Intensity of Workload per Exam and Academic Outcomes*

Carmen Aina[†] Koray Aktaş[‡] Giorgia Casalone[§]

May 2019

Abstract

This paper studies the causal effect of the workload per exam on the academic outcomes of students. We employ a difference-in-difference approach by exploiting exogenous variation—induced by a nationwide higher education reform—in the number of exams, holding constant the total credits for graduation. Using administrative data of a public Italian university, our results show that students enrolled in science degrees fail significantly more when they have to take fewer exams to achieve same amount of credit, which translates into an increase in first-year dropouts (22 pp.) and a less pronounced decrease in graduation rates (11 pp.). However, students enrolled in social sciences and humanities, are not affected by changes in the number of exams. Although their probability of failing an exam increases, this does not lead to any significant changes in their dropout and graduation rates. Finally, the new curricula do not affect the human capital accumulation of graduates as we do not find any impact on the graduation marks, wages, and employment probabilities.

JEL code: I23
Keywords: number of exams, college dropout, graduation, workload per exam, higher education.

*Koray Aktaş acknowledges financial support from the Economics Department of the Universitá del Piemonte Orientale. We thank the participants at various conferences and workshops. We gratefully thank Tommaso Colussi and Diogo G.C. Britto for their insightful comments and suggestions through the development of this paper. This paper has also benefited from conversations with Ingo Isphording, Ulf Zölitz, Anne Boring, Enrico Cantoni, Fabian Waldinger, and Daniela Sonedda.

[†]Department of Economics, Universitá del Piemonte Orientale, Novara, Italy; carmen.aina@uniupo.it

[‡]Department of Economics and Finance, Universitá Cattolica del Sacro Cuore, Milan, Italy; koray.aktas@unicatt.it

[§]Department of Economics, Universitá del Piemonte Orientale, Novara, Italy; giorgia.casalone@uniupo.it

1 Introduction

Despite the vast literature on education, there is no empirical evidence to help us understand whether students perform better over the course of their degree when they must study more material per exam but take fewer exams or when they take more exams but must study less material per exam. We aim to fill this gap in the literature by estimating the causal effects of increased workload per exam on the academic outcomes of students, while holding constant the number of credits required for degree completion. Our work carries important implications for policy makers and academics as understanding student behaviors toward academic curricula is a prerequisite for establishing efficient education systems.

A substantial body of literature reveals the importance of education on many fronts. Studies show that individuals with more education earn more during their life time (Card (1999); Heckman et al. (2006)). Furthermore, investments in higher education pay off in terms of life time earnings (Oreopoulos and Petronijevic (2013)). Greater education also reduces the likelihood of committing crime and increases health (Lochner (2011)). In light of this evidence and other studies, policy makers have shown interest in the efficiency of education systems (for instance, adjustments in the length of school years for different education levels). One of the objectives of policy makers is to decrease the age of labor market participation of individuals without compromising their human capital accumulation. This objective is spurred by the aging populations (see Marcus and Zambre (2018) for detailed discussion on the subject). Accordingly, education reforms have sought to decrease the starting age of schooling (Bedard and Dhuey (2012)), reduce the years of schooling (Card (1999); Pischke (2007); Krashinsky (2014)), and reduce the required years to obtain college degrees (Cappellari and Lucifora (2009)). However, evidence in the literature suggests that these reforms mostly generated unexpected negative effects on individuals' labor market outcomes, such as low earnings.

Marcus and Zambre (2018) investigate the effects of the G8 education reform in Germany, which reduced the school year of academic high schools but increased the instruction hours per week. Contrary to expectations, the study estimates that students who complete high school early are more likely to delay their college enrollment. Another recent study by Arteaga (2018) provides evidence that a reduction in the amount of coursework required for degree completion leads to decreased wages.

To study how the number of exams affects the key outcomes in the students' performance, we use an administrative data of the Università del Piemonte Orientale (hereafter UPO), which is a public university in northern Italy, for the academic years from 2002–03 through 2010–11. We exploit the exogenous variations in the number of exams, induced by a nationwide higher education reform introduced in 2007, to estimate the relative changes in the several academic outcomes. This reform sought to standardize the organization of degree courses across the country and to overcome an issue of delayed graduation of students in Italy. It set a maximum threshold for the total number of exams for degree completion (for instance, 20 exams to complete a 3-year bachelor degree) without making any changes to the number of credits needed for the degree (180 credits).[1]

At UPO, the reform was implemented differently in the various degree courses. For example, students enrolled in bachelor of science programs (e.g. biotechnology, biology, mathematics) used to take on average 11 exams in their first year; this number fell to 6 after the reform, but the number of credits required to continue to the second year remained at the same level (60 credits). Similar but smaller decreases occurred for bachelor degrees in social sciences (e.g. economics, political science) and humanities (e.g. literature), which dropped from an average of 10 exams to 8 exams in the first academic year. On the other hand, no changes were made in the number of exams required for bachelor degrees in medical school because the

[1] Details of the reform and the Italian higher education system will be discussed in the next section.

exam numbers were already at the level required by the reform.

In a difference-in-difference framework, we estimate the effect of taking fewer exams on academic outcomes of students enrolled at bachelor degrees in science, social sciences and humanities and consider the students enrolled in bachelor degrees in the medical school as a counterfactual group. Our results show a substantial increase, an average 22 percentage points (pp.), in the first-year dropout rate of students in science programs (the average first-year dropout rate 22% in these programs before the reform). We also observe a reduction of 11 percentage points on average in the likelihood of graduation for these students. We show that the mechanism that accounts for this significant worsening in academic performances is the increase in the probability of not passing any exam, which is about 24 percentage points. The latter effect indicates that it is more difficult for students to pass concentrated exams and that students who fail all exams decide to drop out. However, for the students who passed at least one exam in their first year, there is no effect on average grades. We show that students from high-income families are significantly less affected by the reform (on average about 12 pp. less) in terms of the effect on the first-year dropout rates. These students are more likely to have access to more resources, and also they might develop more non-cognitive skills during early life. Although the differential finding on the students of high-income families could indicate that the new curricula lead an increase in the inequality of opportunity, the total effect on these students is still positive and significant.

On the other hand, conditional on graduation, we do not find any significant changes in the time to graduate or in the final graduation marks. Moreover, we examine the effects on the labor market outcomes a year after the graduation by using labor market survey data in which we identify 85% of the graduates in our working sample. We do not find any impact on the wages and employment probabilities, suggesting that graduates' human capital is not affected by the new curricula.

The results concerning the struggles students have passing unified exams are consistent with the evidence in the literature on the task management of workers. Coviello et al. (2015) show that when judges work on many cases simultaneously rather than subsequently, their productivity decreases and it takes more time for them to conclude cases. In our context, given the time constraint for the exams, students must learn more material per unified exam, and this leads them to fail more compared to their performance when they have to study less material per exam but deal with more exams.

The findings for social sciences and humanities students differ from the ones for science students. We indeed observe a significant increase—about 10 percentage points—in the relative probability of not passing any exam in the first year. However, the latter effect does not translate into an increase in the first-year dropout rate. Although we estimate a 3.5 percentage points increase in the dropout rate, it is not statistically different than zero once we control for the observable characteristics of students. We do not find any effects on the other outcomes of students either—graduation, time to graduation, and final graduation marks. We attribute the finding of heterogeneous effects between science and social sciences and humanities to the smaller reduction in the total number of exams in the degrees of the latter group and to the arguably more challenging curricula of science degrees. Moreover, Stinebrickner and Stinebrickner (2013) show that students who choose science subjects as a major are significantly overconfident about their skills compared to students in other degree programs, which can also explain, in our analysis, why students in science programs dropped out after failing their exams in the first academic year.

We also present the results obtained from an event-study specification and show that the outcomes of interest between our so-called treatment and control groups followed parallel trends prior to the reform. Additionally, we perform exact matching on the observable characteristics of students to control for changes in students' composition. The results from the matched samples are in line with the ones obtained from the main estimation samples, indicating the estimated effects are not driven by the compositional changes.

Our study is also linked to a longstanding problem of the low efficiency of higher education institutes in Italy.[2] According to the OECD (2018), despite an increase in educational achievement—mainly for recent cohorts—Italy remains in the bottom of the education distribution for OECD countries and shows a persistent gap with other developed countries.[3]

Recently, discussions in Italy have focused on the introduction of selective admission procedures (such college entrance tests).[4] However, the evidence is rather mixed on the subject, as some studies show improvement in academic outcomes (e.g. Carrieri et al. (2015); Aktaş and Cappellari (2017)) while others (e.g. Francesconi et al. (2011)) do not find any effect from restrictive access to tertiary degrees. In this paper, the difference in effects between first-year dropout (22% increase) and graduation (11% decrease) rates in science degrees after introducing concentrated exams suggests that more than half of the students who dropped out at the end of their first year would have dropped out anyway in subsequent years. The latter is suggestive evidence that increased workload per exam in the first years can be used as an ex-post selection procedure. Nevertheless, it is important to find an optimal level in the intensity of exams because there is also an adverse effect on the students, who have a potential of graduating.

The rest of the paper is organized as follows. In Section 2, we explain the Italian higher education system and details of the reform. Section 3 contains information

[2] See Triventi and Trivellato (2009) for the historical trends in higher education outcomes in Italy.

[3] For instance, in 2015, the fraction of Italy's population aged 25—64 with tertiary education was 18%, while the average in OECD countries was around 35%; for the youngest cohort (aged 25—34), the figures were 25% and 42%, respectively. The small number of graduates is strongly related to persistent high dropout rates from the Italian university system despite efforts by the Italian Ministry of Education over the years, to increase retention.

[4] In Italy, students can choose tertiary education institutes without restriction, except for degrees in medicine and a few other exceptions (for example, Bocconi University).

on the data and provides descriptive statistics. Section 4 outlines the identification strategy and the econometric specifications. In Section 5, we discuss the results. We explain the matching procedure and the results obtained from matched samplein Section 6. Section 7 concludes.

2 The Italian University System and the Reform

The Italian university system is organized in three cycles, according to the Bologna structure:[5] the first cycle is the 3-year Bachelor's degree (*Laurea*), the second cycle is the 2-year master's degree (*Laurea Magistrale*), and the third cycle is the 3-year PhD programs (*Dottorato di Ricerca*). For particular programs (medicine, law, pharmacy, etc.) there is a "single cycle degree" (*Laurea a ciclo unico*) that lasts 5 or 6 years. Those who hold an upper secondary school degree are eligible to enroll in first cycle or single cycle degree programs. Some degree programs have selective admission tests that can be national or local. Degree programs are structured in university credits (*Crediti Formativi Universitari* (CFU) corresponding to the European Credit Transfer System (ECTS)). University credits represent a proxy for the workload an average student needs to achieve the expected learning outcomes. Each credit approximatively corresponds to 25 hours of student workload, including hours spent in class and hours of study at home. Students should achieve 60 credits by each academic year to graduate within the degree programs' intended duration. The first cycle degree is awarded to students who earn at least 180 credits, while second cycle and single cycle degrees are awarded to students who achieve at least 120 and 300 credits (360 for 6-year degrees in medicine and dentistry), respectively.

Degree programs are organized into several subjects. Each subject is assigned a number of credits that the student earns if he or she passes the final (written or

[5] See Bratti et al. (2008) and Cappellari and Lucifora (2009) for the detailed discussion of Bologna Process.

oral) exam. Other activities (such as internships, final dissertation, project works, etc.) allow students to accumulate credits needed for the degree. Exams are graded according to a scale ranging from 0 to 30, with 18 as a passing mark. Some courses (for instance basic computing skills, foreign languages, etc.) have no numerical grade, but rather are taken on a pass/fail basis.

Students take the exams in specific "exams sessions", generally at the end of the semester (term) in which they took the corresponding classes. The final exam covers the entire curriculum of the course, even if an unofficial mid-term was held. Official exams sessions last for several weeks, and there are many dates on which exams can be taken. Students can retake exams as they wish, even if they have already passed them and want to improve their grade. There is no cap on the number of times in which student can retake an exam, but each university/department can autonomously set limits provided that they do not infringe the rights of the students.

The changes introduced by Ministry Decree 155/2007 (known as *Mussi's Reform*) can be considered "minor" as the main features of the Italian university system, (e.g. the 3+2 system, the number of credits required to get the degrees, and the organization of degree courses according to the main field), remained unaltered from prior laws.[6] However, Mussi's Reform introduced some nonnegligible changes to the organization of the university teaching activity that affected the design of the degree programs, and, potentially, students' behavior and outcomes.

The aspect of the reform with the clearest and most immediate impact on students behaviors and outcomes is the setting of a cap on the number of subjects (and consequently, exams) required to earn the degree and the resulting reorganization of all study plans. Up until the 2007 reform, universities were allowed to freely determine the number of subjects (exams) needed to earn that university's degrees, provided that the total number of credits was of 180, 120, or 300, depending on

[6]D.M. 509/1999 and D.M.270/2004.

the cycle. This autonomy had two main consequences. First, the degree programs became excessively fragmented, with a huge number of small and barely coordinated subjects (exams). Second, due to the differences in the setting of the study plans across universities, it was difficult (nearly impossible) for students to move from one university to another, even within the same degree classes. The reform's cap was set at 20 subjects (exams) for first cycle degrees, and at 12 for second cycle degrees. Degree programs that exceeded the cap were forced to modify their study plan by combining two or more subjects or by eliminating small subjects, leaving only larger subjects. This change was explicitly meant to "improve the efficacy, quality and consistency of the degree programs", by promoting "the cooperation between professors of different subjects".[7] Obviously, as the number of credits did not change, the unforeseen consequence of this reform has been an increase of the average workload (in terms of ECTS credits) for each exam. The documents related to the Ministry Decree and the reform guidelines do not explain why the cap was set to these values. Probably the idea was that an average student can take between 3 and 4 classes each semester. However, there are no indications in the specialized literature, as far as we know, concerning the most effective organization of the university degree programs, and the choice seems to be rather arbitrary.

Furthermore, to allow for student mobility, universities that receive students from other universities are expected to recognize at least 50% of the exams already passed in the university of origin, provided that the student moves from a degree program of the same degree class.

As the reform introduced several changes to the organization of the degree courses, universities were expected to apply the new regulation within the two following academic years (2008–2009 and 2009–2010). The only constraint was that all the degree programs belonging to the same degree class in the same university had to

[7]Ministry Guidelines for the implementation of the new degree classes.

apply the reform in the same academic year,[8] and that students who had enrolled in the degree programs before the reform were subject to prereform curricula.

3 Data and Descriptive Statistics

The data we use in this paper are provided by the administrative office of the Università del Piemonte Orientale. UPO is a public university based in the Piemonte region of Italy with three campuses in the provinces of Alessandria, Novara and Vercelli. The university offers students a wide range of courses in 3-year bachelor programs; 2-year master programs and 5- or 6-year programs. During the academic years we work with, the teaching activities of the university are organized by 7 faculties: science (mathematics, physics, and natural sciences), medicine, pharmacy, law, political science, economics, literature and philosophy. These 7 faculties are aggregated into 4 scientific areas: medicine and pharmacy faculties into medicine (*area sanitaria*); science faculty into science (*area scientifica*); law, economics, political science into social sciences (*area sociale*); literature and philosophy into humanities (*area umanistica*). As previously mentioned, the reform required changes in the curricula and in the number of exams at the degree program level. However, at UPO, the reform has taken place at a more aggregate level—the scientific area level. Degree programs within the same scientific area (science, social sciences, humanities, and medical school) simultaneously adjusted their curricula.

All degrees in medical school (3-year degree programs as well as the 6-year medicine degree) perform a mandatory admission test prior to enrollment and only the students who pass this test can enroll in these degrees. For the science degrees, only biotechnology program requires students to take the admission test. On the other hand, all the social sciences and humanities degrees are open to any

[8] For instance, all the degree courses belonging to the degree class L35 (mathematics) in the same university had to apply the new regulation in the same academic year.

students holding an upper secondary school diploma. It has been confirmed by the administrative officials at UPO, during the academic years we study in this paper there has been no changes in any of the degree program's admission policy. As we work in a difference-in-difference framework, the latter confirmation is reassuring because otherwise our identification strategy could have been contaminated.

We restrict our sample to first-year bachelor degree students in order to have more homogeneous and comparable treatment and control groups because students who enroll in master programs and in 5- or 6-year degrees (pharmacy and law, medical school, respectively) are more likely to be more skilled with different motives than the bachelor's degree students.

The timing of the introduction of the reform to the degree programs varies across the scientific areas at UPO, which enables us to have an identification strategy to estimate the effects of the changes in curricula. The degree programs in social sciences and humanities introduced the reform in the 2008–09 academic year, while science degree programs introduced it in 2009–10. Finally, in the 2011–12 academic year, the medical school degree programs made the changes.

In Figure 1, we highlight the average number of exams per scientific area over the years (figure on the top displays the corresponding numbers for the first academic year, while the one at the bottom for the entire degree). These numbers are calculated by only taking into account the exams taken by students who completed their degree within the intended duration, which is 42 months. Otherwise, these numbers would have been inflated due to optional classes being taken and due to failed exams. As we see in the Figure 1, there is a significant reduction in the number of exams for degrees in social sciences and humanities in the academic year 2008–09, while we observe a similar but bigger fall in the number of exams for degrees in science at the beginning of the academic year 2009–10. As for the medical school degrees, we do not see any sort of changes in the number of exams, which is reassuring for our identification strategy as these students define our counterfactual group. The latter also explains the reason why the medical school waited to apply the changes required by the reform until the academic year 2011–12. Basically, the number of exams in these programs were already at the required level, and the changes occurred for them only on paper as a formality.

For convenience, we split our working sample into two. The first sample consists of students in the science programs and medical school. The second sample contains information about students in the social sciences, humanities and medical school programs. The former sample has information about 7,822 students and the latter has information about 12,385 students, while both samples cover the academic years 2002–03 to 2010–11. The restriction on the final year of the sample because the medical school applied the changes in curricula and exams at the beginning of 2011–12. Although, as explained above, these changes occurred only as formality, in the same academic year, another nationwide higher education reform restructured the organization of departments and faculties in universities. That could cloud our identification strategy for the estimates after the corresponding academic year.

Our data contain information at the student level. Specifically, we have information on the exact date of enrollment, the enrolled degree, exact date of exit from the degree with the reason for exit (dropout or graduation), the date of birth, gender, students final high school mark, the type of high school diploma, and the province of the high schools that the students graduated from. We also have student-level data on the exams. For each student in a given academic year, we have information as to which exams were passed, along with the grades earned from these exams. However, it is noteworthy that in exam data, the information about exams appears only if student passed the corresponding exam. Students must earn a grade of 18 (out of 30) to pass an exam.

We focus on seven different student outcomes in our analysis. Namely, these outcomes are the probability of not passing an exam in the first-year, the probability of dropping out in the first-year without passing any exam, the probability of first-year dropout, the probability of first-year "official" dropout, the probability of graduation, time to graduation (conditional on graduation), and the final graduation mark (conditional on graduation). There are two definitions for dropouts in our data. We label the first one as "official dropout", which occurs if a student withdraws her enrollment by giving a notice to the administrative office of the university. The second one occurs when students leave the university without informing administrative officials. Although this type of dropout takes some time to detect by the administrative officials, we have confirmed that the information on dropouts without notice is up to date during the academic years our sample covers. Nevertheless, for transparency, we report results for the official dropout rate as well as for the general dropout rates. As for the graduation variable, we set a maximum length to consider student's graduation status based the longest period of graduation of the youngest cohort in our sample. Specifically, among the students who enrolled at UPO in the academic year 2010/11, the latest graduation takes place in 68 months. Accordingly, for every other cohort in our data we do not take into account students who complete their degrees longer than 68 months when it comes to the regressions on graduation probability.

Table 1 reports the mean statistics of the academic outcomes this study focuses on. When we look medical school students, we see that academic outcomes on average are rather stable over the years. These students have the lowest first-year dropout rate (an average of 14%) and higher graduation rates (on average 67%) with respect to student outcomes in other scientific areas. The first-year dropout rate of students in science degrees before the changes in exam numbers (from 2002–03 through 2008–09) is about 22% on average, while during the following two academic year, this number increases to 43%. This already gives us a hint about the how students are affected by having more concentrated exams in these programs. Similar worsening in outcomes can be seen for the probability of not achieving credit, and the probability of dropping out without achieving any credit in the first year. We also see some changes in the outcomes of students in social sciences and humanities through the introduction of reform in 2008–09. The difference in outcomes before and after the reform in these programs is not as big as what is seen for the science programs. However, the first-year dropout rate in 2008–09 is the highest among other years in these degrees at 33.2%. It is also worth to remembering that in Figure 1, we observe a smaller decrease in the number of exams in these programs with respect to the decrease in exam numbers in science programs. Of course these are only raw numbers; we will investigate the changes in these outcomes with a proper econometric approach later in the paper.

We report the mean statistics of students' observable characteristics by scientific areas over the years in Table 2. These variables will be used as covariates in regressions. In the medical school, females make up 68% of the program participants. However, females make up only 46% of the students in science programs. In the third group, social science and humanity, the average female student percentage is about 60%. These numbers are consistent with the international student profiles in these degrees (i.e. low female participation in STEM subjects and high female participation in social sciences and in medicine-oriented subjects as nursing). The average age of students in the first and third groups is around 22, while it is 20 for the second group. Furthermore, the average high school final mark for students in medical school is 74, for students in science degrees is 77, and for students in social sciences and humanities is 76. Although final high school marks can be considered as a proxy for the students ability, its meaning is heterogeneous across the type of high schools as well as across regions of Italy.[9] Therefore, we also show the share of students who completed high schools with science curriculum (high-ability

[9]For example, students who graduated from scientific high schools might have lower final marks, but this might simply be an outcome of studying a more challenging curriculum.

track schools).[10] As one would expect, the share of these students is higher in science programs, with an average of 25%, and in medical school (22% on average) as compared to social science-humanities (18%). We also observe that a big share of students in all three groups come from the high schools that are located in the Piemonte region. In general, the differences in levels across scientific areas are not so relevant to our identification (at least in our case, the differences are in a reasonable range) but the trends in these differences are. In Section 6, we will present results obtained from working samples that are built upon matching these observable student characteristics.

4 Identification and Empirical Framework

We employ a difference-in-difference approach to identify the effects of exogenous changes in the number of exams on the outcomes of interest. We have two sources of variation. The first one is the variation across the timing of the implementation of the changes in curricula. As stated previously, the degree courses in social sciences and humanities adjusted their curricula at the beginning of the 2008–09 academic year, and the science degree courses applied the required changes in the academic year of 2009–10. Bachelor's degrees in medical school, on the other hand, waited until the academic year 2011–12 to execute these changes given that the number of exams in these courses were already at the required level (as shown in Figure 1). Therefore, we restrict our estimation sample to the academic years between 2002–03 and 2010–11 in order to use students enrolled in medical school programs as a counterfactual group, which allows us to identify the effect for social sciences and humanities during the academic years 2008–09 and 2010–11, and for the science degrees during the periods 2009–10 and 2010–11.

[10] For the regressions, we categorize high school diplomas into 6 types: scientific, professional, technical, classical, linguistic, and other.

The second source is the information on outcomes of treatment groups before and after the reform. We split the main working sample into two groups. The first one consists of medical school students as a control group and the science program students as the treatment group. In the second sample, social sciences and humanities make up the treatment group and once again the medical school is the control group.

4.1 Baseline econometric specification

The econometric specification we use is as follows:

$$Y_i = a + \gamma SA_i + \lambda POST_t + \delta(SA_i \times POST_t) + X_i\beta + \epsilon_i; \tag{1}$$

where Y_i is the outcome of interest, SA_i is a dummy variable that is equal to 1 if i is a student of the treated scientific area, $POST$ is a dummy and is equal to 1 if the academic year t is a post-reform year, X_i is the observable characteristics of student i; the parameter of interest we want to estimate is δ, which provides us the average effects. We replicate Equation 1 for each academic outcome of each working sample.

Standard errors are clustered at degree course and academic year level (Carrieri et al. (2015)). For the sample of science and medical school, there are 90 clusters, while there are 99 clusters for the second sample. In general, the ideal way of clustering standard errors in a difference-in-difference framework is to do it at the "treatment" level. In the cases of an insufficient number of clusters, the evidence shows that wild bootstrapping yields convenient results (Cameron et al. (2008)). However, MacKinnon and Webb (2017) provide evidence that if the number of clusters in regressions is less than 12, wild bootstrapping severely under rejects the null hypothesis of being equal to zero. In our case, the treatment occurs at the degree program level; we have 10 degree courses in first estimation sample and the second sample consists of 11 courses. Furthermore, in a context of an academic environment, the curricula and assignment of lecturers is updated at the beginning of each academic year, so two-way clustering at the degree program and academic year levels is much less restrictive than clustering at, say, state and calender year levels (e.g. to evaluate a labor market reform occurs at state level).[11]

4.2 Event-study Specification

In addition to our baseline model, Equation 1, we also set up an event-study specification to check whether the common trends assumption of difference-in-difference approach is satisfied in our estimation samples. This also allows us to estimate the effects year by year. To do so, we simply interact the year dummies with the treated scientific field dummies. We choose the year before the intervention as a baseline, which is the academic year 2008–09 for the science and medical school sample, and 2007–08 for the social sciences, humanities and medical school sample.

We set up the following model as an event-study specification:

$$Y_i = \alpha + \gamma SA_i + \sum_{k=2002}^{2010} \lambda_k P_k + \sum_{k=2002}^{2010} \delta_k (SA_i \times P_k) + X_i\beta + \epsilon_i; \qquad (2)$$

where $k = 2002, 2003, ..2010$, P_k are the dummy variables which are equal to 1 in year k, SA_i is equal to 1 if student i is in the treated scientific area, X_i is the observable characteristics of student i, the coefficient estimates of interaction terms, δ_k, are the parameters of interest in the model. As is in our main specification, we cluster the standard errors at degree the program level and academic years.

[11]Nevertheless, we have produced our results by clustering at the degree course level and performing a wild bootstrap procedure. For some outcomes of interest, we lose power in our treatment effect estimates and they become borderline insignificant. This is, as discussed above, due to an extremely small number of clusters in the analysis.

5 Results

In this section we discuss the estimation results obtained from the econometric specifications outlined in the previous section.

We report the results from baseline model—Equation 1—in Table 3 (for the science and medical school sample) and Table 4 (for the social sciences, humanities and medical school sample). In both tables panel A shows the results when the covariates for the observable student characteristics are not included in the model, while panel B shows the results when these control variables are included. As can be seen by comparing the parameter estimates across panel A and panel B, there is no important change in the main results after the controlling for the characteristics, although the covariates used in regressions are statistically meaningful. Nevertheless, the following discussions arise based on the estimation results reported in panel B.

Table 3 shows that, on average, there is a substantial worsening in the academic performances of students when they have to take fewer exams. In general, the first-year dropout rate increases by 22 percentage points. More specifically, we see an average increase of 24 pp. in the relative probability of not achieving any credit in the first year, and an increase of about 20 pp. in the probability of first-year dropout without achieving any credit. The latter finding shows that students fail significantly more exams; this translates into dropping out of college, and possibly indicates discouragement created by not earning any credit. We also observe that the increase in first-year dropout rate occurs as students officially terminate their enrollment at the university. As for the graduation rate, we see a decrease of 11 pp., on average. This result is rather interesting as is about the half of the first-year dropout rate in size, suggesting that half of the students among those who dropped out at the end of their first academic year would have dropped out anyway during the subsequent academic years. This also implies that, to some extent, having more-concentrated exams during the first academic years can be used as an ex-post selection procedure; it is very relevant from a policy standpoint given concerns about the efficiency of Italy's higher education system. Nevertheless, the other half of these students from the pool of first-year dropouts could have completed their degree courses. Finally, students' time to graduation increases, on average, about two months (conditional on graduation), while the final graduation marks do not change.

The coefficient estimates of covariates used in regressions are in line with other studies in the higher education literature. Students who graduated from professional, technical, classical, linguistic and other (as a category) high schools are more likely to drop out in their first year and thus less likely to graduate with respect to the students who graduated from science high schools. Their time to graduation also takes longer and it seems that they graduate from college with lower marks than students from science high schools. High school graduation marks are also important in explaining the students' academic performances. A one unit increase in the high school final mark leads a statistically significant reduction in first-year dropout rates by 0.3% and about a 0.8% increase in the probability of graduation. We also see that older students are more likely to drop out and less likely to graduate. However, once we look at the time to graduation and final graduation marks conditional on graduation,we observe that older students graduate in a slightly shorter period with slightly higher final marks; however, the size of these coefficients are very small. As for the gender effect, the coefficient estimates are mostly not statistically significant with exception of graduate probability, which is 6% higher for female students.

Table 4 reports the results for the social science, humanities and medical school sample. Unlike the previous results discussed above, we see a completely different picture when it comes to the behavior of students from the social sciences and humanities. There is an increase of about 10 pp. in the relative probability of not achieving any credit in the first year of the degrees. However, the increase in the first-year dropouts without achieving any credit is about half of the latter effect, while there is no statistically significant effect on general first-year dropouts.[12] This means that students perform worse in terms of achieving credit, but it does not discourage them or translate into dropping out. One possible explanation for this finding is that as shown in Figure 1, the reduction in the number of exams in the first year for students in social sciences and humanities is lower with respect to the changes in the number of exams for science programs. Moreover, the curricula in social sciences and humanities are not as difficult as the ones in science degrees, which mean that failing exams in the first year probably does not have the same impact on students expectations regarding their performance in the upcoming academic years.

The correlations between the observable student characteristics and academic outcomes are mostly consistent with the previous finding for the science programs. The only difference is in the performance of female students. This time we observe that female students are less likely to drop out in their first year and more likely to graduate (with slightly better final marks) than male students.

Figure 2 highlights the results obtained from Equation 2 for the science and medical school sample. The results clearly show that the common trend assumption in our identification strategy perfectly holds for every outcome we investigate in this study. The estimated effects on the first-year dropout rate and on the graduation rate vary differently across the academic years of 2009–10 (14 pp. with a p-value of 0.079) and 2010–11 (27 pp. with a p-value of 0.006). The bigger effect in the second year of the reduced number of exams reassures that the teaching staff of the university did not lower the difficulty level of the exams in 2010–11 after observing the decline in performances of previous cohort. In 2009–10 there is literally no effect on the graduation rate while the first-year dropout rate of the same cohort significantly increases by 14%. The latter provides a clearer understanding of how more-concentrated exams in first year can be considered an ex-post monitoring

[12] However, it is worth to note that when we exclude the last year from sample we estimate an average of 5.8 percentage points increase with a p-values of .030.

mechanism. Of course at this stage, this finding cannot be more than suggestive evidence given that in 2010–11 the graduation rate is significantly reduced as well (but the reduction is smaller than the increase in first-year dropouts).

We also extend the science and medical school sample one more academic year (2011–12) and estimate the effects on the official dropout rates. As discussed previously, one of the reasons we set the final year to 2010–11 in the estimation samples is that the bachelor degrees in medical school officiate the reform in 2011–12. However, we also discussed that in these programs there should be no change in real terms. This allows us to estimate the effects for one more academic year so that we can check if there is any adjustment by the lecturers to stop the ongoing worsening in students performance. Accordingly, in Figure 4 we reshow results on the official dropout rates by including an additional academic year.[13] As we see in the figure the official official dropout rate is still significantly higher in 2011–12 with a coefficient estimate of .106 (p-value .027).

We display the results of the event-study specification for our second estimation sample in Figure 3. Once again we observe common trends between the outcomes of the treatment and control groups across the academic years before the reform. The probability of not achieving any credit significantly increases during the first two years (2008–09 and 2009–10) but this effect disappears in the third year. This might be due to some adjustments in the difficulty of exams by lecturers after observing that students failed more exams in the previous year. More importantly, the increase in probability of failing exams does not translate into a dropping out in either of those years.

In addition to outcome variables discussed above, we present the results for the exam grades of students in their first year (conditional on passing an at least one exam). Figure 5 highlights these results for both of our samples. We do not observe

[13]We choose specifically the official dropout rate because the information of general dropout rate is not up to date for the degrees in medical school after the academic year 2010–11.

any significant changes. This is another piece of evidence suggesting that the main channel of the effects we observe on the first year dropout rate is from not passing any exam.

5.1 Results by Family Income

In this section we look at the heterogeneous results by family income of students. In theory, students coming from high income families can develop more non-cognitive skills, and these students can also have more resources available to them with respect to the students of low income families. Therefore, students of high income families can be expected to cope up with the concentrated exams better.

Our data contains information on the tuition fees that students are supposed to pay in each year of their degrees. The amount of fee that students pay is calculated based on the student's family income. The higher the student's declared family income is the higher the tuition fee student pays. Hence, tuition fees are very good proxy for family income.

In our estimation sample for the degrees in science and medical school, tuition fees vary from 88 Euro to 3333 Euro. Over the years our sample covers we do not see any significant changes in the composition of students in terms of average tuition fee they pay (see Figure 12, upper-left).

By splitting the distribution of tuition fees into three quantiles we define three categories: low (if tuition fee is less than or equal to 610 Euro), medium (if tuition fee is between 610 and 1228 Euro), and high (if tuition fee is greater than or equal to 1228 Euro) family income. Then we split our sample into these three categories and estimate the effect of the reform for each sample separately, using the econometric specification outlined in Equation 2. The results are highlighted in Figure 12. As can be seen from the Figure, for each sample, we estimate significant increase in the first year dropout rate after the introduction of the reform. However, for the sample that consists of students of high income families the estimated effects on the first year dropout probability is smaller compared to the low and medium income families. To identify the latter differential effect by family income, and see whether it is statistically significant, we extend Equation 2 and introduce triple interaction into our econometric specification.

$$
\begin{aligned}
Y_i &= \alpha + \gamma SA_i + \tau High_i + \pi(SA_i \times High_i) \\
&+ \sum_{k=2002}^{2010} \lambda_k P_k + \sum_{k=2002}^{2010} \theta_k (P_k \times High_i) \\
&+ \sum_{k=2002}^{2010} \delta_k (SA_i \times P_k) + \sum_{k=2002}^{k=2010} \omega_k (SA_i \times High_i \times P_k) \\
&+ X_i \beta + \epsilon_i;
\end{aligned} \tag{3}
$$

where Y_i is a binary dependent variable that takes value 1 if student i dropouts out at the end of her first year and 0 otherwise, $k = 2002, 2003, ..2010$, P_k are the dummy variables which are equal to 1 in year k, SA_i is equal to 1 if student i is in the treated scientific area, $High_i$ is a dummy variable and is equal to 1 if student i is in the high family income category, otherwise 0. The coefficients δ_k give us the total effects of the reform on the students who are in the low and medium family income categories, while the coefficients ω_k provide the differential effects on the students who are in the high family income category.

We highlight the results obtained from Equation 3 in Figure 13. The first graph on the top of the Figure presents the coefficient estimates of δ_k, and the second graph on the bottom of the Figure shows the coefficient estimates of ω_k. The total estimated effects on the students from low and medium income families are 15.6 and 33.4 percentage points in 2009 and 2010, respectively, and they are statistically significant. In 2009 the differential effect estimated for the students from the high income families is −3.1 pp. but not statistically significant (with a p-value 0.76), while in 2010 we observe a statistically significant effect of −18.3 percentage points. Although the latter finding suggests that these students are less affected by the changes in their curricula with respect to the students from lower income families, the cumulative effect is positive and significant.

5.2 Effects on Labor Market Outcomes

Labor market effects of the reform under scrutiny are investigated by adding to the subsample of UPO graduates information drawn from *AlmaLaurea*, which is a consortium of Italian Universities whose aim is to provide employers with data on graduates.

Graduates fill in a questionnaire at the completion of their three-year degree (*Profilo dei Laureati* survey) and are monitored after 1 year from graduation (*Condizione Occupazionale dei Laureati* survey). The response rate is about 80% for each cohort of graduates. For graduates not enrolled in further education (i.e. master degrees), the survey collect information on their employment condition; namely, time to get a job, occupational characteristics, and wages. This set of variables is matched with students' details contained in the universities' administrative data registers.

3684 people compose our final sample of graduates who completed their degrees in science or medical school, excluding nonrespondents and those who are still in education. The latter number corresponds to 85% of our working sample conditional on graduation. To analyze labor market performance of our graduates, we define the following two dependent variables: net monthly salary that is the amount of salary in euros received by graduates, and employment rate. Wages are recorded in the data as intervals, and each interval is in the range of 250 Euro (i.e. 0-250, 250-500, 500-750 and so on). We use the midpoints of these intervals as a proxy to the actual wages. As for the employment status, there are three categories in the data:

employed, searching for a job, and not searching for a job. Our dependent variable on the employment status is a binary dummy, which takes value 1 if the student is employed, and takes value 0 if the student is searching for a job at the time of interview. Among the 3684 people identified in the labor market data, 2334 of them provided answer to the wage question, and 2771 of them answered the employment status question.

Using our baseline model outlined in Equation 2, we estimate the effects of the reform on the labor market outcomes of graduates. Results are shown in Figure 15. Our findings show that neither wages nor employment status of graduates is affected by the changes in their curricula during their degree programs. As we discussed early in this section, the reform does not have any impact on the academic outcomes of graduates (time to graduation and final graduation marks). The results on the labor market outcomes of graduates also reassure us that the reform does not improve (or diminish) the human capital accumulation during degree programs.

6 Exact Matching

In addition to the common trend assumption, which has been shown to hold in our estimation samples, the difference-in-difference framework requires stable composition across treatment and control groups when working with repeated cross-section data. In this section, we create estimation samples by performing exact matching on the observable student characteristics. We then highlight the trends in student composition over the years and determine whether there are any significant differences in those trends to see the quality of the matching procedure. Finally, we present the results obtained from the matched samples.

6.1 Matching procedure

We perform a two-stage matching process. First, in a given academic year (for years both before and after the reform), we exact match treated students with the counterfactual group according to gender, age and the type of high school diploma. For these matched students, we calculate the differences in their high school graduation marks; we keep students in the control group if the difference between their final marks and their counterpart's final mark is not greater or lower than one unit.[14] In the last stage, we randomly pick two students from the pool of matched control students for each treated student. Considering that our main estimation samples are modest in terms of sample size, in some cases there is no proper match for the treated students, in which case those treated students are excluded from sample, or the number of matched control is only one student. Furthermore, in some cases, one control student is a match for multiple treated students, but this is not problematic in our set up because we work with repeated cross-section data in which we observe each student only for one year. Nevertheless, we also show results from a matched sample in which we allow the observations of control students to repeat based on how many times they are matched with a treated student. This is broadly equivalent to weighted propensity score matching and increases the precision of the matching procedure.

In the end, our first matched sample consists of 1,703 students from science programs (treated) and 1,476 students (control) from medical school, while the second matched sample contains 3,784 students (treated) from social sciences and humanities programs and 2,430 students (control) from medical school.

[14]A matching procedure on high school marks can be also executed as an exact match, but this leads us to an extremely small sample. Therefore, we perform matching by accepting a small gap between the final marks. High school final marks range between 60–100 so allowing a maximum of a one unit gap for the differences in these marks is reasonable.

6.2 Results from matched samples

Using the econometric specification (Equation 2) outlined in Section 4.2, we estimate the relative changes in the composition of treated students. The results are highlighted in Figure 6 for the science and medical school sample, and in Figure 7 for the social sciences, humanities and medical school sample. As can be seen from these graphs, the matching procedure did a good job of balancing the student compositions between the treatment and control groups over the years for every variable that is used in matching.

The results on the academic outcomes are displayed in Figure 8 and in Figure 9. All results are perfectly in line with what has been presented in this study so far. In Figure 8, we see a small but significant deviation in the common trend in the 2003–04 academic year. However, this occurs 6 years before the reform; regarding the most recent academic years before the reform, the trends are parallel. One interesting finding in Figure 9 is that, unlike the results we have seen for the main social sciences, humanities and medical school sample, the probability of dropping out in the first year without achieving any credit significantly increases in the 2008–09 and 2009–10 academic years. Yet the probability of first-year dropout is not affected.

Figure 10 and Figure 11 present the results from the matched sample when the observation of matched control students is allowed to repeat if they are a match to many treated students. Once again, we see that the results are consistent with the previously outlined findings. In fact, results on the probability of not passing any exam, dropping out without passing any exam, and the general first-year dropout rates are more pronounced for the social sciences and humanities from this sample. That is, if anything, slight differences in student compositions actually work against our results rather than derive them.

6.3 Heterogeneous Results by High-school Type and Ability

Using our matched sample explained previously, we investigate whether certain students are affected differently by the reform. First, we look at the differential effects on the students who completed academic-track high-schools (*licei*), which, in our data, include scientific, classical and linguistic high-school types. The rest (technical, professional and others) lies in the category of nonacademic-track high schools.

We use the econometric specification in Equation 3, replacing the dummy variable $High_i$ with a dummy variable that takes value 1 if a student graduated from an academic-track high-school. Results are highlighted in Figure 16. The estimated differential effects on the first-year dropout rate of academic-track high-school graduates are negative, on average it is about -15 percentage points, but statistically not significant. Nevertheless, it is worth to note that the differential effect in 2010 is borderline insignificant. As a conclusion, even though the difference is not so strong, the findings suggest that the general effects that we have discussed throughout the paper is mainly driven by the students who completed nonacademic-track high-schools.

We also estimate heterogeneous effects by the abilities of students. Our ability measure is the final high-school graduation marks of students. We categorize students who completed their high-school with a final mark that is lass than or equal to 75 (these marks vary from 60 to 100 in Italy) as *low-ability* students. However, these final marks can represent different skill levels in different high-school tracks. Therefore, we run the heterogeneous analysis by ability separately for students who completed academic-track high-schools and for students who completed nonacademic-track high-schools.

Results for the sample consists of students from nonacademic-track high-schools are presented in Figure 17, while in Figure 18 we display the results on the students from academic-track high-schools. We do not observe any sort of differential effects on the *low-ability* students in neither of the samples, suggesting that the high-school track choice of students is more relevant than the performances of these students during their high-secondary education.

7 Conclusions

In this study, we have investigated how students perform if they take fewer exams during their college degree, but need to study more material per exam. Economists have overlooked this question in the literature. We use an administrative data set of a public university located in Italy to exploit exogenous variation generated by a national higher education reform regarding the number of exams required in degree programs. We employ a difference-in-difference approach. Our findings show a tremendous increase in the first-year dropout rate of students in science programs, which has the highest workload intensity per exam. The effect on graduation rates is less pronounced. We pointed out that the difference between the increase in first-year dropouts and the decrease in graduation rates can be considered, to some extent, as a post-enrollment monitoring mechanism.

On the other hand, we have not estimated any significant effect on first-year dropout or graduation rates of students in social sciences and humanities, even though their probability of passing any exam in their first year significantly decreases. We have discussed several explanations for this differential effect across the different scientific areas. The most obvious one is that the social sciences and humanities programs were subject to less reduction in the number of exams than other programs. The second explanation is the evidence in the literature that students who choose science subjects as major are overoptimistic.

As a final note, this paper does not aim to evaluate the reform in a general sense because we do not know how this reform affected different programs of different universities in terms of the changes in the number of exams. Nevertheless, our findings present several recommendation. From a policy point of view, national reforms on such a sensitive subject should not take place based on arbitrary choices. It may be that reforms at the national level should not take place prior to gathering information from the institutes that would be affected by reforms. A more coordinated effort between the institutes and policy makers would likely result in better student outcomes and educational systems.

References

Aktaş, K. and L. Cappellari (2017). Selective Admissions and Freshmen's Academic Outcomes. *Manuscript*.

Arteaga, C. (2018). The effect of human capital on earnings: Evidence from a reform at Colombia's top university. *Journal of Public Economics 157*, 212–225.

Bedard, K. and E. Dhuey (2012). School-entry policies and skill accumulation across directly and indirectly affected individuals. *Journal of Human Resources 47*(3), 643–683.

Bratti, M., D. Checchi, and G. De Blasio (2008). Does the expansion of higher education increase the equality of educational opportunities? Evidence from Italy. *Evolution of Labour Market Inequalities 22*(s1), 53–88.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics 90*(3), 414–427.

Cappellari, L. and C. Lucifora (2009). The "Bologna Process" and college enrollment decisions. *Labour Economics 16*(6), 638–647.

Card, D. (1999). The causal effect of education on earnings. In *Handbook of labor economics*, Volume 3, pp. 1801–1863. Elsevier.

Carrieri, V., M. D'Amato, and R. Zotti (2015). On the causal effects of selective admission policies on students performances: evidence from a quasi-experiment in a large Italian university. *Oxford Economic Papers 67*(4), 1034–1056.

Coviello, D., A. Ichino, and N. Persico (2015). The inefficiency of worker time use. *Journal of the European Economic Association 13*(5), 906–947.

Francesconi, M., C. Aina, and L. Cappellari (2011). Student Performance may not Improve when Universities are Choosier. *Evaluation & Research in Education 24*(4), 231–242.

Heckman, J. J., L. J. Lochner, and P. E. Todd (2006). Earnings functions, rates of return and treatment effects: The mincer equation and beyond. *Handbook of the Economics of Education 1*, 307–458.

Krashinsky, H. (2014). How would one extra year of high school affect academic performance in university? Evidence from an educational policy change. *Canadian Journal of Economics/Revue canadienne d'économique 47*(1), 70–97.

Lochner, L. J. (2011). Nonproduction benefits of education: Crime, health, and good citizenship. In *Handbook of the Economics of Education*, Volume 4, pp. 183–282. Elsevier.

MacKinnon, J. G. and M. D. Webb (2017). Wild bootstrap inference for wildly different cluster sizes. *Journal of Applied Econometrics 32*(2), 233–254.

Marcus, J. and V. Zambre (2018). The effect of increasing education efficiency on university enrollment: Evidence from administrative data and an unusual schooling reform in Germany. *Journal of Human Resources*, 1016–8324R.

OECD (2018). Education at a Glance 2018: OECD Indicators. *OECD Publishing, Paris*.

Oreopoulos, P. and U. Petronijevic (2013). Making college worth it: A review of the returns to higher education. *The Future of Children*, 41–65.

Pischke, J.-S. (2007). The impact of length of the school year on student performance and earnings: Evidence from the German short school years. *The Economic Journal 117*(523), 1216–1242.

Stinebrickner, R. and T. Stinebrickner (2013). A major in science? Initial beliefs and final outcomes for college major and dropout. *The Review of Economic Studies 81*, 426–472.

Triventi, M. and P. Trivellato (2009). Participation, performance and inequality in Italian higher education in the 20th century. *Higher Education 57*(6), 681–702.

Tables

Table 1: **Academic outcomes of students over the years by scientific field**

	Medical							Science							Social sc. & Humanity						
	Pr-fail	Drop(1)	Drop(2)	Drop(3)	Grad(1)	Grad(2)	Mark-G	Pr-fail	Drop(1)	Drop(2)	Drop(3)	Grad(1)	Grad(2)	Mark-G	Pr-fail	Drop(1)	Drop(2)	Drop(3)	Grad(1)	Grad(2)	Mark-G
2002	.096	.091	.132	.070	.630	41.8	96.6	.157	.139	.196	.109	.419	44.5	99.9	.185	.166	.222	.098	.515	45.9	98.4
2003	.114	.108	.165	.102	.660	42.2	98.5	.158	.143	.2	.092	.522	43.7	101.	.171	.145	.239	.086	.484	45.2	98.8
2004	.179	.156	.210	.134	.627	41.3	97.3	.190	.165	.248	.125	.430	44.3	101.	.25	.217	.303	.146	.394	45.0	97
2005	.128	.102	.163	.093	.675	41.4	95.8	.222	.188	.245	.157	.465	46.6	100.	.225	.197	.274	.118	.450	44.6	96.7
2006	.113	.096	.144	.090	.680	41.5	96.4	.170	.139	.228	.156	.468	47.1	101.	.221	.192	.284	.128	.402	44.8	95.9
2007	.100	.083	.145	.093	.618	42.4	98.6	.163	.131	.247	.191	.401	46.0	101.	.234	.191	.264	.138	.438	44.3	96.7
2008	.087	.080	.133	.076	.706	42.6	98.7	.154	.128	.211	.150	.492	47.3	99.9	.346	.270	.332	.180	.409	47.1	94.2
2009	.073	.062	.122	.071	.667	41.8	97.7	.313	.262	.350	.271	.418	48.6	99.2	.308	.236	.287	.143	.411	45.6	94.3
2010	.067	.057	.119	.081	.739	41.4	98.9	.462	.379	.496	.401	.313	48.0	100.	.205	.165	.238	.116	.461	42.8	96.0

Notes: Table 1 reports the mean statistics of the outcome variables in our estimation samples. Pr-fail stands for the probability of failing all exams in first academic year. Drop(1) is the probability of dropping out in the first year without passing any exam. Drop(2) is the probability of dropping out in the first year, Drop(3) is the probability of official dropout in the first year, Grad(1) is the probability of graduation. Grad(2) is time to graduation (conditional on graduation). Mark-G is the final graduation mark (conditional on graduation).

Table 2: **Composition of students over the years by scientific field**

	Medical								Science								Social sc. & Humanity							
	Female	Age	Age(SD)	Mark	Mark(SD)	HS Sc	Pie.	N	Female	Age	Age(SD)	Mark	Mark(SD)	HS Sc	Pie.	N	Female	Age	Age(SD)	Mark	Mark(SD)	HS Sc	Pie.	N
2002	.70	22.7	4.92	74.0	11.4	.15	.78	727	.40	22.0	5.50	77.0	12.1	.22	.90	436	.63	22.5	5.97	77.3	12.1	.18	.84	800
2003	.71	22.1	5.03	74.6	11.3	.15	.77	508	.44	21.3	4.67	78.2	13.0	.23	.91	334	.63	22.8	6.38	76.5	12.4	.17	.81	847
2004	.71	22.4	5.53	74.1	11.1	.19	.77	511	.41	20.8	4.36	79.8	13.1	.21	.90	325	.58	23.6	8.19	77.2	12.8	.12	.84	877
2005	.71	22.0	5.21	74.3	11.7	.17	.75	456	.40	20.7	4.53	77.8	13.1	.28	.91	354	.58	23.0	7.60	77.2	12.7	.16	.84	771
2006	.68	23.2	6.57	74.3	11.8	.21	.67	483	.45	21.1	5.26	79.0	12.2	.23	.94	294	.59	22.7	7.08	77.3	12.9	.15	.85	855
2007	.66	22.8	6.89	73.1	10.9	.21	.64	551	.49	20.3	4.00	78.8	12.2	.25	.89	251	.61	22.8	7.17	75.8	12.0	.22	.84	931
2008	.66	22.5	6.25	74.7	11.4	.26	.64	496	.44	20.6	4.89	77.6	11.2	.27	.89	264	.61	22.3	6.00	75.4	11.5	.18	.84	901
2009	.63	22.5	5.90	74.1	10.8	.29	.72	531	.50	20.5	4.69	75.4	11.4	.27	.88	354	.58	22.1	6.38	75.5	11.1	.20	.81	864
2010	.64	22.0	5.62	75.4	10.8	.34	.70	503	.57	20.2	3.98	74.1	10.9	.29	.88	449	.59	22.4	7.11	74.8	10.9	.19	.78	782
Total	.68	22.5	5.79	74.3	11.3	.223	.721	4,766	.462	20.9	4.7	77.3	12.3	.254	903	3,061	607	22.7	7.11	76.3	12.2	.18	.833	7,628

Notes: Table 2 reports the mean statistics of the observable characteristics of students. Female is the percentage of female students. Age is the average age at enrollment, and Age(SD) is the standard deviation in age. Mark is the final high school mark, and Mark(SD) is the standard deviation of high-school final marks. HS Sc. is the percentage of students holding scientific high school degree, Pie. is the percentage of students who completed high school in the Piemonte region, N stands for the number of students.

Table 3: **Estimation Results: Science vs. Medical**

Panel-A

	Fail (1)	No crdt (2)	Drop (3)	Drop (4)	Grad (5)	Time Grad (6)	Grad Mark (7)
POST*SCIENCE	.264***	.217***	.238***	.220***	-.144***	3.03**	-1.95
	(.077)	(.066)	(.067)	(.079)	(.049)	(1.26)	(2.05)
SCIENCE	.0588***	.047***	.069***	.0449***	-.201***	3.66***	3.44***
	(.017)	(.015)	(.017)	(.013)	(.027)	(.657)	(1.18)
POST	-.044***	-.041***	-.033**	-.015*	.046	-.334	.892
	(.013)	(.012)	(.013)	(.009)	(.034)	(.533)	(1.81)
Constant	.115***	.101***	.154***	.092***	.656***	41.95***	97.47***
	(.012)	(.010)	(.011)	(.008)	(.016)	(.332)	(1.07)

Panel-B

	Fail (1)	No crdt (2)	Drop (3)	Drop (4)	Grad (5)	Time Grad (6)	Grad Mark (7)
POST*SCIENCE	.242***	.201***	.221***	.208***	-.11**	2.04*	-.72
	(.07)	(.06)	(.06)	(.07)	(.04)	(1.1)	(1.6)
SCIENCE	.088***	.068***	.088***	.052***	-.25***	4.76***	.841
	(.01)	(.01)	(.01)	(.01)	(.02)	(.64)	(1.0)
POST	-.02**	-.02**	-.01	-.007	.024	-.06	.263
	(.01)	(.01)	(.01)	(.00)	(.02)	(.44)	(1.3)
HS Mark	-.003***	-.003***	-.003***	-.001***	.008***	-.16***	.295***
	(.0006)	(.0005)	(.0004)	(.0003)	(.0006)	(.01)	(.01)
Professional	.125***	.117***	.118***	.063***	-.17***	2.19***	-4.9***
	(.02)	(.01)	(.01)	(.01)	(.01)	(.42)	(.53)
Technical	.067***	.066***	.068***	.032***	-.07***	1.72***	-3.3***
	(.01)	(.01)	(.01)	(.01)	(.01)	(.31)	(.29)
Classical	.038*	.040**	.058***	.057***	-.08***	1.20***	-1.0*
	(.02)	(.02)	(.01)	(.01)	(.02)	(.43)	(.58)
Linguistic	.062***	.058***	.067***	.048**	-.07***	2.37***	-3.5***
	(.01)	(.01)	(.02)	(.01)	(.02)	(.42)	(.46)
Others	.045***	.045***	.049***	.025**	-.07***	1.45***	-2.9***
	(.01)	(.01)	(.01)	(.01)	(.02)	(.39)	(.49)
Age	.003***	.003***	.003***	-.0006	-.007***	-.14***	.176***
	(.003)	(.003)	(.001)	(.0009)	(.0009)	(.03)	(.02)
Female	-.01	-.006	-.01	.011	.066***	-.51	-.16
	(.01)	(.009)	(.01)	(.01)	(.01)	(.25)	(.28)
Constant	.292***	.214***	.320***	.175***	.217**	56.9***	73.9***
	(.07)	(.07)	(.06)	(.05)	(.09)	(1.5)	(1.9)
Region fixed effects	YES	YES	YES	YES	YES	YES	YES
N	7822	7822	7822	7822	7230	4326	4326

Notes: Table 3 reports the estimation results from Equation 1 for the sample of degrees in science and in medical school. Panel A shows the results when the control variables are not included into regressions, while Panel B shows the results when the control variables are included. (1) is the probability of not passing any exam in the first year. (2) is the probability of dropping out in the first year without passing any exam. (3) is the probability of dropping out in the first year. (4) is the probability of official dropout in the first year. (5) is the probability of graduation. (6) is time to graduation (conditional on graduation), (7) is the final graduation mark (conditional on graduation). Standard errors are in parentheses and clustered at the degree program and academic year level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: **Estimation results: Social Sciences and Humanities vs. Medical**

Panel-A

	Fail (1)	No crdt (2)	Drop (3)	Drop (4)	Grad (5)	Time Grad (6)	Grad Mark (7)
POST*SOC.&HUM.	.116***	.071***	.048**	.049***	-.072*	-.313	-3.15
	(.026)	(.023)	(.023)	(.016)	(.036)	(.927)	(2.21)
SOC.&HUM.	.095***	.080***	.105***	.031***	-.214***	3.70***	2.14
	(.015)	(.014)	(.015)	(.012)	(.023)	(.432)	(1.44)
POST	-.043***	-.037**	-.031*	-.018	.056*	.161	1.00
	(.015)	(.015)	(.016)	(.011)	(.029)	(.565)	(1.64)
Constant	.120***	.104***	.157***	.095***	.648***	41.81***	97.38***
	(.013)	(.011)	(.012)	(.009)	(.018)	(.307)	(1.240)

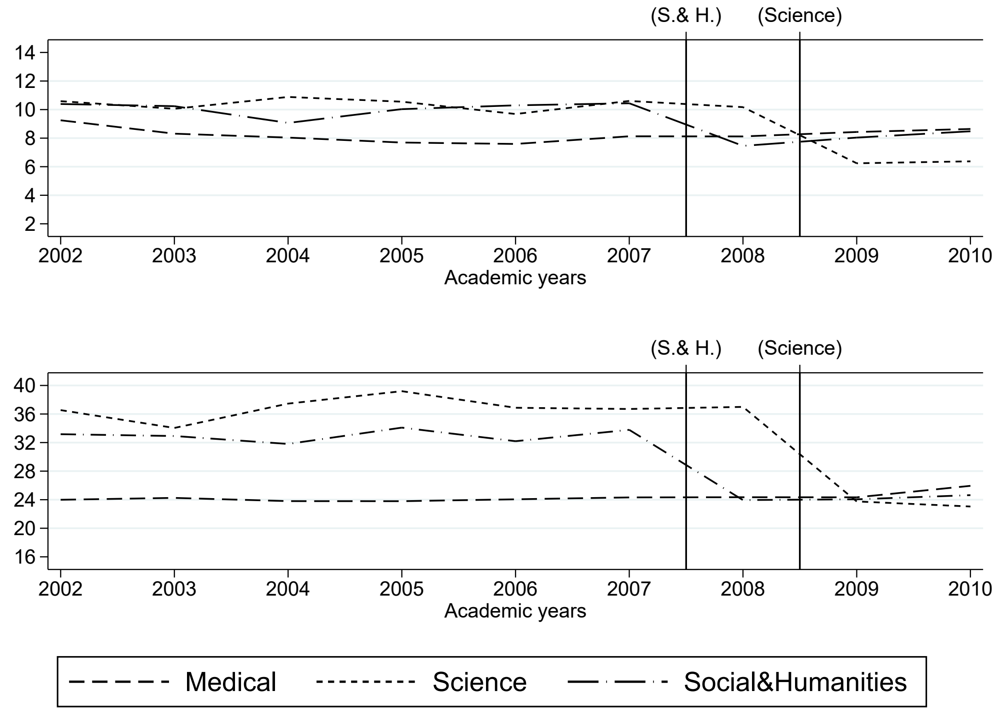
Panel-B

	Fail (1)	No crdt (2)	Drop (3)	Drop (4)	Grad (5)	Time Grad (6)	Grad Mark (7)
POST*SOC.&HUM.	.101***	.059***	.035	.041**	-.04	-1.0	-2.1
	(.02)	(.02)	(.02)	(.01)	(.02)	(.87)	(1.8)
SOC.&HUM.	.110***	.089***	.112***	.034***	-.24***	4.93***	-.10
	(.01)	(.01)	(.01)	(.01)	(.02)	(.44)	(1.2)
POST	-.02*	-.02*	-.01	-.01	.032	.517	.685
	(.01)	(.01)	(.01)	(.01)	(.02)	(.49)	(1.3)
HS Mark	-.004***	-.003***	-.004***	-.001***	.009***	-.18***	.313***
	(.0004)	(.0003)	(.0002)	(.0003)	(.0002)	(.01)	(.01)
Professional	.130***	.115***	.121***	.058***	-.22***	3.01***	-4.6***
	(.01)	(.01)	(.01)	(.01)	(.02)	(.51)	(.44)
Technical	.076***	.066***	.070***	.038***	-.11***	1.85***	-3.6***
	(.01)	(.01)	(.01)	(.009)	(.01)	(.28)	(.27)
Classical	.022	.020	.043*	.029**	-.05***	.734*	.471
	(.01)	(.01)	(.01)	(.01)	(.01)	(.41)	(.46)
Linguistic	.061***	.053***	.065***	.033***	-.12***	2.90***	-1.8***
	(.01)	(.01)	(.01)	(.01)	(.02)	(.45)	(.41)
Others	.059 ***	.049***	.048***	.014	-.10***	1.96***	-2.6***
	(.01)	(.01)	(.01)	(.01)	(.01)	(.29)	(.39)
Age	.004***	.003***	.006***	-.001	-.01***	-.05**	.302***
	(.0008)	(.0007)	(.0009)	(.0002)	(.001)	(.02)	(.03)
Female	-.03***	-.03***	-.03***	-.009	.044***	-.19	1.26***
	(.008)	(.008)	(.008)	(.007)	(.01)	(.18)	(.37)
Constant	.343***	.277***	.358***	.229***	.248***	55.1***	70.1***
	(.06)	(.05)	(.05)	(.03)	(.06)	(1.0)	(2.2)
Region fixed effects	YES	YES	YES	YES	YES	YES	YES
N	12385	12385	12385	12385	11467	5957	5957

Notes: Table 4 reports the estimation results from Equation 1 for the sample of degrees in social sciences, in humanities and in medical school. Panel A shows the results when the control variables are not included into regressions, while Panel B shows the results when the control variables are included. (1) is the probability of not passing any exam in the first year. (2) is the probability of dropping out in the first year without passing any exam. (3) is the probability of dropping out in the first year. (4) is the probability of official dropout in the first year. (5) is the probability of graduation. (6) is time to graduation (conditional on graduation), (7) is the final graduation mark (conditional on graduation). Standard errors are in parentheses and clustered at the degree program and academic year level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

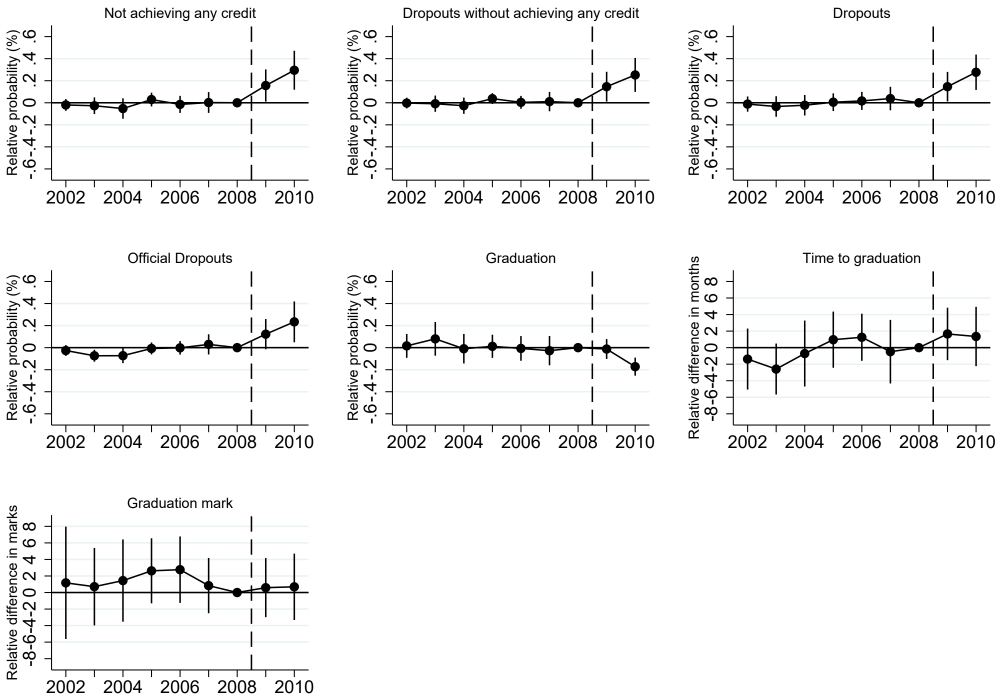
Figures

Figure 1: **Number of exams**



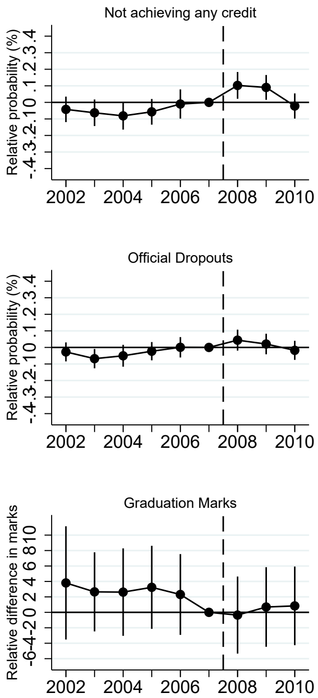
Notes: Figure 1 plots the average number of exams taken by students who completed their degrees on time, across scientific fields, over the academic years. The figure on top highlights the number of exams in the first year; the figure in the bottom shows the number of exams for the entire duration of the degree program. Vertical lines represents the introduction of the reform for programs in social sciences and humanities (S.& H.) and in science.

Figure 2: **Event-study Specification: Science vs. Medical**



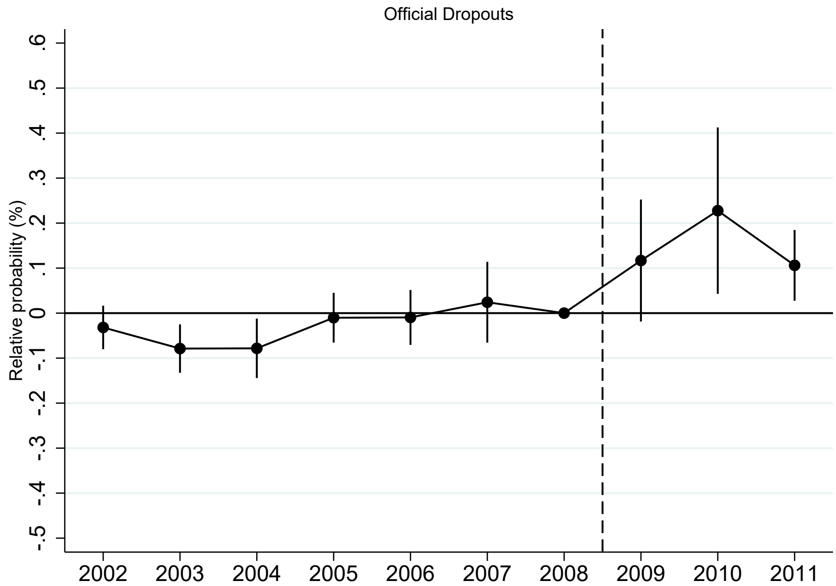
Notes: Figure 2 highlights the coefficient estimates of δ_k specified in Equation 2 for the sample of degrees in science and in medical school. Confidence intervals are at 90% level.

Figure 3: **Event-study specification: Social Sciences and Humanities vs. Medical**



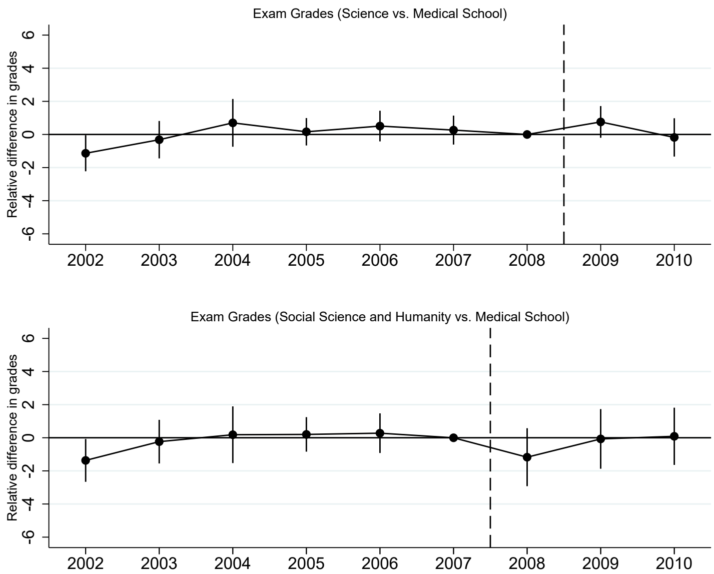
Notes: Figure 3 highlights the coefficient estimates of δ_k specified in Equation 2 for the sample of degrees in social sciences, in humanities and in medical school. Confidence intervals are at 90% level.

Figure 4: **Event-study Specification: Science vs. Medical**



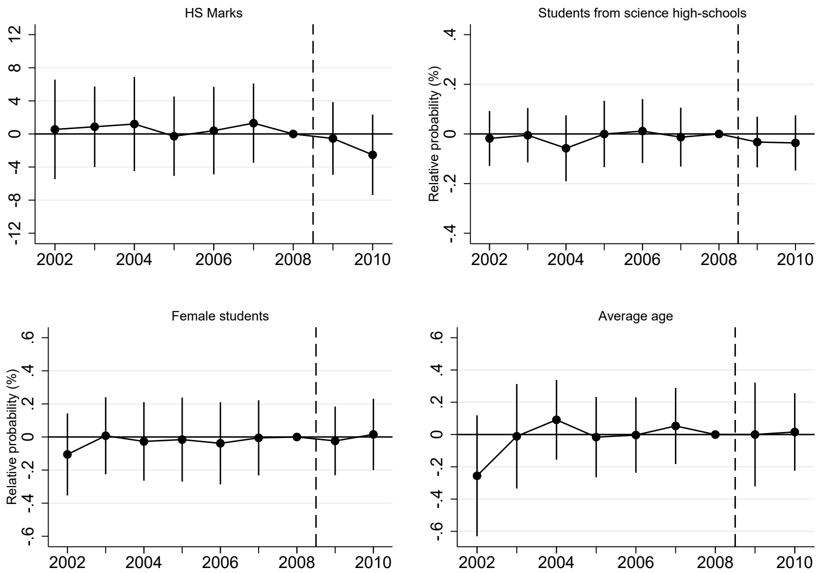
Notes: Figure 4 highlights the coefficient estimates of δ_k specified in Equation 2 for the sample of degrees in science and in medical School. The outcome variable is the first year official dropout. Confidence intervals are at 90% level.

Figure 5: **Event-study Specification on Grades**



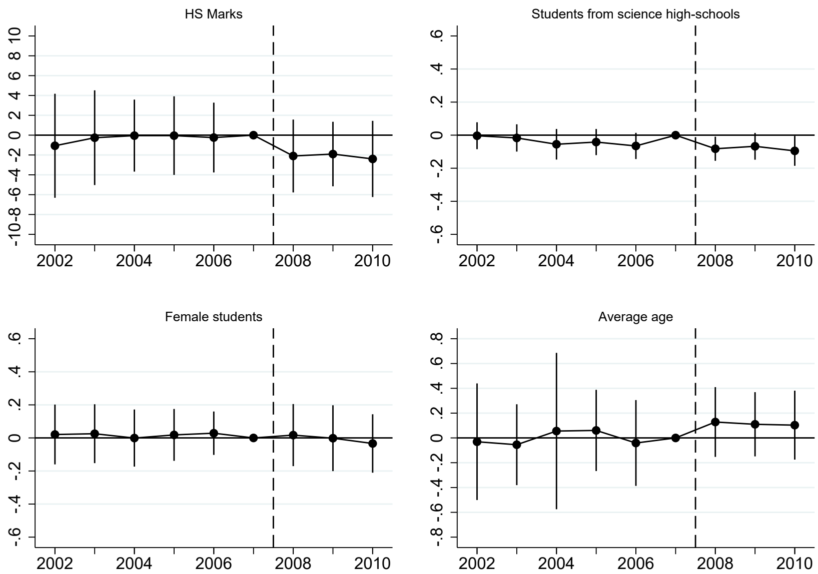
Notes: Figure 5 highlights the coefficient estimates of δ_k specified in Equation 2 for both of the estimation samples. The outcome variable is the grades of students (conditional on passing exam). Confidence intervals are at 90% level.

Figure 6: **Event-study Specification for Student Composition: Science vs. Medical**



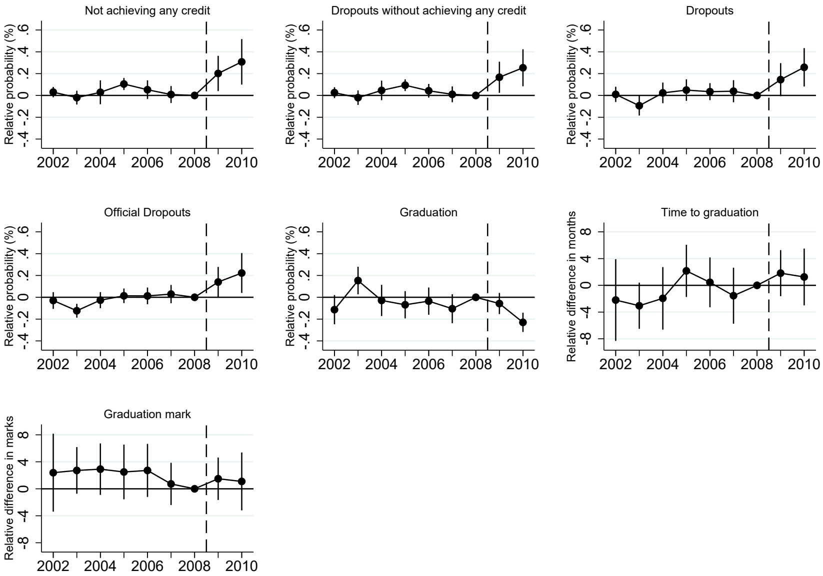
Notes: Figure 6 highlights the coefficient estimates of δ_k specified in Equation 2 for the sample of degrees in science and medical School. Confidence intervals are at 90% level.

Figure 7: **Event-study specification for students composition: Social sciences and Humanities vs. Medical**



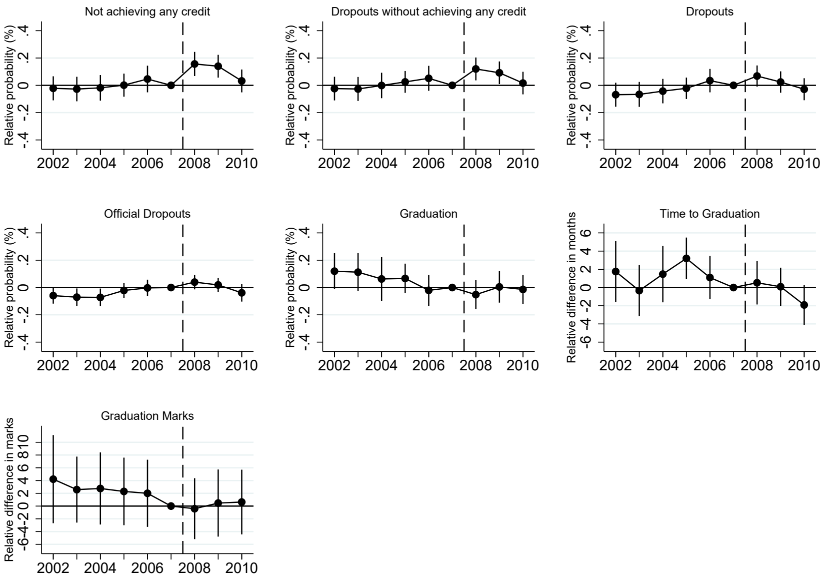
Notes: Figure 7 highlights the coefficient estimates of δ_k specified in Equation 2 for the sample of degrees in Social Sciences, in Humanities and in Medical School.

Figure 8: **Event-study specification (matched sample): Science vs. Medical**



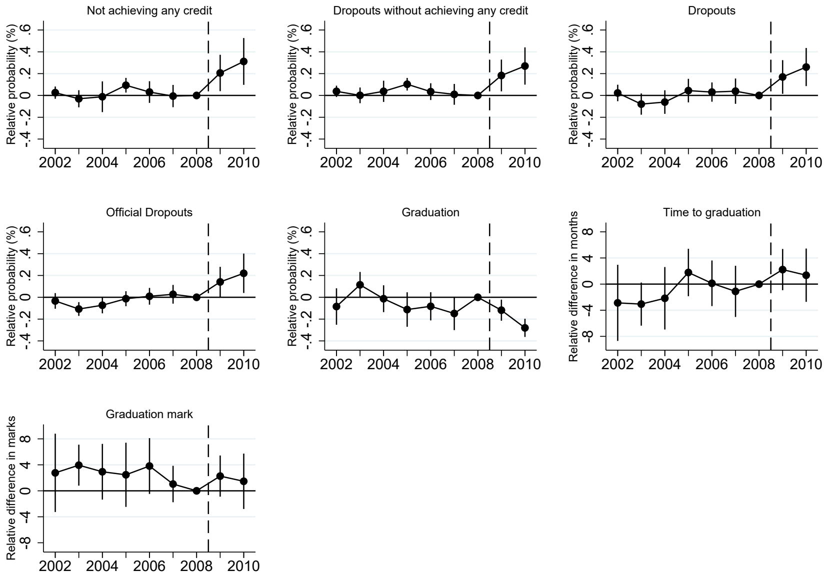
Notes: Figure 8 highlights the coefficient estimates of δ_k specified in Equation 2 for the sample of degrees in Science and in Medical School.

Figure 9: **Event-study Specification (Matched Sample): Social Sciences and Humanities vs. Medical**



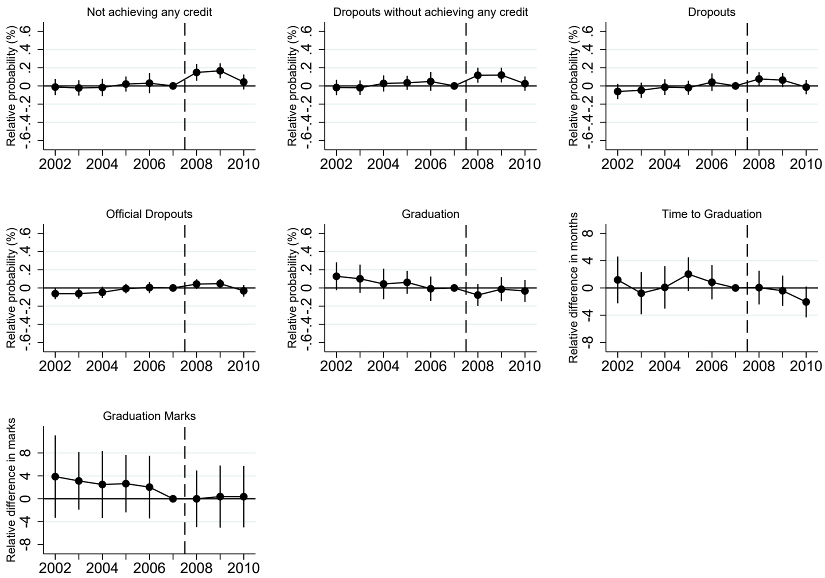
Notes: Figure 9 highlights the coefficient estimates of δ_k specified in Equation 2 for the sample of degrees in social sciences, in humanities and in medical School. Confidence intervals are at 90% level.

Figure 10: **Event-study specification (matched sample): Science vs. Medical**



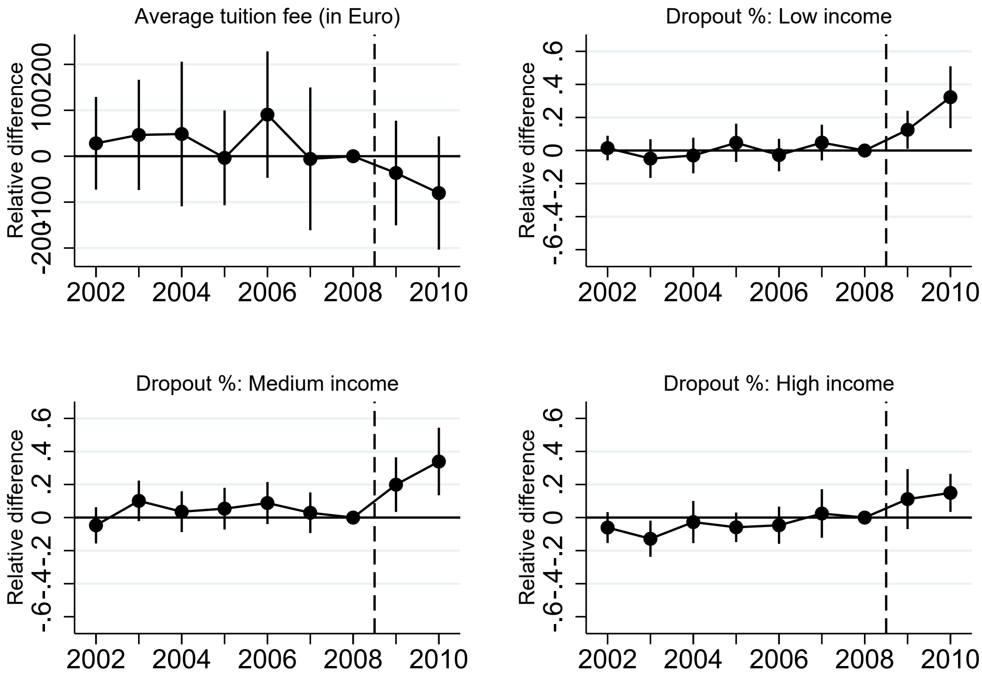
Notes: Figure 10 highlights the coefficient estimates of δ_k specified in Equation 2 for the sample of degrees in Science and in Medical School. Confidence intervals are at 90% level.

Figure 11: **Event-study Specification (Matched Sample): Social Sciences and Humanities vs. Medical**



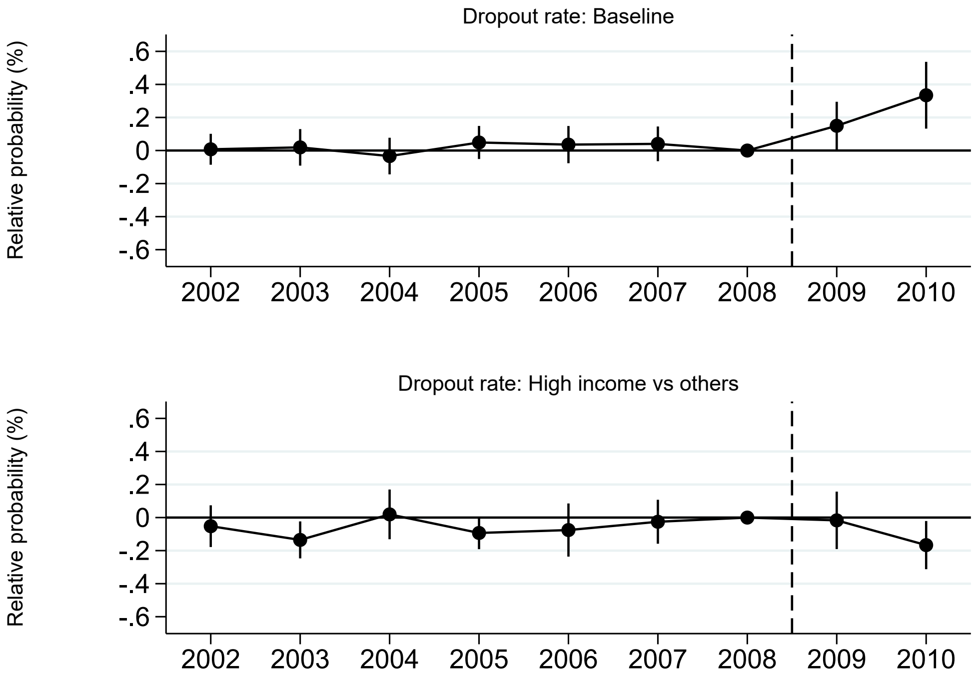
Notes: Figure 11 highlights the coefficient estimates of δ_k specified in Equation 2 for the sample of degrees in social sciences, in humanities and in medical School. Confidence intervals are at 90% level.

Figure 12: **Event-study Specification : Science vs. Medical**



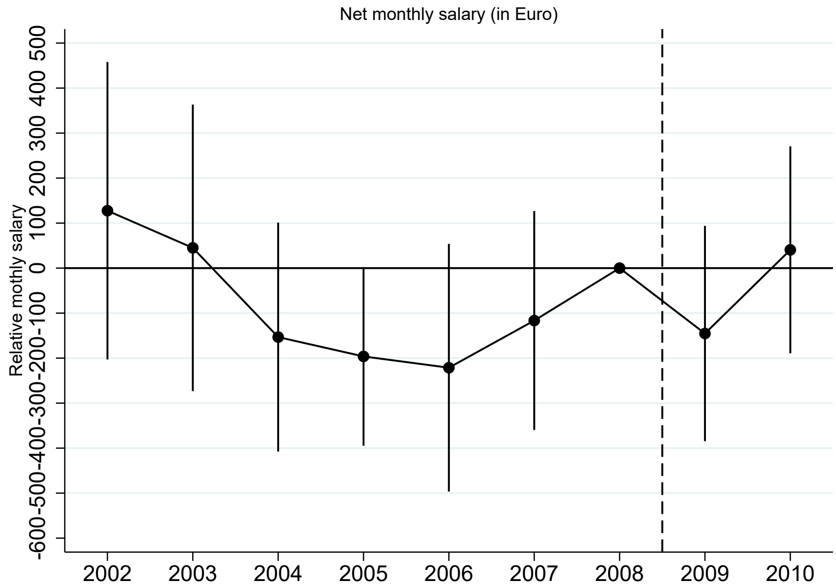
Notes: Figure 12 highlights the coefficient estimates of δ_k specified in Equation 2 for the sample of degrees in Science and in Medical School. Confidence intervals are at 90% level.

Figure 13: **Event-study Specification (by high income) : Science vs. Medical**



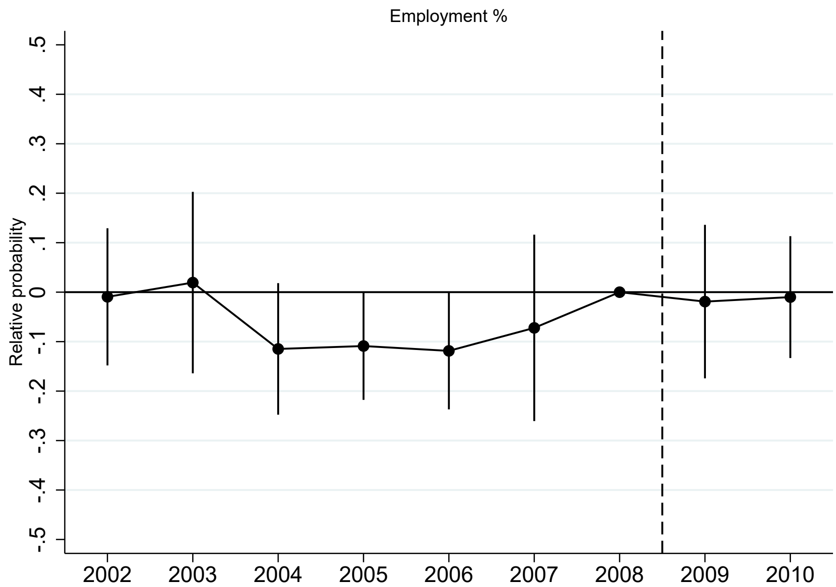
Notes: Figure 13 highlights the coefficient estimates of δ_k (graph on the top) and the coefficient estimates of ω_k (graph on the bottom) specified in Equation 3 for the sample of degrees in Science and in Medical School. Confidence intervals are at 90% level.

Figure 14: **Event-study Specification : Science vs. Medical**



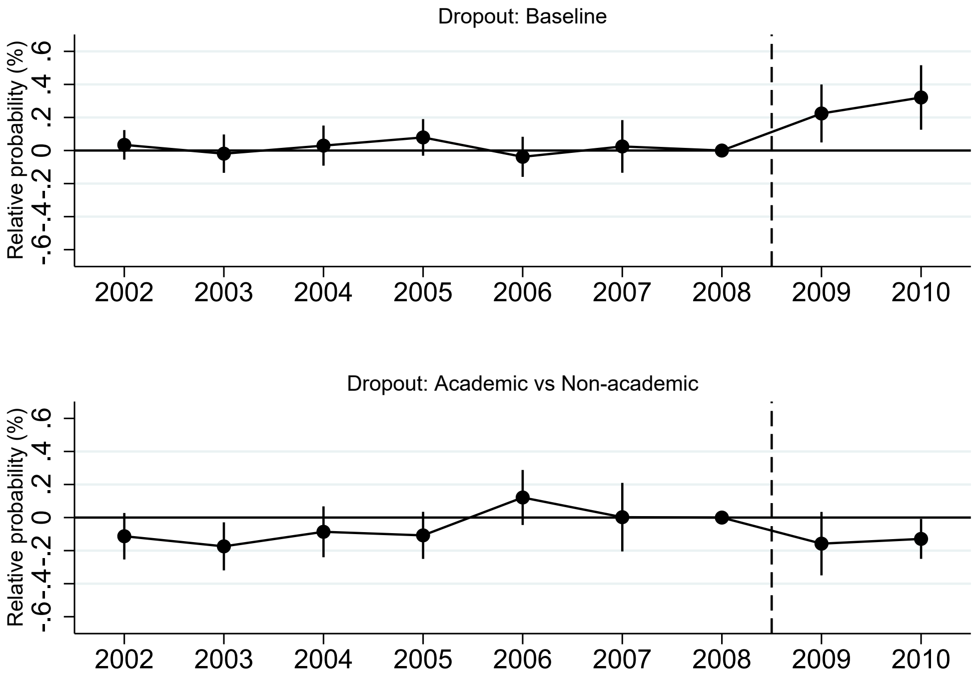
Notes: Figure 14 highlights the coefficient estimates of δ_k specified in Equation 2 for the sample of degrees in Science and in Medical School. Confidence intervals are at 90% level.

Figure 15: **Event-study Specification : Science vs. Medical**



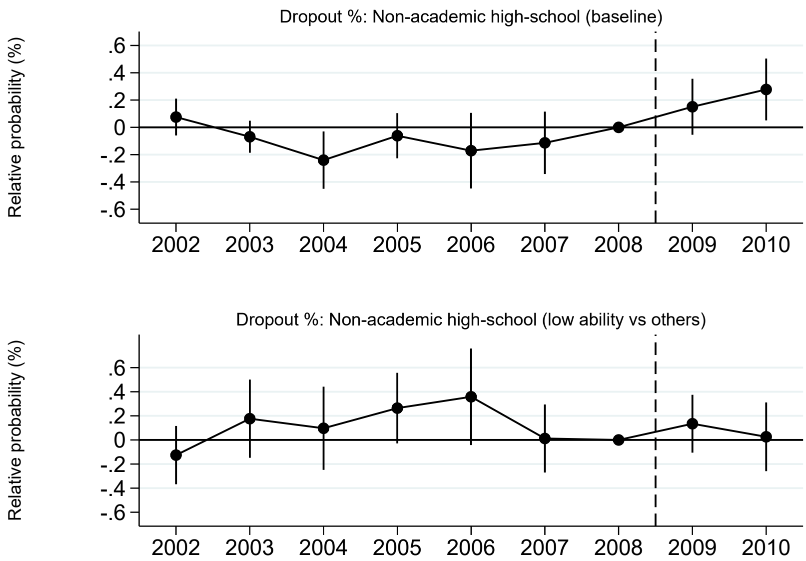
Notes: Figure 15 highlights the coefficient estimates of δ_k specified in Equation 2 for the sample of degrees in Science and in Medical School. Dependent variable is the employment status at the moment of interview. Confidence intervals are at 90% level.

Figure 16: **Event-study Specification : Science vs. Medical**



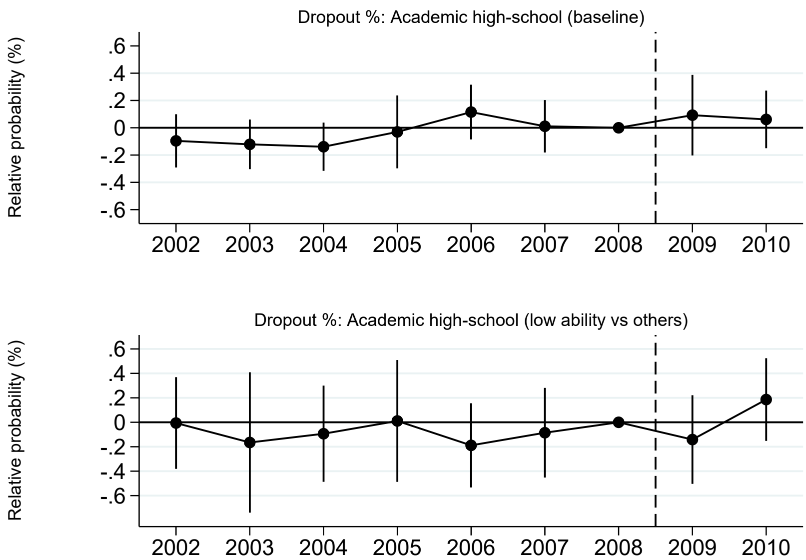
Notes: Figure 16 highlights the coefficient estimates of δ_k (graph on the top) and the coefficient estimates of ω_k (graph on the bottom) specified in Equation 3 for the sample of degrees in Science and in Medical School. Confidence intervals are at 90% level.

Figure 17: **Event-study Specification : Science vs. Medical**



Notes: Figure 17 highlights the coefficient estimates of δ_k (graph on the top) and the coefficient estimates of ω_k (graph on the bottom) specified in Equation 3 for the sample of degrees in Science and in Medical School. Confidence intervals are at 90% level.

Figure 18: **Event-study Specification : Science vs. Medical**



Notes: Figure 18 highlights the coefficient estimates of δ_k (graph on the top) and the coefficient estimates of ω_k (graph on the bottom) specified in Equation 3 for the sample of degrees in Science and in Medical School. Confidence intervals are at 90% level.