost three quarters of our sample. Following the guidelines from the Office of National Statistics, we group the seven categories present in our data into the following three:

- High SES: includes higher and lower managerial, administrative and professional occupations;
- Intermediate SES: includes intermediate occupations, plus small employers and own account workers;

- Low SES: includes lower supervisory and technical, semi-routine, and routine occupations.

To compare the sub-sample for which we have this information to the whole sample, in columns 4 and 5 of table 10 we report results from our preferred specification for the whole sample, corresponding to column 3 of table 6, and for this sub-sample. Albeit slightly smaller, a negative class size effect is present also in this sub-sample. In the first three columns of table 10, we then estimate the same specification for the three socio-economic groups shown separately. Not surprisingly, there are more students from a high socio-economic background than from intermediate and low SES. Comparing the implied effect size, it appears that the negative effect of larger classes is present across the three groups, but much larger for the students with a low socio-economic background.[11] Table 11 repeats the same exercise for the two fields. For STEM, students with low or intermediate SES are more affected by class size than students with high SES, and a marginally significant negative effect appears for non-STEM students with low socio-economic background.

Repeating a similar exercise on the basis of parental higher education qualification (see table 16 in the appendix) does not reveal stark differences between students based on whether parents do have a higher education degree or not.

Table 10: Class Size Effects by Parent's Socio Economic Classification

	High SES	Intermediate SES	Low SES	SES Sample	Overall Sample
	(1)	(2)	(3)	(4)	(5)
Class Size	-0.00442**	-0.00689***	-0.0119***	-0.00621***	-0.00755***
	(0.00188)	(0.00220)	(0.00242)	(0.00186)	(0.00193)
Implied Effect Size	-0.0252**	-0.0393***	-0.0680***	-0.0354***	-0.0431***
	(0.0107)	(0.0126)	(0.0138)	(0.0106)	(0.0110)
Student Fixed Effect	Yes	Yes	Yes	Yes	Yes
Observations	91,365	26,796	20,418	138,579	190,231
Adjusted R-squared	0.460	0.465	0.467	0.462	0.476
Number of Students	12,192	3,643	2,791	18,626	25,442

Notes: Dependent variable is final exam scores. Course-year clustered standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. In all columns, we control for whether the course is a compulsory or elective course for the student. In addition, we also control for the peer group characteristics–share of women, the ethnic fragmentation among students, the fragmentation of students by academic unit and share of British students. The implied effect size is defined as the effect on mean final grade of a one standard deviation increase in class size divided by the standard deviation of grades.

This result on socio-economic background is important in terms of an assessment of the social impact of large class sizes. This is indeed not uniform, but rather more severe

[11] We formally test the difference of the class size effects across SES sub-samples in Tables 10 and 11 by running fully interacted models and confirm that the differences we highlight are statistically significant.

Table 11: Class Size Effects by Parent's Socio Economic Classification and Discipline

	STEM	STEM	STEM	Non-STEM	Non-STEM	Non-STEM
	High SES	Intermediate SES	Low SES	High SES	Intermediate SES	Low SES
	(1)	(2)	(3)	(4)	(5)	(6)
Class Size	-0.00954***	-0.0196***	-0.0159***	0.00133	0.00286	-0.00531*
	(0.00333)	(0.00388)	(0.00408)	(0.00204)	(0.00240)	(0.00283)
Implied Effect Size	-0.0474***	-0.0973***	-0.0790***	0.00887	0.0191	-0.0354*
	(0.0165)	(0.0193)	(0.0203)	(0.0136)	(0.0160)	(0.0189)
Student Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	41,991	11,608	9,075	49,374	15,188	11,343
Adjusted R-squared	0.489	0.489	0.497	0.418	0.439	0.432
Number of Students	4,749	1,313	1,017	7,443	2,330	1,774

Notes: Dependent variable is final exam scores. Course-year clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. In all columns, we control for whether the course is a compulsory or elective course for the student. In addition, we also control for the peer group characteristics–share of women, the ethnic fragmentation among students, the fragmentation of students by academic unit and share of British students. The implied effect size is defined as the effect on mean final grade of a one standard deviation increase in class size divided by the standard deviation of grades.

for students from a disadvantaged background, therefore exacerbating inequality. This finding is consistent with previous evidence in De Giorgi et al. (2012), which finds bigger negative class size effects for lower-income students. On the other hand, Bandiera et al. (2010) conduct a related exercise when looking at differences between students residing in private residences in postcodes over or under the median values of house price sales and they find identical class size effects.

5.5 Heterogeneous Class Size Effects by Ability and Gender

Finally, we analyse heterogeneity by gender and by differentiating between students who obtained at least two A, A* or AA grade in the A-level subjects (we refer to them as A students) and students who did not (we refer to them as not-A students).[12]

We see in table 12 that a bit more than half of the students are A students. The negative class size effect is present only for these students (albeit only with 10% statistical significance), while not-A students do not seem to be affected. Looking at STEM (columns 3 and 4), we see that only A students are adversely affected by larger classes. The results for non-STEM (columns 5 and 6), on the other hand, indicate no evidence of adverse class size effects for students of any ability type.

Previous evidence on the effect of class size for students of different ability are mixed.

[12] In the UK higher education system, A-levels are the primary route into higher education and A-level grades are the main criteria that universities use in admissions. In our sample, this information is available for slightly below 80% of the students. Missing A-level information is partly due to students having attended high school outside of the UK.

Our findings of a larger effect of class size for high-ability students are consistent with Bandiera et al. (2010), who, using a quantile regression methodology, find that "those students at right tail of the conditional distribution of test scores - whom we refer to as high-ability students - are more affected by increases in class size". On the other hand, some previous studies (De Paola et al., 2013; Diette and Raghav, 2015) find that the negative class size effect is significantly bigger for low-ability students. Our results are different from these findings, albeit we should keep in mind that they are based on different methodology and measure of ability and are in a very different institutional context.

Table 12: Class Size Effects by A-Level grades and Discipline

	All	All	STEM	STEM	Non-STEM	Non-STEM
	A stud.	Not-A stud.	A stud.	Not-A stud.	A stud.	Not-A stud.
	(1)	(2)	(3)	(4)	(5)	(6)
Class Size	-0.00387*	0.000728	-0.00922***	0.000660	0.00171	0.00221
	(0.00206)	(0.00196)	(0.00347)	(0.00364)	(0.00234)	(0.00213)
Implied Effect Size	-0.0221*	0.00415	-0.0458***	0.00328	0.0114	0.0147
	(0.0118)	(0.0112)	(0.0173)	(0.0181)	(0.0156)	(0.0142)
Student Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	82,769	65,467	41,732	27,720	41,037	37,747
Adjusted R-squared	0.456	0.402	0.485	0.427	0.383	0.382
Number of Students	10,948	8,921	4,706	3,264	6,242	5,657

Notes: Dependent variable is final exam scores. Course-year clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. In all columns, we control for whether the course is a compulsory or elective course for a student. In addition, we also control for the peer group characteristics–share of women, the ethnic fragmentation among students, the fragmentation of students by academic unit and share of British students. The implied effect size is defined as the effect on mean final grade of a one standard deviation increase in class size divided by the standard deviation of grades.

Finally, turning to heterogeneity by gender in table 13, we see that, while overall the gender composition is quite balanced, this is no longer the case when splitting by discipline. Unsurprisingly, females are only one third of students in the STEM sample, whereas they represent almost two thirds in non-STEM disciplines.[13] The class size effect is larger for males both overall and for STEM,[14] while the results for non-STEM again indicate no evidence of significant class size effects. The finding of a larger class size effect for male students is consistent with previous evidence in De Giorgi et al. (2012).

[13] A burgeoning literature focuses on understanding the factors behind the underrepresentation of women in STEM fields (Kahn and Ginther, 2017).

[14] We formally test the difference of the class size effects by gender by running fully interacted models and confirm that the differences we highlight are statistically significant.

Table 13: Class Size Effects by Gender and Discipline

	All	All	STEM	STEM	Non-STEM	Non-STEM
	Male	Female	Male	Female	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)
Class Size	-0.0107***	-0.00410**	-0.0211***	-0.00937***	0.00272	-0.000369
	(0.00237)	(0.00185)	(0.00361)	(0.00329)	(0.00248)	(0.00209)
Implied Effect Size	-0.0612***	-0.0234**	-0.105***	-0.0466***	0.0181	-0.00246
	(0.0135)	(0.0106)	(0.0179)	(0.0164)	(0.0166)	(0.0140)
Student Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	104,806	85,425	63,151	26,511	41,655	58,914
Adjusted R-squared	0.495	0.445	0.513	0.476	0.444	0.422
Number of Students	12,991	12,451	6,855	3,156	6,136	9,295

Notes: Dependent variable is final exam scores. Course-year clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. In all columns, we control for whether the course is a compulsory or elective course for a student. In addition, we also control for the peer group characteristics–share of women, the ethnic fragmentation among students, the fragmentation of students by academic unit and share of British students. The implied effect size is defined as the effect on mean final grade of a one standard deviation increase in class size divided by the standard deviation of grades.

6 Conclusion

We employ administrative data from a large UK university to examine the policy relevant question of the effect of class size on student academic performance in tertiary education, exploring the difference between STEM and non-STEM fields. We also highlight the heterogeneity of the effect along the dimensions of students' socio-economic status, ability, and gender.

Our findings indicate that allocating resources to reduce class sizes would have a significant impact on student achievement in STEM disciplines but not in non-STEM subjects. Moreover, smaller class sizes in those disciplines would be particularly beneficial to certain categories of students, therefore affecting not only the level, but also the distribution of academic achievement. In particular, we have examined the (admittedly not independent) dimensions of social status, ability and gender, finding that in STEM disciplines smaller classes appear to be particularly beneficial for students from a low socio-economic background, with higher attainment in A-levels and to male students. In light of the policy drive to increase enrolment in subjects like engineering, mathematics and natural sciences, it is thus important to allocate enough resources to avoid a deterioration in student achievement due to congested classes.

References

Angrist, J. D. and V. Lavy (1999). Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *The Quarterly Journal of Economics 114*, 533–575.

Bandiera, O., V. Larcinese, and I. Rasul (2010). Heterogeneous class size effects: New evidence from a panel of university students. *The Economic Journal 120*, 1365–1398.

Bedard, K. and P. Kuhn (2008). Where class size really matters: Class size and student ratings of instructor effectiveness. *Economics of Education Review 27*, 253–265.

Bettinger, E., C. Doss, S. Loeb, A. Rogers, and E. Taylor (2017). The effects of class size in online college courses: Experimental evidence. *Economics of Education Review 58*, 68–85.

Cheng, D. A. (2011). Effects of class size on alternative educational outcomes across disciplines. *Economics of Education Review 30*, 980–990.

Cuseo, J. (2007). The empirical case against large class size: Adverse effects on the teaching, learning, and retention of first-year students. *The Journal of Faculty Development 21*(1), 5–21.

De Giorgi, G., M. Pellizzari, and W. G. Woolston (2012). Class size and class heterogeneity. *Journal of the European Economic Association 10*(4), 795–830.

De Paola, M., M. Ponzo, and V. Scoppa (2013). Class size effects on student achievement: heterogeneity across abilities and fields. *Education Economics 21*(2), 135–153.

Department for Business, E. and I. Strategy (2017). Industrial strategy: building a Britain fit for the future.

Diette, T. M. and M. Raghav (2015). Class size matters: Heterogeneous effects of larger classes on college student learning. *Eastern Economic Journal 41*, 273–283.

Feng, A. and G. Graetz (2017). A question of degree: the effects of degree class on labor market outcomes. *Economics of Education Review 61*, 140–161.

Gaggero, A. and G. Haile (2019). Does class size matter in postgraduate education? *The Manchester School*.

Hanushek, E. A. (2003). The failure of input-based schooling policies. *The Economic Journal 113*(485), F64–F98.

Ho, D. E. and M. G. Kelman (2014). Does class size affect the gender gap? A natural experiment in law. *The Journal of Legal Studies 43*(2), 291–321.

Hoxby, C. M. (2000). The effects of class size on student achievement: New evidence from population variation. *The Quarterly Journal of Economics 115*(4), 1239–1285.

Huxley, G., J. Mayo, M. W. Peacey, and M. Richardson (2018). Class size at university. *Fiscal Studies 39*(2), 241–264.

Jones, W. A. (2012). Variation among academic disciplines: An update on analytical frameworks and research. *Journal of the Professoriate 6*(1), 9–26.

Kahn, S. and D. Ginther (2017). Women and STEM. Technical report, National Bureau of Economic Research.

Kokkelenberg, E. C., M. Dillon, and S. M. Christy (2008). The effects of class size on student grades at a public university. *Economics of Education Review 27*, 221–233.

Kokkelenberg, E. C. and E. Sinha (2010). Who succeeds in STEM studies? an analysis of Binghamton University undergraduate students. *Economics of Education Review 29*(6), 935–946.

Krueger, A. B. (1999). Experimental estimates of education production functions. *The Quarterly Journal of Economics 114*(2), 497–532.

Krueger, A. B. (2003). Economic considerations and class size. *The Economic Journal 113*(485), F34–F63.

Lenton, P. (2015). Determining student satisfaction: An economic analysis of the National Student Survey. *Economics of Education Review 47*, 118–127.

Lindblom-Ylanne, S., K. Trigwell, A. Nevgi, and P. Ashwin (2006). How approaches to teaching are affected by discipline and teaching context. *Studies in Higher Education 31*(3), 285–298.

Mandel, P. and B. Sussmuth (2011). Size matters. the relevance and Hicksian surplus of preferred college class size. *Economics of Education Review 30*, 1073–1084.

Monks, J. and R. M. Schmidt (2011). The impact of class size on outcomes in higher education. *The B.E. Journal of Economic Analysis and Policy 11*(1 (Contributions)), Article 62.

Naylor, R., J. Smith, and S. Telhaj (2015). Graduate returns, degree class premia and higher education expansion in the UK. *Oxford Economic Papers 68*(2), 525–545.

Neumann, R. (2001). Disciplinary differences and university teaching. *Studies in Higher Education 26*(2), 135–146.

Neumann, R., S. Parry, and T. Becher (2002). Teaching and learning in their disciplinary contexts: A conceptual analysis. *Studies in Higher Education 27*(4), 405–417.

Sapelli, C. and G. Illanes (2016). Class size and teacher effects in higher education. *Economics of Education Review 52*, 19–28.

Wakeham, W. (2016). Wakeham review of STEM degree provision and graduate employability. *Innovation and Skills: Department for Business. Retrieved from https://www.gov. uk/government/uploads/system/uploads/attachment_data/file/518582/ind-16-6-wakeham-review-stem-graduateemployability. pdf*.

Walker, I. and Y. Zhu (2011). Differences by degree: Evidence of the net financial rates of return to undergraduate study for England and Wales. *Economics of Education Review 30*(6), 1177–1186.

7 Appendix

Figure 2: Distribution of Grade

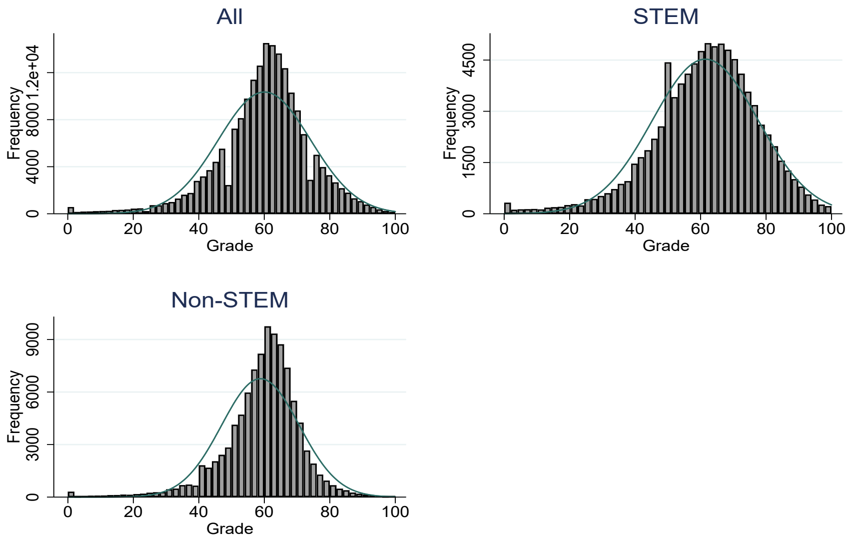


Figure 3: Distribution of Class Size

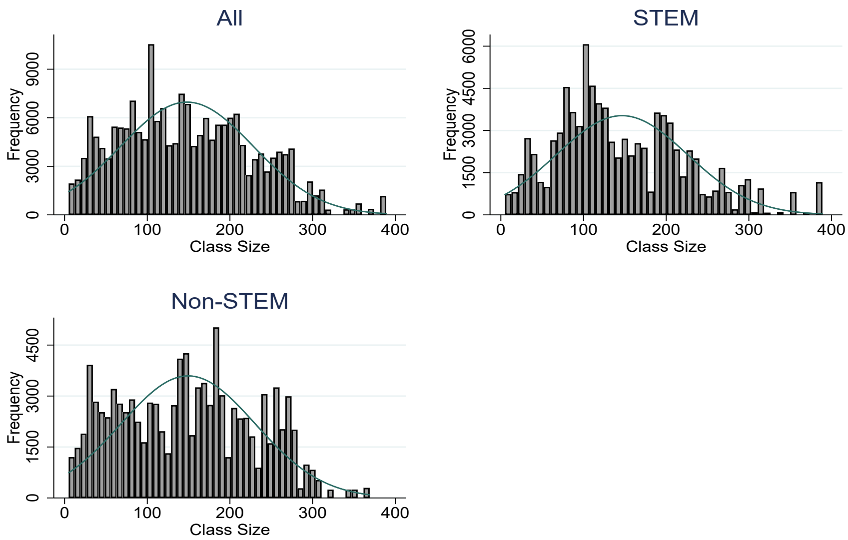


Table 14: Descriptive Statistics on Grades and Class Sizes by Academic Units

	Grades			Class Sizes					
	Mean	Between Std. Dev.	Within Std. Dev.	Mean	Between Std. Dev.	Within Std. Dev.	No of Obs.	No of Stud.	Av. Mod. per stud.
STEM									
Biological Sciences	59.4	9.03	8.61	170.6	35.49	51.55	12714	1550	8.2
Chemistry	60.2	12.00	10.30	92.1	22.71	38.97	5654	735	7.7
Electronics and Computer Science	63.1	12.82	12.11	91.6	20.46	26.88	17453	1628	10.7
Acoustical Eng.	60.2	14.23	11.84	83.9	109.25	79.82	1456	149	9.8
Aerospace Eng.	63.3	13.98	12.70	197.7	63.93	70.41	6315	614	10.3
Civil and Env Eng.	62.1	11.75	11.81	89.8	58.15	54.32	3660	452	8.1
Environmental Sci.	59.9	8.64	8.52	141.9	26.03	94.59	2687	362	7.4
Maritime Eng.	58.5	12.07	12.75	179.0	61.29	88.28	2414	227	10.6
Mechanical Eng	63.0	12.77	12.63	214.7	70.70	61.74	7544	800	9.4
Mathematical Sciences	59.9	15.06	9.12	196.1	29.42	50.53	11125	1386	8.0
Ocean and Earth Science	61.2	9.93	9.34	149.3	21.57	63.95	10530	1295	8.1
Physics and Astronomy	62.6	13.23	11.10	119.1	20.09	34.28	8110	813	10.0
Non-STEM									
Archaeology	57.6	9.27	7.35	63.6	23.36	46.20	3128	410	7.6
Art	56.5	7.45	7.02	230.8	57.33	31.01	8338	2041	4.1
English	60.7	6.41	6.22	141.9	20.60	62.00	6578	1172	5.6
Film	60.1	7.38	6.27	83.8	46.21	57.78	1505	353	4.3
History	60.5	6.27	5.38	96.8	21.32	81.37	8077	1324	6.1
Modern Languages	61.9	8.19	7.42	69.2	26.89	39.88	5959	959	6.2
Music	60.8	7.04	6.56	69.0	18.00	21.77	3362	426	7.9
Philosophy	60.0	9.47	11.22	103.9	38.62	49.30	4031	597	6.8
Management	58.8	10.05	10.22	176.5	24.09	73.49	10125	1259	8.0
Law	54.5	8.75	7.18	193.7	30.73	11.38	5845	1355	4.3
Education	56.7	9.37	9.30	69.0	23.16	31.81	3449	427	8.1
Geography	59.3	6.56	7.85	200.9	38.42	51.16	8909	1183	7.5
Psychology	61.7	7.03	7.07	160.0	20.38	25.31	8424	1056	8.0
Soc. Sciences	57.4	10.01	9.65	162.0	35.84	65.41	22839	2869	8.0

Table 15: Class Size Effects

	Unconditional (1)	Within Student (2)	Class Composition (3)
Class Size	-0.00794***	-0.00906***	-0.00755***
	(0.00170)	(0.00190)	(0.00193)
Compulsory Course		1.831***	1.879***
		(0.205)	(0.205)
Female Share			2.235***
			(0.858)
British Share			-13.60***
			(1.788)
Ethnic Index			-6.538***
			(1.343)
Academic Index			-0.590
			(0.660)
Student Fixed Effect	No	Yes	Yes
Adjusted R-squared	0.002	0.474	0.476
Observations	190,231	190,231	190,231
Number of Students	25,442	25,442	25,442

Notes: Dependent variable is final exam scores. Course-year clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. In columns 2 and 3 we control for whether the course is a compulsory or elective course for a student. In column 3 we also control for the peer group characteristics – the share of women, the ethnic fragmentation among students, the fragmentation of students by academic unit, and the share of British students. The implied effect size is defined as the effect on mean final grade of a one standard deviation increase in class size divided by the standard deviation of grades. Only in column 1, are these standard deviations calculated over all students; in the other columns the standard deviations refer to the within student values.

Table 16: Regressions by Parent's Higher Education Qualification by Discipline

	All	All	STEM	STEM	Non-STEM	Non-STEM	
	PHE	PHE	PHE	PHE	PHE	PHE	PHE
	Yes	No	Yes	No	Yes	No	Sample
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Class Size	-0.00838***	-0.00759***	-0.0162***	-0.0164***	0.000798	-0.000346	-0.00818***
	(0.00204)	(0.00205)	(0.00333)	(0.00363)	(0.00214)	(0.00228)	(0.00196)
Implied Effect Size	-0.0478***	-0.0433***	-0.0808***	-0.0817***	0.00532	-0.00231	-0.0466***
	(0.0116)	(0.0117)	(0.0165)	(0.0181)	(0.0143)	(0.0152)	(0.0112)
Student Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	111,069	52,238	55,456	22,162	55,613	30,076	163,307
Adjusted R-squared	0.471	0.472	0.499	0.505	0.419	0.433	0.472
Number of Students	14,590	7,110	6,146	2,512	8,444	4,598	21,700

Notes: Dependent variable is final exam scores. Course-year clustered standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. In all columns, we control for whether the course is a compulsory or elective course for a student. In addition, we also control for the peer group characteristics including share of women, the ethnic fragmentation among students, the fragmentation of students by academic unit and share of British students. The implied effect size is defined as the effect on mean final grade of a one standard deviation increase in class size divided by the standard deviation of grades.