Received: 7 May 2019 | Revised: 9 July 2019 | Accepted: 8 May 2020

DOI: 10.1002/jae.2781

RESEARCH ARTICLE

WILEY Journal of APPLIED ECONOMETRICS

The price of forced attendance

Sacha Kapoor[1] | Matthijs Oosterveen[2] | Dinand Webbink[1,3,4]

[1]Department of Economics, Erasmus School of Economics, Erasmus University Rotterdam, Rotterdam, The Netherlands

[2]Advance/CSG, Lisbon School of Economics and Management, University of Lisbon, Lisbon, Portugal

[3]Tinbergen Institute, Rotterdam, The Netherlands

[4]IZA Institute of Labor Economics, Bonn, Germany

Correspondence
Sacha Kapoor, Burgemeester Oudlaan 50, 3062 PA Rotterdam, The Netherlands.
Email: kapoor@ese.eur.nl

Summary

We draw on a discontinuity at a large university, wherein second-year students with a low first-year grade point average are allocated to a full year of forced, frequent, and regular attendance, to estimate the causal effect of additional structure on academic performance. We show that the policy increases student attendance but has no average effect on grades. The effects differ, however, depending on how course instructors handled unforced students, such that we observe significant grade decreases in courses where unforced students were given full discretion over their attendance. Our evidence suggests that grades decrease in these courses because the policy prevented forced students from picking their desired mix of study inputs.

1 | INTRODUCTION

For many people their first real encounter with autonomy happens at college or university. Many students use this new-found autonomy to skip class, especially in the early years of their undergraduate education, choosing instead to focus on extracurricular activities such as student government or leisure with their friends. To combat the rampant absenteeism this new-found autonomy begets,[1] and because of the returns to college performance and graduation (Cunha, Karahan, & Soares, 2011; Jones & Jackson, 1990; Oreopoulos & Petronijevic, 2013), university administrators and instructors often mandate frequent and regular class attendance among their students.[2] These attendance policies provide students with structure, helping them to circumvent behavioral predispositions towards nonacademic activities and ultimately avoid decisions that can be bad for their lifetime utility (Lavecchia, Liu, & Oreopoulos, 2014). By this token, and as long as attendance is valuable, additional structure should be good for academic performance. At the same time, however, additional structure constrains choices (e.g., time on self-study) that are important for grades and, by doing so, precludes sensible students from choices that best serve their own self-interest. This can be bad for academic performance.

We draw on a natural experiment at a large European university to estimate the causal effects of a full year of forced, frequent, and regular attendance. The experiment requires students who average less than 7 (out of 10) in their first year to attend 70% of tutorials in each of their second-year courses. It imposes heavy time costs on students, as they can expect to spend 250 additional hours traveling and attending tutorials over a full academic year, amounting to approximately seven additional hours per week. Students who fail to meet the attendance requirement face a stiff penalty, not being allowed to write the final exam for their course, and having to wait a full academic year before they can

[1]Student absenteeism can be upwards of 60% of classes (Desalegn, Berhan, & Berhan, 2014; Kottasz, 2005; Romer, 1993).

[2]An early discussion of mandatory attendance in economics can be found in the correspondence section of the *Journal of Economic Perspectives* in 1994 (Correspondence, 1994).

take the course again. Because students have imprecise control over their *average* grade in first year, the experiment facilitates a regression discontinuity design (Lee, 2008; Lee & Lemieux, 2010) for identifying the effects of forced attendance.

What does it mean to be forced? Our working definition is that a person is forced if a higher authority unilaterally takes away some of their potential choices. Or, more formally, if the authority imposes a heavy, sometimes infinite, penalty on a particular choice. The policy we study is well within the confines of this definition. The policy asks students to come to campus frequently and regularly—choices that are normally under the purview of the student—and imposes a heavy penalty when they fail to do so. In addition to fitting well with a natural definition for economists, students perceived the policy as one where their attendance was forced because this was how it was communicated to them by the university. Our data support the notion that attendance was forced, as below-7 students collectively failed to meet the 70% criterion in less than one half of 1% of their courses. A more severe penalty—automatic expulsion, for example—would have increased participation by less than half a percent, in other words.

Our estimates imply the policy had no effect on second-year performance, on average, across all courses. The point estimate is negative, however, and allows us to rule out positive effects larger than 0.1 standard deviations with reasonable confidence. We document that this average effect hides effect heterogeneity across courses. While the university required all students below 7 to attend 70% of tutorials in all their second-year courses, it had no policy on how students above 7 should be treated. Several courses overlaid their own attendance initiatives on to the university policy, each differing in the intensity of the attendance constraint they imposed on students who scored above 7 in the first year. Some courses penalized absenteeism by any student, others strongly intimated and explained why all students should attend, while a third group of courses followed the university policy and left attendance decisions up to above-7 students. We observe the same students in all three scenarios because students have no discretion over course choice in the second year.

The university policy had its largest effects in the third group of “attendance-voluntary” courses. For these courses the attendance of forced students increased by more than 50%, while their grades decreased by 0.16–0.26 standard deviations. We delve into mechanisms behind the decrease. We show first that the policy had its largest effects on the attendance of students who live far from campus and who had a greater propensity to miss tutorials in the first year. We use course evaluations to show next that the policy generated an increase in lecture attendance similar to the increase in tutorial attendance, without having a measurable impact on total study time. The first result is consistent with students making calculated decisions about their attendance. The second result is consistent with the policy altering time spent on self-study. The results together suggest that the policy prevented students from attaining their desired mix of study inputs, in line with existing evidence on the importance of time use for student performance (Stinebrickner & Stinebrickner, 2008). We rule out several other mechanisms, including the importance of an increase in exposure to other forced (and relatively low-achieving) peers in the tutorials, as well as the possibility of course heterogeneity in tutorial usefulness, or heterogeneity in course design more generally.

The university policy was abolished in the last year of our sample. The abolition came as a surprise, as students only learned of it after the start of their second year. We find no grade difference between above- and below-7 students in the abolition cohort, that the grades of above-7 students were the same in the abolition and treated cohorts, and that the grades of below-7 students in attendance-voluntary courses were higher in the abolition cohort compared to the treated cohorts. The abolition cohort evidence supports continuity of mean grades near 7 in the absence of treatment, a key identifying assumption in our regression discontinuity design. The evidence is also consistent with the grade decrease in attendance-voluntary courses being driven by grade decreases among forced students alone, and thus with the treatment having no (negative) spillover effects on the grades of unforced students.

Our study contributes to an expanding literature on incentives in education. Recent work analyzes the effects of interventions that reward students financially for “good” choices or better academic performance (Angrist, Oreopoulos, & Williams, 2014; Castleman, 2014; Cohodes & Goodman, 2014; De Paola, Scoppa, & Nistico, 2012; Dynarski, 2008; Leuven, Oosterbeek, & van der Klaauw, 2010).[3] We instead analyze the effect of an intervention that penalizes students heavily for “bad” choices, where the penalty is in terms of time rather than money.

Our findings contribute to debates over the merits of mandatory attendance in higher education (Romer, 1993). The argument for mandatory attendance is based on a robust positive correlation between grades and attendance.[4] The

[3]For more comprehensive lists, at all levels of education, see Lavecchia et al. (2014) and Gneezy, Meier, and Rey-Biel (2011).

[4]For some of the many examples, see Romer (1993), Durden and Ellis (1995), Kirby and McElroy (2003), Stanca (2006), Lin and Chen (2006), Marburger (2001), Martins and Walker (2006), Chen and Lin (2008), and Latif and Miles (2013).

argument has been reinforced by studies that use classroom or course-level evidence to show positive correlations between mandatory attendance and grades (see e.g., Dobkin, Gil, & Marion, 2010; Marburger, 2006; Snyder, Lee-Partridge, Jarmoszko, Petkova, & D'Onofrio, 2014). We build on these studies by estimating the causal effect of a large-scale and year-long mandatory attendance policy.

There are several plausible explanations for why we find negligible to negative effects whereas positive effects have been reported in a wide range of contexts. One explanation relates to identification concerns that we are able to resolve, such as selection bias relating to cohort-specific unobservables or gaming for the purposes of avoiding mandatory attendance policies. A second explanation may simply be that our negative to negligible effects are not inconsistent with the positive effects researchers have found. It could be that the (average) treatment effect is positive and that our effects are specific to the types of students who would be at 7 in the context we study.

This article contributes, more generally, to debates over the role of structure in higher education (Lavecchia et al., 2014; Scott-Clayton, 2011). Arguments for additional structure usually focus on student predispositions towards nonacademic activities, emanating from behavioral biases such as impatience, or imperfect information about behaviors that engender success at university. Our findings imply that additional structure does not increase performance for students with a grade point average (*GPA*) of 7 (out of 10) at a large public university in the Netherlands.

2 | CONTEXT

Our venue is the economics undergraduate program of this university. The economics program itself is large; in the 2013–14 academic year alone, the program saw an influx of approximately 700 students. Students have no discretion over their courses in the first two years of the program, as all students follow the same ten courses per year, covering basic economics, business economics, and econometrics (see Table A.1 in the Supporting Information Appendix). Students have discretion over their courses in the third year and, in line with this, declare a minor and major specialization (e.g., Accounting and Finance) which they can subsequently continue through to a Master's program.[5] The economics program is given in Dutch or English. The only other difference between the programs is that the Dutch program has approximately 2.5 times more students.

Academic years are divided into five blocks, of 8 weeks each (7 weeks of teaching and 1 week of exams). First- and second-year students have one light and one heavy course in each block, where they get four credits for the light course and eight for the heavy course.[6] Heavy courses have two to three large-scale lectures per week, while light courses have one to two. Lecture attendance is always voluntary. Heavy courses have two small-scale tutorials ($\sim$30 students) per week, while light courses have one. Lectures and tutorials both last for 1 hr and 45 min. Unlike lectures, but much like what may be found in structured college programs, tutorials require preparation and active student participation, via, for example, discussions of assignments and related materials.

Second-year courses each have several time slots for tutorials and students can choose the one they wish to attend. Students register for slots a few weeks before the block begins. At registration time, students are unaware of the teaching assistant (TA) that will teach each tutorial group, which are mostly senior undergraduate and PhD students. Students cannot switch their tutorial group after the registration period ends. All students must register for a tutorial. We observe for which group and at which time the student registered and can evaluate whether there were systematic differences in registration patterns for forced and unforced students.

Grades range from 1 to 10. Students fail a course if their grade is below 5.5. The *GPA* in the first year is weighted by the credits of the course. Note that a *GPA* of 8.25 or more at the end of the first year is awarded *cum laude*.

[5]The Dutch and North American systems differ in two important ways. First, majors are defined more narrowly, as students decide to pursue economics, political science, sociology, and other social sciences before entering university. Second, they do 3 rather than 4 years of bachelor's before a Master's.

[6]In Europe study credits are denoted by ECTS, which is an abbreviation for European Transfer Credit System. This is a common measure for student performance to accommodate the transfer of students and grades between European Universities. One ECTS is supposed to be equivalent to roughly 28 hr of studying. 60 ECTS account for 1 year of study.

2.1 | University policy

Second-year students that scored a *GPA* of less than 7 in the first year were forced to attend 70% of tutorials for all second-year courses. Failure to fulfill the 70% attendance requirement precludes students from writing the final exam for the course. They must in turn wait a full year before they can take the course again to obtain these required credits.

Students that failed to complete the first year within year one, however, were forced to attend 70% of the second-year tutorials irrespective of their *GPA*. This implies there is only variation in the assignment to forced attendance for students near 7 that completed the first year on time. To complete the first year on time a student must score: (i) 5.5 or higher in each of the 10 first-year courses; or (ii) 5.5 or higher in most courses and use their high scores in these courses to compensate for grades of 4.5–5.4 in their remaining courses. The 10 first-year courses are assigned to one of three groups and students can only compensate one course within each group (Table A.1 in the Supporting Information Appendix). For example, a student who receives an 8 in microeconomics and 4.5 in macroeconomics can complete the first year by taking 1 point from their micro grade and use it towards their macro grade. The on-time completion rate for students near 7 is 92%. Half do this via criterion (i), while the other half do this via criterion (ii).

Note that the on-time completion rule has no bearing on causal identification. The rules for completing the first year apply to both above- and below-7 students such that there is no sample selection. Consistent with this, we observe no statistical imbalance in the first-year completion rate nor in the use of the compensation method near 7. Throughout the paper we thus restrict the sample to students who completed the first year on time, which contains 92% of all students near 7.[7] In this sample, the mean and standard deviation of first-year *GPA* are 6.99 and 0.70. The analogues in the unrestricted sample are 6.65 and 0.79.

The policy imposes sizable time costs on students. Forced students must spend 26 hr per block (3.5 hr per week) in tutorials. Once we account for travel time, about 45 minutes each way on average,[8] forced students must spend 50 hr per block traveling to and attending tutorials. All costs are in terms of time rather than money because student travel is fully subsidized in the Netherlands.

The introduction of the policy had nothing to do with the historical grade distribution of first-year students. It was introduced as part of a university-wide initiative to personalize education via small-scale tutorials. The initiative came about for three reasons: first, the university had grown to a scale that made education impersonal; second, the tutorials encourage active participation; third, the tutorials facilitate student involvement in the university community. Forced attendance was made part of the initiative to ensure a return on the university's sizable investment in small-scale tutorials.

2.2 | Course policies

While the university forced the attendance of below-7 students in all their second-year courses, courses differed in how they dealt with above-7 students. Table A.2 in the Supporting Information Appendix provides a detailed overview on the courses and on how they dealt with these students. There were three types of courses: (i) two attendance-voluntary or $7^{+}vol$ courses; (ii) three attendance-encouraged or $7^{+}enc$ courses; (iii) three absence-penalized or $7^{+}for$ courses. In $7^{+}vol$ courses attendance was voluntary for above-7 students. In $7^{+}enc$ courses their attendance was strongly encouraged. In $7^{+}for$ courses their absence was penalized. In this last set of courses, students had tutorial assignments that made up 5–30% of their final grade. By not attending, students received a zero on this part of the course, meaning that at most they could obtain a 7–9.5 (rather than 10). The remaining two courses had no tutorials, and the final grade (mostly) consists of writing a research report in groups. Accordingly, these two courses are excluded from our analysis.[9] Ultimately, the course policies provide us with three counterfactuals: the grades of above-7 students whose attendance is voluntary, strongly encouraged, and forced. The three counterfactuals help us sort through mechanisms.

[7]Note that the high first-year completion rates prevent us from estimating a local difference-in-difference, which would compare changes in the grades of students near the cutoff who completed the first year, with the changes in the grades of students near the cutoff who did not complete the first year.

[8]The average student lives 22.9 km from campus. From the Dutch student survey we learn that more than 70% of university students travel by public transport (https://www.studentenmonitor.nl/). We then used the Dutch public transport website (https://9292.nl/) to get an idea of travel times between the university and a few cities within a radius of 20–30 km of the university.

[9]There is no difference in grades near 7 for these two courses. Note that they do not provide credible placebo tests as final grades are largely determined via group work.

2.3 | Abolition

The policy lasted 5 years, starting in 2009–10 and ending 2013–14. The 2008–09 cohort was the first to be subjected to the policy in their second year; the 2012–13 cohort was the last. The policy was abolished in 2014–15 because the student body and faculty, rightfully, as this paper shows, lobbied against it. The abolition came as a surprise to the 2013–14 cohort, as they were only made aware of it *after* their second year had started, in the first block of the academic year 2014–15. They had the same incentive to score above 7 in first year as earlier cohorts, even though below-7 students were ultimately given discretion over their attendance in the second year.

3 | DATA

Our main information source is the university's administrative data. Our sample ranges from the 2008–09 academic year until 2014–15. We observe grades at the level of the student for all three undergraduate years, tutorial attendance for the first 2 years, course evaluations, and various personal characteristics. As discussed in Section 2.1, we will restrict the analysis to students who completed their first year on time. After further restricting the sample to students that score a first-year *GPA*within 0.365 grade points of 7, our baseline estimation sample, we have 524 students and 3,585 (second-year) course–student observations. All but one of the second-year exams are made up out of multiple-choice questions. This precludes TAs from having a direct effect on grades.

The university uses attendance lists to track tutorial attendance. Students must sign in and teaching assistants must upload the attendance data to the university's online portal. The uploaded data are then used by the exam administration to verify that the attendance requirement is met.[10]

We observe the attendance of each student at each tutorial session. We expect little measurement error because instructors required teaching assistants to prevent fraudulent sign-ins via student counts. The attendance statistics for above-7 students reinforces the point. These students attend 55–60% of their tutorial sessions. We show later that they also attend roughly 55–60% of their lectures. The similarity between tutorial and lecture attendance, together with the idea that students incur sunk costs of visiting campus, suggests tutorial attendance is measured accurately.

Our data include information from course evaluations. One week before the exam, students are invited by email to evaluate the course anonymously. They are reminded of the evaluations shortly after the exam. All evaluations have the same 16 core questions, grouped into the general opinion of the course, structure, fairness, quality of lecturer and TA, and usefulness of lectures. Importantly, students are asked about their lecture attendance, as well as time spent on their studies in total (see Table A.4 of the Supporting Information Appendix for comprehensive details). Note that the evaluations are filled out by 20% of the students. The response rate is the same just left and right of 7.

Our personal characteristics data include gender, age, distance from their residence to the university (in kilometers), and whether they are from the European Economic Area (EEA). For Dutch students, roughly 80% of our baseline estimation sample, we also have information on high school performance. Their grade for each of their high school courses is a 50:50 weighted average of the grade they earned in the course and the grade they earned on a nationwide exam for that course.

3.1 | Basic descriptives

Table 1 summarizes the data. It compares students (who completed their first year within year one) with a first-year *GPA* between 6.635 and 7 to students whose *GPA* was between 7 and 7.365. The top panel restricts the sample to second-year courses, where the unit of observation is the student–course combination. The student is the unit of observation otherwise.

Forced students score 0.48 standard deviations worse despite being 13 percentage points more likely to attend tutorials. The bottom panel implies students left and right of 7 are roughly similar. The lone statistical difference is for high school *GPA*, wherein poor-performing students appear to be overrepresented to the left of 7. Note, however, that the difference is statistically insignificant according to our main balancing tests presented later.

[10]The match rate for the attendance and administrative data was 93% (in our baseline sample). We compare matched and unmatched observations in Table A.3 of the Supporting Information Appendix and find no evidence of selection. Therefore, we work with this 93% sample throughout the paper.

TABLE 1 Basic descriptives

Variable	First-year *GPA* [6.635, 7)	First-year *GPA* [7, 7.365]	Diff.
Course level (second year)			
Grade	6.33	6.81	0.481***
	(1.33)	(1.19)	(0.059)
Tutorial attendance	0.90	0.77	−0.130***
	(0.12)	(0.29)	(0.011)
Observations	1827	1758	3585
Student level (all students)			
Distance to university (km)	23.18	22.12	−1.061
	(31.66)	(28.81)	(2.649)
Age	20.28	20.16	−0.126
	(1.07)	(1.20)	(0.099)
Gender (female = 1)	0.30	0.31	0.008
	(0.46)	(0.46)	(0.040)
European Economic Area	0.94	0.92	−0.015
	(0.24)	(0.27)	(0.022)
Observations	269	255	524
Student level (Dutch students)			
High school *GPA*	6.68	6.92	0.237*
	(1.33)	(1.34)	(0.128)
Observations	225	206	431

Notes.
1. Sample is from all eight eligible courses.
2. Grades and high school *GPA* range from 1 to 10.
3. Each high school grade is a 50:50 weighted average of the grade the high school assigned and the grade the student received on a national exam for the course.
4. First two columns have standard deviations in parentheses. Last column has standard errors in parentheses.
5. Asterisks denote statistical significance for difference in means, standard errors clustered on student level.
6. Significance levels: *<10%; **<5%; ***<1%.

3.2 | Preview of baseline results

The left-hand column of Figure 1 examines the attendance effect for the three course types, where attendance is simply the percentage of tutorials attended (per course). In particular, the figures plot the second-year attendance rate against first-year *GPA*, where the difference at a *GPA* of 7 measures the policy impact. In $7^{+}vol$ courses (Figure 1a) this difference in attendance was between 30 and 35 percentage points. This translates into five extra tutorials for an eight-credit course (three for a four-credit course), or about 13 hr of extra schooling per block. In $7^{+}enc$ courses (Figure 1b) the difference at a *GPA* of 7 was approximately 13 percentage points. There was no attendance difference in $7^{+}for$ courses (Figure 1c).

Figure 1 suggests the $7^{+}vol$ and $7^{+}enc$ attendance rates for forced students are higher than necessary. Forced students attend roughly 90% of tutorials for both of these courses, whereas the requirement is 70%. What explains the discrepancy? One explanation relates to the discrete number of tutorials. Light (4-credit) courses have seven tutorials. Attending five of seven tutorials would give students a 71% attendance rate. Going from five to six tutorials, however, increases the completion rate to 86%. If there is some uncertainty about the completion rate, relating for example to how it is recorded, then risk-averse students may attend an additional tutorial just to make sure. In Section 6.2 we further document heterogeneous effects of the policy on attendance that support this interpretation.

The right-hand column of Figure 1 examines the unconditional effect on grades. Grades in $7^{+}vol$ courses decrease by roughly 0.2 standard deviations. For the other courses there seems to be no effect on grades. The attendance and grade effects suggest grades might only decrease if the additional constraint on choices is especially severe.

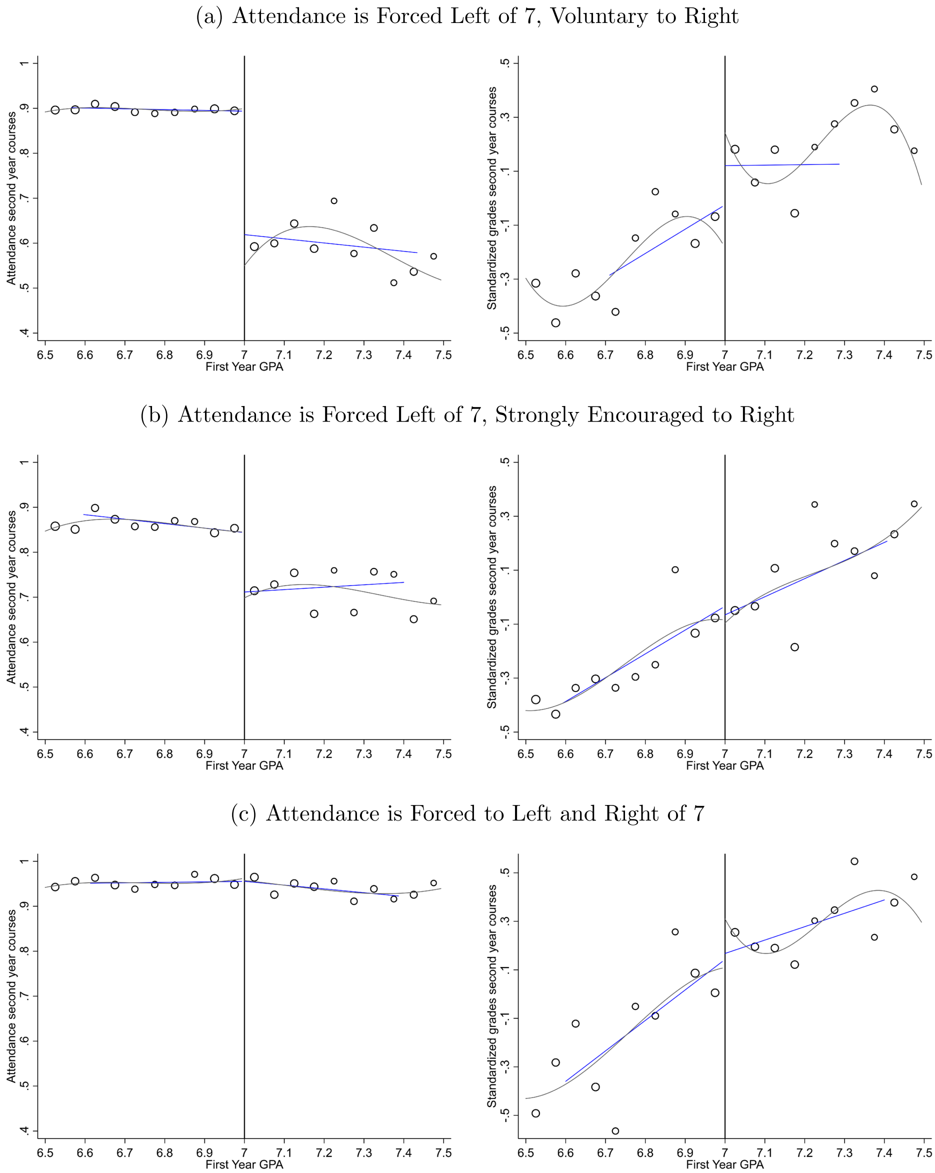


FIGURE 1 Second-year attendance and grades, by course type. (a) Attendance is forced left of 7, voluntary to right. (b) Attendance is forced left of 7, strongly encouraged to right. (c) Attendance is forced to left and right of 7 [Colour figure can be viewed at wileyonlinelibrary.com] [Colour figure can be viewed at wileyonlinelibrary.com]

Notes:

1. Locally linear and cubic scatterplots for attendance or second-year grades against first-year GPA.
2. The local linear polynomial is estimated upon the optimal bandwidth for each outcome relative to a MSE criterion (Calonico, Cattaneo, Farrell, & Titiunik, 2017). The cubic polynomial is estimated upon a bandwidth of 0.5, which is the same across all figures.
3. Dots are based on local averages for a binsize of 0.05. Dot sizes reflect the number of observations used to calculate the average.
4. Binsizes for local averages are selected via F-tests from regressions of second-year grades on K bin dummies and $2K$ bin dummies for the first-year GPA.

4 | EMPIRICAL SPECIFICATION

Let $G_j(D)$ denote the student's second-year grade in course j under regime D, where D indicates whether first-year GPA is less than 7. We are interested in the parameter

$$\tau = \mathbb{E}[G_j(1) - G_j(0) | GPA = 7], \tag{1}$$

the effect of forced attendance at 7. The adoption and use of the forced attendance policy suggest $\tau > 0$. The constraining effects of the policy on choices suggests $\tau < 0$.

We assume the conditional expectations $\mathbb{E}[G_j(1)|GPA=7]$ and $\mathbb{E}[G_j(0)|GPA=7]$ are continuous at 7 (Hahn, Todd, & Van der Klaauw, 2001). Under this assumption τ is identified by

$$\lim_{x \to^- 7} \mathbb{E}[G_j|GPA=x] - \lim_{x \to^+ 7} \mathbb{E}[G_j|GPA=x], \tag{2}$$

where x is a realization of GPA, G_j is the observed grade, and $-$ and $+$ indicate whether GPA approaches 7 from below or above. The continuity assumption can fail if students have precise control over their first-year GPA (Cattaneo, Idrobo, & Titiunik, 2019b; Lee, 2008). Because students were made aware of the policy early in their first year, they could try to avoid forced attendance in the second year. Our identification strategy works as long as first-year grades are somewhat outside of the student's control.

The above is generally a weak identifying assumption (Lee, 2008) and is reasonable in our setting. The assignment to forced attendance is based on the student's *average* grade. As students accumulate grades they lose control over the average. Importantly, first-year adjustments to the threat of second-year forced attendance, such as the practice of asking professors for grade increases, have less of an effect on first-year GPA than on the grade of any one course. Limited control over the average favors the continuity of conditional expectations (for potential outcomes) at 7.

We use weighted least squares to estimate Equation 2 via the locally linear regression specification (Cattaneo, Idrobo, & Titiunik, 2019b; Imbens & Lemieux, 2008):

$$G_{ij} = \beta_0 + \beta_1 D_i + f_+(GPA_i - 7) + f_-(GPA_i - 7)D_i + \varepsilon_{ij}, \tag{3}$$

where i denotes the student. Second-year grades G_{ij} are measured in standard deviations ($1\sigma = 1.45$), $f_+(\cdot)$ and $f_-(\cdot)$ are normalized linear polynomials in first-year GPA_i, and ε_{ij} is a random variable reflecting unobserved differences in second-year grades. We allow the polynomial to differ across 7 (see the discussion by Lee & Lemieux, 2010) and weight observations by a triangular kernel, which (linearly) assigns less weight to observations further from the cutoff. Our main estimates are based on the sample of students within a bandwidth of 0.365 of 7. This is the optimal bandwidth for student grades relative to an MSE criterion (Calonico, Cattaneo, Farrell, & Titiunik, 2017) for the full sample of student–course observations (i.e., when including all three course types). Our decision to use a common bandwidth for the main estimates stems from the panel data structure. A common bandwidth ensures the sample of students is the same across all specifications.

Inference is based on standard errors clustered at the level of the student. We rely mostly on conventional (clustered) standard errors because of our preference for a consistent sample across specifications. Since conventional standard errors are invalid for inference by construction (Calonico, Cattaneo, & Farrell, 2019), we report results based on robust bias-corrected standard errors and MSE-optimal bandwidths unique to each specification in the Supporting Information Appendix. A comparison shows that the estimates and statistical significance reported in the main text are conservative.

4.1 | Continuity near the cutoff

We examine the validity of the continuity assumption. We test for discontinuities in predetermined personal characteristics as well as the density of students near 7.

Table 2 presents estimates of Equation 3 on the student level, where instead of grades the dependent variables are personal characteristics. Students to the left and right of the cutoff are similar in whether they come from the European Economic Area, age, distance from the university (in kilometers), and high school GPA. This conclusion holds if we select the bandwidth MSE-optimally for each background characteristic (Table A.5 in the Supporting Information Appendix). It also holds if we consider grade differences for various high school courses separately (Figure A.1 in the Supporting Information Appendix).

We next draw specifically on the test developed in Cattaneo, Jansson, and Ma (2018) and Cattaneo, Jansson, and Ma (2019) to test for a discontinuity in the probability density for GPA at 7 (McCrary, 2008). If students can manipulate their GPA, then we could observe bunching just above 7. The results of the test are summarized in Figure 2. The figure shows no evidence of bunching. The bias-corrected discontinuity test statistic is 0.25 with a p-value of 0.80, implying that we cannot reject the null hypothesis of no discontinuity at 7. This supports the absence of manipulation around the cutoff.

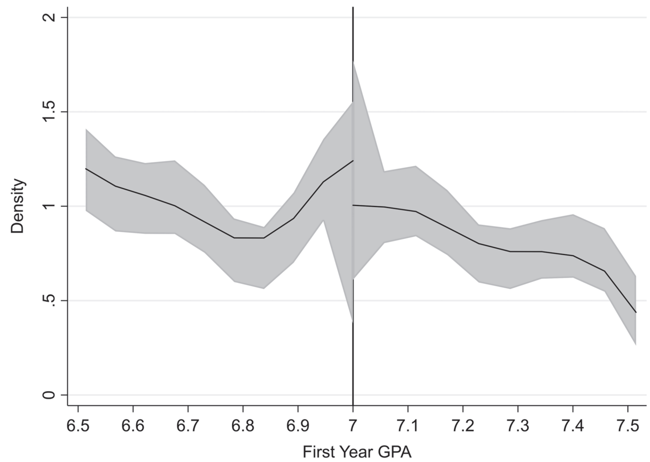
TABLE 2 Balancing tests around the cutoff

	Distance to uni. (km) (1)	Age (2)	Gender (3)	European Economic Area (4)	High school *GPA* (5)
1st-year GPA is below 7	3.213 (6.134)	0.247 (0.178)	0.173* (0.083)	−0.024 (0.051)	−0.428 (0.310)
Mean dep. var.	22.979	20.289	0.290	0.938	6.882
Observations	524	524	524	524	431

Notes.

1. Unit of observation is the student. The outcome variable is displayed at the top of each column. The outcome variables are not standardized.
2. The regressions include a first-order polynomial which is interacted with the treatment. The bandwidth is 0.365 and the kernel is triangular.
3. Standard errors are clustered on the student and in parentheses.
4. Significance levels: *<10%; **<5%; ***<1%.

FIGURE 2 Discontinuity in density test



Although much of the evidence favors continuity, column (3) of Table 2 indicates that women are underrepresented just to the right of the cutoff. The gender imbalance is problematic if it reflects men having a general tendency to ask for and obtain higher grades and, importantly, if this tendency generates a discontinuity in the conditional expectations for potential outcomes at 7. To test for this, Figure A.2 in the Supporting Information Appendix breaks down the density manipulation test by gender. The results imply that the probability density for both males and females around 7 is continuous, though graphically the support is strongest for males. Later we will show that our results are robust to controls for gender.

4.2 | Abolition

We use the abolition cohort to further test the continuity assumption. For this cohort, the treatment regime D equals 0 across all realizations of *GPA*, such that there is no treatment effect at 7 for this cohort. To analyze whether this is the case, we plot second-year grades against first-year *GPA* as in Figure 1, but this time for the 2013–14 cohort only. Figure A.3 in the Supporting Information Appendix documents the results and provides strong support for no difference in second-year grades at the *GPA* of 7 across all three course types. We confirm these zeros more formally in Table A.6 of the Supporting Information Appendix, which reports estimates of Equation 3 using only the abolition cohort.[11]

[11]Note that we cannot analyze the implications for attendance because the university stopped registering attendance in the abolition year.

4.3 | Sample attrition

The policy may have incentivized students to drop courses if and once they fail the 70% attendance requirement. Attrition of this sort could threaten identification because dropouts are not graded. Accordingly, we test for a policy effect on the number of second-year courses for which a student obtained a valid grade. The results in Table A.7 (column (1)) of the Supporting Information Appendix imply that the policy has no effect on the number of completed courses, consistent with the fact that students near the cutoff tend to complete most of their second-year courses (more than 9 out of 10 on average).

Students near 7 may differ in their propensity to complete course evaluations and thus compromise the use of course evaluations in our analysis. Columns (2)–(4) of Table A.7 of the Supporting Information Appendix report estimates of the policy effect on an indicator for whether students completed the course evaluation for all three course types. We find no statistical differences in the propensity to complete the evaluation near 7. As with course completion, our evidence suggests no differential selection into course evaluations.

4.4 | Mass points

One remaining concern relates to whether first-year GPA has enough mass points to warrant a continuity-based RD design. To this end, note that there are 168 unique GPA values for the 524 students in our estimation sample of 6.635–7.365, amounting to approximately one GPA value for every three students. This coverage of the support for GPA is usually sufficient for a continuity-based design.[12]

5 | BASELINE RESULTS

Table 3 reports estimates for student grades based on pooled data from the eight affected courses. Pooling is advantageous because it lets us account for across-course error correlation within students. Average effect estimates are found in columns (1)–(3). Columns (2) and (3) show that the estimates do not change when controlling for fixed effects for the course–cohort combination and for personal characteristics. The point estimates are negative, but not statistically different from zero, and imply that the university-wide policy had little to no average effect on student performance.

5.1 | Course-level attendance policies

Table 4 evaluates the policy effect for the three course types separately. Moving from left to right, the table reports estimates for $7^{+}vol$ courses, $7^{+}enc$ courses, and $7^{+}for$ courses. The table starts with the effect on tutorial attendance in the top panel. The estimates show that the policy increased the attendance of forced students in $7^{+}vol$ courses by 31 percentage points ($p < 0.01$), increased attendance by 13 percentage points in $7^{+}enc$ courses ($p < 0.01$), and had no statistical effect on attendance in $7^{+}for$ courses. The estimates in the top panel show that the policy had a first-order effect on student choices.

The middle panel of Table 4 evaluates the effect on grades. The policy decreased grades by 0.18 standard deviations in $7^{+}vol$ courses ($p < 0.1$). On the Dutch grading scale this amounts to approximately 0.3 grade points ($\approx 0.18 \times 1.45$). The grades of forced students were 0.04 standard deviations higher in $7^{+}enc$ courses and 0.03 standard deviations lower in $7^{+}for$ courses. The latter two estimates are statistically insignificant. Note that columns (4)–(6) of Table 3 show that similar conclusions are reached with pooled data and interactions between the treatment and course type.

Whereas forced students obtain lower grades in $7^{+}vol$ courses, this does not necessarily mean that they also obtain lower passing rates. We explore whether passing rates are affected in the bottom panel of Table 4, which is equivalent to checking whether the grade decreases occur at 5.5, the threshold for passing a course. The results show that the probability of passing is 7 percentage points lower in $7^{+}vol$ courses. The estimate is insignificant at conventional significance levels, however ($p \approx 0.12$). Columns (2) and (3) show that there is effectively no difference in passing rates for $7^{+}enc$ and $7^{+}for$ courses. We conclude that the impact on passing rates is small to negligible.

[12]Cattaneo, Idrobo, and Titiunik (2019a) analyze an example where for every 110 observations one unique value for the running variable is observed. They conclude that continuity-based analysis might be possible in this context.

TABLE 3 Student performance for all eight eligible courses

	Grade (standardized)					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
1st-year *GPA* is below 7	−0.04 (0.08)	−0.02 (0.07)	−0.01 (0.07)	0.04 (0.10)	0.07 (0.10)	0.07 (0.10)
Attendance is voluntary × treatment				−0.22* (0.13)	−0.23* (0.12)	−0.23* (0.12)
Absence is penalized × treatment				−0.07 (0.12)	−0.08 (0.11)	−0.08 (0.10)
Course–cohort FE	No	Yes	Yes	No	Yes	Yes
Personal characteristics	No	No	Yes	No	No	Yes
Observations	3,585	3,585	3,585	3,585	3,585	3,585

Notes.
1. Grades are standardized, where one standard deviation equals 1.45 grade points on the Dutch grading scale.
2. Controls for personal characteristics include distance to the university, age, gender, and European Economic Area.
3. The regressions include a first-order polynomial that is interacted with the treatment. The bandwidth is 0.365, the (MSE) optimal bandwidth for the baseline RD specification with all eight eligible courses and no controls. The kernel is triangular. In columns (4)–(6) the treatment effect and the polynomials are allowed to differ by course type.
4. Standard errors are clustered on the student and in parentheses.
5. Significance levels: *<10%; **<5%; ***<1%.

TABLE 4 Student attendance and performance by course type

	(1)	**(2)**	**(3)**
Attendance rate			
1st-year GPA is below 7	0.31*** (0.04)	0.13*** (0.03)	0.00 (0.01)
Grade (standardized)			
1st-year GPA is below 7	−0.18* (0.11)	0.04 (0.10)	−0.03 (0.11)
Passes course			
1st-year GPA is below 7	−0.07 (0.05)	0.01 (0.05)	−0.03 (0.04)
Course type	7^{+}*vol*	7^{+}*enc*	7^{+}*for*
Observations	927	1,424	1,234

Notes.
1. The outcome variable is displayed at the top of each panel. "Attendance rate" is the percentage of tutorials attended. "Passes course" is a binary variable where pass = 1 and fail = 0.
2. "Course type" refers to how individual courses dealt with above-7 students. "7^{+}*vol*" means above-7 students had full discretion over their attendance. "7^{+}*enc*" means above-7 students were strongly encouraged to attend. "7^{+}*for*" means that above- and below-7 students were penalized for being absent, effectively forcing both groups to attend.
3. The regressions include a first-order polynomial that is interacted with the treatment. The bandwidth is 0.365 and the kernel is triangular.
4. Standard errors are clustered on the student and in parentheses.
5. Significance levels: * < 10%; ** < 5%; *** < 1%.

5.2 | Robustness

We analyzed the robustness of the heterogeneous policy effects across the three course types. Table A.8 in the Supporting Information Appendix tests whether the effects are robust to the inclusion of course–cohort fixed effects and personal characteristics. Table A.9 further includes high school *GPA*, which we only observe for Dutch students.

Both tables show that the baseline results in Table 4 are robust. This is reassuring, especially with respect to the possible gender imbalance at the cutoff.

Table A.10 in the Supporting Information Appendix reports estimates of specifications that use bandwidths that are optimal for each course type. The bandwidths are MSE optimal when estimating the discontinuity at 7. They are CER (coverage error) optimal for the purpose of robust bias-corrected inference, as recommended in Cattaneo, Idrobo, and Titiunik (2019b). The results across all outcomes are very similar, where the estimate in column (1) of the middle panel implies that grades of forced students decrease by 0.26 standard deviations in $7^{+}vol$ courses, which is statistically significant at the 5% level.

Figure A.4 in the Online Appendix explores whether the estimate for $7^{+}vol$ courses is robust to the bandwidth choice. It shows that the estimate on student grades hovers between −0.15 and −0.30 while using bandwidths between 0.10 (first-year GPA of 6.9–7.1) and 0.50 (first-year GPA of 6.5–7.5). Unsurprisingly, the confidence intervals are too wide to reject a null estimate with very small bandwidths. The baseline estimate, however, is statistically significant at bandwidths between 0.15 and 0.40, where the *p*-value is slightly above 10% for the largest bandwidths. We also tested for discontinuities at the fake cutoffs of 6, 8, 8.25 (*cum laude*), and 9 in all our main outcomes for $7^{+}vol$ courses. Table A.11 in the Supporting Information Appendix documents an absence of significant discontinuities across all student outcomes and all fake cutoffs.

5.3 | Abolition cohort

Table 5 reports mean unstandardized $7^{+}vol$ grades for just below and just above 7 students in the treated and abolition cohorts. The top row shows a grade difference of 0.37 (on a 10-point scale) across below- and above-7 students in treated cohorts. The bottom row shows a grade difference of 0.13 for the abolition cohort. The grade difference for the abolition cohort is statistically insignificant. It is approximately one third of its analog for treated cohorts. The evidence is consistent with no grade difference in the abolition cohort or with a grade difference that is abnormally small.

The left-hand column shows that below-7 students from the abolition cohort have grades that are 0.35 points higher than the grades of below-7 students from earlier treated cohorts. The across-cohort difference in the left-hand column is similar to the within-cohort difference of 0.37 in the top row. The grade decrease we observe therefore reflects behavioral changes by forced students rather than behavioral changes by unforced students.

The right-hand column of Table 5 supports this conclusion, showing that the grades of above-7 students from the abolition cohort are 0.11 points higher than the grades of above-7 students from earlier treated cohorts. The difference is statistically insignificant and small relative to other differences in the table. If cohort-specific differences are negligible, then no grade difference for above-7 students would be consistent with no spillovers from forced to unforced students (Dong & Lewbel, 2015). This would suggest it is the behavior of forced students themselves that drives the grade decrease in $7^{+}vol$ courses.

TABLE 5 Unstandardized grades above and below 7, both before and after the abolition

	First-year GPA		
Cohort	**[6.9–7.0]**		**[7.0–7.1]**
2009–2013	6.40	$p = 0.004$***	6.77
	($N = 161$)		($N = 146$)
	$p = 0.126$		$p = 0.487$
2014	6.75	$p = 0.651$	6.88
	($N = 38$)		($N = 61$)

Notes.

1. Local averages of unstandardized grades for a bandwidth of 0.1. The number of observations used to calculate the averages are displayed in parentheses.
2. Averages are for the $7^{+}vol$courses only, which are the courses where above-7 students had full discretion over their attendance during the policy.
3. The *p*-values refer to two-sided significance tests for the difference means.
4. Significance levels: *<10%; **<5%; ***<1%.

6 | BASELINE MECHANISMS

6.1 | Tutorial quality

The grade decrease in $7^{+}vol$ courses may be attributable to especially poor tutorial quality. This might also explain why there is a grade decrease in $7^{+}vol$ courses and no grade difference in $7^{+}enc$ courses. We investigate this possibility using TA evaluations as proxy for tutorial quality, regressing these evaluations on indicators for the three different course types. If $7^{+}vol$ tutorials are indeed poor or ineffective, then we expect lower TA evaluations in these courses. Estimates are found in Table A.12 of the Supporting Information Appendix. Note that we use TA evaluations from the abolition year to circumvent concerns about whether evaluations are contaminated by forced attendance.

Column (1) shows $7^{+}vol$ TAs score 0.21 points higher than $7^{+}enc$ TAs ($p < 0.1$) on the question "TA gives good tutorials." They score about the same relative to TAs in $7^{+}for$ courses. Column (2) implies that the TAs across all three course types provide similar levels of assistance. We conclude that TA quality is in fact moderate to relatively high in $7^{+}vol$ courses.

The grade decrease in $7^{+}vol$ courses may be attributable more broadly to course design. Course instructors may give above-7 students discretion over tutorial attendance and consequently ensure that all students could obtain everything they needed to know via the plenary lectures alone. In this scenario, the TAs for $7^{+}vol$ courses can be excellent yet contribute little to student performance. Two pieces of evidence contradict this possibility. Section 6.4 will show first that the university policy generated parallel increases in lecture and tutorial attendance. If the lectures for $7^{+}vol$ courses were exceptionally useful then grades should have been higher, rather than lower, for forced students. Second, we regress student perceptions of lecturer quality on indicators for the three different course types again using data from the abolition year (columns (3) and (4) of Table A.12 in the Supporting Information Appendix). If lecturers made their courses lecture-heavy, then we expect higher perceived lecturer quality in these courses. Yet we find that the perceived lecturer quality is the same across the three course types.

6.2 | Attendance price and propensity

We investigate whether our treatment effects differ depending on the distance to the university and on the propensity to attend first-year tutorials. We first estimate

$$A_{ij} = \gamma_0 + \gamma_{1i} D_i + f_{+}(GPA_i - 7) + f_{-}(GPA_i - 7) D_i + \varepsilon_{ij}, \tag{4}$$

where A_{ij} is the percentage of tutorials attended in the second year. If γ_{1i} is large then the student's desired attendance is low, such that they would have attended far fewer tutorials in the absence of forced attendance. Alternatively, a small γ_{1i} implies attendance is desirable, such that the student attends the same number of tutorials with or without forced attendance. In the parlance of the treatment effects literature (Angrist & Pischke, 2008), students who otherwise prefer not to attend (large γ_{1i}) are compliers. Students who would attend anyways (small γ_{1i}) are always takers. There are no never-takers or defiers by the very definition of the policy, as it leaves students with no choice but to attend tutorials when their first-year *GPA* is below 7. Indeed, below-7 students collectively failed to meet the 70% criteria in only 0.44 percent of their courses.[13]

We interpret distance to the university as a proxy for the price of attendance and average tutorial attendance in the first year as a proxy for the additional utility from attendance. Distant students pay a higher attendance price because they have to spend more time traveling to campus. Students with a high attendance propensity in the first year presumably derive additional utility from attendance in the second year. We thus operationalize γ_{1i} via treatment interactions with our proxies for the price of and additional utility from attendance. Estimates are found in the first three columns

[13]One might argue that the grade for never-takers is never observed, as they cannot write the exam. However, in Section 4.3 we showed that students generally participate in every second-year course, and that their near-perfect course participation is unaffected by the treatment (leaving no room for never-takers).

TABLE 6 Heterogeneous effects by distance and first-year attendance

	Attendance rate			Grade (standardized)			Passes course		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1st-year GPA is below 7	0.34***	0.14***	0.00	−0.19*	0.03	−0.02	−0.08*	0.01	−0.03
	(0.04)	(0.03)	(0.01)	(0.10)	(0.10)	(0.11)	(0.04)	(0.05)	(0.04)
Distance to university (standardized)	−0.05**	−0.05***	−0.01	0.04	−0.06	−0.00	0.02	−0.01	−0.00
	(0.02)	(0.02)	(0.01)	(0.04)	(0.07)	(0.04)	(0.02)	(0.03)	(0.01)
Distance × treatment	0.06**	0.05***	−0.01	−0.02	0.09	−0.02	0.00	0.02	−0.00
	(0.02)	(0.02)	(0.02)	(0.06)	(0.08)	(0.07)	(0.02)	(0.04)	(0.02)
Attendance in first year (standardized)	0.15***	0.07***	0.02***	−0.05	−0.05	0.09	−0.02	−0.01	0.03*
	(0.02)	(0.02)	(0.01)	(0.06)	(0.04)	(0.05)	(0.02)	(0.02)	(0.02)
Attendance in first year × treatment	−0.13***	−0.04**	−0.00	−0.01	0.02	−0.01	0.01	0.02	−0.02
	(0.02)	(0.02)	(0.01)	(0.08)	(0.06)	(0.07)	(0.03)	(0.02)	(0.03)
Course type	7^+vol	7^+enc	7^+for	7^+vol	7^+enc	7^+for	7^+vol	7^+enc	7^+for
Observations	927	1,424	1,234	927	1,424	1,234	927	1,424	1,234

Notes.

1. "Attendance" rate is the percentage of tutorials attended.
2. "Course type" refers to how individual courses dealt with above-7 students. "7^+vol" means above-7 students had full discretion over their attendance. "7^+enc" means above-7 students were strongly encouraged to attend. "7^+for" means that above- and below-7 students were penalized for being absent, effectively forcing both groups to attend.
3. The regressions include a first-order polynomial which is interacted with the treatment. The bandwidth is 0.365 and the kernel is triangular.
4. Distance and attendance in first year are standardized, where the standard deviations are 30.3 km for distance and 0.07 for attendance (on a scale from 0 to 1).
5. Standard errors are clustered on the student and in parentheses.
6. Significance levels: *<10%; **<5%; ***<1%.

of Table 6, where column (1) focuses on 7^+vol courses. Note that distance and first-year attendance are standardized, where the standard deviations are 30.3 km for distance and 0.07 for attendance (on a scale from 0 to 1).

Three patterns stand out. First, the direct effect of the proxy is always opposite, but similar in magnitude, to its interaction effect. This suggests the interactions pick up the student's counterfactual attendance had the policy not been in place. Second, the policy had a larger effect on students who live far from campus. The attendance effect increases by 6 percentage points for students that live one standard deviation further from campus. This suggests distant students have a greater propensity to attend less in the absence of forced attendance. Third, the policy had a smaller effect on students who have a higher attendance propensity. The attendance effect decreases by 13 percentage points for students who attended one standard deviation more tutorials in the first year. The last two patterns are consistent with students making calculated attendance decisions.

6.3 | Differential grade effects

The differential effects on tutorial attendance are consistent with the university policy constraining calculated decisions by forced students. We check for similar differential effects on academic performance. Our idea is that, if the additional constraint on attendance drives the grade decreases in 7^+vol courses, then grades should decrease by more for students who live far from campus and who have a low propensity for tutorial attendance in the first year. Columns (4)–(6) of Table 6 show the heterogeneity results for grades and columns (7)–(9) do so for passing rates. Columns (4) and (7) focus on the sample of 7^+vol courses.

The results imply that the interaction effects for distance and attendance propensity on academic performance are both statistically and substantively small. While the estimates fail to support a mechanism where grades decrease because the policy constrains student behavior, it is not necessarily inconsistent with this mechanism. Students may be compensating for the lost time and energy in a variety of unobserved ways. For example, distant students may use their additional travel time to study the material.

TABLE 7 Lecture attendance and total study time

	(1)	(2)	(3)
Attended lectures			
1st-year *GPA* is below 7	0.25	0.08	−0.05
	(0.17)	(0.09)	(0.07)
Intercept	0.59***	0.87***	0.95***
	(0.13)	(0.07)	(0.04)
Observations	170	292	272
Total study time			
1st-year *GPA* is below 7	1.98	4.54	2.12
	(3.53)	(3.71)	(3.40)
Intercept	11.00***	15.13***	13.44***
	(2.54)	(1.97)	(2.09)
Observations	170	292	272
Course type	$7^{+}vol$	$7^{+}enc$	$7^{+}for$

Notes.
1. "Attended Lectures" is a binary variable based on the answer to "Have you attended lectures?" "Total study time" is an ordinal variable based on the answer to "Average study time (hours) for this course per week (lectures + tutorials + self-study)?" The maximum for the interval was used to convert the categories into hours.
2. The regressions include a first-order polynomial that is interacted with the treatment. The bandwidth is 0.365 and the kernel is triangular.
3. Intercepts approximate the outcome mean near the threshold of students right of seven.
4. Standard errors are clustered on the student and are in parentheses.
5. Significance levels: *<10%' **<5%; ***<1%.

6.4 | Other input choices

To better understand the policy impact on student input choices, Table 7 investigates the effect on self-reported lecture attendance and total study time. The top panel reports the effect on an indicator for whether the student attended lectures. The bottom panel reports the effect on total study time (lectures+tutorials+self study).

Forced students are 25 percentage points more likely to attend lectures in $7^{+}vol$ courses ($p \approx 0.11$), 8 percentage points more likely in $7^{+}enc$ courses ($p > 0.10$), and 5 percentage points less likely in $7^{+}for$ courses ($p > 0.10$). The slope estimates, while insignificant, align well with how tutorial attendance changed across the three course types (top panel of Table 4). The slope estimates for lecture and tutorial attendance are both largest in $7^{+}vol$ courses and smallest in $7^{+}for$ courses, and have similar orders of magnitudes. The intercept estimates of Table 7 also align well with the intercepts for tutorial attendance (left-hand panel of Table 1, right of 7). The similarities between the slopes and intercepts suggest the policy forces students to pay a time cost that becomes sunk after they arrive at campus, such that lecture attendance is relatively cheap when the student is already there.

The bottom panel of Table 7 also shows that the policy increased total study time by about 2 hr in $7^{+}vol$ courses, 4.5 hr in $7^{+}enc$ courses, and 2 hr in $7^{+}for$ courses. The estimates are all statistically insignificant, implying that we cannot rule out no effect of the university policy on total study time. A null or small positive effect on total study time, together with large attendance increases for tutorials and lectures, would imply that the university policy decreased time spent on self study. Less time on self-study would further suggest that inputs other than tutorial attendance were affected by the policy.[14] This explanation for the grade decrease fits well with the careful time use study of Stinebrickner and Stinebrickner (2008). They show that a 1 hr reduction in self-study (in the first semester) causes *GPA* to decrease by 0.36 points.

[14]Although our estimates suggest a decline in self-study, we do not make a precise calculation because course evaluations are completed by 20% of students and because we converted the categories of the total study time variable (1 = 0 hr, 2 = 1–5 hr, and 10 = more than 40 hr) into hours based on the maximum for the interval.

TABLE 8 Peer exposure and peer effects

	(1)	(2)	(3)	(4)	(5)	(6)
Exposure to forced peers						
1st-year *GPA* is below 7	0.24***	0.02	0.12***	0.03**	−0.01	−0.01
	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)	(0.03)
Attendance rate		0.70***		0.68***		0.46***
		(0.02)		(0.01)		(0.05)
Mean dep. var.	0.56	0.56	0.56	0.56	0.61	0.61
Observations	926	926	1421	1421	1231	1231
Grades (standardized)						
1st-year *GPA* is below 7	−0.18*	−0.17	0.03	0.03	−0.03	−0.04
	(0.11)	(0.11)	(0.11)	(0.10)	(0.11)	(0.11)
Peer average 1st-year *GPA*	0.01		−0.03		0.06*	
	(0.04)		(0.04)		(0.03)	
Peer avg. GPA × treatment	−0.04		0.06		−0.00	
	(0.06)		(0.06)		(0.05)	
Peer average registration time		−0.01		0.01		0.01
		(0.01)		(0.01)		(0.01)
Peer avg. registration time × treatment		0.00		−0.01		−0.01
		(0.01)		(0.01)		(0.01)
Observations	927	927	1424	1424	1234	1234
Course type	$7^{+}vol$	$7^{+}vol$	$7^{+}enc$	$7^{+}enc$	$7^{+}for$	$7^{+}for$

Notes.
1. The exposure variable is missing if nobody within a tutorial group attended any of the sessions. This explains the slightly fewer number of observations compared to the baseline regressions by course type (compared to, e.g., the bottom panel).
2. The regressions include a first-order polynomial that is interacted with the treatment. The bandwidth is 0.365 and the kernel is triangular.
3. "Attendance rate" refers to the percentage of tutorials attended (top panel).
4. Peer group averages are leave-out means (bottom panel). Peer average 1st-year GPA is standardized with mean 0 and standard deviation 1 and average peer tutorial registration time is measured in differences in days from the course mean registration time.
5. Standard errors are clustered on the student and in parentheses.
6. Significance levels: *<10%; **<5%; ***<1%.

6.5 | Peer effects

By forcing tutorial attendance, the policy increases the exposure of forced students to other forced and therefore relatively low-achieving students. Additional exposure to low achievers can also explain the grade decrease in $7^{+}vol$ courses. As a first step towards understanding the importance of this mechanism, we evaluate whether there are indeed differences in exposure to forced students. We use our rich attendance data to construct an exposure measure for student i in course j:

$$\text{Exposure}_{ij} = \frac{1}{S_j}\sum_{s=1}^{S_j} \mathbb{1}[A_{isj}=1]\left(\frac{\mathcal{A}^{F}_{-isj}}{\mathcal{A}_{-isj}}\right),$$

where S_j is the total number of tutorial sessions in course j, A_{isj} is the attendance of i in session s, and $\mathbb{1}$ denotes the indicator function. $\frac{\mathcal{A}^{F}_{-isj}}{\mathcal{A}_{-isj}}$ is the leave-out proportion of forced students who attended a specific tutorial session, where $\mathcal{A}_{-isj}$ is the number of students besides i who attended session s, and $\mathcal{A}^{F}_{-isj} = \sum_{k\neq i}^{\mathcal{A}_{-isj}} \mathbb{1}[A_{ksj}=1]\mathbb{1}[D_k=1]$ is the number of forced students besides i who attended session s. We then use the treatment effect on Exposure_{ij} to quantify the additional exposure of forced students.

Estimates are found in the odd-numbered columns of the top panel of Table 8. The policy increased exposure by 24 percentage points in $7^{+}vol$ courses ($p < 0.01$), by 12 percentage points in $7^{+}enc$ courses ($p < 0.01$), and had no effect on exposure in $7^{+}for$ courses.

Our exposure measure stresses two channels for the increased exposure in $7^{+}vol$ and $7^{+}enc$ courses. One relates to the simple fact that forced students are more likely to attend tutorials. The other channel relates to the possibility that forced students may be more likely to attend tutorials with other forced students even after conditioning on attendance probabilities. This can be the case if forced students deliberately register for the same tutorial group or attend the same tutorial sessions. These sorts of coordination can foster bad peer influence among forced students.

We assess these channels separately by estimating specifications that control for the course-specific attendance rate of the student in the even-numbered columns of the top panel of Table 8. The exposure differences are much smaller (close to 0 in fact) once we control for attendance rates, consistent with the unconditional treatment effect reflecting a mechanical increase in attendance rates rather than increased and deliberate coordination with other low-achieving peers. The evidence suggests in turn that if peer effects are present, they are not operating through the coordination decisions of forced students.

We also evaluated the potential importance of peers using the most common peer effects specification (Booij, Leuven, & Oosterbeek, 2017). More specifically, we reestimated our baseline equation while including treatment interactions with measures of peer quality, the average first-year GPA of the peer group, and the average peer registration time for tutorials. Both are leave-out means, where the average first-year GPA is standardized, and the tutorial registration time is measured in differences in days from the course mean registration time and subsequently averaged across one's peers. The latter measure reflects the idea that weak students might leave tutorial registration to the last minute.

The bottom panel of Table 8 shows the results, where all the effects of treatment interactions with peer quality are modest. All the estimates are statistically insignificant at conventional levels, while the main treatment effect estimate is unchanged compared to our baseline specifications. Negligible peer effects are unsurprising given recent discussions in the literature (Booij et al., 2017; Feld & Zölitz, 2017; Sacerdote, 2014). Altogether, the evidence suggests that relatively heavy exposure to forced peers is not an important mechanism for the grade decrease in $7^{+}vol$ courses.

7 | CONCLUSION

We draw on a discontinuity at a large public Dutch university, wherein second-year students with a first-year *GPA* below 7 were allocated to a full year of forced, frequent, and regular attendance, to estimate the causal effect of additional structure on academic performance. Our estimates imply that forced students, with a first-year *GPA* at 7, cannot expect a positive effect on their *GPA* in the second year. The average null estimate masks differential effects that are attributable to how course instructors dealt with above-7 students. The policy had its largest effects in courses where above-7 students were allowed to decide their attendance, as the attendance of forced students increased by more than 50%, and their grades decreased by about 0.16–0.26 standard deviations.

We find some evidence that the grade decreases are explained by the constraining effects of the policy—that is, that the policy prevented students from attaining their desired mix of study inputs. We rule out several other mechanisms, including the importance of an increase in exposure to other forced (and relatively low-achieving) peers, as well as the possibility of course heterogeneity in the usefulness of tutorials, or heterogeneity in course design more generally.

ACKNOWLEDGMENTS

We thank Suzanne Bijkerk, Robert Dur, Julian Emami Namini, Johanna Posch, and Philip Oreopoulos for helpful comments and suggestions. The paper has also benefited from the comments and suggestions of participants at EEA-ESEM 2016, Erasmus University Rotterdam Seminar Series, IZA Summer School 2017, IZA Workshop on the Economics of Education, and the Tinbergen Institute Seminar Series. The authors have no relevant or material financial interests that relate to the research described in this paper. Oosterveen acknowledges financial support from the FCT—Fundação para a Ciência e Tecnologia (grant PTDC/EGE-OGE/28603/2017). All omissions and errors are our own.

OPEN RESEARCH BADGES

This article has earned an Open Data Badge for making publicly available the digitally-shareable data necessary to reproduce the reported results. The data is available at [http://qed.econ.queensu.ca/jae/datasets/kapoor001/].

REFERENCES

Angrist, J. D., Oreopoulos, P., & Williams, T. (2014). When opportunity knocks, who answers? New evidence on college achievement awards. *Journal of Human Resources*, *1*(1), 1–29.

Angrist, J. D., & Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton, NJ: Princeton University Press.

Booij, A. S., Leuven, E., & Oosterbeek, H. (2017). Ability peer effects in university: Evidence from a randomized experiment. *Review of Economic Studies*, *84*(2), 547–578.

Calonico, S., Cattaneo, M. D., & Farrell, M. H. (2019). Optimal bandwidth choice for robust bias corrected inference in regression discontinuity designs (Working paper).

Calonico, S., Cattaneo, M. D., Farrell, M. H., & Titiunik, R. (2017). rdrobust: Software for regression-discontinuity designs. *The Stata Journal*, *17*(2), 372–404.

Castleman, B. L. (2014). Prompts, personalization, and pay-offs: Strategies to improve the design of college and financial aid information (Working paper). Washington, DC: The George Washington University Graduate School of Education and Human Development.

Cattaneo, M. D., Idrobo, N., & Titiunik, R. (2019a). *A practical introduction to regression discontinuity designs: Extensions*. Cambridge Elements: Quantitative and Computational Methods for Social Science. Cambridge, UK: Cambridge University Press.

Cattaneo, M. D., Idrobo, N., & Titiunik, R. (2019b). *A practical introduction to regression discontinuity designs: Foundations*. Cambridge Elements: Quantitative and Computational Methods for Social Science. Cambridge, UK: Cambridge University Press.

Cattaneo, M. D., Jansson, M., & Ma, X. (2018). Manipulation testing based on density discontinuity. *The Stata Journal*, *18*(1), 234–261.

Cattaneo, M. D., Jansson, M., & Ma, X. (2019). Simple local polynomial density estimators. *Journal of the American Statistical Association*. DOI https://doi.org/10.1080/01621459.2019.1635480. Advance online publication.

Chen, J., & Lin, T.-F. (2008). Class attendance and exam performance: A randomized experiment. *Journal of Economic Education*, *39*(3), 213–227.

Cohodes, S., & Goodman, J. (2014). Merit aid, college quality and college completion: Massachusetts' Adams scholarship as an in-kind subsidy. *American Economic Journal: Applied Economics*, *6*(4), 251–285.

Correspondence (1994). Correspondence: Should class attendance be mandatory? *Journal of Economic Perspectives*, *8*(3), 205–216.

Cunha, F., Karahan, F., & Soares, I. (2011). Returns to skills and the college premium. *Journal of Money, Credit and Banking*, *43*(s1), 39–86.

De Paola, M., Scoppa, V., & Nistico, R. (2012). Monetary incentives and student achievement in a depressed labor market: Results from a randomized experiment. *Journal of Human Capital*, *6*(1), 56–85.

Desalegn, A. A., Berhan, A., & Berhan, Y. (2014). Absenteeism among medical and health science undergraduate students at Hawassa University, Ethiopia. *BMC Medical Education*, *14*(1), 81.

Dobkin, C., Gil, R., & Marion, J. (2010). Skipping class in college and exam performance: Evidence from a regression discontinuity classroom experiment. *Economics of Education Review*, *29*(4), 566–575.

Dong, Y., & Lewbel, A. (2015). Identifying the effect of changing the policy threshold in regression discontinuity models. *Review of Economics and Statistics*, *97*(5), 1081–1092.

Durden, G. C., & Ellis, L. V. (1995). The effects of attendance on student learning in principles of economics. *American Economic Review*, *85*(2), 343–346.

Dynarski, S. (2008). Building the stock of college-educated labor. *Journal of Human Resources*, *43*(3), 924–937.

Feld, J., & Zölitz, U. (2017). Understanding peer effects-on the nature, estimation and channels of peer effects. *Journal of Labor Economics*, *35*(2), 387–428.

Gneezy, U., Meier, S., & Rey-Biel, P. (2011). When and why incentives (don't) work to modify behavior. *Journal of Economic Perspectives*, *25*(4), 191–209.

Hahn, J., Todd, P., & Van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, *69*(1), 201–209.

Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, *142*(2), 615–635.

Jones, E. B., & Jackson, J. D. (1990). College grades and labor market rewards. *Journal of Human Resources*, *25*(2), 253–266.

Kirby, A., & McElroy, B. (2003). The effect of attendance on grade for first year economics students in University College Cork. *Economic and Social Review*, *34*(3), 311–326.

Kottasz, R. (2005). Reasons for student non-attendance at lectures and tutorials: An analysis. *Investigations in University Teaching and Learning*, *2*(2), 5–16.

Latif, E., & Miles, S. (2013). Class attendance and academic performance: A panel data analysis. *Economic Papers*, *32*(4), 470–476.

Lavecchia, A. M., Liu, H., & Oreopoulos, P. (2014). Behavioral economics of education: Progress and possibilities (NBER Working Paper No. 20609). Cambridge, MA: National Bureau of Economic Research.

Lee, D. S. (2008). Randomized experiments from non-random selection in us house elections. *Journal of Econometrics*, *142*(2), 675–697.

Lee, D. S., & Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, *48*(2), 281–355.

Leuven, E., Oosterbeek, H., & van der Klaauw, B. (2010). The effect of financial rewards on students' achievement: Evidence from a randomized experiment. *Journal of the European Economic Association*, *8*(6), 1243–1265.

Lin, T.-F., & Chen, J. (2006). Cumulative class attendance and exam performance. *Applied Economics Letters*, *13*(14), 937–942.

Marburger, D. R. (2001). Absenteeism and undergraduate exam performance. *Journal of Economic Education*, *32*(2), 99–109.

Marburger, D. R. (2006). Does mandatory attendance improve student performance? *Journal of Economic Education*, *37*(2), 148–155.

Martins, P. S., & Walker, I. (2006). Student achievement and university classes: Effects of attendance, size, peers, and teachers (IZA Discussion Paper series No. 2490). Bonn, Germany: Institute of Labor Economics.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, *142*(2), 698–714.

Oreopoulos, P., & Petronijevic, U. (2013). Making college worth it: A review of research on the returns to higher education (NBER Working Paper No. 19053). Cambridge, MA: National Bureau of Economic Research.

Romer, D. (1993). Do students go to class? Should they? *Journal of Economic Perspectives*, *7*(3), 167–174.

Sacerdote, B. (2014). Experimental and quasi-experimental analysis of peer effects: Two steps forward? *Annual Review Economy*, *6*(1), 253–272.

Scott-Clayton, J. (2011). The shapeless river: Does a lack of structure inhibit students' progress at community colleges? (CCRC Working Paper No. 25). New York, NY: Community College Research Center.

Snyder, J. L., Lee-Partridge, J. E., Jarmoszko, A. T., Petkova, O., & D'Onofrio, M. J. (2014). What is the influence of a compulsory attendance policy on absenteeism and performance? *Journal of Education for Business*, *89*(8), 433–440.

Stanca, L. (2006). The effects of attendance on academic performance: Panel data evidence for introductory microeconomics. *Journal of Economic Education*, *37*(3), 251–266.

Stinebrickner, R., & Stinebrickner, T. R. (2008). The causal effect of studying on academic performance. *B.E. Journal of Economic Analysis and Policy*, *8*(1), 1–55.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Kapoor S, Oosterveen M, Webbink D. The price of forced attendance. *J Appl Econ*. 2021;36:209–227. https://doi.org/10.1002/jae.2781