Labour Economics 90 (2024) 102559

Contents lists available at ScienceDirect

Labour Economics

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

Effects of workload allocation per course on students' academic outcomes: Evidence from STEM degrees[☆]

Carmen Aina [a], Koray Aktaş [b,*], Giorgia Casalone [c,d]

[a] Department of Sustainable Development and Ecological Transition, Università del Piemonte Orientale, Italy
[b] Department of Economics, Università Ca' Foscari Venezia, Italy
[c] Department of Economics and Business, Università del Piemonte Orientale, Italy
[d] Dondena Centre for Research on Social Dynamics and Public Policy, Università Bocconi, Italy

ARTICLE INFO

JEL classification:
I23
I26
I28
J21
J24

Keywords:
Number of courses
University dropout
Graduation
Credits per course
Intensity of workload
Procrastination

ABSTRACT

We investigate how the allocation of workload across university courses affects students' outcomes. Using a difference-in-differences design, we provide novel evidence that reducing the number of courses in a degree, while keeping the total course work unchanged, strongly reduces students' performance and increases first-year dropout rates. We show procrastination accounts for these effects, suggesting that students struggle to adjust their study time to handle the intensified courses. We also show that the adverse impacts on dropouts are significantly stronger for students from less affluent families, indicating that the reform likely increases inequality. On the other hand, post-reform graduates exhibit better labor market outcomes. The discussion on potential mechanisms suggests that the reform enhanced the skills of the graduates who successfully navigated the unified exams.

1. Introduction

When preparing for examinations, students must allocate their time efficiently to ensure that they study a certain amount of material proposed by their course teachers. Several studies have investigated the impact of the number of years of schooling at different educational levels (Card 1999, Pischke 2007, Cappellari and Lucifora 2009, Krashinsky 2014, Marcus and Zambre 2018), the amount of coursework during college degrees (Arteaga, 2018), and organization of academic calendars (Bostwick et al., 2022) on students' outcomes. However, no evidence exists on how workload allocation across university courses within degree programs influences college students' academic performance.

Let us consider a simple situation wherein a college student takes two mathematics courses—say, mathematics I and II; each course is associated with five credits.[1] Accordingly, they sit two final exams to successfully complete these courses at the end of the semester. In a different situation, a student takes a mathematics course with ten credits and sits an overall examination, and hence they have to study more course material per exam. A priori, the effects of the increased

☆ Koray Aktaş acknowledges financial support from the Università del Piemonte Orientale, Italy. We thank the three anonymous referees and the various conference and workshop participants for their comments and suggestions. We are thankful to Tommaso Colussi and Diogo G.C. Britto for their insightful comments and suggestions during this paper's development. Furthermore , this paper has benefited from conversations with Ingo Isphording, Ulf Zölitz, Anne Boring, Enrico Cantoni, Fabian Waldinger, Daniela Sonedda, Marco Manacorda, Paolo Ghinetti, Maria De Paola, and Piergiovanna Natale. We also thank the administrative staff of the Università del Piemonte Orientale for providing us the student university records – especially Andrea Turolla and Emanuela Rosetta – and the AlmaLaurea consortium for the data on employment conditions of graduates at the Università del Piemonte Orientale. Finally, we sincerely thank the colleagues of the Department of STEM degree programs - Mauro Botta, Jean Daniel Coisson, Lorella Giovannelli, Leonardo Marchese and Paolo Paiuzzi - and of the Medical School - Simone Binello and Daniela Gentile - for their valuable support in reconstructing both study plans and the methods of managing courses after the introduction of the reform.

* Correspondence to: Palazzo Moro, Cannaregio, 2978, 30123, Venice, Italy.
E-mail addresses: carmen.aina@uniupo.it (C. Aina), koray.aktas@unive.it (K. Aktaş), giorgia.casalone@uniupo.it (G. Casalone).

[1] University credit is the metric used in European universities to evaluate the workload an average student needs to achieve the expected learning outcomes (see Section 3 for further detail).

https://doi.org/10.1016/j.labeco.2024.102559
Received 9 January 2023; Received in revised form 29 April 2024; Accepted 29 April 2024
Available online 3 May 2024
0927-5371/

workload per course are ambiguous. On the one hand, studying for the unified exam can help students to enhance their knowledge of the subject, while sitting fewer exams can also mean less stress and lower fatigue for students (Goulas and Megalokonomou, 2020). This can foster non-cognitive skills, as students must learn to better manage their time via self-imposed deadlines. On the other hand, evidence reveals that students procrastinate and therefore benefit more from externally imposed deadlines (Ariely and Wertenbroch, 2002). Tuckman (1998) considers frequent testing to be a motivational equalizer. When students sit frequent exams, they receive greater feedback (De Paola and Scoppa, 2011). The results of an experiment by Buser and Peter (2012) demonstrate that individuals who multitask and are free to organize their own schedules perform worse than those who are forced to work on tasks sequentially. Moreover, students must study and process a large amount of information for a unified exam, which can culminate in worsening academic performance.

This study documents the first evidence in the literature on how workload reallocation among university courses, which increases the average workload per course, affects college students' key academic outcomes. Accordingly, we exploit the exogenous variations in the number of courses, holding constant the total number of credits that students must achieve to graduate across degree programs and academic years. Further, we estimate the effects of the reform on graduates' early labor market performance.

Our identification strategy builds on the exogenous variation in the number of courses (i.e., exams)[2] induced by the nationwide tertiary education reform in Italy. The reform sought to standardize degree programs' organization across country and reduce the time taken to complete a degree.[3] To homogenize degree programs and avoid the excessive fragmentation of curricula, the reform sets a maximum cap for the total number of courses required for degree completion (for instance, 20 courses to complete a three-year bachelor degree) without altering the number of credits required to achieve the degree (180 credits for a three-year bachelor degree). To comply with the reform requirements, degree programs with more courses than the maximum number set by the law had to change their curricula, predominantly by integrating two or, in some cases, three courses into one. Hence, the workload intensity per course increased mechanically.

We use the administrative dataset of Università del Piemonte Orientale (UPO), which contains all of the details regarding the students' enrollment records and performance during their degrees, as well as information regarding these students' pre-enrollment characteristics. At the UPO, the reform affected degree programs differently. For example, students enrolled in the bachelor STEM programs required an average of 33 courses to attain their degrees and eight courses in their first year,[4] whereas after the reform these numbers fell to 20 for degree completion and five in the first academic year. However, the number of credits required to graduate (180 credits) as well as their distribution across years (about 60 credits per year) did not change. By contrast, the number of courses required for bachelor's degrees in healthcare sciences organized by the medical school remained unchanged because the course numbers were already at the level required by the reform.[5]

Therefore, a quasi-experimental environment was established following the reform's implementation. Such a setup enables identifying causal effects and renders the UPO dataset's use appropriate. The law passed in 2007 compelled all universities to revise their degrees by the beginning of the 2009/10 academic year. In our setup, STEM degrees reorganized their course schedules from the 2009/10 academic year. In the difference-in-differences framework, we consider the three-year degree programs in STEM subjects as treated units, whereas the three-year degree programs offered by the medical school are considered as the control group. We estimate the effect of the reallocation of workload among university courses – and the consequent increase in the average workload per course – by comparing the differences in the outcomes of the treatment and control groups before and after the reform's implementation. Throughout the paper, we present results on varied academic outcomes, focusing on the effects on first-year dropout and degree completion rates. These two academic outcomes are the most policy-relevant indicators of higher education efficiency (Aina et al., 2022). Additionally, they are not subject to endogeneity problems as they are available to all students. Our working sample covers the academic years at first enrollment (i.e., matriculation) from 2002/03 to 2014/15.

Our results demonstrate that after the introduction of unified exams, students in STEM degrees are, on average, up to 22 percentage points (p.p.) more likely to fail all their exams in their first year, which increases the first-year dropout probability by 20 p.p. We provide further evidence that the effects are more pronounced for high-procrastinators, suggesting that students fail to effectively organize their study time for exam preparations.[6] The increase in the first-year dropout rate translates into a decrease in the graduation rate by approximately 23 p.p. Event-study analysis reveals that the parallel trend assumption holds in regressions. These estimates further demonstrate that our findings are not temporary – for example, driven by the transition effect – because the increase in the first-year dropout rate remains significant; even after six years. These results are also robust to controlling for student composition as we construct a sample by exact matching on key pre-enrollment characteristics of students. Furthermore, we show that the intensity of treatment matters for the timing of the dropout decision of students. The first-year dropout rate of students in STEM degrees that are less affected by the reform does not significantly change after the reform, although their graduation rate decreases. On the other hand, students in more affected STEM degrees, who experience greater changes in their exams, have a significantly higher probability of dropping out at the end of their first year, suggesting that the changes in the exams discouraged these students earlier. However, we do not observe any significant effect on the other outcomes related to degree completion, namely time to graduation (measured in months) and final graduation marks.

Finally, we examine the early labor market performance of students who achieved their degree after the reform using the survey data of *AlmaLaurea* on Italian graduates' employment conditions one year after graduation.[7] We observe sizable effects of unified exams throughout their degrees on the employment rates of graduate students. For graduates with STEM degrees, we estimate an average 25.6 p.p. increase in the employment rate. We provide direct and indirect evidence of two potential mechanisms behind this finding: the peer effect channel, and the improved skills channel. We conclude that the latter better explain our results.

[2] In the Italian university system, all university courses are assessed via an exam, as explained in Section 3. Accordingly, in this paper, we use the terms "course" and "exam" interchangeably. Details regarding the reform and Italian higher education system are discussed in Section 3.

[3] For an analysis of the time-to-degree in Italy, see Aina et al. (2011).

[4] STEM stands for Science, Technology, Engineering, and Mathematics. UPO offers STEM bachelor degrees in biotechnology, biology, mathematics, environmental sciences, materials science, and computer science.

[5] For simplicity, we refer to students enrolled in healthcare sciences bachelor degrees as "medical school students" throughout the paper. Our working sample invariably comprises students enrolled in three-year bachelor degrees only (in both the treatment and control groups).

[6] We follow the work by De Paola and Scoppa (2015) to compute a proxy of procrastination in the Italian higher education context. This procedure's details are discussed in Section 7.

[7] *AlmaLaurea* – an inter-university consortium representing 90% of Italian graduates – surveys the profile and employment status of graduates after one, three, and five years. It is funded by the member universities, the Ministry of University and Research, and companies and bodies that use the services offered.

The remainder of this paper proceeds as follows. In Section 2, we frame our analysis within the literature. In Section 3, we provide details regarding the Italian tertiary education system and the reforms that we leverage to identify our estimates. In Section 4, we present the dataset and descriptive evidence. In Section 5, we discuss the identification strategy, outline the econometric specifications, and explain the exact matching procedure. In Section 6, we present the results on academic outcomes, sensitivity tests, and the findings on labor market outcomes Section 7 provides further evidence on underlying mechanisms Section 8 presents heterogeneous analysis by observable characteristics of students; in Section 9, we conclude the study.

2. Related studies and contribution to the literature

Our study contributes to several strands of literature. First, we provide novel empirical evidence on students' preferences regarding task management pertaining to preparing for final exams that students must pass to earn the corresponding university credits. De Paola and Scoppa (2011)'s study, which is related to our work, reveals that students who sit midterm exams are more likely to pass the final exam and obtain higher grades than students who only sit the final exam.[8] Goulas and Megalokonomou (2020) provide evidence that exam scheduling has a significant effect on the exam performance of high-school students, especially in STEM subjects. Their results suggest that students' performance in exams increases with the number of exams that they take, which the authors call the *warm-up effect*. They also show that the greater the time that students have between their exams, the worse they perform (*fatigue effects*). We specifically analyze the final academic outcomes of students in addition to exam performance, assessing the effects of the intensity of workload per exam rather than the effects of taking mid-term exams or the effects of exam scheduling.

Second, we contribute to studies on the organization of academic calendars. While Bostwick et al. (2022) demonstrate that switching from quarters to semesters negatively affects college students, increasing the first-year dropout rates and reducing the rate of graduating on time, we provide evidence on a narrower aspect of the organization of academic calendars in terms of the number of final exams that students sit, holding constant the duration of the calendar and total credits to be achieved by students. Another recent study by Arteaga (2018) provides evidence that a reduction in the amount of coursework required for college degree completion adversely affects human capital accumulation and leads to reduced wages. We contribute to the literature by documenting that coursework per exam severely affects students' performance.

Third, our work contributes to the ever-growing literature on STEM majors in tertiary education that produce skilled labor, which is crucial for economic growth (Peri et al., 2015).[9] As documented by Chen (2013), the dropout rate in STEM majors is extremely high, and students in these degrees are more likely to switch their majors. These students are also known to be overconfident regarding their skills at the time of college enrollment (Stinebrickner and Stinebrickner, 2013). We demonstrate that, in addition to these students' personal traits, the organization of degree courses in these scientific fields is crucial and significantly affects academic performance.

Fourth, we contribute to the literature exploring the role of students' ability – generally evaluated based on previous school achievements – and family income in university dropout and completion. Overall, the evidence suggests that high-school type and grade are strong predictors of college academic retention. For example, Danilowicz-Gösele et al. (2017) find that final high-school marks are the most important determinants of the probability of obtaining both a degree and a high final grade. Similarly, a large body of economic literature investigates the effect of parental financial resources on university enrollment and completion by suggesting that family income becomes substantially more significant over time (Belley and Lochner, 2007). However, this relationship is driven by the tertiary education system's specific nature. For instance, Stinebrickner and Stinebrickner (2003) demonstrate that less-privileged students exhibit higher dropout rates, even in the absence of direct education costs. Nevertheless, Stinebrickner and Stinebrickner (2008) find that dropout is only partially driven by financial conditions, as it remains persistent once credit constraints are eliminated. Additionally, Bound et al. (2012) find that the increasing difficulties in financing university education, especially for students from low-income families, have increased student employment to cover college costs, thus extending their time in university. Our heterogeneous results confirm that affluent students are more likely to persist and attain degrees. Moreover, high-ability students are less affected by the reform, as they are significantly more capable of managing this intensive workload per course.

Furthermore, our study contributes to the policy debate about higher education institutes' low efficiency in Italy.[10] According to OECD (2018), despite an increase in educational achievement, predominantly observed in recent cohorts, Italy remains at the bottom of the education distribution for OECD countries and exhibits a persistent gap compared with other developed countries.[11] Our findings have policy implications related to higher education efficiency in Italy, as we demonstrate that excessively increasing the course workload significantly negatively affects dropout and graduation rates.

3. Institutional settings and the reform

The Italian university system is organized into three cycles according to the Bologna reform process[12]: the first cycle is a three-year bachelor's degree, the second cycle is a two-year master's degree, and the third cycle is a three-year PhD program. For specific programs (e.g. medicine, law, pharmacy), there is a "single-cycle degree" lasting for five to six years. All individuals holding an upper secondary school diploma are eligible to enroll in first- or single-cycle degree programs. Some degree programs have selective admission procedures that can be implemented at either the university or national level.

Degree programs are structured in university credits according to the European Credit Transfer System. Each credit corresponds to approximately 25 h of student workload, including both lectures and

[8] The authors conducted a randomized field experiment involving only a cohort of Economics students attending the introductory course of microeconomics and macroeconomics in a public university in southern Italy to test the effect of examination frequency and interim feedback provision on achievement. These students were divided into the treatment group, which could split the exam into two parts (i.e. an intermediate test and a final one), and the control group, which could only take the final comprehensive exam. The results demonstrated that receiving intermediate feedback and breaking down the exam led to better outcomes. In practice, they estimate a consequence of the Mussi's reform, namely the merging of two or more courses with the option to take only a comprehensive exam, but not the reform itself. In our paper, unlike theirs, we analyze the impact of this reform on various academic outcomes to assess its short and medium-term effects as well as the early labor market outcomes.

[9] In our setup, the science degree courses include science, technology, and mathematics, but not engineering.

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

[11] For example, 18% of Italy's population aged 25–64 had completed tertiary education in 2015, while the average in OECD countries was approximately 35%. Regarding the youngest cohort (aged 25–34), the figures were 25% and 42%, respectively. The low number of graduates is significantly related to persistent high dropout rates from the Italian university system, despite efforts to increase retention by the Italian Ministry of Education over the years.

[12] See Bratti et al. (2008) and Cappellari and Lucifora (2009) for further details on the Bologna process.

home study hours. To graduate within a degree program's legal duration, a student must achieve on average 60 credits per academic year.[13] The first-cycle degree is awarded to students who earn at least 180 credits, while second- and single-cycle degrees are awarded to students who achieve at least 120 and 300 credits (360 for six-year degrees in medicine and dentistry), respectively.

Within this framework, Ministry Decree 155/2007 (known as *Mussi's reform*) introduced a change that potentially affected students' behavior and outcomes. The change was to set a cap on the number of courses required to be taken (and, consequently, exams to be passed) to earn the degrees.[14] Until the 2007 reform, universities were allowed to freely determine the number of courses that composed the degree program curricula, as long as the minimum number of credits needed to earn the degree was 180, 300, or 120, depending on the university cycle. The autonomy allowed for universities outlying degree programs entailed two main consequences: first, degree programs became excessively fragmented, with a large number of small and barely coordinated courses; second, the heterogeneity in the setting of the study plans across universities made it difficult for students to move from one university to another, even within the same field of study. Such heterogeneous design of undergraduate curricula contrasted the spirit of the Bologna process, aiming "to bring more coherence to higher education systems across Europe [....], to facilitate student and staff mobility".[15]

The reform set a cap of 20 and 12 courses for first-cycle three-year degrees and second-cycle two-year degrees, respectively. In order to be compliant with the reform, degree programs that exceeded this threshold were forced to revise their study plans by integrating two or more courses, while keeping the contents of the educational programs unchanged. This change was explicitly meant to "improve the efficacy, quality and consistency of the degree programs" by promoting "the cooperation between professors of different courses".[16] Nevertheless, by keeping the total number of credits needed to graduate constant, this reform has had the unforeseen consequence of increasing the average workload (in terms of credits) for each course. Neither the Ministry Decree nor the reform guidelines provide any explanation regarding why the cap was set at this level. To the best of our knowledge, no indications exist regarding the most effective organization of university degree programs in the literature, and therefore the latter constraint seems rather arbitrary.

Universities were expected to apply Mussi's reform within two academic years (2008/09 and 2009/10) following its approval. The only constraint was that the whole degree programs offered at the university level in the same "degree class" had to apply the reform simultaneously.[17] Only post-reform freshmen were affected. The changes driven by the Ministry Decree 155/2007 had different impacts on various degree programs, even within the same disciplinary group. For instance, with reference to the STEM degrees at the UPO, the reform has exhibited varying levels of treatment intensity. Overall, despite the varying intensity of the treatment, in the post-reform reorganization the number of credits per academic year remained largely unchanged, and courses were grouped based on their similarities, with only some slight adjustments in the year when they were offered. To highlight these disparities, as an example; we present the study plans of two degree programs: materials science and computer science.[18] As shown in Table 1, the former is among the strongly affected degrees (i.e. *more affected* degrees), as the comparison before (i.e., 2008/09) and after (i.e., 2009/10) the reform reveals that during the first academic year the number of courses decreased from fifteen to seven, corresponding to an increase in the average number of credits per course from 4 to 8.4. The changes predominantly entailed the unification of small courses, inducing that students had to study (on average) 210 h (after reform) instead of 100 h (before reform) per exam as well as some reallocation across academic years. Similarly, in the second year, we also observe a reduction in the number of exams by seven, coupled with an average increase in credits of 5.4. On the other hand, Table 2 presents the study plan of computer science, which has experienced a moderate impact from the reform (i.e. *less affected* degrees) compared to the degree program discussed above. As can be seen, in both the first and second years of this degree program the number of courses has decreased by four with the reform, while the average number of credits has increased in each academic year by 2.7 and 3.4, respectively. Accordingly, in the first year the average study hours per course increased by 50 h, and in the second year by 85 h. Therefore, the reform induced a significant change in the organization and planning of study activities. In Fig. 1, we plot the number of academic staff with permanent contracts by gender for STEM degrees (Panel A) and in medical school (Panel B).[19] We distinguish three types of permanent occupations: full professors, associate professors and assistant professors.[20] We observe an increase in the number of male full professors and female assistant professors in the medical school but in the years before the reform (up to 2008). We do not observe any changes in the composition of lecturers that could be associated with the reform's implementation. We also plot the figure by only academic positions in Figure C1 (Appendix C).

To foresee the effect of the degree programs' new setting introduced by the 2007 law on students' outcomes, providing information on how the assessment activity is organized in the university under analysis, which is representative of the Italian university system, can be useful. At the end of each teaching semester, there is an "examination session": for each course taught in the semester, professors set three dates when students can sit the exam, and exams must be scheduled fourteen days apart. For example, if the first exam date of course X is 10th January, the second date must be set no earlier than 24th January. Once they comply with the pre-requisite under the degree program regulation, students are free to choose when to sit the exam. Moreover, in each examination session, students can sit either the exams of the courses taught during the semester or those of previous semesters that they have not yet passed. For example, during the June 2023 examination session, students can not only take the exam of a course taught in

[13] In the Italian university system, for each cycle a minimum – but not a maximum – period to graduate is defined.

[14] Another novelty of the reform is that at least 90 (out of 180) credits should "normally" (the prescription is not binding) be taught by tenured teachers. However, this requirement applies only to the new degree programs that universities aim to launch. As no new degree program has been introduced by UPO during the analyzed period, this prescription does not apply to our case. Additional aspects concerning the reform were vaguer and difficult to verify. Universities were expected to implement: a broader offer of curricula (i.e. coordinated sets of courses that can help student to delve into specific aspects of the study program), especially at the master degree level; an effective and reliable definition of educational objectives of each degree program; an increased collaboration with stakeholders in the labor market and liberal professions; a "clear and coherent" allocation of the courses across first and second level degrees; an improved alignment between high school study programs and the academic world. However, no specific targets were set regarding these objectives, making it difficult to assess what has actually been done and how these aspects impacted students' careers.

[15] https://education.ec.europa.eu/education-levels/higher-education/inclusive-and-connected-higher-education/bologna-process.

[16] Ministry guidelines for the new degree classes' implementation.

[17] For example, all of the degree courses belonging to the degree class L35 (mathematics) in the same university had to be reformed in 2008/09 or 2009/10.

[18] For each degree program, we only show the first two academic years since the courses are mandatory and identical for all students. In the third year, aside from a few mandatory courses, each student can personalize their own path by adding elective courses, making it impossible to display the changes imposed by the reform.

[19] The information on the number of staff members is available by department and field at the MUR website link to the website.

[20] Until 2012, assistant professors in Italy were hired with a permanent contract, after a public competition. Since 2012, such positions have become fixed-term, but for a specific type (RTDB) they follow a tenure track system, typically advancing to the associate professor position after three years.

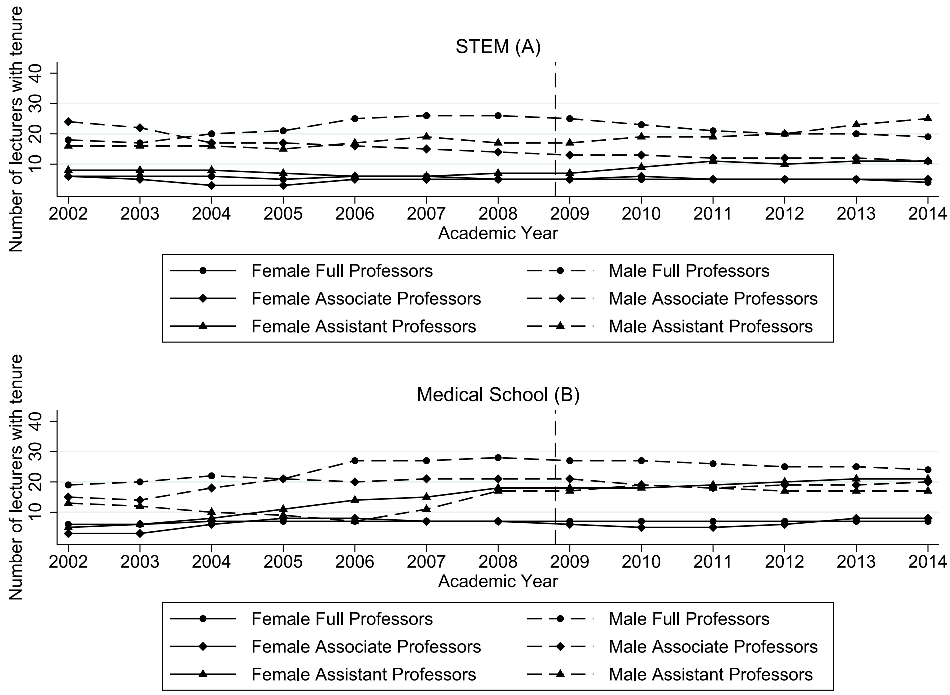

Fig. 1. Number of permanent academic staff over years.
Notes: The figure plots the number of academic staff holding permanent contracts in STEM degrees (Panel A) and in Medical School degrees (Panel B), by gender and type of contract (Full professors, Associate professors, and Assistant professors).

Table 1
Study plan of the first two academic years of materials science (i.e. chemistry) before and after the reform's implementation.

Before reform (2008/09)				After reform (2009/10)			
ID	Course name	N. courses	ECTS	ID	Course name	N. Courses	ECTS
A	General Chemistry	1	4	A+B+C	General, Inorganic Chemistry and Inorg. Lab.	1	10
B	Inorganic chemistry	1	3	D+E	Organic Chemistry + Lab.	1	10
C	Inorganic chemistry: laboratory	1	3	F+G+H	General Physics I, measure method and data analysis	1	10
D	Organic chemistry	1	7	I+L+M	General Physics II + Lab.	1	10
E	Organic chemistry: laboratory	1	3	N+O	Mathematics I and II	1	11
F	Electromagnetic and Optics	1	4	P	English language	1	5
G	Electrostatic and Electrodynamic	1	3	ZZ	Computer laboratory	1	5
H	Measure methods and data analysis	1	3				
I	Fluids and Thermodynamic	1	3				
L	Mechanics and Wave	1	4				
M	Physics: laboratory	1	4				
N	Mathematics I	1	6				
O	Mathematics II	1	5				
P	English language	1	5				
YY	Thermodynamic chemistry	1	3				
	Total 1st year	*15*	*60*		*Total 1st year*	*7*	*61*
	Average 1st year		*4*		*Average 1st year*		*8.4*
ID	Course name	N. courses	ECTS	ID	Course name	N. Courses	ECTS
A	Analytical chemistry of materials I	1	5	A+B+C	Analytical chemistry of materials + laboratory	1	13
B	Analytical chemistry of materials II	1	4	D+E	Quantum mechanics and maths for physics.	1	10
C	Analytical chemistry of materials: laboratory	1	3	F+G	Polymer chemistry + laboratory	1	9
D	Quantum mechanics	1	5	H	Calculation laboratory	1	5
E	Maths for physics	1	4	I+L	Matter physics and materials physics laboratory	1	13
F	Polymer chemistry	1	6	M+N+YY	Physical chemistry + laboratory and thermodynamic chemistry	1	10
G	Polymer chemistry: laboratory	1	3				
H	Calculation laboratory	1	5				
I	Matter physics	1	6				
L	Materials physics: laboratory	1	7				
M	Physical chemistry	1	4				
N	Physical chemistry: laboratory	1	3				
ZZ	Computer Laboratory	1	5				
	Total 2nd year	*13*	*60*		*Total 2nd year*	*6*	*60*
	Average 2nd year		*4.6*		*Average 2nd year*		*10*

Notes: The table reports the comparison of the study plan of the first two academic years of a representative STEM degree at UPO before and after Mussi's reform, which has been significantly impacted by it. The links of the courses before and after the reform are reported for each academic year (see ID column). The third academic year is not detailed as students can personalize their study plan by choosing specific courses within a list.

Table 2
Study plan of the first two academic years of computer science before and after the reform's implementation.

Before reform (2008/09)				After reform (2009/10)			
ID	Course name	N. courses	ECTS	ID	Course name	N. Courses	ECTS
A	Algebra	1	4	A+B	Algebra and Geometry	1	9
B	Geometry	1	5	C	Mathematics	1	6
C	Mathematics	1	5	D+E	Computer Architecture	1	12
D	Computer Architecture I	1	5	F	Information Economics	1	5
E	Computer Architecture I I	1	5	G	Physics	1	5
F	Information Economics	1	5	H+I	Programming and laboratory I	1	9
G	Physics	1	5	L+M	Programming and laboratory II	1	9
H	Programming I	1	5	N	English language	1	5
I	Programming I: laboratory	1	5				
L	Programming II	1	5				
M	Programming II: laboratory	1	5				
N	English language	1	5				
	Total 1st year	*12*	*59*		*Total 1st year*	*8*	*61*
	Average 1st year		*4.9*		*Average 1st year*		*7.6*
ID	Course name	N. courses	ECTS	ID	Course name	N. Courses	ECTS
A	Algorithms and data structure I	1	10	A	Algorithms and data structure I	1	9
B	Mathematics: logic	1	3	B+C	Logic, algorithms and data structure II	1	9
C	Algorithms and data structure II	1	5	D+E	Databases and Information Systems	1	9
D	Databases and Information Systems I: foundations	1	5	F+G	Fundamental of statistics	1	9
E	Databases and Information Systems I: experiments	1	5	H	Computer systems I	1	6
F	Probability	1	4	I	Operating systems I	1	10
G	Statistics	1	4	L+M	Operating systems and laboratory II		10
H	Computer systems I	1	5				
I	Operating systems I	1	10				
L	Operating systems II	1	5				
M	Operating systems II: laboratory	1	5				
	Total 2nd year	*11*	*61*		*Total 2nd year*	*7*	*62*
	Average 2nd year		*5.5*		*Average 2nd year*		*8.9*

Notes: The table reports the comparison of the study plan of the first two academic years of a representative STEM degree at UPO before and after Mussi's reform, which has been less impacted by it. The links of the courses before and after the reform are reported for each academic year (see ID column). The third academic year is not detailed as students can personalize their study plan by choosing specific courses within a list.

the second semester of 2022/23 but also a course taught in the first semester of the same academic year, or the previous academic years; if they failed to pass those exams. Finally, in September, immediately before the new academic year's commencement, they still have an additional exam session. Overall, students have seven different dates on which to sit each exam throughout the academic year.[21] In this way, students are granted considerable flexibility in their study planning and organization.

Further, with reference to materials science at the UPO (see Table 1), in which the number of courses offered in the first year was reduced from 15 (i.e., 7.5 per semester) to 7 (i.e. 3.5 per semester), considering the aforementioned organization of the exam sessions, with seven courses per semester and three fourteen-day slots in which students can sit exams at the end of the semester, students have to sit several exams in all slots (e.g., two exams in the first slot, two exams in the second slot, and three exams in the last slot). With three courses per semester, students must sit only one exam in each fourteen-day exam slot. In the former case, students have to switch from one course to another in a few days and sit multiple exams. In the latter case, they can stay focused on the same course for more days, which is more demanding in terms of the topics covered. Procrastinators – namely, those who study in the run-up to the exam – can probably manage small courses, but they can hardly handle courses with several credits. Additionally, less capable students may find it difficult to switch from small to large courses, as there is more content to learn.

Given the potential difficulties arising from the unified exams, doubts can emerge that only a formal restyling of the study plans was made. This might be the case if professors of single modules (the old shorter courses) could administer a test at the end of their module and the final grade is given by the (weighted) average of the different scores. Although we cannot rule out such behavior, the regulation of a post-reform degree courses clearly reported "In the case of integrated teaching (consisting of several modules) a single coordinated exam will be held among the teachers of the integrated course" (art. 37 of the Biological Sciences degree course regulation 2009/2010). Consequently, any exam administered outside of this regulation would have been formally invalid.[22]

A further issue is that the reform's timing coincided with the economic recession starting in 2008, which can relatively affect student composition at enrollment. According to statistics provided by the Ministry of Education (MIUR), the number of first-year undergraduates in Italy decreased by approximately 12.5% from the 2003/04 academic year (338,036) to the 2009/10 academic year (293,149). Nevertheless, the latter reduction in college enrollment occurred rather smoothly compared with the sharp effect of the economic recession.[23] The decline in university enrollment was due to a diminution in the transition rate

[21] University guidelines state that students can sit the exam a maximum of three times during the academic year. Given that it is not possible to automatically count the number of times students that sat an exam during the period under analysis, we do not know how many professors complied with this guideline.

[22] In Section 6.1, we discuss the effects of the potential non-compliance with the regulation's guidelines on our estimates. Another aspect of coordination issues is addressed by Villalobos et al. (2021) where they consider the assignment of different professors to teach the same topics, which can impact the students' learning process. However, in our case, each professor is responsible of a module of the unified course, which is not replicated by other professors in the same degree, eliminating the need for coordination in aspects such as syllabus, textbook selection, and the agreement on examination methods.

[23] The enrollment numbers are 338,036 in 2003/04, 331,893 in 2004/05, 323,930 in 2005/06, 308,185 in 2006/07, 307,586 in 2007/08, 294,932 in 2008/09, 293,149 in 2009/10, 288,876 in 2010/11, 280,119 in 2011/12, and 269,888 in 2012/13. These data are publicly available at MIUR's website: link to the website.

from school to university, indicating a pro-cyclical pattern of university enrollments in Italy.[24]

To address the concerns regarding potential changes in student composition, we present regression results obtained from a sample created through an exact matching procedure on students' pre-enrollment characteristics. Furthermore, we include tuition fees in regressions, as a proxy of family income, to consider potential consequences of the economic crisis. Additionally, our data enable us to estimate the reform's effects up to six years after the reform (at least for the first-year outcomes) to understand whether the effect is transitory.

4. Data and descriptive evidence

We use administrative data from Università del Piemonte Orientale (UPO), a public university based in the Piedmont region of Italy with three campuses in the provinces of Alessandria, Novara, and Vercelli. The university offers students a wide range of courses in three-year bachelor's programs, two-year master's programs, and five- or six-year single-cycle programs. During the academic years with which we work, the university's teaching activities were organized into seven faculties: science (biotechnology, biology, mathematics, environmental sciences, materials science, and computer science), medicine, pharmacy, law and political science, business and economics, literature, and philosophy. These are aggregated into four areas: medical and pharmacy faculties into medical school; science faculty into science, which we refer to as STEM in this paper; law and political science, business and economics into social sciences; and literature and philosophy into humanities.[25] As mentioned earlier, the reform required changes in the curricula and number of courses at the degree class level. However, at the UPO, the reform occurred at a more aggregate level, namely, the scientific area level. Degree programs within the same scientific area (STEM, social sciences, humanities, and medical school) adjusted their curricula simultaneously.

All degrees offered by the medical school (three-year degree programs as well as the six-year medicine degree) require passing a mandatory admission test before enrollment. Regarding STEM degrees, only biotechnology programs required students to take admission tests. By contrast, at UPO all of the social sciences and humanities degrees are open to any student holding an upper secondary school diploma. UPO administrative officials confirmed that no changes to the admission policy occurred during the academic years analyzed in this study. As we utilize a difference-in-differences framework, this confirmation verifies that our identification strategy has not been contaminated.

We restrict our sample to bachelor's degree (three-year) students whose degrees require 180 credits for graduation to have homogeneous and comparable treatment and control groups. Students who enroll in master's programs and five- or six-year programs (pharmacy, law, medicine, and surgery) are likely to be more skilled and motivated than bachelor's degree students. Moreover, our working sample includes only first-time enrollees at UPO (i.e., not coming from other degree programs at UPO). Therefore, if a student drops out from a degree program and enrolls in another, the information regarding the second (or further) enrollment is not covered in our final estimation sample. However, descriptive aggregate statistics of student transfers from UPO to other universities, provided by UPO administrative staff, show that this phenomenon, regardless of the field of study, is negligible, especially during the first academic year. Data from the first academic year, in fact, indicate values below 1% for STEM, social science and humanities, and medical degree students. This percentage slightly increases in the second academic year for STEM students with an average of 3.67% in the years 2008/09-2015/16. In conclusion, this evidence reassures that our estimates, especially regarding first-year dropouts, genuinely represent exits from the university system and are not influenced by the phenomenon of transfers.

The timing of the application of the reform to the degree programs varies across the scientific areas at the UPO, which enables us to adopt an identification strategy that estimates the effects of the changes in the curricula. The degree programs in STEM introduced the reform in the 2009/10 academic year, while degree programs in social sciences and humanities introduced it in 2008/09.[26] For convenience, we split the working sample into two parts: the first sample comprises students in STEM programs and medical school, while the second sample contains information regarding students in social sciences, humanities, and medical school programs. The former sample contains information regarding about 10,608 students and the latter contains information of about 15,867 students, with both samples covering the 2002/03 to 2014/15 academic years. However, for the sake of brevity, we solely present the results for STEM degrees in the main body of this article, while those related to social sciences and humanities can be found in Appendix D. This is the case because these latter degree programs had study plans with a number of courses not significantly distant from the cap imposed by the reform, thus limiting its intensity, as evidenced by the lower impact shown by the estimates on students' outcomes.

Finally, the medical school degree programs implemented the reform in the 2011/12 academic year. However, these adjustments were simply formalities and predominantly entailed a change in exam identifiers, whereby exams in Italy must have universal identifiers required by the Ministry of Education. In the subsequent section, we demonstrate that the number of exams in medical school degrees was already at 20, as required by the reform. Therefore, in practice students in medical school were not affected by the reform in our setup, even after the 2011/12 academic year.

Our data contain information at the student level. Specifically, we have information on the exact date of enrollment, degree enrolled, exact date of exit from the degree with the reason for exit (dropout or graduation), date of birth, gender, final high-school mark, type of high-school diploma, and province of the high school from which they graduated. Additionally, we obtain student-level data on exams. For each student in a given academic year, we possess information concerning which courses were passed along with the corresponding grades earned. However, noteworthily, in the exam data information regarding courses only appears if the student passes the corresponding exam. Students must earn a grade of 18 (out of 30) to pass the exam.

4.1. *Descriptive evidence*

In this section, we present descriptive evidence from a pre-matched sample on the number of courses, credit per exam, and the main academic outcomes of interest.

4.1.1. *Effects on number of exams and workload per exam*

In Fig. 2, we highlight the average number of exams (Panel A) and average credits per exam (Panel B) by scientific area over the years. These numbers are calculated by solely considering the courses taken by students who completed their degree within 43 months, assuming that these students are more likely to have passed their courses "on time".[27]

[24] The high school graduates were 452,726 in 2003/2004, 450,150 in 2009/2010 and 450,169 in 2012/2013 (Serie storiche Istat, Licenziati e diplomati delle scuole di secondo grado). The transition rate from high school to university then decreased from 74.7% in 2003/2004, to 65.1% in 2009/2010, to 60% in 2012/2013.

[25] We provide further information regarding UPO in Appendix E.

[26] These academic years reflect the enrollment years of students. Throughout the paper, when we specify an academic year, it invariably refers to the corresponding cohort's enrollment year.

[27] Bachelor degree graduates in Italy are defined as "on time" if they graduate within the month of April of their 4th year of enrollment, i.e. 43 months after their first enrollment.

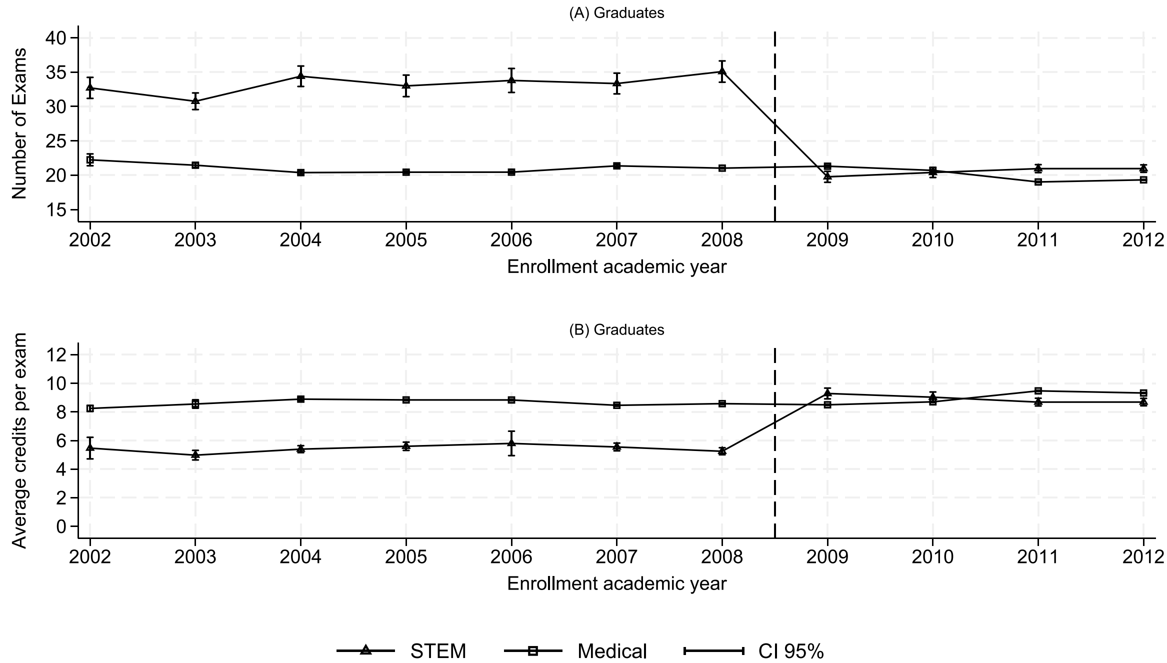

Fig. 2. The effects of reform on workload allocation.
Notes: The figure plots the average number of courses (Panel A), and the average credits per course (Panel B), taken by students who completed their degrees within 43 months, across scientific fields, over the academic years (at enrollment). The sample comprises 2884 graduates (545 in STEM). Vertical lines represent the reform's introduction for programs in STEM subjects (in 2009/10). Confidence intervals are at the 95% level.

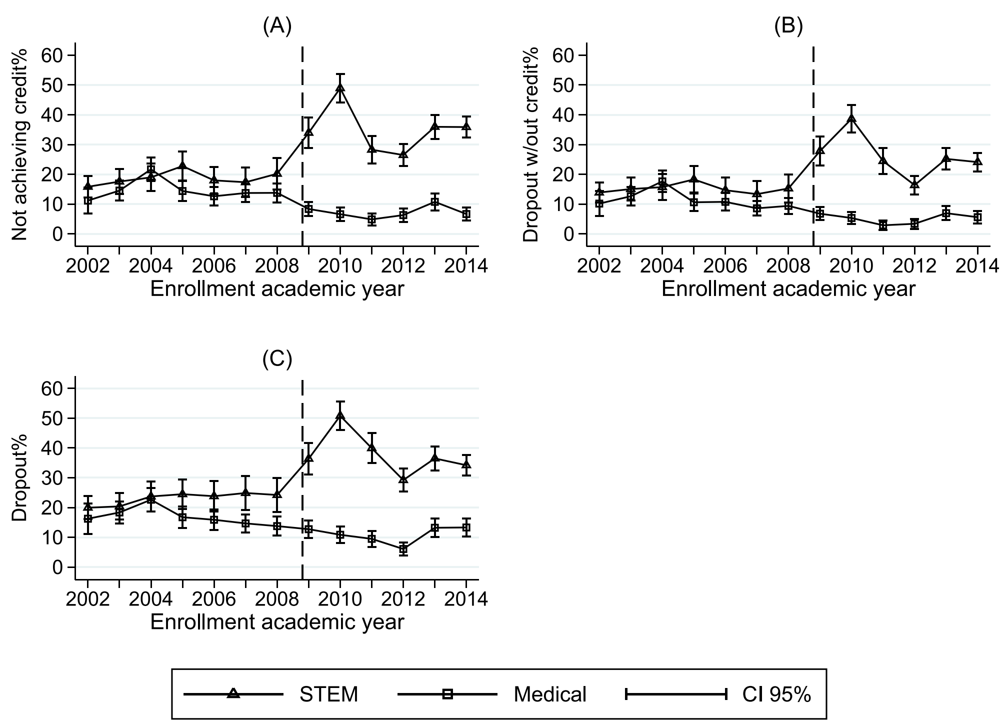

Fig. 3. First-year outcomes' descriptive statistics.
Notes: The figure highlights the averages of first-year outcomes of students pursuing STEM degrees (triangle) and in medical school (square). Panel (A) presents the average probability of not achieving credit at the end of the first-year; Panel (B) presents the average dropout rate without achieving any credit at the end of the first-year; Panel (C) presents the average first-year dropout rates. The number of observations is 10,608. Confidence intervals are at the 95% level.

As presented in Fig. 2 (Panel A), a sharp decline in the number of STEM degree exams occurs at the beginning of the 2009/10 academic year. For example, before the reform, a student with a STEM degree would take an average of 33 courses, whereas the corresponding number is 20 courses after the reform.

Regarding medical school degrees, we do not see any change in the number of courses, which is reassuring for our identification strategy, given that these students define our counterfactual group. Essentially, the number of courses in these programs was already at the required level.

To quantify how changes in the number of courses translate into an increase in workload per course, we plot the average credits per course in Panel (B). The reform did not change the total credits that students had obtain to complete their degrees (i.e., 180 credits), and hence the changes in the number of courses mechanically increased the average number of credits. We note that a student in a STEM degree has to study approximately eight credits per course throughout their degrees after the reform, while six credits were required before the reform.

4.1.2. Academic outcomes

We focus on six main academic outcomes, three of which are first-year outcomes: the probability of not achieving credits in the first academic year, the probability of dropping out without achieving credits at the end of the first year, and the first-year dropout rate (with or without credits).[28]

Our data are compiled using different sources provided by the UPO administrative office. In the data providing information regarding exams, grades, and credits, we can only observe students if they pass at least one exam. Therefore, we create a dummy variable that equals 1 if a student does not pass any exam during freshman enrollment year, otherwise it is equal to 0. If this variable is equal to 1, it means that the student did not pass any exams in their first year—that is, they did not achieve any credit.[29] Regarding the second variable, we create a dummy that equals 1 if a student drops out at the end of their first year without achieving any credits. The third variable occurs, which is the main outcome of interest in this study, is a dummy that equals 1 if a student drops out at the end of their first academic year, and 0 otherwise.

The other three outcomes are related to the overall degree, namely graduation probability, time to graduation (in months), and degree grade.[30] First-year dropout and graduation rates represent the most policy-relevant outcomes in our study. Nevertheless, we also present results for the other outcomes as they help to better understand the overall effects of the reform.

In Fig. 3 (first-year outcomes) and Fig. 4 (degree outcomes), we present these variables' mean statistics, along with the confidence intervals at the 95% level for our pre-matched sample. The Fig. 3 illustrates that the first-year outcomes of students in both STEM degrees and medical school display similar trends before the reform, with the exception of Panel C where the first-year dropout rate of medical students slightly decreases starting from the 2004/05 academic year, while it remains stable for students in STEM degrees until the reform. However, in Section 6.1, we will present results from an event-study specification to check the validity of parallel trend assumption. During the pre-treatment periods, STEM students' average dropout rate is approximately 23%, while the average dropout rate for students at a medical school is around 17%. We observe a substantial increase with the introduction of the new curricula, whereby the first-year dropout rate raises to 36% for the cohort of the 2009/10 academic year in STEM degrees and reaches nearly 51% in the 2010/11 academic year. Although these figures are unconditional averages and we control for students' key observable characteristics in our regression analysis, such a substantial surge in dropouts during the treatment period already reveals students' struggles under the new curricula. After 2010/11, we observe a decreasing pattern. However, the first-year dropout rate does not fall below the 30% range, which is the initial level immediately after the reform. This long-term effect indicates that the administrative transition period cannot explain the observed effect. Panels (A) and (B) exhibit similar levels and patterns of not achieving credits and dropping out without achieving credits. It also suggests that changes in the first-year dropout rate are likely to be associated with the exam-related difficulties that students experienced after the reform.

Regarding degree outcomes, we employ a sample of students who enrolled in their degrees from the 2002/03 academic year to 2011/12. Further, we place a cap on the duration of graduation, namely 54 months from enrollment. The latter duration for graduation is the longest observed in our data for the 2011/12 cohort. This restriction ensures that the effects on graduation are not driven by a mechanical relationship between the enrollment years and the year in which our data are constructed (September 2016).[31]

A similar but reverse pattern is observed for the graduation rates (Panel A in Fig. 3). We note that graduation rates seem to have declined from 40% to 20% after the reform's introduction in STEM degrees. By contrast, for the other two figures in Fig. 4 concerning the time to graduation and graduation marks, we do not observe any clear patterns or fractions after the reform, at least descriptively.

5. Identification and empirical framework

We employ a difference-in-differences approach to identify the effects of exogenous changes in the number of courses on the outcomes of interest. As mentioned earlier, the degree courses in STEM applied the required changes in the 2009/10 academic year. However, bachelor's degrees in medical school are not affected by the reform, considering that the number of courses was already at the required level (as elucidated in Fig. 2). Therefore, we use students enrolled in medical school programs as a counterfactual group, which allows us to identify the effects on STEM degrees from the 2009/10 to 2014/15 academic years. We aim to compare the differences in the outcomes of interest between STEM and medical school degrees before and after the reform's implementation.

Concerns may arise regarding the suitability of using students enrolled in medical degree programs as a counterfactual. However, while students attending six-year degrees offered by the medical school (medicine and surgery) are certainly highly selected due to both the high-stake national test and the long duration of the study program, our counterfactual sample appears to be less selected. First, it comprises students enrolled in first-level (three-year) programs in health professions that award degrees in, e.g., nursing, obstetrics, physiotherapy, and dental hygiene. These professions are crucial for the health systems and patients' care, but they are not as socially and economically prestigious as general practitioners or specialized doctors. According to *AlmaLaurea* survey on Italian graduates in 2012,[32] the social origins of the students enrolled in first-level medical degrees in UPO do not appear better than their peers enrolled in STEM courses, while they are poorer than the graduates in six-year degree programs.[33] Second, we observe that

[28] In Appendix B, we also considered the average grades of first-year students and presented the results. However, as we discussed earlier, this variable is left-censored and requires Tobit regressions (see De Paola and Scoppa 2011), which provides relatively less credible results in our set up as the reform can affect the probability of passing exams.

[29] However, we cannot possibly know whether these students actually sat an exam during their first year.

[30] Evidently, we can only observe the final grade and time to graduation for the students who complete their degrees, which can raise concerns regarding selection bias. This issue has been addressed by performing exact matching among graduates.

[31] We yield the same findings without this cap, and they are perfectly consistent with the results reported in this section.

[32] The *AlmaLaurea* survey on Italian graduates "Profilo dei laureati" outlines the characteristics and performance of graduates who obtained their degrees at *AlmaLaurea* member universities.

[33] The share of students whose parents are not college graduates is 86.3% for graduates in first-level degrees of the medical school, 78.3% for STEM graduates, and 69.8% for graduates in the six-year degree program in medicine and surgery. Further, the share of graduates with a working-class (blue collar) social background is 38% for the graduates in first-level degrees of the medical school, 33% for the graduates in STEM, 16.3% for the graduates in the six-year degree programs in medicine and surgery. These data are publicly available on AlmaLaurea's website: link to the website.

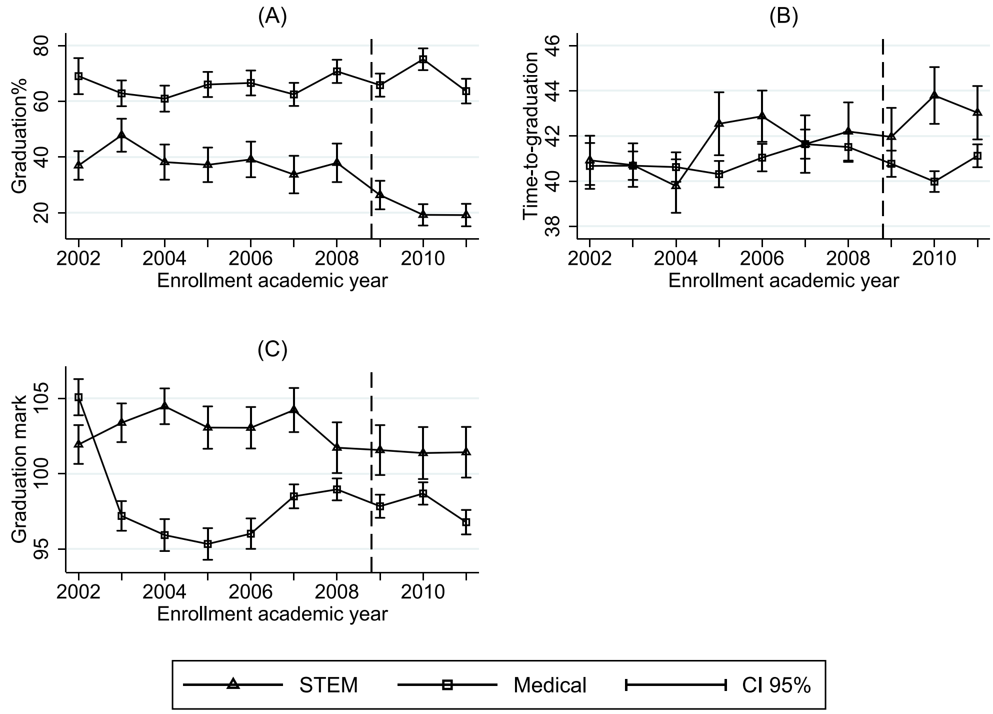

Fig. 4. Degree outcomes' descriptive statistics.
Notes: The figure highlights the averages of degree outcomes of students pursuing STEM degrees (triangle) and in medical school (square). Panel (A) presents the average graduation rate, N: 6770. Panel (B) presents the average duration (in months) of degree completion of graduates, N: 3452. Panel (C) presents average final marks of graduates, N: 3452. Confidence intervals are at the 95% level.

they do not appear better selected than their peers in STEM degrees, at least for the school performance before university enrolment, as their average final high-school mark is no higher than their peers enrolled in STEM courses.[34] Assuming that the final high-school mark is a reliable proxy for students' ability, this comparison suggests that the admission tests for bachelor's degrees offered by the medical school do not select "top" students but rather exclude those who do not have the minimum level to succeed in their studies. Therefore, students who passed the test for such degrees have similar schooling abilities to their peers admitted (without the test) in STEM degrees. STEM degree programs probably have less need to set up admission tests to make the same kind of selection (of ill-prepared students) because the well-known difficulty of these courses can already induce a self-selection of the students.

5.1. Baseline econometric specification

The econometric specifications used are as follows:

$$Y_i = a + \gamma STEM_i + \lambda Post_t + \delta(STEM_i \times Post_t) + X_i\beta + \epsilon_i; \quad (1)$$

where Y_i is the outcome of interest, $STEM_i$ is a dummy variable equal to 1 if i is a student of a STEM degree,[35] $Post_t$ is a dummy equal to 1 if the enrollment academic year t is during and after the 2009/10 enrollment academic year, and X_i is a vector of pre-enrollment observable characteristics of student i—specifically, the final high-school mark, the type and region of the high-school from which the student graduated, age, gender, and tuition fee at enrollment. The parameter of interest that we aim to estimate is δ, which provides us with the difference-in-differences estimate.

Standard errors are clustered at the degree and academic-year levels (Carrieri et al., 2015). Ideally, standard errors in a difference-in-differences framework are clustered at the "treatment" level. In the case of an insufficient number of clusters, evidence reveals that wild bootstrapping yields convenient results (Cameron et al., 2008). However, MacKinnon and Webb (2017) provide evidence that if the number of clusters in the regressions is less than twelve, wild bootstrapping severely underrejects the null hypothesis of being equal to zero. In our case, the treatment occurs at the degree-program level, and we have ten degree courses in our estimation sample. Nevertheless, we also reproduced our main results by clustering at the degree course level and performing a wild bootstrap procedure. These results are consistent with the following description, and are available in Appendix A.[36]

5.2. Event-study specification

In addition to our baseline model, Eq. (1), we established an event-study specification to ascertain whether the difference-in-differences approach's common trend assumption is satisfied in our estimation samples, and whether the reform has long-term implications. For this purpose, we simply interact the year dummies with the treated STEM degree dummies. We choose the year before the intervention as the baseline, namely the 2008/09 academic year.

[34] We present average statistics on students' observable characteristics in Appendix F, Table F1. These numbers are computed before performing exact matching.

[35] We also present results obtained from a model with degree course fixed-effects in Appendix C. As we will explain in this section, our matching procedure takes place at STEM degree level, which can create relative instability across degree courses over the years.

[36] Notably, in an academic environment's context, the curricula and assignment of lecturers is updated at the beginning of each academic year. Therefore, two-way clustering at the degree-program and academic-year levels is significantly less restrictive than clustering at, for example, state and calendar-year levels (e.g., to evaluate a labor market reform occuring at the state level).

We establish the following model as an event study specification:

$$Y_i = \alpha + \gamma STEM_i + \sum_{\substack{k=2002/03 \\ k\neq 2008/09}}^{2014/15} \lambda_k P_k + \sum_{\substack{k=2002/03 \\ k\neq 2008/09}}^{2014/15} \delta_k (STEM_i \times P_k) + X_i \beta + \epsilon_i; \quad (2)$$

where P_k are the dummy variables, which equal 1 in year k, $STEM_i$ equals 1 if student i is in the STEM degree, X_i is the vector of pre-enrollment observable characteristics of student i as specified in Eq. (1). Interaction terms' coefficient estimates, δ_k, are the parameters of interest in the model. As in our main specification, we cluster the standard errors at the degree-program and academic-year levels.

5.3. Matching procedure

The difference-in-differences framework requires a stable composition across treatment and control groups when working with repeated cross-sectional data. In this section, we create estimation samples by performing exact matching on the observable characteristics. This allows us to run our difference-in-differences analysis with a balanced student composition between the treatment and control group.

We perform a two-stage matching process. First, in a given academic year (for years both before and after the reform's implementation), we exact match the treated students with the counterfactual group according to gender, age, type of high-school diploma, and the region of high school from which a student graduated. For these matched students, we calculate the differences in their high-school graduation marks, whereby 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 two units.[37]

In the final stage, we randomly select four students from a 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 no proper match exists for the treated students, in which case the treated students are excluded from the sample. Furthermore, in some cases, one control student is a match for multiple treated students. However, this is not problematic in our setup because we work with repeated cross-sectional data whereby for each outcome analyzed we observe each student only once. Further, we allow the observations of control students to be repeated based on how many times they are matched with a treated student (i.e., matching with replacement). This is broadly equivalent to weighted propensity score matching, and also increases the matching procedure's precision.[38]

In the end, our first matched sample comprises 2629 observations from STEM degrees and 5416 from medical school. Table 3 presents the mean statistics in matched samples used in this section. We note that the standardized differences are considerably small both before and after the implementation of the reform. Imbens and Rubin (2015) suggest that the corresponding figure should be smaller than .20 for the variables related to the compositions of the treatment and control groups.

Moreover, we present the trends in student composition by age, gender, high-school diploma, final high-school marks, tuition fees, and the share of students who went to high school in Piemonte region in Fig. 5 for the matched sample. We observe that our matching procedure successfully balanced the composition of the students at enrollment.

We adopt a slightly different approach for graduation outcomes, namely time to graduation and graduation mark. While the matching procedure remains the same, we change its timing. Instead of matching students at enrollment, we match them conditional on graduation considering that we observe significantly large effects on first-year outcomes, which can also alter the composition of those who remained enrolled under the new curricula.[39]

6. Results

First, we present the estimation results in Fig. 6 on the number of courses that students take and their average credits.[40] These results are obtained from Eq. (2), and provide greater formal information pertaining to the descriptive evidence discussed in Section 4.1 regarding how the reform affected the number of exams that the graduates sat during their degree courses.[41] Panel (A) of Fig. 6 reveals that after the reform's introduction in the 2009/10 academic year, STEM degree students had a reduction of thirteen courses compared to medical school students. Moreover, after the reform, students with STEM degrees exhibited an increase in the average number of credits per exam compared to medical school students.

6.1. Effect on academic outcomes

We present the results obtained from Eq. (1) for the average effects of the workload increase per exam on STEM degree students' academic outcomes in Table 4.

Columns (1)–(3) report the findings of students' first-year outcomes. We find a significant increase of 20 p.p. in first-year students' dropout probability. Additionally, we observe that the probability of not achieving credit increases (i.e. failing all the exams) by 22 p.p. Similar but slightly smaller effects are estimated for the first-year dropout probability without achieving any credit. These findings are consistent with the effect on first-year dropouts, indicating that the latter effect is associated with the exam-related issues faced by these students after the reform's implementation.[42]

Column (4) presents the results for graduation rates. Consistent with the effects on the first-year dropout probability, we observe negative effects on the graduation rates by approximately 23.7 p.p. Recall that, in Section 4.1.2, for this outcome we place a cap on the duration of graduation, which is set to 54 months from the time of enrollment. This duration for graduation is the longest observed in our data for the 2011/12 cohort.

However, we do not observe statistically significant changes in the time needed to complete a degree course or final graduation marks (columns 5 and 6). The result on the time to gain a degree is controversial because a primary reason for placing a cap on the number of exams was helping students to complete their degrees *on time*.[43] Nevertheless, these two outcomes can only be observed among students who have successfully completed their studies. We already know that our treatment significantly affects the first-year dropout rates, and hence the

[37] A matching procedure on high-school marks can also be executed as an exact match. However, this yields an extremely small sample. Therefore, we perform matching by accepting a small gap between the final marks. Final high-school marks range between 60 and 100. Therefore, allowing a maximum of a two-unit gap for the differences in these marks is reasonable. Nevertheless, we perform comprehensive robustness tests in Appendix C and show that our results do not depend on any sample selection criteria of our matching procedure.

[38] This final adjustment in creating the matched samples insignificantly alters the results. The results obtained from matched samples without replacement are available upon request.

[39] We present our treatment's heterogeneous effects in Section 8.

[40] These results are obtained from regressions without using any control variables on students' characteristics.

[41] As stated in Section 4.1, to calculate the number of courses, we consider the students who completed their degrees within 43 months.

[42] In Appendix B, we also present results on the average grades of freshmen, whereby we estimate significant negative effect on the grade. However, as we discuss in Appendix B, this outcome is left-censored as we can only observe it for the exams that students passed, and thus we place these findings in Appendix.

[43] Section 1.3 "System goals" of the ministerial guidelines for the reform's implementation.

Table 3
Summary statistics in matched sample before and after the reform's implementation.

	(1)	(2)	(3)	(4)	(5)	(6)
	Before reform (2002/03–2008/09)			After reform (2009/10–2014/15)		
	STEM	Medical	Std. dif.	STEM	Medical	Std. dif.
Variables						
Academic track%	0.34	0.36	−0.04	0.59	0.61	−0.06
High-school marks	78.14	77.82	0.02	74.54	74.71	−0.02
Female%	0.56	0.62	−0.14	0.61	0.65	−0.07
Age	19.31	19.20	0.11	19.38	19.26	0.09
Tuition fee (euro)	989.23	1011.85	−0.05	1064.83	1098.94	−0.05
Dropout%	0.22	0.18	0.10	0.37	0.13	0.60
Num. of observations	1038	2073		1591	3343	

Notes: The table reports the summary statistics of variables in the matched sample (STEM vs. Medical). Columns (1) and (2) report the mean values of variables for degrees in STEM and Medical school before the implementation of the reform, respectively. Columns (4) and (5) report the values of variables after the reform. Columns (3) and (6) present the standardized mean differences calculated by dividing the mean difference between the outcomes of treated and control units by the variable's standard deviation.

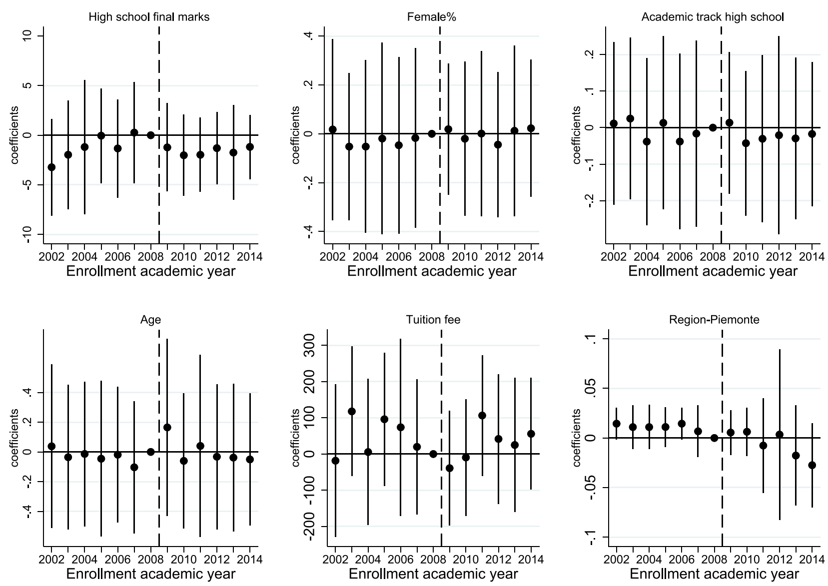

Fig. 5. Event study specification for students' composition: STEM vs. Medical.
Notes: The figure highlights the coefficient estimates of δ_k specified in Eq. (2) for the matched sample of degrees in STEM and medical school, N: 8045. Standard errors are clustered at the degree course and academic year level. Confidence intervals are at the 95% level.

coefficient estimates on these two outcomes conditional on graduation might be relatively biased owing to changes in the compositions of graduates after the reform. Furthermore, as we discuss in the rest of the paper, the effect on first-year dropout rate might change the composition of students who continue their degrees (e.g. high-skilled students survive the first year) or dealing with intensified courses can improve certain skills of students. These two factors can offset the difficulties of unified exams, hence we do not observe effect on graduation marks and time to graduation.

Regarding the control variables, apart from the statistically insignificant coefficient estimate for gender, all other coefficient estimates are consistent with initial expectations.[44] For example, the higher the students' high-school graduation marks, the lower their likelihood to drop out at the end of their first year. Moreover, students who graduated from high schools with a scientific curriculum are less likely to drop out than those who completed high-school with other curricula.

We present the results from our event-study specification outlined in Eq. (2) for the first-year dropout and graduation rates in Figs. 7 and 8, respectively. These figures depict that the parallel trend assumption holds in our data, which is crucial for our identification strategy.

In Fig. 7, we observe that the largest effects occur during the early post-reform years, especially for students enrolled in the 2010/11 academic year. Nevertheless, the effect for students in STEM degrees remains statistically significant even six years after the reform's implementation. This finding suggests that the reform's effect cannot be accounted for merely by an undesirable administrative transition period or lecturers "mishandling" the difficulties of bundled exams. While lecturers could make some adjustments (e.g., preparing relatively easier exams) based on the significant negative impacts that they experienced during the early years of the new curricula, they would only attenuate the estimated effects that we present rather than drive them (i.e., attenuation bias).

A potential concern regarding the accurate interpretation of our results may arise from the fact that a few professors may have administered unofficial tests at the end of their modules, contravening the degree program regulations. Provided that we cannot verify the actual behavior of all teachers, it is reassuring that our findings are conservative. De Paola and Scoppa (2011) demonstrated that administering intermediate tests (instead of unified exams) increases the likelihood of passing the exams. We can then argue that if some professors were not compliant with the law prescription, our findings would underestimate the effect of the reform.

[44] The literature reveals that female students are more likely to drop out of STEM subjects. See, among others, Griffith (2010), Chen (2013), and Isphording et al. (2019).

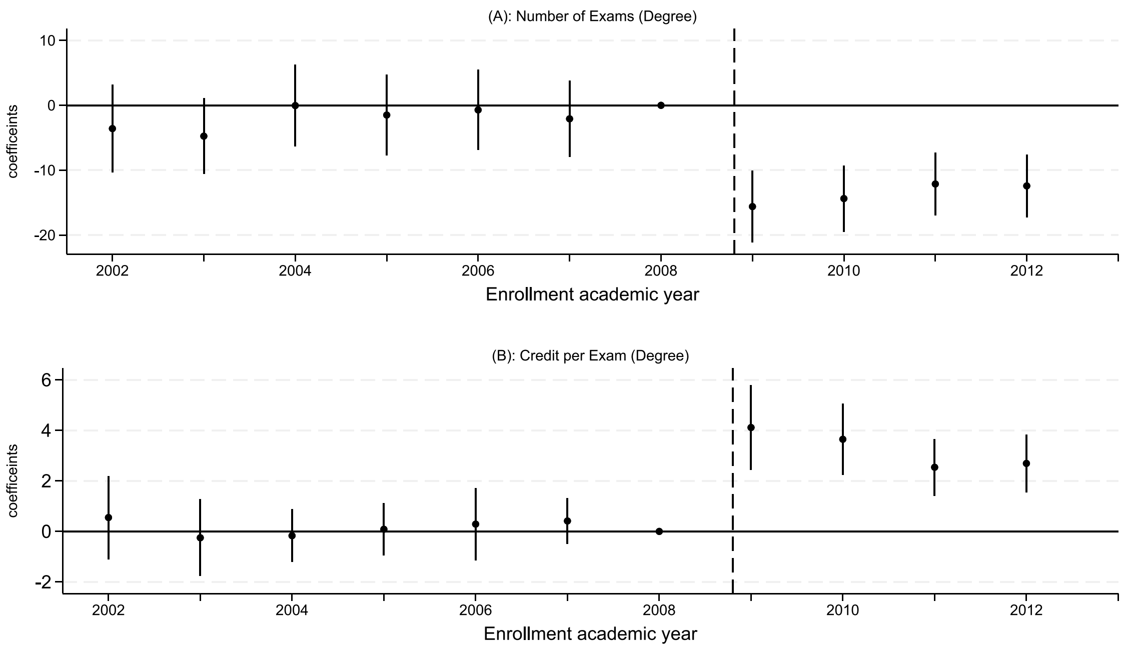

Fig. 6. Event study specification: Effect on workload.
Notes: The figure highlights the coefficient estimates of δ_k specified in Eq. (2). Panel (A) plots the average number of exams, and Panel (B) plots the average credits per exam taken by students who completed their degrees within 43 months over the academic years (at enrollment). The sample of STEM and Medical degrees comprises 2884 graduates (545 in STEM). Vertical lines represent the introduction of the reform for programs in STEM subjects (in 2009/10). Confidence intervals are at the 95% level.

Table 4
Results on academic outcomes estimated from matched sample.

	(1) No credit	(2) Drop w/out credit	(3) Dropout	(4) Graduate	(5) Grad. Mark	(6) Time-to-grad.
STEM X Post	0.220***	0.161***	0.202***	−0.237***	1.897	0.393
	(0.046)	(0.037)	(0.044)	(0.057)	(1.815)	(0.974)
Post	−0.088***	−0.064***	−0.051*	0.077*	−1.591	−0.055
	(0.027)	(0.020)	(0.029)	(0.039)	(1.465)	(0.490)
STEM	0.013	0.022	0.039	−0.209***	1.671	0.923
	(0.029)	(0.023)	(0.027)	(0.044)	(1.030)	(0.569)
HS. Marks	−0.004***	−0.003***	−0.004***	0.011***	0.334***	−0.133***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.027)	(0.020)
High school type. Omitted category: Scientific						
Professional	0.142***	0.112***	0.136***	−0.276***	−5.084***	2.922**
	(0.040)	(0.032)	(0.041)	(0.054)	(1.391)	(1.457)
Technical	0.037**	0.043***	0.042**	−0.131***	−4.228***	1.307**
	(0.015)	(0.014)	(0.018)	(0.026)	(0.578)	(0.544)
Classical	0.072**	0.081***	0.081**	−0.152**	−0.173	0.889
	(0.031)	(0.028)	(0.040)	(0.060)	(1.170)	(1.694)
	(0.023)	(0.020)	(0.032)	(0.038)	(0.878)	(0.674)
Other HS.	0.044*	0.041**	0.040	−0.132***	−3.790***	1.620**
	(0.024)	(0.019)	(0.027)	(0.041)	(0.741)	(0.706)
Female	−0.026*	−0.018	−0.010	0.058**	−1.143	−0.353
	(0.014)	(0.012)	(0.016)	(0.023)	(0.838)	(0.635)
Tuition fee	.0000239	.000017	.0000218	7.42e−06	.0008919*	.0003544
	(.0000174)	(.0000154)	(.0000188)	(.000024)	(.0004549)	(.0002859)
Region dummies	yes	yes	yes	yes	yes	yes
Age dummies	yes	yes	yes	yes	yes	yes
Mean dep. var.	0.15	0.12	0.20	0.59	101.78	40.94
N	8045	8045	8045	4748	1318	1318

Notes: The table reports the estimation results from Eq. (1) for the matched sample of STEM and medical school degrees. The samples in (1)–(3) cover the academic years at enrollment from 2002/03 to 2014/15. The samples in (4)–(6) cover the academic years of enrollment from 2002/03 to 2011/12. (1) is the probability of not achieving credit by the end of the first year and (2) is the probability of dropping out without achieving credit at the end of the first year. (3) represents the probability of dropping out in the first year. (4) is the probability of graduation. (5) Graduates' final graduation marks. (6) is the duration (in months) of graduates' degree completion. HS. Marks represent the final score on the high school exam. *Post* is a dummy that equals 1 if the academic year of enrollment is after reform (i.e. from 2009/10 to 2014/15). Regressions include dummies for the region of the high school that students graduated from, and for the age of students at enrollment. Standard errors are in parentheses and are clustered at the degree and academic year levels. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

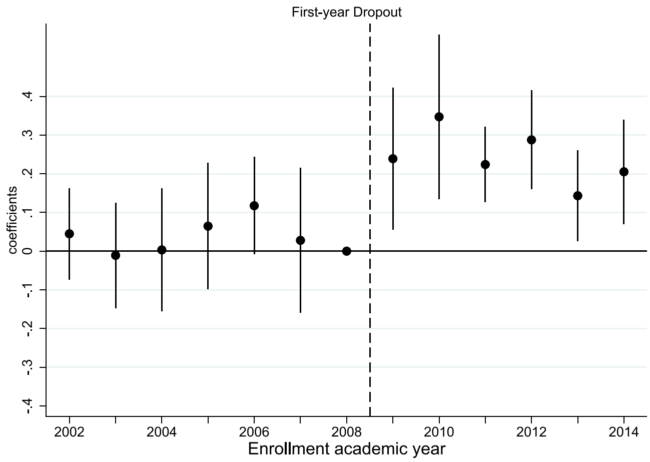

Fig. 7. Event study specification: STEM vs. Medical.
Notes: The figure highlights the coefficient estimates of δ_k specified in Eq. (2). Number of observations is 8045. The dependent variable is the dropout probability at the end of the first year. The control variables include age, gender, tuition fee, type of high school, final high school marks, and the region of high school. Standard errors are clustered at the degree course and academic year level. Confidence intervals are at the 95% level.

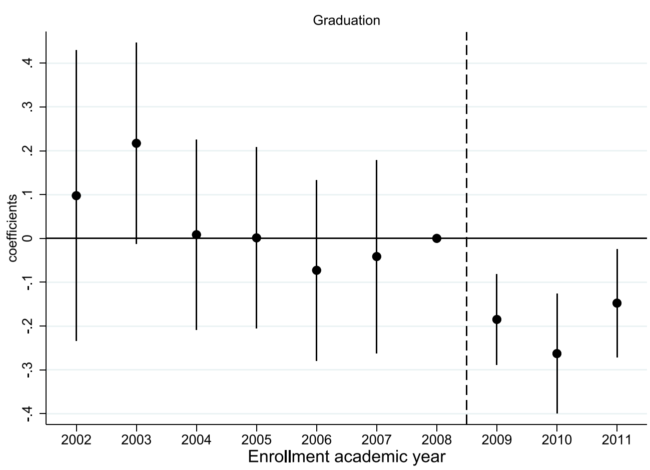

Fig. 8. Event study specification: STEM vs. Medical.
Notes: The figure reports the findings on the probability of graduating. The figure highlights the coefficient estimates of δ_k specified in Eq. (2). The number of observations is 4748. The control variables include age, gender, tuition fee, type of high school, final high school marks, and the region of high school. Standard errors are clustered at the degree course and academic year level. Confidence intervals are at the 95% level.

Another potential concern with our findings is that the simultaneous economic crisis could have impacted student outcomes, potentially clouding the interpretation of our results. Notably, the Great Recession in Italy was a double-dip recession and lasted until 2014, the period of observation of our analysis. Hereafter, we will discuss why we can reasonably conclude that the interpretation of our results remains valid even when considering the effects of the prolonged economic crisis.

First, figures showing a decrease in freshmen enrollments in Italian universities (reported in Section 3) suggest that our findings (increasing dropout) are unlikely to be attributed to a poorer selection at the entrance of universities, which would be the case if enrollment to tertiary education were counter-cyclical. According to Ghignoni (2017), during the Great Recession, the incoming freshmen cohorts in Italy were indeed more selected in terms of family social class, cultural background, type of high school attended, and other factors. Hence, the rich set of control variables on the pre-enrollment characteristics of students that our econometric models include is very important. In Section 6.2 we also include the average tuition fees paid by students during their degree.

A further possible effect of the economic crisis could be a change in the composition of the students enrolled by field of study: an increase in enrollees in STEM degree programs, due to the financial crisis and the higher unemployment rate in STEM-related occupations, could represent an alternative explanation to the increasing dropout that we document. However, enrollment data by field of study at the national level do not register this increase, as the share of freshmen in STEM degree programs remained rather stable during the years around the 2008 crisis.[45] On the other hand, a reduction in university enrollment was recorded for social sciences and humanities degrees.

Lastly, regardless of university enrollment dynamics, the crisis could directly impact the probability of student dropout. A recession can affect student outcomes in two opposite ways. On the one hand, it could increase student persistence and reduce dropout rates. This is because students might have less incentive to leave university due to the heightened risk of unemployment. On the other hand, the probability of student withdrawal could augment, as the parents of students are more at risk of facing job dismissal during a recession. This, in turn, would diminish the financial resources available to support their children's university education. According to existing literature on the Italian case (Di Pietro 2006, Adamopoulou and Tanzi 2017), the first channel appears to be more relevant: deteriorating labor market conditions contribute to a decrease in dropout rates. These findings are consistent with (Becker, 2006) who demonstrates how, in the context of relatively low university tuition fees, such as in Italy, students may consider the academic experience a kind of *parking lot* in presence of high unemployment rates. In contrast, our analysis reveals an increase in the dropout rate among STEM students after the reform, namely during the years of recessions.

6.2. Sensitivity analysis

In this section, we present an additional robustness test to verify our findings' validity. These sensitivity checks are performed using the matched sample explained in Section 5.3.

6.2.1. Intensity of treatment

We focus on the differential effect of the reform across degrees in STEM subjects. The sample of graduates, which is also used in Fig. 2, reveals that the reform affected some STEM degrees relatively less than the others. We consider the two degrees affected the least, namely computer science and mathematics, as *less affected* degrees. On the other hand, the other four degrees, environmental sciences, biology, biotechnology, and materials science are considered as *more affected* degrees.

In Table 5, we report the difference-in-differences coefficient estimates obtained from Eq. (1) for these groups. *Less affected* degrees reduce the number of exams on average by 6.5 exams and increase the

[45] The number of freshmen at the national level in first level STEM degrees offered by UPO (i.e. excluding Engineering) ranges from 32,494 in 2008/2009 to 31,884 in 2012/2013, which corresponds to a share of 14% and 15.6% respectively of the entire freshmen cohorts. Unfortunately, time series by field of study are only available from the 2008/2009 academic year, and therefore we lack comparable information for the years before. However, the economic crisis strongly impacted the Italian economy since 2009 (with a decline of the real GDP of 6.6% in one year), and therefore if university enrolments in STEM fields were counter-cyclical, we should expect an increase by the 2009/2010 academic year. Also, the statistics about freshmen by high school track and field of study show that the composition of enrollees remains stable over time. These data are publicly available at these websites; linktoISTATdata; linktotheMURdata.

Table 5
Heterogeneous results by intensity of treatment: effects on workload.

	Full sample		Less affected		More affected	
	(1) Exam	(2) Credits per exam	(3) Exam	(4) Credits per exam	(5) Exam	(6) Credits per exam
STEM X Post	−11.775***	3.223***	−6.481***	1.678***	−13.075***	3.596***
	(1.208)	(0.445)	(0.829)	(0.442)	(1.322)	(0.468)
STEM	12.121***	−3.262***	6.863***	−1.769***	13.418***	−3.630***
	(0.899)	(0.234)	(0.530)	(0.325)	(1.003)	(0.235)
Post	−0.763	0.299	−0.763	0.299	−0.763	0.299
	(0.470)	(0.202)	(0.472)	(0.203)	(0.470)	(0.202)
N	2884	2884	2426	2426	2797	2797

Notes: The table reports the effects of reform on the number of exams and credits per exam from Eq. (1) for the graduates on time in our main sample. Less affected degrees are computer science and mathematics, while more affected degrees are environmental sciences, biology, biotechnology, and materials science. Standard errors are in parentheses and are clustered at the degree and academic year levels. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6
Heterogeneous results by intensity of treatment: effects on academic outcomes.

	(1) No credit	(2) Drop w/out credit	(3) Dropout	(4) Graduate	(5) Grad. Mark	(6) Time-to-grad.
Panel A: Less affected vs. Medical						
STEM X Post	0.080	0.027	−0.014	−0.178***	−0.009	1.931
	(0.049)	(0.046)	(0.061)	(0.057)	(3.237)	(2.225)
Mean dep. var.	0.12	0.09	0.16	0.64	101.21	40.70
N	5953	5953	5953	3632	952	952
Panel B: More affected vs. Medical						
STEM X Post	0.261***	0.198***	0.254***	−0.270***	1.951	0.320
	(0.050)	(0.039)	(0.044)	(0.064)	(1.871)	(1.014)
Mean dep. var.	0.15	0.11	0.20	0.61	103.28	40.86
N	7499	7499	7499	4416	1246	1246
Panel C: More affected vs. Less affected						
More X Post	0.187***	0.172***	0.271***	−0.086	1.278	−2.016
	(0.058)	(0.054)	(0.065)	(0.069)	(2.971)	(2.491)
Mean dep. var.	0.26	0.20	0.31	0.37	103.26	41.71
N	2620	2620	2620	1443	438	438

Notes: The table reports the results from Eq. (1) on academic outcomes. Panel A shows the results on students in STEM degrees that are less affected by the reform, Panel B shows them for the students whose degrees are more affected, and Panel C shows the results for the more affected students compared to less affected students. All regressions include control variables on gender, tuition fee, high school final marks, type of high school, age, and region. Less affected degrees are computer science and mathematics, while more affected degrees are environmental sciences, biology, biotechnology, and materials science. Standard errors are in parentheses and are clustered at the degree and academic year levels. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

credit per exam by 1.7 credits, while the figures are thirteen exams and 3.6 credits per exam for the *more affected* degrees, respectively.

Table 6 presents the results on academic outcomes. In Panel A, when we compare the *less affected* degrees with degrees in medical school, we see that the first-year outcomes of students in *less affected* degrees are not significantly affected by the reform. However, the probability of graduation significantly decreases by 18 p.p. for these students. This result suggests that the increase in workload per exam in these degrees was not sufficiently large to discourage students early in their first year but rather the intensity of exams became overwhelming in the following years. On the other hand, in Panel B, we see that students in *more affected degrees* are significantly influenced in terms of both dropout and graduation rates.

In Panel C, we compare the students in *more affected degrees* (treatment group) with the ones in *less affected* degrees (control group). Despite being imperfect considering that the reform also has an impact on the graduation rates of students in *less affected*, this robustness exercise reassures that the main findings presented in the paper are not solely driven by some peculiarities of medical school as a control group.

6.2.2. Biotechnology vs. Medical school

The differences in admission procedures between the treatment and control groups might be a potential concern. As explained in Section 4, the degrees organized by the medical school (i.e., our control group) perform ex-ante admission selection tests, while the degrees in our treatment groups do not conduct such a selection procedure, with the exception of biotechnology degrees.[46] Although these admission policies of degree programs do not change over time and we control for the time-invariant heterogeneity in our empirical setup, we undertake this further robustness test by exploiting the unique situation of biotechnology degrees in our estimation sample.

At the UPO, the biotechnology degree is offered by the medical school, although based on the definition decided by the Italian Ministry of Education, these degrees are officially labeled as a science degree and subject to any change in the regulations regarding STEM degrees. Therefore, students with biotechnology degrees represent a perfect treatment group in terms of comparison with the students in our control group, namely the medical school students. We develop our working

[46] The admission test for medical school degrees is regulated at the national level, while for biotechnology at the local level by each university. The reason behind this difference is that medical school degrees give access to professions regulated at the national level, while biotechnology degree does not. Despite these elements of difference, we maintain that the two tests are still comparable. First, in both cases, students compete for a spot within the maximum capacity allocated for each course of study at that university. Second, in the analyzed University both tests are administered by the School of Medicine. Third, the competences of those who prepare the test for biotechnology and for medical degrees are the same, as the content of the tests is similar (both have core questions on mathematics, biology, physics, and chemistry).

Table 7
Results on academic outcomes: Biotechnology vs. Medical school.

	(1) No credit	(2) Drop w/out credit	(3) Dropout	(4) Graduate	(5) Grad. Mark	(6) Time-to-grad.
Biotech. X Post	0.427***	0.315***	0.339***	−0.454***	0.887	1.857
	(0.083)	(0.064)	(0.069)	(0.065)	(1.617)	(1.769)
Post	−0.053	−0.013	0.003	0.047	−1.457	0.171
	(0.037)	(0.023)	(0.028)	(0.039)	(1.507)	(0.671)
Biotech.	−0.028	0.001	0.046	−0.060	−0.753	0.822
	(0.039)	(0.025)	(0.031)	(0.052)	(0.927)	(0.815)
HS. Marks	−0.003**	−0.003**	−0.003***	0.008***	0.329***	−0.110***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.033)	(0.029)
High school type. Omitted category: Scientific						
Professional	0.072	0.063	0.093	−0.263***	−5.194**	2.798
	(0.076)	(0.058)	(0.063)	(0.083)	(2.008)	(2.020)
Technical	0.039*	0.040*	0.027	−0.101***	−4.272***	0.772
	(0.023)	(0.022)	(0.027)	(0.027)	(0.885)	(0.644)
Classical	0.033	0.060	0.055	−0.058	0.888	1.672
	(0.058)	(0.055)	(0.059)	(0.066)	(1.907)	(2.441)
Linguistic	0.054	0.063	0.036	−0.054	−4.525***	1.313*
	(0.045)	(0.043)	(0.047)	(0.046)	(1.009)	(0.679)
Other HS.	0.045	0.058	0.061	−0.176***	−5.389***	1.953*
	(0.040)	(0.036)	(0.040)	(0.057)	(1.027)	(1.055)
Female	−0.029	−0.017	−0.017	0.069**	−1.437	0.670
	(0.023)	(0.017)	(0.021)	(0.033)	(1.120)	(0.710)
Tuition fee	−.0000327	−.0000278	−9.65e−06	8.46e−06	.0009075*	.000437
	(.0000251)	(.0000229)	(.0000247)	(.0000286)	(.0005146)	(.0003619)
Region dummies	yes	yes	yes	yes	yes	yes
Age dummies	yes	yes	yes	yes	yes	yes
Mean dep. var.	0.16	0.12	0.19	0.65	100.87	40.84
N	1912	1912	1912	1494	514	514

Notes: The table reports the estimation results from Eq. (1) for the matched sample of Biotechnology degree and degrees in medical school. The samples in (1)–(3) cover the academic years at enrollment from 2002/03 to 2014/15. The samples in (4)–(6) cover the academic years of enrollment from 2002/03 to 2011/12. (1) is the probability of not achieving credit by the end of the first year and (2) is the probability of dropping out without achieving credit at the end of the first year. (3) represents the probability of dropping out in the first year. (4) is the probability of graduation. (5) Graduates' final graduation marks. (6) is the duration (in months) of graduates' degree completion. HS. Marks represent the final score on the high school exam. *Post* is a dummy that equals 1 if the academic year of enrollment is after the reform's implementation (i.e., after 2008/09). Regressions include dummies for the region of the high school that students graduated from, and for the age of students at enrollment. Standard errors are clustered at the degree and academic year levels. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

sample for this sensitivity test by performing the matching procedure explained in Section 5.3 on students in the biotechnology degree and our usual control group.

We replicate our main results for a matched sample of students in the biotechnology degree (treatment) and medical school (control). Table 7 reports these results, which are remarkable as we observe an average increase (decrease) of 34 p.p. (45.4 p.p.) in the first-year dropout (graduation) rate, which was 20 p.p. (23.7 p.p.) in the main sample. This finding reveals that the existence of the admission test in degrees in our control group does not generate our results.

6.2.3. Further robustness checks

We conduct robustness tests and present them in Appendix C. Table C1 Panel A shows that excluding students who are enrolled in the 2010/11 academic year, for which we observe the highest first-year dropout rate, from the sample does not alter the results. Panel B shows the results while including degree fixed effects and results are aligned with the main findings.

In Table C2, we re-estimate our main findings by excluding one STEM degree at a time from the sample to ensure that the findings are not driven by the presence of a peculiar degree in treatment group. The results are always in line with the ones presented in the paper.

We also perform a series of additional robustness tests on the matching procedure and present the findings in Table C3. Panel A reports the findings when the final high-school marks of students enter the exact matching rather than allowing a gap of two points between the treatment and control group. Panel B shows the results when the latter gap is set as one point rather than two, and we observe that the results do not alter. Furthermore, Panel C presents the estimates obtained from matching in which tuition fees take place.[47] Although the number of observations decreases, the estimates remain very similar compared to the main findings. In Panel D, we show the results obtained from matching that comprises exact matching categorical variables and propensity score matching based on continuous variables (high school final marks and tuition fees).[48] The results are once again perfectly in line with the main findings of this paper. Finally, we replicate our findings on academic outcomes while controlling for the average tuition fee a student pays throughout their degree and present the results in Table C4. While the tuition fees that students paid after their first enrollment is subject to some level of endogeneity, it is still important to understand to what extent our estimates are affected by the choice of proxies for family background.

6.3. Labor market outcomes

In this section, we examine graduates' early labor market outcomes. We combine our administrative dataset with the *AlmaLaurea* labor market survey, which covers almost all graduates in Italian universities.[49] The merging is based on the students' fiscal codes. Students are contacted for the survey one year after their graduation. We investigate two labor market outcomes of students, namely, employment and net

[47] We split the tuition fee paid in a given academic year by 25th, 50th, 75th percentiles and add the selection to our preferred exact matching procedure in the paper.

[48] We compute propensity scores by running logit regressions. Among the exactly matched students we pick one control unit with the closest propensity score to treated unit.

[49] We only had access to the information regarding the graduates of the UPO.

Table 8
Results on labor market outcomes.

	(1) Employed	(2) Searching job	(3) Not searching job	(4) Monthly salary
STEM X Post	0.257***	−0.262***	0.005	324.199**
	(0.082)	(0.066)	(0.083)	(130.258)
Post	−0.181***	0.138***	0.043	−113.212**
	(0.029)	(0.046)	(0.037)	(49.713)
STEM	−0.690***	0.192***	0.498***	−727.691***
	(0.037)	(0.037)	(0.034)	(82.086)
Control variables	yes	yes	yes	yes
Region FE	yes	yes	yes	yes
Age dummies	yes	yes	yes	yes
Mean dep. var.	0.67	0.12	0.21	1276.18
N	918	918	918	600

Notes: The table reports the results for the matched sample of graduates in STEM and medical school degrees. The sample covers the academic years of enrollment from 2004/05 to 2011/12, where (1) is the probability of being employed, (2) is the probability of searching for a job, (3) is the probability of not searching for a job, and (4) represents the net monthly earnings. *Post* is a dummy that equals 1 if the academic year of enrollment is after the implementation of the reform (i.e., from 2009/10 to 2011/12). Regressions include control variables for the region of the high school that students graduated from, age of students at enrollment, type of high school that students graduated from, final high school grades, gender, and tuition fees at enrollment. Standard errors are clustered at the degree and academic year levels. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

monthly labor earnings,[50] which are the most consistently observed outcomes in the data. Notably, in our labor market survey data, we identify several important insights into student employment or unemployment status. Students can describe their employment status as employed, searching for a job or not searching for a job. We run linear probability regressions for each category separately using Eq. (1).

We observe an extremely high proportion of these students in the labor market dataset, approximately 90% of the graduates, depending on the degree course. Hence, sample coverage only slightly decreases. Ultimately, after excluding "no response" in employment status, the final sample comprises 664 observations of STEM degrees and 1946 observations of medical school from the 2004/05 to 2011/12 enrollment academic years.[51]

However, we present the results for the matched sample among graduates. As discussed earlier, our treatment heterogeneously changed the student composition after their first year of enrollment. Therefore, to balance the sample in terms of student composition among graduates, we perform matching conditional on graduation. This creates a comparable sample by eliminating our treatment's initial effects on first-year dropout rates. The matched sample comprises both STEM degree and medical school students, with 918 observations (308 STEM, 610 medical school).

We present the results for monthly earnings and employment in Table 8. We note a significant relative increase in the employment probability of STEM degree graduates who were subject to new curricula after the reform. Column (1) reveals that the estimated effect is 25.7 p.p. We confirm that no change arises in the probability of not searching for a job (column 3). The sizable effect that we estimate on employment is driven by shifts from searching for a job to being employed. We also find, on average, a significant increase in the net monthly earnings (column 4), although the latter effect is rather weaker and insignificant once estimated year-by-year.[52]

Employment outcomes of graduates in healthcare professions are peculiar with respect to other graduates, as they are more likely employed in the public sector and, consequently, their wages are more rigid. Our identification strategy allows to control for time invariant differences across treated and control groups. As for possible time-varying differences across groups, we might expect a decrease in the wages of the STEM graduates (with respect to the medical control group) over the course of the crisis years, as they are more subject to the market forces. In this case, our findings would be conservative as the moderately positive effect on STEM wages that we find would be net of the potential negative effect of the Great Recession.

7. Underlying mechanisms

In this section, we investigate the potential underlying mechanisms behind the findings that we have presented so far. First, we test the procrastination channel that could play a role in explaining the worsened performance of students in terms of the probability of passing exams, which eventually leads them to dropout at the end of their first-year. If students procrastinate and start to study for unified exams only during the exam weeks, it is more likely that they will fail these exams. Second, we conduct additional analyses to account for the positive effect on labor market outcomes of post-reform graduates. In theory, two broader channels can account for our findings on labor market outcomes: first, the selection at the end of first academic year, which can increase the peer quality to which these students are exposed in the following academic years; and second, improved cognitive and/or non-cognitive skills of graduates as they have to better organize themselves via self-imposed deadlines (e.g. non-cognitive skills) and have to understand and create links between several subjects for unified exams (e.g. cognitive skills).

7.1. Procrastination

We create a proxy for procrastination by following a methodology similar to De Paola and Scoppa (2015), which considers the time difference between when college students (at the University of Calabria in Italy) are informed of admission decisions and when they finalize their enrollment as a proxy for procrastination. The longer that it takes students to finalize their enrollment, the higher procrastinators they are. The authors reveal that such a measure of procrastination is significantly correlated with students' observable pre-enrollment characteristics (e.g., negatively associated with the final high-school marks) and academic performance.

Unfortunately, our data do not contain information regarding the date of admission, and hence we follow a slightly different approach. We consider the earliest enrollment date of a student in a given degree program and enrollment year as a benchmark and calculate the difference (in days) from this date for all other students in the corresponding degree (in the same enrollment year). Our procrastination measure entails an average of eighteen days with a standard deviation of eighteen, ranging from 0–125.[53] In this analysis, we assume that if students procrastinate in one aspect of their lives, they are also expected to procrastinate in other aspects.

We run a descriptive OLS regression of the procrastination measure on students' observable characteristics. The results are reported in column (1) of Table 9. Our procrastination measure is negatively associated with final high-school marks. Further, we observe significant differences by the type of high school from which these students graduated. We do not observe any significant correlation between the

[50] The net monthly earnings is reported in the data at intervals of 250 Euro, starting from 0-250 to 3000+ Euro. We consider the mid points of these intervals as our monthly earning measure (e.g. if it is 1000–1250 Euro, we consider it as 1125 Euro).

[51] We exclude the students of the 2002/03 and 2003/04 cohorts from this analysis due to very few observations.

[52] The results obtained from Eq. (2) on these outcomes are presented in Figure C2 in Appendix C. We see in the figure that the parallel trend assumption holds for the labor market outcomes as well.

[53] We trim 74 observations above the 99th percentile (125 days) of this measure to avoid considering students enrolled several months after the first student (e.g. transfers from other universities in the middle of an academic year).

Table 9
Heterogeneous results by procrastinators.

			Bottom 20%		Bottom 40%		Top 40%		Top 20%	
	(1) Proc.	(2) Proc.	(3) Dropout	(4) No credit	(5) Dropout	(6) No credit	(7) Dropout	(8) No credit	(9) Dropout	(10) No credit
STEM X Post		6.780	−0.069	0.022	0.140**	0.195***	0.231***	0.234***	0.280**	0.330***
		(5.602)	(0.082)	(0.065)	(0.060)	(0.052)	(0.067)	(0.069)	(0.106)	(0.107)
Post		−6.578	0.114***	0.061**	−0.016	−0.078**	−0.068	−0.098*	−0.094	−0.167*
		(4.874)	(0.041)	(0.026)	(0.039)	(0.037)	(0.054)	(0.052)	(0.097)	(0.092)
STEM	20.578***	16.270***	0.115*	0.091*	−0.020	−0.050	0.065	0.040	0.041	−0.013
	(2.021)	(4.244)	(0.062)	(0.049)	(0.034)	(0.037)	(0.054)	(0.052)	(0.090)	(0.090)
HS. Marks	−0.138***	−0.127***	−0.000	−0.000	−0.003**	−0.002**	−0.004***	−0.006***	−0.005***	−0.007***
	(0.027)	(0.043)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
High school type. Omitted category: Scientific										
Professional	3.698*	6.230**	-0.064	0.179*	0.054	0.068	0.212***	0.209***	0.284***	0.208***
	(1.880)	(2.510)	(0.056)	(0.098)	(0.057)	(0.054)	(0.059)	(0.052)	(0.067)	(0.062)
Technical	2.170***	3.597***	0.022	0.082**	0.030	0.014	0.039	0.051**	0.045	0.074**
	(0.670)	(1.125)	(0.030)	(0.034)	(0.026)	(0.024)	(0.035)	(0.025)	(0.044)	(0.036)
Classical	0.339	−1.219	0.078	0.056	0.059	0.045	0.075*	0.096**	0.138**	0.129**
	(0.788)	(1.085)	(0.101)	(0.059)	(0.065)	(0.043)	(0.045)	(0.041)	(0.055)	(0.063)
Linguistic	1.324	2.789*	0.113	0.044	0.046	0.001	0.087*	0.111***	0.163***	0.215***
	(0.895)	(1.567)	(0.113)	(0.036)	(0.050)	(0.040)	(0.052)	(0.039)	(0.055)	(0.045)
Other HS.	0.548	1.347	-0.064	-0.009	-0.010	-0.014	0.109***	0.117***	0.164***	0.156***
	(0.775)	(1.027)	(0.060)	(0.031)	(0.031)	(0.036)	(0.039)	(0.037)	(0.060)	(0.048)
Female	0.487	−0.868	0.011	0.005	−0.037	−0.028	−0.008	−0.028	−0.063**	−0.079**
	(0.643)	(0.858)	(0.033)	(0.026)	(0.024)	(0.019)	(0.026)	(0.026)	(0.031)	(0.037)
Tuition fee	.0000161	−.0005863	.000083***	.0000835***	.0000435*	.0000399*	.0000216	.0000171	.0000374	.0000457
	(.0003514)	(.0003774)	(.0000208)	(.0000175)	(.0000242)	(.0000217)	(.0000284)	(.00003)	(.000039)	(.0000408)
Region dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Age dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Mean dep. var.	17.74	17.74	0.16	0.09	0.15	0.10	0.27	0.22	0.33	0.28
N	8045	8045	1421	1421	3250	3250	3231	3231	1626	1626

Notes: (1) presents results of a regression of procrastination on observable characteristics of students. (2) presents the DID results for procrastination. (3) and (4) show the results on probabilities of dropping out and not achieving any credit in first-year for the students who are at he bottom 20% of procrastination measure, respectively. (5) and (6) show the results for the bottom 40%, (7) and (8) show them for the top 40% and (9) and (10) show the results for the top 20% procrastinators in the final matched sample. The sample covers the academic years at enrollment from 2002/03 to 2014/15. HS. Marks represent the final score on the high school exam. *Post* is a dummy that equals one if the academic year of enrollment is after reform (i.e. from 2009/10 to 2014/15). Regressions include dummies for the region of the high school that students graduated from, and for the age of students at enrollment. Standard errors are in parentheses and are clustered at the degree and academic year levels. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

tuition fees and procrastination. This is reassuring because at the UPO, a student's enrollment date is the date when they pay their tuition fee. Theoretically, one concern is that students from high-income families can make their tuition fee payments earlier, although we do not observe such a situation in our data.

In column (2) of Table 9, we report the results regarding whether the reform affects student composition in terms of procrastination in STEM degrees. This result is obtained from Eq. (1) with procrastination as the dependent variable. On average, STEM degree students procrastinate more than medical school students, and the coefficient estimate of the interaction term is not statistically significant, indicating that the reform did not affect student composition in STEM degrees – compared with that in medical school – in terms of procrastination. Therefore, we can conduct a heterogeneous analysis of academic outcomes by student rank in the distribution of our procrastination measure.

In columns (3) and (4), we observe that students who are at the bottom 20% of procrastination measure are not affected by the reform in terms of probability of dropping out and failing all of the exams at the end of their first year, respectively. Considering the bottom 40% of procrastination in columns (5) and (6), the estimated effects are statistically significant but smaller than the total effects reported in Table 4. As for the top 40% and 20%, we observe larger effects, especially the effect on the probability of failing all exams (column 10). These findings suggest that students who procrastinate more are more strongly affected by sitting unified exams, which is consistent with theoretical expectations.

7.2. Peer effect vs. improved skills

In this section we provide direct and indirect evidence on the potential underlying mechanisms behind the effects on labor market outcomes. We first investigate whether the peer effect channel plays a role by looking at the changes in the composition of students who did not drop out at the end of their first year. To do so, we split the sample by first-year dropouts and non-dropouts, and estimate the diff-in-diff parameter outlined in Eq. (1) for 3 outcomes; namely, share of female students, share of students who graduated from academic track high schools, and high school final marks. Results are presented in Table 10. We do not see statistically significant changes in the skill composition of either group. We also accompany these estimates by looking at the effect of the reform on the share of high-skilled students among the graduates of each cohort, using the specification outlined in Eq. (2).[54] Fig. 9 shows that there is no change in the composition of graduates in terms of the presence of high-skilled students. Taken together, these results indicate the absence of peer effect channel.

As for the indirect evidence, we look at the effects by the intensity of treatment on labor market outcomes of graduates (Table 11). In Panels A and B, we observe significant increases in employment and monthly salaries of graduates from both *less affected* and *more affected* degrees. Recall that the reform significantly affected both the first-year dropout and graduation rates in *more affected degrees* but it only affected graduation rates in *less affected* degrees (see Table 6). Since we observe positive effects on labor market outcomes for the students of each group when compared to students in medical school, the results suggest that

[54] We consider the students who are at the top 10% in the distribution of final high school marks as highly-skilled students.

Table 10
Results on composition of students by first-year dropout decision.

Dependent variable:	Female		Academic HS		HS marks	
	(1) Dropouts	(2) Non-Dropouts	(3) Dropouts	(4) Non-Dropouts	(5) Dropouts	(6) Non-Dropouts
STEM X Post	0.077	0.008	−.00005	−0.029	−1.582	−0.115
	(0.094)	(0.088)	(0.092)	(0.068)	(2.005)	(1.302)
Post	0.112**	0.005	0.377***	0.228***	−0.056	−3.743***
	(0.054)	(0.054)	(0.063)	(0.046)	(1.694)	(1.118)
STEM	−0.066	−0.062	0.026	−0.023	0.567	0.345
	(0.079)	(0.058)	(0.064)	(0.047)	(1.609)	(1.056)
Type of high-school	NO	NO	NO	NO	YES	YES
Mean dep. var.	0.58	0.63	0.47	0.52	73.74	76.47
N	1618	6427	1618	6427	1618	6427

Notes: Dependent variables: Female is the share of female students, Academic HS is the share of students who graduated from academic track high schools, HS Marks is the final high school marks. Columns (1), (3), and (5) show the results for the students who drop out at the end of their first-year, while the columns (2), (4), and (6) show the results for the students who continue their degree after the first-year. Standard errors are in parentheses and are clustered at the degree and academic year levels. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 11
Heterogeneous results by intensity of treatment: effects on labor market outcomes.

	(1) Employed	(2) Searching job	(3) Not searching job	(4) Monthly salary
Panel A: Less affected vs. Medical				
STEM X Post	0.264*	−0.315***	0.051	526.921*
	(0.138)	(0.087)	(0.135)	(291.484)
Mean dep. var.	0.83	0.09	0.08	1349.41
N	660	660	660	532
Panel B: More affected vs. Medical				
STEM X Post	0.269***	−0.263***	−0.006	323.836**
	(0.088)	(0.071)	(0.091)	(141.431)
Mean dep. var.	0.69	0.12	0.19	1284.15
N	868	868	868	582
Panel C: More affected vs. Less affected				
More X Post	0.046	0.062	−0.108	−292.639
	(0.158)	(0.094)	(0.149)	(402.881)
Mean dep. var.	0.28	0.18	0.54	769.25
N	307	307	307	86

Notes: The table reports the results from Eq. (1) on labor market outcomes. Panel A shows the results on students in STEM degrees that are less affected by the reform, Panel B shows them for the students whose degrees are more affected, and Panel C shows the results for the more affected students compared to less affected students. All regressions include control variables on gender, tuition fee, high school final marks, type of high school, age, region. Less affected degrees are computer science and mathematics, while more affected degrees are environmental sciences, biology, biotechnology, and materials science. Standard errors are in parentheses and are clustered at the degree and academic year levels. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

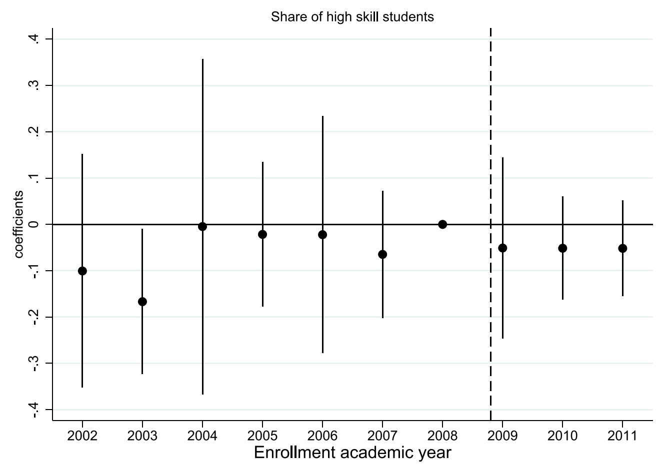

Fig. 9. Event study specification: STEM vs. Medical.
Notes: The figure reports the findings on the share of high-skilled students among the graduates of STEM and medical school degrees. The figure highlights the coefficient estimates of δ_k specified in Eq. (2). The number of observations is 1318. The control variables include the type and region of high schools from which students graduated. Standard errors are clustered at the degree course and academic year level. Confidence intervals are at the 95% level.

the latter effect is more likely to be caused by improved skills during their degrees due to students' effort to pass unified exams.[55]

In the light of these direct and indirect evidence, the peer effect channel, which could be triggered by the increased first-year dropout rate, does not seem to be a factor driving the results on labor market outcomes.

8. Heterogeneous effects

We complement the previous analysis by investigating whether heterogeneous effects of the treatment on several subgroups exist. This analysis can reveal policy relevant implication if students with disadvantaged background are affected worse. Using our matched sample, we assess the results according to gender, high-school track, ability, and family economic situation.

To this end, we expand our baseline econometric model outlined in Eq. (1) and introduce a triple interaction term as follows:

$$\begin{aligned} Y_i = \; & a + \gamma STEM_i + \omega Het_h + \lambda Post_t \\ & + \psi(STEM_i \times Het_h) + \delta(STEM_i \times Post_t) + \Omega(Het_h \times Post_t) \\ & + \pi(STEM_i \times Het_h \times Post_t) \\ & + X_i\beta + \epsilon_i; \end{aligned} \tag{3}$$

where Het_h is a dummy representing the sub-group of interest (female, academic-track high-school, high-ability, high-income family). The definitions of other variables in Eq. (3) remain as in Eq. (1). The parameter providing us the heterogeneous treatment effects is π. We replicate the analysis for each sub-group.

We focus on first-year dropouts and graduates. Ultimately, these are the most important academic outcomes and are free of any concerns regarding selection bias because they are possible for any student enrolled in any academic year.

We present our results in Table 12. We do not observe any significantly different results for female students (columns 1 and 2), suggesting that they suffer from increased workload per exam as much as male students. Considering that a growing body of literature exists on

[55] Additionally, in Panel C of Table 11, where the *more affected* degrees are the treatment group and *less affected* degrees are the control group, we do not observe a significant changes in labor market outcomes. If the graduates in both groups improved their skills by handling unified exams throughout their degrees, it accounts for the latter insignificant result.

Table 12
Heterogeneous effects on first-year dropout and graduation rates.

Het groups:	Female		Academic HS		High-ability		High-income	
	(1) Dropout	(2) Graduate	(3) Dropout	(4) Graduate	(5) Dropout	(6) Graduate	(7) Dropout	(8) Graduate
Post X Het X STEM	0.068	−0.086	0.053	−0.115	−0.206***	0.185	−0.115	0.225**
	(0.066)	(0.076)	(0.061)	(0.073)	(0.073)	(0.115)	(0.081)	(0.092)
STEM X Post	0.164**	−0.190**	0.170***	−0.191***	0.219***	−0.237***	0.228***	−0.304***
	(0.069)	(0.076)	(0.055)	(0.064)	(0.048)	(0.060)	(0.055)	(0.069)
Het	−0.048	0.056	−0.101***	0.108**	−0.162***	0.011	−0.010	0.045
	(0.039)	(0.050)	(0.032)	(0.042)	(0.031)	(0.051)	(0.048)	(0.052)
STEM	0.048	−0.256***	0.036	−0.236***	0.029	−0.220***	0.026	−0.195***
	(0.052)	(0.060)	(0.036)	(0.048)	(0.033)	(0.048)	(0.033)	(0.051)
STEM X Het	−0.020	0.082	0.013	0.065	0.049	0.052	0.066	−0.059
	(0.053)	(0.060)	(0.041)	(0.055)	(0.044)	(0.055)	(0.058)	(0.073)
Post X Het	0.050	−0.027	0.089**	0.003	0.129**	−0.229***	0.091	−0.167**
	(0.046)	(0.058)	(0.038)	(0.057)	(0.051)	(0.069)	(0.055)	(0.065)
Post	−0.082	0.093	−0.090***	0.079	−0.066**	0.089**	−0.080**	0.125***
	(0.052)	(0.060)	(0.033)	(0.049)	(0.032)	(0.041)	(0.031)	(0.045)
Region dummies	yes	yes	yes	yes	yes	yes	yes	yes
Age dummies	yes	yes	yes	yes	yes	yes	yes	yes
N	8045	4748	8045	4748	8045	4748	8045	4748

Notes: The table reports the estimation results from Eq. (3) for the sample obtained from exact matching for degrees in STEM and medical schools. In columns (1), (3), (5), and (7), the sample covers academic years at enrollment from 2002/03 to 2014/15. In columns (2), (4), (6), and (8), the sample covers the academic years at enrollment, from 2002/03 to 2011/12. Het stands for heterogeneous groups, which takes the value of 1 if the student is female in (1)–(2), if the student graduated from an academic track school in (3)–(4), if the student's final high school mark is in the top 10% in (5)–(6), and if the tuition fee that the student is supposed to pay is in the top 30% in (7)–(8). $Post$ is a dummy that takes a value of one if the academic year is after the reform at enrollment. All regressions include control variables for gender, age, type of high school, final high school marks, and the region of the high school that the students graduated from, and tuition fees. Standard errors are in parentheses and are clustered at the degree and academic year levels. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

the gender gap in STEM- subjects,[56] our findings regarding female students' performance carry further implications for understanding their academic performance in STEM subjects.

Columns (3) and (4) do not reveal any significant heterogeneous effects for students graduating from academic-track high schools. Further, we examine students' ability, using final high-school marks as a proxy for ability. We consider the top 10% of the distribution of high school marks, which corresponds a score above 93 in our data, as high-ability students.[57] Recall that these results are obtained from a matched sample wherein high school marks and tracks are balanced over the years across the treatment and control groups, while we control for school tracks in the regression. Columns (5) and (6) present the results for high-ability students. Consistent with theoretical expectations, these students are relatively less affected by sitting unified exams. They are 20.6 p.p. less likely to drop out and 18.5 p.p. more likely to graduate compared with low- and mid-ability students. Moreover, this result possibly helps us to better explain not finding an effect on graduation marks and time to graduate. Overall, the increased difficulty of exams is leveled off by the high-skilled students' ability, and therefore no effect is noted among graduates.

Additionally, we investigate whether any differential effects exist for students from high-income families. Accordingly, we consider the tuition fees at enrollment. Tuition fees in Italy are based on (equivalent) family economic situation (income and assets). Tuition fees are set by each university according to national regulations. The lower the family income, the lower the tuition fee that students have to pay, up to a complete exemption. Regarding students from low-income families, scholarships are also available (e.g., exemption from tuition fees, food stamps, and free accommodation). We consider the top 30% of the tuition fee distribution as representing high-income families. These students might have access to extra resources during their degrees, which allows them to cover university costs. Our results corroborate these arguments (see columns 7 and 8), as students from more affluent families seem less likely to drop out by the end of their first year by 11.5 p.p., although this finding is not statistically significant. The heterogeneous effect is even more pronounced for the graduation rate with a significant coefficient estimate of 22.5 p.p. Having access to extra resources seems to help these students to mitigate the difficulties in sitting unified exams. From a policy perspective, this result suggests that an increase in the inequality of opportunity is an unexpected effect of bundling exams.

9. Concluding remarks

In a world of growing task complexity, it has become relevant to investigate whether time allocation across assignments is efficient, as it entails the level of productivity attained. Nevertheless, this link has predominantly been explored for workers and has not been analyzed for university students (e.g. Coviello et al. 2014, 2015). To address this gap, this study documents the first evidence of the relationship between the intensity of workload per course and the key academic outcomes of university students by exploiting an exogenous variation in the number of courses induced by a national tertiary education reform. Accordingly, we investigate the effect of sitting unified exams (i.e., increased workload per course) by comparing the differences in the college outcomes of the treatment (STEM degrees) and control (medical school) groups before and after the reform's implementation. Our findings are stark: for students in STEM degrees, we find that the reform's introduction increased the first-year dropout probability by 20 p.p. and the effect persisted after several years. Concerning the graduation outcome, we note a negative effect (23 p.p.) on the graduation rate of STEM students.

Additionally, the results obtained from event-study specifications confirm the existence of parallel trends in the analyzed outcomes before the new curricula's introduction. Moreover, our estimation sample is created by performing exact matching on the observable student characteristics, which reassures that these results are not driven by changes in student composition. To assess the potential heterogeneous effects on dropout and graduation outcomes, we run estimates by gender, academic track, ability, and family economic situation. The findings do not reveal any gender differences in the effect of increasing workload per course, however they demonstrate that an unintended consequence of bundling exams is that students from more affluent families are favored, thus increasing the inequality of opportunity.

[56] See, e.g., Kahn and Ginther (2018).

[57] This variable ranges from 60 to 100.

Unequal access to higher education is a universal concern, far beyond the Italian context. Therefore, our results have policy implications for the organization of degree courses in terms of workload intensity. Finally, considering graduates' early labor market outcomes, we find that graduates exposed to the reform have better job market prospects, as their employment probability is significantly greater.

The search for possible mechanisms of our results led us investigate the impact of the "attitude of procrastination" on students' ability to organize their study time. Accordingly, we tested whether students who tend to delay commitments/decisions have more difficulty handling curricula with intensified courses. The findings demonstrate that the probability of dropping out is significantly higher for students who are at the top of the procrastination distribution, i.e. those who have a more pronounced procrastination behavior.

Additional mechanisms can be at play in explaining post-reform labor market outcomes, which can be only partially tested due to data availability. First, students in our treatment group have to better organize their schedules and study-time to pass more comprehensive exams. This could thus enhance the non-cognitive skills of post-reform students, such as an increased ability to handle more complex situations, optimize and balance different work and non-work activities, making them more easily aligned with the demands of the labor market.[58] Second, studying for a single exam that covers a greater quantity of concepts enhances both the depth of knowledge in the subject matter and, at the same time, the ability to make connections both within various topics and across different disciplines (i.e. improved cognitive skills). This increases the overall human capital endowment of students exposed to the reform in both the university outcomes and the job market. It is then plausible to believe that for students who graduated after the reform, the improvement of the skills mentioned earlier may contribute to explaining the better employment outcome of these cohorts. A third channel could be an average increase of peer quality in degree courses.[59] However, our investigation on the intensity of treatment reveals that even the students in *less affected* degrees, whose decisions to dropout at the end of their first year remain unaffected, experience improvement in their labor market outcomes. We do not observe any changes in the share of high-skilled students among the graduates after the reform either. These evidence thus corroborate the first two mechanisms.

Our findings suggest that the impact of the reform was not temporary in two respects. First, the effects remain significant over several years and across different student cohorts, indicating that challenges faced by the initial treated group persisted for subsequent cohorts. Second, the exploration of labor market outcomes reveals that treated cohorts exhibit improved employment prospects, suggesting a positive impact of the reform. The analysis of potential mechanisms further supports the argument that post-reform graduates acquire valuable cognitive and non-cognitive skills, contributing to their positive outcomes in the labor market.

As for the external validity of our findings, the analysis of the degree programs in social sciences and humanities shows a lesser or null impact of the reform, depending on the outcome. While we cannot exclude that this may depend on the different learning processes within these fields of study, we believe that the varying intensity of the treatment may account for the different effects of the reform as well. Indeed, degree programs in the socio-humanistic fields had to make less drastic changes to their curricula compared to STEM degrees. Therefore, we cannot generalize our findings across other fields in higher education. However, STEM degrees are the keystone for fostering the development of skilled-labor, hence our study's contribution remains relevant.

All in all, our results indicate a trade-off for educational policymakers. On the one hand, intensified courses act as effective selection mechanisms of students enrolled in open admission degree programs. This selection mechanism is probably fairer than an admission test where students have a single chance to be admitted to the degree program chosen. On the other hand, this selection might be too harsh and undesirable especially in a country such as Italy, which struggles to increase the number of university graduates.[60] Moreover, if this selection not only harms the weakest students, namely those who are less able and less organized, but at the same level of ability, primarily those who are less affluent, this creates a glaring problem of equality of opportunity in tertiary education.

CRediT authorship contribution statement

Carmen Aina: Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Koray Aktaş:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Giorgia Casalone:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization.

Data availability

The authors do not have permission to share data.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.labeco.2024.102559.

References

Adamopoulou, E., Tanzi, G.M., 2017. Academic drop-out and the great recession. J. Hum. Cap. 11 (1), 35–71.
Aina, C., Baici, E., Casalone, G., 2011. Time to degree: students' abilities, university characteristics or something else? Evidence from Italy. Educ. Econ. 19 (3), 311–325.
Aina, C., Baici, E., Casalone, G., Pastore, F., 2022. The determinants of university dropout: A review of the socio-economic literature. Socio-Econ. Plan. Sci. 79, 101102.
Anelli, M., Peri, G., 2019. The effects of high school peers' gender on college major, college performance and income. Econ. J. 129 (618), 553–602.
Anelli, M., Shih, K., Williams, K., 2023. Foreign students in college and the supply of STEM graduates. J. Labor Econ. 41 (2), 511–563.
Ariely, D., Wertenbroch, K., 2002. Procrastination, deadlines, and performance: Self-control by precommitment. Psychol. Sci. 13 (3), 219–224.
Arteaga, C., 2018. The effect of human capital on earnings: Evidence from a reform at Colombia's top university. J. Public Econ. 157, 212–225.
Becker, S., 2006. Introducing time-to-educate in a job-search model. Bull. Econ. Res. 58 (1), 61–72.
Belley, P., Lochner, L., 2007. The changing role of family income and ability in determining educational achievement. J. Hum. Cap. 1 (1), 37–89.
Bostwick, V., Fischer, S., Lang, M., 2022. Semesters or quarters? The effect of the academic calendar on postsecondary student outcomes. Am. Econ. J.: Econ. Policy 14 (1), 40–80.
Bostwick, V.K., Weinberg, B.A., 2022. Nevertheless she persisted? Gender peer effects in doctoral STEM programs. J. Labor Econ. 40 (2), 397–436.
Bound, J., Lovenheim, M.F., Turner, S., 2012. Increasing time to baccalaureate degree in the United States. Educ. Finance Policy 7 (4), 375–424.
Bratti, M., Checchi, D., De Blasio, G., 2008. Does the expansion of higher education increase the equality of educational opportunities? Evidence from Italy. Labour 22, 53–88.
Brenøe, A.A., Zölitz, U., 2020. Exposure to more female peers widens the gender gap in stem participation. J. Labor Econ. 38 (4), 1009–1054.
Buser, T., Peter, N., 2012. Multitasking. Exp. Econ. 15 (4), 641–655.

[58] Heckman et al. (2006) provide evidence that non-cognitive skills are significantly associated with individuals' labor market outcomes.

[59] See e.g., among others, evidence on the effect of peer compositions on students' performance and decisions (Anelli and Peri, 2019), Brenøe and Zölitz (2020), Bostwick and Weinberg (2022), Anelli et al. (2023).

[60] In 2021, the Italian share of 25–34 year-olds with a tertiary education was 28.3%, whereas on average across the OECD this proportion was 47.5% (OECD, 2022).

Cameron, A.C., Gelbach, J.B., Miller, D.L., 2008. Bootstrap-based improvements for inference with clustered errors. Rev. Econ. Stat. 90 (3), 414–427.

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

Card, D., 1999. The causal effect of education on earnings. In: Handbook of Labor Economics. Vol. 3, Elsevier, pp. 1801–1863.

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

Chen, X., 2013. STEM Attrition: College Students' Paths into and Out of STEM Fields. Statistical Analysis Report. NCES 2014-001, National Center for Education Statistics.

Coviello, D., Ichino, A., Persico, N., 2014. Time allocation and task juggling. Amer. Econ. Rev. 104 (2), 609–623.

Coviello, D., Ichino, A., Persico, N., 2015. The inefficiency of worker time use. J. Eur. Econom. Assoc. 13 (5), 906–947.

Danilowicz-Gösele, K., Lerche, K., Meya, J., Schwager, R., 2017. Determinants of students' success at university. Educ. Econ. 25 (5), 513–532.

De Paola, M., Scoppa, V., 2011. Frequency of examinations and student achievement in a randomized experiment. Econ. Educ. Rev. 30 (6), 1416–1429.

De Paola, M., Scoppa, V., 2015. Procrastination, academic success and the effectiveness of a remedial program. J. Econ. Behav. Organ. 115, 217–236.

Di Pietro, G., 2006. Regional labour market conditions and university dropout rates: Evidence from Italy. Reg. Stud. 40 (6), 617–630.

Ghignoni, E., 2017. Family background and university dropouts during the crisis: The case of Italy. Higher Educ. 73 (1), 127–151.

Goulas, S., Megalokonomou, R., 2020. Marathon, hurdling, or sprint? The effects of exam scheduling on academic performance. BE J. Econ. Anal. Policy 20 (2), 20190177.

Griffith, A.L., 2010. Persistence of women and minorities in STEM field majors: Is it the school that matters? Econ. Educ. Rev. 29 (6), 911–922.

Heckman, J.J., Stixrud, J., Urzua, S., 2006. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. J. Labor Econ. 24 (3), 411–482.

Imbens, G.W., Rubin, D.B., 2015. Causal Inference in Statistics, Social, and Biomedical Sciences. Cambridge University Press.

Isphording, I., Qendrai, P., et al., 2019. Gender Differences in Student Dropout in STEM. IZA Research Report87.

Kahn, S., Ginther, D., 2018. Women and Science, Technology, Engineering, and Mathematics (STEM). In: The Oxford Handbook of Women and the Economy.

Krashinsky, H., 2014. How would one extra year of high school affect academic performance in university? Evidence from an educational policy change. Can. J. Econ./revue Can. Econ. 47 (1), 70–97.

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

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

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

OECD, 2022. How many students complete tertiary education? Education at a Glance 2022: OECD Indicators. OECD Publishing, Paris.

Peri, G., Shih, K., Sparber, C., 2015. STEM workers, H-1B visas, and productivity in US cities. J. Labor Econ. 33 (S1), S225–S255.

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

Stinebrickner, R., Stinebrickner, T.R., 2003. Working during school and academic performance. J. Labor Econ. 21 (2), 473–491.

Stinebrickner, R., Stinebrickner, T., 2008. The effect of credit constraints on the college drop-out decision: A direct approach using a new panel study. Amer. Econ. Rev. 98 (5), 2163–2184.

Stinebrickner, R., Stinebrickner, T., 2013. A major in science? Initial beliefs and final outcomes for college major and dropout. Rev. Econ. Stud. 81, 426–472.

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

Tuckman, B.W., 1998. Using tests as an incentive to motivate procrastinators to study. J. Exp. Educ. 66 (2), 141–147.

Villalobos, C., Kim, H.W., Huber, T.J., Knobel, R., Setayesh, S., Sasidharan, L., Galstyan, A., Balogh, A., 2021. Coordinating STEM core courses for student success. PRIMUS 31 (3–5), 316–329.