I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14559

# Compulsory Class Attendance versus Autonomy

Sofoklis Goulas
Silvia Griselda
Rigissa Megalokonomou

JULY 2021

DISCUSSION PAPER SERIES

IZA DP No. 14559

# Compulsory Class Attendance versus Autonomy

**Sofoklis Goulas**
*Stanford University*

**Silvia Griselda**
*Bocconi University*

**Rigissa Megalokonomou**
*University of Queensland and IZA*

JULY 2021



**IZA – Institute of Labor Economics**

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

# ABSTRACT

## Compulsory Class Attendance versus Autonomy*

Understanding the impact of the COVID-19 pandemic on education requires a solid grasp of the impact of student autonomy on learning. In this paper, 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 in randomly formed classrooms. Identification comes from a natural experiment that allowed higher-achieving students to miss 30 percent more classes without penalty. Using a difference-in-difference-in-differences approach, we find that allowing higher-achieving students to skip class more often improves their performance in high-stakes subjects and increases their university admission outcomes. Higher-achieving students in more academically diverse classrooms exerted more autonomy when allowed to.

**JEL Classification:** I26

**Keywords:** COVID-19, learning autonomy, school attendance, returns to education, natural experiment

**Corresponding author:**
Sofoklis Goulas
Hoover Institution
Stanford University
450 Serra Mall
Stanford, CA 94305
USA
E-mail: goulas@stanford.edu

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

# 1 Introduction

During the COVID-19 pandemic, officials in many countries used school distancing to mitigate the spread of the virus (Goulas and Megalokonomou, 2020; Richardson, Hannah and Sellgren, Katherine, 2020). School distancing was often implemented partially (e.g., in Australia and the United Kingdom) with students allowed to attend school if they chose to (Richardson, Hannah and Sellgren, Katherine, 2020). While the pandemic distanced students from the school, officials are interested in evaluating the consequences of the pandemic-induced student autonomy (OECD, 2020; World Bank, 2020).

Student autonomy is likely to have different effects on students with different performance history. While previous literature has analyzed the negative return of absences for low-performing students (Dobkin et al., 2010; Marburger, 2001; Snyder et al., 2014; Kapoor et al., 2021), in this paper we focus on how higher autonomy, in the form of a lower class attendance requirement, affects higher-achieving students' performance. There are several reasons the returns to attendance of higher-performing students may differ from those of lower-performing students. First, students on a trajectory of higher academic performance 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 offer less 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).

In this paper, we draw on a reform implemented by the Greek Ministry of Education, intended to encourage students' autonomy (Gov. Gazette 65/A/30-3-2006). The increased autonomy policy allowed high school students with a prior-year grade point average (GPA) greater than 75% to miss up to 30 percent more classes than before without penalty. Targeted students now have greater flexibility to make choices that best serve their own interests (i.e., time on self-study or leisure). Attendance requirements were unaffected for students with a prior-year GPA lower than 75%. We exploit this natural experiment to estimate the effects of increased time autonomy on student short- (i.e., high school performance) and longer-term outcomes (i.e., performance on university admission exam and quality of enrolled degree).

We develop a simple theoretical model that provides insights into students' time allocation problem. Students maximize their utility by allocating time between leisure and study (either in class or at home) of subjects with differing utility weights. For example, students may spend more time studying *high-stakes subjects* that matter for university admission, and less time studying *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 have to allocate to class study. Relaxing attendance requirements allows students to reallocate time from class study to leisure or home study. Our framework provides intuition for the drivers of student attendance decisions and provides hypotheses, which we test empirically.

Our empirical strategy employs a *difference-in-difference-in-differences* (DDD) methodology to identify the intention-to-treat effects of autonomy on targeted students. This method compares changes in school attendance, performance, and university admission outcomes for students from grade 11 to grade 12, in unaffected (control group) and affected (treatment group) cohorts, for both targeted (those with a prior-year GPA above 75%) and non-targeted students (those with a prior-year GPA below 75%). By using multiple cohorts and conditioning on student fixed effects, we are able to control for unobserved factors that might confound the propensity to attend class. We enhance the credibility of the identification strategy by examining the existence of common trends between treated and control groups. We also validate the triple-differences estimates through a set of matching methodologies.

Our results show that providing increased autonomy decreases the attendance of targeted students by roughly four classes per year and increases their performance in *high-stakes subjects* by 0.07 standard deviations. The autonomy effect on targeted higher-achieving students' high-stakes performance is comparable to a reduction in class size of about 8% (Krueger, 2003), an increase in teacher quality by 1.4 standard deviations (Carrell and West, 2010), or attending 50 additional days of schooling (Hanushek et al., 2012). In contrast, targeted students' performance in *low-stakes subjects* remains unaffected. Our results reveal that higher-performing students perform better on university admission exams and are admitted to university degree programs of higher selectiveness when they are allowed more autonomy in the form of class absences.

We find varying effects of autonomy based on class heterogeneity. In particular, we show that targeted higher-performing students miss more classes when quasi-randomly assigned to more academically diverse classrooms. This points to potentially higher incentives to skip class for higher-performing students when they are in more heterogeneous peer environments. Our finding that higher-performing students demand more autonomy in academically diverse settings may be associated with higher-performing students' potentially increased readiness to self-learn and manage learning material independently.

We contribute to the existing literature in two important ways.[1] Parents, policymakers, and

[1]Our study is also related to the literature on the effect of school absences on performance. 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;

researchers are interested in the effect of class attendance requirements on students from different parts of the ability distribution (NECTL, 2005). The literature thus far has primarily focused on the association between attendance requirements and performance for lower-performing students (Dobkin et al., 2010; Marburger, 2001; Snyder et al., 2014; Kapoor et al., 2021). Our paper is the first, to our knowledge, to examine the impact of class attendance requirements on higher-performing students.

There are three reasons increased autonomy may be optimal for higher-performing students. First, talented students may have a propensity to better self-regulate, and therefore the ability to acquire more knowledge on their own (Zimmerman, 2008). Second, higher-performing students may be allocating their time across subjects and material more efficiently (Reis and Renzulli, 2010). Third, forced class attendance may be potentially detrimental for high-achieving students. The Report of the National Education Commission in the U.S. points out that high-ability 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."

We also contribute to the literature on the association between peer group diversity and student outcomes. Aucejo et al. (2021) show that teacher effectiveness drops in more academically diverse classrooms. This suggests that a more diverse classroom may offer an environment that is less conducive to learning for higher-performing students. Thus, optimal class attendance may be lower for higher-performing students in a more diverse classroom. Our analysis provides corroborating evidence of this finding by examining the impact of autonomy on attendance and performance in quasi-randomly formed classrooms.

The COVID-19 pandemic changed the way students attend school and allowed for greater learning autonomy. The pandemic has also contributed to greater academic diversity at school (Kuhfeld et al., 2020; Raymond et al., 2020; Pier et al., 2021). Understanding how high-performing students respond to autonomy—especially in academically diverse settings—can provide focus for recovery strategies from pandemic-related learning losses and predict their success.

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

---

Arulampalam et al., 2012; Latif and Miles, 2013; Gaete, 2018; Liu et al., 2021).

## 2 A Simple Model of Time Allocation

In this section, we develop a theoretical framework to study students' time allocation problem. In this framework, a student spends time learning or enjoying leisure. Students can learn in either a classroom environment or outside the classroom, potentially at home, studying independently. Students take two types of subjects: high- and low-stakes subjects. Suppose the representative student faces the following additively separable utility function:

$$U = u(l, s(c, h, a)), \tag{1}$$

which is a function of leisure, $l$ and the following education production function:

$$s = s(c, h, a), \tag{2}$$

where $s$ is a measure of a student's academic performance, $c$ is the amount of time she spends in class, $h$ is the amount of time the student spends in out-of-class learning activities, and $a$ captures individual characteristics such as ability, motivation, and effort. Assume that $c$ and $h$ in the production function are neither complements nor substitutes, but independent. Suppose, for simplicity and without loss of generality, that the marginal utility of $s$ is one and the marginal utility of leisure is constant for every unit of time outside the classroom.[2]

The student maximizes utility (1) by allocating her time efficiently among leisure, in-class study, and out-of-class study, given her time constraint:

$$c^* + o^* \leq T^* + l^* \leq D, \tag{3}$$

where $c^*$ is the time students devote to classroom learning, $o^*$ represents class absences, and $T^*$ is the time designated by the school for classroom learning. The remainder of the time unit considered, $D$, is spent on leisure, $l^*$. We assume there is no coordination between students in their decision regarding time allocation. Thus, any peer effects in attendance are random and not the result of collective action. Marginal products may vary from person to person. Assume that students have accurate information the relative marginal productivities of the inputs in their education production function. This assumption may be more likely to hold for high school students, as they may be more experienced learners. Assume also that the marginal products of study time

[2]The assumption of constant marginal utility of leisure is not crucial. Here is an example where we relax this assumption. Consider the following production function: $s = s(c, h, a)$. Suppose the utility function takes the following form: $U = u(s, l) = s(c, h, a) + \gamma\sqrt{l} = \alpha\sqrt{c} + \beta\sqrt{h} + \gamma\sqrt{l}$. Maximizing utility under the time constraint gives the following optimal time allocation: $\{c^*, h^*, l^*\} = \{\frac{\alpha^2}{\alpha^2+\beta^2+\gamma^2}, \frac{\beta^2}{\alpha^2+\beta^2+\gamma^2}, \frac{\gamma^2}{\alpha^2+\beta^2+\gamma^2}\}$.

in class and out of class are positive but exhibit diminishing returns and are independent of each other and of ability: $\frac{\partial s}{\partial c} := mpc > 0$, $\frac{\partial s}{\partial h} := mph > 0$, $\frac{\partial^2 s}{\partial c^2} < 0$, $\frac{\partial^2 s}{\partial h^2} < 0$, $\frac{\partial^2 s}{\partial c \partial h} = 0$, $\frac{\partial^2 s}{\partial c \partial a} = 0$, and $\frac{\partial^2 s}{\partial c \partial a} = 0$. We also assume $\frac{\partial s}{\partial a} > 0$. We can represent diagrammatically the solution to the problem of the utility-maximizing student.

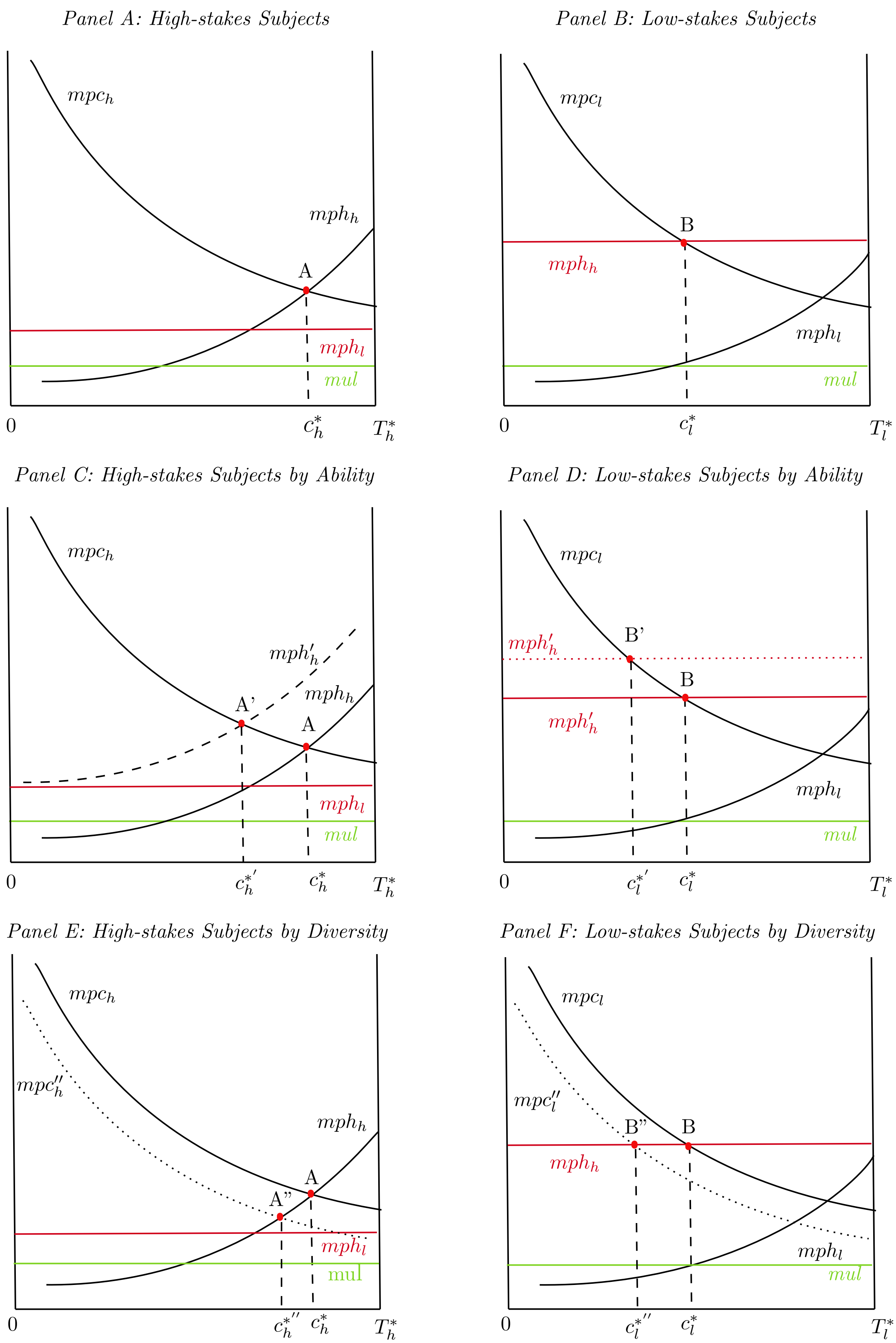
Figure 1 Panel A shows students' optimization problem for high-stakes subjects. The figure depicts the marginal productivity of study time in the classroom, $mpc_h$, and the marginal productivity of out-of-class learning, $mph_h$, for high-stakes subjects. $T_h^*$ represents the total amount of classroom learning for high-stakes subjects determined by schools. In Figure 1 Panel A, the $mpc_h$ ($mph_h$) is decreasing (increasing) in the amount of time spent in the classroom. Each student faces constant marginal productivity of out-of-school learning for low-stakes subjects $mph_l$, and constant marginal utility of leisure $mul$.

A utility-maximizing student will optimize at point $A$ in Figure 1 Panel A, choosing to attend $c_h^*$ class periods. Therefore, the student would choose to take $T_h^* - c_h^* = o_h^*$ class absences and spend the time in out-of-class learning activities (to the right of optimization point A, where $mph_h \geq mph_l$ and $mph_h \geq mul$). If the intersection point between $mph_h$ and $mpc_h$, A, lies outside the graph, students would not take any absences. If $mph_l$ is above point A, students will choose to take absences and devote out-of-class learning time to low-stakes subjects ($mph_l \geq mph_h$ and $mph_l \geq mul$).

Figure 1 Panel B depicts students' optimization problem for low-stakes subjects. The figure plots the marginal productivity of study time in class, $mpc_l$, and the marginal productivity of home study, $mph_l$, for low-stakes subjects. $T_l^*$ represents the school-prescribed time in class for low-stakes subjects. The marginal productivity of out-of-school learning for high-stakes subjects, $mph_h$, and the marginal utility of leisure, $mul$, are independent of the time spent in class. In this case, the utility-maximizing student optimally chooses to attend $c_l^*$ class periods and take $T_l^* - c_l^*$ class absences to study for high-stakes subjects outside the classroom. If the intersection point between $mph_l$ and $mpc_l$ lies outside the graph, the student will not take any absences. If the intersection point between $mph_l$ and $mpc_l$ lies above the $mph_h$ line, the student will take absences. Whether the student would spend this time on out-of-school activities related to low-stakes subjects or devote it to leisure depends on the relative positions of $mph_h$ and $mul$.

We relax now our assumption that factor inputs are independent and allow the marginal product of out-of-school learning to be positively correlated with ability. In particular, suppose the marginal productivity of out-of-school learning for high-stakes subjects increases with ability, motivation, and effort $a$. Figure 1 Panel C shows the time-optimization problem for high-stakes subjects.

Figure 1: A Model of Time Allocation



For a high level of ability $a$, the marginal productivity of out-of-school learning for high-stakes subject increases to $mph'_h$. At the new intersection point between $mph'_h$ and $mpc_h$, A', students would choose to attend a lower number of classes, $c_h^{*'}$, and take a greater number of absences $T_h^* - c_h^{*'} > T_h^* - c_h^*$. Figure 1 Panel D displays the optimization problem for low-stakes subjects. As the marginal productivity of out-of-school learning $mph'_h$ increases, a utility-maximizing student would attend fewer classes, $c_l^{*'} < c_l^*$, and take a higher number of absences $T_l^* - c_l^{*'} > T_l^* - c_l^*$.

This framework can be extended to consider the determining factors of in-school and out-of-school marginal learning productivity. As an example, the literature suggests that classroom academic diversity may be a key determining factor of in-school learning productivity (Aucejo et al., 2021). If academic diversity in the classroom is negatively associated with learning productivity at school, as the literature finds, the model suggests that students in more academically diverse peer environments may miss school more often than students in more academically homogeneous peer groups, ceteris paribus. In other words, if the *mpc* curve of students in more academically diverse classrooms is to the left of the *mpc* curve of students in less academically diverse classrooms, the optimal class attendance of the former will be lower than that of the latter. This scenario is depicted in Panels E and F of Figure 1. When classroom diversity decreases learning productivity in class ($mpc$ shifts to $mpc''$), optimal class attendance decreases to $c_h^{*''}$ and $c_l^{*''}$ for high- and low-stakes subjects, respectively.

In summary, this model shows that students who maximize their utility with respect to in-classroom and outside-classroom learning, as well as leisure, may optimally choose to take a positive number of absences. This optimal number of absences may be positively correlated with student ability as the marginal productivity of out-of-classroom learning increases with $a$. The optimal number of absences may also be negatively associated with classroom academic diversity to the extend that diversity translates to lower classroom learning productivity. Whether students devote their-out-of-school time to independent learning activities or leisure depends on their marginal productivity of out-of-class learning and their marginal utility of leisure.

The theoretical arguments presented here postulate a set of empirical research questions: Does relaxing the class attendance requirement for high achievers change their attendance and performance in high- and low-stakes subjects? Do higher-performing students in academically diverse peer groups miss more classes when allowed to? We hypothesize that when higher-achieving students are allowed more autonomy, they skip class more often. We also hypothesize that higher-performing students in academically diverse classrooms skip more classes when permitted to. Our empirical analysis investigates these hypotheses.

# 3 Institutional Framework

## 3.1 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] 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). Attendance policies are strict. Until the end of the 2005-06 school year, every student was allowed 50 unexcused and 64 excused class absences in a year. An absence is a missed school period. Missing one day of school equals as many absences as the number of school periods in that day. The penalty for exceeding the number of allowed absences is to repeat the grade.[4] Absences can be excused only by a doctor or a guardian. Usually families must submit a doctor's note to the school principal that justifies a health condition that prevented the student from attending class. Only whole days of absence can be excused. For example, if a student goes to school late in the morning or if they decide to leave school at midday, their absences that day cannot be excused. 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. The assignment of students and teachers to classrooms in each school is random.[5] 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).[6] 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 randomized peer influences, which we show later. Teachers rotate between classrooms to teach classes in their subject.

High school students have little choice regarding the classes they take. A typical high school student is required to take 10 grade-specific subjects and three or four compulsory electives. Around 60% of students' instructional time is spent in compulsory core education classes. General ed-

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

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

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

[6]See Government Gazette of the Hellenic Republic 167 A/1566/1985. See also Education Ministry Bulletin of the Hellenic Republic 100749/Γ2/17-09-07.

ucation subjects include Greek Language, Math, Physics, and History. The remaining 40% of instructional time is allocated to specialization classes. Students choose a specialization track at the beginning of the $11^{\text{th}}$ and $12^{\text{th}}$ grade. The available tracks are Classics, Science, and Information Technology (IT). Each track requires that students take different sets of specialization subjects. For the Classics track, specialization subjects 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.[7] The average university admission score of all admitted candidates in a university degree program is used to infer that degree's *quality/selectiveness*.[8]

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

## 3.2 The Increased Autonomy Policy

Near the end of the 2005-06 school year, the Ministry of Education implemented a policy change intended to encourage students' autonomy. The new policy provided eligible students with 50 additional excused class absences. Every student who had received a raw GPA higher than 75% the year before was eligible to take up more absences in the current year. The rationale for this policy change was that targeted students—those with a prior GPA above 15/20—would 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).

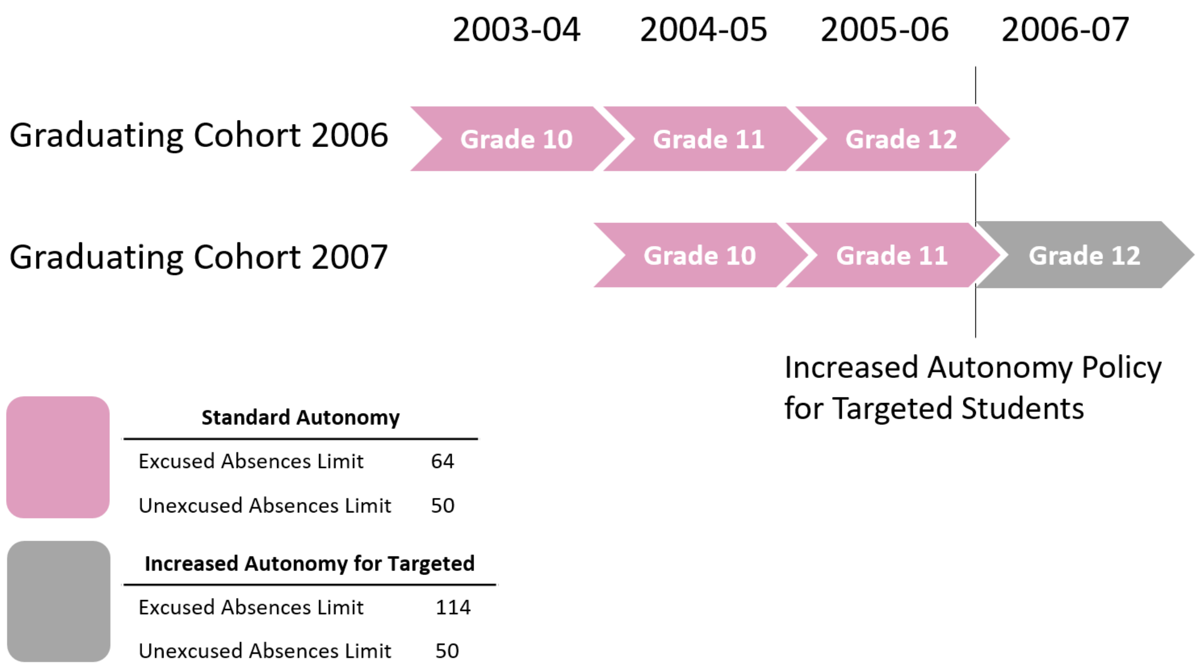
We consider the cohort graduating high school in 2006 to be the control group and the cohort

[7]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 the student's teacher.

[8]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.

graduating in 2007 the treated group. The cohort graduating in 2007 was subject to the increased autonomy policy, introduced in 2006-07, in grade 12. Students could not manipulate in advance their eligibility for the increased autonomy policy, since the autonomy policy was unanticipated and eligibility depends on prior-year GPA. Figure 2 displays the timeline of the reform and the affected cohorts.

Figure 2: Timeline of the Introduction of the Increased Autonomy Policy




*Notes:* The figure shows the timing of the introduction of the increased autonomy policy. Students in the cohort graduating in 2006 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 excused absences in a year before retention.

## 4 Data Sources and Description

We investigate the effect of the increased autonomy policy on education outcomes by combining data from two administrative sources. First, we use manually collected demograhic, classroom assignment, class and track enrollment, attendance, and transcript data from a sample of 107 public high schools we visited across Greece, corresponding to roughly 10% of the public high schools in the nation.[9] Our data include student records from all three grades of high school

[9] Goulas et al. (2018) and Goulas and Megalokonomou (2021) show that the sampled schools are representative of the general school population.

for the cohorts graduating between 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. Second, we have 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 the setting studied. Figure 1 shows that students tend to increase their school absences substantially between grade 11 and grade 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 prepare for university admission exams and may potentially wish to substitute time at school for out-of-school study time. Alternatively, students may choose to miss school more frequently in grade 12 to allocate time to activities unrelated to furthering their human capital or improving their chances of university admission. This setup of increased demand for autonomy at the end of high school allows for estimation of the effect of relaxing the school-time allocation constraint on attendance and performance.

Figure 2 shows the distribution of total class absences by cohort. Vertical lines indicate the upper absences thresholds before retention under the strict (old limit) and the increased autonomy regime (new limit). 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 careful to not exceed the upper absences limit.

Figure 3 plots the distribution of excused absences (Panel A) and unexcused absences (Panel B) for the unaffected and affected cohorts. The increased autonomy policy altered the number of excused absences students are allowed to take but left the upper number of unexcused absences unchanged. The distribution of excused absences also shows a shift to the right for the 2007 cohort (affected cohort) relative to the 2006 cohort (unaffected cohort), while the ones of unexcused absences for the unaffected and affected cohorts are similar. This reassures us that the shifts in the distributions of total and excused absences that we observe in Figure 2 and Figure 3 (Panel A) can be attributed to the increased autonomy policy.

Table 1 presents summary statistics of the analytic data. The average prior-year GPA—the running variable for targeted status—is 71.17 out of 100, and 37% of students are targeted.[10] Each student takes on average 64 class absences a year, 30 of which are excused and the rest unexcused.

[10] Figure A1 plots the full distribution of prior-year GPA, the running variable for increased autonomy eligibility.

The average standard deviation of prior-year GPA in the classroom is 13.36 standard deviations and the average interquartile range (IQR) of prior-year GPA in the classroom is 10.61 standard deviation. 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. Performance on high-stakes subjects in grade 12 matters for university admission. Low-stakes subjects include general education Math, Physics, and History. High-stakes subjects include general education 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. On average, admitted students enroll in a degree program between the $55^{\text{th}}$ and $56^{\text{th}}$ percentiles of quality.

# 5 Identification Strategy

In this section, we empirically investigate the impact of providing increased autonomy to targeted students, ceteris paribus. To investigate the average effect of autonomy on targeted students, we employ a difference-in-difference-in-differences framework. Table 1 summarizes our identification strategy. In this triple-differences framework, there are three dimensions of comparison. The first is between grade 11 and grade 12 of students on both sides of the eligibility cutoff in either the 2006 or 2007 graduating cohorts. The second dimension of comparison is between students above (targeted) and below (non-targeted) the eligibility cutoff 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 cohort of 2006 (standard autonomy policy) and the graduating cohort of 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

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 + ‘$07_{12}$ + $T_{12}$ | $(T_{12} - NT_{12})$ + 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 + ‘$06_{12}$ + $NT_{12}$+ D | | |
| | | 11 | $I^{NT}_{‘06}$ | | | |

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

of the increased autonomy policy on targeted students is captured by $D$. We identify the impact of increased autonomy using the following specification:

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

where outcome $y_{icg}$ includes total absences, excused absences, unexcused absences, performance in low-stakes subjects, and performance in high-stakes subjects for student $i$ in classroom $c$. Scores are standardized at the school-grade-subject level. We control for student fixed effects in $\eta_i$. Student fixed effects allow us to account for student unobservables; these could include ability, family background, and resources. We control for classroom characteristics, which include classroom diversity, measured by the standard deviation of prior-year GPA in the classroom; class size;[11] and the proportion of female peers in the classroom in $X_c$.[12] Variable $G_g$ is an indicator for grade 12. $T_{ig}$ 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). Variable $C_i$ is an indicator for the treatment cohort of 2007. Coefficient $\beta_1$ corresponds to the effect of the increased autonomy policy. Standard errors

[11] We include indicators for each class size value for estimation precision, given that class size is positively associated with academic diversity in the classroom ($\rho = 0.109$, $p < 0.001$).

[12] 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).

are clustered at the school level to allow for heteroskedasticity and serial correlation in student outcomes within each school. The identification assumption is that in the absence of the autonomy policy, the differences in outcomes between targeted students in the treated cohort and not-targeted ones in the same cohort would otherwise have trended similarly, on average, to the differences in outcomes between those two groups in the control cohort.[13]

In the natural experiment studied, assignment to treatment does not coincide with receiving treatment.[14] The fact that it is possible for non-targeted students to take on more absences, even when they are not eligible for increased autonomy in the 2006-07 school year, suggests that our estimates constitute an intention-to-treat effect and a lower bound for the effect of increased autonomy on treated students.

# 6 Validity of the Identification Strategy

We are able to interpret the DDD estimator as the causal 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 may be 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 but also to differences in their peer environments, which may drive differential demand for autonomy. Moreover, if students above and below the eligibility cutoff 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.

The second assumption is that the trend in student attendance and performance between grades

[13]Because specification (4) controls for student fixed effects, coefficient $\beta_5$ captures the outcome component associated with students whose eligibility/targeted status changes between grade 11 and grade 12 (roughly 8%).

[14]As the autonomy policy provides an exogenous eligibility cutoff at prior-year GPA of 75%, one might be inclined to consider a *regression discontinuity design* (RDD) identification framework. However, an RDD approach ignores substantial impacts of the policy change on eligible students far from the eligibility cutoff. Appendix Section 14 provides an empirical investigation of the effects of the increased autonomy policy around the eligibility cutoff, in which we find limited evidence of the effects of the autonomy policy around the eligibility cutoff.

would be the same in both the control and treatment cohorts in the absence of the increased autonomy policy. If this common trends assumption is violated, the difference in attendance and performance observed in the year the increased autonomy policy was introduced may not be fully attributable to the autonomy policy.

A third assumption requires that treatment 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 cutoff, 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 finish above the eligibility cutoff. This manipulation of the running variable is not possible in the institutional setting we exploit in this study. In our quasi-experimental environment, 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.[15]

## 6.1 Randomized Peer Influences

One might worry that peer characteristics influence not only individual performance—and consequently targeted status—but also impact the propensity to use the increased autonomy policy. Being targeted at grade 12 may be associated with being placed with higher-quality peers' in grade 10. At the same time, being among higher-performing peers may lower the need for autonomy due to potentially lower classroom disruption. Our setting allows us to mitigate concerns that differences in the need for autonomy between cohorts and between targeted and non-targeted students may be associated with differences in peer group characteristics. A key feature of the setting studied is consistent, quasi-random peer group formation (Goulas et al., 2020). Quasi-random peer group formation allows us to causally interpret the estimated effect of increased autonomy on targeted students to the extent that targeted students in the treated year do not differ from non-targeted students or control cohort students in classroom-specific influences.

Quasi-random assignment to classrooms at the beginning of high school guarantees that peer groups remain stable throughout high school and mitigates concerns regarding potential bias in the estimated effect of increased autonomy on the targeted. Table 2 uses attendance and performance information from grade 10, the instance closest to peer group formation, to provide evidence that the alphabetical assignment to classrooms is practically random.[16] Specifically, Table 2 shows that

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

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

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 eligibility cutoff for increased autonomy.

## 6.2 Common Trends

In this section, we empirically investigate the common trends assumption required for the identification of the autonomy policy effect. We compare mean excused and unexcused absences and performance in high- and low-stakes subjects in grades 10, 11, and 12 of students above (targeted) and below (non-targeted) the eligibility cutoff in the graduating cohorts of 2006 and 2007.[17] We estimate the following specification:

$$y_{ig\tau} = \alpha + \beta_{g\tau} C_i + \gamma_{11\tau} C_i + \kappa_{g\tau} + \lambda_i + \varepsilon_{ig\tau}, \tag{5}$$

where $g \in \{10, 11, 12\}$ and $\tau \in \{Targeted, Non-targeted\}$. Variable $y_{ig\tau}$ denotes the outcome of student $i$ (i.e., attendance or performance) in grade $g$ of targeted status $\tau$. Vector $\kappa_{g\tau}$ represents fixed effects at grade $g$ and targeted status $\tau$ (i.e., the Kronecker product of the set of grades with the set of targeted status values). Coefficients $\beta_{g\tau}$ reflect the mean difference between control and treatment cohort for targeted and non-targeted students in each grade. The mean difference between control and treatment cohort for students in each treatment status in $11^{\text{th}}$-grade outcomes, captured by $\gamma_{11\tau}$, is the benchmark. We control for student fixed effects, $\lambda_i$. Standard errors are clustered at the school level. We plot the estimates for absences, excused and unexcused, as well as for performance, in high- and low-stakes subjects, along with the corresponding confidence intervals in Figures 4 and 5, respectively. We find that the pre-treatment trajectories of performance measures and school absences of non-targeted (targeted) students in the 2006 cohort are similar to those of non-targeted (targeted) students in the 2007 cohort.

### 6.2.1 Falsification

We perform falsification tests of the effect of the increased autonomy policy on attendance and performance. The autonomy policy relaxed the upper limit of excused absences but not the limit of unexcused absences. Thus, the autonomy policy is expected to influence student excused absences

[17] 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.

but not unexcused absences. Figure 4 provides evidence of the common trend assumption for both excused and unexcused absences. This allows us to identify the impact of the increased autonomy policy on excused absences, while unexcused absences serve as a placebo.

We also perform a falsification investigation of performance in subjects of different stakes. One might expect increased autonomy to improve the performance of targeted students relative to non-targeted students in high-stakes subjects but less so in low-stakes subjects. Figure 5 shows that the common trend assumption is satisfied for both high- and low-stakes subjects, which allows us to explore the impact of increased autonomy on both types of subjects. Investigation of the effect of increased autonomy on low-stakes subjects serves as a placebo.

# 7 Results

Table 3 presents our main results. We report the impact of the increased autonomy policy on total class absences, excused and 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 $T_{ig} \times C_i \times G_g$, $\beta_1$, in specification (4). 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 $C_i \times G_g$, $\beta_2$, in specification (4). Our estimates account for student fixed effects, grade 12 indicator, an indicator taking the value of one for students with prior performance above the eligibility cutoff (targeted status indicator), an interaction between the targeted status indicator and the grade 12 indicator, and an interaction between the targeted status indicator and the treatment cohort indicator. We also control for classroom size (using indicators for each value), classroom diversity, and the proportion of female peers in the classroom (leave-out mean).

Targeted students in the treated cohort decreased their attendance and improved their performance in high-stakes subjects as a result of the increased autonomy policy. In particular, targeted students increased their total (excused) absences by 0.09 (0.13) standard deviations—roughly four (three) additional classes relative to non-targeted students during the year the increased autonomy policy was in effect. Targeted students' performance in high-stakes subjects increased by 0.07 standard deviations due to the increased autonomy policy. Targeted students' performance in low-stakes subjects and overall GPA remained unaffected by the increased autonomy policy.

The policy is also associated with an increase of roughly 0.12 standard deviations in non-targeted students' excused absences (fewer than 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.

# 8 Heterogeneous Effects by Classroom Diversity

In this section, we examine heterogeneity in the effects of the increased autonomy policy using quasi-random variation in ability diversity in the classroom. The policy allows higher-performing students to distance themselves from settings with lower learning productivity. Academically more diverse classrooms may offer lower learning productivity (Aucejo et al., 2021). Thus, one might expect that targeted students could exploit the increased autonomy policy more when they are in more academically diverse classrooms. We investigate whether targeted students skip class more when randomly assigned to a more diverse class and the impact on their performance. We use the standard deviation of previous-year overall GPA as a measure of classroom academic diversity. We also consider an alternative measure of classroom academic diversity, the interquartile range (IQR) of previous-year overall GPA.[18] Figure A2 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 (4) with breakout binary indicator variables of targeted status in each quartile of classroom academic diversity, $T_{ig}^q$, where $q \in \{1, 2, 3, 4\}$. 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. Our specification is as follows:

$$y_{icg} = \beta_0 + \sum_{q=1}^{4} \beta_1^q T_{ig}^q \times C_i \times G_g + \beta_2 C_i \times G_g + \sum_{q=1}^{4} \beta_3^q T_{ig}^q \times G_g + \sum_{q=1}^{4} \beta_4^q T_{ig}^q \times C_i + \sum_{q=1}^{4} \beta_5^q T_{ig}^q + \beta_6 G_g + X_c + \eta_i + \varepsilon_{icg}, \tag{6}$$

where $\beta_1^q$ captures the impact of the increased autonomy on targeted students in the treated cohort in classrooms in diversity quartile $q$. Table 4 shows that higher quartiles of classroom academic diversity are associated with lower attendance 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

[18] We define the IQR of previous-year GPA as follows: $IQR = \frac{Q_3 - Q_1}{2}$, where $Q_1$ and $Q_3$ represent, respectively, the first and third quartiles of previous-year overall GPA.

classes relative to non-targeted students. The attendance of targeted students in classrooms with academic diversity in the bottom quartile of the diversity distribution is statistically similar to the attendance of non-targeted students.

Classroom diversity is associated with average performance among targeted and non-targeted student groups.[19] 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 (7) to control directly for prior-year performance (GPA) in a robustness investigation. Our results on Table A1 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.

We 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 A2 shows the estimated heterogeneous effects of the increased autonomy policy on targeted students using the IQR, and the results show patterns that are substantially similar to the results using the standard deviation of prior performance in the classroom.

# 9 Robustness Check

One might worry that overestimation of the outcomes trajectory between grades 11 and 12 of the counterfactual could bias the estimated effect of the increased autonomy policy (Cunningham, 2021). Differences in the outcomes trajectories between grades 11 and 12 may be driven by differences in the baseline outcome values in grade 11. For example, students with different absences levels in grade 11 might differ in their propensity to demand further autonomy and increase their absences in grade 12.[20] This suggests that 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. Consequently, if students in the control cohort have lower baseline absences level in grade 11 than students in the treatment cohort, it is likely that the estimated effect of the increased autonomy policy on

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

[20]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.

absences will be downward biased.[21]

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 A3 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 cutoff, 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.[22] We apply equal weighting, which averages over multiple records in the control cohort for each record in the treatment cohort.[23]

Table 5 shows the estimates from our matching approach. To alleviate concerns about bias in

[21]Table A3 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.

[22]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.

[23]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 A4 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.

the matching estimator due to selective matching, we employ inverse probability weights (IPWs) to control for differences in obtaining matches for students with different pre-treatment characteristics. The increased autonomy policy is found to be associated with lower attendance (more excused absences) and higher performance in high-stakes subjects for targeted students. In particular, increased autonomy is found to be associated with an increase in excused absences of 3.762 (compared with 3.094 from DDD) and an increase in performance in high-stakes subjects by 0.151 standard deviations (compared with 0.073 from DDD).[24]

# 10 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. 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.[25] We estimate the following value-added model:

$$y_{iycts} = \beta_0 + \beta_1 C_{iy} + T_i + W_i + X_c + + \zeta_y + \theta_t + \eta_s + \varepsilon_{iycts}, \tag{7}$$

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 (proxied by the standard deviation of prior-year GPA). We account for year (equivalent to cohort), track, and school fixed effects in $\zeta_y$, $\theta_t$, and $\eta_s$, respectively. Coefficient $\beta_1$ captures the effect of the autonomy policy on targeted students in the treatment cohort (graduating cohort of 2007). Standard errors are clustered at the school level.

[24] 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 A5 presents estimates from a matching approach with the expanded matching criteria. The estimated effect of increased autonomy on excused absences and performance in high-stakes subjects for targeted students remains substantial and statistically significant.

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

If students act strategically, using their autonomy to improve their performance in high-stakes subjects, one might expect a positive impact of the increased autonomy policy on targeted students' university admission scores. Table 6 shows our estimates. We find significantly higher university admission scores among targeted students in the cohort with increased autonomy 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 admission score of 0.13 standard deviations. Table 6 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.

# 11 Conclusion

Understanding how autonomy affects education outcomes is essential in assessing the academic standing of students as schools reopen following the COVID-19 pandemic. Differential levels of learning productivity away from traditional schools during the pandemic are expected to widen educational gaps and render school settings more academically diverse (Kuhfeld et al., 2020; Raymond et al., 2020; Pier et al., 2021). At the same time, strategies that aim to recover pandemic-related learning losses rely on assumptions about students' learning productivity at the traditional school (Raymond, 2021). Academic diversity may lead to lower in-school learning productivity, especially for high performers. Quantifying the impact of autonomy on student performance can inform learning recovery programs and policies that aim to bring all students up to speed following the pandemic. Additionally, determining the conditions under which some students can learn effectively away from school can widen the array of pandemic recovery policies that allow resources to be allocated where they are most needed and help policymakers predict their success.

In this paper, estimate the causal effects of autonomy on short- (i.e., high school performance) and longer-term outcomes (i.e., university admission score). We exploit an innovative nationwide policy in Greece that allowed higher-performing grade 12 students to skip 30 percent more classes without penalty. The policy targeted students with a prior-year GPA above 75%. Targeted students could decide how to spend their time with increased flexibility.

We deploy a triple-differences identification strategy that compares student attendance and performance trajectories between grades of targeted and non-targeted students in a control cohort (graduating high school in 2006) and a treated cohort (graduating in 2007). Student fixed effects allow us to control for unobserved factors that influence student performance and propensity to attend class. We are also able to enhance the credibility of this identification strategy by examining the existence of common trends between treated and control groups. A complementary estimation approach through matching supports the validity of the estimated autonomy effects.

When grade 12 students are offered discretion over their attendance, their performance on high-stakes exams increases substantially. This increase in high-stakes exam performance due to autonomy is equivalent to reducing class size by 0.23 standard deviations (or 8%) (Krueger, 2003);[26] having a teacher of quality of 1.4 standard deviations above average (Carrell and West, 2010); or attending 50 additional days of schooling (Hanushek et al., 2012). In contrast, their performance in low-stakes subjects is unaffected. Higher-performing students also obtain a higher university admission score and are admitted to university degree programs of higher quality when they are allowed to skip more classes. Further investigation reveals that higher-performing students have higher performance gains associated with autonomy in more academically diverse classrooms. This suggests that learning productivity may be lower in academically heterogeneous contexts. Lower-performing students' scores remain statistically unchanged when higher-performing students exercise autonomy.

Our results demonstrate that autonomy in the form of relaxed school attendance for higher-performing 12$^{\text{th}}$-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.

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

# References


Arulampalam, W., R. Naylor, and J. Smith (2012). Am I missing something? *Economics of Education Review 31*(4), pp. 363–375.

Aucejo, E. M., P. Coate, J. Fruehwirth, S. Kelly, and Z. Mozenter (2021). Teacher effectiveness and classroom composition: Understanding match effects in the classroom. *CEPR Discussion Paper Number DP13166*.

Carrell, S. E. and J. E. West (2010). Does professor quality matter? Evidence from random assignment of students to professors. *Journal of Political Economy 118*(3), 409–432.

Caviglia, H. (2006). Attendance and achievement in economics: Investigating the impact of attendance policies and absentee rates on student performance. *Journal of Economics and Finance Education 4*(2), pp. 470–476.

Cohn, E. and E. Johnson (2006). Class Attendance and Performance in Principles of Economics. *Education Economics 14*(2), 211–233.

Cunningham, S. (2021). *Causal inference: The mixtape.* New Haven, CT: Yale University Press.

Dehejia, R. H. and S. Wahba (1999). Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs. *Journal of the American Statistical Association 94*(448), 1053–1062.

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

Gaete, G. (2018). Follow the leader: Student strikes, school absenteeism and long-term implications for education outcomes. Mimeo, 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. *American Educational Research Journal 47*(2), 434–465.

Goulas, S., S. Griselda, and R. Megalokonomou (2020). Comparative advantage and gender gap in STEM. *IZA Discussion Paper Number 13313*.

Goulas, S. and R. Megalokonomou (2020). School attendance during a pandemic. *Economics Letters 193*, 109275.

Goulas, S. and R. Megalokonomou (2021). Knowing who you actually are: The effect of feedback on short-and longer-term outcomes. *Journal of Economic Behavior & Organization 183*, 589–615.

Goulas, S., R. Megalokonomou, and Y. Zhang (2018). Does the girl next door affect your academic outcomes and career choices? *IZA Discussion Paper Number 11910*.

Guryan, J., K. Kroft, and M. J. Notowidigdo (2009). Peer effects in the workplace: Evidence from random groupings in professional golf tournaments. *American Economic Journal: Applied Economics 1*(4), 34–68.

Hanushek, E. A., P. E. Peterson, and L. Woessmann (2012). Achievement growth: International and US state trends in student performance. pepg report no.: 12-03. *Program on Education Policy and Governance, Harvard University*.

Hennessy, S., W. B. Bilker, J. A. Berlin, and B. L. Strom (1999). Factors influencing the optimal control-to-case ratio in matched case-control studies. *American journal of epidemiology 149*(2), 195–197.

Kapoor, S., M. Oosterveen, and D. Webbink (2021). The price of forced attendance. *Journal of Applied Econometrics 36*(2), 209–227.

Krohn, G. A. and C. M. O'Connor (2005). Student effort and performance over the semester. *Journal of Economic Education 36*(1), 3–28.

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

Kuhfeld, M., J. Soland, B. Tarasawa, A. Johnson, E. Ruzek, and J. Liu (2020). Projecting the potential impact of COVID-19 school closures on academic achievement. *Educational Researcher 49*(8), 549–565.

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

Lavy, V. and R. Megalokonomou (2019, June). Persistency in teachers' grading bias and effects on longer-term outcomes: University admissions exams and choice of field of study. *NBER Working Paper* (26021).

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

Liu, J., M. Lee, and S. Gershenson (2021). The short-and long-run impacts of secondary school absences. *Journal of Public Economics 199*, 104441.

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

Martins, P. S. and I. Walker (2006). Student achievement and university classes: Effects of attendance, size, peers, and teachers. *IZA Discussion Paper Number 2490*.

Moore, R. (2006). Class attendance: How students' attitudes about attendance relate to their academic performance in introductory science classes. *Research and Teaching in Developmental Education 23*(1), 19–33.

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 (2015). *Education at a Glance 2015.* OECD Publishing, Paris.

OECD (2018). *Education for a Bright Future in Greece.* OECD Publishing, Paris.

OECD (2020, March). Education responses to COVID-19: Embracing digital learning and online collaboration. *Technical Report*.

Pier, L., H. J. Hough, M. Christian, N. Bookman, B. Wilkenfeld, and R. Miller (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. (2021). Learning losses - what to do about them. Technical report, Hoover Institution, Stanford University, Stanford, CA.

Raymond, M. E., C. Han, S. Goulas, W. F. Lee, and C. Waeiss (2020). Estimates of learning loss in the 2019-2020 school year. *Center for Research on Education Outcomes at Stanford University*.

Reis, S. M. and J. S. Renzulli (2010). Is there still a need for gifted education? an examination of current research. *Learning and Individual Differences 20*(4), 308–317.

Richardson, Hannah and Sellgren, Katherine (2020, March). Which children are still going to school? *BBC News Education Reporters*.

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

Scott, M. M., J. Shakesprere, and K. Porter (2020, October). Why schools should embrace flexibility and innovation beyond COVID-19. Technical report, Urban Institute, Washington, D.C.

Singer, N. (2021, April). Online schools are here to stay, even after the pandemic. *The New York Times*.

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

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science: a Review Journal of the Institute of Mathematical Statistics 25*(1), 1.

World Bank (2020, July). How countries are using edtech (including online learning, radio, television, texting) to support access to remote learning during the COVID-19 pandemic. *Technical Report*.

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal 45*(1), 166–183.

Table 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 |

*Notes:* This table reports summary statistics for the prior-year GPA (the running variable for targeted status) and targeted status (Panel A), attendance (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 previous-year GPA across all students in the classroom. [2] Classroom IQR is the interquartile range of previous-year GPA in the classroom. Observations: 24,542.

## Table 2: Random Assignment of Students into Classrooms

| | Av. Excused Absences | Av. Unexcused Absences | Av. GPA | GPA Diversity | Targeted Share | Av. Excused Absences of Targeted | Av. Unexcused Absences of Targeted | Av. GPA of Targeted | Av. Excused Absences of Non-targeted | Av. Unexcused Absences of Non-targeted | Av. GPA of Non-targeted |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
| Class Number=2 | -0.027 | 1.739 | -0.011 | 0.002 | -0.014 | -0.212 | -0.266 | -0.012 | 0.185 | 2.005 | 0.001 |
| | (0.508) | (1.297) | (0.028) | (0.012) | (0.012) | (0.290) | (0.280) | (0.016) | (0.468) | (1.316) | (0.015) |
| Class Number=3 | -0.567 | 1.478 | 0.001 | 0.017 | -0.008 | -0.288 | 0.015 | 0.007 | -0.279 | 1.464 | -0.006 |
| | (0.490) | (1.463) | (0.030) | (0.015) | (0.014) | (0.268) | (0.291) | (0.017) | (0.478) | (1.484) | (0.017) |
| Class Number=4 | -0.419 | 0.455 | 0.014 | -0.010 | 0.004 | -0.495 | -0.104 | -0.004 | 0.076 | 0.560 | 0.018 |
| | (0.634) | (1.563) | (0.041) | (0.019) | (0.019) | (0.406) | (0.397) | (0.023) | (0.582) | (1.591) | (0.022) |
| Class Number=5 | 0.676 | 2.282 | -0.027 | -0.036 | -0.029 | 0.218 | -0.322 | -0.033 | 0.458 | 2.604 | 0.006 |
| | (0.864) | (1.594) | (0.068) | (0.031) | (0.035) | (0.612) | (0.725) | (0.044) | (0.928) | (1.941) | (0.032) |
| Class Number=6 | -1.349 | 2.642 | -0.062 | -0.025 | -0.057 | -0.520 | -0.467 | -0.055 | -0.828 | 3.109 | -0.007 |
| | (1.747) | (3.672) | (0.060) | (0.048) | (0.035) | (0.653) | (0.892) | (0.044) | (1.605) | (4.288) | (0.037) |
| Observations | 598 | 598 | 598 | 598 | 598 | 598 | 598 | 598 | 598 | 598 | 598 |
| Y Mean | 24.89 | 33.72 | 0.00 | 0.98 | 0.36 | 8.39 | 8.88 | 0.37 | 16.50 | 24.85 | -0.36 |
| Y Standard Deviation | 14.45 | 14.16 | 0.25 | 0.12 | 0.15 | 7.68 | 4.89 | 0.16 | 10.02 | 14.04 | 0.17 |
| School x Year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| F-stat. for joint significance | 1.27 | 0.74 | 0.38 | 0.80 | 0.97 | 0.56 | 0.27 | 0.82 | 0.70 | 0.77 | 0.30 |
| P-value for joint significance | 0.284 | 0.596 | 0.861 | 0.554 | 0.438 | 0.734 | 0.931 | 0.536 | 0.621 | 0.573 | 0.912 |

*Notes:* 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. 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. All specifications include School × Year FE. Standard errors clustered at the school level are reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 3: Effect of the Increased Autonomy Policy on Attendance and Performance

| | Class Absences | | | Performance | |
|---|---|---|---|---|---|
| | Total | Excused | Unexcused | High-stakes Subjects | Low-stakes Subjects |
| Targeted in Treated Cohort [1] *relative to non-targeted* | 3.908** | 3.094*** | 0.814 | 0.073*** | -0.015 |
| | (1.550) | (1.049) | (0.981) | (0.018) | (0.020) |
| Non-targeted in Treated Cohort [2] *relative to control cohort* | 1.516 | 2.812** | -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 | 64.22 | 29.38 | 34.84 | 0.01 | 0.01 |
| Y Standard Deviation | 43.57 | 23.88 | 34.44 | 0.88 | 0.85 |
| Student FE | ✓ | ✓ | ✓ | ✓ | ✓ |
| P-value for H0: [1] + [2] = 0 | 0.01 | 0.00 | 0.60 | 0.00 | 0.97 |

*Notes:* 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. 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. 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. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 4: Heterogeneous Effects of Increased Autonomy Policy by Classroom Diversity (Standard Deviation of Prior Performance)

| | Class Absences | | | Performance | |
|---|---|---|---|---|---|
| | Total | Excused | Unexcused | High-stakes Subjects | Low-stakes Subjects |
| Targeted in Treated | 8.772*** | 7.227*** | 1.545 | 0.035 | 0.006 |
| Cohort × Class SD Top Quartile | (3.069) | (2.328) | (1.459) | (0.022) | (0.026) |
| Targeted in Treated | 4.892* | 4.155** | 0.738 | 0.075*** | -0.018 |
| Cohort × Class SD Second Quartile | (2.716) | (1.980) | (1.241) | (0.025) | (0.032) |
| Targeted in Treated | 2.744 | 0.741 | 2.004 | 0.085*** | -0.029 |
| Cohort × Class SD Third Quartile | (2.410) | (1.802) | (1.212) | (0.024) | (0.030) |
| Targeted in Treated | -0.121 | 0.813 | -0.934 | 0.094*** | -0.017 |
| Cohort × Class SD Bottom Quartile | (3.340) | (2.496) | (1.465) | (0.027) | (0.028) |
| Non-targeted in | 1.092 | 2.507** | -1.415 | -0.020 | 0.016 |
| Treated Cohort | (1.835) | (1.185) | (0.905) | (0.012) | (0.011) |
| Observations | 24,542 | 24,542 | 24,542 | 24,542 | 24,542 |
| Y Mean | 64.22 | 29.38 | 34.84 | 0.01 | 0.01 |
| Y Standard Deviation | 43.57 | 23.88 | 34.44 | 0.88 | 0.85 |
| Student FE | ✓ | ✓ | ✓ | ✓ | ✓ |

*Notes:* 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. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 5: Effect of the Increased Autonomy Policy on Attendance 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 | 67.69 | 31.74 | 35.96 | 0.02 | -0.10 |
| Y Standard Deviation | 26.92 | 21.81 | 11.04 | 0.77 | 0.75 |
| Student FE | ✓ | ✓ | ✓ | ✓ | ✓ |
| P-value for H0: [1] + [2] = 0 | 0.00 | 0.00 | 0.57 | 0.00 | 0.15 |

*Notes:* 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. 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. 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. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

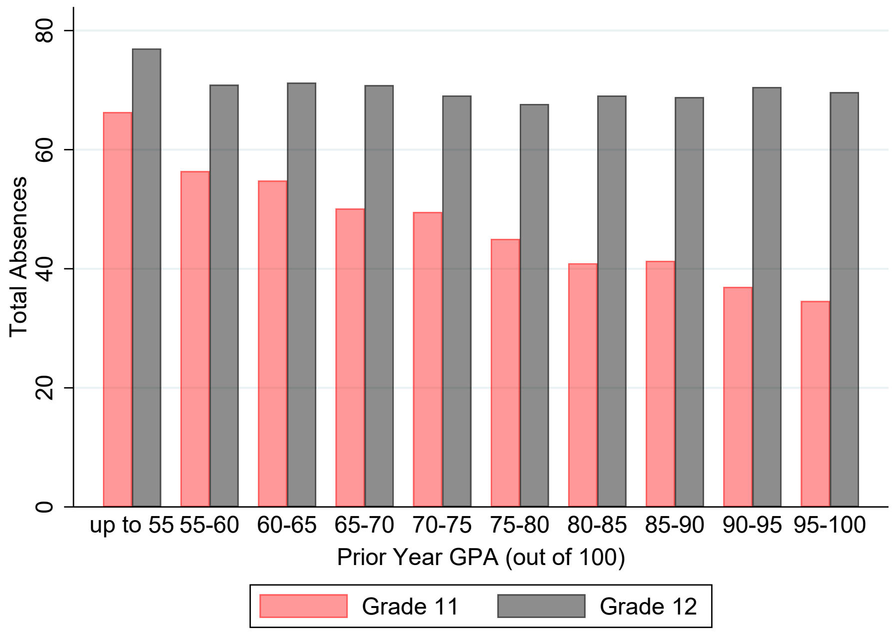
Table 6: EFFECT OF THE INCREASED AUTONOMY POLICY ON LONGER-TERM OUTCOMES

| | (1) University Admission Score | (2) Quality of Enrolled Degree |
|---|---|---|
| Targeted in Treated Cohort | 0.127*** | 2.109* |
| | (0.026) | (1.086) |
| Observations | 9,362 | 6,897 |
| Y Mean | 0.00 | 55.68 |
| Y Standard Deviation | 1.00 | 27.75 |
| Track FE | ✓ | ✓ |
| School FE | ✓ | ✓ |
| Year FE | ✓ | ✓ |

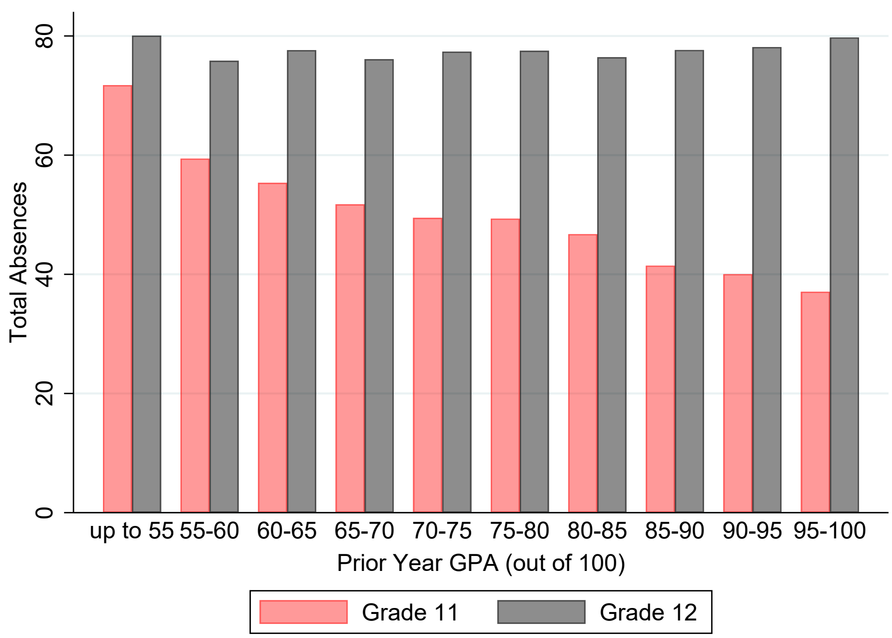
*Notes:* This table shows the estimated effect of increased autonomy on longer-term outcomes. The sample is restricted to students graduating between 2006 and 2007 in grade 12. Coefficient *Targeted in Treated Cohort*, [1], represents the effect of increased autonomy policy on targeted senior students 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, previous-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. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Figure 1: Total Absences Across Grades and Cohorts

Panel A: Control Cohort: Graduating in 2006 (Standard Autonomy Policy)

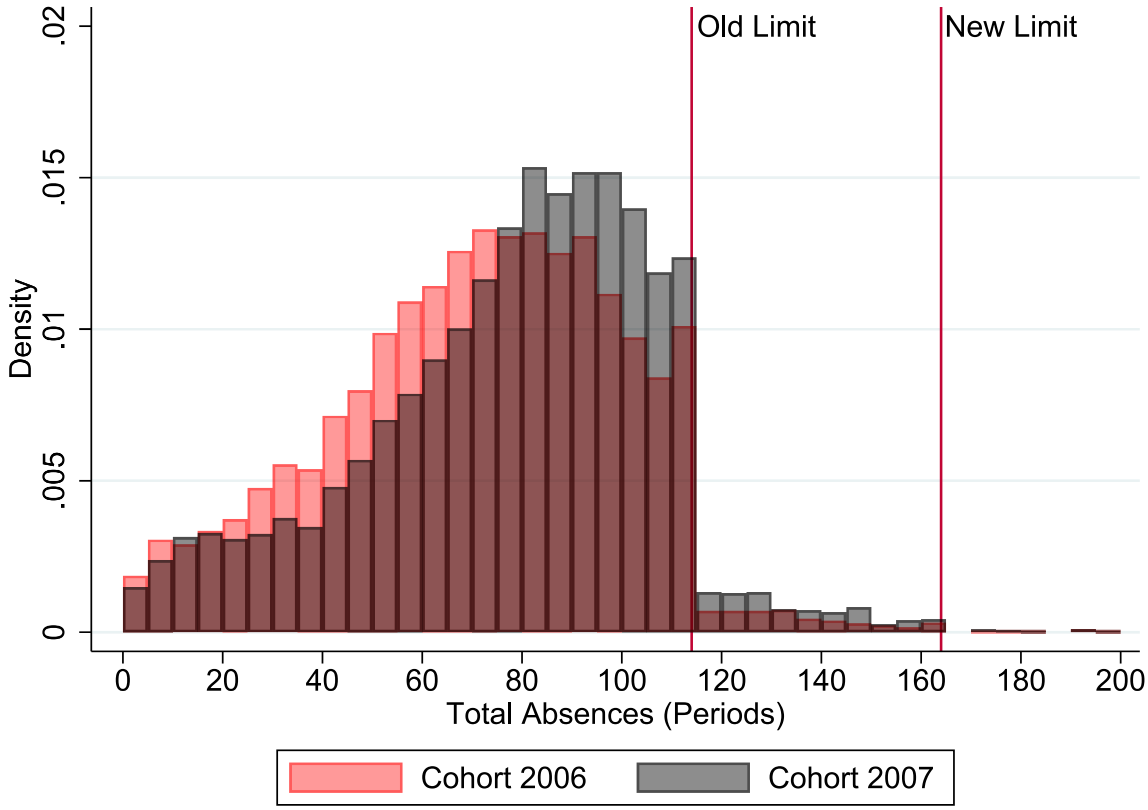



Panel B: Treated Cohort: Graduating in 2007 (Increased Autonomy Policy)




*Notes:* Graphs 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.

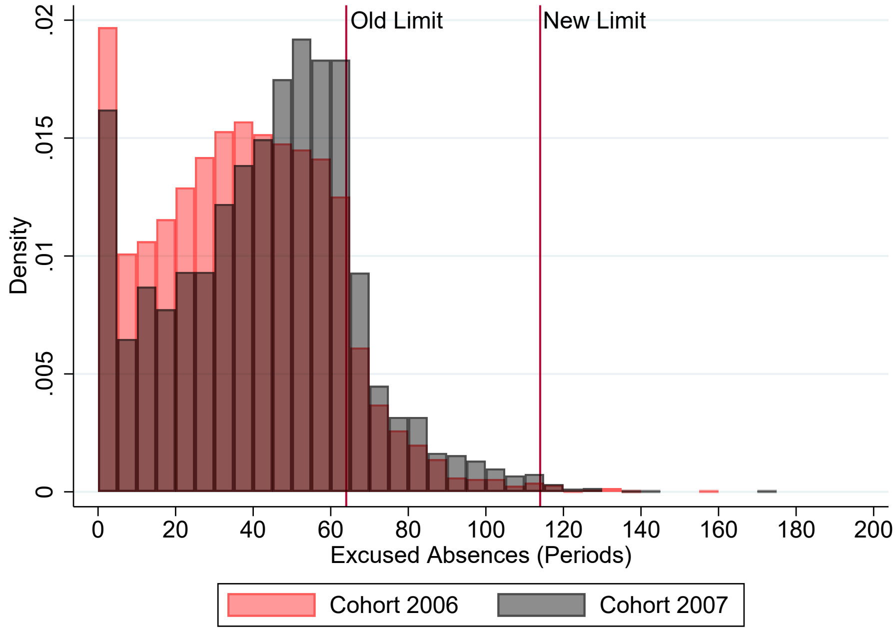
Figure 2: Total Absences Distribution in Grade 12 by Cohort



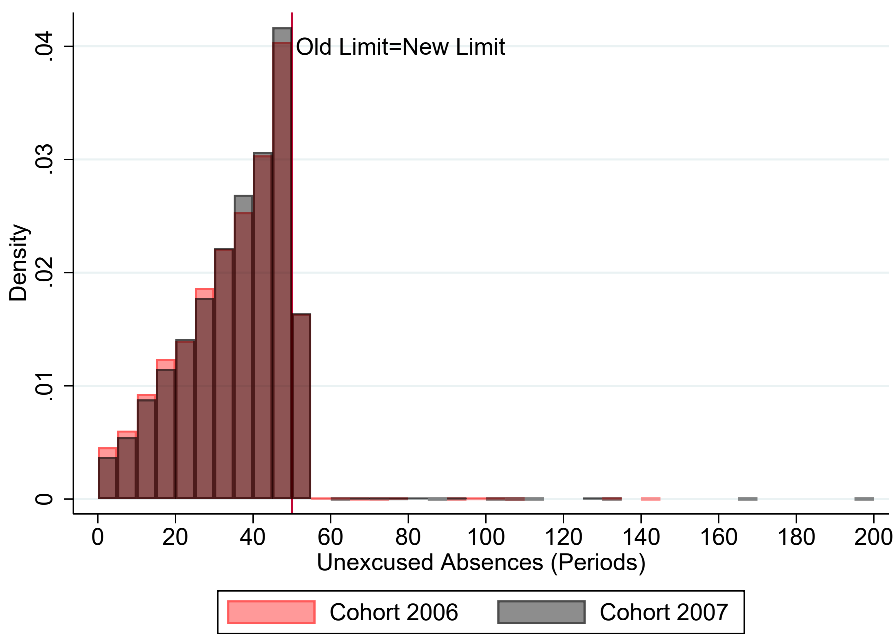

*Notes:* This figure shows the distribution of total class absences for the graduating cohorts of 2006 and 2007 in grade 12. Students graduating in 2006 could take up to 114 class absences in total before retention. Students graduating in 2007 could take up to 164 class absence in total before being retained.

Figure 3: Distributions of Absences in Grade 12 by Cohort

Panel A: Excused Class Absences by Cohort



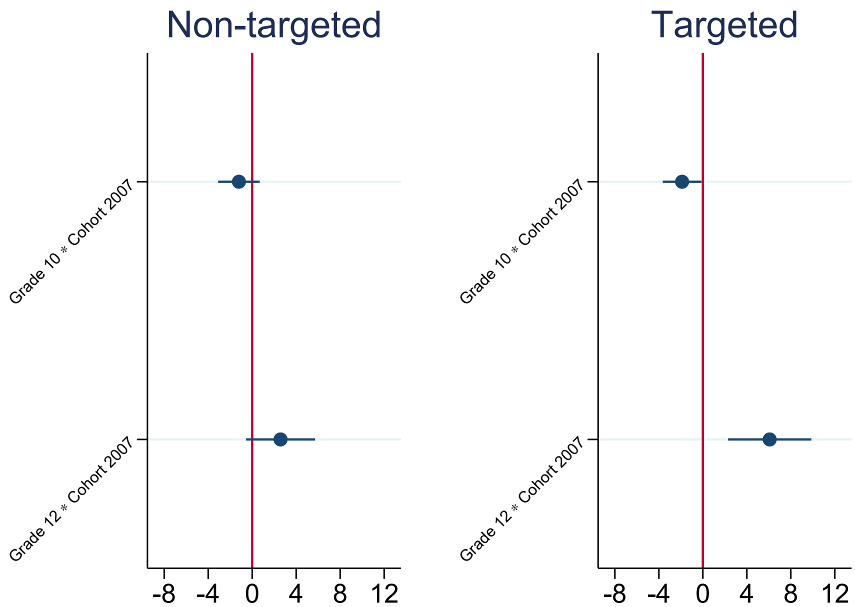

Panel B: Unexcused Class Absences by Cohort



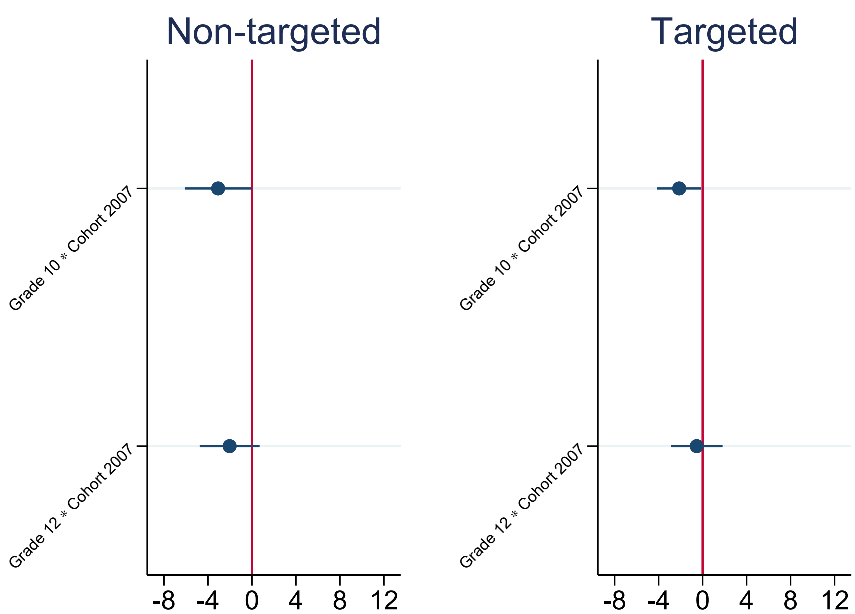

*Notes:* Panel A and B show the distribution of excused and unexcused class absences, respectively, for the graduating cohorts of 2006 and 2007 in grade 12. Students graduating in 2006 could take up to 64 excused class absences and 50 unexcused class absences before retention. Students graduating in 2007 could take up to 114 excused class absences and 50 unexcused class absences before being retained.

Figure 4: Common Trends Investigation of Attendance

Panel A: Excused Absences



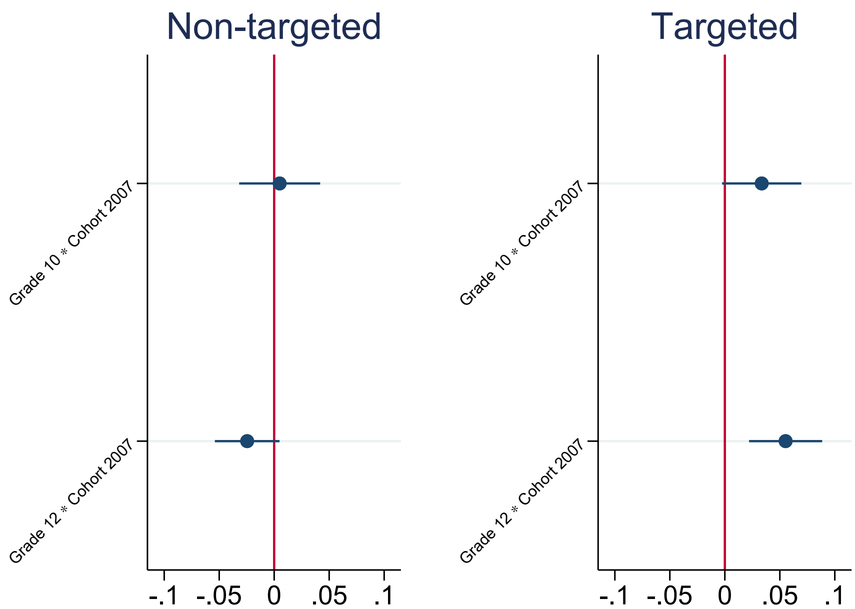

Panel B: Unexcused Absences



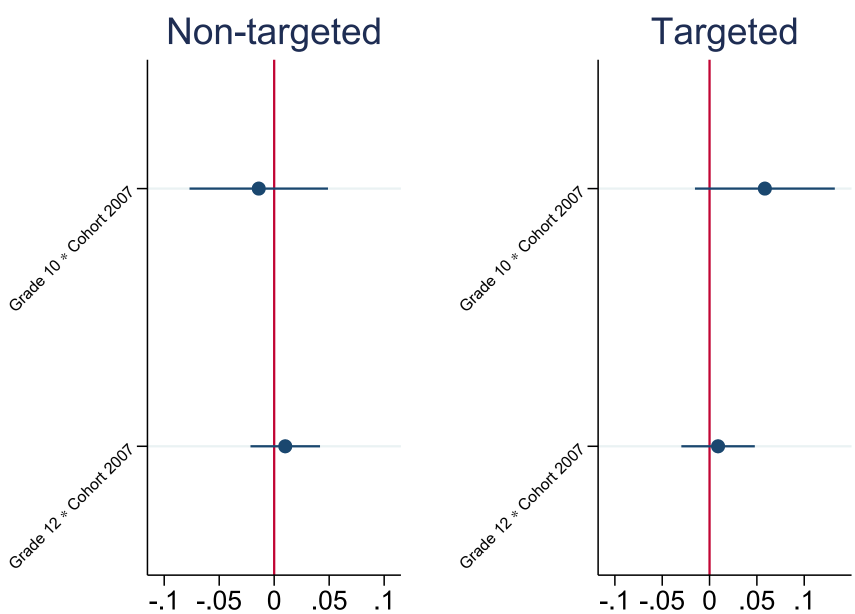

*Notes:* Each panel plots the estimated coefficients for the interactions between indicators for grades 10 and 12 and a cohort 2007 indicator for non-targeted (left) and targeted (right) students using specification (5).

Figure 5: Common Trends Investigation of Performance

Panel A: High-stakes Subjects




Panel B: Low-stakes Subjects




*Notes:* Each panel plots the estimated coefficients for the interactions between indicators for grades 10 and 12 and a cohort 2007 indicator for non-targeted (left) and targeted (right) students using specification (5).

# 12 Appendix Tables

Table A1: Heterogeneous Effects of the Increased Autonomy Policy 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 | 7.538** | 6.740*** | 0.799 | 0.028 | 0.003 |
| Cohort × Class SD Top Quartile | (2.966) | (2.385) | (1.217) | (0.023) | (0.027) |
| Targeted in Treated | 3.755 | 3.757* | -0.002 | 0.073*** | -0.007 |
| Cohort × Class SD Second Quartile | (2.794) | (2.052) | (1.220) | (0.025) | (0.033) |
| Targeted in Treated | 0.937 | -0.198 | 1.135 | 0.083*** | -0.026 |
| Cohort × Class SD Third Quartile | (2.561) | (1.810) | (1.208) | (0.025) | (0.030) |
| Targeted in Treated | -0.889 | 0.432 | -1.322 | 0.080*** | -0.026 |
| Cohort × Class SD Bottom Quartile | (3.221) | (2.487) | (1.204) | (0.028) | (0.029) |
| Non-targeted in | 1.841 | 2.680** | -0.839 | -0.011 | 0.019 |
| Treated Cohort | (1.920) | (1.235) | (0.932) | (0.014) | (0.013) |
| Observations | 21,948 | 21,948 | 21,948 | 21,948 | 21,948 |
| Y Mean | 64.22 | 29.38 | 34.84 | 0.01 | 0.01 |
| Y Standard Deviation | 43.57 | 23.88 | 34.44 | 0.88 | 0.85 |
| Student FE | ✓ | ✓ | ✓ | ✓ | ✓ |
| Control for Prior-year GPA | ✓ | ✓ | ✓ | ✓ | ✓ |

*Notes:* 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. 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, 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. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A2: Heterogeneous Effect of the Increased Autonomy Policy by Classroom Diversity (Interquartile Range)

| | Class Absences | | | Performance | |
|---|---|---|---|---|---|
| | Total | Excused | Unexcused | High-stakes Subjects | Low-stakes Subjects |
| Targeted in Treated | 6.011** | 6.332*** | -0.321 | 0.036 | 0.039 |
| Cohort × Class IQR Top Quartile | (2.672) | (2.097) | (1.318) | (0.023) | (0.029) |
| Targeted in Treated | 6.041*** | 3.705** | 2.336* | 0.079*** | -0.049* |
| Cohort × Class IQR Second Quartile | (2.241) | (1.540) | (1.208) | (0.020) | (0.029) |
| Targeted in Treated | 5.019* | 3.233 | 1.786 | 0.095*** | -0.003 |
| Cohort × Class IQR Third Quartile | (2.832) | (2.001) | (1.455) | (0.026) | (0.028) |
| Targeted in Treated | -1.479 | -0.689 | -0.790 | 0.076*** | -0.039 |
| Cohort × Class IQR Bottom Quartile | (3.121) | (2.432) | (1.368) | (0.028) | (0.029) |
| Non-targeted in | 1.685 | 2.958** | -1.273 | -0.022* | 0.015 |
| Treated Cohort | (1.802) | (1.145) | (0.902) | (0.012) | (0.011) |
| Observations | 24,542 | 24,542 | 24,542 | 24,542 | 24,542 |
| Y Mean | 64.22 | 29.38 | 34.84 | 0.01 | 0.01 |
| Y Standard Deviation | 43.57 | 23.88 | 34.44 | 0.88 | 0.85 |
| Student FE | ✓ | ✓ | ✓ | ✓ | ✓ |

*Notes:* 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. 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 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. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A3: Descriptive Characteristics of Matched Samples

| | Full Sample | | Matched Sample 1 (Table 5) | | | Matched Sample 2 (Table A4) | | | Matched Sample 3 (Table A5) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 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 | 5,843 | 5,720 | | 4,349 | 4,349 | | 4,349 | 4,349 | | 2,052 | 2,052 |
| Sampling Weights | | | Equal | | | Distance-Inverse | | | Equal | | |

*Notes:* 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 previous-year GPA across all students in the classroom.

Table A4: Effect of the Increased Autonomy Policy on Attendance and Performance Using Matching with Weighted Controls

| | Class Absences | | | Performance | |
|---|---|---|---|---|---|
| | Total | Excused | Unexcused | High-stakes Subjects | Low-stakes Subjects |
| Targeted in Treated Cohort [1] | 5.596*** | 5.299*** | 0.296 | 0.180*** | -0.085*** |
| *relative to non-targeted* | (1.879) | (1.606) | (1.049) | (0.044) | (0.031) |
| Non-targeted in Treated Cohort [2] | 3.703** | 4.970*** | -1.267* | -0.016 | 0.027 |
| *relative to control cohort* | (1.532) | (1.119) | (0.767) | (0.036) | (0.017) |
| Observations | 10,426 | 10,426 | 10,426 | 10,426 | 10,426 |
| Y Mean | 67.74 | 31.75 | 36.00 | 0.03 | -0.10 |
| Y Standard Deviation | 27.35 | 22.03 | 11.73 | 0.78 | 0.75 |
| Student FE | ✓ | ✓ | ✓ | ✓ | ✓ |
| P-value for H0: [1] + [2] = 0 | 0.00 | 0.00 | 0.31 | 0.00 | 0.03 |

*Notes:* 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. 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. 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. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

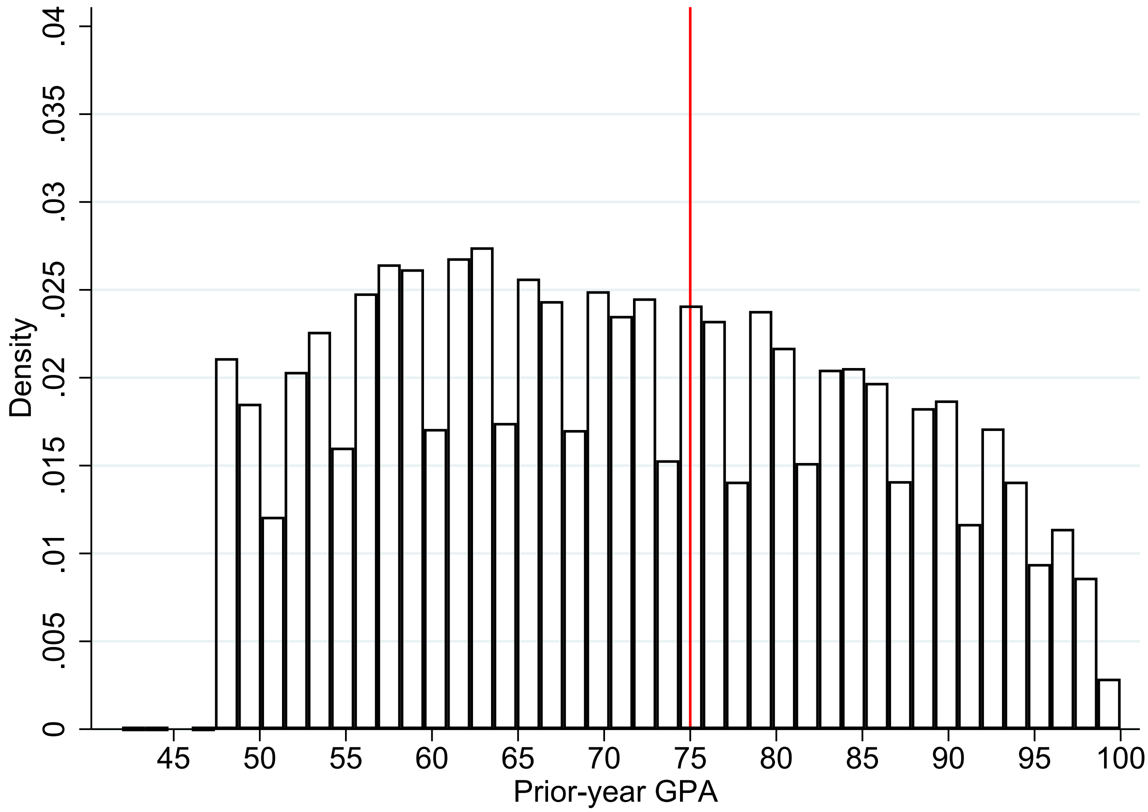
Table A5: Effect of the Increased Autonomy Policy on Attendance and Performance Using Matching with Additional Criteria

| | Class Absences | | | Performance | |
|---|---|---|---|---|---|
| | Total | Excused | Unexcused | High-stakes Subjects | Low-stakes Subjects |
| Targeted in Treated Cohort [1] | 5.894** | 6.259*** | -0.364 | 0.118** | -0.108** |
| *relative to non-targeted* | (2.378) | (2.127) | (1.812) | (0.059) | (0.051) |
| Non-targeted in Treated Cohort [2] | 3.217 | 3.921*** | -0.703 | -0.016 | -0.010 |
| *relative to control cohort* | (1.960) | (1.373) | (1.044) | (0.045) | (0.028) |
| Observations | 4,690 | 4,690 | 4,690 | 4,690 | 4,690 |
| Y Mean | 69.53 | 32.50 | 37.03 | 0.01 | -0.11 |
| Y Standard Deviation | 27.44 | 22.53 | 10.73 | 0.81 | 0.79 |
| Student FE | ✓ | ✓ | ✓ | ✓ | ✓ |
| P-value for H0: [1] + [2] = 0 | 0.00 | 0.00 | 0.55 | 0.05 | 0.00 |

*Notes:* 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. 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. 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. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

## 13 Appendix Figures

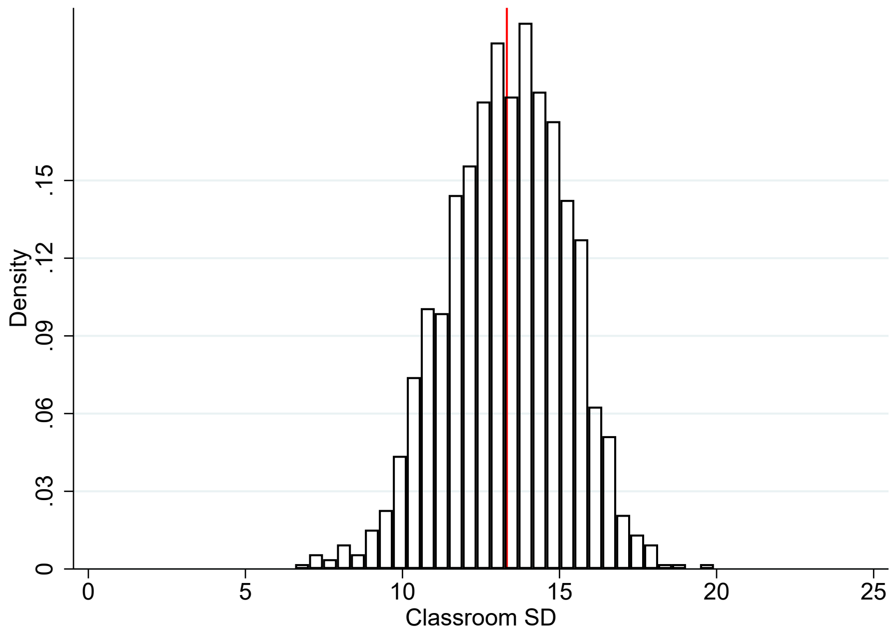
Figure A1: Distribution of Prior-year GPA



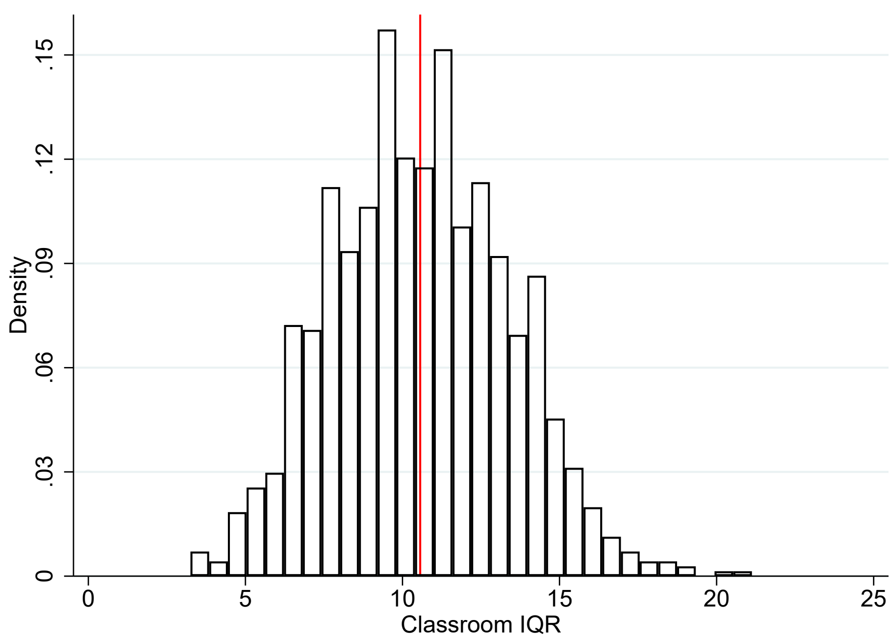

*Notes:* This figure plots the distribution of prior-year GPA across all students. Students with a prior-year GPA above the 75% (reflected by a vertical red line) were eligible for increased autonomy in the form of additional excused class absences in the 2006-07 school year.

Figure A2: Distribution of Classroom Diversity

Panel A: Classroom SD of Prior-year GPA




Panel B: Classroom IQR of Prior-year GPA




*Notes:* 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.

# 14 Appendix: Effect of the Increased Autonomy Policy around the Eligibility Cutoff

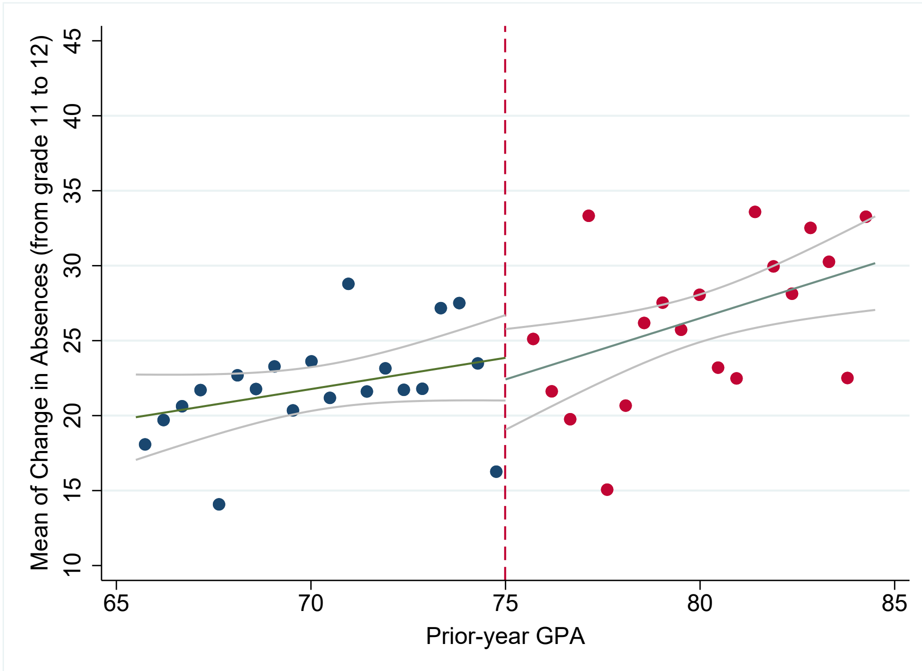
In this section, we investigate the absences trajectories of students above and below the eligibility cutoff for increased autonomy in the year prior to the implementation of the increased autonomy policy and the year the increased autonomy policy was in effect. A regression discontinuity approach to the investigation of the impact of the increased autonomy policy would rely on a significant jump in total absences student trajectories around the eligibility cutoff.

Following Lee and Lemieux (2010), Figure A3 plots the change in total absences between grade 11 and 12 of students in the control cohort (panel A) and treated cohort (panel B) using binned local averages. We find no substantial difference between the absences grade-over-grade trajectory of students who are right below and right above the eligibility cutoff of 75% in the previous-year GPA, for either the control or the treated cohort.

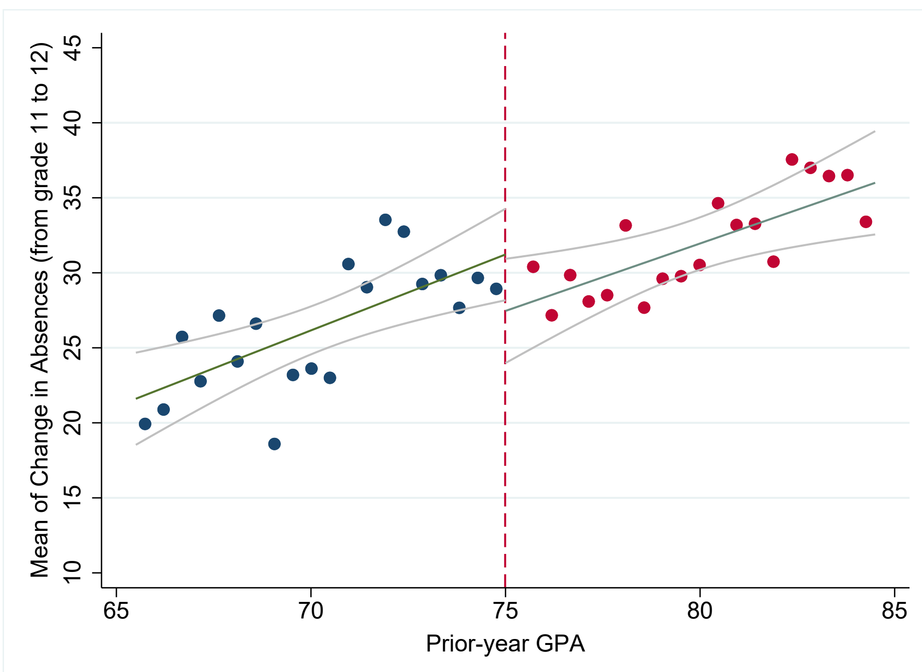
The findings of non-substantial jump in absences trajectories around the eligibility cutoff may be driven by two non-mutually exclusive mechanisms. As shown in Figure 1, the absences constraints may not binding for all students. With the absences constraint not binding, it is possible for students that are not eligible (non-targeted) for the increased autonomy policy to still increase their absences during the autonomy policy regime. At the same time, targeted students with a prior-year GPA further from the eligibility cutoff may have higher incentives to make use of the increased autonomy policy compared with the students closer to the cutoff as they may potentially have higher learning productivity away from public school (for example, due to access to private resources).

Figure A3: Increased Autonomy Effect around the Eligibility Cutoff

Panel A: Control Cohort: Graduating in 2006 (Standard Autonomy Policy)




Panel B: Treated Cohort: Graduating in 2007 (Increased Autonomy Policy)




*Notes:* Graphs A and B display the change in total absences between grades 11 and 12 for students graduating in 2006 and 2007, respectively.