The Price of Forced Attendance

Sacha Kapoor* Matthijs Oosterveen† Dinand Webbink‡

July 5, 2018

Abstract

Recent scholarship has argued that structure - forced, frequent, and regular attendance *e.g.* - may be good for academic performance in higher education. We draw on a natural experiment at a large university to estimate the causal effect of a full year of forced, frequent, and regular attendance on academic performance, and find no evidence for positive effects. Our evidence instead implies students have lower grades and less leisure, both because of direct effects of the policy and because of spillovers on performance in other courses. The policy has enduring effects, as students have lower grades later on, even after regaining discretion over their attendance.

JEL: I23, D12, J22
Keywords: Structure, Mandatory Attendance, Academic Performance, Higher Education

*Department of Economics, Erasmus School of Economics, Erasmus University Rotterdam. kapoor@ese.eur.nl

†Department of Economics, Erasmus School of Economics, Erasmus University Rotterdam. oosterveen@ese.eur.nl

‡Department of Economics, Erasmus School of Economics, Erasmus University Rotterdam, Tinbergen Institute, IZA Bonn. webbink@ese.eur.nl

§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, IZA Summer School 2017, IZA Workshop on the Economics of Education, and the Tinbergen Institute. All omissions and errors are our own.

For many people their first real encounter with autonomy happens at college or university. Out from under the roofs of their parents and high school teachers, how they manage their lives is now largely up to them. Many students use their newfound autonomy to skip class, especially in the early years of their undergraduate education, choosing instead to focus on extracurricular activities, such as student government, watching March Madness, or chasing other young men and women. To combat the rampant absenteeism this newfound autonomy begets,[1] and because of the substantial returns to college performance and graduation [Cunha, Karahan, and Soares, 2011, Jones and Jackson, 1990, Oreopoulos and Petronijevic, 2013], university administrators and instructors often mandate frequent and regular class attendance among their students.[2][3] These attendance policies provide students with structure, helping them to circumvent behavioral predispositions towards non-academic activities, and ultimately to avoid decisions that can be bad for their lifetime utility [Lavecchia, Liu, and 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) which 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.

In this article we argue that additional structure is, in fact, bad for the performance of relatively good students. To make this argument, 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 percent 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

[1] Student absenteeism can be upwards of 60 percent of classes [Desalegn, Berhan, and Berhan, 2014, Kottasz et al., 2005, Romer, 1993].

[2] An early discussion of mandatory attendance can be found in the correspondence section of the Journal of Economic Perspectives in 1994 [Correspondence, 1994]. Motivated by Romer [1993], it consists of short letters by economics professors detailing their use of mandatory attendance.

[3] American universities spend 33 percent of their total budget on student instruction. This amounts to 56.7 billion dollars (for private nonprofit universities, years 2013-2014). Obtained via NCES: `https://nces.ed.gov/programs/digest/d15/tables/dt15_334.40.asp`, retrieved on 15-02-2017.

traveling and attending tutorials over a full academic year, amounting to approximately 7 additional hours per week. Students who fail to meet the attendance requirement face a stiff penalty, as they are not allowed to write the final exam for their course and must wait a full academic year before they can take the course again. Because students have imprecise control over their *average* grade in the first year, the experiment facilitates a regression discontinuity design [Lee, 2008, Lee and Lemieux, 2010] for identifying the various 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.[4] The policy we study is well within confines of this definition.[5] The policy asks students to come to campus frequently and regularly, choices which 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, as this was how it was communicated to them by the university. Our data supports the notion that attendance was forced, as below-7 students collectively failed to meet the 70 percent criteria in less than one half of one percent of their courses. A more severe penalty, death *e.g.*, would have increased participation by less than half a percent, in other words.

Our estimates imply that forced students can expect a GPA decrease of 0.20 standard deviations over the remainder of their undergraduate degree. They can expect a decline of 0.15 standard deviations in their second year, when their attendance is forced, and a decline of 0.25 deviations in their third and last year, when they regain the right to decide their attendance. While the negative effects on second and third year performance are marginally significant or insignificant at conventional levels, we are able to rule out

[4] Our paper is about more than just the role of sticks versus carrots in university education. A stick is typically defined as a penalty on performance, which itself is determined by choices and luck. Sticks constrain choices only implicitly, as the decision maker still has the freedom to make "bad" choices, and can simply hope that good luck helps them avoid penalties for poor performance.

[5] Our definition differs from the notion of labor coercion, which focuses on how physical force or the threat of it influences labor market institutions and outcomes.

positive effects in the ranges of 0.05 and 0.1 standard deviations.

The average effect on second-year performance aggregates differential effects across all courses. While the university required all students below 7 to attend 70 percent of tutorials in all their second-year courses, it had no policy on how students above 7 should be treated. Unsurprisingly, several courses overlaid their own attendance initiatives onto the university policy, each differing in the intensity of the attendance constraint they imposed on students who scored above 7 in first year. Some courses penalized absenteeism by any student (absence-penalized courses), others strongly intimated and explained why all students should attend (attendance-encouraged courses), while a third group of courses followed the university policy and left the attendance decision up to above-7 students (attendance-voluntary courses). We observe the same students in all three scenarios because students have no discretion over course choice in second year.

The university policy had its largest effects in courses where attendance was voluntary for above-7 students. In these attendance-voluntary courses, it increased attendance by more than 50 percent, significantly decreased grades by 0.35 standard deviations, and significantly decreased the chances of passing by more than 10 percentage points. Self reports of total study time suggest further that forced students spend less time on non-academic activities such as leisure.

We delve into mechanisms behind the significant grade decreases in attendance-voluntary courses. We argue that the university policy forces students to spend a substantial number of hours in a specific way, leaving them less time for other activities, including activities which are important for grades. Grades decrease because the grade loss from spending less time on other academic activities outweighs the grade gains from additional attendance. What we observe fits with a model where students care about their grades and make informed decisions about their attendance. The latter is reinforced by our complier analysis, which identifies the most affected students, and shows that the largest policy effects are on the attendance of students who live far from campus or had a greater propensity to miss tutorials in first year.

In addition to aggregating the differential effects across all courses, the average second-year effect aggregates spillovers across courses taken concurrently. Forced students have slightly lower grades and passing rates in absence-penalized courses, even though they are not disadvantaged in their attendance decisions, having the same attendance rates as above-7 students. We explain that absence-penalized courses are always taken concurrently with a course where forced students are at an attendance disadvantage, arguing in turn that the grade decreases are consistent with negative spillovers from these courses. The spillovers, together with our results for activities other than attendance (*e.g.* leisure), suggest that the policy effect on grades does not operate through attendance alone.[6]

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 show there was no grade difference near 7 for the abolition cohort. No grade difference for this cohort provides additional evidence against differential sorting of forced students into second year. More generally, it helps us rule out shocks other than the policy as drivers of worse performance just below the threshold. It also helps us show our results are driven by worse performance among forced students rather than better performance among above-7 students. Finally, it supports the presence of spillovers during years when the policy was in place.

Our study contributes to an expanding literature on incentives in education. A good deal of recent work analyzes the effects of interventions that reward students financially for "good" choices or better academic performance [Angrist, Oreopoulos, and Williams, 2014, Castleman, 2014, Cohodes and Goodman, 2014, De Paola, Scoppa, and Nistico, 2012, Dynarski, 2008, Leuven, Oosterbeek, and van der Klaauw, 2010].[7] We instead analyze the effect of an intervention which penalizes students heavily for "bad" choices, where the penalty is in terms of time rather than money.

[6]We use all three courses, including the attendance-encouraged courses, to investigate other mechanisms, such as direct policy effects on self-perception or identity or stigmatization by other students, general discontent with the policy itself, negative peer effects, or course-level differences in the usefulness of tutorials. Our results imply these other mechanisms are unimportant.

[7]For more comprehensive lists, at all levels of education, see Lavecchia, Liu, and Oreopoulos [2014] and Gneezy, Meier, and Rey-Biel [2011].

Our findings contribute to debates over the merits of mandatory attendance in higher education [Romer, 1993].[8] The argument for mandatory attendance is based on a robust positive correlation between grades and attendance.[9] The argument has been reinforced by studies showing positive correlations between mandatory attendance and grades (see *e.g.* Marburger [2006] and Snyder et al. [2014]). We estimate the causal effects of a large-scale mandatory attendance policy and find negative effects.

One explanation for the discrepancy relates to the weight of the constraint imposed by the policy we study. A hefty constraint, spanning a full academic year, makes a negative finding more plausible. Another explanation relates to identification concerns in other studies. Previous research has relied on either year-over-year comparisons of students from different cohorts, or a discontinuity that allocates students to mandatory attendance later on in the same course. These strategies are problematic because the estimated effects of mandatory attendance may instead be attributable to heterogeneity across cohorts or students' initial efforts to avoid mandatory attendance later on. Our context allows for within-cohort comparisons and lets us deal with anticipation effects.

This article contributes, more generally, to debates over the role of structure in higher education [Lavecchia, Liu, and Oreopoulos, 2014, Scott-Clayton, 2011]. Arguments for additional structure usually focus on student predispositions towards non-academic activities, emanating from behavioral biases such as impatience, or imperfect information about behaviors that engender success at university. Our findings imply structure is detrimental to students with a GPA of 7, as well as to students with a GPA around 7 [Cerulli et al., 2017]; above-average students at a prominent university in the Netherlands. From this perspective, our contribution is in showing that the cost of structure in higher education is lower academic performance among relatively good students.

[8]Our study has an indirect link with the compulsory schooling literature [Angrist and Krueger, 1991, Oreopoulos, 2007]. We also examine the effect of a policy that penalizes people for specific choices. We differ in that our focus is on attendance at university, with steep and enforced penalties for absenteeism, and that we show that such policies can be very costly for students.

[9]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].

1 Context

Our venue is the economics undergraduate program at a large public university in the Netherlands. 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 the courses they take 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 Appendix). Students have discretion over their courses in 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 Masters program.[10] The economics program is given in both Dutch and English. The only difference between the programs is that the Dutch program has approximately 2.5 times more students.

Academic years are divided into five blocks, of eight weeks each (seven weeks of teaching and one 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.[11] Heavy courses have three large-scale lectures per week, while light courses have two. Lecture attendance is always voluntary. Heavy courses have two small-scale tutorials ($\approx$ 30 students) per week, while light courses have one. Lectures and tutorials both last for 1 hour and 45 minutes. Unlike lectures, but much like what may be found in structured college programs, tutorials require preparation and active participation of the student, via *e.g.* 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 the time of registration, students are unaware of the teaching assistant (TA)

[10]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 three rather than four years of bachelors before a Masters.

[11]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 hours of studying. 60 ECTS account for one year of study.

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, including the ones scored above 7 in the first year. 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 voluntary students.

Grading is done on a scale that ranges from 1 to 10. Students fail a course if their grade is below 5.5. The average grade in the first year is weighted by the amount of credits the student gets for completing the course.

1.1. University Policy. Second-year students must attend 70 percent of tutorials for all of their second-year courses if they:

1. had an average grade (weighted by course credits) of less than 7 in first year;

2. failed at least one of their 10 first year courses.[12]

The table summarizes the students who had to comply with the policy. Students were not allowed to write the final exam and had to wait a full year before retaking the course if they failed to fulfill the 70 percent attendance requirement.

Our analysis focuses on the sample of students who completed the first year on time because first year completion rates for students around the cutoff is 92 percent.[13] In this our primary estimation 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 means imply that the university policy assigns above-average students to forced attendance and, because the university is one of the more prominent universities in the Netherlands, that our findings apply to populations of relatively good students.

[12]Courses are grouped (Table A.1) such that a student can compensate a failing grade of between 4.5 and 5.4 from one course with a passing grade from another. This applies to all students, whether they are above or below the threshold of 7. A student who receives an 8 in microeconomics and 4.5 in macroeconomics can, in effect, take 1 point from their micro grade and use it towards their macro grade.

[13]In principle, one could estimate a local difference-in-difference, comparing changes in the grades of these students, around the cutoff, with changes in the grades of students who did not complete the first year. We did not do this because completion rates were so high near seven.

Completed first year	GPA< 7	GPA ≥ 7
Yes	Forced	Free
No	Forced	Forced

The policy imposes sizeable time costs on students. Forced students must spend 26 hours per block (3.5 hours per week) in tutorials.[14] Once we account for the travel time of the average student, about 45 minutes each way,[15] forced students must spend 50 hours per block traveling to and attending tutorials.[16] All costs are in terms of time rather than money because student travel is fully subsidized in the Netherlands.

Students were made aware of the policy in their first year. Incoming students are assigned to tutors who, among other things, explain the policy to them. Student awareness facilitates adjustments in anticipation of forced attendance in the second year. As we explain later, our identification strategy is robust to anticipation effects as long as the average grade in first year is somewhat outside the student's control.

The introduction of the policy had nothing to do with the historical grade distribution of first-year students. The policy was introduced as part of a university-wide initiative to personalize education via small-scale tutorials. The initiative came about for three reasons: (i) the university had grown to a scale that made education impersonal; (ii) tutorials encourage active participation; (iii) 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 sizeable investment in small-scale tutorials.

1.2. Course Policies. While the university forced the attendance of all below-7 students in all their second-year courses, courses differed in how they dealt with above-7 students. Table 1 provides a detailed overview on the courses and, in particular, on how they dealt

[14] This is based on the fact that there are 3 tutorials of 1.75 hour per week, 7 non-exam weeks in a block, and that students must attend 70 percent of tutorials.

[15] The average student lives 22.9 kilometers from campus. From the Dutch student survey "Studenten Monitor" we observe that more than 70 percent of university students travel by public transport (http://www.studentenmonitor.nl/). To get an idea of the travel time, we used the Dutch public transport website (http://9292.nl/) to check travel times between the university and the few larger cities within a radius of 20 and 30 kilometers of the university.

[16] 50 hours is a lower bound, as it ignores the preparation time for active participation in tutorials.

with these students. Attendance was voluntary in two of the courses. Three courses strongly encouraged these students to attend. Three courses penalized them, and in fact also the below-7 students, for not attending. In this last set of courses, students had to complete assignments at the tutorials that made up five to thirty percent 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 to 9.5 (rather than 10). The remaining two courses had no tutorials, and the final grade (mostly) consist of writing a research report in groups. Accordingly, these two courses are excluded from our analysis.[17]

Note that because second-year students have no discretion over course choice, the pool of treated (and control) students is the same across the three types of courses. The lack of choice leaves no room for differential selection of voluntary students into one type of course or another.[18] Ultimately, the course policies provide us with three counterfactuals: the grades of students whose attendance is voluntary, strongly-encouraged, and penalized. The three counterfactuals help us sort through various mechanisms which can generate and foster a relationship between forced attendance and academic performance.

1.3. Abolition. The policy lasted five years, starting in 2009-10 and ending 2013-14. Thus, the 2008-09 cohort was the first to be subjected to the policy in their second year, while 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 second year.

[17]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.

[18]Table 1 also shows multiple choice questions are used on the exams of all but one course. This precludes TAs from having a direct effect on grades.

2 Data

Our main information source is the administrative data of the university. Our sample ranges from the 2008-09 academic year until 2014-15. We observe grades at the level of the student for all three of their undergraduate years, tutorial attendance for the first two years, course evaluations, and various personal characteristics. After restricting the sample to be within 0.5 grade points of 7, our main estimation sample, we have 5000 course-student observations based on more than 700 students.

The university uses attendance lists to track the attendance of students at tutorials. Students must sign in and teaching assistants must upload the attendance data to the university's online portal. The uploaded data is then used by the exam administration to verify that the attendance requirement is met.[19]

Our attendance variable is the percentage of tutorials the student attends (per course). It was measured quite accurately because teaching assistants were tasked with preventing fraudulent sign-ins, as instructors required them to count the number of students present. The attendance statistics for voluntary students reinforces this point. On average these students attend tutorials 55 percent of the time. We can show that they also attend roughly 55 percent of their lectures. The match between tutorial and lecture attendance, together with the idea that students incur sunk costs of visiting campus, suggests that tutorial attendance is measured accurately.

Our data includes information from course evaluations. One week before the exam, students are invited by email to anonymously evaluate the course online at the university portal. They are reminded of the evaluations shortly after the exam. All evaluations contain the same set of 21 core questions, which are grouped into the general opinion of the course, structure, fairness, quality of lecturer and tutor, and usefulness of the lectures.

[19]While matching attendance with the administrative data (*e.g.* grades and demographics), we experienced a match rate of 93 percent (in our main sample). We compared the matched observations with the non-matched observations and find that: (i) grades do not differ between the two groups; (ii) the treatment effect on grades is not different between the two groups; (iii) scoring below a seven in the first year could not explain whether or not a record is matched (See Table A.2 in the Appendix). Therefore we work with this 93-percent sample throughout the paper.

Importantly, students are asked about their attendance at lectures, as well as the time they spend on their studies in total. Together with the data on tutorial attendance, we can infer how students adjust their time use between classes and studying on their own in response to forced attendance.[20] Note that the evaluations are filled out by 20 percent of the students. Later we will show that the response rate is the same just to the left and right of a first-year GPA of 7.

Our data on the personal characteristics of students includes information on their gender, age, distance from their residence to the university (in kilometers), and whether they are from the European Economic Area (EEA).[21] For Dutch students we also have information on their performance in high school. 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.

2.1. Basic Descriptives. Table 2 provides a basic summary of the data. The table compares students with an average first-year grade between 6.5 and 7 to students whose average grade was between 7 and 7.5. The unit of observation in the top panel is the student-course combination. The unit of observation in the bottom panel is the student. Second-year grades are measured in standard deviations.

The top panel shows forced students score 0.42 standard deviations worse than their peers. This is despite the fact that they attend tutorials 14 percentage points more of the time. The bottom panel implies students on one side of the cutoff are roughly similar to students on the other. The main difference being that poor performing students are likely to be over-represented to the left of 7 as visualized by their GPA in high school. Accordingly, we will account for this in our more flexible regression specifications by focusing on changes near 7.

2.2. Preview of Baseline Results. The leftmost column of Figure 1 examines the effect on attendance for the three types of courses. In courses where above-7 students

[20] For comprehensive details of the course evaluations see Table A.3 in the Appendix.

[21] Tuition fees are based on the student's EEA classification. Students who enroll in 2017-18, for example, pay €2,006 if they are from inside the EEA and €8,900 if not.

were given the option to attend, the difference in attendance between students above and below 7 was more than 30 percentage points. This translates into five extra tutorials for an eight credits course (three for a four credit course), or about 13 hours of extra schooling per block. In courses where the above-7 students were encouraged to attend, the difference was 12 percentage points. There was no attendance difference in courses that had their own penalty for being absent.

The middle column of Figure 1 examines the unconditional effect on grades. In courses where above-7 students had the option to attend, grades decreased by 0.35 of a standard deviation. In the other courses the effect on grades is a statistical zero. A grade comparison for attendance-voluntary and attendance-encouraged courses suggests that grades might only decrease if the additional time constraint is severe.

The figure at the bottom of the middle column shows a small grade decrease in absence-penalized courses. Although this grade difference across the cutoff is small and statistically insignificant, it appears to be puzzling at first sight as there is no attendance difference in these courses. Section 5 presents evidence against the decrease reflecting a direct effect of the university policy on grades (*e.g.* via self perception of the student) together with across-course differences in the value of tutorial attendance. Section 6 provides evidence that this difference reflects negative spillovers from adjacent courses where above-7 students have discretion over their attendance. This suggests that the university policy does not operate through attendance alone.

2.3. Abolition Results. The rightmost column of Figure 1 plots grade distributions in the abolition year, 2014-15. The figures show little to no difference in grades around the cutoff for the three types of courses. We observe that both the direct (top right) and spillover (bottom right) effects of the university policy have disappeared. Appendix Table A.4 shows formally that the differences are all statistically insignificant.[22]

Appendix Table A.5 compares mean grades above and below 7, before and after the

[22]We cannot plot attendance because the university stopped registering attendance in the abolition year.

abolition, in attendance-voluntary courses alone. The table shows that the unstandardized grades of students who expected to be forced in 2014-15, but ultimately were not, are 0.35 points (on a 10-point scale) higher than the grades of forced students from earlier cohorts. It also shows that this across-cohort difference is similar to the within-cohort difference of 0.37. In addition to providing further evidence that it is the forced attendance which decreases grades, Table A.5 implies that our estimates are being generated by lower performance of forced students, rather than by better performance of unforced students.

3 Empirical Specification

The second-year grade $g_{ijc}^{(2)}$ of student i in course j and cohort c is given by

$$g_{ijc}^{(2)} = \beta_0 + \beta_1 D_{ic} + f(\bar{g}_{ic}^{(1)} - 7) + f(\bar{g}_{ic}^{(1)} - 7)D_{ic} + C_{jc}^{(2)} + \mathbf{X}_i\mathbf{\Gamma} + \varepsilon_{ijc}^{(2)} \tag{1}$$

where D_{ic} equals 1 if first-year GPA is below 7, $\bar{g}_{ic}^{(1)}$ is their GPA in first year, $f(\cdot)$ is some polynomial expansion in $\bar{g}_{ic}^{(1)}$, $C_{jc}^{(2)}$ are course-cohort fixed effects, and $\mathbf{X}_i$ includes personal characteristics such as age. We allow the polynomial to differ from the left to the right of 7 (see the discussion in Lee and Lemieux [2010]), in part because it allows us to later analyze the external validity of our estimates [Cerulli et al., 2017]. Our primary interest is β_1, the effect of forced attendance near 7. The adoption and use of the forced attendance policy suggests $\beta_1 > 0$. The constraint it imposes on choices suggests $\beta_1 < 0$.

We can interpret estimates of β_1 causally if [Lee, 2008]:

> **Identifying Assumption:** Students have imprecise control over their average grade in the first year, meaning that conditional on their characteristics, the distribution for average grades is continuous around 7.

Because students were made aware of the policy early on and throughout their first year, they could try to take actions to avoid forced attendance in the second year. Our identification strategy will still work as long as first-year grades are at least 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,[23] have less of an effect on first-year GPA than on the grade of any one course.[24] The lack of control, together with the presence of aggregate shocks to the first-year performance of the individual student, should be enough for generating random assignment around 7.

To gain intuition for the identification argument, let

$$g^{(1)}_{ijc} = e^{(1)}_{ijc} + a_{ijc} + \delta^{(1)}_{jc} + \eta^{(1)}_{ijc}$$

denote the student's grade in first-year course j, a_{ijc} is their ability, $\delta^{(1)}_{jc}$ is something particular about the course-cohort combination (such as the professor or teaching assistant), and $\eta^{(1)}_{ijc}$ is the idiosyncratic component of the first-year grade. $e^{(1)}_{ijc}$ encapsulates any choice that affects grades, including the intensity of effort, study hours, tutorial and class attendance, or requests for grade increases. Second-year tutorial attendance is mandatory if:

$$\bar{e}^{(1)}_{ic} + \bar{a}_{ic} + \bar{\delta}^{(1)}_{c} + \bar{\eta}^{(1)}_{ic} < 7$$

where the bars indicate that the variable is averaged over all first-year courses j.

[23] Asking professors for grade increases, or any other such practice, can effect treatment assignment only when cumulative GPA is very close to 7.

[24] We are developing a companion article that studies adjustments to the threat of forced attendance. Our evidence shows that the threat does elicit a response but that, as expected, the response is almost never enough to get out of forced attendance. This claim is supported by various randomization and McCrary tests, as well as the null effects for abolition year. Nonetheless, because of this concern, we will use models that exclude potentially problematic neighbourhoods around 7 (donut-hole RD models) to demonstrate the robustness of our results.

The argument has three parts. The first is the student has limited control, $\bar{e}_{ic}^{(1)}$, over their average performance, as the effect of *e.g.* grade manipulation is smaller in the aggregate. The second is that there are aggregate shocks to first-year performance, $\bar{\eta}_{ic}^{(1)}$, such as bad luck across the exams they wrote that year. Shocks like these ensure that two students, with similar ability and average effort, end up on either side of the cutoff. As a result, the conditional distribution of first-year GPA is continuous and the variation in treatment status will be random in a neighborhood of 7. The third is that randomization near the cutoff takes place cohort by cohort. The student pool near 7 in one cohort may differ from the student pool near 7 in another. The presence of $\bar{\delta}_{c}^{(1)}$ suggests we should control for differences across cohorts.

3.1. Continuity Near the Cutoff. Local randomization of the treatment near the cutoff gives us two testable implications: (i) observed characteristics are identical from one side of the cutoff to the other; (ii) the probability density for GPA is continuous. We evaluate the implications one by one.

Table 3 presents estimates of our main empirical specification (Equation (1)) where instead of grades the dependent variables are student characteristics. The table presents results for local linear regressions (panel A) and a third order polynomial for $f(\cdot)$, with our main estimation sample (panel B) and the full sample (panel C).

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 in their performance in high school (level, track, and average grade).[25] This conclusion holds if we select the bandwidth optimally for each background characteristic (Appendix Table A.6). It also holds if we consider grade differences for various high school courses (Appendix Table A.7). Although much of the evidence supports the local randomization interpretation, in two of the three specifications of Table 3 the estimates indicate that

[25]A Dutch high school student might have followed two different levels before enrolling at university (easy=0, difficult=1). They might have followed one of 4 tracks within each level (1=least prestigious, 4=most prestigious). For the latter track variable, the results are unchanged if we account for the ordered nature of the variable.

women are underrepresented just to the right of the cutoff, consistent with the idea that women are manipulating grades less than men. The gender imbalance near 7, and residual concerns for grade manipulation more generally, further motivates estimation of donut-hole RD models.

We examine whether the probability density for GPA is continuous around 7 [McCrary, 2008]. If students can manipulate their GPA here, then we could observe bunching just above 7. To check we estimated Equation (1) using normalized counts of the number of students as the dependent variable.[26] Figure 2 summarizes the results, showing no evidence of bunching above the threshold. Table A.8 in the Appendix verifies this, formally showing that we are unable to reject the null of continuity near the cutoff.

3.2. Sample Attrition. The policy may have incentivized students to drop courses if and once they fail the 70 percent 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 Appendix Table A.9 (Columns 1-3) imply the policy has no effect on the number of completed courses. The intercepts support this conclusion, as they show students near the cutoff complete almost every course (nine out of ten).

Students near 7 may differ in their propensity to complete course evaluations and thus compromise the use of course evaluations in our analysis. Appendix Table A.9 (Columns 4-7) reports estimates of the policy effect on an indicator for whether students completed the course evaluation, for all eight courses, and separately for attendance-voluntary courses. 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.

3.3. Estimation and Inference. While discussing the results we present two specifications for Equation (1): a linear and third order polynomial for $f(\cdot)$ with bandwidths of

[26] To count the number of students we select bin sizes in accordance with the proposed strategy of McCrary [2008]. The results are robust to the bin size.

0.2 (6.8 to 7.2) and 0.5 (6.5 to 7.5) respectively. For all specifications we cluster standard errors at the level of the student.[27]

We elaborate on how we settled on our preferred specifications. Because local randomization implies that local comparisons provide an unbiased estimate of β_1, our starting point will almost always be specifications with a narrow bandwidth, where $f(\cdot)$ is taken to be linear. We follow Imbens and Lemieux [2008] and refer to these as locally linear regressions. For these specifications, we use a bandwidth of 0.2 because it is optimal according to the cross-validation method, and confirmed via the various bandwidth selectors provided by Calonico et al. [2016]. Our second set of specifications use a wider bandwidth of 0.5 together with the third-order polynomial for $f(\cdot)$. A wider bandwidth and flexible functional form lets us use more data and approximate what a locally-randomized experiment would have shown [Van der Klaauw, 2002]. To select $f(\cdot)$ for this larger estimation sample, we estimated Equation (1) while adding equal-sized bin dummies of GPA and including higher-order polynomials until the bin dummies were jointly insignificant.[28] We did this for multiple bandwidth choices. To select the 0.5 bandwidth, we again made use of the various bandwidth selectors of Calonico et al. [2016].[29]

One remaining concern relates to whether GPA has enough mass points to warrant a continuity-based RD design, which allows for the possibility that average potential outcomes vary with the running variable (GPA). To this end, note that there are 228 unique GPA values for the 717 students in our estimation sample of 6.5 to 7.5, amounting to approximately one GPA value for every 3 students. This amount of coverage of the support for GPA is usually sufficient for a continuity-based design.[30]

[27] We do not cluster on the tutorial group because peer composition differs from course to course. However, we show that our results are robust to including tutorial fixed-effects.

[28] We ran various regressions while changing the number of bins, but our preferred specification includes the number of bins (8) for which we first stopped rejecting the small (few dummies) versus the big model while choosing the binsize for the local averages for the RD graphs (see Figure 1).

[29] See Appendix Figure A.1 and Table A.10 for more details on the optimal bandwidth selection. Note that we use the equation between student grades and first-year GPA for selecting the bandwidth and polynomial order. This seems reasonable as the relationship between attendance and first-year GPA is relatively flat to left and right of 7. In the latter case we would expect the polynomial to be linear and the optimal bandwidth to be wide.

[30] Cattaneo, Idrobo, and Titiunik [2018] analyze an example where for every 110 observations one unique value for the running variable is observed. They conclude that continuity-based analysis might

4 Baseline Results

Table 4 reports estimates that are based on pooled data from the 8 affected courses. Basing estimates on the pooled data allows us to account for across-course error correlation within students. Estimates of the average effect are found in Columns (1) and (2). Panels A and B report the estimated effects for attendance and grades.

The university-wide policy increases the attendance of forced students by 15 percentage points ($p < 0.01$). It decreases their grades by 0.15 standard deviations. While we are unable to reject the null hypothesis of no average effect on grades, we are able to reject null hypotheses of positive effects of 0.05 and 0.1 standard deviations, with p-values of 0.09 and 0.03, respectively.

4.1. Course-Level Attendance Policies. Table 5 evaluates the policy effect on attendance for the three types of courses. Columns (1) and (2) of Table 5 report estimated effects on attendance in courses which gave above-7 students the option to attend, (3) and (4) report effects for courses where attendance was strongly encouraged, and (5) and (6) report effects for courses where everyone was penalized for being absent.

Forced students attended 29 to 34 percentage points more tutorials than above-7 students in attendance-voluntary courses ($p < 0.01$). They attended 11 to 15 percentage points more in attendance-encouraged courses ($p < 0.01$). They had the same attendance as above-7 students in absence-penalized courses.

Analogous estimates for grades are found in Table 6. Columns (1) and (2) show forced students have grades which are 0.34 to 0.43 standard deviations lower in attendance-voluntary courses ($p < 0.01$). Columns (3) and (4) shows little to no grade difference in attendance-encouraged courses. Columns (5) and (6) show the grades of forced students are 0.14 to 0.17 standard deviations lower in absence-penalized courses, though these differences are statistically insignificant at the 10 percent level. Note that columns (3) and (4) of Table 4 show the estimates in Tables 5 and 6 are similar to the estimates

be possible in this context.

we would obtain with pooled data and interactions between the treatment variable and course type.

Students can actually be better off with lower grades if their goal is to pass and forced attendance makes passing equally or more likely, perhaps because tutorials give students a better overview of the minimum they need to know. They may, in other words, achieve their desired result (passing) with less effort.

Columns (1) and (2) of Table 7 suggest this is not the case. Forced attendance decreases the probability of passing by 10 to 13 percentage points. The narrow-bandwidth estimates are statistically significant at the 5 percent level. The wide-bandwidth estimates have p-values which are a bit above 10 percent.[31] Columns (3) and (4) show there is effectively no difference in passing rates for attendance-encouraged courses. Columns (5) and (6) show passing rates which are 7 percentage points lower, with p-values which fluctuate around 10 percent.

4.2. Robustness. We analyzed the robustness of the result that forced attendance lowers grades in courses where attendance was voluntary for students to the right of 7. We estimated Equation (1) with the third order polynomial while varying the size of the bandwidth from 0.3 until 1.0. Appendix Figure A.2 shows the estimates hover around -0.4 and -0.3 and are significant across the whole range of optimal bandwidths using the various bandwidth selectors of Calonico et al. [2016].

We tested for significance at fake cutoffs. We estimated our main specification using the third order polynomial and a bandwidth of 0.5, while implementing fake cutoffs at every 0.005 points for GPA between 6.5 and 7.5, where the true cutoff is at 7. Appendix Figure A.3 presents a histogram and probability density of the β_1 estimates. The distribution mean is 0.02. The estimate at the true cutoff is extreme relative to the mean, having an empirical p-value that ranges between 3 and 6 percent (depending on whether normality is assumed).

We used donut-hole RD models to address concerns about the potential for manip-

[31] A probit analysis with a third-order polynomial yields similar but stronger (statistically) results.

ulation and gender imbalance near the cutoff. Appendix Figure A.4 shows the effect on grades is more negative as observations near 7 are removed. Note that this is consistent with forced attendance being relatively more costly to students who try to avoid it.[32]

We tested whether our results change if we restrict the linear polynomial $f(\cdot)$ to be the same on both sides of the cutoff. Appendix Table A.11 shows our results, where the estimates are virtually unchanged for all groups of courses. Finally, the robustness is supported by negligible effects in the abolition year. See the right panel of Figure 1 and Appendix Table A.4 for details.

4.3. External Validity. Our RD estimates apply to students with a GPA of 7. A valid question relates to the applicability of the estimates to students with a GPA other than 7. To speak to this question, Cerulli et al. [2017] recommend examining the coefficient on the interaction of the linear polynomial term and the treatment ($(\bar{g}_{ic}^{(1)} - 7)^1 D_{ic}$ in Equation (1)). A non-zero coefficient implies that students with a GPA just around, but not equal to, 7 can expect different treatment effects. It is effectively the Treatment Effect Derivative (TED) at 7.

We examine the TEDs for attendance and grades in Appendix Table A.12, for all 8 courses, as well as for attendance-voluntary courses, where the effect sizes are largest. We find that the TED estimates are all statistically insignificant. Similar implications follow from Figure 1, as it shows similar curves to the left and right of 7, especially in the case where attendance is the dependent variable. We also follow the suggestion of Cerulli et al. [2017] to consider the relative TED, *i.e.* the treatment effect divided by the TED and multiplied by the bandwidth. If the relative TED is less than one in absolute value, then the treatment effect changes sign somewhere in the estimation sample defined by the bandwidth. Appendix Table A.12 shows the relative TEDs have values which are above one for five out of the eight specifications. For the remaining three specifications we cannot reject the null hypothesis that the relative TED is equal to one. This suggests

[32]We consider donut holes with a maximum size of 6.95-7.05. To see why, suppose the GPA of the student is 6.95. To get to 7 they would need to receive a grade increase of more than 0.376 (0.752) for an eight (four) credits course. These sorts of increases are large and unlikely.

that our RD estimates apply to students whose GPA differs slightly from 7.

5 Baseline Mechanisms

We explore several potential mechanisms behind lower grades and passing rates in courses where attendance was voluntary for above-7 students.

5.1. Peer Effects. If the performance decline is driven by lower quality peers and TAs, then we would expect a more moderate or negligible decline if our estimates were based on comparisons of forced and voluntary students who attend the same tutorial. Appendix Table A.13 considers these comparisons, presenting treatment effect estimates for grades which are conditional on fixed effects for the tutorial group. The estimates are similar to our baseline estimates, suggesting that peer and TA quality are relatively unimportant for the effect of forced tutorial attendance on performance.

Appendix Table A.13 also evaluates peer effects more specifically, using the most common peer effects specifications in the literature [Booij, Leuven, and Oosterbeek, 2017], and focusing on whether the policy effect differs depending on the peer. The table reports effects of treatment interactions with the average 1^{st}-year grade for the peer group, as well as interaction effects for the average peer registration time for tutorials, measured in differences in days from the course mean registration time. The interaction effects account for the possibility that students coordinate tutorial times with their most preferred peers, which for forced students might very well include other low achievers. It also helps with the possibility that weak students coalesce simply because registration is left to the last minute.

The effects of treatment interactions with peer quality are modest. All the estimates are statistically insignificant at conventional significance levels, while the main treatment estimate is unchanged compared to our baseline specifications. Negligible peer effects are unsurprising given recent discussions and results in the literature [Sacerdote, 2014].[33]

[33] Feld and Zölitz [2017] is especially relevant. They estimate positive but small peer effects in tutorials for economics students at another Dutch university.

5.2. Attendance is Useful in Some Courses, but not Others? The effectiveness of tutorials provides another alternative explanation for why performance is worse in attendance-voluntary courses. To justify our thought process on this, we will draw on estimates from the attendance-encouraged and absence-penalized courses. Notice that the grades of forced students were about 0.15 standard deviations lower in courses where all students were penalized for missing tutorials. There was no grade difference in attendance-encouraged courses. The patterns may reflect the combined influence of a direct effect of the university policy (*e.g.* via self perception) and course-specific heterogeneity in the usefulness of attendance. Grades may be similar in attendance-encouraged courses because attendance is useful which cancels out the 0.15 reduction in grades. Students may have lower grades in attendance-voluntary courses because attendance is useless. Useless attendance can reinforce the 0.15 reduction from the direct effect, taking it down to the 0.35 reduction we observe in the data.

If our results are driven by differences in the usefulness of attendance, then TA and Lecturer quality should be highest in attendance-encouraged courses. Appendix Table A.14 uses data from the abolition year to investigate this possibility, reporting estimates of the relationship between perceived TA/Lecturer quality and fixed effects for the different types of courses, the baseline being the courses where students to right of 7 were encouraged. Data from the abolition year circumvents concerns about whether the course evaluations are contaminated by participation in forced attendance.

Appendix Table A.14 suggests, if anything, that TA quality is lowest in courses where attendance was encouraged (the base group). It also shows no statistical difference in lecturer quality across the three types of courses. The evidence suggests our results are not explained by a direct negative effect combined with course-specific heterogeneity in the usefulness of attendance.

5.3. It's About Time. If the policy has its largest effects on students who pay a high price for or derive little additional utility from attendance, then our results would be consistent with a model where students care about grades, where they think carefully about their attendance, and, importantly, where their time is being constrained by the policy. These students are really forced, being pushed further away from the choices they would make in the absence of the policy.

We estimate

$$A^{(2)}_{ijc} = \gamma_0 + \gamma_{1ic} D_{ic} + \varepsilon^{(2)}_{ijc} \tag{2}$$

where $A^{(2)}_{ijc}$ is the percentage of tutorials attended. If γ_{1ic} 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_{1ic} 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 and Pischke, 2008], students who otherwise prefer not to attend (large γ_{1ic}) are compliers. Students who would attend anyways (small γ_{1ic}) 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, of the courses from students with a first-year GPA below 7, we observe only 0.44 percent with an attendance rate below 70 percent.[34]

We operationalize γ_{1ic} via treatment interactions with proxies for the price of and additional utility from attendance. Our price proxy is distance to the university. Distant students pay a higher price for attendance because they have to spend more time travelling to campus. Our proxy for the additional utility of attendance is students' average tutorial attendance in first year. Students with a high propensity to attend in first year presumably derive additional utility from attendance in second year.[35]

Estimates are found in Table 8. From left to right the panel reports interaction effects for distance to the university, average attendance in the first year, and both together. Distance and first-year attendance are standardized, where the standard deviations are 30.9 kilometers for distance and 0.102 for attendance (on a scale from 0 to 1).

[34] One might argue that the grade for never takers are never observed, as they cannot write the exam. However, in Section 3.2 we showed 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).

[35] This proxy is implied by the assumption that preferences over tutorial attendance are stable from first to second year. Our results are consistent with the assumption.

Three patterns stand out. First, the direct effect of the characteristic is always opposite, but similar in magnitude, to the effect of its interaction with the treatment. This suggests the interactions pick up the student's counterfactual attendance had the policy not been in place. Second, Column (1) indicates the policy had a larger effect on students who live far from campus. The effect on attendance increases by 4.4 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, Column (2) shows the policy had a smaller effect on students who have a higher attendance propensity. The effect on attendance decreases by 13 percentage points for students that attended one standard deviation more tutorials in first year. The results are fairly stable when both interactions are included together (Column (3)).

5.4. Less Time for Leisure. Table 9 uses data from course evaluations to provide more direct evidence of the effect on time use. The left panel reports the effect on an indicator for whether the student attended lectures. The right panel reports the effect on total study time (lectures+tutorials+self study).[36] Note that the distance control accounts for direct influences of travel time on student responses.

Columns (1) to (2) show forced students are 28 to 45 percentage points more likely to attend lectures. The estimates, while marginally insignificant, are in line with the increases in tutorial attendance. The intercept and slope in the lecture attendance regressions are similar to the intercepts and slopes in the tutorial attendance regressions. This suggests 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 sunk cost interpretation is reinforced when analyzing the lecture attendance for courses that penalize all students for absence. The average student, forced or otherwise, will attend 90 percent of lectures and 90 percent of tutorials for this group of courses.

[36]Total study time is measured in 10 categories (1=0 hours, 2=1 to 5 hours, and 10=more than 40 hours). We used the maximum for the interval to convert the categories into hours, where the category 10 is assigned 45 hours. Only the intercepts change if we use the minimum or the mean.

Columns (3) to (4) shows results for total study time. We refrain from interpreting the exact magnitudes of the estimates, as our study-time measure is discrete with bins of 5 hours. The signs suggest, however, that forced attendance increases total study time. While the statistical significance is marginal, the estimates are consistent with reduced time for other courses and leisure. Later, when we investigate spillovers across courses, we will show estimates which are in fact consistent with reduced time for leisure. Lower grades and less leisure implies students are worse off under forced attendance.

5.5. Self-Study Time and Efficiency. Given our results, we feel that a reduction in self-study time or efficiency is the most reasonable explanation for the performance decline in attendance-voluntary courses. If the production function for grades is increasing in attendance, self-study hours, as well as self-study efficiency, and the policy increases attendance, then the only way for grades to decrease is if students spend less time on self study or become less efficient at it. If students were studying less or less efficiently, then our results would be consistent with a model where students care about their grades and leisure, where they think about attendance and self study carefully, and where the policy constrains self study indirectly via the additional constraint on attendance. Note, however, that our data does not allow us to show this explicitly.

Our argument fits well with the discussion of higher-education production in the careful time use study of Stinebrickner and Stinebrickner [2008]. They show that one additional hour of study (in the first semester) causes GPA to increase by 0.36 points. Our results are consistent with this mechanism, but also with a mechanism where there is a decline in study efficiency, and quite possibly with other mechanisms that fall outside the traditional scope of education production.

6 Spillovers Across Courses

The estimates for absence-penalized courses support the idea that the policy has effects outside of its direct influence on attendance. There is no attendance difference in these courses, yet there are differences in grades and passing rates. Why would this be the case? The results in Section 5 show peer effects and other direct impacts of the policy, such as adverse influences on self perception or identity that discourage students from doing their best, are unimportant for the effect on performance. In addition, Appendix Table A.15 uses our data on course evaluations to evaluate the effect of the policy on student perceptions of course attributes such as course structure and fairness. It shows no evidence of differential perceptions around the cutoff.

Another explanation is that the grade decrease in absence-penalized courses reflects negative spillovers from other courses. Being forced to spend extra time on one course could come at the cost of performance in another course situated in the same block.

Recall that the students in our setting all take the same 10 courses, in pairs of two, with one heavy (8 credit) and one light (4 credit) course in each block. While the three absence-penalized courses eliminate the disadvantage of forced students, they are adjacent to courses where forced students are disadvantaged. In Block 1 the heavy course is International Economics and the light course is Ageing and Fiscal Economics. The light course, Ageing and Fiscal Economics, penalizes absenteeism indiscriminately, whereas the heavy course, International Economics, encourages but does not force the attendance of above-7 students. The same scenario plays out in Block 2, but with Finance I as the heavy course, and Applied Statistics II as the light course. In Block 4 the scenario is slightly different. The heavy course is Methods and Techniques and penalizes absenteeism indiscriminately, where the light course is Behavioral Economics in which above-7 attendance is voluntary.[37] The extra time forced students spend on adjacent courses (International Economics, Finance I, and Behavioral Economics) should eat into the time they have for absence-penalized courses. This would provide an explanation for the small grade decrease observed in absence-penalized courses.

We look for further evidence of spillovers of this sort. To do this we investigate how self-reported total study hours differs depending on the course and block. If additional

[37]Block 3 and 5 both contain one course without tutorials. Grades for these courses cannot be credibly analyzed as they are largely determined via group work.

time on attendance-encouraged or attendance-voluntary courses crowds out self-study time for adjacent absence-penalized courses, and since absence-penalized courses have the same class attendance across the 7-threshold, then we should observe an increase in total study hours for forced students in courses where attendance is voluntary or encouraged and a decrease in total study hours in the adjacent absence-penalized courses.

Estimates are found in Table 10. The table organizes the estimates by block (1 to 5) and then by the course weight (light or heavy). Our hypothesis stipulates that (i) we should observe positive effects in Columns (3) and (4) for Blocks 1 and 2, as well as in Columns (1) and (2) for Block 4, and (ii) negative effects in Columns (1) and (2) for Blocks 1 and 2, as well as in Columns (3) and (4) for Block 4.

From Table 10 we find evidence supporting the first hypothesis. Columns (3) and (4) of Block 1 and 2 document positive effects on total study hours for forced students. While the estimates are borderline significant, the patterns support an increase in total study hours.[38] Columns (1) and (2) of Block 4 are somewhat inconclusive, the estimates are positive but imprecise. Support for the second hypothesis is moderate, as the increase in total study time does not seem to crowd out study hours for the adjacent courses. The estimates in Columns (1) and (2) for Blocks 1 and 2 and Columns (3) and (4) for Block 4 are statistical zeros. While the zeros do not provide strong evidence for a crowding-out of study hours in absence-penalized courses, the contrast with the positive estimates in the adjacent attendance-voluntary and attendance-encouraged courses is notable.

A couple of factors can explain the lack of support for the crowding-out of study hours. One is that our measure of total study hours is too imprecise from the perspective of identifying spillovers from other courses. Another more economically substantive factor is that crowding-out may operate along another margin such as study efficiency. Our view is that this mechanism is consistent with the fairly stable patterns observed in Columns (3) and (4) of Block 1, 2, 3, and 5 and Columns (1) and (2) of Block 4, namely that forced students spend more time on their courses where attendance was voluntary or encouraged

[38] To this end we can show that pooling the data for these courses yields statistically significant increases in total study time (for both bandwidths).

for above-7 students. This increase in total study time might make the remaining study time in the adjacent absence-penalized course less effective.

Our analysis here warrants further comment. First, to properly quantify spillover effects it would have been useful to have adjacent courses which are identical, with credit weights of 6 and 6 rather than 8 and 4, apart from their attendance policies for above-7 students. Second, whether the decrease in absence-penalized courses is generated by spillovers or not has no bearing on our capacity to answer our research question, *i.e.* to quantify the effect of a full year of forced attendance on academic performance. Third, the estimates in Table 10 imply that forced students enjoy less leisure under the forced attendance policy. The estimates reinforce our earlier claim that students are worse off under forced attendance.

7 Long-Run Performance

We investigate the effect of forced attendance in second year on performance in third year when, according to the university, tutorial attendance was once again under the purview of the student. Table 11 reports the effects of forced attendance on third-year grades. Columns (1) and (3) report estimates without controlling for course-cohort fixed effects, where Columns (2) and (4) include them as controls. Although we realize course-cohort fixed effects might potentially be bad controls, they are informative about why the performance of forced students is worse in third year.

Students who were forced in the past have lower grades on average. Column (1) shows a decline of 0.25 standard deviations ($p < 0.1$) for students near 7. Column (3) shows the decline is 0.18 standard deviations ($p > 0.1$) if the larger bandwidth of 0.5 is used. Column (3) rejects a positive effect of 0.1 with a p-value of 8 percent.

We find evidence that performance is worse even after students regain the right to decide their attendance. Why would this be the case? One explanation relates to course choice and the grades students expect to receive. In addition to retaining decision rights over attendance, students had the right to pick their courses in third year. At the same time, they carry their second-year grades with them. The historical grades in third-year courses and their own historical performance are information they can use to select courses. The courses they select can drive down their third-year performance. The estimates in Columns (2) and (4) are consistent with this, as they show course-cohort fixed effects eliminate roughly half of the negative effect.

8 Conclusion

We estimate the causal effects of a full year of forced, frequent, and regular attendance on the academic performance of the above-average student at a large public university. Our estimates imply that forced students, with a first-year GPA at or around [Cerulli et al., 2017] 7, can expect a GPA decrease of 0.20 standard deviations over the remainder of their undergraduate degree.[39] The aggregate effect consists of a decrease of 0.15 standard deviations in second year, when attendance is forced, and a decrease of 0.25 standard deviations in third year, when they regain discretion over their attendance. While the negative effects on second and third year performance are marginally significant or insignificant, we are able to rule out positive effects in the ranges of 0.05 and 0.1 standard deviations.

The effects on second-year performance are moderated by the attendance policies of individual courses. The largest effects are in courses where the attendance advantage of above-7 students was greatest, where they had full discretion over their attendance. Forced attendance decreases grades in these courses by 0.35 standard deviations and the chances of passing by more than 10 percentage points. The smallest, and statistically negligible, effects are in courses where the attendance of above-7 students was strongly encouraged, suggesting the effects may depend on the degree of the attendance disad-

[39]The decrease is a weighted average of the effect on all courses in the second year (Column (1) of Table 4) and third year (Column (1) of Table 11). This point estimate is statistically significant at the 5%-level.

vantage of forced students. We find intermediate effects in courses that eliminated the attendance advantage of above-7 students via absenteeism penalties that applied equally to all students. We argue that these intermediate effects reflect negative spillovers from adjacent courses where forced students are disadvantaged.

Our evidence suggests forced students enjoy less leisure in second-year. We also showed that grades are lower in third-year, when all students regain the right to decide their attendance. The decrease in grades in second and third year, together with the reduction in leisure, imply that the university policy makes forced students worse off. The moderating-effects of course-level attendance policies suggests we are underestimating their loss relative to a counterfactual policy that leaves above-7 students with full discretion over attendance in all their courses.

References

Angrist, J. D. and A. B. Krueger (1991). Does compulsory school attendance affect schooling and earnings? *The Quarterly Journal of Economics 106*(4), 979–1014.

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

Angrist, J. D. and J.-S. Pischke (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

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

Calonico, S., M. D. Cattaneo, M. H. Farrell, and R. Titiunik (2016). rdrobust: Software for regression discontinuity designs. *Unpublished manuscript available at: http://faculty. chicagobooth. edu/max. farrell/research/Calonico-Cattaneo-Farrell-Titiunik_2016_Stata. pdf*.

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

Cattaneo, M. D., N. Idrobo, and R. Titiunik (2018). A practical introduction to regression discontinuity designs: Part ii. *Preparation for Cambridge Elements: Quantitative and Computational Methods for Social Science, Cambridge University Press*.

Cerulli, G., Y. Dong, A. Lewbel, and A. Poulsen (2017). Testing stability of regression discontinuity models. In *Regression Discontinuity Designs: Theory and Applications*, pp. 317–339. Emerald Publishing Limited.

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

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

De Paola, M., V. Scoppa, and R. Nistico (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., A. Berhan, and Y. Berhan (2014). Absenteeism among medical and health science undergraduate students at hawassa university, ethiopia. *BMC medical education 14*(1), 81.

Durden, G. C. and L. V. Ellis (1995). The effects of attendance on student learning in principles of economics. *The 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. and U. Zölitz (2017). Understanding peer effects-on the nature, estimation and channels of peer effects. *Journal of Labor Economics 35*(2).

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

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

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

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

Kottasz, R. et al. (2005). Reasons for student non-attendance at lectures and tutorials: An analysis. *Investigations in university teaching and learning 2*(2), 5–16.

Latif, E. and S. Miles (2013). Class attendance and academic performance: a panel data analysis. *Economic Papers: A journal of applied economics and policy 32*(4), 470–476.

Lavecchia, A. M., H. Liu, and P. Oreopoulos (2014). Behavioral economics of education: Progress and possibilities. Technical report, 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. and T. Lemieux (2010). Regression discontinuity designs in economics. *Journal of economic literature 48*(2), 281–355.

Leuven, E., H. Oosterbeek, and B. van der Klaauw (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. and J. Chen (2006). Cumulative class attendance and exam performance. *Applied Economics Letters 13*(14), 937–942.

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

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

Martins, P. S. and I. Walker (2006). Student achievement and university classes: Effects of attendance, size, peers, and teachers. Technical report.

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. (2007). Do dropouts drop out too soon? wealth, health and happiness from compulsory schooling. *Journal of Public Economics 91*(11), 2213–2229.

Oreopoulos, P. and U. Petronijevic (2013). Making college worth it: A review of research on the returns to higher education. *NBER Working Paper No. 19053*.

Romer, D. (1993, September). 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? *Annu. Rev. Econ. 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*.

Snyder, J. L., J. E. Lee-Partridge, A. T. Jarmoszko, O. Petkova, and M. J. D'Onofrio (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. *The Journal of Economic Education 37*(3), 251–266.

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

Van der Klaauw, W. (2002). Estimating the effect of financial aid offers on college enrollment: A regression–discontinuity approach. *International Economic Review 43*(4), 1249–1287.

Table 1: **Attendance Policies of Second-Year Courses.**

Course	ECTS	Tutorials	Policy	Years	Tutorial Description	Exam Qs.	Block
International Economics	8	Yes	Encourage	2009/13	Students explicitly told to attend 10 of 13 tutorials. Discussion of exercises that are hand in before tutorial. No direct influence on final grade.	MC	1
Ageing or Fiscal Economics	4	Yes	Penalize	2010/13	Economics of Ageing: Exercises + Presentations, Accounts for (roughly) 30 percent of their final grade. Fiscal Economics: Exercises, Accounts for 25 percent of final grade. Absence implies a 0 out of respectively 30 and 25.	MC	1
Finance I	8	Yes	Encourage	2009/13	Exercises, Outside tutorials there are weekly quizzes that account for 20 percent of final grade.	MC	2
Applied Statistics II	4	Yes	Penalize	2009/13	Exercises, Accounts for 15 percent of final grade. Absence implies a 0 out of 15.	Open	2
Applied Microeconomics	8	Yes	Encourage	2009/13	Draws on tutorial exercises for two interim tests which account for 20 percent of the final grade.	MC	3
History of Economic Thought	4	No	None	2009/13	Group and individual research projects.		3
Methods & Techniques	8	Yes	Penalize	2009/13	Exercises in Computer Lab, Accounts for 5 percent of final grade. Absence implies a 0 out of 5.	MC	4
Behavioral Economics	4	Yes	Voluntary	2010/13	Exercises, Actual Experiments, No direct influence on final grade.	MC	4
Intermediate Accounting	8	Yes	Voluntary	2009/13	Exercises, No direct influence on final grade.	MC	5
Research Project	4	No	None	2009/13	Group research projects.		5

Notes:
1. The description is extracted from course guides.
2. The Policy column indicates whether the course layered their own attendance policy over top of the forced attendance policy of the university. Voluntary indicates that attendance was voluntary for above-7 students. Penalized indicates that above-7 students (and all students) were penalized for missing tutorials. Encouraged indicates that attendance was strongly encouraged for above-7 students.

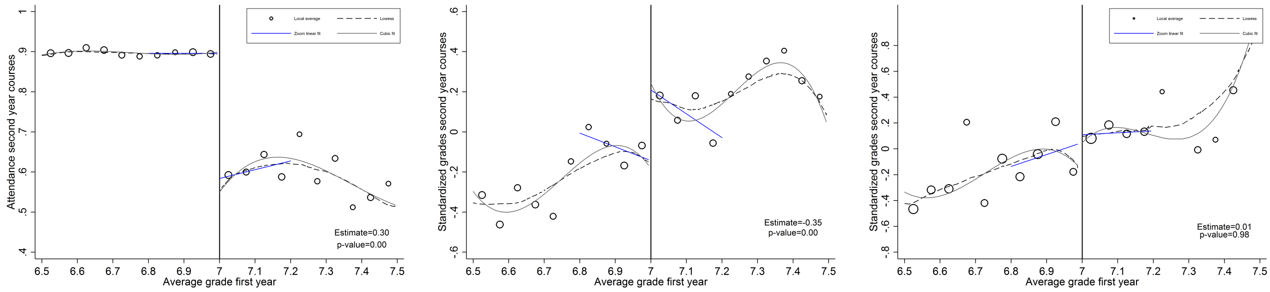
Table 2: **Basic Descriptives (All 8 Eligible Courses).**

	Grade Range		
	[6.5-7)		[7-7.5]
Course level (second year)			
Observations	2610		2291
Grade (*s.d.*)	-0.23	***	0.19
Attendance tutorials	0.90	***	0.76
Student level			
Observations	386		331
Distance to university (km)	24.13		22.04
Age	20.28		20.23
Gender (1=female)	0.30		0.31
European Economic Area (1=yes)	0.93		0.93
High-School Grade (*s.d.*)	-0.10	***	0.12

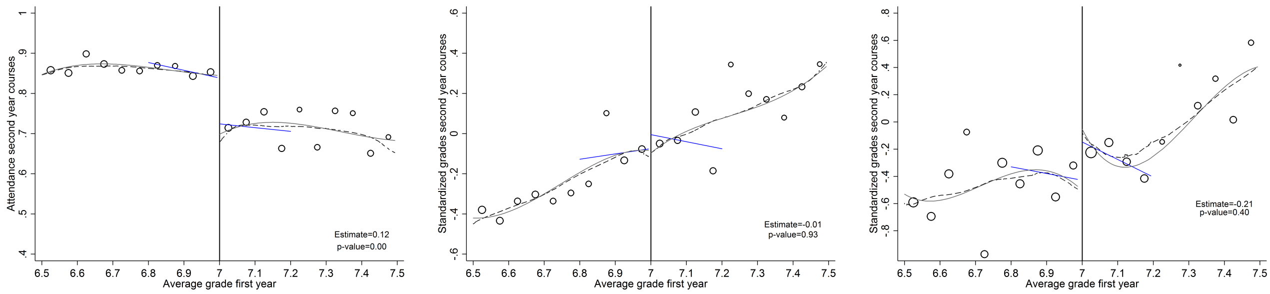
Notes:
1. Each high-school grade is a 50-50 weighted average of the grade the high school assigned and the grade the student received on a national exam for the course.
2. *s.d.* denotes measurement in standard deviations.
3. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
4. Stars denote the statistical significance for the difference in means, standard errors are clustered on the student level.

Figure 1: **Second Year Attendance and Grades, by Course Type.** The Left and Middle Panel show Attendance and Grades during the Policy (2009-14) respectively, Panel on the Right shows Grades after the Policy is Abolished (2014-15).

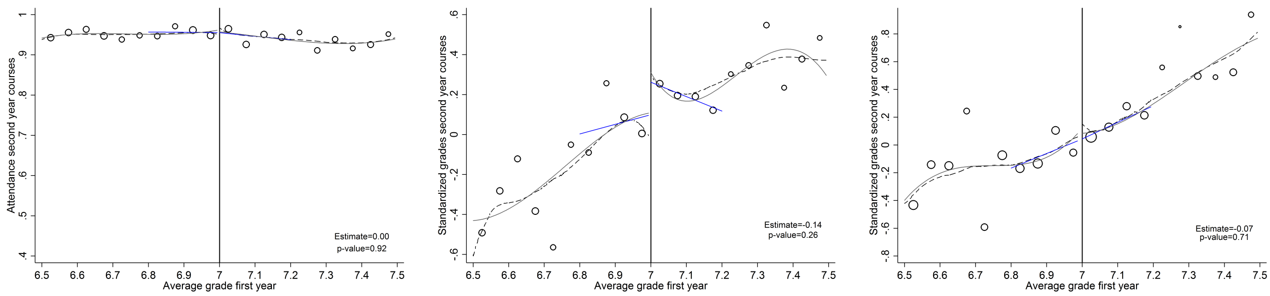
(a) Attendance is Forced to Left of 7, Voluntary to the Right



(b) Attendance is Forced to Left of 7, Strongly Encouraged to the Right



(c) Attendance is Forced to Left of 7, Absence is Penalized to the Right



Notes:

1. Locally linear, cubic and weighted scatterplots (lowess) for attendance or 2^{nd}-year grade against average 1^{st}-year grade.
2. Dots are based on local averages for a binsize of 0.05. Dot sizes reflect the number of observations used to calculate the average.
3. Linear and cubic fits are chosen according to our preferred specifications (see Section 3). Lowess makes no assumption on functional form (estimated with a bandwidth of $0.8N$).
4. Binsizes for local averages are selected via F-tests for regressions of 2^{nd}-year grades on K bin dummies and $2K$ bin dummies for the average 1^{st} year grade.

Table 3: **Balancing Tests around the Cutoff.**

	Distance to University	Age	Gender	European Economic Area	High School Level	High School Track	High School Grade
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			A: Local linear regression				
Average 1^{st}-year Grade is Below 7	1.796	0.257	0.257**	-0.032	0.021	0.241	-0.209
	(0.24)	(1.22)	(2.50)	(-0.53)	(1.14)	(1.14)	(-0.83)
Observations	310	310	310	310	248	248	248
Adjusted R^2	0.000	-0.000	0.001	-0.004	0.006	0.274	0.003
			B: Third order polynomial				
Average 1^{st}-year Grade is Below 7	-3.999	0.357	0.278**	0.024	0.004	0.170	-0.348
	(-0.42)	(1.27)	(2.13)	(0.31)	(0.21)	(0.63)	(-0.82)
Observations	717	717	717	717	592	592	592
Adjusted R^2	0.008	-0.005	-0.000	0.005	0.017	0.258	0.008
			C: Third order polynomial on whole sample				
Average 1^{st}-year Grade is Below 7	1.349	0.248	0.084	-0.005	0.042*	0.245	-0.313
	(0.24)	(1.37)	(1.06)	(-0.12)	(1.72)	(1.47)	(-1.29)
Observations	1420	1420	1420	1420	1176	1176	1176
Adjusted R^2	0.009	-0.001	0.002	0.008	0.011	0.350	0.065
Mean Outcome Var.	22.978	20.289	0.290	0.938	0.979	2.501	6.882

Notes:
1. The unit of observation is the student.
2. Regressions include cohort fixed effects.
3. Top panel displays local linear regressions with the optimal bandwidth of 0.2 around first-year grade of 7. Middle panel shows regressions for the optimal bandwidth of 0.5 with the third order polynomial. Bottom panel includes all observations. Polynomial is interacted with the treatment.
4. Column (5) until (7) only use the sample of Dutch students.
5. t-statistics in parentheses, standard errors are robust.
6. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 2: **No Bunching Just Above 7.** RD plot of the density for the number of students.

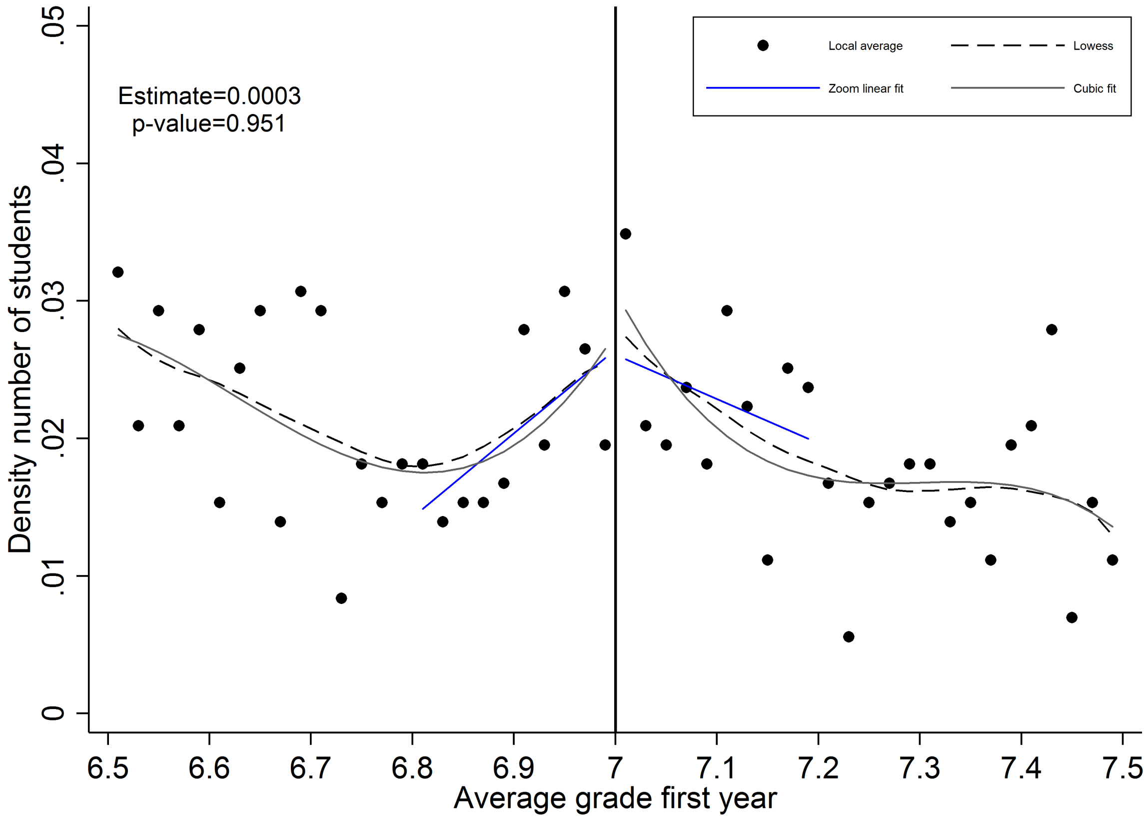


Table 4: **RD for All 8 Eligible Courses.**

	Average Effect		Marginal Effects by Course Type	
	(1)	(2)	(3)	(4)
	A: Attendance (% Tutorials Attended)			
Average 1^{st}-year Grade is Below 7	0.151***	0.147***	0.151***	0.146***
	(4.30)	(4.28)	(2.96)	(2.94)
Attendance is Voluntary × Treatment			0.193***	0.195***
			(3.91)	(3.96)
Absence is Penalized × Treatment			-0.151***	-0.149***
			(-2.99)	(-2.97)
Adjusted R^2	0.306	0.311	0.365	0.370
	B: Grade (Standardized)			
Average 1^{st}-year Grade is Below 7	-0.153	-0.154	0.0293	0.0262
	(-1.26)	(-1.28)	(0.18)	(0.16)
Attendance is Voluntary × Treatment			-0.451**	-0.447**
			(-2.36)	(-2.35)
Absence is Penalized × Treatment			-0.188	-0.185
			(-1.07)	(-1.06)
Observations	4901	4901	4901	4901
Adjusted R^2	0.210	0.210	0.210	0.210
Controls	No	Yes	No	Yes

Notes:
1. Regressions include course-cohort fixed effects.
2. Controls include distance to the university, age, gender, and European Economic Area.
3. All regressions use a third-order polynomial, as well as their interactions with the treatment, with a bandwidth of 0.5.
4. t- statistics in parentheses, standard errors are clustered on the student level.
5. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: **Forced Students Attend More Often.**

	Attendance (% Tutorials Attended)					
	Courses where Attendance is Forced to the Left of 7 and where to the Right					
	Attendance is Voluntary		Attendance is Encouraged		Absence is Penalized	
	(1)	(2)	(3)	(4)	(5)	(6)
	A: Local linear regression					
Average 1^{st}-year Grade is Below 7	0.296***	0.285***	0.122***	0.107***	0.002	0.000
	(6.25)	(6.21)	(3.01)	(2.78)	(0.10)	(0.01)
Observations	547	547	847	847	742	742
Adjusted R^2	0.366	0.376	0.153	0.174	0.154	0.180
	B: Third order polynomial					
Average 1^{st}-year Grade is Below 7	0.344***	0.335***	0.151***	0.145***	0.000	-0.000
	(5.79)	(5.72)	(2.97)	(2.92)	(0.01)	(-0.01)
Observations	1275	1275	1965	1965	1661	1661
Adjusted R^2	0.408	0.412	0.172	0.184	0.146	0.151
Controls	No	Yes	No	Yes	No	Yes

Notes:
1. Regressions include course-cohort fixed effects.
2. Controls include distance to the university, age, gender, and European Economic Area.
3. Top panel uses a bandwidth of 0.2 around a first-year grade of 7. Bottom panel uses a bandwidth of 0.5. Polynomial is interacted with the treatment.
4. t-statistics in parentheses, standard errors are clustered on the student level.
5. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: **Forced Students Perform Worse** in courses where attendance is voluntary for students scoring above 7 in first year.

	Grade (Standardized)					
	Courses where Attendance is Forced to the Left of 7 and where to the Right					
	Attendance is Voluntary		Attendance is Encouraged		Absence is Penalized	
	(1)	(2)	(3)	(4)	(5)	(6)
	A: Local linear regression					
Average 1^{st}-year Grade is Below 7	-0.349***	-0.342***	-0.011	0.010	-0.143	-0.164
	(-2.80)	(-2.77)	(-0.09)	(0.08)	(-1.12)	(-1.27)
Observations	547	547	847	847	742	742
Adjusted R^2	0.177	0.174	0.201	0.200	0.096	0.099
	B: Third order polynomial					
Average 1^{st}-year Grade is Below 7	-0.422***	-0.426***	0.029	0.041	-0.158	-0.169
	(-2.65)	(-2.74)	(0.18)	(0.26)	(-0.97)	(-1.02)
Observations	1275	1275	1965	1965	1661	1661
Adjusted R^2	0.216	0.216	0.251	0.250	0.156	0.158
Controls	No	Yes	No	Yes	No	Yes

Notes:
1. Regressions include course-cohort fixed effects.
2. Controls include distance to the university, age, gender, and European Economic Area.
3. Top panel uses a bandwidth of 0.2 around a first-year grade of 7. Bottom panel uses a bandwidth of 0.5. Polynomial is interacted with the treatment.
4. t-statistics in parentheses, standard errors are clustered on the student level.
5. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: **Forced Students are Less Likely to Pass.**

	Passes the Course					
	Courses where Attendance is Forced to the Left of 7 and where to the Right					
	Attendance is Voluntary		Attendance is Encouraged		Absence is Penalized	
	(1)	(2)	(3)	(4)	(5)	(6)
	A: Local linear regression					
Average 1^{st}-year Grade is Below 7	-0.125**	-0.117**	0.000	0.003	-0.066	-0.072*
	(-2.16)	(-2.02)	(0.01)	(0.05)	(-1.56)	(-1.74)
Observations	547	547	847	847	742	742
Adjusted R^2	0.074	0.070	0.093	0.088	0.050	0.051
	B: Third order polynomial					
Average 1^{st}-year Grade is Below 7	-0.106	-0.103	0.029	0.034	-0.070	-0.071
	(-1.47)	(-1.45)	(0.40)	(0.47)	(-1.29)	(-1.33)
Observations	1275	1275	1965	1965	1661	1661
Adjusted R^2	0.082	0.082	0.120	0.118	0.092	0.097
Controls	No	Yes	No	Yes	No	Yes

Notes:
1. Regressions include course-cohort fixed effects.
2. Controls include distance to the university, age, gender, and European Economic Area.
3. Top panel uses a bandwidth of 0.2 around a first-year grade of 7. Bottom panel uses a bandwidth of 0.5. Polynomial is interacted with the treatment.
4. t-statistics in parentheses, standard errors are clustered on the student level.
5. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: **Differential Effects on Attendance.**

	Attendance (% Tutorials Attended)		
	(1)	(2)	(3)
Average 1^{st}-year Grade is Below 7	0.337***	0.389***	0.390***
	(5.83)	(7.15)	(7.27)
Distance to University	-0.040**	-0.013	-0.036**
	(-2.43)	(-1.59)	(-2.26)
Distance×Treatment	0.044***		0.041**
	(2.61)		(2.51)
Attendance in First Year (Standardized)		0.152***	0.151***
		(8.39)	(8.43)
Attendance in First Year × Treatment		-0.133***	-0.130***
		(-7.24)	(-7.21)
Observations	1275	1275	1275
Adjusted R^2	0.417	0.485	0.490

Notes:
1. Courses where attendance was voluntary for students scoring above 7 in first year.
2. Regressions include course-cohort fixed effects, a polynomial in first-year grade, its interaction with the treatment, distance to university, age, gender, and European Economic Area.
3. Distance and attendance in first year are standardized, where the standard deviations are 30.9 kilometers for distance and 0.102 for attendance (on a scale from 0 to 1).
4. Bandwidth is 0.5.
5. t statistics in parentheses, standard errors are clustered on the student level.
6. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 9: **Less Time for Leisure or Non-Academic Activities?**

	Attended Lectures		Total Study Time	
	(1)	(2)	(3)	(4)
Average 1^{st}-year	0.282	0.455	5.170	8.391*
Grade is Below 7	(1.21)	(1.47)	(1.60)	(1.81)
Intercept	0.575***	0.435**	8.726***	5.720*
	(3.27)	(2.08)	(3.04)	(1.68)
Polynomial	1^{st}	3^{rd}	1^{st}	3^{rd}
Bandwidth	0.2	0.5	0.2	0.5
Observations	89	235	89	235
Adjusted R^2	-0.093	0.045	0.404	0.315

Notes:
1. Courses where attendance was voluntary for students scoring above 7 in first year.
2. The dependent variable in the left panel is the answer to the question "Have you attended lectures?". The dependent variable on the right is the answer to the question "Average study time (hours) for this course per week (lectures+tutorials+self study)?" where we used the maximum for the interval to convert the categories into hours.
3. Regressions include course-cohort fixed effects, a polynomial in first-year grade, its interaction with the treatment, distance to university, age, gender, and European Economic Area.
4. The intercepts are calculated via regressions which exclude course-cohort fixed effects and controls. They approximate the outcome mean near the threshold of students right of seven.
5. t-statistics in parentheses, standard errors are clustered on the student level.
6. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 10: **Total Study Time Across Courses.**

		Total Study Time			
		Light Course		Heavy Course	
		(1)	(2)	(3)	(4)
Block 1: Absence-Penalized Light Course, Attendance-Encouraged Heavy Course	Average 1^{st}-year Grade is Below 7	2.945	-1.831	5.874	5.916
		(0.54)	(-0.36)	(1.31)	(1.30)
	Observations	42	94	59	160
Block 2: Absence-Penalized Light Course, Attendance-Encouraged Heavy Course	Average 1^{st}-year Grade is Below 7	-1.368	1.021	8.023	12.04*
		(-0.24)	(0.20)	(1.50)	(1.96)
	Observations	50	130	48	119
Block 3: No Tutorials for Light Course, Attendance-Encouraged Heavy Course	Average 1^{st}-year Grade is Below 7	NA	NA	1.854	8.995
				(0.28)	(1.16)
	Observations			50	121
Block 4: Attendance-Voluntary Light Course, Absence-Penalized Heavy Course	Average 1^{st}-year Grade is Below 7	0.832	5.117	0.169	4.723
		(0.19)	(0.99)	(0.03)	(0.65)
	Observations	43	115	61	146
Block 5: No Tutorials for Light Course, Attendance-Voluntary Heavy Course	Average 1^{st}-year Grade is Below 7	NA	NA	10.17*	10.93
				(1.91)	(1.45)
	Observations			46	120
	Polynomial	1^{st}	3^{rd}	1^{st}	3^{rd}
	Bandwidth	0.2	0.5	0.2	0.5

Notes:
1. The dependent variable is the answer to the question "Average study time (hours) for this course per week (lectures+tutorials+self study)?" where we used the maximum for the interval to convert the categories into hours.
2. Attendance-Encouraged, Absence-Penalized, Attendance-Voluntary refer to how courses treated above-7 students. Below-7 students are forced in all these courses.
2. Regressions include course-cohort fixed effects, a polynomial in first-year grade, its interaction with the treatment, distance to university, age, gender, and European Economic Area.
3. Columns with an odd number use a bandwidth of 0.2 around a first-year grade of 7 and the even columns a bandwidth of 0.5. Polynomial is interacted with the treatment.
4. t-statistics in parentheses, standard errors are clustered on the student level.
5. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 11: **Performance Decline in Third Year.**

	Grade (Standardized)			
	Local linear regression		Third order polynomial	
	(1)	(2)	(3)	(4)
Average 1^{st}-year Grade is Below 7	-0.247*	-0.134	-0.176	-0.0750
	(-1.86)	(-1.05)	(-1.12)	(-0.52)
Observations	1869	1869	4236	4236
Adjusted R^2	0.003	0.223	0.022	0.254
Course-Cohort FE	No	Yes	No	Yes

Notes:
1. No student is required by the university to attend tutorials in third year.
2. Regressions include a polynomial in first-year grade, its interaction with the treatment, distance to university, age, gender, and European Economic Area.
3. Column (1) and (2) use a bandwidth of 0.2 around 7, whereas column (3) and (4) use a bandwidth of 0.5.
4. t-statistics in parentheses, standard errors are clustered on the student level.
5. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The Price of Forced Attendance

Online Appendix

Sacha Kapoor   Matthijs Oosterveen   Dinand Webbink

July 5, 2018

Table A.1: **Overview of Program.**

Group	First Year Courses	Second Year Courses
A	Microeconomics Macroeconomics Organisation and Strategy	Applied Microeconomics International Economics History of Economic Thought
B	Financial Information Systems Marketing Financial Accounting	Intermediate Accounting Behavioral Economics Finance I
C	Mathematics I Mathematics II Applied Statistics I ICT	Methods & Techniques Research Project Applied Statistics II Economics of Ageing (Eng) or Fiscal Economics (Dutch)

Notes:
1. The Economics of Ageing is taught in the English program. The Dutch program substitutes this for Fiscal Economics.
2. Students can compensate an insufficient grade (between a 4.5 and 5.4) with grades from other courses in the same group if: the other grades are sufficient (above 5.5) and the (weighted) average within the cluster is above 5.5. This applies to all students, whether they are above or below the threshold for the forced attendance policy.

Table A.2: **No Sample Selection when Matching Grades with Attendance.**

	Grade (standardized)				Matched	
	(1)	(2)	(3)	(4)	(5)	(6)
Matched	-0.0511	0.0219	0.491	0.628**		
	(-0.88)	(0.56)	(1.17)	(2.29)		
Average 1^{st}-year			-0.132	-0.133	0.00161	0.00112
Grade is Below 7			(-0.98)	(-0.98)	(1.05)	(0.47)
Their Interaction			-0.0276	-0.0346		
(Matched×Treatment)			(-0.26)	(-0.49)		
Polynomial	-	-	1^{st}	3^{rd}	1^{st}	3^{rd}
Bandwidth	0.2	0.5	0.2	0.5	0.2	0.5
Observations	2298	5297	2298	5297	2298	5297
Adjusted R^2	-0.000	-0.000	0.168	0.211	0.994	0.984

Notes:
1. Matched is a variable which equals 1 if the grade record found a match with the attendance data and 0 otherwise.
2. Columns (1) and (2) regress second year grades on a constant and the matched-variable and shows that grades are similar for matched and nonmatched records.
3. Columns (3) and (4) show the reduced-form effect is not different between matched and nonmatched records (Matched×Treatment). The final two columns regress the matched-variable upon scoring below 7 in the first year and thereby show the policy is unable to explain whether or not a record is matched.
4. Columns (3) until (6) include course-cohort fixed effects.
5. t-statistics in parentheses, standard errors are clustered on the student level.
6. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.3: **Overview of Categories and Questions in Course Evaluations.**

Question	Measurement scale	Category
Objectives of course are clear	1-5	General
Course is relevant for my studies	1-5	General
Course is interesting	1-5	General
Course is well organized	1-5	Structure
Course material is understandable	1-5	Structure
Can be completed within allocated study points	1-5	Fairness
Time needed to complete exam is enough	1-5	Fairness
Exam reflects course content	1-5	Fairness
Exam questions are clearly defined	1-5	Fairness
Total study time (lectures+tutorials+self study)	1-10	Total study time
Have you attended lectures?	0-1	Lecture attendance
Lectures are useful	1-5	Lectures useful
Lecturer is competent	1-5	Quality lecturer(s)
Lecturer makes you enthusiastic	1-5	Quality lecturer(s)
Lecturer has good command of English	1-5	Quality lecturer(s)
Lecturer can be easily contacted	1-5	Quality lecturer(s)
Lecturer provides sufficient assistance	1-5	Quality lecturer(s)
TA gives good tutorials	1-5	Quality TA
TA can be easily contacted	1-5	Quality TA
TA provides sufficient assistance	1-5	Quality TA
TA has good command of English	1-5	Quality TA

Notes:
1. Questions are measured on a Likert scale, where 1 equals strongly disagree and 5 equals strongly agree, with the two exceptions being total study time (1 being 0 hours, 2 being $[1-5]$ hours, 3 being $[6-10]$ hours and 10 being ≥ 40 hours) and lecture attendance (1 being yes and 0 being no).
2. We take the mean for questions within a category, ignoring potential missing values within a category. The more sophisticated approach of calculating the principal components gives qualitatively similar results.

Table A.4: **Negligible Effects when Forced Attendance is Abolished.**

	Grade (Standardized)					
	Courses where Attendance was Previously Forced to the Left of 7 and where to the Right					
	Attendance was Voluntary		Attendance was Encouraged		Absence was Penalized	
	(1)	(2)	(3)	(4)	(5)	(6)
	A: Local linear regression					
Average 1^{st}-year Grade is Below 7	0.00815	-0.0355	-0.210	-0.299	-0.0746	-0.216
	(0.02)	(-0.10)	(-0.85)	(-1.28)	(-0.37)	(-1.12)
Observations	190	190	292	292	292	292
Adjusted R^2	0.177	0.167	0.025	0.060	0.208	0.242
	B: Third order polynomial					
Average 1^{st}-year 7 Grade is Below 7	-0.121	-0.141	-0.403	-0.428	-0.0665	-0.161
	(-0.28)	(-0.31)	(-1.30)	(-1.43)	(-0.27)	(-0.64)
Observations	384	384	585	585	575	575
Adjusted R^2	0.236	0.240	0.089	0.106	0.269	0.279
Controls	No	Yes	No	Yes	No	Yes

Notes:
1. Regressions include course-cohort fixed effects.
2. Controls include distance to the university, age, gender, and European Economic Area.
3. Top panel uses a bandwidth of 0.2 around a first-year grade of 7. Bottom panel uses a bandwidth of 0.5. Polynomial is interacted with the treatment.
4. t-statistics in parentheses, standard errors are clustered on the student level.
5. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.5: **It is the Forcing that Worsens Performance.**

Cohort	GPA$\in [6.9 - 7.0)$		GPA $\in [7.0 - 7.1]$
2009 - 2013	6.40	$p = 0.004^{***}$	6.77
	$p = 0.126$		$p = 0.487$
2014	6.75	$p = 0.651$	6.88

Notes:
1. Local averages of raw grades for a bandwidth of 0.1.
2. Courses where attendance was normally voluntary for students scoring above 7 in first year.
3. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.6: **Additional Balancing Tests.** Optimal Bandwidth Selected per Background Characteristic for both the Local Linear Regression and the Third Order Polynomial using Calonico et al. [2016].

	Distance to University	Age	Gender	European Economic Area	High School Level	High School Track	High School Grade
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			A: Local linear regression				
Average 1^{st}-year Grade is Below 7	2.706	0.274	0.168*	-0.0328	0.0317*	0.142	-0.236
	(0.52)	(1.50)	(1.86)	(-0.77)	(1.75)	(0.99)	(-1.24)
Observations	585	643	386	574	523	537	493
Adjusted R^2	0.002	0.003	-0.007	0.005	0.025	0.279	0.011
			B: Third order polynomial				
Average 1^{st}-year Grade is Below 7	-6.912	0.214	0.267**	0.0378	0.00521	0.246	-0.366
	(-0.64)	(0.85)	(2.20)	(0.46)	(0.25)	(1.02)	(-0.84)
Observations	599	871	817	607	603	724	558
Adjusted R^2	0.010	-0.001	0.002	0.003	0.016	0.301	0.005
Mean Outcome Var.	22.978	20.289	0.290	0.938	0.979	2.501	6.882

Notes:
1. The unit of observation is the student.
2. Regressions include cohort fixed effects.
3. Top panel displays local linear regressions with an optimal bandwidth that is calculated for every background characteristic separately using MSERD of Calonico et al. [2016], where the bottom panel displays this for the third order polynomial. Polynomial is interacted with the treatment.
4. Column (5) until (7) only use the sample of Dutch students.
5. t-statistics in parentheses, standard errors are robust.
6. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.7: **Additional Balancing Tests.** Using Secondary School Grades as Dependent Variable.

	Grade (Standardized)							
	Dutch	English	Economics	General science	Civic education	History	German	Man. and org.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	A: Local linear regression							
Average 1^{st}-year Grade is Below 7	-0.247	-0.179	0.0918	-0.0128	0.208	0.0788	0.423	0.0582
	(-0.96)	(-0.67)	(0.43)	(-0.05)	(0.79)	(0.26)	(1.49)	(0.20)
Observations	241	234	228	232	229	187	165	154
Adjusted R^2	-0.001	0.004	-0.003	-0.018	0.005	-0.030	0.020	-0.001
	B: Third order polynomial							
Average 1^{st}-year Grade is Below 7	-0.324	-0.190	0.166	0.0385	0.424	0.0394	0.574	0.183
	(-0.82)	(-0.50)	(0.53)	(0.09)	(0.97)	(0.09)	(1.39)	(0.43)
Observations	578	560	554	552	543	457	393	381
Adjusted R^2	-0.004	0.004	0.013	-0.007	0.011	-0.006	0.035	0.007
	C: Third order polynomial on whole sample							
Average 1^{st}-year 7 Grade is Below 7	-0.260	-0.131	-0.0288	0.0312	0.0244	0.125	0.191	-0.0317
	(-1.21)	(-0.63)	(-0.15)	(0.13)	(0.10)	(0.63)	(0.87)	(-0.12)
Observations	1143	1112	1094	1086	1074	916	809	751
Adjusted R^2	0.037	0.035	0.095	0.036	0.046	0.050	0.083	0.077

Notes:
1. Regressions use secondary school grades as the dependent variable.
2. Data is at the student level. Regressions include cohort fixed effects. Column (7) controls for track choice via a binary indicator.
3. Panel A displays local linear regressions with the optimal bandwidth of 0.2 around 7. Panel B shows regressions for the optimal bandwidth of 0.5 with the third order polynomial. Panel C includes all observations. Polynomial is interacted with the treatment.
4. Observations differ per column, as not all students have followed the same courses.
5. t-statistics in parentheses, standard errors are robust.
6. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.8: **No Bunching Just Above 7.** Tested through the method proposed by McCrary [2008].

	Counts of Number of Students		
	Local linear regression	Second order polynomial	Third order polynomial
	(1)	(2)	(3)
	A: Binsize as suggested by McCrary [2008]		
Average 1^{st}-year Grade is Below 7	0.000363	-0.00203	-0.00294
	(0.06)	(-0.39)	(-0.41)
Observations	20	50	50
Adjusted R^2	0.127	0.211	0.203
	B: Bins two times as small		
Average 1^{st}-year Grade is Below 7	-0.0000178	-0.00119	-0.00205
	(-0.01)	(-0.39)	(-0.50)
Observations	40	100	100
Adjusted R^2	0.000	0.088	0.078
	C: Bins four times as small		
Average 1^{st}-year Grade is Below 7	-0.0000519	-0.000632	-0.00108
	(-0.03)	(-0.39)	(-0.51)
Observations	80	200	200
Adjusted R^2	-0.009	0.032	0.026

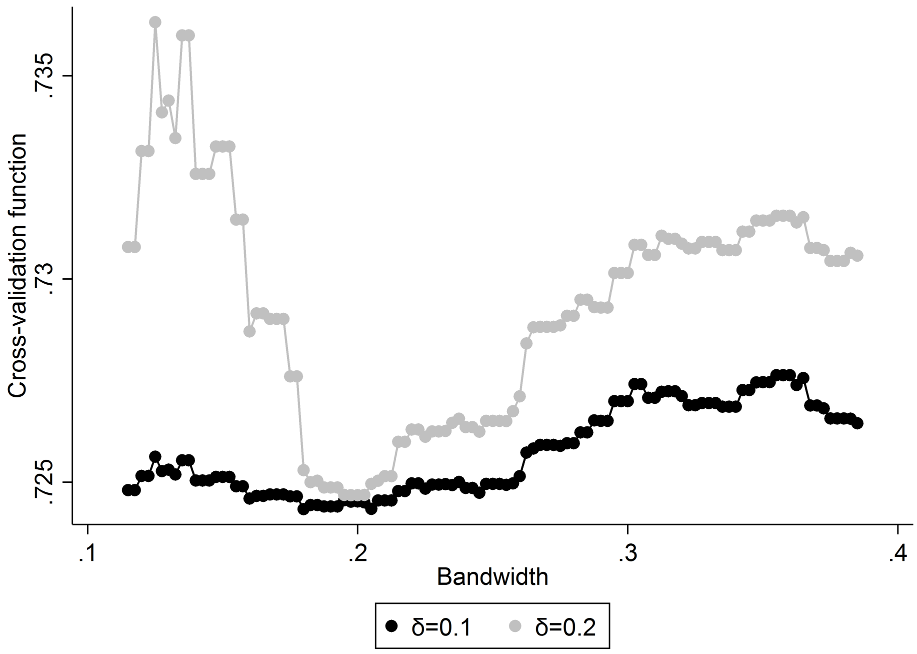
Notes:
1. The local linear regression is estimated on the optimal bandwidth of 0.2 around a first-year grade of 7, whereas the second- and third order polynomial is estimated on the optimal bandwidth of 0.5. Polynomial is interacted with the treatment.
2. The panels refer to the different binsize as to compute the histogram for the number of students. Panel A uses the plug-in estimate of McCrary [2008], panel B and C subsequently undersmooth and compute bins two and four times as small respectively. Results are robust to the binsize.
3. t-statistics in parentheses, standard errors are robust.
4. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.9: **Sample Selection.**

	Number of Courses			Completed Course Evaluation			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Average 1^{st}-year Grade is Below 7	0.126	-0.013	0.039	-0.091	-0.059	-0.079	-0.073
	(0.48)	(-0.08)	(0.11)	(-1.51)	(-0.80)	(-1.19)	(-0.93)
Intercept	9.168***	9.199***	9.165***	0.204***	0.185***	0.181***	0.169***
	(54.03)	(85.74)	(40.76)	(4.80)	(3.88)	(4.17)	(3.39)
Polynomial	1^{st}	1^{st}	3^{rd}	1^{st}	3^{rd}	1^{st}	3^{rd}
Bandwidth	0.2	0.5	0.5	0.2	0.5	0.2	0.5
Observations	310	717	717	2136	4901	547	1275
Adjusted R^2	0.040	0.035	0.031	0.055	0.072	-0.010	0.018

Notes:
1. Columns (1) until (3) include cohort fixed effects, whereas column (4) until (7) include course-cohort fixed effects. No further controls are included. Polynomial is interacted with the treatment.
2. The intercepts are calculated via regressions which exclude course-cohort fixed effects. They approximate the outcome mean near the threshold of students right of seven.
3. t-statistics in parentheses, standard errors are robust (columns (1) until (3)) or clustered on the student level (columns (4) until (7)).
4. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure A.1: **Selection of Optimal Bandwidth for the Local Linear Regression.**



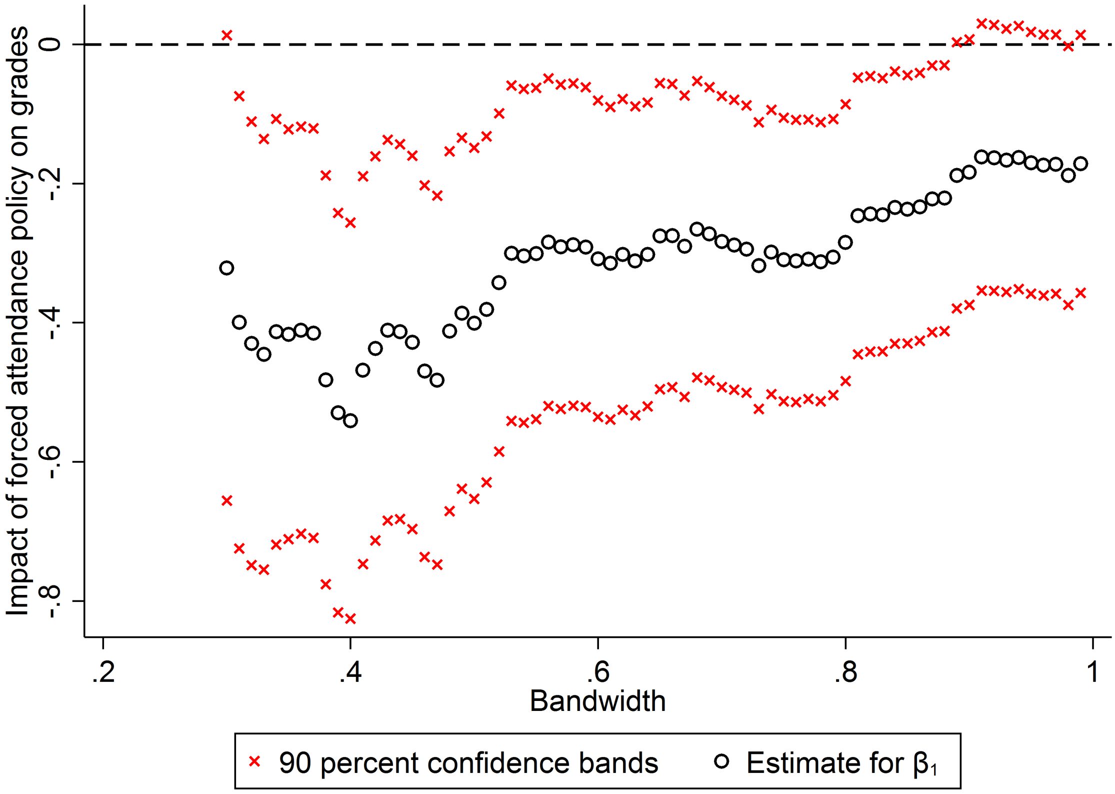
Notes: We follow Imbens and Lemieux [2008] to obtain predicted grades on either side of the cutoff and use the predictions to define a cross-validation criterion for selecting the bandwidth. δ denotes the distance from the grade of the student to cutoff and appears in the criterion function. δ equal to 0.1 and 0.2 roughly correspond to 10 and 20 percent of the observations at both sides of the cutoff. For both values the criterion is minimized at a bandwidth of 0.2.

Table A.10: **Calculations of the Optimal Bandwidth Using Methods of Calonico et al. [2016].**

		First order polynomial		Third order polynomial	
		Left of 7	Right of 7	Left of 7	Right of 7
	MSE rd	0.220	0.220	0.413	0.413
Mean	MSE two	0.226	0.226	0.437	0.639
squared	MSE sum	0.327	0.327	0.491	0.491
error	MSE comb1	0.220	0.220	0.423	0.413
	MSE comb2	0.226	0.327	0.437	0.491
	CER rd	0.139	0.139	0.248	0.248
Coverage	CER two	0.168	0.263	0.263	0.384
error	CER sum	0.207	0.207	0.295	0.295
rate	CER comb1	0.139	0.139	0.248	0.248
	CER comb2	0.168	0.207	0.263	0.295

Notes:
1. Optimal bandwidth sizes for both the local linear regressions and the third order polynomial.
2. For the local linear regression the result corresponds with the cross-validation method depicted in Figure A.1, the desired bandwidth hovers around 0.2 for both MSE- and CER methods.
3. For the third order polynomial the optimal bandwidth is between 0.4 and 0.6 for the MSE methods, while being significantly smaller for the CER methods. As such, for the third order polynomial we start out with a bandwidth of 0.5, but check for robustness.

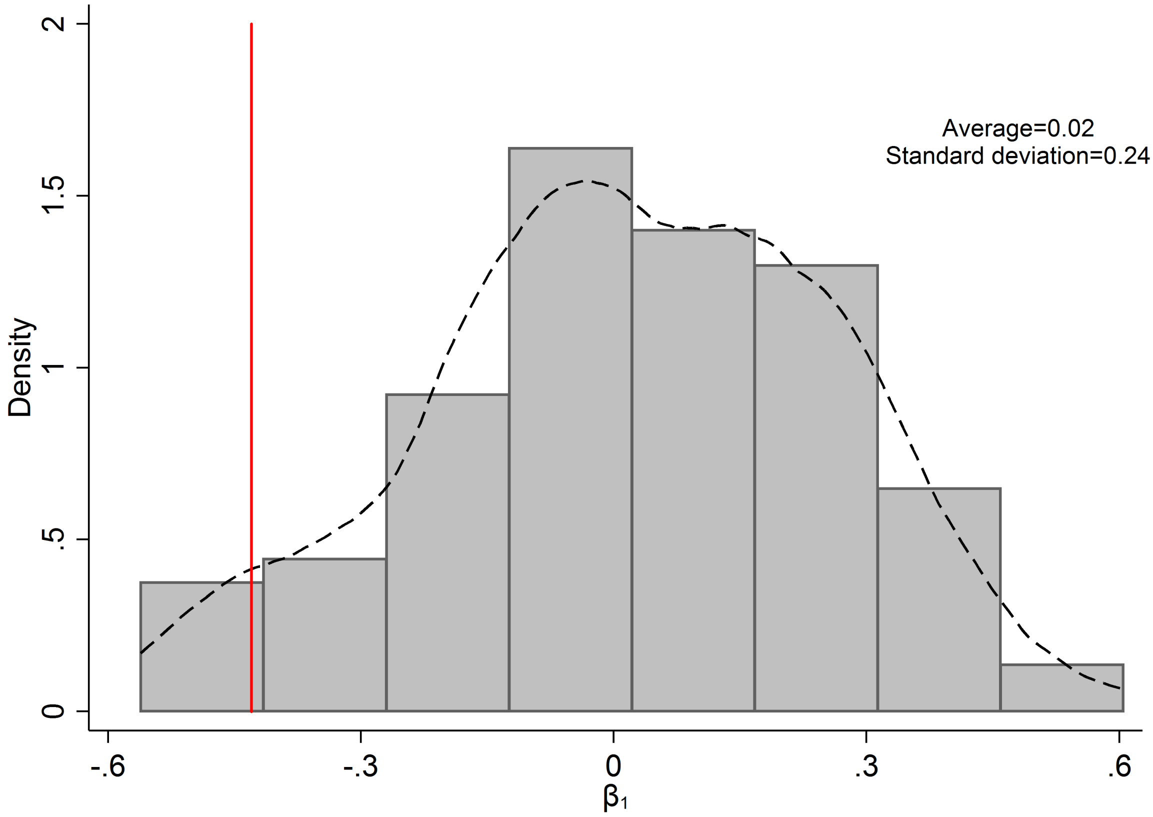
Figure A.2: **Estimate Insensitive to Bandwidth Choice.** Courses where attendance was voluntary above 7.



Notes:

1. The figure plots the estimates of the policy effect on grades for different bandwidths, against the bandwidth used to estimate the treatment effect.
2. The estimates are saddled by their confidence intervals.
3. The bandwidth ranges from 0.3 until 1.0.
4. Estimates based on specifications that control for a third order polynomial in the first year grade, its interactions with a treatment dummy at the cutoff, fixed effects for the course-cohort combination, distance to university, age, gender, and European Economic Area.

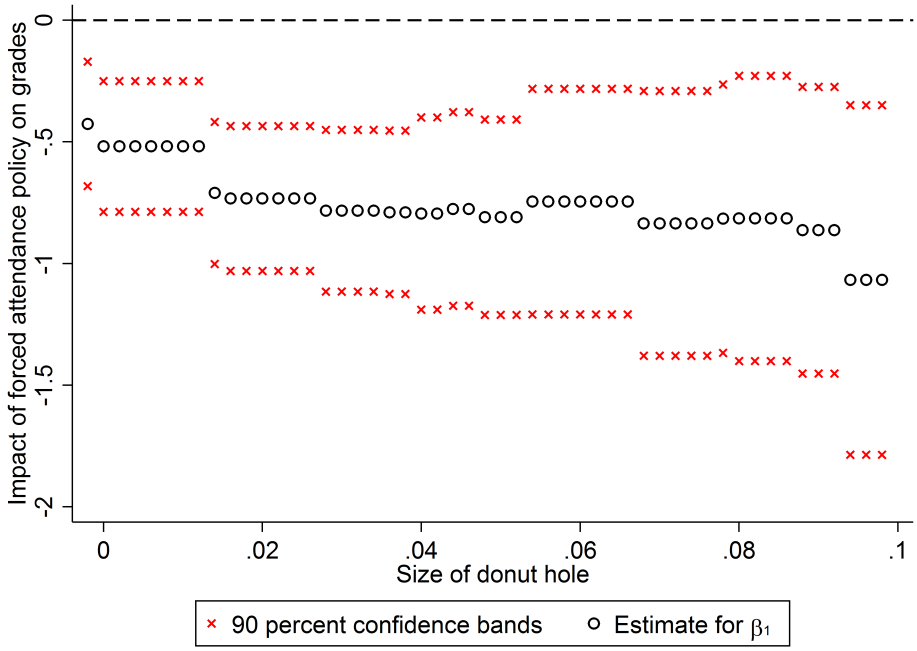
Figure A.3: **Fake Cutoffs.**



Notes:

1. Histogram for the estimates of the policy on grades at cutoffs that are arbitrarily assigned by us to every 0.005-points for GPA between 6.5 and 7.5.
2. Estimates use the sample of courses where attendance was voluntary for students scoring above 7 in first year.
3. Estimates based on specifications that control for a third order polynomial in the first year grade, its interactions with a treatment dummy at the fake cutoff, fixed effects for the course-cohort combination, distance to university, age, gender, and European Economic Area.
4. Vertical red line identifies the estimate at the true cutoff of 7.
5. Bandwidth for estimation is 0.5.

Figure A.4: **Robustness of Estimate Against a Donut Hole RD.**



Notes:

1. The figure plots the estimates of the policy effect on grades for different ranges of removed observations near the cutoff (the donut hole), against the size of the donut hole.
2. The estimates are saddled by their confidence intervals.
3. The donut hole ranges from 0 unto 0.1.
4. Estimates based on specifications that control for a third order polynomial in the first year grade, its interactions with a treatment dummy at the cutoff, fixed effects for the course-cohort combination, distance to university, age, gender, and European Economic Area.

Table A.11: **Results of Local Linear Regressions.** Restricting the polynomial to be similar on both sides of cutoff.

	All Courses		Courses where Attendance is Forced to the Left of 7 and where to the Right					
			Attendance is Voluntary		Attendance is Encouraged		Absence is Penalized	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
			A: Forced Attendance Increases Attendance					
Average 1^{st}-year Grade is Below 7	0.128***	0.120***	0.301***	0.291***	0.123***	0.108***	0.000893	-0.000203
	(4.31)	(4.26)	(5.87)	(5.86)	(2.86)	(2.66)	(0.06)	(-0.01)
Adjusted R^2	0.305	0.316	0.366	0.375	0.154	0.175	0.155	0.181
			B: Forced Attendance Decreases Grades					
Average 1^{st}-year Grade is Below 7	-0.154*	-0.150	-0.357***	-0.350***	-0.0196	0.000265	-0.157	-0.178
	(-1.66)	(-1.63)	(-2.94)	(-2.92)	(-0.16)	(0.00)	(-1.28)	(-1.44)
Observations	2136	2136	547	547	847	847	742	742
Adjusted R^2	0.165	0.166	0.178	0.176	0.202	0.200	0.096	0.099
Controls	No	Yes	No	Yes	No	Yes	No	Yes

Notes:
1. Regressions include course-cohort fixed effects.
2. Controls include distance to the university, age, gender, and European Economic Area.
3. Estimated by local linear regression with a bandwidth of 0.2 around first-year grade of 7, polynomial is restricted to be similar on both sides of the cutoff.
4. t- statistics in parentheses, standard errors are clustered on the student level.
5. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.12: **Testing the External Validity of the RD Estimate.** Using the Method of Cerulli et al. [2017].

	All Courses		Courses Where Attendance to the Right is Voluntary	
	(1)	(2)	(3)	(4)
	A: Attendance (% Tutorials Attended)			
Average 1^{st}-year Grade is Below 7	0.118***	0.147***	0.285***	0.335***
	(4.48)	(4.28)	(6.21)	(5.72)
Treatment Effect Derivative (TED)	-0.127	-0.255	-0.427	-1.190
	(-0.56)	(-0.40)	(-1.09)	(-1.09)
|Relative TED|	4.67	1.16	3.33	0.56
P-value of: |Relative TED|=1				0.338
Observations	2136	4901	547	1275
Adjusted R^2	0.316	0.311	0.376	0.412
	B: Grade (Standardized)			
Average 1^{st}-year Grade is Below 7	-0.139	-0.154	-0.342***	-0.426***
	(-1.49)	(-1.28)	(-2.77)	(-2.74)
Treatment Effect Derivative (TED)	0.825	1.112	0.523	0.719
	(0.93)	(0.49)	(0.47)	(0.24)
|Relative TED|	0.84	0.28	3.27	1.19
P-value of: |Relative TED|=1	0.892	0.272		
Observations	2136	4901	547	1275
Adjusted R^2	0.167	0.210	0.174	0.216
Polynomial	1^{st}	3^{rd}	1^{st}	3^{rd}
Bandwidth	0.2	0.5	0.2	0.5

Notes:
1. Regressions include course-cohort fixed effects, a polynomial in first-year grade, its interaction with the treatment, distance to the university, age, gender, and European Economic Area.
2. Columns with an odd number use a bandwidth of 0.2 around a first-year grade of 7 and the even columns a bandwidth of 0.5. Polynomial is interacted with the treatment.
3. The TED is defined as the linear term on the running variable that is interacted with the treatment variable. It measures whether the treatment effect changes while moving away from the cutoff.
4. The relative TED divides the treatment effect by the absolute TED, while multiplying the TED with the size of the bandwidth. If the absolute value is smaller than 1, it means that the treatment effect changes sign somewhere in the estimation sample considered by the bandwidth.
5. t- statistics in parentheses, standard errors are clustered on the student level.
6. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.13: **Negligible Effects of Low-Achieving Peers.**

	Grade (Standardized)		
	(1)	(2)	(3)
Average 1^{st}-year Grade is Below 7	-0.401** (-2.46)	-0.423** (-2.30)	-0.430*** (-2.75)
Average 1^{st}-year Grade Among Peers		0.087 (0.81)	
Their Interaction (Treatment×Peers)		0.008 (0.05)	
Average Registration Time Among Peers			0.002 (0.30)
Its Interaction with Treatment			0.001 (0.13)
Observations	1275	1275	1275
Adjusted R^2	0.209	0.215	0.215

Notes:
1. Courses where attendance was voluntary for students scoring above 7 in first year.
2. Column (1) includes tutorial fixed effects. The remaining regressions include course-cohort fixed effects.
3. Regressions use a third order polynomial in first-year grade, as well as their interactions with the treatment and include distance to the university, age, gender, and European Economic Area.
4. The peer group average is the leave-out mean.
5. Bandwidth is 0.5.
6. t-statistics in parentheses, standard errors are clustered on the student level.
7. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.14: **Attendance is Useful in Some Courses, but Not Others? Evidence from the Abolition Year.**

	TA Quality		Lecturer Quality	
	(1)	(2)	(3)	(4)
Courses where Attendance was Voluntary (Right of 7)	0.187	-0.021	-0.122	-0.012
	(0.63)	(-0.08)	(-0.46)	(-0.08)
Courses where Absence was Penalized (Right of 7)	0.135	0.271*	-0.133	0.024
	(0.84)	(1.96)	(-0.84)	(0.25)
Intercept	4.165***	4.094***	3.837***	3.826***
	(33.90)	(36.23)	(35.43)	(49.80)
Bandwidth	0.2	0.5	0.2	0.5
Observations	94	199	89	184
Adjusted R^2	-0.011	0.015	-0.014	-0.011
P-value for Difference Between Rows 1 and 2	0.866	0.239	0.955	0.777

Notes:
1. Sample is from year when forced attendance was abolished.
2. TA and Lecturer Quality are the averages of questions which are measured on a 5-point likert scale (1 equals strongly disagree and 5 equals strongly agree). Questions include, for example, "Lecturer is competent". See Appendix Table A.3 for detailed definitions of the dependent variables.
3. The p-value indicates whether the course dummies are significantly different from each other.
4. t-statistics in parentheses, standard errors are clustered on the student level.
5. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.15: **Absence of Other Channels.** Using All 8 Eligible Courses.

	General	Structure	Fairness	Usefulness Lectures
	(1)	(2)	(3)	(4)
Average 1^{st}-year Grade is Below 7	-0.118	-0.340	-0.334	-0.0618
	(-0.58)	(-1.53)	(-1.39)	(-0.13)
Constant	4.064***	3.963***	3.698***	3.483***
	(22.43)	(27.76)	(17.47)	(15.28)
Observations	1003	1005	910	603
Adjusted R^2	0.220	0.243	0.244	0.041

Notes:
1. The dependent variables are drawn from the course evaluations using all 8 eligible courses. See Table A.3 for detailed definitions of the dependent variables.
2. Regressions include course-cohort fixed effects, a third order polynomial in first-year grade, its interaction with the treatment, distance to university, age, gender, and European Economic Area.
3. Bandwidth is 0.5 around first-year grade of 7.
4. The intercepts are calculated via regressions which exclude course-cohort fixed effects and controls. They approximate the outcome mean near the threshold of students right of seven.
4. t-statistics in parentheses, standard errors are clustered on the student level.
6. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.