Journal of Economic Behavior and Organization 212 (2023) 935–981

Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

journal homepage: www.elsevier.com/locate/jebo

Compulsory class attendance versus autonomy☆

Sofoklis Goulas[a,*], Silvia Griselda[b], Rigissa Megalokonomou[c]

[a] Economic Studies, Brookings Institution, USA
[b] AXA Research Lab, Bocconi University, Italy
[c] Department of Economics, Monash University, Australia

ARTICLE INFO

Article history:
Received 11 March 2022
Revised 11 June 2023
Accepted 14 June 2023

JEL classification:
I26

Keywords:
Learning autonomy
School attendance
Returns to education
Natural experiment

ABSTRACT

We estimate the effect of an increased autonomy policy for higher-performing students on short- and longer-term school outcomes. We exploit an institutional setting with high demand for autonomy. Identification comes from a nationwide natural experiment that allowed higher-achieving students to miss 44 percent more classes with parental approval. Using a difference-in-difference-in-differences approach, we find that allowing higher-achieving students to skip more classes increases their performance in subjects that matter for university admission and improves the quality of their enrolled college degree. Top-performing students and students in more academically diverse classrooms demand more autonomy when it is offered.

1. Introduction

Most people encounter learning autonomy only after high school. Until the end of high school, school administrators and teachers usually mandate strict class attendance among students because of the returns to school attendance and graduation (Angrist and Keueger, 1991; Oreopoulos, 2006; 2007; Messacar and Oreopoulos, 2012; Brough et al., 2021). These attendance policies provide students with structure and help them overcome behavioral tendencies towards nonacademic activities and ultimately refrain from decisions that can harm lifetime welfare (Lavecchia et al., 2016). By this token, and as long as attendance is beneficial, additional structure should improve academic performance. At the same time, structure limits choices (e.g., time on independent study) that impact performance and precludes sensible students from choices that best serve their own self-interest. By doing so, time structure hinders academic progress.

Student autonomy is likely to have different effects on low- and high-performing students. While previous studies have analyzed the returns to absences for low-performing students (Marburger, 2001; Dobkin et al., 2010; Snyder et al., 2014; Kapoor et al., 2021), in this paper we investigate how increased autonomy, in the form of a lower class attendance requirement, affects higher-achieving high school students. There are two main reasons the optimal level of attendance of higher-performing students may differ from that of lower-performing students. First, students of higher academic perfor-

☆ The authors wish to thank seminar participants at the University of Melbourne, the XXXI Annual Conference of the Italian Society of Public Economics, the Applied Econometrics Workshop at the University of Wellington, and the 6th IZA Workshop on the Economics of Education. Any errors or omissions are our own. Silvia Griselda acknowledges financial support for this research from the University of Melbournes FBE Doctoral Program Scholarship.

* corresponding author.
E-mail addresses: sgoulas@brookings.edu (S. Goulas), silvia.griselda@unibocconi.it (S. Griselda), rigissa.megalokonomou@monash.edu (R. Megalokonomou).

https://doi.org/10.1016/j.jebo.2023.06.018
0167-2681/

mance may have stronger self-regulation, and therefore it may be easier for them to acquire knowledge on their own rather than in a classroom setting (Zimmerman, 2008). Second, classroom-based instruction may not offer sufficiently challenging material to higher-performing students, who could potentially learn more from targeted projects and tasks tailored to their knowledge capital (Reis and Renzulli, 2010).

To investigate the impact of autonomy, we use new data for more than 12,000 students from a sample of 107 public high schools in Greece. We draw on a nationwide increased autonomy policy implemented by the Ministry of Education, which intended to encourage high school students autonomy and provide them with greater discretion over their time management.[1] The increased autonomy policy allowed high school students with a prior-year grade point average (GPA) greater than 75% to miss 44 percent more classes than before with parental approval. Targeted students were given greater flexibility to make choices that best served their own interests (i.e., time on self-study or leisure). Attendance requirements were unaffected for students with a prior-year GPA equal to or lower than 75%.

Students maximize their utility by allocating time between leisure and study (either in class or at home) of subjects with differing utility weights. We consider two types of subjects: *high-stakes subjects*, which matter for university admission, and *low-stakes subjects*, which do not play a role in university admission. This optimization problem is subject to time constraints. Attendance requirements determine the time students allocate to class study. Relaxing attendance requirements enables students to reallocate time from class study to home study or leisure. Intuitively, students may skip school classes on low-stakes subjects to study for high-stakes subjects out of school. Additionally, we hypothesize that the optimal number of absences may be higher for students with higher out-of-classroom learning productivity than in-classroom learning productivity, such as higher-performing students or students. We empirically investigate these hypotheses.

We discuss the requirements for the identification of the autonomy policy effect. Our main empirical strategy employs a *difference-in-difference-in-differences* (DDD) methodology to estimate the intention-to-treat effects of autonomy on targeted students' short- (i.e., high school performance) and longer-term outcomes (i.e., performance on university admission exam and quality of enrolled degree). We compare changes in school absences, performance, and university admission outcomes for students from grade 11 to grade 12 for both targeted (those with a prior-year GPA above 75%) and non-targeted students (those with a prior-year GPA at or below 75%) in unaffected and affected cohorts. By using multiple cohorts and conditioning on student fixed effects, we are able to control for unobserved factors that confound the propensity to attend class. We solidify the credibility of the identification assumptions by examining the existence of common trends between affected and unaffected cohorts. We also validate the triple-differences results through a set of matching methodologies.

Our results show that providing increased autonomy to targeted students decreases their overall attendance by roughly four classes per year relative to non-targeted students. We estimate the impact of autonomy on targeted students' performance in subjects that matter for university admission (*high-stakes subjects*) to be 0.07 standard deviations. The estimated effect of autonomy on targeted students' performance may include a motivational effect of treating students as independent learners. In contrast, targeted students' performance in *low-stakes subjects*, which do not matter for university admission, remains unaffected. Our results also reveal that higher-performing students score higher on university admission exams and are admitted to more selective university degree programs when they are allowed more autonomy in the form of parent-approved class absences.

We find varying effects of autonomy based on prior performance and class heterogeneity. In particular, we show that targeted students of higher prior performance use more autonomy but their performance gains do not differ substantially from those of other targeted students. Additionally, targeted students miss more classes when quasi-randomly assigned to more academically diverse classrooms. As long as classmate heterogeneity is related to disruption in classroom learning, our results suggest that higher-performing students may want to distance themselves from disruptive learning environments. Targeted students in more academically diverse classrooms do not post increased performance gains associated with their additional class absences. Our finding that higher-performing students demand more autonomy in academically diverse settings may point to higher-performing students' increased readiness to self-learn and manage material independently.

We contribute to the literature in several important ways. Our study is related to the literature on the effect of school absences on performance. Absenteeism is a prevalent phenomenon. Data from the Snyder and Dillow (2015) and Liu et al. (2021) show that in the US, high school students skip school 12–15 days per year on average.[2] The literature documents no (Krohn and O'Connor, 2005; Caviglia, 2006; Martins and Walker, 2006) or a negative association between absenteeism and performance (Romer, 1993; Moore, 2006; Cohn and Johnson, 2006; Gottfried, 2010; Arulampalam et al., 2012; Latif and Miles, 2013; Gaete, 2018; Goulas and Megalokonomou, 2020; Liu et al., 2021). Most studies examine the effect of compulsory class attendance at the college level (Romer, 1993; Marburger, 2001; Moore et al., 2003; Marburger, 2006; Martins and Walker, 2006; Arulampalam et al., 2012; Latif and Miles, 2013; Kapoor et al., 2021). This study is the first one, to our knowledge, to examine the impact of learning autonomy on students as young as high schoolers.

The literature has also focused on the effect of attendance requirements primarily on lower-performing students (Marburger, 2001; Dobkin et al., 2010; Snyder et al., 2014; Lavecchia et al., 2016; Kapoor et al., 2021). Parents and policymakers are interested though in the effect of class attendance requirements on students in different parts of the ability

[1] More than 90% of high-school students attend public schools.

[2] Liu and Salvo (2018) document differences in absenteeism across countries. The authors find that students in China skip class more than their peers in the US, Canada, or Europe.

distribution. For example, the Report of the National Education Commission in the U.S. points out that forced class attendance may be detrimental for high-achieving students. This is because high-ability high school students often experience boredom, low motivation, and frustration when forced to spend more time than they need to on a curriculum designed for students of moderate ability (NECTL, 2005); the report characterizes high-performing students as "prisoners of time." Our paper contributes to a novel strand of the literature that examines the impact of compulsory attendance on higher-performing students (Kapoor et al., 2021).

We also build on the literature on the association between classroom composition and learning productivity in class. Previous studies have found that teacher effectiveness drops in more academically diverse classrooms (Steinberg and Garrett, 2016; Aucejo et al., 2021; Salvati, 2021). Thus, a more academically diverse learning setting may be less conducive to learning. We examine the impact of autonomy on attendance and performance in classrooms where prior performance of peers varies quasi-randomly. Our findings suggest that optimal class attendance may be lower for higher-performing students in a more diverse classroom.

Overall, this paper contributes to a better understanding of students' performance maximization problem subject to time allocation constraints. We show that constraints on how students spend their time play a key role in school performance and shape the quality of their higher education, and consequently their careers and income.

2. Institutional framework

2.1. The education system in greece

The education system in Greece is highly centralized (OECD, 2018). More than 90% of students attend traditional public schools (Goulas and Megalokonomou, 2021).[3] The education system in Greece has a very strict attendance policy. All high-school students attend classes back to back with short recesses in between from 8 am to 2 pm, Monday through Friday. Among OECD countries, Greek high school students are among those who spend more time in class every year (OECD, 2015). In the Greek system, there are two types of class absences: excused and unexcused absences. Absences can be excused only by parents or guardians, or by a doctor in case of sickness, and only if they involve entire school days. Unexcused absences do not have a permission from parents or doctors and are associated with tardiness or expulsion.[4]

Until the end of the 2005-06 school year, every student was allowed a maximum of 50 unexcused and 64 hours (about 10 days) of excused class absences in a year. One absence is equal to one missed school period. Missing one day of school equals as many absences as the number of school periods in that day (approximately six absences). The penalty for exceeding the number of allowed absences is to repeat the grade.[5] By design, periods of the same subject are spread out within the week's schedule. Thus, excused absences cannot be used to skip only specific subjects.

Students are assigned to the high school that serves the zone of their residential address (Goulas et al., 2018). The assignment of students and teachers to classrooms in each school is quasi-random.[6] In particular, in accordance with a strictly enforced law, in the beginning of high school, students are assigned to classrooms in alphabetical order based on their last name (Goulas et al., 2020; 2022; 2023).[7] Students with a last name starting with a letter earlier in the alphabet are given a classroom number smaller than the classroom number given to students with a last name starting with a letter later in the alphabet. Students are not allowed to switch classrooms. The alphabetical classroom assignment allows for exogenous peer group formation.[8] Teachers rotate between classrooms to teach classes of their expertise.

High school students have limited choice over their courses. There are two types of high school courses: general education and specialization courses. All 10th-grade courses are general education courses. Students choose a specialization track in grades 11 and 12. The typical student in grade 11 (12) takes 10 general education courses and three (four) specialization courses. Around 60% of instructional time is spent on compulsory general education courses. General education courses include Modern Greek Language, Math, Physics, and History. The remaining 40% of instructional time is allocated to specialization courses.[9] The available specialization tracks are Classics, Science, and Information Technology (IT). Each track requires that students take a different set of specialization courses. For the Classics track, specialization courses include Latin and Ancient Greek, while for the Science and IT tracks the specialization subjects include Math and Physics.

Post-secondary education is free of tuition fees. Twelfth-graders take end-of-the-year national standardized exams for university admission in a subset of the subjects taught. These include general education Modern Greek and track-specific subjects. Every student gets a *university admission score*, which is their average performance on nationally tested subjects.[10]

3 Roughly 2% of students attend public experimental (charter) schools and approximately 8% attend private high schools.

4 Arriving late to class or leaving school at midday results in unexcused absences. The principal or a teacher can expel a student.

5 Among high school students in our sample, only 1.30% were retained (i.e., held back) due to excessive absences.

6 Evidence of the random teacher assignment in that context can be found in Lavy and Megalokonomou (2019).

7 See Government Gazette of the Hellenic Republic 167 A/1566/1985.

8 Table A.5 provides evidence of quasi-random peer group formation.

9 Students attend general education classes with all their alphabetically assigned peers and they attend specialization classes with a subset of their alphabetically assigned peers based on specialization track choices.

10 The format of the national exams is the same as the within-school end-of-year exams, but the former are externally graded and proctored. School exams are usually graded by teachers using grading guidelines from the Ministry of Education under the principal's supervision.

Table 1
Empirical Identification Design.

Cohort	Group	Grade	Outcomes	D_1	D_2	D_3
'07	Targeted	12	$I^{T}_{'07}$ + G + '07_{12} + T_{12} + D	G + '07_{12} + T_{12} + D	($T_{12} - NT_{12}$) + D	D
		11	$I^{T}_{'07}$			
	Non-Targeted	12	$I^{NT}_{'07}$ + G + '07_{12}+ NT_{12}	G + '07_{12} + NT_{12}		
		11	$I^{NT}_{'07}$			
'06	Targeted	12	$I^{T}_{'06}$ + G + '06_{12}+ T_{12}	G + '06_{12} + T_{12}	$T_{12} - NT_{12}$	
		11	$I^{T}_{'06}$			
	Non-Targeted	12	$I^{NT}_{'06}$ + G + '06_{12}+ NT_{12}	G + '06_{12} + NT_{12}		
		11	$I^{NT}_{'06}$			

We use a triple-differences identification strategy that compares targeted (i.e., students with prior-year GPA above the eligibility threshold) 12th-grade students in the 2007 graduating cohort (treatment cohort) with non-targeted (i.e., students with prior-year GPA below the eligibility threshold) 12th-grade students in the 2007 graduating cohort; targeted 12th-grade students in the 2006 graduating cohort; and targeted and non-targeted 11th-grade students in the 2006 and 2007 graduating cohorts.

Table 2
Effect of the Increased Autonomy Policy on Targeted Students' Absences and Performance.

	Class Absences			Performance	
	Total	Excused	Unexcused	High-stakes Subjects	Low-stakes Subjects
Targeted in Treated Cohort [1]*relative to non-targeted*	3.907**	3.093***	0.815	0.073***	-0.015
	(1.550)	(1.050)	(0.981)	(0.018)	(0.020)
Non-targeted in Treated Cohort [2]*relative to control cohort*	1.511	2.807**	-1.295	-0.020	0.015
	(1.811)	(1.172)	(0.891)	(0.012)	(0.011)
Observations	24,542	24,542	24,542	24,542	24,542
Y Mean (Non-targeted)	68.22	29.41	38.81	-0.44	-0.44
Y St. Dev. (Non-targeted)	48.64	23.16	41.36	0.71	0.60
Y Mean (Targeted)	57.41	29.32	28.09	0.79	0.78
Y St. Dev. (Targeted)	32.08	25.06	14.84	0.53	0.62
Student FE	✓	✓	✓	✓	✓
P-value for H0: [1] + [2] = 0	0.01	0.00	0.60	0.00	0.97

Low-stakes subjects include general education Mathematics, History, and Physics. High-stakes subjects include general education Modern Greek, Ancient Greek, Latin for students in the Classics track, and Mathematics and Physics for students in the Science and IT tracks. Performance in high- and low-stakes subjects is standardized at the school-grade-cohort-subject level. Coefficient *Targeted in Treated Cohort*, [1], represents the effect of increased autonomy policy on targeted students in grade 12 in treated school year 2006-07 relative to non-targeted students in grade 12 in the same cohort (β_1 in specification (1)). Coefficient *Non-targeted in Treated Cohort*, [2], captures the effect of the increased autonomy policy on non-targeted students in grade 12 in school year 2006-07 relative to non-targeted students in the control school year 2005-06 (β_2 in specification (1)). We test the hypothesis that the full effect of the increased autonomy policy on targeted students in grade 12 in 2006-07 (relative to non-targeted students in the control cohort) is equal to zero. All specifications include student fixed effects, an indicator for being targeted in a given year, and classroom-level controls, such as classroom size, standard deviation of prior performance in the classroom, and the proportion of females in the classroom. Standard errors clustered at the school level are reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < .01$.

The average university admission score of all admitted candidates in a university degree program is used to infer that degree's *quality/selectiveness*.[11]

We consider 12th-grade subjects that matter for high school graduation and for university admission and their 11th-grade equivalents to be high-stakes subjects. We consider 12th-grade subjects that matter for high school graduation but not for university admission and their 11th-grade equivalents to be low-stakes subjects.

2.2. The increased autonomy policy

Before the beginning of the 2006-07 school year, the Ministry of Education announced a nationwide policy to encourage students' autonomy (Gov. Gazette 65/A/30-3-2006). The policy provided eligible students with 50 additional excused class

[11] Following the national exams, students submit a preference list of degree programs to the Ministry of Education. Candidates are ranked based on their admission scores. The admission algorithm admits the top candidate to his/her top choice, and the algorithm admits each candidate to their most preferred degree program that has not reached its admission capacity before moving to the next candidate. Candidates can include as many degree programs as they want on their preference list. No fees are charged in the admission process. Goulas et al. (2018) and Goulas and Megalokonomou (2021) provide a detailed description of the university admission process.

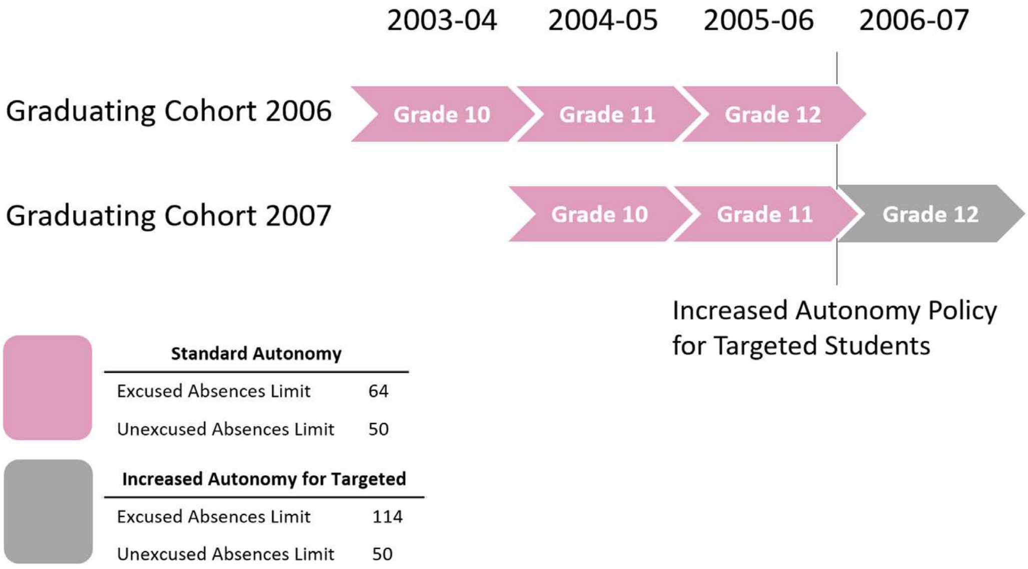


Fig. 1. Timeline of the Introduction of the Increased Autonomy Policy This figure shows the timing of the increased autonomy policy. Students in the 2006 graduating cohort were subject to standard autonomy policy throughout their high school career. Twelfth-graders in 2007 with a prior-year GPA above 75% were eligible for increased autonomy in the form of 50 additional parent-approved (excused) absences in a year before retention..

absences. Every student who had received a raw GPA higher than 75% the year before was eligible to take up more parent-approved absences in the current year. The policy intended for targeted students—those with a prior-year GPA above 75%—to have greater flexibility in making decisions related to their class attendance that best serve their own interests (i.e., time on self-study or leisure). We consider the cohort graduating high school in 2006 to be the control group and the cohort graduating in 2007 to be the treated group. The cohort graduating in 2007 was subject to the increased autonomy policy when they were in grade 12 in school year 2006-07.[12] Students could not manipulate in advance their eligibility for the increased autonomy policy, since the autonomy policy was unanticipated and eligibility depended on prior-year GPA. Fig. 1 displays the timeline of the policy and the affected students.

3. Data sources and description

We investigate the effect of the increased autonomy policy on education outcomes by combining data from 2 administrative sources. First, we collected attendance, transcript, demographic, classroom assignment, and class enrollment data from 107 public high schools we visited across Greece, corresponding to roughly 10% of the public high schools in the country.[13] Our data include student records from all three grades of high school for the cohorts graduating in 2006 and 2007. Attendance records contain the number of excused, unexcused, and total class absences (in school periods) for each student in each school year.[14] Second, we obtained national exam performance and university admission data from the Ministry of Education for all students in the nation in the graduating cohorts of 2006 and 2007.

There is demonstrated demand for increased autonomy in our setting. Fig. A.1 shows that students tend to increase their school absences substantially between grades 11 and 12. The increase in absences in grade 12 is particularly pronounced for higher-performing students. Increased school absences in grade 12 suggest a demand for time flexibility or autonomy that may in some cases exceed supply. Students in grade 12 usually prepare for university admission exams and often wish to substitute time at school for out-of-school study time. This setup of increased demand for autonomy at the end of high school enables us to estimate the effect of relaxing the school-time allocation constraint on attendance and performance.

Fig. 2 plots the distribution of excused absences (Panel A) and unexcused absences (Panel B) for the unaffected (2006) and affected (2007) cohorts. Vertical lines indicate the upper absences thresholds before retention under the strict (old limit)

[12] The Government Gazette contains no information on the intended consequences of the laws. The term “increased autonomy policy” constitutes a researchers’ interpretation of the law that allowed qualified students to skip school more often.

[13] Using data from the same context, we have shown that the sample is nationally representative with regard to several important variables, such as student performance and university admission outcomes (Goulas et al., 2018; Goulas and Megalokonomou, 2021).

[14] Absences records are generated as follows. A student in each classrooms serves as the record keeper, who will prepare the attendance sheet at the beginning of each period (Goulas et al., 2023). Each teacher reviews, validates, and signs the attendance sheet in the beginning of each period. Another teacher transfers the attendance records to the school’s attendance book. The same teacher is also responsible for updating the electronic attendance records on the school’s database, where our data were retrieved from.

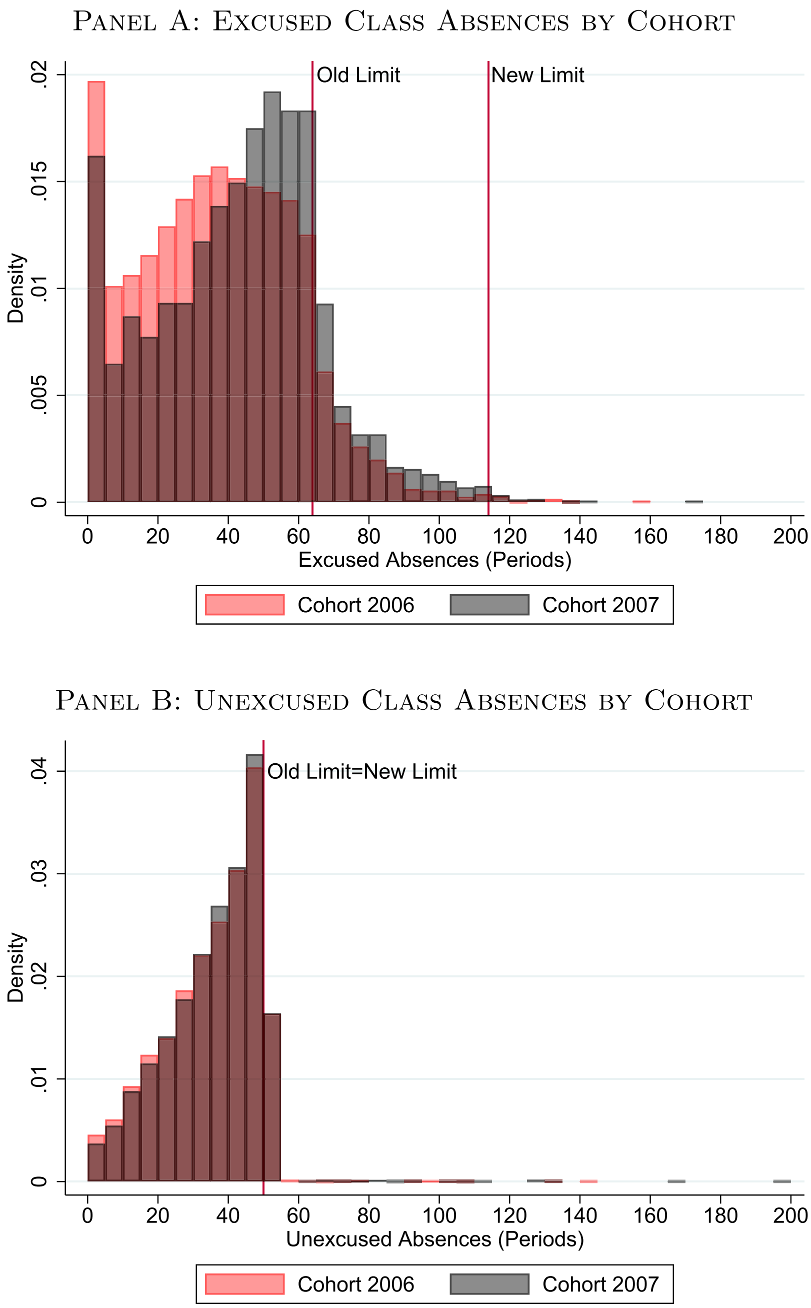


Fig. 2. Distributions of Absences in Grade 12 by Cohort Panels A and B show the distribution of excused and unexcused class absences, respectively, for the graduating cohorts of 2006 and 2007 in grade 12. Students graduating in 2006 could take up to 64 excused class absences and 50 unexcused class absences before retention. Targeted students graduating in 2007 could take up to 114 excused class absences and 50 unexcused class absences before being retained.

and the increased autonomy regime (new limit). The distribution of parent-approved (excused) absences shows a shift to the right for the 2007 cohort (affected cohort) relative to the 2006 cohort (unaffected cohort), while the ones of non-parent-approved (unexcused) absences for the unaffected and affected cohorts are similar. This reassures us that the shifts in the distributions of excused absences that we observe in Panel A of Fig. 2 can be attributed to the increased autonomy policy.[15]

[15] Fig. A.2 shows the distribution of total class absences by cohort. The figure shows a shift in the distribution of the number of total absences to the right for the 2007 cohort (increased autonomy policy) relative to the 2006 cohort (standard autonomy policy). The distributions suggest that students are

Table A.1 presents summary statistics. The average prior-year GPA—the running variable for targeted status—is 71.17 out of 100, and 37% of students are targeted.[16] Each student takes on average 64 class absences a year, 29 of which are excused and the rest unexcused. The average classroom-level standard deviation of prior-year GPA—our measure of academic diversity in the classroom—is 13.36. The final sample includes 12,240 unique students and 24,542 observations.

We consider final exam performance in two groups of subjects in grade 11 and grade 12: *high-* and *low-stakes* subjects. High-stakes exam performance in grade 12 matters for university admission. Low-stakes subjects include general education Math, Physics, and History. High-stakes subjects include general education Modern Greek Language and track-specific subjects. We consider track subjects offered in both grades 11 and 12. In particular, we use Ancient Greek and Latin for the Classics track, Mathematics and Physics for the Science track, and Mathematics and Physics for the IT Track. Subject-specific exam performance is standardized at the school-grade-subject level. University admission scores are standardized at the year-track level. Degree quality reflects the prestige/selectiveness of a degree. This is calculated as a degree's ranking based on the average admission score of each university department across the sample years and takes values from 0 to 100, with 100 being the highest. Table A.1 shows that, on average, admitted students enroll in a degree program between the 55th and 56th percentiles of quality.

4. Identification strategy

Given the sharp eligibility threshold at 75% of prior-year GPA, one is inclined to use a regression discontinuity design (RDD) to evaluate the effect of the increased autonomy policy on targeted students. To gauge the suitability of an RDD approach, we investigate the impact of the increased autonomy policy on attendance near and far the targeted status threshold.

Fig. A.5 shows the differential increases in absences from grade 11 to grade 12 between the treatment and the control cohort at different values of the prior-year GPA. We find that the increased autonomy policy is associated with an increase in parent-approved (excused) absences (Panel A) for students at every value of the prior-year GPA. It was possible for students below the eligibility threshold to take more absences during the autonomy policy regime as long as they did not go over their absences limits. We find no impact of the increased autonomy policy on non-parent-approved (unexcused) absences at any level of the prior-year GPA (Panel B).

The results in Fig. A.5 may point to the autonomy policy acting as a signal for all students interested in pursuing further education that they could use class absences to increase their learning autonomy and increase their likelihood of university admission. Policymakers intended autonomy to be used by higher-performing students but could not preclude lower-performing from increasing their class absences as long as their absences remained below retention level. Fig. A.5 indicates the potential presence of treatment spillover effects on the control group of non-targeted students. Treatment spillover is prevalent in interventions where treatment is perceived as beneficial (Keogh-Brown et al., 2007; Cook et al., 2010; Reichardt, 2019). The treatment spillover on non-targeted students limits the capacity of a regression discontinuity approach to capture the average treatment effect on the targeted.

To confirm our hypothesis that treatment spillover eliminates the autonomy policy effect around the targeted status threhold, we follow Lee and Lemieux (2010) and plot the estimated policy effects on the change in total absences between grades 11 and 12 for students in the control cohort (panel A) and treatment cohort (panel B). Fig. A.6 plots the results using binned local averages. We find no substantial difference between the grade-over-grade excused absences trajectory for students who are right below and right above the targeted status threshold of 75% in the prior-year GPA for either the control or the treatment cohort.

Treatment spillover produces non-zero autonomy policy effects on non-targeted students. Thus, identification strategies, such as an RDD approach, that rely on the comparison of outcomes between targeted and non-targeted students will lead to biased estimates. On the other hand, a DDD approach features a comparison of both targeted and non-targeted students in the treated cohort to a control cohort. This additional dimension of comparison allows us to obtain un-biased estimates of the autonomy policy effects on targeted students. Moreover, unlike RDD, a DDD approach measures the policy effect on all students, not only those around the eligibility threshold. Capturing the impact of autonomy on all students across the entire prior performance distribution is particularly relevant to school administrators and policymakers, who are interested in both first-order (i.e., on targeted students) and second-order (i.e., on non-targeted students) policy effects (Imbens and Wooldridge, 2009).[17] Estimating general equilibrium policy effects is also valuable in debates regarding the widening of school achievement gaps as a result of school policies that address individual student needs rather than treating all students the same (Pane et al., 2015).

Our main approach to investigate the average effect of autonomy on targeted students is a triple-differences framework.[18]

careful not to exceed the upper absences limit. Fig. A.3 shows the distribution of excused and excused absences for targeted and non-targeted students in the 2006 and 2007 graduating cohorts.

[16] Fig. A.4 plots the full distribution of prior-year GPA, the running variable for increased autonomy eligibility.

[17] Evaluations of education interventions often report the impact on both treatment-eligible and treatment-ineligible students in a given setting (Abdulkadiroğlu et al., 2016; Goulas et al., 2017).

[18] A simulation exercise in Section Appendix D of the Appendix shows the statistical power of our design (i.e., the probability our sample is capable of detecting the estimated DDD effects of the increased autonomy policy).

Table 1 summarizes our identification strategy. There are three dimensions of comparison. The first is between grade 11 and grade 12 of students on both sides of the targeted status threshold in either the 2006 or 2007 graduating cohorts. The second dimension of comparison is between students above (targeted) and below (non-targeted) the targeted status threshold of the prior year's GPA of 75% in either the 2006 and 2007 graduating cohorts. The third dimension of comparison is between the graduating cohorts of 2006 (standard autonomy policy) and 2007 (increased autonomy policy).

Vectors I^T_{07} (I^T_{06}) and I^{NT}_{07} (I^{NT}_{06}) in Table 1 reflect the set of student-specific components of targeted and non-targeted students, respectively, in the 2007 (2006) cohort. Quantity G represents the average outcome change between grade 11 and grade 12 in the population. Parameters $'07_{12}$ and $'06_{12}$ represent the mean outcome change between grade 11 and grade 12 of students in the 2007 and 2006 cohort, respectively. Quantities T_{12} and NT_{12} reflect the mean outcome change between grade 11 and grade 12 of targeted and non-targeted students, respectively. The impact of the increased autonomy policy on targeted students is captured by D. We identify the impact of increased autonomy using the following specification:

$$y_{icg} = \beta_0 + \beta_1 T_{ig} \times C_i \times G_g + \beta_2 C_i \times G_g + \beta_3 T_{ig} \times G_g + \beta_4 T_{ig} \times C_i + \beta_5 T_{ig} + \beta_6 G_g + X_c + \eta_i + \varepsilon_{icg}, \quad (1)$$

where outcome y_{icg} includes total absences, excused absences, unexcused absences, high-stakes exam performance, and low-stakes exam performance for student i in classroom c in grade g. Scores are standardized at the school-grade-subject level.

We control for student fixed effects in η_i. Student fixed effects allow us to account for student unobservables; these include ability, family background, and resources. It is important to account for student fixed effects as family characteristics may influence students' likelihood to take up more parent-approved absences. We also control for classroom characteristics, which include classroom diversity, measured by the standard deviation of prior-year GPA in the classroom; class size;[19] and the proportion of female peers in the classroom in X_c.[20] Variable G_g is an indicator for grade 12.[21] T_{ig} is an indicator taking the value one when student i in grade g has a prior-year GPA above the eligibility threshold (i.e., is targeted).[22] Variable C_i is an indicator for the treatment cohort of 2007.

Coefficient β_1 corresponds to the effect of the increased autonomy policy on targeted students. The fact that non-targeted students can take on more absences even when they are not eligible for increased autonomy in the 2006-07 school year suggests that our estimates of β_1 constitute an intention-to-treat effect of the increased autonomy policy.[23] Standard errors are clustered at the school level to allow for heteroskedasticity and serial correlation in student outcomes within each school.[24]

5. Validity of the identification strategy

We are able to interpret the DDD estimator as the causal intention-to-treat effect of the increased autonomy policy on targeted students, distinct from the effect of student-related idiosyncratic influences and peer effects, under three assumptions.

The first assumption requires that student characteristics are not correlated with school or classroom characteristics that influence learning productivity at school and consequently the likelihood of demanding increased autonomy. Demand for autonomy may be associated with learning productivity at school, which in turn is often associated with peer characteristics. If student assignment to peer environments differs from cohort to cohort, the differences in attendance and performance between them may not be fully attributable to the increased autonomy policy. Moreover, if students above and below the targeted status threshold are found in systematically different peer environments (for example, because of tracking), the differences in attendance and performance between them in the year the increased autonomy was introduced may not be fully attributable to the policy but also to systematic differences in their peers" characteristics.

Quasi-random peer group formation in our setting mitigates concerns that differences in the demand for autonomy between cohorts and between targeted and non-targeted students may be associated with differences in peer group characteristics (Goulas et al., 2020). Quasi-random assignment to classrooms at the beginning of high school guarantees that peer groups remain stable throughout high school. Table A.5 uses attendance and performance information from grade 10, the

[19] We include indicators for each class size value for estimation precision, given that class size is positively associated with academic diversity in the classroom ($\rho = 0.109$, $p < .001$).

[20] We use the leave-out mean of the female indicator to account for the mechanical relationship between a student's gender and their peers" (Guryan et al., 2009).

[21] Grade level may be related to attrition. Differential attrition between targeted and non-targeted students can lead to selection bias. The compulsory schooling age in Greece is 15 years, the age at which most students graduate from 9th grade (OECD, 2021). This suggests that, potentially, students who are likely to drop out of school may do so before they start 10th grade. Table A.2 compares the dropout rates of targeted and non-targeted students in the control (2006) and treatment (2007) cohorts. We find no statistically significant differences in the drop out rates between the 2006 and 2007 cohorts either overall or for targeted and non-targeted students separately.

[22] Because specification (1) controls for student fixed effects, coefficient β_5 captures the outcome component associated with students whose eligibility/targeted status changes between grade 11 and grade 12 (roughly 8%).

[23] Section Appendix B in the Appendix explores the effect of the increased autonomy policy on students who chose to increase their parent-approved (excused) absences during the autonomy policy regime. We find significant effects of the increased autonomy policy on high-stakes exam performance for those who chose to take more parent-approved absences during the autonomy policy regime. The magnitude of the estimated effect in high-stakes exam performance for students who chose to increase their excused absences during the autonomy policy regime is roughly two thirds of the corresponding estimated effect on targeted students. We find no evidence of a substantial impact of the increased autonomy policy on low-stakes exam performance.

[24] Tables A.3 and A.4 report estimates using clustering at different levels.

instance closest to peer group formation when data are available, to provide evidence that the alphabetical assignment to classrooms is practically random.[25] Specifically, Table A.5 shows that classroom numbers are not systematically associated with differences in attendance (excused and unexcused absences), GPA, academic diversity (measured by standard deviation of GPA), or the share of students above the targeted status threshold for increased autonomy.

The second assumption is that the outcomes of targeted students in the treatment cohort would not differ from the outcomes of non-targeted students in the same cohort of targeted students in the control cohort in grades prior to grade 12 in school year 2006-07, when the increased autonomy policy was in effect. If this common trends assumption is violated, the difference in targeted students" attendance and performance observed in the year the increased autonomy policy was introduced may not be fully attributable to the autonomy policy.

We investigate the common trends assumption necessary for the identification of the autonomy policy effect on targeted students. We compare mean excused and unexcused absences and performance in high- and low-stakes subjects of students above (targeted) and below (non-targeted) the eligibility threshold in the graduating cohorts of 2006 and 2007 in grades 10, 11, and 12.[26] We use the mean difference in outcomes in grade 11 outcomes as the benchmark.[27]

Figs. A.7 and A.8 plot the estimates for class absences and performance, respectively. We find that the pre-treatment trajectories of class absences and performance of targeted students in the treatment cohort in grade 10 are statistically indistinguishable from those in grade 11, suggesting that the common trend assumption is satisfied. Only the outcomes of excused absences and performance in high-stakes subjects for targeted students in the treatment cohort in grade 12 are statistically different from those in grade 10 or 11.

The third assumption requires treatment to be uncorrelated with the error term conditional on the cohort indicator, grade indicator, indicator of having a prior-year GPA above the increased autonomy eligibility threshold, classroom controls, and student fixed effects. This assumption would be violated if students anticipated the increased autonomy discretion provided to eligible students in the treatment year and exerted more effort in the prior year to obtain a GPA above the targeted status threshold. This manipulation of the running variable is not possible in the institutional setting we exploit in this study. The legislation that provided for the increased autonomy policy was not discussed, prepared, or published until near the end of the 2005-06 school year.[28]

We provide falsification tests of the effect of the increased autonomy policy on attendance and performance. The estimated effects of increased autonomy on non-parent-approved (unexcused) absences and on performance in low-stakes exams, which do not play a role in university admission, serve as placebos.[29]

6. Results

Table 2 presents our main results. We report the estimated impact of the increased autonomy policy on total class absences, parent-approved (excused) and non-parent-approved (unexcused) class absences, and performance in high- and low-stakes subjects for two groups. First, we show the effect of the increased autonomy policy on targeted students in the treatment cohort compared with non-targeted students in the same cohort. This effect corresponds to the coefficient of the triple interaction $T_{ig} \times C_i \times G_g$, β_1, in specification (1). Second, we report the effect of the increased autonomy policy on non-targeted students in the treatment cohort compared with students in the control cohort. This effect is captured by the coefficient of the interaction term $C_i \times G_g$, β_2, in specification (1).

Targeted students in the treated cohort took more absences and improved their high-stakes exam performance as a result of the increased autonomy policy. In particular, targeted students increased their total (excused) absences by four (three) additional class hours—roughly 0.09 (0.13) standard deviations—relative to non-targeted students during the year the increased autonomy policy was in effect. Targeted students" high-stakes exam performance increased by 0.07 standard deviations—or 0.019 standard deviations per missed period—due to the increased autonomy policy. Targeted students" low-stakes exam performance remained unaffected by the increased autonomy policy.

Intuitively, when students and families demand academic autonomy they may do so because they expect improved outcomes from it. Thus, granting autonomy in the form of parent-approved (excused) absences may lead to improved academic performance, especially in subjects that matter more (i.e., high-stakes subjects). At the same time, establishing an increased autonomy policy signals that autonomy may be academically beneficial. This signal is potentially impactful on its own accord. The estimated effects of the autonomy effect may reflect both an improved learning productivity away from disruptive settings but also the motivational effect of being treated as an autonomous learner.

[25] Using data from grade 10 mitigates contamination of outcomes from peer influences over the years.

[26] We carry backward the targeted status from grade 11 to grade 10 to impute missing targeted status information from grade 10. Grade 10 high-stakes subjects include Ancient Greek, Modern Greek, Algebra, and Physics. We consider "Technology" to be a low-stakes subject in grade 10 because it was compulsory: students submitted written work and took exams in it, but it did not count toward the GPA.

[27] An estimation using grade 10 as benchmark produces similar results (see Table A.6). Grade 12 is the treatment grade.

[28] See Government Gazette of the Hellenic Republic 65 A/30-3-2006.

[29] Fig. A.7 provides evidence that the effect of the autonomy policy on unexcused absences constitutes a valid placebo check for the impact of the autonomy policy on excused absences because the common trends assumption is satisfied for both types of absences. Fig. A.8 shows the impact of the autonomy policy on low-stakes exams can serve as falsification check for the impact of the autonomy policy on high-stakes exams because the common trends assumption is satisfied for both types of exams.

The policy is also associated with an increase of roughly 0.12 standard deviations in non-targeted students' excused absences (roughly three additional excused class absences) in the year the increased autonomy policy was introduced. The effect of the policy on non-targeted students' total absences or on school performance is not statistically significant. The unexcused absences of targeted and non-targeted students are unaffected by the increased autonomy policy.

It is possible that non-targeted students absences and performance patterns may change in response to the signal of positive value of academic autonomy associated with the autonomy policy announcement. This signal may induce non-targeted students to take more class absences. The impact of non-targeted students additional absences on performance depends on their relative learning productivity at and away from school. We find that even though non-targeted students take more parent-approved absences when the autonomy policy for targeted students is in effect, their scores decrease. This suggests that non-targeted students may imitate autonomy policy treatment status but autonomy does not increase their performance.

Longer-term outcomes exams in high-stakes subjects in grade 12 are centrally designed, organized, and graded. The equivalent grade-11 subjects to grade-12 high-stakes subjects are administered and graded at the school level.[30] Our standardization of scores in grades 11 and 12 at the school-grade-subject level allows us to remove school-specific influences in scores additionally to variation components associated different grades and subjects. We have explored the robustness of our findings to alternative standardization approaches at the year-subject-grade and subject-grade levels as well as when no standardization is applied. Our results in Table A.7 are comparable to the main results in Table 2.

7. Heterogeneous effects

7.1. By prior performance

We investigate heterogeneity in the effects of the increased autonomy policy by prior performance. If performance is associated with learning independence, one may expect students with higher prior performance to have higher performance gains when they take more class absences. We use prior-year GPA to measure prior performance. We explore heterogeneous increased autonomy effects by prior performance by replacing the targeted status indicator (T_{ig}) in specification (1) with breakout binary indicator variables of targeted status in each quartile of prior performance, T_{ig}^q, where $q \in \{1, 2, 3, 4\}$. The quartiles of prior performance for targeted students reflect parts of the prior-year GPA distribution above 75%, the targeted status threshold. Therefore, the sum of the T_{ig}^q indicators across quartiles equals the targeted indicator of student i in grade g, T_{ig} (i.e., $\sum_{q=1}^4 T_{ig}^q = T_{ig}$). This approach allows us to estimate nonlinear heterogeneous effects by prior performance. Our specification is the following:

$$y_{icg} = \beta_0 + \sum_{q=1}^{4} \beta_1^q T_{ig}^q \times C_i \times G_g + \beta_2 C_i \times G_g + \sum_{q=1}^{4} \beta_3^q T_{ig}^q \times G_g + \sum_{q=1}^{4} \beta_4^q T_{ig}^q \times C_i + \sum_{q=1}^{4} \beta_5^q T_{ig}^q + \beta_6 G_g + X_c + \eta_i + \varepsilon_{icg}, \quad (2)$$

where β_1^q captures the impact of the increased autonomy on targeted students in the treated cohort of prior-performance in quartile q.

Table 3 shows that higher quartiles of prior performance are associated with more absences during the increased autonomy regime relative to lower quartiles of prior performance. Specifically, targeted students with prior performance in the top and second quartile of the prior performance distribution skip roughly 10 and five more classes relative to non-targeted students, respectively. The absences of targeted students in classrooms with prior performance in the third and bottom quartile of the prior performance distribution is statistically similar to the absences of non-targeted students.[31]

Table 3 also reveals that the impact of the autonomy policy on high-stakes exam performance is positive and significant in every quartile of prior performance even though a pattern of differential gains as prior performance increases may not be evident. The impact of the increased autonomy policy on high-stakes exam performance is the highest for targeted students in the bottom quartile of prior performance followed by targeted students in the top quartile of prior performance. The impact of the increased autonomy policy on low-stakes exam performance is positive and significant only for targeted students in the bottom quartile of targeted students' prior performance. The results suggest that students with prior performance marginally above the policy eligibility threshold may use absences to improve their performance not only in high-stakes exams but also in low-stakes exams.

The increased autonomy policy has a negative impact on low-stakes exam performance for students in the second and third quartiles of prior performance with the effect being significant only in the second quartile. This finding may reflect a

[30] Details on how grade 11 assessments are graded can be found in Goulas and Megalokonomou (2021).

[31] We also employ an alternative approach in which we compare absences patterns in the treatment cohort of students in different parts of the prior-year performance distribution, regardless of targeted status, to those of students in the control cohort, while accounting for grade- and student-level unobservables. We plot the estimates for excused and unexcused absences in Panels A and B of Fig. A.9, respectively. We find that as prior-year performance increases, the impact of the increased autonomy policy on excused absences increases. We find no statistically significant effect of the increased autonomy policy on unexcused absences. A further investigation breaks out the effects by gender. We plot the estimates in Fig. A.10. We find no differential effect of the increased policy autonomy by gender in any part of the prior performance distribution.

Table 3
Heterogeneous Effects of Increased Autonomy Policy on Targeted Students by Prior Performance.

	Class Absences			Performance	
	Total	Excused	Unexcused	High-stakes Subjects	Low-stakes Subjects
Targeted in Treated Cohort × Prior Performance Top Quartile	10.256***	7.450***	2.806**	0.084***	0.011
	(2.020)	(1.431)	(1.306)	(0.019)	(0.027)
Targeted in Treated Cohort × Prior Performance Second Quartile	5.086**	4.097**	0.989	0.046**	-0.092***
	(1.962)	(1.601)	(1.144)	(0.020)	(0.027)
Targeted in Treated Cohort × Prior Performance Third Quartile	1.803	1.644	0.158	0.063***	-0.019
	(1.806)	(1.403)	(0.989)	(0.023)	(0.026)
Targeted in Treated Cohort × Prior Performance Bottom Quartile	-2.105	-1.231	-0.874	0.102***	0.061**
	(1.924)	(1.467)	(1.086)	(0.028)	(0.029)
Non-targeted in Treated Cohort	1.832	3.050***	-1.218	-0.023*	0.014
	(1.784)	(1.142)	(0.886)	(0.013)	(0.011)
Observations	24,542	24,542	24,542	24,542	24,542
Y Mean (Non-treated)	68.22	29.41	38.81	-0.44	-0.44
Y St. Dev. (Non-treated)	48.64	23.16	41.36	0.71	0.60
Y Mean (Treated)	57.42	29.32	28.09	0.79	0.78
Y St. Dev. (Treated)	32.08	25.06	14.84	0.53	0.62
Student FE	✓	✓	✓	✓	✓

Low-stakes subjects include general education Mathematics, History, and Physics. High-stakes subjects include general education Modern Greek, Ancient Greek, Latin for students in the Classics track, and Mathematics and Physics for students in the Science and IT tracks. Performance in high- and low-stakes subjects is standardized at the school-grade-cohort-subject level. Coefficients of the interactions of *Targeted in Treated Cohort* with the indicators *Prior Performance Top Quartile—Bottom Quartile* represent the effect of increased autonomy policy on targeted students in different quartiles of prior-year GPA relative to non-targeted students in the treatment year (β_1^q in specification (1)). Coefficient *Non-targeted in Treated Cohort*, captures the effect of the increased autonomy policy on non-targeted students in grade 12 in school year 2006-07 relative to non-targeted students in the control school year 2005-06 (β_2 in specification (2)). All specifications include student fixed effects, an indicator for being targeted in a given year, and classroom-level controls. Classroom controls include classroom size, quantiles of standard deviation of prior performance in the classroom, and the proportion of females in the classroom. Standard errors clustered at the school level are reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < .01$.

Table 4
Heterogeneous Effects of Increased Autonomy Policy on Targeted Students by Classroom Diversity (Standard Deviation of Prior Performance).

	Class Absences			Performance	
	Total	Excused	Unexcused	High-stakes Subjects	Low-stakes Subjects
Targeted in Treated Cohort × Class SD Top Quartile	8.765***	7.226***	1.539	0.035	0.006
	(3.068)	(2.328)	(1.459)	(0.022)	(0.026)
Targeted in Treated Cohort × Class SD Second Quartile	4.882*	4.151**	0.731	0.074***	-0.018
	(2.714)	(1.979)	(1.241)	(0.025)	(0.032)
Targeted in Treated Cohort × Class SD Third Quartile	2.734	0.738	1.997	0.085***	-0.029
	(2.411)	(1.803)	(1.212)	(0.024)	(0.030)
Targeted in Treated Cohort × Class SD Bottom Quartile	-0.131	0.810	-0.941	0.094***	-0.017
	(3.340)	(2.496)	(1.465)	(0.027)	(0.028)
Non-targeted in Treated Cohort	1.096	2.509**	-1.412	-0.019	0.017
	(1.835)	(1.185)	(0.905)	(0.012)	(0.011)
Observations	24,542	24,542	24,542	24,542	24,542
Y Mean (Non-treated)	68.22	29.41	38.81	-0.44	-0.44
Y St. Dev. (Non-treated)	48.64	23.16	41.36	0.71	0.60
Y Mean (Treated)	57.42	29.32	28.09	0.79	0.78
Y St. Dev. (Treated)	32.08	25.06	14.84	0.53	0.62
Student FE	✓	✓	✓	✓	✓

Low-stakes subjects include general education Mathematics, History, and Physics. High-stakes subjects include general education Modern Greek, Ancient Greek, Latin for students in the Classics track, and Mathematics and Physics for students in the Science and IT tracks. Performance in high- and low-stakes subjects is standardized at the school-grade-cohort-subject level. Coefficients of the interactions of *Targeted in Treated Cohort* with the indicators *Class SD Top Quartile—Bottom Quartile* represent the effect of increased autonomy policy on targeted students in the treatment year who are in classrooms of different diversity, as captured by quartiles of standard deviation of prior performance in the classroom relative to non-targeted students in the same year (β_1^q in specification (1)). Coefficient *Non-targeted in Treated Cohort*, captures the effect of the increased autonomy policy on non-targeted students in grade 12 in school year 2006-07 relative to non-targeted students in the control school year 2005-06 (β_2 in specification (2)). All specifications include student fixed effects, an indicator for being targeted in a given year, and classroom-level controls. Classroom controls include classroom size, quantiles of standard deviation of prior performance in the classroom, and the proportion of females in the classroom. Standard errors clustered at the school level are reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < .01$.

time substitution as targeted students in the middle part of the prior performance range may skip classes in low-stakes subjects and study for high-stakes subjects independently. The increased autonomy policy has no significant effect on low-stakes exam performance for targeted students in the top quartile of prior performance. This suggests that the top performers may be able to benefit academically from autonomy in high-stakes subjects without sacrificing their performance in low-stakes exams.

7.2. By classroom diversity

The policy may affect students differently depending on learning productivity in their classroom. Academically more diverse classrooms may offer lower learning productivity (Aucejo et al., 2021). Thus, one might expect targeted students to exploit the increased autonomy policy more when they are in more academically diverse classrooms. We investigate whether targeted students skip class more when quasi-randomly assigned to a more diverse class and estimate the impact on their performance. We use the standard deviation of prior-year GPA to measure classroom academic diversity.[32] Fig. A.11 reveals substantial variation in classroom diversity.

To empirically investigate heterogeneous increased autonomy effects by classroom academic diversity, we replace the targeted status indicator (T_{ig}) in specification (1) with breakout binary indicator variables of targeted status in each quartile of classroom academic diversity, T_{ig}^q, where $q \in \{1, 2, 3, 4\}$ (similar to specification (2)). The sum of the T_{ig}^q indicator vectors across quartiles equals the targeted indicator of student i in grade g, T_{ig} (i.e., $\sum_{q=1}^{4} T_{ig}^q = T_{ig}$). This approach allows us to estimate nonlinear heterogeneous effects by classroom diversity.

Table 4 shows that higher quartiles of classroom academic diversity are associated with more absences during the increased autonomy regime relative to lower quartiles of classroom academic diversity. In particular, targeted students in classrooms with academic diversity in the top quartile of the diversity distribution skip roughly nine more classes relative to non-targeted students. The absences of targeted students in classrooms with academic diversity in the bottom quartile of the diversity distribution are statistically similar to the absences of non-targeted students.[33]

We provide insights into our findings by comparing the absences and performance of targeted students, class size, and the percentage of targeted students in classrooms of different academic diversity prior to the increased autonomy policy. Table A.8 presents our results. Our findings suggest the existence of (1) a channel consistent with the *Boutique* model of peer effects, and (2) ceiling effects. According to the Boutique model, students benefit from being in a class with students of similar ability Sacerdote (2011). Table A.8 shows that, even before the autonomy policy took place, targeted students in more heterogeneous classrooms were taking more absences (columns 2–3) and had higher performance (column 1), compared to targeted students in more homogeneous classrooms. This suggests that, to the extent allowed prior to the autonomy policy, higher-performing students may have been maximizing their performance by taking more absences when found in more academically diverse classrooms. Consistent with ceiling effects, academically diverse classrooms may be more likely to have top performing targeted students, who may have less room for further score improvements, compared with more homogeneous classrooms.

Classroom diversity affects the academic performance of both targeted and non-targeted students.[34] One might worry that the estimated differential impact of the increased autonomy policy on targeted students by classroom diversity could be confounded by unobservables related to individual prior-performance. We mitigate this concern by augmenting specification (2) to control directly for prior-year performance (GPA) in a robustness investigation. Our results on Table A.10 show that the estimates of differential impact of the increased autonomy policy by classroom diversity remain similar when accounting for the influence of prior performance.[35] Our result provide suggestive evidence of the positive effect of tracking on student performance (Duflo et al., 2011; Card and Giuliano, 2016).

8. Robustness investigation

One might worry that pre-treatment imbalances between the treatment and control cohorts could bias the estimated effect of the increased autonomy policy (Cunningham, 2021). Specifically, if students in the control cohort have lower baseline absences level in grade 11 than students in the treatment cohort, the estimated effect of the increased autonomy policy on

[32] We also consider an alternative measure of classroom academic diversity, the interquartile range (IQR) of prior-year overall GPA. We define the IQR of prior-year GPA as follows: $IQR = \frac{Q_3 - Q_1}{2}$, where Q_1 and Q_3 represent, respectively, the first and third quartile of prior-year GPA.

[33] A further investigation examines the impact of the autonomy policy on non-targeted students in classrooms of different academic diversity (Table A.9). We find that the among non-targeted students, the autonomy policy has a greater impact on the excused absences of non-targeted students in more academically diverse classrooms (i.e., *Class SD Top Quartile*). We find no impact of the autonomy policy on the performance of any subgroup of non-targeted students. Our results suggest that students may seek autonomy when they face a highly diverse classroom environment, regardless of targeted status.

[34] The correlation between classroom diversity and the prior-year GPA of targeted and non-targeted students in the classroom is ρ =0.202 (p-value<0.001) and ρ =0.276 (p-value<0.001), respectively.

[35] We further investigate the robustness of our estimates when classroom diversity is measured using the IQR instead of the standard deviation of prior performance in the classroom. Table A.11 shows the estimated heterogeneous effects of the increased autonomy policy on targeted students using the IQR. The results show patterns that are similar to the results using the standard deviation of prior performance in the classroom.

absences may be downward biased.[36] This is because students with a lower level of total absences in grade 11 are more likely to have a higher increase in absences between grade 11 and grade 12.[37]

We deploy a matching methodology as a robustness exercise to gauge potential bias in the DDD estimates that stems from baseline grade 11 differences in outcomes in the control and treatment cohorts. We search for matches in the control cohort for each student in the treatment cohort, and perform caliper matching with replacement. Matching with replacement reduces bias, because control records that look similar to many treated records can be used multiple times (Stuart, 2010). This is particularly useful in settings in which there may be few control individuals comparable to the treated individuals (Dehejia and Wahba, 1999).

Our matching approach uses the entire control record pool to form a robust counterfactual record for each treatment record. This improves balance in the pre-treatment characteristics between the treatment and counterfactual records. Table A.12 compares the grade 11 baseline characteristics of matched treatment cohort records and their controls. We match on gender, age, having a prior-year GPA above or below the eligibility threshold, grade 11 GPA with a caliper width of $\pm$ 0.1 standard deviations, and grade 11 excused and unexcused absences with a caliper width of $\pm$ 20%. For treatment cohort records with more than five matches, the best five matches are kept with the highest proximity to the treatment record's grade 11 GPA and grade 11 excused and unexcused absences.[38] We apply equal weighting, which averages over multiple records in the control cohort for each record in the treatment cohort.[39]

Table A.14 shows the estimates from our matching approach. To alleviate potential bias in the matching estimator due to selective matching, we employ inverse probability weights (IPW) to account for differences in pre-treatment characteristics of matched and unmatched records. The increased autonomy policy is found to be associated with an increase in parent-approved absences of 3.762 (compared with 3.090 from DDD) and an increase in high-stakes exam performance by 0.151 standard deviations (compared with 0.073 from DDD).[40]

9. Longer-term outcomes

In this section, we investigate the impact of the increased autonomy policy on students' university admission exam score and the quality/selectiveness of their enrolled university degree program. The university admission score is a weighted average of the midterm and final exam performance in all high-stakes subjects in grade 12. University admission scores are standardized at the year-track level. Quality of enrolled degree is increasing in quality and is calculated as a degree's ranking based on the average university admission score of admitted students of each university department across the sample years.[41] We estimate the following value-added model:

$$y_{iycts} = \beta_0 + \beta_1 C_{iy} + T_i + W_i + X_c + + \zeta_y + \theta_t + \lambda_s + \varepsilon_{iycts}, \tag{3}$$

where y_{iycts} represents the outcome of grade 12 student i in cohort y in classroom c in track t in school s. We control for student gender, year of birth indicators, and grade 11 GPA in W_i. Indicator T_i captures being targeted by the increased autonomy policy (i.e., having prior-year GPA above 75%). Indicator C_{iy} takes the value one for targeted students in the graduating cohort of 2007. Vector X_c includes classroom-level controls: class size, share of females, and academic diversity. We account for cohort, track, and school fixed effects in ζ_y, θ_t, and λ_s, respectively. Coefficient β_1 captures the effect of the autonomy policy on targeted students in the treatment cohort. Standard errors are clustered at the school level.

If students act strategically and use their autonomy to improve their high-stakes exam performance, one might expect a positive impact of the increased autonomy policy on targeted students' university admission scores. Table 5 reveals significantly higher university admission scores among targeted students in the treatment cohort relative to non-targeted students in the same cohort. In particular, the increased autonomy policy is found to be associated with an increase in university

[36] Table A.12 shows the baseline characteristics in grade 11 of students in the control and treatment cohorts. If students in the control cohort had a higher level of absences in grade 11, closer to the starting absences in grade 11 of students in the treatment cohort, they might have had a lower increase in absences between grade 11 and grade 12 than their realized increase. This suggests that the DDD approach may overestimate the increase in absences between grade 11 and 12 of the counterfactual condition for students in the treatment cohort. This means that the DDD estimates may constitute a lower bound of the effect of the increased autonomy policy on students in the treatment cohort.

[37] The association between grade 11 absences and the change in absences between grades 11 and 12 is substantially negative ($\rho = -0.515$, p-value<0.001) across the entire sample. The correlation between grade 11 excused absences and the increase in excused absences between grades 11 and 12 is $\rho = -0.487$ (p-value<0.001) and $\rho = 0.552$ (p-value<0.001) for students in the control and treatment cohorts, respectively.

[38] We follow the practice in the literature of using a match-to-case ratio of five (Hennessy et al., 1999). Our results remain similar when we change the number of control cohort records per treated cohort records to four or six.

[39] One drawback of using multiple control cohort record matches for each record in the treatment cohort is the increased variance from control cohort matches that are less similar to the treatment cohort record of interest. To reduce variance from less similar control cohort records, we apply a weighting technique in which control cohort records that are further away in similarity from their corresponding treatment cohort record are assigned a lower weight. Table A.13 shows the estimated effect of the increased autonomy policy in the matched data with weighted controls. Matched control records are ranked based on their similarity to the treatment cohort student of interest and assigned weights equal to the inverse of the rank position times the weight of the control record with the highest similarity. The weights of matched control records for each treatment cohort record sum to one.

[40] As a sensitivity check, we expand the set of matching criteria to include having a GPA above the eligibility threshold in grade 11, classroom size, classroom diversity (proxied by the standard deviation of prior-year GPA in the classroom), and the proportion of female peers in the classroom. Table A.15 presents estimates from a matching approach with the expanded matching criteria. The estimated effects of increased autonomy on excused absences and high-stakes exam performance for targeted students remain economically strong and statistically significant.

[41] The quality of enrolled degree takes values between 0 and 100, with 100 being the highest.

Table 5
Effect of the Increased Autonomy Policy on Targeted Students' Longer-Term Outcomes.

	(1) University Admission Score	(2) Quality of Enrolled Degree
Targeted in Treated Cohort	0.127***	2.109*
	(0.026)	(1.086)
Observations	9362	6897
Y Mean (Non-treated)	-0.57	38.30
Y St. Dev. (Non-treated)	0.83	22.71
Y Mean (Treated)	0.71	67.37
Y St. Dev. (Treated)	0.69	24.53
Track FE	✓	✓
School FE	✓	✓
Year FE	✓	✓

This table shows the estimated effect of increased autonomy on longer-term outcomes. The sample is restricted to students graduating between 2006 and 2007 in grade 12. Coefficient *Targeted in Treated Cohort*, [1], represents the effect of increased autonomy policy on targeted senior students in treated school year 2006-07 relative to non-targeted senior students in the same cohort. The university admission score is a weighted average of midterm and final exam performance in high-stakes subjects in grade 12. University admission scores are standardized at the year-track level. Quality of enrolled degree, which reflects the prestige/selectiveness of a degree, is calculated as a degree's ranking based on the average admission score of each university department first across students and second across the sample years, and takes values from 0 to 100 with 100 being the highest. All specifications control for student gender, year of birth indicators, prior-year GPA, class size, female share in the classroom, classroom academic diversity, as well as track, school, and year FE. Standard errors clustered at the school level are reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < .01$.

admission score of 0.13 standard deviations. Table 5 also shows that increased autonomy is associated with being admitted to university degree programs of higher quality/selectiveness—an improvement roughly equivalent to two percentiles in the distribution of degree quality.

Table A.16 and Table A.17 show the effect of increased autonomy policy on the longer-term outcomes by targeted students' prior-year GPA, and by classroom academic diversity, respectively. Targeted students across the prior performance distribution increased their university admission scores due to the increased autonomy policy. The quartile of top performing targeted students is significantly associated with university admission to departments of higher quality (Table A.16). We do not find heterogeneous longer-term effects effects of the increased autonomy policy by classroom academic diversity (Table A.17).

10. Mechanisms

The results presented so far show that the increased autonomy policy increases both student absences and student performance in high-stakes subjects. The estimated effect of the increased autonomy policy on targeted students' performance is large in comparison to the effect on targeted students' absences. We consider our estimated effects of the autonomy policy on performance to be composite of several mechanisms. A first mechanism through which increased autonomy could influence performance is because of the attendance-performance association, which may also differ across students (Liu et al., 2021). A second mechanism through which increased autonomy could affect performance is related to the effective class size. As students take more class absences due to the increased autonomy policy, the class size decreases and student performance increases. A third mechanism may be related to changes in peer characteristics due to the increased autonomy policy. As targeted students skip class more often, the effective peer characteristics may change, impacting student performance (Lavy et al., 2012). A fourth channel of influence of increased autonomy on performance is through teachers. As targeted students skip class more often because of the increased autonomy policy, instruction might become more effective (Blatchford and Russell, 2019; Wright et al., 2019). A fifth channel of influence of the increased autonomy policy on performance is related to student motivation from agency as suggested by the Self Determination Theory (Deci and Ryan, 2012; Ryan and Deci, 2017; 2020). Students targeted by the increased autonomy policy might feel they are being treated as responsible individuals, capable of making optimal decisions on their own, potentially feeling motivated to perform higher (Evans and Boucher, 2015; Hattie et al., 2020; Ryan and Deci, 2020). The likelihood of each mechanism depends on the strength of peer influences in the classroom and the association between student characteristics and instruction quality.

One might expect that the channels of influence of the autonomy policy on performance through the effective class size, effective peer characteristics, and the instruction quality might be more pronounced for non-targeted students than targeted. In particular, non-targeted students in classrooms with a higher percentage of targeted students might experience greater changes in effective class size, effective peer characteristics, and the instruction quality due to the autonomy policy. We empirically investigate this hypothesis. We group classrooms in three quantiles (i.e., terciles) based on the percentage of targeted students. We compare the differential impact of the autonomy policy on non-targeted students in the treated cohort in classrooms in the top and bottom tercile of the proportion of targeted students, relative to the middle tercile. Table A.18 presents our results. The effects of the autonomy policy on non-targeted students' absences or performance in classrooms in the top tercile of the proportion of targeted students are statistically indistinguishable from those in classrooms in the middle tercile. We do not find evidence of substantial aggregate performance effects associated with the mechanisms of effective class size, effective peer characteristics, and the instruction quality.

We perform further empirical investigations to understand the relative importance of the motivation channel and other channels relative to the absences channel through which the autonomy policy impacts targeted students' performance. Specifically, we re-produce our main results using the student performance outcomes and controlling for the change in absences associated with the increased autonomy policy. Table A.19 shows our results. We find that the estimated effect of the increased autonomy policy on performance in high-stakes subjects decreased by 36 percent and becomes statistically insignificant. This finding suggests that the absences channel may remain a substantial explanatory factor of the influence of the autonomy policy on performance even if additional mechanisms cannot be fully identified. At the same time, the motivation and encouragement students might feel when given agency and being treated as responsible individuals may also contribute to the magnitude of the estimated effect of the autonomy policy on targeted students' performance.

The institutional setting may also contribute to the magnitude of the estimated effects of the autonomy policy. Societal pressure for high performance in school may contribute to student stress and render opportunities for decompression appealing. Students in Greece often report unhappiness and lack of motivation (OECD, 2014; 2018). Stress may be more common among economically disadvantaged students. At the same time, students in Greece spend substantial amount of time attending after-school classes or private tutoring, as shown in Fig. A.12, even though it falls short of countries like Indonesia and South Korea (OECD, 2013; 2016; Liodaki and Liodakis, 2016).[42] This suggests that students who skip regular school may attend private educational facilities. The potentially high teaching efficacy of after-school classes or private tutors may lead to substantial test score gains with a limited sacrifice in regular school attendance.

One might expect the increased autonomy policy to have differential effect by student socio-economic background. We obtain income information at the school postcode level from the Ministry of Finance and empirically investigate this hypothesis. Table A.20 presents our results. We find comparable increased autonomy policy effects on targeted students' attendance and performance in neighborhoods with below-median or above-median income. At the same time, we find differential increased autonomy policy effects on non-targeted students by neighborhood income. Non-targeted students in below-median-income neighborhoods experience a larger increase in their excused absences and a stronger decline in their performance due to the autonomy policy compared with their counterparts in above-median-income neighborhoods. These results suggest that learning autonomy may be detrimental for lower-performing students in less affluent neighborhoods, which may potentially offer fewer out-of-school learning resources.

We collected supplementary data on the prevalence of learning and leisure facilities in each school's postcode. Table A.21 shows that neighborhoods with higher income also have more libraries and afternoon schools (i.e., learning facilities) as well as more parks and sports facilities (i.e., leisure facilities). We investigate differential effects of the increased autonomy policy in neighborhoods with above- and below-median number of learning and leisure facilities. Tables A.22 and A.23 present the estimated heterogeneous effects by learning and leisure neighborhood facilities, respectively. Table A.22 shows that targeted students in the treated cohort in neighborhoods with above-median learning facilities post slightly larger impacts of the increased autonomy policy on both excused absences and performance in high-stakes subjects than their counterparts in neighborhoods with below-median learning facilities. Non-targeted students in neighborhoods with above-median learning facilities experience a slightly smaller autonomy policy effect on excused absences than non-targeted students in neighborhoods with below-median learning facilities. Table A.23 shows that targeted students in the treated cohort in neighborhoods with above-median leisure facilities post slightly larger increased autonomy policy effects on both excused absences and performance in high-stakes subjects than targeted students in neighborhoods with below-median learning facilities.[43] Non-targeted students in neighborhoods with below-median learning facilities have a slightly larger autonomy policy effect on excused absences than non-targeted students in neighborhoods with above-median learning facilities. The increased autonomy policy increase non-targeted students' excused absences increase more in neighborhoods with below-median learning facilities than it did in neighborhoods with below-median leisure facilities. The results suggest that initiatives that provide learning autonomy to higher-performing students might be more likely to reduce school attendance of lower-performing students in neighborhoods with limited out-of-school learning resources.[44] Practices and policies aimed at increasing students' learning autonomy should be coupled with strategies to increase the supply of learning facilities, especially in economically disadvantaged neighborhoods.

[42] Supplementary education is popular in Asia, Southeast Asia, Southern Europe, as well as North America but to a lesser extent (CCL, 2007; Aurini et al., 2013; Bray, 2013).

[43] The pattern of estimated heterogeneous autonomy policy effects on targeted students' excused absences and high-stakes scores by learning facilities is similar to that of heterogeneous effects by leisure facilities. This implies substantial correlation in the geographical prevalence of learning and leisure facilities.

[44] Learning facilities are more prevalent in neighborhoods in urban settings as shown on Table A.21. Table A.24 shows heterogeneous autonomy policy effects by urban and non-urban setting and shows larger policy effect on targeted students' high-stakes exam performance in urban settings.

As discussed in this section, there are multiple mechanisms through which increased student autonomy might affect performance. Although we cannot fully distinguish them or exclude other channels of influence, our results suggest that that student attendance is an important driver of our results. At the same time, more research is required to understand exactly how student behavior changes when increased learning autonomy is provided and how autonomy influences non-targeted students or instruction quality.

11. Conclusion

In this paper, we estimate the effect of time autonomy on short- and longer-term school outcomes. We exploit an innovative nationwide policy in Greece that allowed grade 12 students with a prior-year GPA above 75% to skip 44 percent more classes with parental approval. Targeted students could decide how to spend their time with increased flexibility.

We employ a triple-differences identification strategy that compares student absences and performance trajectories between grades of targeted and non-targeted students in a control cohort and a treatment cohort. When students are offered discretion over their absences, their performance on exams that matter for university admission increases substantially. In contrast, targeted students' performance in exams that do not matter for university admission is unaffected. Targeted students also obtain a higher university admission score and are admitted to university degree programs of higher quality when they are given more autonomy.

The estimated effect of autonomy on targeted students' performance can be compared with effects documented by a large economics literature on schools' production function, learning time, and time allocation. In particular, the estimated effect of autonomy on targeted students' performance documented in this paper is equivalent to reducing class size by 0.23 standard deviations (or 8%) (Krueger, 2003);[45] having a teacher of quality of 1.4 standard deviations above average (Carrell and West, 2010); attending 50 additional days of schooling (Hanushek et al., 2012), an increase the class learning time between 1 to 2 hours per week (Lavy, 2015; Rivkin and Schiman, 2015), and an increase in school effectiveness, measured as frequent teacher feedback, data-driven instruction, frequent tutoring, increased instructional time, and focus on academic achievement, by 1 standard deviation (Dobbie and Fryer Jr, 2013; Fryer Jr, 2014). In contrast, targeted students' performance in exams that do not matter for university admission is unaffected. Targeted students also obtain a higher university admission score and are admitted to university degree programs of higher quality when they are given more autonomy.

Further investigation reveals that higher-performing targeted students and students in more academically diverse classrooms exercise greater autonomy when allowed to. This suggests that optimal level class absences may be higher for students who can learn independently or for those in settings that may be less conducive to learning.

Our results demonstrate that autonomy in the form of relaxed school attendance for higher-performing 12^{th}-graders may improve their performance and that, in the context of high-stakes exams, it may have significant long-term consequences on careers. More generally, the results highlight how compulsory class attendance can lead to allocative inefficiency. The allocation of student time to potentially less productive contexts (i.e., in academically diverse classrooms or lower-gravity classes) could result in poorer university and career placements and lower labor productivity. Our findings lend empirical support to the request by many families for schooling options that can provide flexibility without sacrificing academic performance, especially post-pandemic (Scott et al., 2020; Singer, 2021). Policymakers should consider providing more flexible schooling options that maximize overall educational productivity without increasing public spending. At the same time, it is important to underscore that the role of school extends beyond knowledge delivery. Schools play a key role in promoting social integration, the development of non-cognitive skills, and creating a safe environment for students.

Differential levels of learning productivity away from traditional schools during the COVID-19 pandemic are expected to widen educational gaps (Kuhfeld et al., 2020; Raymond et al., 2020; Pier et al., 2021). Our estimate of the impact of autonomy on student performance guides learning recovery programs and policies that aim to bring all students up to speed following the pandemic. Understanding the conditions under which students can learn effectively away from school widens the array of pandemic recovery policies and helps policymakers predict their success.

Declaration of Competing Interest

None.

Data Availability

The data that has been used is confidential.

[45] The average class size is roughly 23 students. The estimated autonomy effect on high-stakes exam performance is comparable to a reduction in class size by approximately two students.

Appendix A

Table A.1
Descriptive Statistics.

	Mean	Std. Dev.	Min.	Max.
Panel A: Targeted Criterion & Status				
Prior-year GPA	71.17	13.83	42.00	100.00
Targeted (1=Yes)	0.37	0.48	0.00	1.00
Panel B: Class Absences				
Total	64.22	43.57	1.00	953.00
Excused	29.38	23.88	0.00	174.00
Unexcused	34.84	34.44	0.00	953.00
Panel C: Performance				
High-stakes Subjects	0.01	0.88	-3.75	2.91
Low-stakes Subjects	0.01	0.85	-2.79	2.85
Panel D: University Admission Outcomes				
University Admission Score	0.00	1.00	-3.54	2.03
Quality of Enrolled Degree	55.68	27.75	0.12	99.78
Panel E: Class Characteristics				
Classroom Size	22.68	3.85	8.00	32.00
% of Females in Classroom	0.55	0.14	0.13	1.00
Classroom SD [1]	13.36	1.86	6.58	19.92
Classroom IQR [2]	10.61	2.78	3.25	21.13

Observations: 24,542. This table reports summary statistics for the prior-year GPA (the running variable for targeted status) and targeted status (Panel A), absences (Panel B), performance (Panel C), university outcomes (Panel D), and class characteristics (Panel E). Absences are measured in school periods. High- and low-stakes subjects performance is the average final exam score in the high- and low-stakes subjects, respectively. Scores are standardized at the school-grade-subject level. University admission scores are standardized at the year-track level. Quality of enrolled degree reflects prestige/selectiveness and is calculated as the ranking of the average university admission score first across admitted students and second across years. [1] Classroom SD refers to the standard deviation of the prior-year GPA across all students in the classroom. [2] Classroom IQR is the interquartile range of prior-year GPA in the classroom.

Table A.2
Dropout Likelihood in Grade 12 by Cohort and Targeted Status.

	Cohort 2006	Cohort 2007	Diff.	P-value
Targeted	0.021	0.021	0.000	0.959
Non-targeted	0.039	0.044	0.005	0.102
All	0.032	0.035	0.003	0.176

This table reports the mean likelihood of dropping out in grade 12 by cohort and targeted status. The p-value corresponds to a two-tail *t*-test of means between cohort 2006 (control) and cohort 2007 (treated).

Table A.3
Effect of the Increased Autonomy Policy on Targeted Students' Absences and Performance Using Clustering at Different Levels.

	Class Absences			Performance	
	Total	Excused	Unexcused	High-stakes Subjects	Low-stakes Subjects
Targeted in Treated Cohort [1]*relative to non-targeted*	3.907	3.093	0.815	0.073	-0.015
	(1.550)**	(1.050)***	(0.981)	(0.018)***	(0.020)
	[1.897]**	[1.368]**	[1.050]	[0.019]***	[0.021]
	2.174*	1.639*	1.128	0.019***	0.022
	⟨1.699⟩**	⟨1.243⟩**	⟨1.030⟩	⟨0.019⟩***	⟨0.022⟩
	((2.307))*	((1.720))*	((1.192))	((0.021))***	((0.026))
Non-targeted in Treated Cohort [2]*relative to control cohort*	1.511	2.807	-1.295	-0.020	0.015
	(1.811)	(1.172)**	(0.891)	(0.012)	(0.011)
	[1.867]	[1.308]**	[0.912]	[0.014]	[0.012]
	{1.998}	{1.431}*	{0.991}	{0.015}	{0.013}
	⟨1.389⟩	⟨0.933⟩***	⟨0.875⟩	⟨0.015⟩	⟨0.018⟩
	((1.382))	((0.926))***	((0.878))	((0.015))	((0.017))
Observations	24,542	24,542	24,542	24,542	24,542
Y Mean (Non-targeted)	68.22	29.41	38.81	-0.44	-0.44
Y St. Dev. (Non-targeted)	48.64	23.16	41.36	0.71	0.60
Y Mean (Targeted)	57.41	29.32	28.09	0.79	0.78
Y St. Dev. (Targeted)	32.08	25.06	14.84	0.53	0.62
Student FE	✓	✓	✓	✓	✓
P-value for H0: [1] + [2] = 0	0.01	0.00	0.60	0.00	0.97

Low-stakes subjects include general education Mathematics, History, and Physics. High-stakes subjects include general education Modern Greek, Ancient Greek, Latin for students in the Classics track, and Mathematics and Physics for students in the Science and IT tracks. Performance in high- and low-stakes subjects is standardized at the school-grade-cohort-subject level. Coefficient *Targeted in Treated Cohort*, [1], represents the effect of increased autonomy policy on targeted students in grade 12 in treated school year 2006-07 relative to non-targeted students in grade 12 in the same cohort (β_1 in specification (1)). Coefficient *Non-targeted in Treated Cohort*, [2], captures the effect of the increased autonomy policy on non-targeted students in grade 12 in school year 2006-07 relative to non-targeted students in the control school year 2005-06 (β_2 in specification (1)). We test the hypothesis that the full effect of the increased autonomy policy on targeted students in grade 12 in 2006-07 (relative to non-targeted students in the control cohort) is equal to zero. All specifications include student fixed effects, an indicator for being targeted in a given year, and classroom-level controls, such as classroom size, standard deviation of prior performance in the classroom, and the proportion of females in the classroom. Standard errors clustered at the school level (in parentheses), at school-year level (in square brackets), at school-year-grade (in curly brackets), at school-year-grade-class level (in angle brackets), and at school-year-grade-targeted level (in double parentheses). * $p < 0.1$; ** $p < 0.05$; *** $p < .01$.

Table A.4
Effect of the Increased Autonomy Policy on Targeted Students' Absences and Performance Using Wild Bootstrap Standard Errors.

	(1) Class Absences	(2)	(3)	(4) Performance	(5)
	Total	Excused	Unexcused	High-stakes Subjects	Low-stakes Subjects
	Estimates/(P-value)/[CI Lower Bound - CI Upper Bound]				
Targeted in Treated Cohort [1]*relative to non-targeted*	3.907	3.093	0.815	0.073	-0.015
	(0.013)	(0.005)	(0.413)	(0.000)	(0.480)
	[0.882, 6.986]	[0.905, 5.256]	[-1.187, 2.748]	[0.038, 0.108]	[-0.054, 0.025]
Non-targeted in Treated Cohort [2]*relative to control cohort*	1.511	2.807	-1.295	-0.020	0.015
	(0.409)	(0.024)	(0.162)	(0.112)	(0.172)
	[-2.099, 5.064]	[0.453, 5.305]	[-3.235, 0.496]	[-0.046, 0.005]	[-0.006, 0.037]
Observations	24,542	24,542	24,542	24,542	24,542

This table shows estimates from Table 2 of the main model, applying the bootstrapping procedure with asymptotic refinement as in Cameron et al. (2008). Standard errors clustered at the school level. P-values appear in parentheses, while confidence intervals appear in square brackets.

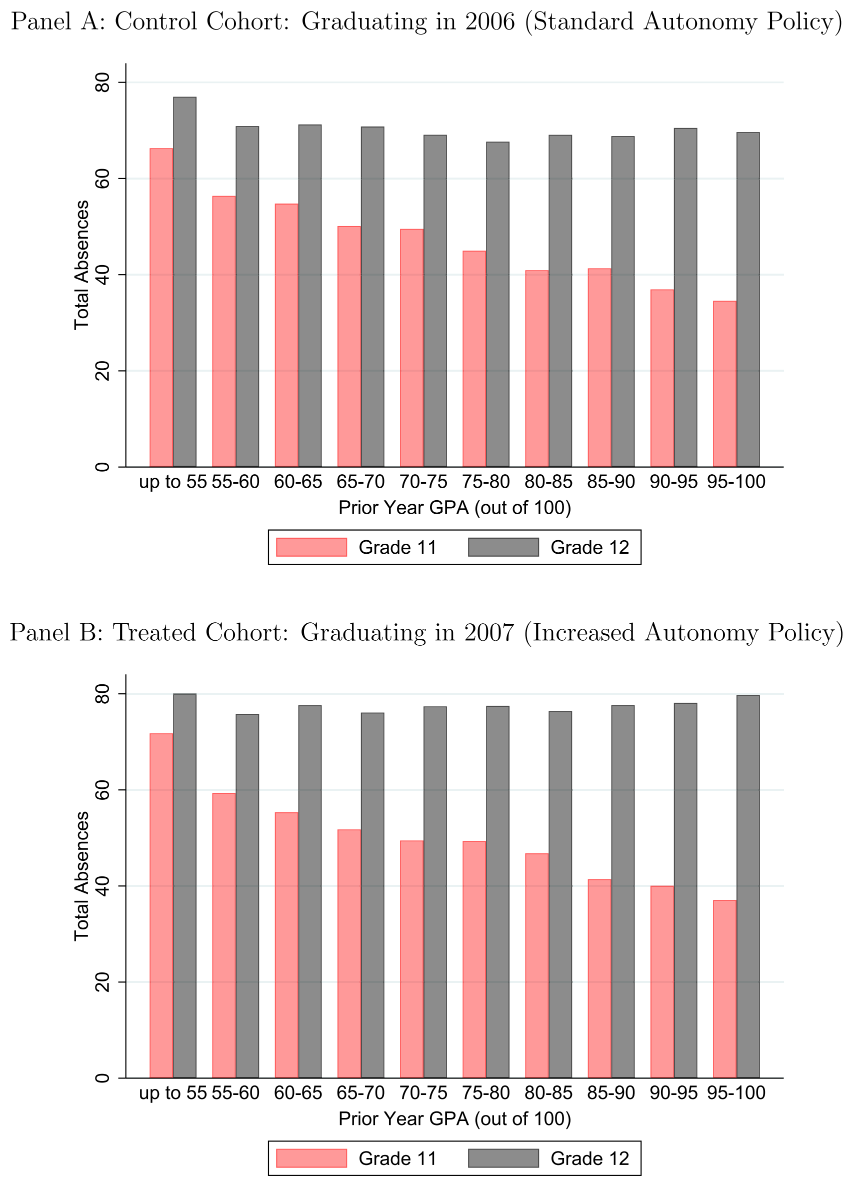


Fig. A.1. Total Absences Across Grades and Cohorts Panels A and B display average total absences in grades 11 and 12 in cohorts graduating in 2006 and 2007, respectively. Absences are measured in class periods..

Table A.5
Random Assignment of Students to Classrooms.

	Av. Excused Absences (1)	Av. Unexcused Absences (2)	Av. GPA (3)	GPA Diversity (4)	Targeted Share (5)	Av. Excused Absences of Targeted (6)	Av. Unexcused Absences of Targeted (7)	Av. GPA of Targeted (8)	Av. Excused Absences of Non-targeted (9)	Av. Unexcused Absences of Non-targeted (10)	Av. GPA of Non-targeted (11)
Class Number=2	0.011	1.739	-0.009	0.002	-0.015	-0.258	-0.274	-0.012	0.269	2.013	0.003
	(0.499)	(1.286)	(0.028)	(0.012)	(0.012)	(0.296)	(0.278)	(0.016)	(0.463)	(1.302)	(0.015)
Class Number=3	-0.633	0.864	0.005	0.015	-0.007	-0.349	0.024	0.008	-0.284	0.840	-0.003
	(0.474)	(1.435)	(0.030)	(0.015)	(0.014)	(0.263)	(0.285)	(0.017)	(0.462)	(1.458)	(0.017)
Class Number=4	-0.468	0.501	0.010	-0.011	0.001	-0.569	-0.157	-0.007	0.101	0.658	0.017
	(0.614)	(1.635)	(0.040)	(0.019)	(0.018)	(0.403)	(0.399)	(0.023)	(0.568)	(1.662)	(0.022)
Class Number=5	0.481	1.700	-0.022	-0.044	-0.028	0.065	-0.351	-0.035	0.416	2.051	0.013
	(0.848)	(1.711)	(0.065)	(0.031)	(0.035)	(0.597)	(0.695)	(0.043)	(0.922)	(2.090)	(0.031)
Observations	606	606	606	606	606	606	606	606	606	606	606
Y Mean	24.81	33.72	0.00	0.98	0.36	8.37	8.86	0.37	16.44	24.86	-0.36
Y Standard Deviation	14.39	14.24	0.25	0.12	0.15	7.68	4.89	0.16	9.91	14.11	0.17
School FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
F-stat. for joint significance	0.92	0.60	0.11	1.14	0.58	0.78	0.35	0.55	0.38	0.75	0.25
P-value for joint significance	0.456	0.664	0.978	0.341	0.679	0.538	0.846	0.703	0.823	0.562	0.908

The table shows results of the estimated effects of the classroom number on a series of classroom-level outcomes in grade 10, the instance of outcomes observation closest to classroom assignment when data are available. We regress average excused absences (column 1), average unexcused absences (column 2), average GPA (column 3), average GPA diversity (measured by standard deviation of prior-year GPA) (column 4), share of targeted (i.e., with a GPA higher than 75%) students (column 5) on classroom number. We also show results for average excused absences, unexcused absences, and GPA for targeted (columns 6–8) and non-targeted students (columns 9–11). Classroom number one is omitted as the reference group. Classroom is the unit of observation. Standard errors clustered at the school level are reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < .01$.

Table A.6
Effect of the Increased Autonomy Policy on Targeted Students' Absences and Performance Using Grade 10 Grade as Benchmark Grade.

	Class Absences			Performance	
	Total	Excused	Unexcused	High-stakes Subjects	Low-stakes Subjects
Targeted in Treated Cohort [1]*relative to non-targeted*	3.857**	3.111***	0.746	0.058***	-0.070
	(1.640)	(1.148)	(0.983)	(0.020)	(0.055)
Non-targeted in Treated Cohort [2]*relative to control cohort*	5.275***	4.054***	1.221	-0.035**	0.025
	(1.857)	(1.146)	(1.068)	(0.016)	(0.027)
Observations	22,652	22,652	22,652	22,652	22,652
Y Mean (Non-targeted)	65.31	27.39	37.92	-0.46	-0.34
Y St. Dev. (Non-targeted)	52.63	23.38	45.92	0.68	0.88
Y Mean (Targeted)	51.30	26.79	24.51	0.83	0.61
Y St. Dev. (Targeted)	32.97	25.16	13.89	0.51	0.59
Student FE	✓	✓	✓	✓	✓
P-value for H0: [1] + [2] = 0	0.00	0.00	0.03	0.11	0.22

Low-stakes subjects include general education Mathematics, History, and Physics. High-stakes subjects include general education Modern Greek, Ancient Greek, Latin for students in the Classics track, and Mathematics and Physics for students in the Science and IT tracks. Performance in high- and low-stakes subjects is standardized at the school-grade-cohort-subject level. The specifications compare the outcomes in grade 12th, compared to grade 10th.Coefficient *Targeted in Treated Cohort*, [1], represents the effect of increased autonomy policy on targeted students in grade 12 in treated school year 2006-07 relative to non-targeted students in grade 12 in the same cohort (β_1 in specification (1)). Coefficient *Non-targeted in Treated Cohort*, [2], captures the effect of the increased autonomy policy on non-targeted students in grade 12 in school year 2006-07 relative to non-targeted students in the control school year 2005-06 (β_2 in specification (1)). We test the hypothesis that the full effect of the increased autonomy policy on targeted students in grade 12 in 2006-07 (relative to non-targeted students in the control cohort) is equal to zero. All specifications include student fixed effects, an indicator for being targeted in a given year, and classroom-level controls, such as classroom size, standard deviation of prior performance in the classroom, and the proportion of females in the classroom. Standard errors clustered at the school level are reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < .01$.

Table A.7
Effect of the Increased Autonomy Policy on Targeted Students' Performance Using Different Levels of Standardization.

	School-Grade-Subject		Year-Subject-Grade		Subject-Grade	
	High-stakes Subjects	Low-stakes Subjects	High-stakes Subjects	Low-stakes Subjects	High-stakes Subjects	Low-stakes Subjects
Targeted in Treated Cohort [1]*relative to non-targeted*	0.073***	-0.015	0.060**	-0.041	0.154***	-0.016
	(0.018)	(0.020)	(0.024)	(0.031)	(0.024)	(0.031)
Non-targeted in Treated Cohort [2]*relative to control cohort*	-0.020	0.015	-0.026	0.008	-0.060**	0.031
	(0.012)	(0.011)	(0.027)	(0.036)	(0.027)	(0.036)
Observations	24,542	24,542	24,542	24,542	24,542	24,542
Y Mean (Non-targeted)	-0.44	-0.44	-0.44	-0.43	-0.44	-0.43
Y St. Dev. (Non-targeted)	0.71	0.60	0.71	0.60	0.71	0.60
Y Mean (Targeted)	0.79	0.78	0.80	0.76	0.80	0.76
Y St. Dev. (Targeted)	0.53	0.62	0.52	0.60	0.52	0.60
Student FE	✓	✓	✓	✓	✓	✓
P-value for H0: [1] + [2] = 0	0.00	0.97	0.17	0.43	0.00	0.71

Low-stakes subjects include general education Mathematics, History, and Physics. High-stakes subjects include general education Modern Greek, Ancient Greek, Latin for students in the Classics track, and Mathematics and Physics for students in the Science and IT tracks. Performance in high- and low-stakes subjects in columns 1–2 are standardized at the school-year-subject-grade level. Performance in high- and low-stakes subjects in columns 3–4 are standardized at the year-grade-subject level.Performance in high- and low-stakes subjects in columns 5–6 are standardized at the grade-subject level. Coefficient *Targeted in Treated Cohort*, [1], represents the effect of increased autonomy policy on targeted students in grade 12 in treated school year 2006-07 relative to non-targeted students in grade 12 in the same cohort (β_1 in specification (1)). Coefficient *Non-targeted in Treated Cohort*, [2], captures the effect of the increased autonomy policy on non-targeted students in grade 12 in school year 2006-07 relative to non-targeted students in the control school year 2005-06 (β_2 in specification (1)). We test the hypothesis that the full effect of the increased autonomy policy on targeted students in grade 12 in 2006-07 (relative to non-targeted students in the control cohort) is equal to zero. All specifications include student fixed effects, an indicator for being targeted in a given year, and classroom-level controls, such as classroom size, standard deviation of prior performance in the classroom, and the proportion of females in the classroom. Standard errors clustered at the school level are reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < .01$.

Table A.8
Student Characteristics in Classrooms of Different Academic Diversity (Standard Deviation of Prior Performance).

	Targeted Students in Treated Cohort Prior to Autonomy Policy				Treated Cohort Prior to Autonomy Policy	
GPA	(1) Total Absences	(2) Excused Absences	(3) Unexcused Absences	(4) Class Size	(5) Prop. of Targeted	(6)
Class SD Top Quartile (Most Heterogeneous, relative to Bottom Quartile)	3.716***	2.184	2.760***	-0.576	1.405***	0.029***
	(0.469)	(1.426)	(1.021)	(0.769)	(0.133)	(0.004)
Class SD Second Quartile (relative to Bottom Quartile)	2.312***	0.937	2.064*	-1.127	1.652***	0.016***
	(0.501)	(1.524)	(1.092)	(0.822)	(0.141)	(0.005)
Class SD Third Quartile (relative to Bottom Quartile)	1.483***	0.269	1.115	-0.846	0.703***	0.027***
	(0.491)	(1.495)	(1.070)	(0.806)	(0.139)	(0.005)
Class SD Bottom Quartile (Most Homogeneous, Benchmark)	82.033***	43.505***	16.375***	27.130***	21.952***	0.357***
	(0.343)	(1.043)	(0.747)	(0.563)	(0.095)	(0.003)
Obs.	2412	2414	2414	2414	6435	6435

This table reports OLS estimates from specifications that include the four indicator variables for the different quartiles of classroom standard deviation. Columns 1–4 report the results for students graduating in 2007 (treated cohort) with a GPA greater than 75% in Grade 11, prior to the increased autonomy policy (i.e., targeted students). Columns 5–6 include all students in 2007 (treated cohort) in Grade 11, prior to the increased autonomy policy. * $p < 0.1$; ** $p < 0.05$; *** $p < .01$.

Table A.9
Heterogeneous Effects of Increased Autonomy Policy on Targeted Students by Classroom Diversity (Standard Deviation of Prior Performance).

	Class Absences			Performance	
	Total	Excused	Unexcused	High-stakes Subjects	Low-stakes Subjects
Targeted in Treated	6.431***	4.864***	1.567	0.047	-0.015
Cohort × Class SD Top Quartile	(2.290)	(1.651)	(1.470)	(0.029)	(0.030)
Targeted in Treated	6.560***	6.298***	0.263	0.060**	-0.026
Cohort × Class SD Second Quartile	(2.247)	(1.411)	(1.645)	(0.028)	(0.033)
Targeted in Treated	3.673*	1.237	2.436**	0.111***	-0.031
Cohort × Class SD Third Quartile	(2.152)	(1.715)	(1.142)	(0.026)	(0.031)
Targeted in Treated	-0.497	0.544	-1.041	0.069**	0.009
Cohort × Class SD Bottom Quartile	(3.057)	(2.074)	(1.545)	(0.033)	(0.034)
Non-targeted in	3.768*	5.236***	-1.468	-0.033	0.038
Treated Cohort × Class SD Top Quartile	(2.179)	(1.767)	(0.990)	(0.026)	(0.027)
Non-targeted in	-1.010	-0.168	-0.842	-0.000	0.025
Treated Cohort × Class SD Second Quartile	(2.845)	(1.831)	(1.852)	(0.022)	(0.025)
Non-targeted in	-0.011	1.939	-1.949*	-0.051**	0.019
Treated Cohort × Class SD Third Quartile	(2.488)	(1.725)	(1.131)	(0.021)	(0.028)
Non-targeted in	1.535	2.843	-1.308	0.008	-0.012
Treated Cohort × Class SD Bottom Quartile	(3.406)	(2.154)	(1.619)	(0.025)	(0.028)
Observations	24,542	24,542	24,542	24,542	24,542
Y Mean (Non-treated)	68.22	29.41	38.81	-0.44	-0.44
Y St. Dev. (Non-treated)	48.64	23.16	41.36	0.71	0.60
Y Mean (Treated)	57.42	29.32	28.09	0.79	0.78
Y St. Dev. (Treated)	32.08	25.06	14.84	0.53	0.62
Student FE	✓	✓	✓	✓	✓

Low-stakes subjects include general education Mathematics, History, and Physics. High-stakes subjects include general education Modern Greek, Ancient Greek, Latin for students in the Classics track, and Mathematics and Physics for students in the Science and IT tracks. Performance in high- and low-stakes subjects is standardized at the school-grade-cohort-subject level. Coefficients of the interactions of *Targeted in Treated Cohort* with the indicators *Class SD Top Quartile—Bottom Quartile* represent the effect of increased autonomy policy on targeted students in the treatment year who are in classrooms of different diversity, as captured by quartiles of standard deviation of prior performance in the classroom relative to non-targeted students in the same year. Coefficient *Non-targeted in Treated Cohort*, captures the effect of the increased autonomy policy on non-targeted students in grade 12 in school year 2006-07 relative to non-targeted students in the control school year 2005-06. All specifications include student fixed effects, an indicator for being targeted in a given year, and classroom-level controls. Classroom controls include classroom size, quantiles of standard deviation of prior performance in the classroom, and the proportion of females in the classroom. Standard errors clustered at the school level are reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < .01$.

Table A.10
Heterogeneous Effects of the Increased Autonomy Policy on Targeted Students by Classroom Diversity (Standard Deviation of Prior Performance) when Controlling for Prior Performance.

	Class Absences			Performance	
	Total	Excused	Unexcused	High-stakes Subjects	Low-stakes Subjects
Targeted in Treated Cohort × Class SD Top Quartile	7.528**	6.737***	0.791	0.028	0.003
	(2.966)	(2.385)	(1.217)	(0.023)	(0.027)
Targeted in Treated Cohort × Class SD Second Quartile	3.739	3.749*	-0.010	0.072***	-0.007
	(2.792)	(2.050)	(1.221)	(0.025)	(0.033)
Targeted in Treated Cohort × Class SD Third Quartile	0.925	-0.201	1.127	0.083***	-0.026
	(2.561)	(1.810)	(1.208)	(0.025)	(0.030)
Targeted in Treated Cohort × Class SD Bottom Quartile	-0.902	0.427	-1.328	0.079***	-0.026
	(3.221)	(2.488)	(1.204)	(0.028)	(0.029)
Non-targeted in Treated Cohort	1.848	2.684**	-0.836	-0.011	0.019
	(1.920)	(1.235)	(0.932)	(0.014)	(0.013)
Observations	21,948	21,948	21,948	21,948	21,948
Y Mean (Non-treated)	68.22	29.41	38.81	-0.44	-0.44
Y St. Dev. (Non-treated)	48.64	23.16	41.36	0.71	0.60
Y Mean (Treated)	57.42	29.32	28.09	0.79	0.78
Y St. Dev. (Treated)	32.08	25.06	14.84	0.53	0.62
Student FE	✓	✓	✓	✓	✓
Control for Prior-year GPA	✓	✓	✓	✓	✓

Low-stakes subjects include general education Mathematics, History, and Physics. High-stakes subjects include general education Modern Greek, Ancient Greek, Latin for students in the Classics track, and Mathematics and Physics for students in the Science and IT tracks. Performance in high- and low-stakes subjects is standardized at the school-grade-cohort-subject level. Coefficients of the interactions of *Targeted in Treated Cohort* with the indicators *Class SD Top Quartile—Bottom Quartile* represent the effect of the increased autonomy policy on targeted students in the treatment year who are in classrooms of different diversity, as captured by quartiles of standard deviation of prior performance in the classroom relative to non-targeted students in the same year (β_1^q in specification (1)). Coefficient *Non-targeted in Treated Cohort*, captures the effect of the increased autonomy policy on non-targeted students in grade 12 in school year 2006-07 relative to non-targeted students in the control school year 2005-06 (β_2 in specification (2)). All specifications include student fixed effects, prior-year GPA, an indicator for being targeted in a given year, and classroom-level controls. Classroom controls include classroom size, quantiles of standard deviation of prior performance in the classroom, and the proportion of females in the classroom. Standard errors clustered at the school level are reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < .01$.

Table A.11
Heterogeneous Effects of the Increased Autonomy Policy on Targeted Students by Classroom Diversity (Interquartile Range).

	Class Absences			Performance	
	Total	Excused	Unexcused	High-stakes Subjects	Low-stakes Subjects
Targeted in Treated Cohort × Class IQR Top Quartile	5.957**	6.277***	-0.321	0.035	0.038
	(2.673)	(2.099)	(1.318)	(0.023)	(0.029)
Targeted in Treated Cohort × Class IQR Second Quartile	6.041***	3.710**	2.332*	0.079***	-0.049*
	(2.241)	(1.540)	(1.209)	(0.020)	(0.029)
Targeted in Treated Cohort × Class IQR Third Quartile	5.042*	3.271	1.771	0.094***	-0.002
	(2.832)	(2.001)	(1.454)	(0.026)	(0.028)
Targeted in Treated Cohort × Class IQR Bottom Quartile	-1.498	-0.700	-0.798	0.076***	-0.039
	(3.121)	(2.432)	(1.368)	(0.028)	(0.029)
Non-targeted in Treated Cohort	1.701	2.970**	-1.269	-0.022*	0.015
	(1.802)	(1.145)	(0.901)	(0.012)	(0.011)
Observations	24,542	24,542	24,542	24,542	24,542
Y Mean (Non-treated)	68.22	29.41	38.81	-0.44	-0.44
Y St. Dev. (Non-treated)	48.64	23.16	41.36	0.71	0.60
Y Mean (Treated)	57.42	29.32	28.09	0.79	0.78
Y St. Dev. (Treated)	32.08	25.06	14.84	0.53	0.62
Student FE	✓	✓	✓	✓	✓

Low-stakes subjects include general education Mathematics, History, and Physics. High-stakes subjects include general education Modern Greek, Ancient Greek, Latin for students in the Classics track, and Mathematics and Physics for students in the Science and IT tracks. Performance in high- and low-stakes subjects is standardized at the school-grade-cohort-subject level. Coefficients of the interactions of *Targeted in Treated Cohort* with the indicators *Class IQR Top Quartile—Bottom Quartile* represent the effect of the increased autonomy policy on targeted students in the treatment year who are in classrooms of different diversity as captured by quartiles of the interquartile range of prior performance in the classroom relative to non-targeted students in the same year (β_1^q in specification (1)). Coefficient *Non-targeted in Treated Cohort* captures the effect of the increased autonomy policy on non-targeted students in grade 12 in school year 2006-07 relative to non-targeted students in the control school year 2005-06 (β_2 in specification (2)). All specifications include student fixed effects, an indicator for being targeted in a given year, and classroom-level controls. Classroom controls include classroom size, quantiles of interquartile range of prior performance in the classroom, and the proportion of females in the classroom. Standard errors clustered at the school level are reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < .01$.

Table A.12
Descriptive Characteristics of Matched Samples.

	Full Sample		Matched Sample 1 (Table A.14)			Matched Sample 2 (Table A.13)			Matched Sample 3 (Table A.15)		
(lr)2–3 (lr)4–6 (lr)7–9 (lr)10–12	Cohort 2006	Cohort 2007	Matched On	Cohort 2006	Cohort 2007	Matched On	Cohort 2006	Cohort 2007	Matched On	Cohort 2006	Cohort 2007
Grade 11											
Gender	0.571	0.568	✓	0.579	0.579	✓	0.579	0.579	✓	0.582	0.582
	(.495)	(.495)		(.494)	(.494)		(.494)	(.494)		(.493)	(.493)
Age	16.86	16.893	✓	16.883	16.883	✓	16.883	16.883	✓	16.922	16.922
	(.548)	(.513)		(.371)	(.371)		(.371)	(.371)		(.299)	(.299)
Unexcused Absences	29.316	30.202	✓	32.234	32.699	✓	32.261	32.699	✓	33.985	34.468
	(13.738)	(13.766)		(11.924)	(12.359)		(12.002)	(12.359)		(11.511)	(11.710)
Excused Absences	18.258	19.933	✓	19.680	20.210	✓	19.683	20.21	✓	19.770	20.297
	(18.554)	(19.377)		(17.684)	(18.462)		(17.710)	(18.462)		(17.904)	(18.565)
GPA	0.014	0.042	✓	-0.029	-0.029	✓	-0.029	-0.029	✓	-0.142	-0.143
	(.994)	(.982)		(.949)	(.949)		(.948)	(.949)		(.971)	(.972)
Targeted (1=yes)	0.403	0.405		0.387	0.386		0.391	0.386	✓	0.339	0.339
	(.491)	(.491)		(.434)	(.487)		(.440)	(.487)		(.474)	(.474)
Grade 12											
Targeted (1=yes)	0.382	0.400	✓	0.368	0.368	✓	0.368	0.368	✓	0.323	0.323
	(.486)	(.490)		(.482)	(.482)		(.482)	(.482)		(.468)	(.468)
Classroom Size	22.619	22.23		22.766	22.267		22.75	22.267	✓	23.106	23.28
	(3.622)	(3.877)		(2.166)	(3.863)		(2.32)	(3.863)		(2.512)	(3.065)
% of Females in Classroom	0.554	0.551		0.548	0.552		0.547	0.552	✓	0.552	0.564
	(.131)	(.141)		(.084)	(.140)		(.086)	(.140)		(.094)	(.110)
Classroom SD [1]	2.774	2.731		2.761	2.732		2.767	2.732	✓	2.783	2.779
	(.353)	(.379)		(.225)	(.374)		(.256)	(.374)		(.261)	(.310)
Observations	5843	5720		4349	4349		4349	4349		2052	2052
Sampling Weights			Equal			Distance- Inverse			Equal		

This table reports mean values for the characteristics used in matching. Standard deviations are reported in parentheses. [1] Classroom SD refers to the standard deviation of the prior-year GPA across all students in the classroom.

Table A.13
Effect of the Increased Autonomy Policy on Targeted Students' Absences and Performance Using Matching with Weighted Controls.

	Class Absences			Performance	
	Total	Excused	Unexcused	High-stakes Subjects	Low-stakes Subjects
Targeted in Treated Cohort [1]*relative to non-targeted*	5.596***	5.299***	0.296	0.180***	-0.085***
	(1.879)	(1.606)	(1.049)	(0.044)	(0.031)
Non-targeted in Treated Cohort [2]*relative to control cohort*	3.703**	4.970***	-1.267*	-0.016	0.027
	(1.532)	(1.119)	(0.767)	(0.036)	(0.017)
Observations	10,426	10,426	10,426	10,426	10,426
Y Mean (Non-treated)	70.59	32.03	38.56	-0.36	-0.51
Y St. Dev. (Non-treated)	26.34	21.06	11.41	0.61	0.49
Y Mean (Treated)	63.83	32.50	31.33	0.70	0.68
Y St. Dev. (Treated)	28.82	24.00	11.09	0.57	0.57
Student FE	✓	✓	✓	✓	✓
P-value for H0: [1] + [2] = 0	0.00	0.00	0.31	0.00	0.03

Matching criteria: gender, age, grade 11 absences (excused and unexcused), grade 11 GPA, and grade 11 targeted status. Matched control records are ranked based on their similarity to the treatment cohort student of interest and assigned weights equal to the inverse of the rank position times the weight of the control record with the highest similarity. The weights of matched control records for each treatment cohort record sum up to one. Low-stakes subjects include general education Mathematics, History, and Physics. High-stakes subjects include general education Modern Greek, Ancient Greek, Latin for students in the Classics track, and Mathematics and Physics for students in the Science and IT tracks. Performance in high- and low-stakes subjects is standardized at the school-grade-cohort-subject level. Coefficient *Targeted in Treated Cohort*, [1], represents the effect of the increased autonomy policy on targeted students in grade 12 in treated school year 2006-07 relative to non-targeted students in grade 12 in the same cohort (β_1 in specification (1)). Coefficient *Non-targeted in Treated Cohort*, [2], captures the effect of the increased autonomy policy on non-targeted students in grade 12 in school year 2006-07 relative to non-targeted students in the control school year 2005-06 (β_2 in specification (1)). We test the hypothesis that the full effect of the increased autonomy policy on targeted students in grade 12 in 2006-07 (relative to non-targeted students in the control cohort) is equal to zero. All specifications include student fixed effects, an indicator for being targeted in a given year, and classroom-level controls, such as classroom size, standard deviation of prior performance in the classroom, and the proportion of females in the classroom. Standard errors clustered at the school level are reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < .01$.

Table A.14
Effect of the Increased Autonomy Policy on Targeted Students' Absences and Performance Using Matching.

	Class Absences			Performance	
	Total	Excused	Unexcused	High-stakes Subjects	Low-stakes Subjects
Targeted in Treated Cohort [1]*relative to non-targeted*	3.428**	3.762***	-0.334	0.151***	-0.034
	(1.635)	(1.431)	(0.898)	(0.036)	(0.026)
Non-targeted in Treated Cohort [2]*relative to control cohort*	4.325***	4.467***	-0.142	-0.018	0.002
	(1.294)	(0.935)	(0.653)	(0.033)	(0.013)
Observations	11,366	11,366	11,366	11,366	11,366
Y Mean (Non-treated)	70.49	32.02	38.47	-0.37	-0.51
Y St. Dev. (Non-treated)	25.81	20.85	10.32	0.60	0.48
Y Mean (Treated)	63.85	32.49	31.36	0.71	0.68
Y St. Dev. (Treated)	28.53	23.72	11.02	0.56	0.57
Student FE	✓	✓	✓	✓	✓
P-value for H0: [1] + [2] = 0	0.00	0.00	0.57	0.00	0.15

Matching criteria: gender, age, grade 11 absences (excused and unexcused), grade 11 GPA, and grade 11 targeted status. Matched control records for the same treated cohort record carry equal weight. Low-stakes subjects include general education Mathematics, History, and Physics. High-stakes subjects include general education Modern Greek, Ancient Greek, Latin for students in the Classics track, and Mathematics and Physics for students in the Science and IT tracks. Performance in high- and low-stakes subjects is standardized at the school-grade-cohort-subject level. Coefficient *Targeted in Treated Cohort*, [1], represents the effect of increased autonomy policy on targeted students in grade 12 in treated school year 2006-07 relative to non-targeted students in grade 12 in the same cohort (β_1 in specification (1)). Coefficient *Non-targeted in Treated Cohort*, [2], captures the effect of the increased autonomy policy on non-targeted students in grade 12 in school year 2006-07 relative to non-targeted students in the control school year 2005-06 (β_2 in specification (1)). We test the hypothesis that the full effect of the increased autonomy policy on targeted students in grade 12 in 2006-07 (relative to non-targeted students in the control cohort) is equal to zero. All specifications include student fixed effects, an indicator for being targeted in a given year, and classroom-level controls, such as classroom size, standard deviation of prior performance in the classroom, and the proportion of females in the classroom. Standard errors clustered at the school level are reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < .01$.

Table A.15
Effect of the Increased Autonomy Policy on Targeted Students' Absences and Performance Using Matching with Additional Criteria.

	Class Absences			Performance	
	Total	Excused	Unexcused	High-stakes Subjects	Low-stakes Subjects
Targeted in Treated Cohort [1]*relative to non-targeted*	5.894**	6.259***	-0.364	0.118**	-0.108**
	(2.378)	(2.127)	(1.812)	(0.059)	(0.051)
Non-targeted in Treated Cohort [2]*relative to control cohort*	3.217	3.921***	-0.703	-0.016	-0.010
	(1.960)	(1.373)	(1.044)	(0.045)	(0.028)
Observations	4690	4690	4690	4690	4690
Y Mean (Non-treated)	72.88	32.91	39.97	-0.39	-0.55
Y St. Dev. (Non-treated)	25.87	21.36	9.50	0.61	0.47
Y Mean (Treated)	62.96	31.69	31.27	0.77	0.74
Y St. Dev. (Treated)	29.19	24.65	10.67	0.56	0.56
Student FE	✓	✓	✓	✓	✓
P-value for H0: [1] + [2] = 0	0.00	0.00	0.55	0.05	0.00

Matching criteria: gender, age, grade 11 absences (excused and unexcused), grade 11 GPA, grade 11 targeted status, grade 12 targeted status, and grade 12 classroom characteristics (size, diversity, female share). Matched control records for the same treated cohort record carry equal weight. Low-stakes subjects include general education Mathematics, History, and Physics. High-stakes subjects include general education Modern Greek, Ancient Greek, Latin for students in the Classics track, and Mathematics and Physics for students in the Science and IT tracks. Performance in high- and low-stakes subjects is standardized at the school-grade-cohort-subject level. Coefficient *Targeted in Treated Cohort*, [1], represents the effect of the increased autonomy policy on targeted students in grade 12 in treated school year 2006-07 relative to non-targeted students in grade 12 in the same cohort (β_1 in specification (1)). Coefficient *Non-targeted in Treated Cohort*, [2], captures the effect of the increased autonomy policy on non-targeted students in grade 12 in school year 2006-07 relative to non-targeted students in the control school year 2005-06 (β_2 in specification (1)). We test the hypothesis that the full effect of the increased autonomy policy on targeted students in grade 12 in 2006-07 (relative to non-targeted students in the control cohort) is equal to zero. All specifications include student fixed effects, an indicator for being targeted in a given year, and classroom-level controls, such as classroom size, standard deviation of prior performance in the classroom, and the proportion of females in the classroom. Standard errors clustered at the school level are reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < .01$.

Table A.16
Heterogeneous Effects of the Increased Autonomy Policy on Targeted Students' Longer-Term Outcomes by Prior Performance.

	(1) University Admission Score	(2) Quality of Enrolled Degree
Targeted in Treated Cohort x Prior Performance Bottom Quartile	0.153***	-0.257
	(0.036)	(1.491)
Targeted in Treated Cohort x Prior Performance Third Quartile	0.111***	2.112
	(0.033)	(1.366)
Targeted in Treated Cohort x Prior Performance Second Quartile	0.085***	0.546
	(0.031)	(1.329)
Targeted in Treated Cohort x Prior Performance Top Quartile	0.165***	6.607***
	(0.029)	(1.225)
Observations	9362	6897
Y Mean (Non-treated)	-0.57	38.30
Y St. Dev. (Non-treated)	0.83	22.71
Y Mean (Treated)	0.71	67.37
Y St. Dev. (Treated)	0.69	24.53
Track FE	✓	✓
School FE	✓	✓
Year FE	✓	✓

This table shows the estimated effect of increased autonomy on longer-term outcomes, by students prior performance. The sample is restricted to students graduating between 2006 and 2007 in grade 12. The coefficient *Targeted in Treated Cohort*, is interacted with four dummies for the quartile of prior performance of targeted students. The coefficients represent the effect of increased autonomy policy on targeted senior students, in that particular quartile of prior performance, in treated school year 2006-07, relative to non-targeted senior students in the same cohort. The university admission score is a weighted average of midterm and final exam performance in high-stakes subjects in grade 12. University admission scores are standardized at the year-track level. Quality of enrolled degree, which reflects the prestige/selectiveness of a degree, is calculated as a degree's ranking based on the average admission score of each university department first across students and second across the sample years, and takes values from 0 to 100 with 100 being the highest. All specifications control for student gender, year of birth indicators, prior-year GPA, class size, female share in the classroom, classroom academic diversity, as well as track, school, and year FE. Standard errors clustered at the school level are reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < .01$.

Table A.17
Heterogeneous Effects of the Increased Autonomy Policy on Targeted Students' Longer-Term Outcomes by Classroom Diversity (Standard Deviation of Prior Performance).

	(1) University Admission Score	(2) Quality of Enrolled Degree
Targeted in Treated Cohort x Class SD Top Quartile	0.110***	1.871
	(0.033)	(1.279)
Targeted in Treated Cohort x Class SD Second Quartile	0.151***	2.144
	(0.035)	(1.312)
Targeted in Treated Cohort x Class SD Third Quartile	0.131***	2.196
	(0.037)	(1.589)
Targeted in Treated Cohort x Class SD Bottom Quartile	0.119***	2.213
	(0.038)	(1.665)
Observations	9362	6897
Y Mean (Non-treated)	-0.57	38.30
Y St. Dev. (Non-treated)	0.83	22.71
Y Mean (Treated)	0.71	67.37
Y St. Dev. (Treated)	0.69	24.53
Track FE	✓	✓
School FE	✓	✓
Year FE	✓	✓

This table shows the estimated effect of increased autonomy on longer-term outcomes, by classroom diversity. Classroom diversity is measured as quartiles of standard deviation of prior performance in the classroom. The sample is restricted to students graduating between 2006 and 2007 in grade 12. Coefficients of the interactions of *Targeted in Treated Cohort* with the indicators Class SD Top QuartileBottom Quartile represent the effect of increased autonomy policy on targeted senior students, in any given quartile of class diversity, in treated school year 2006-07, relative to non-targeted senior students in the same cohort. The university admission score is a weighted average of midterm and final exam performance in high-stakes subjects in grade 12. University admission scores are standardized at the year-track level. Quality of enrolled degree, which reflects the prestige/selectiveness of a degree, is calculated as a degree's ranking based on the average admission score of each university department first across students and second across the sample years, and takes values from 0 to 100 with 100 being the highest. All specifications control for student gender, year of birth indicators, prior-year GPA, class size, female share in the classroom, classroom academic diversity, as well as track, school, and year FE. Standard errors clustered at the school level are reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < .01$.

Table A.18
Effect of the Increased Autonomy Policy on Non-targeted Students' Absences and Performance by the Proportion of Targeted Students in Classroom.

	Class Absences			Performance	
	Total	Excused	Unexcused	High-stakes Subjects	Low-stakes Subjects
Non-targeted in Treated Cohort × *Prop. Targeted Top Tercile (Highest)*	2.626	1.142	1.484	-0.040	0.043
	(4.290)	(3.008)	(2.576)	(0.049)	(0.053)
Non-targeted in Treated Cohort × *Prop. Targeted Bottom Tercile (Lowest)*	-0.059	0.871	-0.931	-0.071*	0.011
	(4.042)	(2.816)	(2.440)	(0.038)	(0.046)
Observations	14,024	14,024	14,024	14,024	14,024
Y Mean (Non-targeted)	68.22	29.41	38.81	-0.44	-0.44
Y St. Dev. (Non-targeted)	48.64	23.16	41.36	0.71	0.60
Student FE	✓	✓	✓	✓	✓

Low-stakes subjects include general education Mathematics, History, and Physics. High-stakes subjects include general education Modern Greek, Ancient Greek, Latin for students in the Classics track, and Mathematics and Physics for students in the Science and IT tracks. Performance in high- and low-stakes subjects is standardized at the school-grade-cohort-subject level. The average percentage of targeted students in the classroom in the top and bottom tercile is 23 and 52 percent, respectively. The average percentage of targeted students in classrooms in the middle tercile (omitted) is 36 percent. All specifications include student fixed effects, an indicator for being targeted in a given year, and classroom-level controls, such as classroom size, standard deviation of prior performance in the classroom, and the proportion of females in the classroom. Standard errors clustered at the school level are reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < .01$.

Table A.19
Effect of the Increased Autonomy Policy on Targeted Students' Absences and Performance when Controlling for Changes in Absences.

	Main Results		Controlling for Changes in Absences	
	(1) High-stakes Subjects	(2) Low-stakes Subjects	(3) High-stakes Subjects	(4) Low-stakes Subjects
Targeted in Treated Cohort [1]*relative to non-targeted*	0.073***	-0.015	0.047	0.033
	(0.018)	(0.020)	(0.045)	(0.048)
Non-targeted in Treated Cohort [2]*relative to control cohort*	-0.020	0.015	-0.060**	0.013
	(0.012)	(0.011)	(0.030)	(0.028)
Observations	24,542	24,542	24,542	24,542
Y Mean (Non-targeted)	-0.44	-0.44	-0.44	-0.44
Y St. Dev. (Non-targeted)	0.71	0.60	0.71	0.60
Y Mean (Targeted)	0.79	0.78	0.79	0.78
Y St. Dev. (Targeted)	0.53	0.62	0.53	0.62
Student FE	✓	✓	✓	✓
P-value for H0: [1] + [2] = 0	0.00	0.97	0.72	0.35

Low-stakes subjects include general education Mathematics, History, and Physics. High-stakes subjects include general education Modern Greek, Ancient Greek, Latin for students in the Classics track, and Mathematics and Physics for students in the Science and IT tracks. Performance in high- and low-stakes subjects is standardized at the school-grade-cohort-subject level. The specifications compare the outcomes in grade 12th, compared to grade 10th.Coefficient *Targeted in Treated Cohort*, [1], represents the effect of increased autonomy policy on targeted students in grade 12 in treated school year 2006-07 relative to non-targeted students in grade 12 in the same cohort (β_1 in specification (1)). Coefficient *Non-targeted in Treated Cohort*, [2], captures the effect of the increased autonomy policy on non-targeted students in grade 12 in school year 2006-07 relative to non-targeted students in the control school year 2005-06 (β_2 in specification (1)). We test the hypothesis that the full effect of the increased autonomy policy on targeted students in grade 12 in 2006-07 (relative to non-targeted students in the control cohort) is equal to zero. All specifications include student fixed effects, an indicator for being targeted in a given year, and classroom-level controls, such as classroom size, standard deviation of prior performance in the classroom, and the proportion of females in the classroom. Specifications in columns (3) and (4) include controls for excused and unexcused absences and their interaction with indicator variables for targeted status, grade level, and cohort. Standard errors clustered at the school level are reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < .01$.

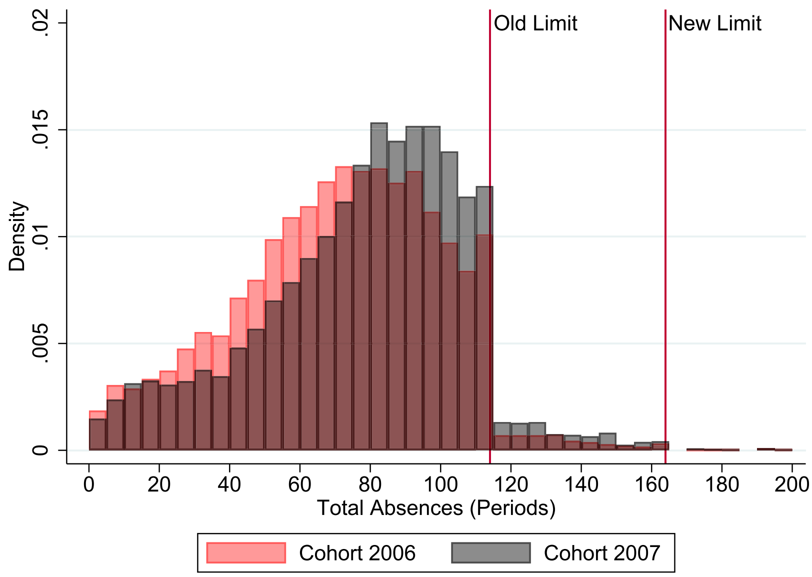


Fig. A.2. Total Absences Distribution in Grade 12 by Cohort This figure shows the distribution of total class absences for the graduating cohorts of 2006 and 2007 in grade 12. Students graduating in 2006 could take up to 114 class absences in total before retention. Students graduating in 2007 could take up to 164 class absence in total before being retained..

Table A.20
Heterogeneous Effects of the Increased Autonomy Policy on Targeted Students by School Postcode Income.

	Below Median Income					Above Median Income				
	(1) Total Absences	(2) Excused Absences	(3) Unexcused Absences	(4) High-stakes Subjects	(5) Low-stakes Subjects	(6) Total Absences	(7) Excused Absences	(8) Unexcused Absences	(9) High-stakes Subjects	(10) Low-stakes Subjects
Targeted in Treated Cohort [1]*relative to non-targeted*	5.178**	3.121**	2.056	0.087***	0.011	2.882	3.278**	-0.396	0.057**	-0.042
	(1.958)	(1.337)	(1.235)	(0.028)	(0.031)	(2.319)	(1.548)	(1.467)	(0.023)	(0.025)
Non-targeted in Treated Cohort [2]*relative to control cohort*	3.369	4.428***	-1.059	-0.038**	-0.005	-0.800	1.084	-1.884	0.005	0.032*
	(2.529)	(1.651)	(1.224)	(0.018)	(0.017)	(2.678)	(1.717)	(1.387)	(0.018)	(0.018)
Observations	12,458	12,458	12,458	12,458	12,458	12,084	12,084	12,084	12,084	12,084
Y Mean (Non-targeted)	67.60	29.18	38.42	-0.44	-0.43	68.87	29.65	39.22	-0.45	-0.44
Y St. Dev. (Non-targeted)	46.62	23.38	38.83	0.69	0.60	50.70	22.92	43.88	0.73	0.60
Y Mean (Targeted)	56.13	28.62	27.51	0.81	0.79	58.70	30.02	28.68	0.77	0.76
Y St. Dev. (Targeted)	32.10	25.30	13.81	0.53	0.62	32.01	24.81	15.80	0.52	0.63
Student FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
P-value for H0: [1] + [2] = 0	0.01	0.00	0.39	0.01	0.75	0.52	0.05	0.12	0.00	0.55

Low-stakes subjects include general education Mathematics, History, and Physics. High-stakes subjects include general education Modern Greek, Ancient Greek, Latin for students in the Classics track, and Mathematics and Physics for students in the Science and IT tracks. Performance in high- and low-stakes subjects is standardized at the school-grade-cohort-subject level. Income information at the school postcode level is in 2009 Euro. Columns 1–5 show the results for students attending school in postcodes with income below the median level, while Columns 6–10 show the results for students attending school in postcodes with income above the median level. All specifications include student fixed effects, an indicator for being targeted in a given year, and classroom-level controls. Classroom controls include classroom size and the proportion of females in the classroom. Standard errors clustered at the school level are reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < .01$.

Table A.21
Correlation of Neighborhood Characteristics.

	Log(Income)	Learning Facilities	Leisure Facilities	Urban (1=yes)
Log(Income)	1.000			
	(0.000)			
Learning Facilities	0.073	1.000		
	(0.480)	(0.000)		
Leisure Facilities	-0.004	0.841	1.000	
	(0.967)	(0.000)	(0.000)	
Urban (1=yes)	0.393	0.444	0.313	1.000
	(0.000)	(0.000)	(0.002)	(0.000)

This table reports pairwise correlation coefficients. P-values for tests of significance are reported in parentheses. Income level at the school postcode level is in 2009 Euro. Learning facilities at the school postcode level reflect libraries and afternoon schools retrieved from yellow pages. Leisure facilities at the school postcode level reflect sports centers, parks, fields, and courts retrieved from yellow pages. Learning and leisure facilities at the school postcode level are standardized. Urban is an indicator taking the value one when school's locale is urban. Information on urban/non-urban locale at the school postcode level was obtained from the Ministry of the Interior.

Table A.22
Heterogeneous Effects of the Increased Autonomy Policy on Targeted Students by Learning Facilities in the Neighborhood.

	Below Median Learning Facilities					Above Median Learning Facilities				
	(1) Total Absences	(2) Excused Absences	(3) Unexcused Absences	(4) High-stakes Subjects	(5) Low-stakes Subjects	(6) Total Absences	(7) Excused Absences	(8) Unexcused Absences	(9) High-stakes Subjects	(10) Low-stakes Subjects
Targeted in Treated Cohort [1]*relative to non-targeted*	3.777	2.999*	0.778	0.065**	-0.002	3.716*	3.344**	0.372	0.078***	-0.026
	(2.467)	(1.646)	(1.448)	(0.026)	(0.030)	(1.956)	(1.400)	(1.305)	(0.024)	(0.026)
Non-targeted in Treated Cohort [2]*relative to control cohort*	1.993	3.229**	-1.236	-0.022	0.003	1.721	3.055	-1.334	-0.016	0.018
	(2.309)	(1.561)	(1.192)	(0.019)	(0.017)	(2.932)	(1.884)	(1.313)	(0.017)	(0.015)
Observations	11,479	11,479	11,479	11,479	11,479	13,063	13,063	13,063	13,063	13,063
Y Mean (Non-targeted)	71.95	30.66	41.29	-0.46	-0.45	65.00	28.33	36.67	-0.43	-0.43
Y St. Dev. (Non-targeted)	57.29	22.79	52.91	0.70	0.60	39.43	23.42	27.61	0.71	0.60
Y Mean (Targeted)	59.22	30.39	28.84	0.79	0.77	55.75	28.34	27.41	0.80	0.78
Y St. Dev. (Targeted)	31.69	25.02	15.93	0.54	0.62	32.35	25.07	13.73	0.52	0.62
Student FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
P-value for H0: [1] + [2] = 0	0.03	0.00	0.72	0.01	0.96	0.13	0.01	0.52	0.00	0.67

Low-stakes subjects include general education Mathematics, History, and Physics. High-stakes subjects include general education Modern Greek, Ancient Greek, Latin for students in the Classics track, and Mathematics and Physics for students in the Science and IT tracks. Performance in high- and low-stakes subjects is standardized at the school-grade-cohort-subject level. Information on the learning facilities at the school postcode level reflects libraries and afternoon schools retrieved from yellow pages. Columns 1–5 show the results for students attending school in postcodes with learning facility below the median level, while Columns 6–10 show the results for students attending school in postcodes with learning facility above the median level. All specifications include student fixed effects, an indicator for being targeted in a given year, and classroom-level controls. Classroom controls include classroom size and the proportion of females in the classroom. Standard errors clustered at the school level are reported in parentheses.
* $p < 0.1$; ** $p < 0.05$; *** $p < .01$.

Table A.23
Heterogeneous Effects of the Increased Autonomy Policy on Targeted Students by Leisure Facilities in the Neighborhood.

	Below Median Leisure Facilities					Above Median Leisure Facilities				
	(1) Total Absences	(2) Excused Absences	(3) Unexcused Absences	(4) High-stakes Subjects	(5) Low-stakes Subjects	(6) Total Absences	(7) Excused Absences	(8) Unexcused Absences	(9) High-stakes Subjects	(10) Low-stakes Subjects
Targeted in Treated Cohort [1]*relative to non-targeted*	2.853	2.183	0.670	0.065**	0.003	4.574**	3.966***	0.608	0.079***	-0.027
	(2.465)	(1.606)	(1.509)	(0.028)	(0.032)	(2.031)	(1.462)	(1.287)	(0.023)	(0.025)
Non-targeted in Treated Cohort [2]*relative to control cohort*	1.409	2.868*	-1.459	-0.021	0.005	1.949	3.129	-1.180	-0.019	0.019
	(2.294)	(1.457)	(1.237)	(0.020)	(0.017)	(2.899)	(1.876)	(1.275)	(0.017)	(0.014)
Observations	10,713	10,713	10,713	10,713	10,713	13,829	13,829	13,829	13,829	13,829
Y Mean (Non-targeted)	72.78	31.04	41.74	-0.46	-0.46	64.74	28.17	36.58	-0.43	-0.42
Y St. Dev. (Non-targeted)	58.39	22.88	54.00	0.70	0.60	39.29	23.30	27.95	0.71	0.60
Y Mean (Targeted)	59.34	30.51	28.84	0.79	0.78	55.84	28.35	27.49	0.79	0.77
Y St. Dev. (Targeted)	31.83	25.04	15.86	0.53	0.62	32.20	25.05	13.93	0.52	0.63
Student FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
P-value for H0: [1] + [2] = 0	0.12	0.01	0.54	0.01	0.69	0.07	0.00	0.70	0.00	0.65

Low-stakes subjects include general education Mathematics, History, and Physics. High-stakes subjects include general education Modern Greek, Ancient Greek, Latin for students in the Classics track, and Mathematics and Physics for students in the Science and IT tracks. Performance in high- and low-stakes subjects is standardized at the school-grade-cohort-subject level. Information on the leisure facilities at the school postcode level reflects sports centers, parks, fields, and courts retrieved from yellow pages. Columns 1–5 show the results for students attending school in postcodes with leisure facilities below the median level, while Columns 6–10 show the results for students attending school in postcodes with leisure facilities above the median level. All specifications include student fixed effects, an indicator for being targeted in a given year, and classroom-level controls. Classroom controls include classroom size and the proportion of females in the classroom. Standard errors clustered at the school level are reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < .01$.

Table A.24
Heterogeneous Effects of the Increased Autonomy Policy on Targeted Students by Urban/Non-urban Locale.

	Urban Neighborhoods					Non-urban Neighborhoods				
	(1) Total Absences	(2) Excused Absences	(3) Unexcused Absences	(4) High-stakes Subjects	(5) Low-stakes Subjects	(6) Total Absences	(7) Excused Absences	(8) Unexcused Absences	(9) High-stakes Subjects	(10) Low-stakes Subjects
Targeted in Treated Cohort [1]*relative to non-targeted*	3.359**	2.464**	0.895	0.076***	-0.019	10.764**	11.857**	-1.093	0.036	0.055
	(1.595)	(1.063)	(1.022)	(0.018)	(0.020)	(4.371)	(4.427)	(3.347)	(0.084)	(0.085)
Non-targeted in Treated Cohort [2]*relative to control cohort*	1.941	3.145***	-1.204	-0.022*	0.015	5.965	7.904	-1.938	-0.027	0.008
	(1.857)	(1.184)	(0.931)	(0.013)	(0.011)	(6.107)	(6.042)	(2.521)	(0.044)	(0.039)
Observations	23,113	23,113	23,113	23,113	23,113	1429	1429	1429	1429	1429
Y Mean (Non-targeted)	68.39	29.30	39.10	-0.45	-0.44	65.44	31.18	34.26	-0.42	-0.43
Y St. Dev. (Non-targeted)	49.42	23.07	42.28	0.71	0.60	33.94	24.52	21.57	0.70	0.57
Y Mean (Targeted)	57.48	29.35	28.14	0.79	0.77	56.20	28.85	27.36	0.79	0.81
Y St. Dev. (Targeted)	32.01	24.97	14.78	0.53	0.62	33.19	26.64	15.85	0.56	0.63
Student FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
P-value for H0: [1] + [2] = 0	0.02	0.00	0.75	0.00	0.82	0.01	0.00	0.32	0.88	0.34

Low-stakes subjects include general education Mathematics, History, and Physics. High-stakes subjects include general education Modern Greek, Ancient Greek, Latin for students in the Classics track, and Mathematics and Physics for students in the Science and IT tracks. Performance in high- and low-stakes subjects is standardized at the school-grade-cohort-subject level. Information on urban/non-urban locale at the school postcode level was obtained from the Ministry of the Interior. Columns 1–5 show the results for students attending school in an urban postcode, while Columns 6–10 show the results for students attending school in a non-urban postcode. All specifications include student fixed effects, an indicator for being targeted in a given year, and classroom-level controls. Classroom controls include classroom size and the proportion of females in the classroom. Standard errors clustered at the school level are reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < .01$.

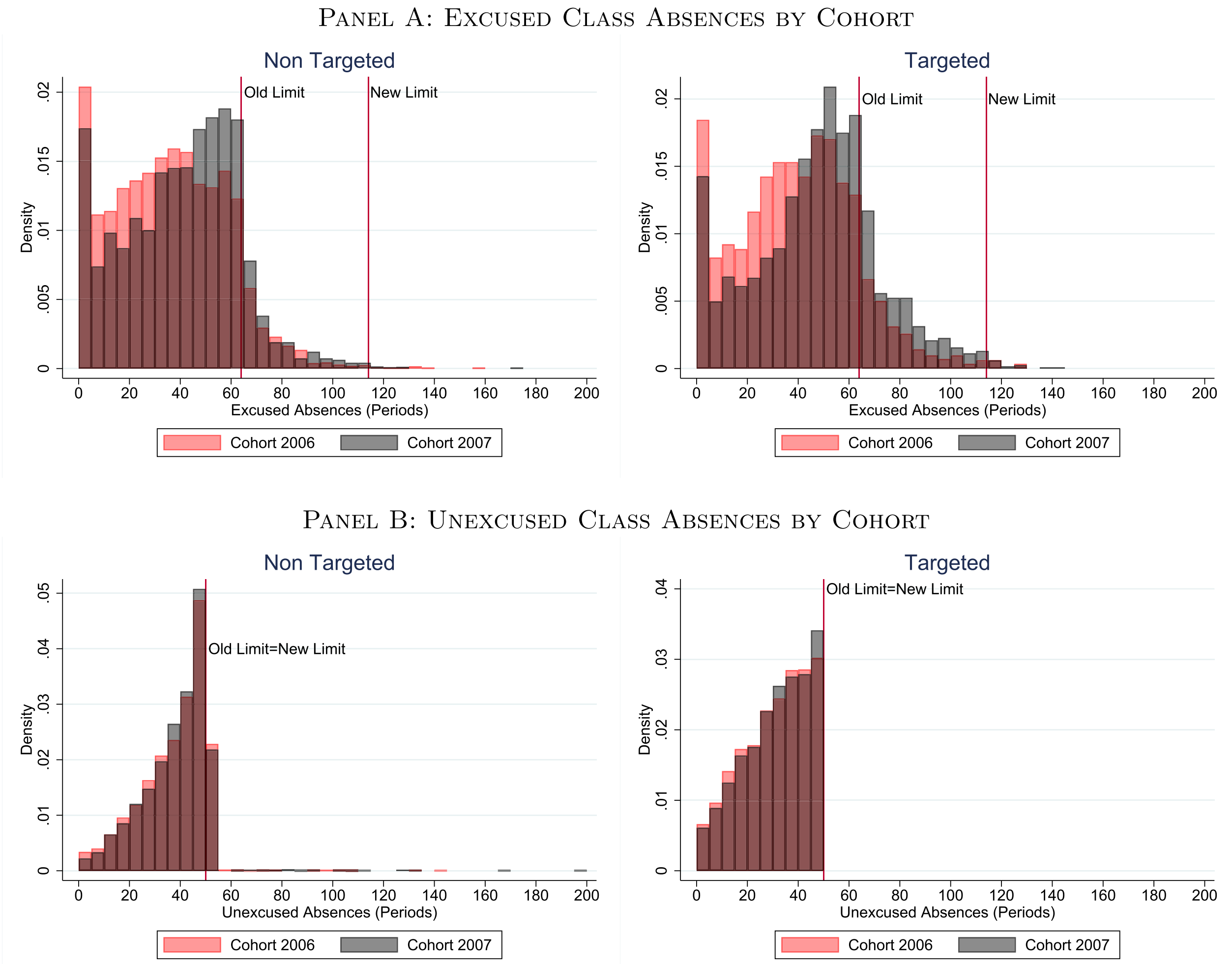


Fig. A.3. Distributions of Absences in Grade 12 by Cohort and Targeted Status Panels A and B show the distribution of excused and unexcused class absences, respectively, for the targeted and non-targeted students in graduating cohorts of 2006 and 2007 in grade 12. Students graduating in 2006 could take up to 64 excused class absences and 50 unexcused class absences before retention. Targeted students graduating in 2007 could take up to 114 excused class absences and 50 unexcused class absences before being retained.

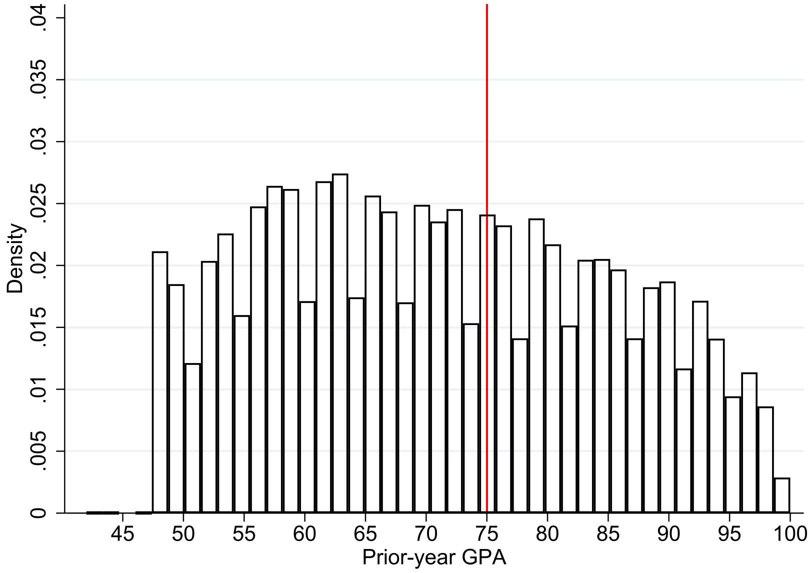
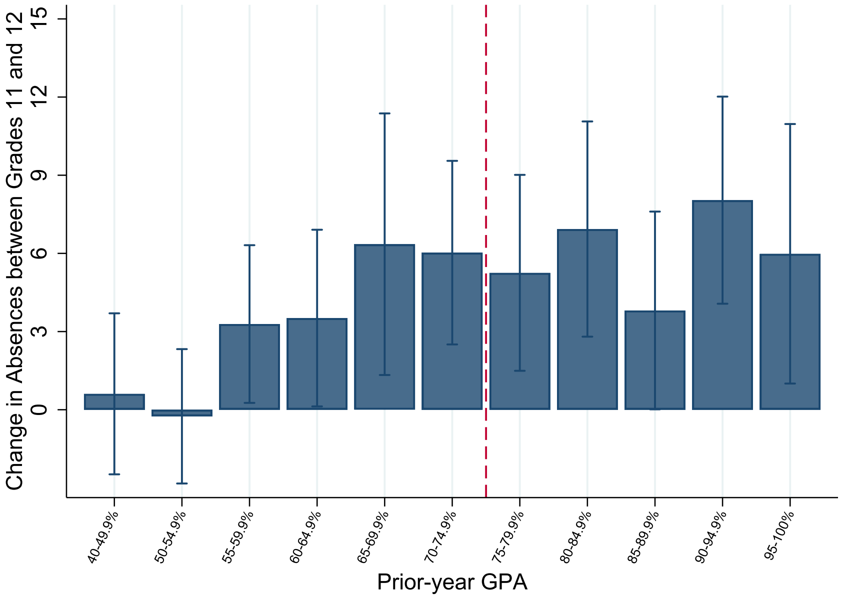


Fig. A.4. Distribution of Prior-year GPA This figure plots the distribution of prior-year GPA across all students. Students with a prior-year GPA above the 75% (reflected by a vertical red line) were eligible for increased autonomy in the form of additional excused class absences in the 2006-07 school year. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Panel A: Excused Class Absences



Panel B: Unexcused Class Absences

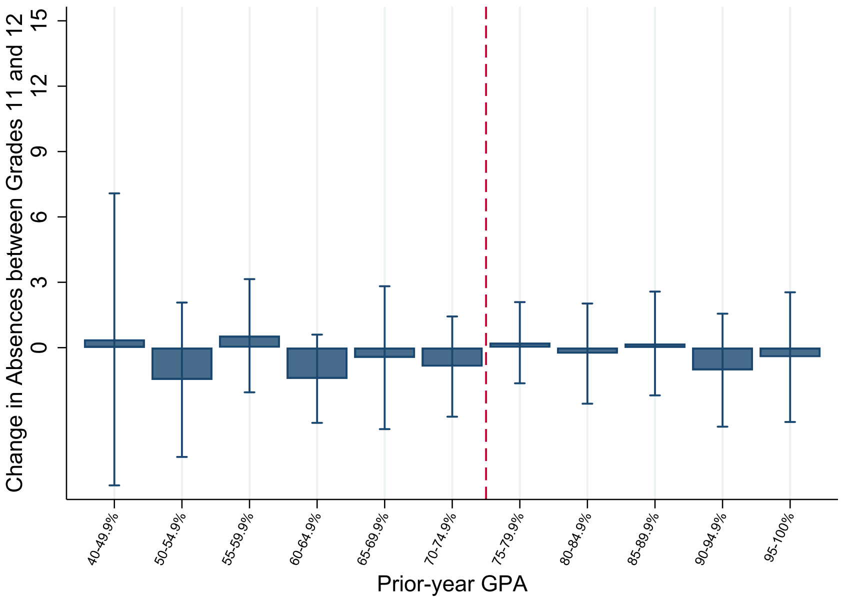
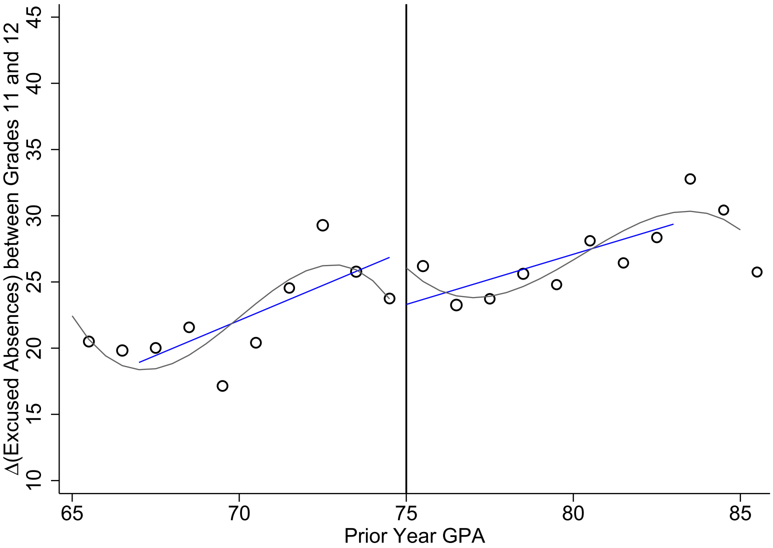


Fig. A.5. Change in Absences between Grades and Cohorts Panels A and B show the differential increases of excused and unexcused absences, respectively, from grade 11 to grade 12 between the treated and control cohorts at different levels of the prior-year GPA..

PANEL A: EXCUSED CLASS ABSENCES



PANEL B: UNEXCUSED CLASS ABSENCES

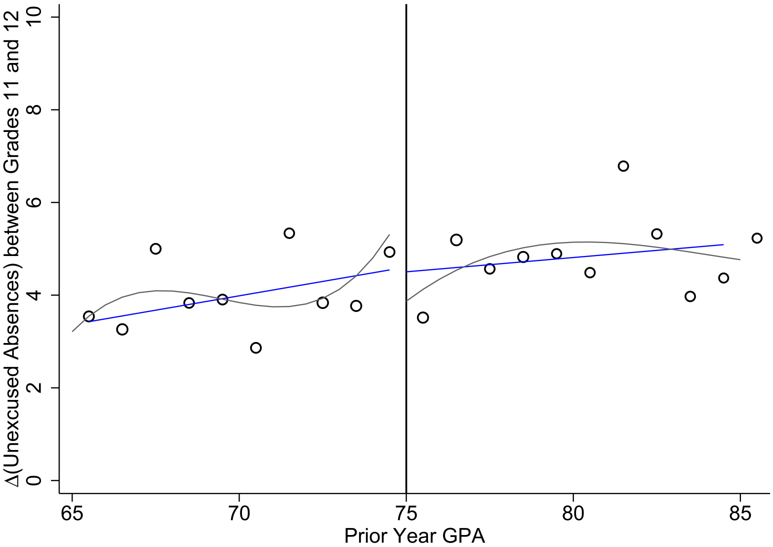


Fig. A.6. Increased Autonomy Effect around the Targeted Status Threshold Panels A and B display locally linear abd cubic scatterplots for the change in parent-approved and non-parent-approved absences, respectively, between grades 11 and 12 around the running variable cutoff for students graduating in 2007 (treatment cohort). The local linear polynomial is estimated using the optimal bandwidth for each outcome based on an MSE criterion (Calonico et al., 2017). The cubic polynomial is estimated using a bandwidth of 1. Dot sizes reflect the number of observations used to calculate the mean. Binsizes for local averages are selected using F-tests from regressions of absences on *K* bin dummies and *2K* bin dummies for the prior-year GPA. .

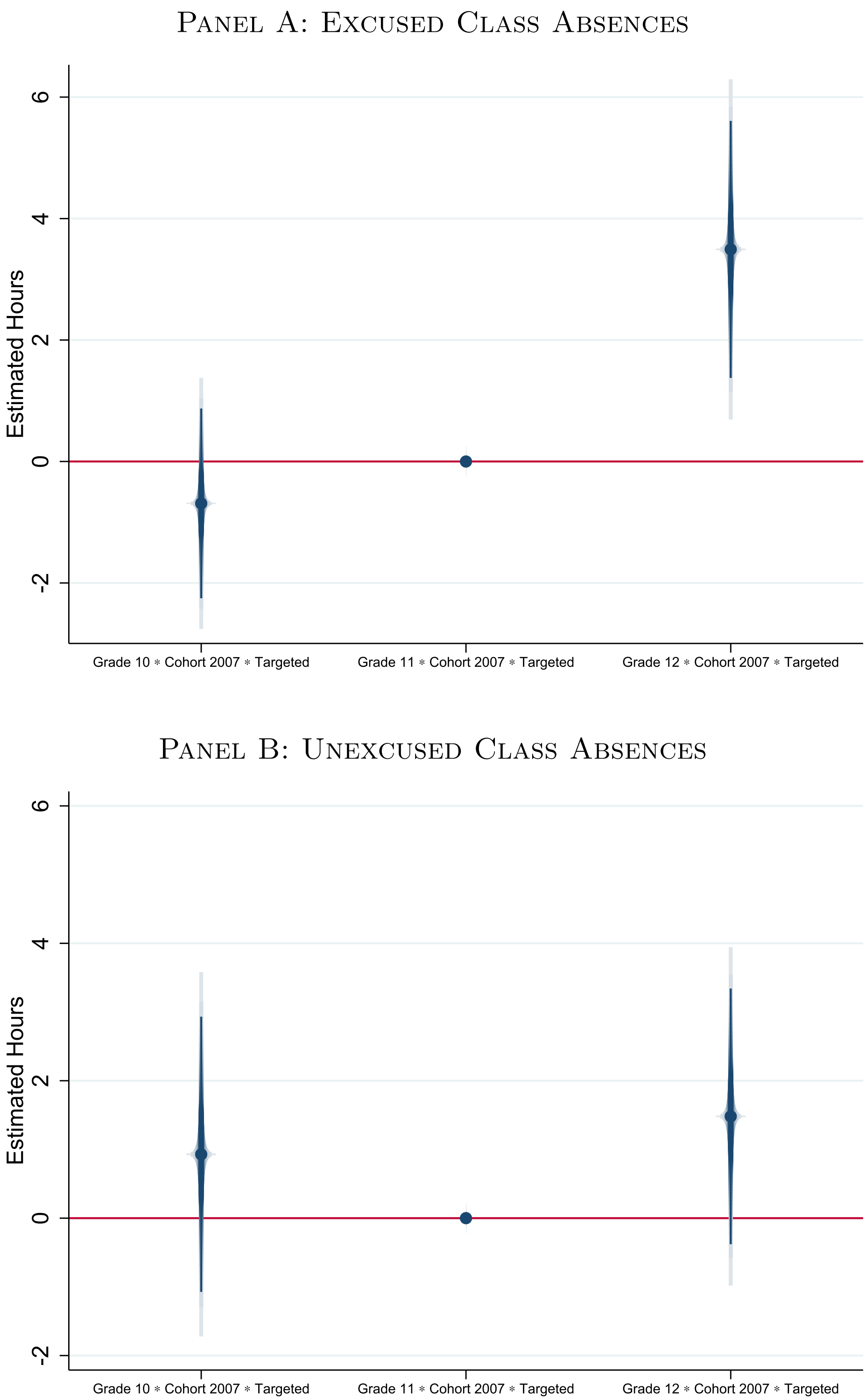


Fig. A.7. Common Trends Investigation of Absences Each panel plots the estimated coefficients for the interactions between indicators for grades 10 and 12, an indicator for the 2007 cohort, and an indicator for targeted status. We plot the 95% confidence interval in dark blue and the 99% confidence interval in light blue.. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

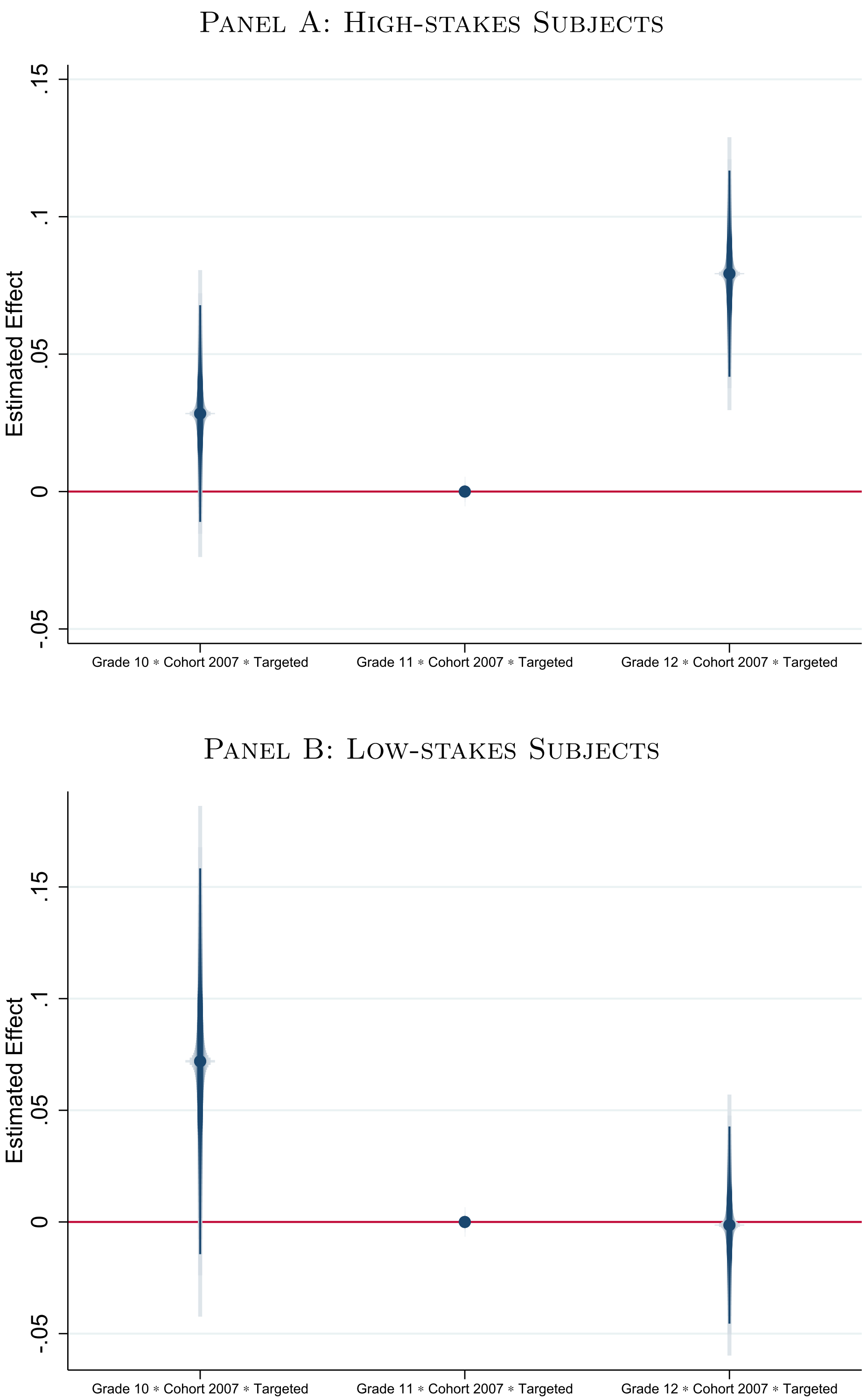
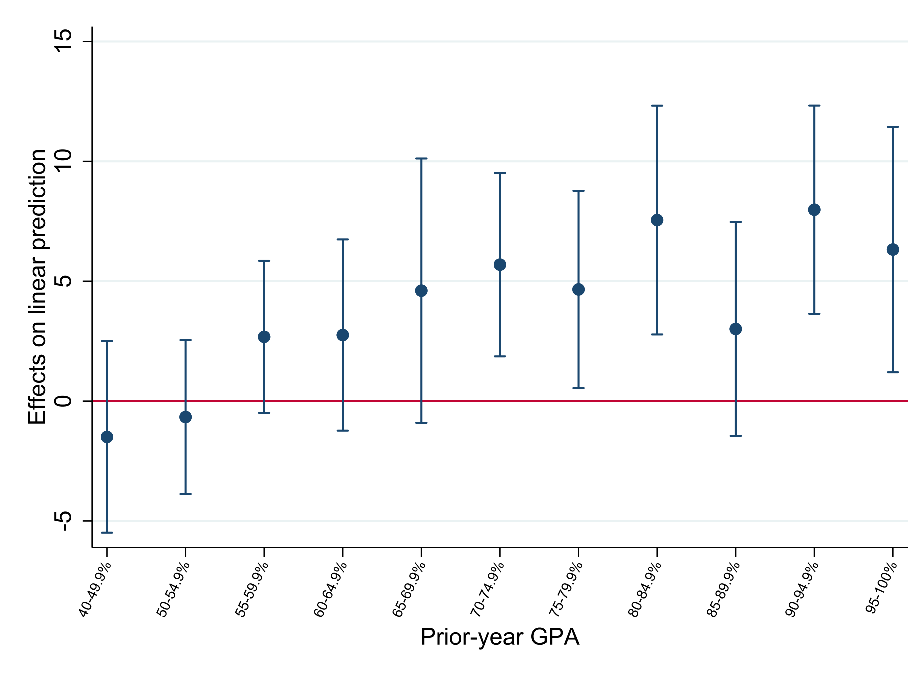


Fig. A.8. Common Trends Investigation of Performance Each panel plots the estimated coefficients for the interactions between indicators for grades 10 and 12, and indicator for the 2007 cohort, and an indicator for targeted status. we plot the 95% confidence interval in dark blue and the 99% confidence interval in light blue.. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

PANEL A: ESTIMATED EFFECT ON EXCUSED CLASS ABSENCES



PANEL B: ESTIMATED EFFECT ON UNEXCUSED CLASS ABSENCES

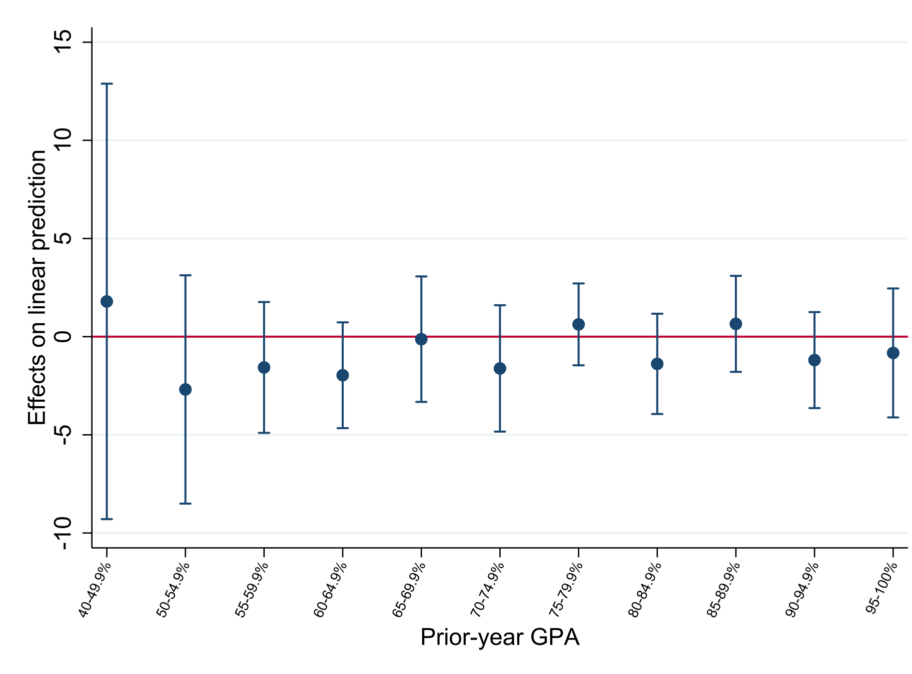


Fig. A.9. Effect of Increased Autonomy on Targeted Students' Absences by Prior Performance Each panel plots the estimated effect of increased autonomy policy on absences captured by the coefficient β_1 in equation (1), by different percentile of prior achievement.

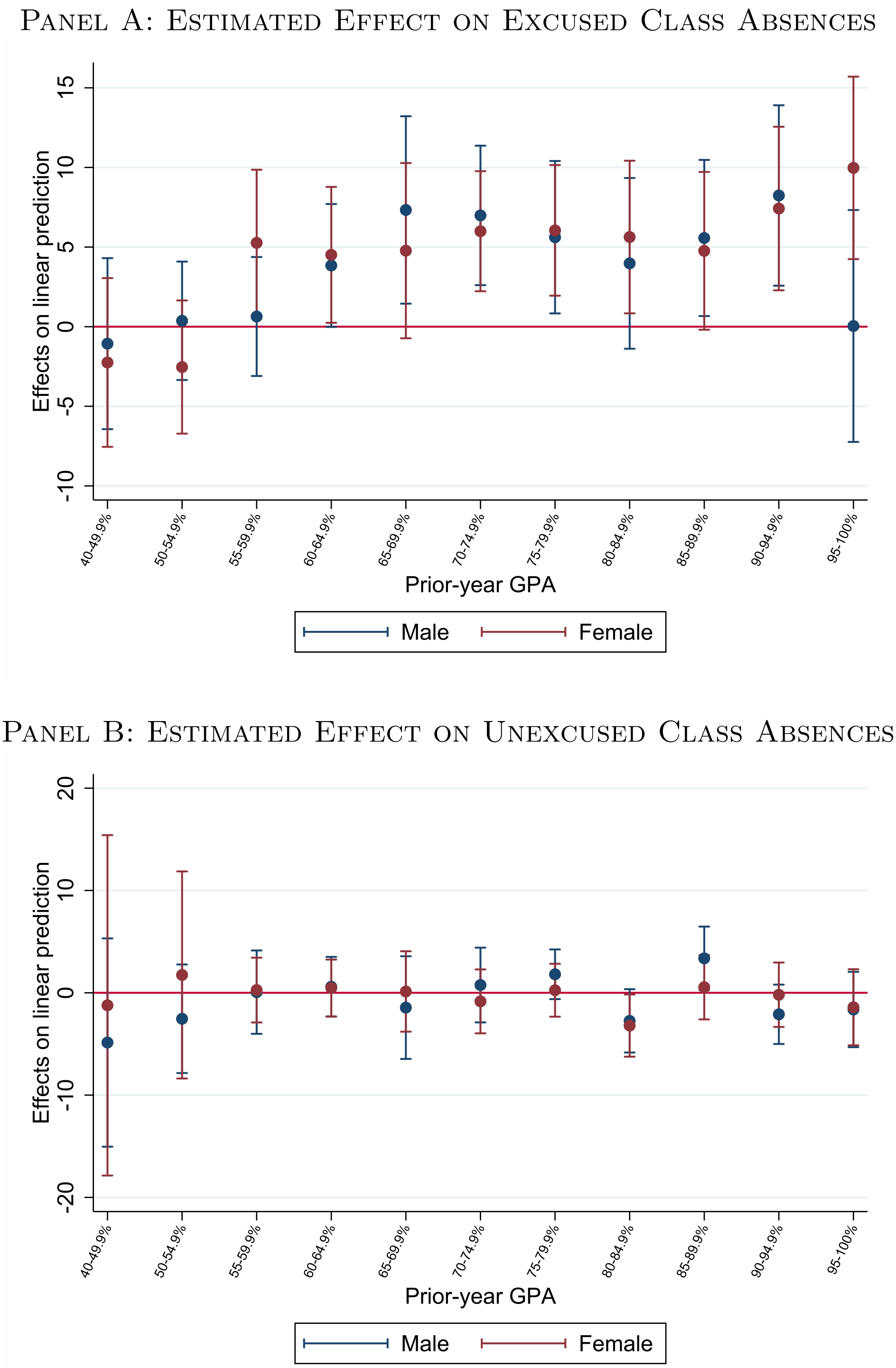


Fig. A.10. Effect of Increased Autonomy on Targeted Students by Previous Performance and Gender Each panel plots the estimated effect of increased autonomy policy captured by the coefficient β_1 in equation (1), by different percentile of prior achievement and gender.

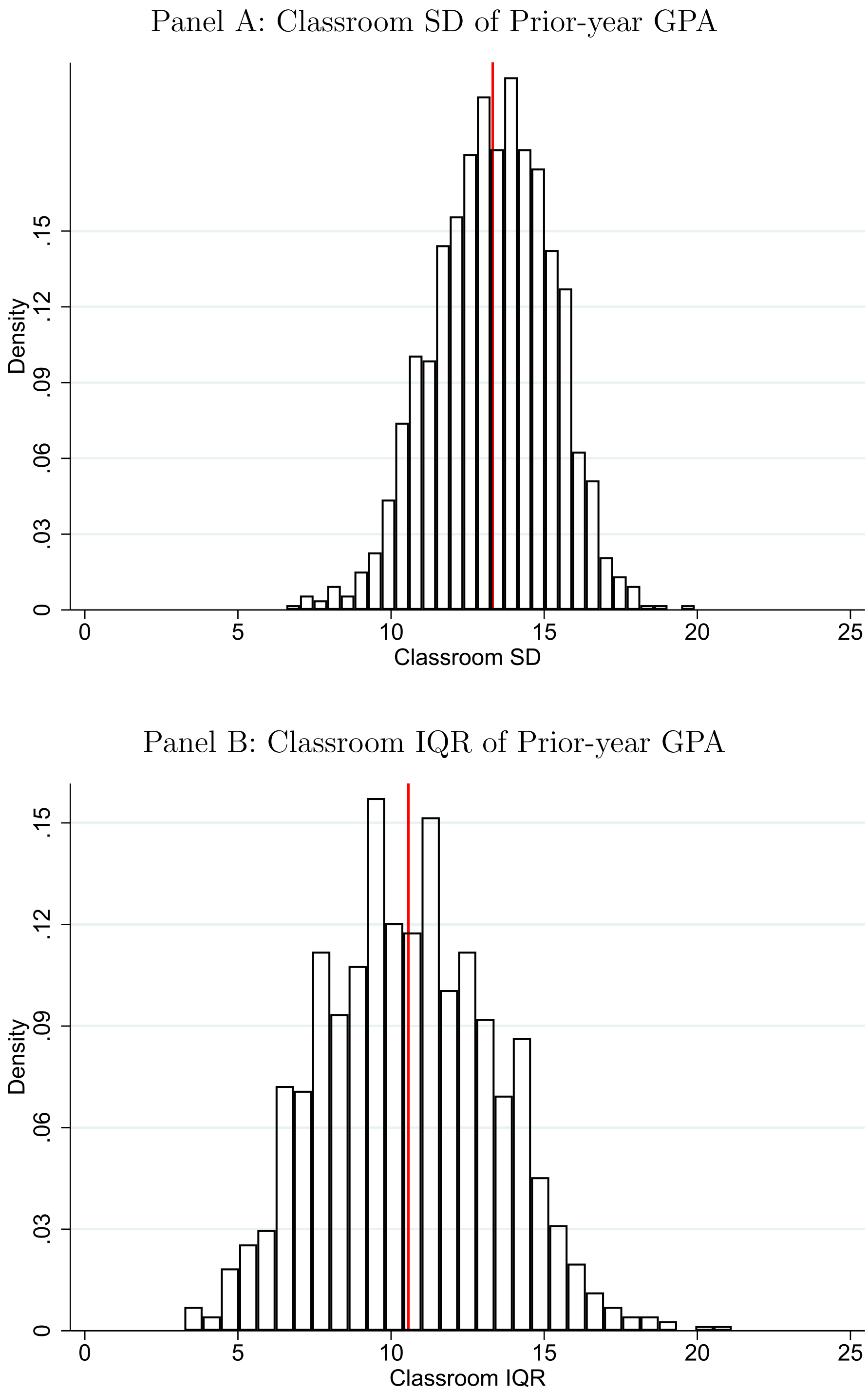


Fig. A.11. Distribution of Classroom Diversity Panel A shows the distribution of classroom diversity measured by the standard deviation (SD) of prior-year GPA in the classroom. Panel B presents the distribution of classroom diversity measured by the interquartile range (IQR) of prior-year GPA in the classroom. Vertical red lines reflect the mean values. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

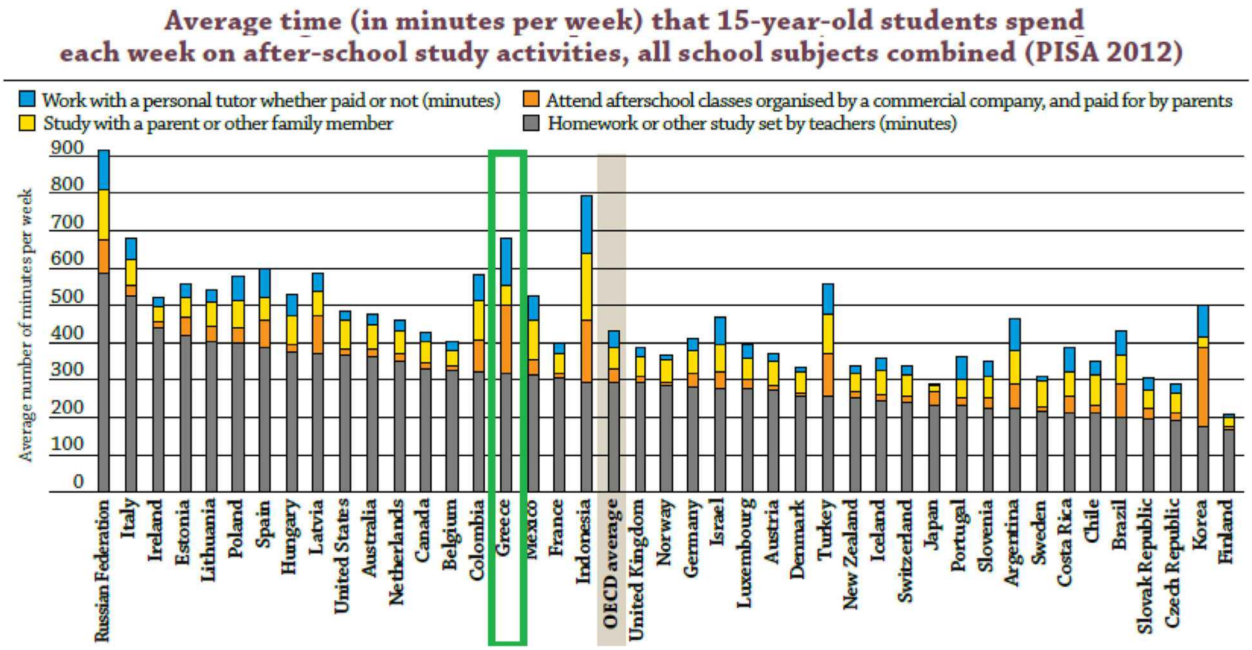


Fig. A.12. Instruction Time in Secondary Education in Greece Countries are ranked in descending order of the number of minutes per week that 15-year-old students spend doing homework or other study set by their teacher. Source: OECD (2016), “How is learning time organised in primary and secondary education?”.

Table A.25
Change in Excused Absences in Grade 12 by Cohort and Targeted Status.

	Targeted		Non-targeted	
More than	Cohort 2006	Cohort 2007	Cohort 2006	Cohort 2007
0 Absences	0.915	0.926	0.866	0.881
6 Absences	0.881	0.904	0.816	0.839
12 Absences	0.845	0.875	0.767	0.791
18 Absences	0.799	0.839	0.713	0.746
24 Absences	0.754	0.799	0.668	0.700
30 Absences	0.709	0.760	0.625	0.660

This table reports the percentage of students who increased their excused absences above different thresholds between grade 11 and grade 12 by targeted status in the control (2006) and treatment (2007) cohorts.

Appendix B. Impact of the Increased Autonomy Policy on Treated Students

In the natural experiment studied, assignment to treatment does not coincide with receiving treatment. The autonomy policy increased the absences allowance of targeted students. The intention was for targeted students to use absences to learn independently. Non-targeted could mimic targeted students' behavior by taking more class absences to the extent allowed by their absences limits. Table A.25 shows this selection into treatment. In particular, Table A.25 shows the percentage of targeted and non-targeted students in each cohort increasing their excused absences above different thresholds. We find that non-targeted students increased their excused absences in grade 12 in both cohorts albeit at a lesser degree than targeted students. In the treated cohort (2007), 76% and 66% of non-targeted and targeted students, respectively, increased their excused absences in grade 12 by more than 30 hours, which is roughly equivalent to a week of schooling.

We investigate the impact of class absences during the autonomy policy regime. We consider students who increased their excused absences from one grade to the next, regardless of whether they were targeted or not, effectively receiving the autonomy policy's treatment. Students who did not increase their excused absences between grades serve as the reference group. We swap out T_{ig} in specification (1) with an indicator for positive change in excused absences by student i in grade g. Table B.1 reports our estimates. We find that students who used the autonomy policy's treatment, regardless of whether they were targeted or not, posted substantial performance gains only in high-stakes exams. Students who did not use the autonomy policy's treatment posted lower high-stakes exam performance during the autonomy policy but the estimated standard errors are large.

Table B.1
Effect of the Increased Autonomy Policy on Treated Students' Absences and Performance.

	Class Absences			Performance	
	Total	Excused	Unexcused	High-stakes Subjects	Low-stakes Subjects
Treated in Treated Cohort [1] *relative to non-treated*	7.464***	4.651***	2.813	0.049**	0.004
	(1.979)	(0.867)	(1.812)	(0.021)	(0.025)
[1em] Non-treated in Treated Cohort [2] *relative to control cohort*	-3.175	-0.458	-2.717	-0.034*	0.002
	(1.980)	(0.536)	(1.840)	(0.018)	(0.021)
Observations	24,542	24,542	24,542	24,542	24,542
Y Mean (Non-treated)	61.87	24.30	37.57	-0.22	-0.19
Y St. Dev. (Non-treated)	62.80	23.31	57.12	0.90	0.83
Y Mean (Treated)	65.01	31.09	33.92	0.09	0.08
Y St. Dev. (Treated)	34.72	23.83	21.93	0.86	0.84
Student FE	✓	✓	✓	✓	✓
P-value for H0: [1] + [2] = 0	0.00	0.00	0.90	0.11	0.44

Low-stakes subjects include general education Mathematics, History, and Physics. High-stakes subjects include general education Modern Greek, Ancient Greek, Latin for students in the Classics track, and Mathematics and Physics for students in the Science and IT tracks. Performance in high- and low-stakes subjects is standardized at the school-grade-cohort-subject level. Coefficient *Treated in Treated Cohort*, [1], represents the effect of increased autonomy policy on students who *voluntarily* took more excused absences in grade 12 than in grade 11 in treated school year 2006-07 relative to non-treated students (those who *voluntarily* did not increase their excused absences between grade 11 and 12) in the same cohort (β_1 in specification (1)). Coefficient *Non-treated in Treated Cohort*, [2], captures the effect of the increased autonomy policy on non-treated students in grade 12 in school year 2006-07 relative to non-treated students in the control school year 2005-06 (β_2 in specification (1)). We test the hypothesis that the full effect of the increased autonomy policy on treated students in grade 12 in 2006-07 (relative to non-treated students in the control cohort) is equal to zero. All specifications include student fixed effects, an indicator for being targeted in a given year, and classroom-level controls, such as classroom size, standard deviation of prior performance in the classroom, and the proportion of females in the classroom. Standard errors clustered at the school level are reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < .01$.

Appendix C. Statistical Power Investigation

We formulate an empirical investigation to gauge the probability that the estimated treatment effect of the analyzed reform is found significant when the true effect is indeed non-zero. We follow Black et al. (2022), as suggested, and perform a simulation exercise to examine how power varies with treatment effects (of varying sizes) induced artificially on real but untreated student records.[46] We consider only students in the control cohort (i.e., students in the 2006 graduating cohort). Using untreated real observations ensures that no treatment effect would influence our results. The simulation exercise is performed as follows. First, we randomly assign a pseudo-treatment regime status to half of the schools in the control cohort. The remaining half constiute a pseudo-control group. Next, we assign a pseudo-treatment status (i.e., pseudo-targeted students) to pseudo-treated students with a prior-year GPA of at least 75% in the pseudo-treated group. Next, we apply the estimated impact of the increased autonomy policy on the pseudo-targeted students. Then, we obtain the p-value of the test of sinificance of the estimated effect on pseudo-targeted students. We run this exercise through 1000 repetitions and calculate the percentage of times the estimated effect was statistically significant at the $1 - \alpha = 95\%$ level.

We consider four different pseudo-treatments of varying magnitudes. The chosen treatment magnitudes come from our main analysis. The idea is to answer the following question: If the true treatment effects are indeed the ones found in the main analysis, what is the probability that our sample would be able to detect them?

Simulation Scenario 1

In the first scenario, we artificially increase the number of excused absences by 6 hours and increase the performance in high-stake subjects by 0.073 standard deviations for all pseudo-targeted students in the pseudo-treated group. These treatment sizes correspond to the average treatment effects of the increased autonomy policy in our main analysis.

Simulation Scenario 2

In the second scenario, we assign half the size of the treatment effects in scenario 1 to all pseudo-targeted students. In other words, we increase the number of excused absences by 3 hours and raise the performance in high-stake subjects by 0.0365 standard deviations for all pseudo-targeted students.

Simulation Scenario 3

In the third scenario, we do not assign the same treatment effect to all pseudo-targeted students. Rather, we draw each student's increase in the number of excused absences from a normal distribution with a mean of 6 hours and a standard deviation of 0.015, matching the point estimate and standard error of the estimated effect of the increased autonomy policy on excused absences.

[46] For complex econometrics models, like in the case of triple difference-in-difference, there is no standard formula for calculating the statistical power does not exist and a simulation approach is necessary.

Table C.1
Simulation Results for Statistical Power.

	Simulated Effect Size on		Calculated Statistical Power		
Scenario	Excused Absences	High-Stake Subject Scores	Excused Absence	High-Stake Subject Scores	Benchmark Treshold of Statistical Power
1	6	0.073	75.4%	83.1%	18-25%
2	3	0.0365	26.0%	30.4%	
3	N(6,1)	N(0.073, 0.018)	76.0%	82.5%	
4	N(3,1)	N(0.0365, 0.018)	26.1%	30.4%	

Notes: Excused absences are measured in class periods. Performance in high-stakes subjects is measured in standard deviation units. The benchmark range of statistical power in economic research is obtained from Zhang and Ortmann (2013) and Ioannidis et al. (2017). Statistical power is the probability of rejecting the null hypothesis when, in fact, it is false. We calculate statistical power in our simulation as the share of simulation iterations in which the null hypothesis of no statistical significance of the parameter of interest is rejected at a significance level $1 - \alpha = 95\%$.

Similarly, we draw each student's treatment effect in performance in high-stake subjects from a normal distribution with a mean of 0.073 and a standard deviation of 0.018, matching the point estimate and standard error of the estimated effect of the increased autonomy policy on performance in high-stakes subjects.

Simulation Scenario 4

In the fourth scenario, we draw each student's increase in the number of excused absences from a normal distribution with a mean of 3 hours and a standard deviation of 1. Similarly, we draw each student's increase in performance in high-stake subjects from a normal distribution with a mean of 0.0365 and standard deviation 0.018 (i.e., half the point estimate and standard error of the estimated effect of the increased autonomy policy on performance in high-stakes subjects).

Table C.1 shows the results from each simulation scenario. In each simulation iteration, we estimate main specification (1). For each of the four simulation scenarios, we report the statistical power, namely the proportion of 1000 simulation iterations that the estimated parameter of interest is found significant at 5% level. In each simulation scenario, we find a statistical power higher than the range of 0.18-0.25 which is proposed in the literature as the typical statistical power in economic research (Zhang and Ortmann, 2013; Ioannidis et al., 2017). Our simulation exercise provides confidence that our sample provides enough power to detect the proposed effects of the increased autonomy policy.

References

Abdulkadiroğlu, A., Angrist, J.D., Hull, P.D., Pathak, P.A., 2016. Charters without lotteries: testing takeovers in new orleans and boston. Am. Econ. Rev. 106 (7), 1878–1920.
Angrist, J.D., Keueger, A.B., 1991. Does compulsory school attendance affect schooling and earnings? Q. J. Econ. 106 (4), 979–1014.
Arulampalam, W., Naylor, R., Smith, J., 2012. Am i missing something? Econ. Educ. Rev. 31 (4), pp.363–375.
Aucejo, E.M., Coate, P., Fruehwirth, J., Kelly, S., Mozenter, Z., 2021. Teacher effectiveness and classroom composition: understanding match effects in the classroom. CEPR Discussion Paper Number DP13166.
Aurini, J., Davies, S., Dierkes, J., 2013. Out of the shadows? an introduction to worldwide supplementary education. Out of the Shadows: The Global Intensification of Supplementary Education xv–xxiv.
Black, B., Hollingsworth, A., Nunes, L., Simon, K., 2022. Simulated power analyses for observational studies: an application to the affordable care act medicaid expansion. J. Public. Econ. 213, 104713.
Blatchford, P., Russell, A., 2019. Class size, grouping practices and classroom management. Int. J. Educ. Res. 96, 154–163.
Bray, M., 2013. Shadow education: comparative perspectives on the expansion and implications of private supplementary tutoring. Procedia-Soc. Behav. Sci. 77, 412–420.
Brough, R., Phillips, D., Turner, P., 2021. The labor market returns to reversing high school dropout. Available at SSRN 3840453.
Calonico, S., Cattaneo, M.D., Farrell, M.H., Titiunik, R., 2017. Rdrobust: software for regression-Discontinuity designs. Stata J. 17 (2), 372–404.
Cameron, A.C., Gelbach, J.B., Miller, D.L., 2008. Bootstrap-based improvements for inference with clustered errors. Rev. Econ. Stat. 90 (3), 414–427.
Card, D., Giuliano, L., 2016. Can tracking raise the test scores of high-ability minority students? Am. Econ. Rev. 106 (10), 2783–2816.
Carrell, S.E., West, J.E., 2010. Does professor quality matter? evidence from random assignment of students to professors. J. Polit. Econ. 118 (3), 409–432.
Caviglia, H., 2006. Attendance and achievement in economics: investigating the impact of attendance policies and absentee rates on student performance. J. Econ. Finance Educ. 4 (2), pp.470–476.
CCL, 2007. Survey of canadian attitudes towards learning: results for elementary and secondary school learning. Canadian Council on Learning.
Cohn, E., Johnson, E., 2006. Class attendance and performance in principles of economics. Educ. Econ. 14 (2), 211–233. doi:10.1080/09645290600622954.
Cook, T.D., Scriven, M., Coryn, C.L., Evergreen, S.D., 2010. Contemporary thinking about causation in evaluation: a dialogue with tom cook and michael scriven. Am. J. Eval. 31 (1), 105–117.
Cunningham, S., 2021. Causal Inference: The Mixtape. Yale University Press, New Haven, CT.
Deci, E.L., Ryan, R.M., 2012. Self-determination Theory. In: Handbook of Theories of Social Psychology, Vol. 1. Sage Publications Ltd, Thousand Oaks, CA, pp. 416–436. 10.4135/9781446249215.n21
Dehejia, R.H., Wahba, S., 1999. Causal effects in nonexperimental studies: reevaluating the evaluation of training programs. J. Am. Stat. Assoc. 94 (448), 1053–1062.
Dobbie, W., Fryer Jr, R.G., 2013. Getting beneath the veil of effective schools: evidence from new york city. Am. Econ. J.: Appl. Econ. 5 (4), 28–60.
Dobkin, C., Gil, R., Marion, J., 2010. Skipping class in college and exam performance: evidence from a regression discontinuity classroom experiment. Econ. Educ. Rev. 29 (4), 566–575.
Duflo, E., Dupas, P., Kremer, M., 2011. Peer effects, teacher incentives, and the impact of tracking: evidence from a randomized evaluation in kenya. Am. Econ. Rev. 101 (5), 1739–1774.
Evans, M., Boucher, A.R., 2015. Optimizing the power of choice: supporting student autonomy to foster motivation and engagement in learning. Mind, Brain Educ. 9 (2), 87–91.
Fryer Jr, R.G., 2014. Injecting charter school best practices into traditional public schools: evidence from field experiments. Q. J. Econ. 129 (3), 1355–1407.
Gaete, G., 2018. Follow the Leader: Student Strikes, School Absenteeism and Long-term Implications for Education OutcomesMimeo, Available at SSRN: http://ssrn.com/abstract=2988825.

Gottfried, M.A., 2010. Evaluating the relationship between student attendance and achievement in urban elementary and middle schools: an instrumental variables approach. Am. Educ. Res. J. 47 (2), 434–465.
Goulas, S., Griselda, S., Megalokonomou, R., 2020. Comparative advantage and gender gap in STEM. IZA Discussion Paper Number 13313.
Goulas, S., Gunawardena, B., Megalokonomou, R., 2023. Same-Gender peer role models in education. Mimeo.
Goulas, S., Megalokonomou, R., 2020. School attendance during a pandemic. Econ. Lett. 193, 109275.
Goulas, S., Megalokonomou, R., 2021. Knowing who you actually are: the effect of feedback on short-and longer-term outcomes. J. Econ. Behav. Org. 183, 589–615.
Goulas, S., Megalokonomou, R., Zhang, Y., 2018. Does the girl next door affect your academic outcomes and career choices? IZA Discussion Paper Number 11910.
Goulas, S., Megalokonomou, R., Zhang, Y., 2022. Female Neighbors, Test Scores, and Careers.
Goulas, S., Raymond, M.E., Bierbaum, L., Bell, L., Snow, W., 2017. The impact of scaling the new orleans charter restart model on student performance. Center for Research on Education Outcomes at Stanford University.
Guryan, J., Kroft, K., Notowidigdo, M.J., 2009. Peer effects in the workplace: evidence from random groupings in professional golf tournaments. Am. Econ. J.: Appl. Econ. 1 (4), 34–68. http://www.aeaweb.org/articles?id=10.1257/app.1.4.34
Hanushek, E.A., Peterson, P.E., Woessmann, L., 2012. Achievement growth: international and US state trends in student performance. PEPG report no.: 12-03. Program on Education Policy and Governance, Harvard University.
Hattie, J., Hodis, F.A., Kang, S.H., 2020. Theories of motivation: integration and ways forward. Contemp. Educ. Psychol. 61, 101865.
Hennessy, S., Bilker, W.B., Berlin, J.A., Strom, B.L., 1999. Factors influencing the optimal control-to-Case ratio in matched case-Control studies. Am. J. Epidemiol. 149 (2), 195–197.
Imbens, G.W., Wooldridge, J.M., 2009. Recent developments in the econometrics of program evaluation. J. Econ. Lit. 47 (1), 5–86.
Ioannidis, J.P., Stanley, T.D., Doucouliagos, H., 2017. The power of bias in economics research. Oxford University Press Oxford, UK.
Kapoor, S., Oosterveen, M., Webbink, D., 2021. The price of forced attendance. J. Appl. Econometr. 36 (2), 209–227.
Keogh-Brown, M., Bachmann, M., Shepstone, L., Hewitt, C., Howe, A., Ramsay, C. R., Song, F., Miles, J., Torgerson, D., Miles, S., et al., 2007. Contamination in Trials of Educational Interventions.
Krohn, G.A., O'Connor, C.M., 2005. Student effort and performance over the semester. J. Econ. Educ. 36 (1), 3–28.
Krueger, A.B., 2003. Economic considerations and class size. Econ. J. 113 (485), F34–F63.
Kuhfeld, M., Soland, J., Tarasawa, B., Johnson, A., Ruzek, E., Liu, J., 2020. Projecting the potential impact of COVID-19 school closures on academic achievement. Educ. Researcher 49 (8), 549–565.
Latif, E., Miles, S., 2013. Class attendance and academic performance: A Panel data analysis. Econ. Paper. 32 (4), pp.470–476.
Lavecchia, A.M., Liu, H., Oreopoulos, P., 2016. Behavioral Economics of Education: Progress and Possibilities. In: Handbook of the Economics of Education, Vol. 5. Elsevier, pp. 1–74.
Lavy, V., 2015. Do differences in schools instruction time explain international achievement gaps? evidence from developed and developing countries. Econ. J. 125 (588), F397–F424.
Lavy, V., Megalokonomou, R., 2019. Persistency in teachers grading bias and effects on longer-Term outcomes: university admissions exams and choice of field of study. NBER Working Paper (26021) doi:10.3386/w26021. http://www.nber.org/papers/w26021
Lavy, V., Silva, O., Weinhardt, F., 2012. The good, the bad, and the average: evidence on ability peer effects in schools. J. Labor. Econ. 30 (2), 367–414.
Lee, D.S., Lemieux, T., 2010. Regression discontinuity designs in economics. J. Econ. Lit. 48 (2), 281–355.
Liodaki, N., Liodakis, N., 2016. Some Effects of the Economic Crisis on Shadow Education in Greece.
Liu, H., Salvo, A., 2018. Severe air pollution and child absences when schools and parents respond. J. Environ. Econ. Manage. 92, 300–330.
Liu, J., Lee, M., Gershenson, S., 2021. The short-and long-run impacts of secondary school absences. J. Public Econ. 199, 104441.
Marburger, D.R., 2001. Absenteeism and undergraduate exam performance. J. Econ. Educ. 32 (2), 99–109.
Marburger, D.R., 2006. Does mandatory attendance improve student performance? J. Econ. Educ. 37 (2), 148–155.
Martins, P. S., Walker, I., 2006. Student Achievement and University Classes: Effects of Attendance, Size, Peers, and Teachers.
Messacar, D., Oreopoulos, P., 2012. Staying in School: A Proposal to Raise High School Graduation Rates.
Moore, R., 2006. Class attendance: how Students' attitudes about attendance relate to their academic performance in introductory science classes. Res. Teach. Dev. Educ. 23 (1), 19–33. http://www.jstor.org/stable/42802273
Moore, R., Jensen, M., Hatch, J., Duranczyk, I., Staats, S., Koch, L., 2003. Showing up: the importance of class attendance for academic success in introductory science courses. Am. Biol. Teach. 65 (5), 325–329. 10.1662/0002-7685(2003)065[0325:SUTIOC]2.0.CO;2
NECTL, 2005. Prisoners of time: Report of the National Education Commission on Time and Learning. Technical Report. National Education Commission on Time and Learning, Washington, DC.
OECD, 2013. Pisa 2012 results: What makes schools successful (volume iv) doi:10.1787/9789264201156-en.
OECD, 2014. Strengthening resilience through education: pisa results.
OECD, 2015. Education at a glance 2015. OECD Publishing, Paris. doi.org/10.1787/eag-2015-en
OECD, 2016. How is learning time organised in primary and secondary education?. 10.1787/5jm3tqsm1kq5-en
OECD, 2018. Education for a bright future in greece. OECD Publishing, Paris. 10.1787/9789264298750-en
OECD, 2018. Pisa results: Greece.
OECD, 2021. Greece.
Oreopoulos, P., 2006. Estimating average and local average treatment effects of education when compulsory schooling laws really matter. Am. Econ. Rev. 96 (1), 152–175.
Oreopoulos, P., 2007. Do dropouts srop out too soon? wealth, health and happiness from compulsory schooling. J. Public. Econ. 91 (11–12), 2213–2229.
Pane, J.F., Steiner, E.D., Baird, M.D., Hamilton, L.S., 2015. Continued progress: Promising evidence on personalized learning. RAND Corporation, Santa Monica, CA doi:10.7249/RR1365.
Pier, L., Hough, H.J., Christian, M., Bookman, N., Wilkenfeld, B., Miller, R., 2021. COVID-19 and the Educational Equity Crisis: Evidence on Learning Loss From the CORE Data Collaborative. Technical Report. Policy Analysis for California Education.
Raymond, M.E., Han, C., Goulas, S., Lee, W.F., Waeiss, C., 2020. Estimates of learning loss in the 2019–2020 school year. Center for Research on Education Outcomes at Stanford University.
Reichardt, C. S., 2019. Quasi-experimentation: A Guide to Design and Analysis.
Reis, S.M., Renzulli, J.S., 2010. Is there still a need for gifted education? an examination of current research. Learn. Individ. Differ. 20 (4), 308–317.
Rivkin, S.G., Schiman, J.C., 2015. Instruction time, classroom quality, and academic achievement. Econ. J. 125 (588), F425–F448.
Romer, D., 1993. Do students go to class? should they? J. Econ. Perspect. 7 (3), pp.167–174. http://www.jstor.org/stable/2138449
Ryan, R.M., Deci, E.L., 2017. Self-determination theory. Basic Psychol. Needs Motivat. Dev. Wellness.
Ryan, R.M., Deci, E.L., 2020. Intrinsic and extrinsic motivation from a self-determination theory perspective: definitions, theory, practices, and future directions. Contemp. Educ. Psychol. 61, 101860.
Sacerdote, B., 2011. Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far? In: Handbook of the Economics of Education, Vol. 3. Elsevier, pp. 249–277.
Salvati, A., 2021. Teaching to the Middle or to the Edges: Classroom Composition, Instructional Choices, and Their Impact on Student Achievement.
Scott, M.M., Shakesprere, J., Porter, K., 2020. Why Schools Should Embrace Flexibility and Innovation Beyond COVID-19. Technical Report. Urban Institute, Washington, D.C..

Singer, N., 2021. Online schools are here to stay, even after the pandemic. The New York Times. https://www.nytimes.com/2021/04/11/technology/remote-learning-online-school.html
Snyder, J.L., Lee-Partridge, J.E., Jarmoszko, A.T., Petkova, O., D'Onofrio, M.J., 2014. What is the influence of a compulsory attendance policy on absenteeism and performance? J. Educ. Bus. 89 (8), 433–440.
Snyder, T.D., Dillow, S.A., 2015. Digest of education statistics 2013. NCES 2015-011. Natl. Center Educ. Stat..
Steinberg, M.P., Garrett, R., 2016. Classroom composition and measured teacher performance: what do TTeacher observation scores really measure? Educ. Eval. Policy Anal. 38 (2), 293–317.
Stuart, E.A., 2010. Matching methods for causal inference: A Review and a look forward. Stat. Sci.: Rev. J. Inst. Math. Stat. 25 (1), 1.
Wright, M.C., Bergom, I., Bartholomew, T., 2019. Decreased class size, increased active learning? intended and enacted teaching strategies in smaller classes. Active Learn. Higher Educ. 20 (1), 51–62.
Zhang, L., Ortmann, A., 2013. Exploring the meaning of significance in experimental economics. UNSW Austral. School Bus. Res. Paper (2013–32).
Zimmerman, B.J., 2008. Investigating self-regulation and motivation: historical background, methodological developments, and future prospects. Am. Educ. Res. J. 45 (1), 166–183.