The Price of Forced Attendance

Online Appendix

Sacha Kapoor[*][†]   Matthijs Oosterveen[‡]   Dinand Webbink[§]

May 6, 2020

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

[†]Corresponding author. Address: Burgemeester Oudlaan 50, 3062 PA Rotterdam, the Netherlands. E-mail: kapoor@ese.eur.nl. Telephone: +31 10 4081406

[‡]Advance/CSG, Lisbon School of Economics and Management, University of Lisbon. oosterveen@iseg.ulisboa.pt

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

Figure A.1: **Additional Balancing Tests with High School Grades.**

kernel = epanechnikov, bandwidth = 0.2411
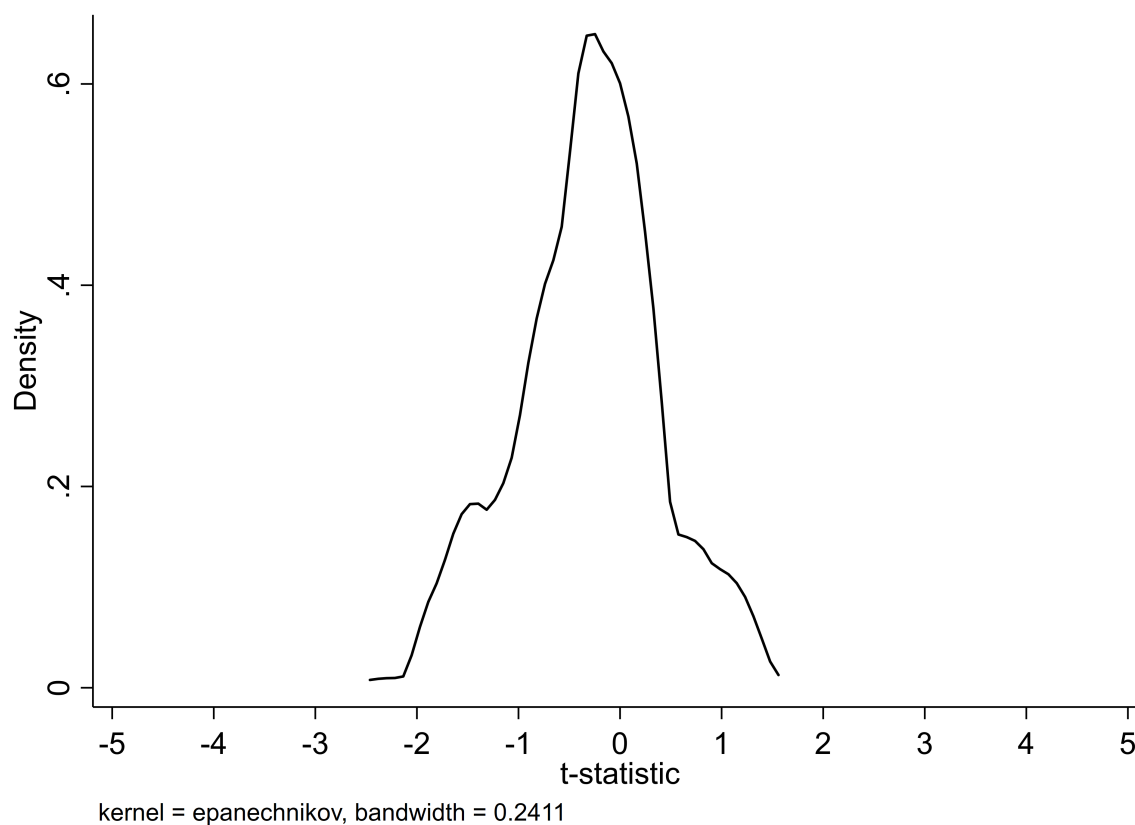
Notes:

1. Each t-statistic is from a balancing test where the dependent variable is the high school grade for a particular subject.

2. The balancing-test regressions include a first-order polynomial which is interacted with the treatment. The bandwidth is 0.365 and the kernel is triangular.

3. The data includes 133 subjects. Our regressions use grades from 44 of these subjects because for several subjects the number of observations was insufficient.

4. The density estimates are weighted by the number of students taking the subject in high school. The average number of students with a grade in these subjects is 111.

Figure A.2: **Discontinuity in Density Test for Females and Males.**

(a) Females

(b) Males
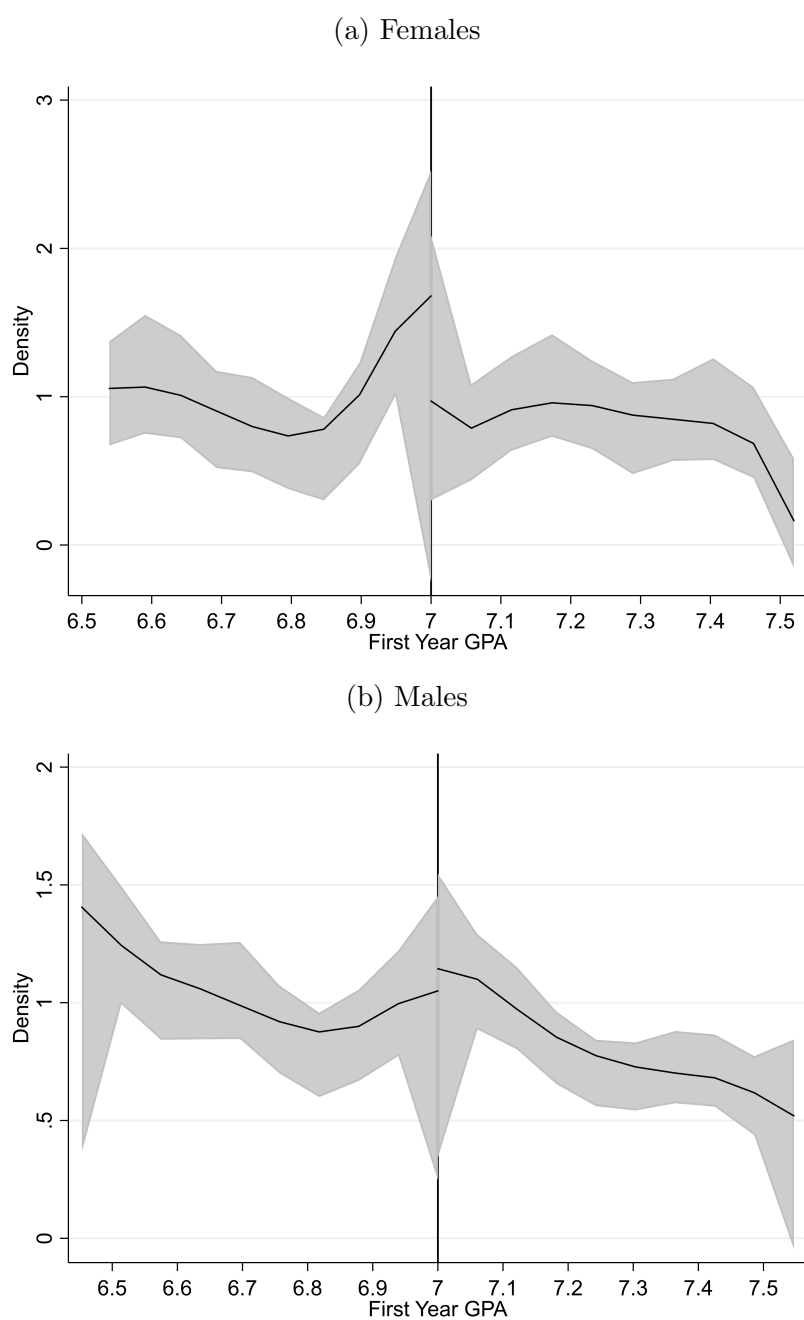
Notes:

1. Density is for the number of students.

2. Figures plot unrestricted manipulation test, where under the null hypothesis the limiting densities of the number of students to the left and right of 7 are the same. Test is unrestricted in that the estimates of densities to left and right of 7 are unrelated.

3. Figures use a second-order polynomial for density estimation and a third-order polynomial for the bias-correction estimate (see Cattaneo, Jansson, and Ma (2018, 2019)). Kernel is triangular. Confidence intervals use jackknifed standard errors.

4. The bias corrected discontinuity test statistic and p-value for females are -1.28 and 0.20. The analogs for males are 0.03 and 0.98. The statistics imply that in both cases we cannot reject the null hypothesis of no discontinuity around the cutoff.

Figure A.3: **Second Year Grades, by Course Type, in the Abolition Year.**
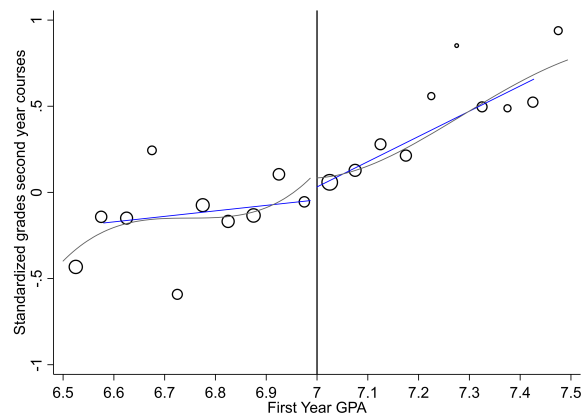
(a) Attendance is Forced Left of 7, Voluntary to Right
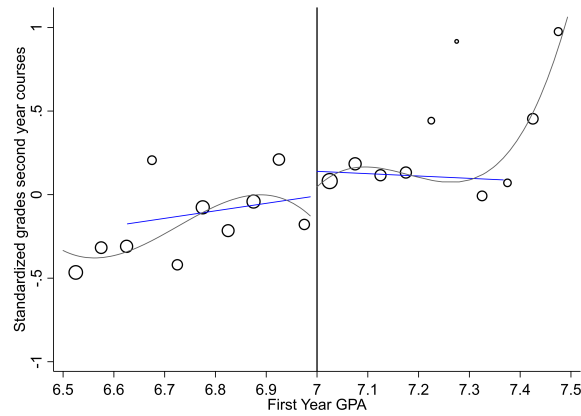
(b) Attendance is Forced Left of 7, Strongly Encouraged to Right
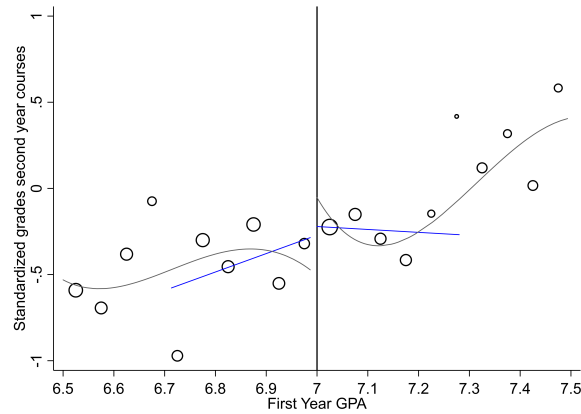

(c) Attendance is Forced Left and Right of 7

Notes:

1. Grades after the university policy was abolished (2014-15).

2. Locally linear and cubic scatterplots for second-year grades against first-year GPA.

3. The local linear polynomial is estimated upon the optimal bandwidth for each outcome relative to a MSE criterion (Calonico, Cattaneo, Farrell, and Titiunik, 2017). The cubic polynomial is estimated upon a bandwidth of 0.5, which is the same across all figures.

4. Dots are based on local averages for a binsize of 0.05. Dot sizes reflect the number of observations used to calculate the average.

5. Binsizes for local averages are selected via F-tests from regressions of second-year grades on K bin dummies and $2K$ bin dummies for the first-year GPA.

Figure A.4: **Sensitivity to Bandwidth for Attendance-Voluntary Courses.**
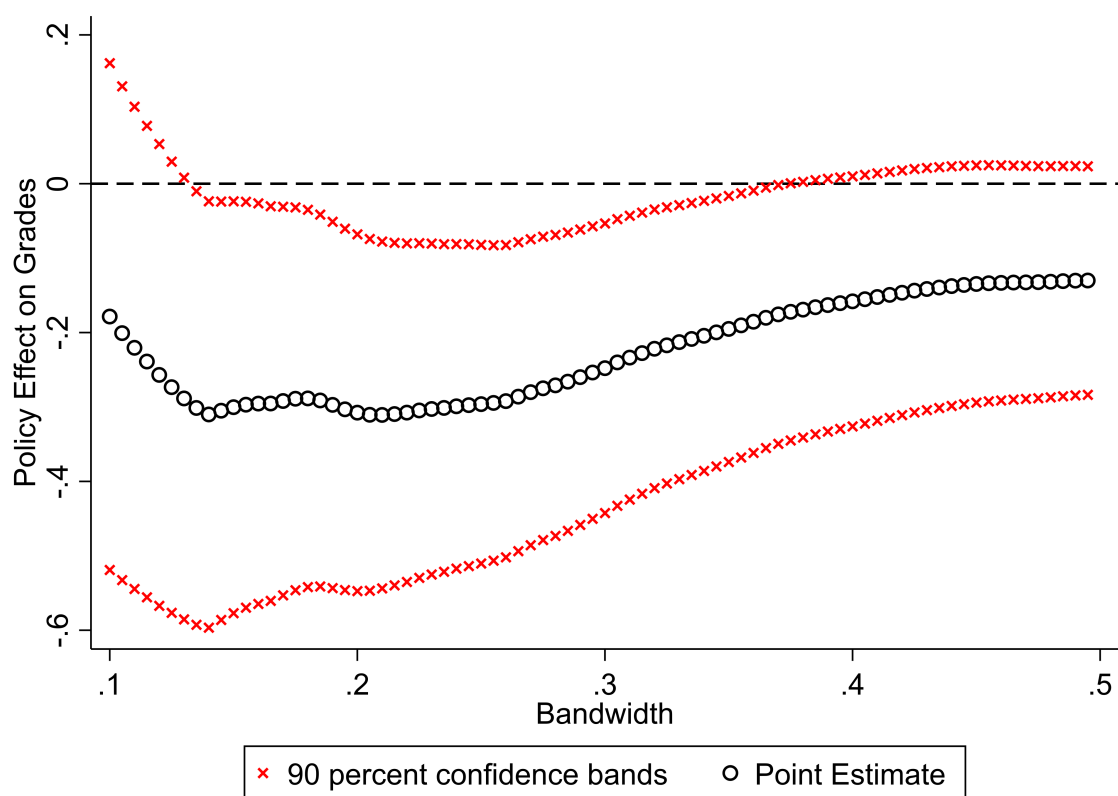
Notes:

1. This figure shows the policy estimate and its confidence interval against different bandwidths for 7^+vol courses only; the courses where above-7 students had full discretion over their attendance.

2. The regressions include a first-order polynomial which is interacted with the treatment. The kernel is triangular.

3. Bandwidth ranges from 0.1 until 0.5.

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. Economics of Ageing is taught in the English program and is replaced by Fiscal Economics in the Dutch program.
2. Students can compensate one 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 group is above 5.5. This applies to all students, whether they are above or below the threshold of the forced attendance policy.

Table A.2: **Attendance Policies of Second-Year Courses.**

Course	ECTS	Tutorials	Policy	Years	Tutorial Description	Exam Qs.	Block
International Economics	8	Yes	7^+enc	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	7^+for	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	7^+enc	2009/13	Exercises, Outside tutorials there are weekly quizzes that account for 20 percent of final grade.	MC	2
Applied Statistics II	4	Yes	7^+for	2009/13	Exercises, Accounts for 15 percent of final grade. Absence implies a 0 out of 15.	Open	2
Applied Microeconomics	8	Yes	7^+enc	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		2009/13	Group and individual research projects.		3
Methods & Techniques	8	Yes	7^+for	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	7^+vol	2010/13	Exercises, Actual Experiments, No direct influence on final grade.	MC	4
Intermediate Accounting	8	Yes	7^+vol	2009/13	Exercises, No direct influence on final grade.	MC	5
Research Project	4	No		2009/13	Group research projects.		5

Notes:
1. The Tutorial Description is extracted from course guides.
2. Following the Tutorial Description, the Policy column summarizes how the course treated the above-7 students. The abbreviations are discussed in more detail in the main text: 7^+vol indicates attendance was voluntary for above-7 students, 7^+enc indicates that attendance was strongly encouraged for above-7 students, and 7^+for indicates absence is penalized for below and above-7 students, effectively forcing both groups to attend.

Table A.3: **No Sample Selection when Matching Grades with Attendance.**

	Grade (Standardized)		Matched
	(1)	(2)	(3)
Matched	0.02	-0.02	
	(0.05)	(0.07)	
1st-year GPA is Below 7		-0.07	0.01
		(0.13)	(0.01)
Matched×Treatment		0.02	
		(0.11)	
Observations	3873	3873	3873

Notes:
1. Matched is a variable which equals 1 if the grade record found a match with the attendance data and 0 otherwise.
2. Column (1) regresses second-year grades on the matched-variable. The column shows grades are similar for matched and nonmatched records.
3. Column (2) shows no difference in the policy effect between matched and nonmatched records.
4. Column (3) regresses the match-variable on a treatment indicator, showing the policy is unable to explain whether or not a record is matched.
5. The regressions in Column (2) and (3) include a first-order polynomial which is interacted with the treatment. The bandwidth is 0.365 and the kernel is triangular.
6. Standard errors are clustered on the student and in parentheses.
7. Significance levels: * < 10% ** < 5% *** < 1%.

Table A.4: **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)
TA gives good tutorials	1-5	Quality TA
TA provides sufficient assistance	1-5	Quality TA

Notes:
1. Most questions are measured on a 5-point scale, where 1 equals strongly disagree and 5 equals strongly agree.
2. Total study time is measured on a 10-point scale, where 1 is 0 hours, 2 is $[1-5]$ hours, 3 is $[6-10]$ hours, and 10 is ≥ 40 hours.
3. Lecture attendance equals 1 if yes and 0 if no.

Table A.5: **Balancing Tests around the Cutoff with MSE optimal Bandwith for each Outcome.**

	Distance to Uni. (km)	Age	Gender	European Economic Area	High School GPA
	(1)	(2)	(3)	(4)	(5)
1st-year GPA is Below 7	3.062	0.224	0.211**	-0.026	-0.383
	(5.799)	(0.167)	(0.098)	(0.049)	(0.270)
Bandwidth	0.41	0.44	0.27	0.40	0.46
Observations	585	643	381	564	554

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

Table A.6: **Policy Effects when Forced Attendance was Abolished.**

	Grade (Standardized)		
	(1)	(2)	(3)
1st-year GPA is Below 7	-0.11 (0.27)	-0.18 (0.21)	-0.00 (0.18)
Course Type	$\tau^+ vol$	$\tau^+ enc$	$\tau^+ for$
Observations	279	430	425

Notes:
1. Sample is restricted to the cohort for which forced attendance was abolished.
2. The regressions include a first-order polynomial which is interacted with the treatment. The bandwidth is 0.365 and the kernel is triangular.
3. Standard errors are clustered on the student and in parentheses.
4. Significance levels: * < 10% ** < 5% *** < 1%.

Table A.7: **Sample Attrition.**

	Number of Courses	Course Evaluation Completed (1=yes, 0=no)		
	(1)	(2)	(3)	(4)
1st-year GPA is Below 7	0.13 (0.21)	-0.03 (0.05)	-0.07 (0.06)	-0.05 (0.06)
Intercept	9.17*** (0.15)	0.15*** (0.04)	0.19*** (0.04)	0.20*** (0.04)
Course Type	-	$\mathcal{T}^{+}vol$	$\mathcal{T}^{+}enc$	$\mathcal{T}^{+}for$
Observations	524	927	1424	1234

Notes:
1. Unit of observation is the student in Column (1). It is the student-course combination in Columns (2) to (4).
2. The regressions include a first-order polynomial which is interacted with the treatment. The bandwidth is 0.365 and the kernel is triangular.
3. Intercepts approximate the outcome mean near the threshold of students right of seven. For Column (1) this shows that students, forced or otherwise, complete (more than) 9 out of 10 courses.
4. Standard errors are robust or clustered on the student and in parentheses.
5. Significance levels: * $< 10\%$ ** $< 5\%$ *** $< 1\%$.

Table A.8: **Student Outcomes by Course Type, with Main Control Variables.**

	(1)	(2)	(3)
	Attendance Rate		
1st-year GPA is Below 7	0.30*** (0.04)	0.12*** (0.03)	0.00 (0.01)
	Grade (Standardized)		
1st-year GPA is Below 7	-0.16* (0.10)	0.08 (0.10)	-0.02 (0.10)
	Passes Course		
1st-year GPA is Below 7	-0.07 (0.04)	0.02 (0.05)	-0.03 (0.03)
Course Type	$\mathcal{T}^{+}vol$	$\mathcal{T}^{+}enc$	$\mathcal{T}^{+}for$
Observations	927	1424	1234

Notes:
1. Main control variables are course-cohort fixed effects, distance to the university, age, gender, and European economic area.
2. Attendance Rate is the percentage of tutorials attended. Passes Courses is a binary variable where pass=1 and fail=0.
3. The regressions include a first-order polynomial which is interacted with the treatment. The bandwidth is 0.365 and the kernel is triangular.
4. Standard errors are clustered on the student and in parentheses.
5. Significance levels: * < 10% ** < 5% *** < 1%.

Table A.9: **Student Outcomes by Course Type, with Main Control Variables and High School GPA.**

	(1)	(2)	(3)
	Attendance Rate		
1st-year GPA	0.28***	0.13***	-0.00
is Below 7	(0.04)	(0.04)	(0.01)
	Grade (Standardized)		
1st-year GPA	-0.17*	0.05	-0.02
is Below 7	(0.10)	(0.12)	(0.12)
	Passes Course		
1st-year GPA	-0.07	0.03	-0.00
	(0.05)	(0.06)	(0.04)
Course Type	$\mathcal{T}^+vol$	$\mathcal{T}^+enc$	$\mathcal{T}^+for$
Observations	762	1166	990

Notes:
1. High school GPA is observed for Dutch students only, which explains the fewer number of observations compared to the baseline regressions by course type.
2. Main control variables are course-cohort fixed effects, distance to the university, age, gender, and European economic area. These regressions additionally control for high school GPA.
3. Attendance Rate is the percentage of tutorials attended. Passes Courses is a binary variable where pass=1 and fail=0.
4. The regressions include a first-order polynomial which is interacted with the treatment. The bandwidth is 0.365 and the kernel is triangular.
5. Standard errors are clustered on the student and in parentheses.
6. Significance levels: * < 10% ** < 5% *** < 1%.

Table A.10: **Student outcomes by Course Type, with MSE Optimal Bandwidth and Robust Bias-Corrected Inference.**

	(1)	(2)	(3)
		Attendance Rate	
1st-year GPA is Below 7	0.30*** [0.00]	0.13*** [0.00]	0.00 [0.91]
MSE RD Bandwidth	0.43	0.41	0.39
CER RD Bandwidth	0.30	0.28	0.27
Observations	1125	1569	1310
		Grade (Standardized)	
1st-year GPA is Below 7	-0.26** [0.03]	0.03 [0.84]	-0.02 [0.79]
MSE RD Bandwidth	0.29	0.41	0.40
CER RD Bandwidth	0.20	0.28	0.28
Observations	724	1598	1350
		Passes Course	
1st-year GPA is Below 7	-0.07 [0.16]	0.01 [0.92]	-0.07* [0.07]
MSE RD Bandwidth	0.40	0.50	0.26
CER RD Bandwidth	0.28	0.35	0.18
Observations	1020	1965	906
Course Type	$\mathcal{T}^+ vol$	$\mathcal{T}^+ enc$	$\mathcal{T}^+ for$

Notes:
1. Attendance Rate is the percentage of tutorials attended. Passes Courses is a binary variable where pass=1 and fail=0.
2. The regressions include a first-order polynomial which is interacted with the treatment. The bandwidth is MSE optimal for each regression, *i.e.*, for each course-type and outcome-variable combination. The kernel is triangular.
4. Standard errors are robust, bias-corrected, and clustered on the student. *p*-values in squared parentheses.
5. Significance levels: * < 10% ** < 5% *** < 1%.

Table A.11: **Fake Cutoffs for Attendance-Voluntary Courses.**

| | Fake Cutoff at | | | |
	6	8	8.25	9
	Attendance Rate			
1st-year GPA	0.01	0.08	-0.01	-0.15
Below Fake Cutoff	(0.02)	(0.09)	(0.10)	(0.44)
	Grade (Standardized)			
1st-year GPA	0.13	-0.24	-0.00	0.18
Below Fake Cutoff	(0.23)	(0.17)	(0.14)	(0.35)
	Passes Course			
1st-year GPA	0.13	-0.01	0.01	0.00
Below Fake Cutoff	(0.14)	(0.03)	(0.01)	(0.00)
Observations	463	339	273	54

Notes:
1. The top of each column indicates at which first-year GPA we set the fake cutoff. Subsequently each column estimates the "policy effect" at that cutoff, for each outcome separately.
2. Sample is restricted to $7^{+}vol$ courses only; the courses where above-7 students had full discretion over their attendance.
3. Attendance Rate is the percentage of tutorials attended. Passes Courses is a binary variable where pass=1 and fail=0.
4. Main control variables are included: course-cohort fixed effects, distance to the university, age, gender, and European economic area.
5. The regressions include a first-order polynomial which is interacted with the treatment. The bandwidth is 0.365 and the kernel is triangular.
6. Standard errors are clustered on the student and in parentheses.
7. Significance levels: $^{*} < 10\%$ $^{**} < 5\%$ $^{***} < 1\%$.

Table A.12: **TA and Lecturer Quality by Course Type.**

	Teaching Assistant		Lecturer	
	Gives Good Tutorials	Provides Sufficient Assistance	Competent	Makes You Enthusiastic
	(1)	(2)	(3)	(4)
$\mathcal{T}^+vol$ Course	0.21*	0.11	0.00	-0.02
	(0.12)	(0.14)	(0.09)	(0.10)
$\mathcal{T}^+for$ Course	0.26***	0.31***	0.09	-0.08
	(0.09)	(0.10)	(0.06)	(0.07)
Intercept	3.95***	3.96***	3.95***	3.65***
	(0.07)	(0.07)	(0.05)	(0.06)
Observations	503	458	470	469
p-value: $\mathcal{T}^+vol = \mathcal{T}^+for$	0.72	0.15	0.21	0.53

Notes:

1. Sample is from year when forced attendance was abolished.

2. Course type refers to how individual courses dealt with above-7 students during the years of the policy. $\mathcal{T}^+vol$ indicates that above-7 students had full discretion over their attendance. $\mathcal{T}^+enc$ indicates that above-7 students were strongly encouraged to attend. $\mathcal{T}^+for$ indicates that above and below-7 students were penalized for being absent, effectively both groups were forced to attend in these courses.

3. The outcome variable is displayed at the top of each column. The questions are measured on a 5-point scale, where 1 is strongly disagree and 5 is strongly agree. See Appendix Table A.4 for more detailed definitions on the dependent variables.

4. Standard errors are clustered on the student and in parentheses.

5. Significance levels: $^*< 10\%$ $^{**}< 5\%$ $^{***}< 1\%$.

References

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

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

Cattaneo, M. D., Jansson, M., & Ma, X. (2019). Simple local polynomial density estimators. *Journal of the American Statistical Association, forthcoming.*