I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14423

When Need Meets Merit: The Effect of Increasing Merit Requirements in Need-Based Student Aid

Tommaso Agasisti
Massimiliano Bratti
Veronica Minaya

MAY 2021

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14423

When Need Meets Merit: The Effect of Increasing Merit Requirements in Need-Based Student Aid

Tommaso Agasisti
Politecnico di Milano

Massimiliano Bratti
Università degli Studi di Milano, IZA and LdA

Veronica Minaya
Columbia University

MAY 2021

IZA – Institute of Labor Economics

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

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

www.iza.org

ABSTRACT

When Need Meets Merit: The Effect of Increasing Merit Requirements in Need-Based Student Aid*

Merit requirements in need-based student aid may exacerbate inequality in higher education but at the same time improve efficiency of aid expenditure by increasing on-time graduation, for instance. Disentangling the effect of the two building blocks of student aid ("need" and "merit") is therefore of key interest to policy makers. In this paper, we seek to estimate the causal effect of tightening the academic requirements embodied in need-based student aid on short-term and long-term student academic performance. This is done leveraging a reform in an Italian region that increased by 40% (i.e. from 25 to 35 out of a maximum of 60) the number of credits to be earned in the first academic year to maintain aid eligibility. Using administrative data from an Italian public university mainly offering STEM degrees, this study reveals that tightening merit requirements had a statistically significant, positive effect on various dimensions of performance of the "average" aid recipient. However, an analysis of treatment heterogeneity unveils winners and losers from the policy: the positive effects are indeed concentrated among higher and medium-ability students, while lower-ability students receiving financial assistance are discouraged from continuing in their studies.

JEL Classification: student financial aid, merit-based requirements, university, difference-in-differences, Italy

Keywords: I21, I22, I28

Corresponding author:
Massimiliano Bratti
Dipartimento di Economia Management e Metodi Quantitativi (DEMM)
Università degli studi di Milano
Via Conservatorio, 7
20122 Milano
Italy

E-mail: massimiliano.bratti@unimi.it

* We gratefully acknowledge comments from participants in seminars and conferences in which earlier versions of this paper were presented, among which the 2019 Association for Education Finance and Policy (AEFP) Conference (Kansas City), IRVAPP (Trento), the 2019 annual conference of the Italian Association of Labour Economists (AIEL), the 2019 workshop "The Economics of Higher Education" ZEW (Mannheim) and the Ninth Italian Congress of Econometrics and Empirical Economics (2021). The usual disclaimer applies.

1. Introduction

In many European countries, equal access to higher education (HE henceforth) is perceived as a fundamental right within the welfare state. Public universities are either free or subject to low tuition fees compared to the US and UK. Despite increasing financial pressures to improve the performance and competitiveness of European HE systems (Aghion et al., 2010), fees are seen as a financial barrier that would prevent students from entering HE. In this context, a key policy question is how to design financial aid packages that promote equal opportunities, particularly for young people from disadvantaged backgrounds.

Need-based financial aid is a widely used tool to help make college more affordable for low-income students. While need-based aid programs are initially based purely on financial need, most of these programs attach merit-based requirements for renewal. Merit-based requirements may increase efficiency of aid expenditure but at the cost of reducing equality of educational opportunities (*equity-efficiency trade-off*). These requirements may increase the risk of losing funding because of failure to renew and consequently increase dropout rates. Therefore, it is important for policymakers to better understand the role of merit-based requirements in need-based aid and to address both efficiency and equity concerns.

Yet, merit-based requirements for need-based aid programs have received little attention from researchers and policymakers. Need-based programs are designed as a "bundled" package, and in studies assessing the effect of financial aid on student performance, it is generally difficult to disentangle the effect of the "need" component from that of the "merit" component. Although the literature finds that minimum merit requirements in need-based aid may hurt persistence of low-income students, arguing that merit requirements promote inequality in HE (Scott-Clayton and Schudde, 2020), there is little evidence on the effect of *increasing merit requirements.* An important exception is Montalbán (2019), who studied the causal effect of receiving the same amount of grant under different intensities of merit requirements on student performance, degree completion, and dropout in an HE institution in Spain. He found positive effects of being eligible for a grant on student performance when combined with stronger merit requirements. His results suggest that merit requirements attached to need-based aid may be an effective tool to improve student performance and aid effectiveness. We contribute to this scant literature by leveraging a policy change introduced in 2011, when the regional government of the most populous Italian region (Lombardy) raised the merit requirements for students to maintain their need-based grant (while the need-based requirements remained practically unchanged). We draw from this reform to explore the consequences of tightening the merit requirements on need-based student aid for cohorts of students enrolled at the *Politecnico di Milano* (PoliMi, hereafter) after the policy change.

In this paper, we estimate the causal impact on a set of short and long-run academic outputs, such as the performance in the first year of bachelor studies (number of credits earned and average grades obtained in the first year, i.e. GPA), student retention and credits/grades later in the academic career, probability of on-time graduation, i.e. within the legal duration of the degree course (three years), or with a maximum of a one-year delay. Our analysis focuses on

first-time-in-college students enrolled in a bachelor's degree (BA) at PoliMi between 2008 and 2013; we follow these cohorts of students until 2017. We use a difference-in-differences (DID) approach to evaluate whether aid-eligible students in cohorts enrolled after the reform experienced changes in academic performance because of the reform itself—namely, due to tightening of the merit component of the financial aid (as separate from criteria based on financial need).

The main results of this study lead to an important conclusion: the policy had on average a positive and statistically significant impact on several dimensions of academic performance. Specifically, students affected by the change in merit requirements earned more credits in the first year and improved their longer-term academic outcomes, such as the probability of obtaining the BA degree on time. This is a very important result, since in Italy many students delay graduation because the system allows them to get better grades while staying longer in higher education (specifically, by repeating exams until getting the desired grade, see Section 2). However, the effects are heterogenous and we find that the positive effects are concentrated among the medium and higher-ability students, namely those who received a score placing them in the second-third or in the highest quartile of the admission test, respectively. The former significantly increased the number of credits (about 3 credits in the first year) and the probability of graduating in 3 or 4 years. The latter benefited in terms of higher first-year GPA and higher final graduation grades. By contrast, the increase in the credit requirements reduced the probability of lower-ability students (first quartile of admission test) enrolling in the second year by 10 percentage points.

In this paper we focus on Italy, which we believe could be an interesting case for several reasons. First, tuition fees at universities in Italy are relatively high compared to other European countries (European Commission, 2018) and have been rising since the introduction of the "Bologna Process," and particularly after the economic crisis in 2008 (Civera et al., 2020). Therefore, need-based financial aid constitutes an important part of Italian HE to ensure that disadvantaged students are not discouraged from accessing HE. Second, Italian HE has long been characterized by a very high proportion of dropouts and graduates behind schedule. Significant numbers of dropouts occur during the first year of enrollment (Mealli and Rampichini, 2012), and this seems to be explained by university costs (tuition fees and housing costs), financial constraints and lack of access to need-based aid (Modena et al., 2020; Ghignoni, 2016). Thus, it could be interesting to understand how students respond to changes in need-based aid programs. Stronger merit requirements to maintain need-based grants may be harmful to low-income students and exacerbate inequality in HE. Third, within the Italian context, PoliMi is an even more intriguing case. It is a prestigious, selective, public university (ranked #1 for Engineering in Italy, and among the top 20 in Europe) whose academic offerings are focused on STEM disciplines, which yield higher earnings returns than degrees in other fields (Altonji et al., 2014). The implications for our analysis are discussed in the conclusions.

This paper makes important contributions towards designing financial aid packages. Results suggest that the specific design of the financial aid package matters in determining its actual effects on student academic performance. Although increasing merit-based requirements provide the right incentives to low-income students for succeeding in HE, our analysis demonstrates highly heterogeneous effects depending on student ability (or academic

preparedness as reflected in the admission test scores), showing that increasing efficiency may come at the cost of also amplifying educational inequalities. Indeed, we find that the reform reduced the re-enrolment probability of lower ability low-income students. Although at first glance, this may seem consistent with the objectives to be achieved, i.e. reducing student aid to "underperforming" students, it raises nonetheless serious concerns. As for efficiency, it must be kept in mind that the lower ability students are nevertheless students that managed to enter a very prestigious and highly selective institution through an admission test. Thus, their inability to keep pace with the requirements is particularly worrying as we might expect stronger negative effects in universities adopting similar policies but having much less selective admission requirements. Moreover, student dropout may have negative effects on the supply of STEM graduates, who are highly demanded in the labor market. As for equity, our analysis demonstrates that unlike those receiving aid, and quite expectedly, lower ability students from advantaged family backgrounds (i.e. not receiving financial aid) were not negatively affected by the reform. Thus, the policy might have reduced the equality of educational opportunities according to socio-economic background. This raises the question of at what level HE institutions should set the bar for merit requirements. Overall, the main message is that effectively balancing criteria based on financial need and merit could be the key to success for guaranteeing equality of opportunities for disadvantaged students but at the same time a high efficiency of student aid.

This paper is organized as follows. Section 2 describes the institutional background, namely the financial aid system in the Italian HE system. Section 3 reviews some related academic literature about the effects of merit-based financial aid on university student performance in Europe and the US. Section 4 provides a simple conceptual framework to discuss the effects of financial aid packages and interpret the empirical results. Section 5 presents the data and discusses identification. Section 6 presents the results from the empirical analyses and Section 7 some robustness checks. Section 8 critically discusses the main findings and draws conclusions.

2. Institutional background: the financial aid system in Italian higher education, *Politecnico di Milano* and the policy change of 2011/12

In Italy, the role of financial aid in promoting equality of educational opportunity is crucial. The Italian financial aid system is regulated at the national level and called *Diritto allo Studio Universitario* (DSU, hereafter). "DSU-aided" students are those receiving grants under the national financial aid system. When national financial resources are not available to cover all eligible students (so creating a list of eligible students who actually do not receive the subsidy), some universities allocate their own resources to fill the gap. PoliMi allocates between 3 and 5 million € every year for this purpose, guaranteeing that 100% of its eligible students receive the grant. This is an important feature of the aid mechanism; in cases like PoliMi's, it is likely that the institutional attention to the effectiveness of grants is particularly high. Moreover, as a substantial amount of funds is allocated by the university itself, it constitutes a potential factor of student attraction.

The financial aid package has three components: grants, free lunch and housing. Grants are by far the major component of financial aid; indeed, in addition to a financial transfer per year, eligible students are exempted from tuition fees (for this reason, the paper explores the measured effects generated by the change in the policy for assigning grants). The amount of the grants depends on whether the student lives in the city where the university is located (*in - site*) or commutes to reach the city (*commuting*) or is an out-of-site student (*out-of-site*). In the years covered by this study, the minimum amount of the grant was an average of around 2,000€, 3,000€ and 5,000€ for the three categories, respectively (every year, the amount is determined by a national government decree and updated to be aligned with inflation). Grants are assigned on the basis of financial need when the students enroll for the first time at a bachelor or master course. Financial need is assessed through an index that is an equivalized economic situation indicator (called ISEE), which is computed on the basis of a student's family income, and the level and composition of wealth. The threshold for eligibility is set by the national government on a yearly basis. The application is submitted to the university, and eligibility is communicated and confirmed some weeks after enrolment (usually in October, while classes start in September).

In order to maintain aid eligibility after the first year of enrollment, students need to meet merit requirements (number of credits—CFUs, hereafter—earned in a given academic year) as well as need-based eligibility. At this point, more information about how PoliMi allocates financial aid may be useful. In their first year, students apply for the grant directly to the university in June/July before the start of the academic year; classes start in September. Students receive provisional information about eligibility in October, and then a final confirmation in December, i.e. after enrolment. At that moment, students receive the first half of the grant. Students are then assessed on August 10th; if they meet the credit threshold, they also receive the second half of the grant, otherwise, they are no longer eligible for that transfer. If students earn the required number of CFUs, they are automatically eligible for the grant in the subsequent year and receive the cash transfer (first half) in September. If the family's economic situation changed substantially, the students must communicate this and a new assessment is carried out to verify that they still meet the need criteria for eligibility. The first-year students also have an additional requirement to fulfil; if they do not reach the merit threshold of the credits by August 10th, they are scrutinized again on November 10th. If by then they do not obtain the cutoff number of credits, they must pay back to the university the whole amount received up to then, i.e., the first instalment of the grant. Thus, the merit requirement is a high-stakes incentive towards reaching satisfactory progression in the first year.

Before the reform, the national regulation identified the threshold for maintaining eligibility for the second year at 25 CFUs (out of 60), to be obtained before August 10th (before the last exam sessions of the first year held in September). The number of CFUs to be obtained before August 10th of the second and third year to maintain eligibility for subsequent academic years was 80 and 135 respectively, so after the first year the students were expected to acquire 55 credits (out of 60) every year.

In 2011, the Lombardy Region (the most populous region of Northern Italy and where PoliMi is located, counting about 10 million inhabitants,) introduced a change in the financial aid system starting in the academic year 2011/12. The regional government reached an

agreement with the local universities to allow them to increase the merit requirement for maintaining eligibility for grants in the second year. After analyzing the data about student performance in the past years, the regional government communicated to universities that the average number of credits obtained by all students (DSU-aided and not aided) was around 35 (out of 60) in the previous three years, which was suggested as a benchmark for those universities modifying the merit-based criteria for grants.

As a result, PoliMI set a new threshold for maintaining aid eligibility in the second year, increasing it from 25 to 35 (out of 60) university credits,[2] a 40% increase for first-year bachelor or master students starting from academic year 2011/2012.[3] The number of CFUs earned to maintain the grant in the second and third year remained unchanged at 80 and 135 credits, respectively. At the same time, PoliMI agreed with student unions that it would have allocated money to provide grants to all eligible students even in the case of insufficient national and regional funds (something that systematically happened every year). The policy and institutional rationale of this choice was that this increase in merit-based requirements would have stimulated grant recipients to study harder and to obtain their credits more rapidly, possibly with a long-term effect of making them able to earn their bachelor's degree more quickly and on time (i.e. within the legal duration of study, three years). All the other rules remained virtually unchanged: the income threshold for obtaining the grant in the first year and maintaining it over time, the amount of the grant (except for inflation adjustment), exemption from fees, etc.

As anticipated, the aim of this paper is to estimate the effect of changing the merit requirement of the DSU program. Such change was equivalent to passing an additional course in the first year (in the first year, most courses are 10 credits, with a teaching load of 100 classroom hours) to meet the DSU renewal requirements. The paper uses three cohorts of students before (academic years from 2008/09 to 2010/11) and two cohorts after (the 2011/12 and 2012/13 academic years) this merit requirement changed.

The design of the reform offers a unique setting to test whether a change limited to the merit component of the grant had an effect on the performance of students after the introduction of the policy. This paper aims to explore this issue by employing a DID estimation strategy comparing cohorts of students before and after the reform.

Given that one of the performance dimensions potentially impacted by the reform is average GPA (in addition to the number of credits), it is useful to clarify a technical aspect of the examination system in Italian universities. The students can sign up to sit the exam at the end of each course, and they can obtain a grade (if positive) between 18 and 30 with honors. If the grade assigned is less than 18 (fail), the students must retake the exam in a later occasion. However, the students can also opt for "refusing" the grade assigned and retake the exam later in the semester or even in subsequent semesters. While the number of available options for

[2] One CFU generally corresponds to 25 hours of student workload.

[3] The general rule described here applies to the Engineering degree courses, which are the majority, while some differences exist for Architecture and Design degrees. When considering Architecture, the number of credits requested for the grant was already higher and equal to 30, so the reform increased the necessary amount only by 5 credits. For the courses in Design, PoliMi set an even higher threshold of credits for obtaining the grant – i.e. 40 instead of 35 (after reform). In this paper, we estimate the average effect of the reform on all fields.

taking exams in a given academic year varies across universities, in many of them there are at least five sessions per academic year. Moreover, in many cases there is neither a limit to the number of times an exam is retaken, nor limits in the specific sequence for taking exams. In this vein, students can retake many exams several times, deferring degree completion until they are satisfied with their GPA (it is worth noting that CFUs are considered obtained once the grade assigned by the teachers is "accepted" by the student). At PoliMi, there are five exam sessions per year—two in the winter term, two in the summer and one in September. No precise statistics are available about the number of exam attempts per student, however retaking the exams after the refusal of a specific grade is quite usual and many students exert this option. In this perspective, average GPA is somehow driven not only by student effort, ability and performance, but also by students' choices (at least partially). This particular feature of the Italian HE system must be kept in mind when observing and commenting on the effect of the policy under scrutiny. A potential trade-off can arise for students: indeed, the incentive of the reform is towards a more rapid acquisition of credits (to obtain a sufficient number of credits for grant eligibility) but this can be in contrast with habits of refusing grades to improve GPA. Hence, the choices about the acceptance (or not) of credits has a direct consequence also on time-to-degree, which is another key performance dimension considered in our evaluation exercise.

3. Related literature

Student aid design matters and varies across different types of students. Dynarski and Scott-Clayton (2013) provide an extensive review of the effectiveness of financial aid programs in the US and suggest that design matters for improving student performance. Specifically, their study suggests that merit-based incentives within the grant/aid systems are helpful for stimulating better performance of eligible students.

The impact of merit-based financial incentives on student performance has been previously studied in the US and European contexts. Scott-Clayton (2011) conducted a rigorous evaluation of the role of merit-based incentives by examining the PROMISE program in West Virginia, which offers free tuition to students who maintain a minimum GPA and course load. Using a regression discontinuity design (RDD), she found that this financial incentive program had significant effects on academic achievement. By exploring the mechanisms behind her results, the author found that the effects were concentrated around the annual merit requirements for scholarship renewal and not around financial constraints.[4]

Todd et al. (2020) studied a change in the rules to receive full tuition scholarships provided to high-achieving students through Georgia's HOPE Scholarship program. Under the new rules, only students meeting more rigorous merit-based criteria would retain the original scholarship covering full tuition, now called Zell Miller, with other students seeing aid reductions. They exploit discontinuities around HS GPA and SAT/ACT cutoffs for Zell Miller eligibility to analyze the effect of partially losing financial aid on college persistence and

[4] Unfortunately, the administrative data we use does not contain information on exams passed but refused or attempted but failed.

graduation of high-achieving students. The authors did not find evidence that the financial aid reduction affected persistence or graduation for these students, which suggests that high-achieving students may be less sensitive to prices once they are in college.

Belot et al. (2007) draw from a major reform in the financial aid system in the Netherlands to study the impact of student support on the performance and time allocation of students. The reform, which was introduced in 1996, reduced the maximum duration of grants by one year and limited it to the nominal duration of studies. The authors employed a DID approach to assess the impact of this reform using cohorts of students between 1995 and 1997 (one year before and after the reform). They found that reform had positive effects on student performance in the first years of HE (students switched less to other programs and obtained higher grades) but did not have an impact on the time allocation of students. This merit-based modification of the rules for obtaining grants helped the financial aid system to work more efficiently and effectively.

The role of merit-based financial incentives was explored in a randomized trial conducted at the University of Amsterdam in 2001/02. A sample of first-year students were eligible to receive a cash amount if they passed all their exams in the year. The results of the experiment, reported by Leuven et al. (2010), revealed that the incentive had a positive effect on high-ability students and a negative effect on those of low ability.[5] This study, albeit giving useful insights on the potential role of merit-based components of financial interventions, cannot be generalized to the problem of designing grants as it was not properly a financial aid program, but instead a "pure incentive" (i.e. a merit-based incentive).

While the literature investigating the effect of introducing merit-based requirements on student performance, where the comparison group is typically made up of students not receiving aid, offers several important contributions, evidence on the effects of tightening the merit component of student financial aid is much less developed. To the best of our knowledge, the only other paper that investigates this effect is the one authored by Montalbán (2019), which focuses on the *Becas de Carácter General* (BCG, the Spanish national financial aid program devoted to low-income students in tertiary education) and the case of Carlos III University students. Leveraging a reform of student aid, introduced in 2013, which changed the minimum merit requirements in terms of credits to be earned and introduced a requirement in terms of GPA, the author compares students who only received a fee waiver with those who were eligible and also received a cash allowance. Montalbán finds that the latter performed significantly better than the former in terms of GPA and credits earned, but only after the reform. By contrast, he does not find significant positive effects from the reform when comparing students who received different amounts of cash allowances and concludes that student aid design rather the simple existence of minimum merit requirements may improve student performance.

Our analysis has some major differences with Montalbán (2019). The first is that he focuses only on *students who applied for student aid* and received some form of aid (at least a

[5] A similar experiment conducted in a Canadian campus led to quite different results. The incentive was designed to offer cash incentives for course grade above 70, i.e. to stimulate a strong merit level. The results reported by Angrist et al. (2014) indicate no strong effects of this incentive scheme, with a null effect on GPA specifically.

fee waiver) and uses an RDD with different income thresholds, which determine different types of treatments within the population of students receiving aid. In his case, all students included in the analysis were subjected to the same increasing merit requirements after the reform. In our paper, identification is based on a DID design in which we contrast changes over time of the performance of students in aid, who were affected by the increase in credit requirements after the reform, with those of students who were not receiving aid and therefore were not subject to the reform. Second, DID allows us to estimate effects that are less "local" than those provided by Montalbán (2019), which apply to students in the vicinity of the income thresholds determining the different types of treatments. Another difference is related to the characteristics of the BCG aid reform that allow Montalbán to isolate the effects of increasing merit requirements only for the fee waiver vs. low-amount grant treatment. The latter however only applies to relatively higher-income students among those receiving aid. We estimate instead an average effect for all students in aid. Finally, there are some institutional differences between Italy's and Spain's university and student aid systems. Unlike in the Spanish system where also first-year students are subjected to some merit-based requirements to receive financial aid (the score of the university entrance exam must be above a given cut-off), in the Italian system eligibility in the first year is purely need-based. Moreover, in Spain undergraduate degrees last four years while in Italy three years. In the case analyzed by Montalbán, students could receive aid up to one year (two years for STEM degrees) after the official length of the degree, while in Italy only for the official length of the degree. Finally, the increase in merit requirements in the case of Montalbán not only involved an increase in the number of credits to be earned but also the introduction of a new requirement on GPA. It is therefore interesting to provide evidence on the effect of reforms that were designed differently but all increased merit requirements, albeit in different higher education systems.

Our paper is also related to recent work by Scott-Clayton and Schudde (2020), who investigate the effect of failing satisfactory academic progress (SAP, hereafter) requirements on student performance in one state in the US. The authors provide evidence in the context of a US state that failing SAP has larger discouragement effects. Their preferred DID estimates point to a decrease in second year enrollment of 6 percentage points, but an increase of GPA of 0.11 grade points for students close to the merit threshold. Moreover, longer-term effects are negative, with reductions in credits attempted and credits earned of 3.4 and 1.4 credits, respectively. This study also sheds light on the potential redistributive effects of failing merit requirements as they also found that dropout effects of SAP failure are stronger for low-income students. These results confirm their earlier findings from examining an SAP policy in a different state (Schudde and Scott-Clayton, 2016). Scott-Clayton and Schudde (2020) used a variety of identification strategies to estimate these effects. In an RDD they estimate the effects of failing SAP within the population of student aid receivers. They also compare the effect of failing SAP for aided students vs. non-aided students in a DID-RDD design. Finally, they use a DID, which enables them to estimate effects also far from the GPA cutoff related to SAP failure. This last specification is the most similar to what we do in our analysis, as it compares aid vs. non-aid students, although our work has two main differences with Scott-Clayton and Schudde's study. First, the latter investigates the consequences of failing the GPA requirements. By contrast, we compare two regimes, leveraging a reform that increased merit

requirements. Hence, we focus on a different policy parameter. Second, we argue that merit requirements (and a reform increasing them) may have other effects on student performance in addition to those mediated by a higher probability of failing these requirements. For instance, in the specific case we consider (Italy), the increased pressure towards meeting the new university credits requirement may reduce student GPA (in Italy students can re-sit exams and "accept" grades only when they are satisfied with their exam results, and they are quite obsessed with their GPA) and induce some students to switch to another degree or institution, or even to drop out from higher education. This effect may materialize even for students that manage to meet the merit requirements. In other words, an increase in merit requirements may have larger effects than those induced by failing the required performance standard.

4. A simple conceptual framework for the analysis of stricter merit requirements

The effect of tightening the merit requirements in merit-based student financial aid can be analyzed using a simple framework of student effort and academic performance for university students.

Let us consider a static model in which an individual maximizes a utility function $U(C, x; a)$, where C is lifetime consumption of a composite good and x is "student effort" (with $x \in [0,1]$), which gives disutility (i.e. a *bad*). We omit the individual subscript to simplify notation. Moreover, we assume that the disutility of studying depends negatively on individual ability a, i.e. time devoted to studying creates a higher disutility to lower-ability individuals. Consumption is equal to individual non-labor income y (e.g. parental transfers) plus labour income, with the life-time labour market reward of student effort (as a university graduate) assumed to be linear in effort, and a scholarship s that is paid only for levels of performance—coinciding with effort for simplicity[6]—above a given threshold. The budget constraint then reads as $C = \alpha x + y + s\, I(x \geq x_0)$, where $I(.)$ is the indicator function. The scholarship s can be thought of as additional lifetime "study" income that is provided to the student if performance (effort in this simple model) is above a given cutoff x_0. The reform consists in increasing the level of performance above which the scholarship is paid from x_0 to x_1.

Firstly, we must distinguish between students eligible and students ineligible for student aid according to family income.[7] For the second group, the change in the merit-based requirements does not change individual incentives and should not produce any effect on student performance.[8] So in our empirical analysis non-aided students will make the natural control group. Conversely, the behavior of students receiving financial aid is likely to be

[6] For ease of exposition, we assume that labor income is linear in academic performance and that the educational production (or performance) function is $f(x) = x$, i.e. only effort matters. Assuming a more general production function $f(x, a)$ would make the budget constraint nonlinear and depend on an individual's ability without changing the main implications of the sketched model.

[7] For instance, in our empirical framework the budget constraint $C = \alpha x + y + s\, \mathbb{I}(x \geq x_0)$ only if $y \leq y_0$ and $C = \alpha x + y$ otherwise.

[8] To be more precise, in case of spillovers (e.g. peer group effects) of the reform from the treated to the untreated individuals also the latter's performance might be affected. In the Italian case being exam grading at university mainly absolute and not relative (e.g. grading on a curve) we do not expect substantial spillover effects.

impacted by the reform. The increase in the merit requirements, in the absence of an increase in the amount of the scholarship, cannot have a positive effect on individual utility because as shown in panel a of Figure A1 (online Appendix A) the set of feasible consumption bundles is reduced by the grey area F after the reform. Thus, in this simple framework individual utility either remains the same or is reduced after the policy change. However, ex-ante we cannot say anything on the direction of the change in the optimal amount of effort x^*, which may increase, decrease or remain the same as before the reform, with effects on academic performance that go in the same direction.

Figure A1 shows five examples of possible effects of the reform. Subfigure (a) depicts the case of a student who before the reform was choosing a low level of effort (and consumption) and is not affected by the reform (i.e. his or her pre-reform optimal bundle is still feasible after the reform). This is likely to happen when indifference curves start to be steep at very low levels of effort, e.g. for lower-ability students who require a high level of consumption to compensate even a relatively small increase of effort. Subfigure (b) displays the situation in which the lower-ability student (steep indifference curves) chooses a higher level of effort after the reform compared to before: academic performance improves, but utility decreases. In particular, even if in the student's optimal choice the student raises effort and performance, he or she loses the scholarship as performance is below the new threshold x_1. Subfigure (c) depicts the case in which the reform induces a lower-ability student to reduce effort and performance and lose the scholarship. Subfigure (d) shows the situation in which the student was choosing a very high level of effort before the reform, e.g. because indifference curves are very flat, i.e. a higher-ability individual. In this case the student is not affected by the increase in the merit-based requirements. Other cases are also possible, for instance some students (most likely students with intermediate ability levels) might increase effort and performance and retain the scholarship after the reform, as shown in Subfigure (e).

This simple representation of a student's choice problem suggests that the effect of the reform can be *heterogeneous* and depend, among other things, on the level of family income (y), the returns to education (α), the amount of student aid (s), how much the merit-based requirements are increased ($x_1 - x_0$) and how indifference curves are shaped with respect to individual ability. Thus, when assessing the issue empirically, it is important to focus not only on the average effect of the reform, but also to investigate effect heterogeneity across different population subgroups (by level of ability, for instance). As for the interplay between the reform and student ability, for instance, according to the conceptual framework above, we put forward that *the reform is more likely to have positively affected the performance of intermediate-ability students and negatively affected those of lower ability, while leaving unaltered the performance of higher-ability students*.

In a fully-fledged dynamic model in which students decide whether to enroll in HE or not and academic performance is not fully deterministic but subject to random shocks, students can decide whether to continue or to drop out from education and we might expect additional effects of tightening the merit requirements of student aid. For instance, some students (most likely the low ability ones) may fail to meet the requirements and decide to drop out from HE. Moreover, policy change may modify the average ability of entrants in HE (*sorting*), especially those of low income (which receive DSU), since the higher effort needed to get the scholarship

after the reform would entail a lower increase in disutility for higher-ability students compared to their lower-ability peers.

5. Data and identification

5.1 Data

This paper uses administrative data from PoliMi.[9] PoliMi's student data-tracking system is very comprehensive and allows us to control for background characteristics (including standardized admission test scores taken during secondary school), information on financial aid, and transcript and degree information such as year-by-year college enrollment (e.g., CFUs, GPA). One thing the financial data do not include before 2013 is the amount of financial aid and income eligibility measures (e.g., ISEE). However, even in these years we observe an indicator of financial aid and university tuition fee levels, and we use this last piece of information in the empirical model as a proxy of socio-economic background (as explained in the subsequent sections).

We focus on first-time-in-college BA students who entered PoliMi between the 2008/2009 and 2013/2014 academic years, and we follow these cohorts until 2017/2018. We further limit the sample to students who entered PoliMi when they were between 17 and 25 years of age (i.e. immature students) and to Italian students. Across these six student cohorts we have about 33,000 observations.

For each student, we distinguish academic outcomes between short-term and long-term outcomes: (i) short-term academic outcomes include credits earned by August 10th of the first year (first deadline to meet merit requirements), credits earned by November 30th (second deadline to meet merit requirements), credits earned in the first year, GPA on August 10th of the first year, GPA on November 30th of the first year, first year GPA, probability of enrolling in the second year of studies; and (ii) long-term academic outcomes including the probability of enrolling in the third year, the probability of graduation within the legal duration, the probability of graduation within the legal duration or with one year delay, the final graduation score (conditional on graduation). Long-term academic outcomes for the pre-policy period are estimated only using the 2008 cohort of entrants to avoid partial treatment issues (i.e. partial exposure to the policy change during some years of their academic career) due to the policy change being introduced in 2011.[10]

Table 1 displays descriptive statistics separately for the periods before and after the change in the merit component of the DSU financial aid program. Almost 3% of the sample received DSU during their first year of enrollment—we label them "DSU-aided" (or simply DSU) students. DSU students are less likely to be female. Moreover, we observe a reduction in female DSU students after the reform. DSU students have higher admission scores compared to non-

[9] We are grateful to the Offices for Student Support and for Information Systems who provided us the data and help in cleaning and interpreting the variables in the dataset.

[10] Students of the 2008/2009 cohort are in their fourth year of studies (one year more than the degree's legal duration) in 2011/2012, the year of the reform, and therefore cannot be affected by student aid, just because they cannot receive it any longer.

DSU and the gap has increased since the reform. Since students know the rules of the student aid package in advance, as they apply for aid at the time of enrollment, a tightening of the merit-based component may produce "cream skimming" of low-income students, attracting those with relatively higher ability. In the empirical analysis, we address this concern.

[Table 1] here

The statistics computed on the raw data shown in Table 1 already hint to some positive effects of the reform on student performance. Both the number of credits achieved in the first year and the first GPA appear to have increased. A significant increase (about 8.4 percentage points) in the probability of graduating within the course's legal duration emerges from the longer-term outcomes. However, these differences may hide changes in student cohort characteristics. For this reason, we implement a DID estimation strategy.

5.2 Identification of the causal effect of the reform

The natural setting to evaluate the effect of the reform is a DID research design. The DID approach enables us to make a comparison between DSU-aided and non-aided students before and after the policy change that tightened the merit requirements. We estimate the effect of raising the merit requirements of the DSU program on the outcome variables described in the previous section using the following parametric specification:

$$Y_{it} = \alpha + \beta(DSU_i * After_t) + \delta Cohort_t + \sigma DSU_i + \gamma X_i + u_{it} \qquad (1)$$

where Y_{it} represents the outcome of student i of cohort t. $After_t$ is a dummy variable that takes the value of zero if the student entered for the first time at PoliMi between 2008 and 2010 (pre-reform period) and one if the student entered between 2011 and 2013 (post-reform period). DSU_j takes the value of one if the student received the DSU grant (the treated group) at first entry and zero if the student did not meet the income eligibility requirements to receive financial aid (the control group). $Cohort_t$ is a vector of student cohort fixed effects, $\mathbf{X_i}$ is a vector of student-level covariates that controls for demographic and pre-college ability and u_{it} an idiosyncratic error term. Covariates included are gender (female), score in the standardized admission test (proxying ability), dichotomous indicators for student fee brackets (proxying income, and in general socio-economic status), residence (mutually exclusive categories: Milan – reference, Lombardy, Other North-Western region, North-Eastern region, Central region, Southern region, Insular region, missing residence), subject field (Engineering – reference, Design, Architecture, Other). β is the DID effect of interest given by the interaction $DSU_i * After_i$.

[Table 2] here

5.3 Identifying assumptions

The critical identifying assumption in the DID approach is that the coefficient on the interaction term $DSU_i * After_t$ from Equation (1) would be zero in the absence of the policy change. Pre- and post-reform cohorts of students may be different, and within each cohort, DSU aided and non-aided students may be different; but nothing can be different about being a post-policy student receiving a DSU grant other than the new merit requirement. In other words, trends in academic achievement of pre-policy cohorts should be good predictors of what would have happened in the absence of this financial aid reform in the post-reform period (*parallel trend assumption*). Moreover, since we use pooled cross-sections, another condition to apply DID is that the composition of the treated and control groups should not have radically changed across cohorts. This may be an issue since, as we observed in the previous section, the reform might have produced "cream skimming" (or sorting) effects.

Evidence on the parallel trend assumption is provided in Figures A1 and A2 in the Appendix, which report the pre- and post-reform trends in the outcome variables for DSU-aided and non DSU-aided students. We have few time observations, but all in all the graphs do not show strong violations of the parallel trend assumption for the short-term outcomes (Figure A1). The parallel trend assumption seems to be more problematic for the long-term outcomes (Figure A2), although it appears to break down especially because of the last pre-reform cohort. As we mentioned, the 2010 and 2009 student cohorts are likely to be partially treated by the reform, and therefore do not make good control cohorts. This may explain the failure of parallel trends. For this reason, our analysis of the long-term outcomes only includes the 2008 pre-reform cohort.[11]

The effect of compositional changes in the student cohort on the outcomes is investigated in Table 2 using the method proposed by Carrell et al. (2018) and recently applied in the context of student-aid by Scott-Clayton and Schudde (2020). The main idea is to test the overall contribution of the change in the observables in the DSU-aided vs. non DSU-aided student groups before and after the reform to the DID estimated coefficient, or in other words to check whether the changes in the observables are sufficient alone to produce spurious evidence of a DID effect of the reform. This is done by running a regression of the outcomes on the control variables, taking the predicted values and running the specification of equation (1) (without the controls X_i) on the predicted outcomes. Table 2 shows the DID coefficients and the standard errors for all outcomes listed in the first column. Reassuringly, the coefficients on $DSU_i * After_t$ estimated on the predicted outcomes are never statistically significant, suggesting that there are no sizeable changes in the composition of the treated and control groups over time that can spuriously produce a statistically significant DID coefficient.

Although these results on the lack of important compositional changes are reassuring, as previously noted, it seems that students receiving the DSU grant are getting slightly better admission test scores over time (0.15 of a standard deviation higher). Thus, it is important to include this background characteristic as a control variable in the analysis. We try different specifications, including this proxy of student ability both linearly and using a flexible

[11] The pre-reform period is too short to implement a "fake DID" which limits the analysis to the period preceding the reform and imputes the reform to a year within this period.

polynomial in some robustness checks. Moreover, in the robustness checks we also report matching-DID estimates.

Finally, the validity of our DID estimates hinges on the assumption that, except for the reform of the merit requirements, individuals were facing the same institutional setting or environment before and after the reform. As we already mentioned, no other key institutional feature of the aid system or PoliMI functioning was touched by the reform. Importantly, the amounts of the scholarships were only marginally increased to adjust for inflation.

6. Baseline results and robustness checks

The baseline results for the estimation of the effects of the policy change on student performance are reported in Table 3. As for the short-term outcomes, the reform increased the number of credits earned in the first year (1.6), the number of credits earned by August 10th and November 30th of the first year, 2.9 and 2, respectively. Consistent with the fact that the first deadline was the most important, as meeting it would have implied receiving the second instalment and also avoiding the risk of failing the second deadline and having to repay the first instalment, effects are highest for CFUs achieved by August. Some of these advantages are lost as students proceed further in the academic year, suggesting that they may simply allocate their time in an optimal way and increase effort especially in the first part of the academic year in order to meet the first deadline, while reducing it in the rest of the year after achieving this important goal. Effects are not only short-term. Indeed, positive effects are also found on the probability of graduating within the degree's legal duration (three years), which increased by about 8.9 percentage points (pp, hereafter) and on the probability of graduating in three or four years (i.e. either "on time" or with one year of delay), which went up by 7 pp. All in all, the reform seems to have produced beneficial effects on the speed of student careers, allowing students to reduce graduation times.

[Table 3] here

This is an important finding, especially in light of the long average graduation times characterizing Italian students, and the high percentage of students delaying graduation (Aina et al. 2011, Garibaldi et al. 2012).

In what follows, we report some robustness checks on the baseline specification. First, in order to better control for student ability, we include a third-degree polynomial in student admission test score. This is important as the reform might have changed the ability profile of DSU students enrolling in PoliMI, this in turn affecting DSU student performance. Results in Table 4 are reassuring. The estimated effects are very close to those in Table 2, namely 1.8, 3 and 2.2 for credits earned in the first year, August 10th and November 30th, respectively. Longer-term outcomes are also very close, 8.7 pp and 7.1 pp on the probability of graduating in three years or with at most one year of delay, respectively.

[Table 4] here

The baseline estimates in Table 2 include students from all family income backgrounds. However, students receiving aid come from low-income households. Thus, we might have a lack of common support between DSU and non-DSU students, and this may be problematic especially if family income has a strong independent effect on student performance over and above student aid. This motivates the robustness check we implement in Table 5, which shows the estimates of equation (1) in different samples that gradually include students with larger family incomes. We compare the estimates when including students in the first four, first five, first six, and all university fee brackets. Two main results stand out in Table 5. First, estimates on CFUs are quite robust to varying the sample composition by family income. Although estimates on number of CFUs attained in the first year, and on August 10th and November 30th are more precise in larger samples, they are comparable in magnitude. Second, when we limit the comparison to students with more similar family incomes (especially the first five or six fee brackets), we also find a significant (generally at 10% level) positive effect on the GPA (between 0.2 and 0.3 points depending on the sample and the date at which GPA is measured). Estimates on longer-term outcomes follow a similar pattern, and are generally larger and more precise when we consider larger samples but are of comparable magnitudes across samples. The results confirm a positive effect of the reform on both on-time graduation and graduation within the legal duration or with a one-year delay.

[Table 5] here

Although the evidence in Section 5.2 shows that the reform did not significantly change the composition of student cohorts, at least in the relevant characteristics impacting on student performance, as a robustness check in Table 6 we implement matched-DID estimates (Villa, 2016). The estimates are based on kernel matching and computed on the common support. The model controls for all the variables in equation (1); the same variables except the fee brackets are used to perform matching.[12] The results are generally consistent with those in Table 2, although the estimates are slightly smaller in magnitude and less precise.

[Table 6] here

7. Heterogeneity effect: Winners and losers from tightening the merit requirement

In this Section we explore whether the effects of the policy change are related to the different initial ability of students, i.e. if the policy has heterogeneous effects over the

[12] Fee brackets are excluded from the matching variable due to the lack of common support (as non-DSU students are on average richer).

distribution of student abilities. To this end, we divide students in three groups using the quartiles of the admission test score distribution (by entry-cohort and field): "lower ability" are those falling in the first quartile, "medium ability" those falling in the second and third quartile, and "higher ability" those falling in the fourth quartile.

The results for the short-term outcomes are reported in Table 7. The most striking finding is that lower-ability students were harmed by the reform. Although the reform does not seem to have had any significant effect on the number of credits achieved, implying that it may have increased the number of students failing the merit requirements due to the higher cutoff, it lowered student GPA and more importantly induced a 10 pp decrease in the probability of enrolling in the second year. In the Italian system students can sit an exam in several exam sessions during the year and can refuse grades and retake the exam if they are not satisfied with their performance. The higher merit requirements may have pressured students, especially those of lower ability, into attempting exams by August 10th and accepting low grades. As for the number of credits achieved, the results also clearly indicate that the reform had the largest effects on medium-ability students, with estimates ranging between 3 and 4 credits depending on the date in which they are measured, and also on higher-ability students, showing some positive effects when credits are measured on August 10th (2.7 credits). In other words, the average effects found in Table 2 were driven by medium and higher-ability students. Higher-ability students seem to have improved their GPA thanks to the reform by 0.3-0.5 points. All in all, our results of inequality-increasing effects of merit requirements are in line with Scott-Clayton and Schudde (2020), although our study specifically focuses on the effect of *raising* these requirements and not of failing them.[13]

Table 8 shows the results on the longer-term outcomes. These estimates suggest that the average positive effect of the reform on the probability of graduating on time or with a one-year delay that was found in Table 2 was mainly driven by higher and medium-ability students. In terms of higher efficiency levels, no significant effect is observed for lower-ability students. For the latter, the reform, if anything, seems to have worsened final graduation score, especially for those graduating within the legal duration. This effect matches what we have found for the first-year GPA of low-income students, who, fearing to lose financial aid, might have reduced their "reservation grades" and become willing to accept lower grades compared to before the reform.[14,15] Lower grades may in turn affect post-graduation outcomes. Indeed, some studies

[13] By contrast, Montalbán (2019) does not find that the increasing merit requirements penalized those students who failed them.

[14] It must be kept in mind that when considering the probability of graduation within a given time interval from enrollment, the outcome takes on a value of zero both for dropouts and for delayers, i.e. those that are still enrolled. This is the reason why one could find a significant negative effect on the probability of enrolling in the second year, and still not find any significant negative effect on the probability of graduation.

[15] We also investigated gender differences in the effect of the reform. Results are reported in Appendix A. In particular, we investigated heterogeneous effects by gender. We find that the reform had similar effects for males and females on the number of credits achieved on August 10th (Table A1). However, compared to men this effect is diluted over time. Indeed, the "speed" premium due to the reform disappears for women when first-year credits or credits on November 30th are considered. The same is observed for longer-term outcomes (Table A2): the reform increased the probability of graduating on-time by 11 pp (significant at 10%) for women and 7.1 pp for

in the UK context find that a lower final graduation score is associated with worse labor market outcomes (e.g. Naylor et al. 2016; Feng & Graetz, 2017), while research on Italy is still scant.

[Tables 7 and 8] here

8. Discussion and concluding remarks

There has been a proliferation of need-based aid programs attaching merit-based requirements to financial aid as incentives to improve student success (Anderson et al., 2020). While the need-based component of these programs helps more disadvantaged students to get access to higher education, the goal of merit-based requirements is to increase the efficiency of student aid by raising student effort and performance. Excessively high merit-based requirements, though, may induce some students to fail the standards set for aid renewal or for not having to repay the aid, and drop out from the university. In other words, merit-based requirements may create an "equity-efficiency trade-off". Despite the importance of this issue, research studying the effects of increasing merit-based requirements in need-based aid is still scant.

In this paper, we aim to fill this gap using quasi-experimental evidence provided by a policy change that increased first-year credit requirements for student aid renewal by 40% in the most populous Italian region (Lombardy). Our analysis on administrative data from an Italian university mainly offering STEM degrees (*Politecnico di Milano*) shows that the reform increased student performance along several dimensions. Our DID estimates show average positive effects of the reform both on first-year student outcomes, namely on the number of credits acquired, but also in the longer term, with an increase of students graduating within the degree's legal duration. Thus, some students are likely to have benefited also economically from the reform, as a shorter time-to-degree is generally associated with better employment and earning outcomes (see, for instance, Aina and Casalone 2020 in the Italian context).[16]

On the other hand, our heterogeneity analysis by student ability (measured through the admission test score) uncovers winners and losers from the policy change. Indeed, the reform decreased the probability of lower-ability students (i.e. those in the bottom quartile of the entry test score) re-enrolling in the second year by about 10 percentage points. We put forward two main hypotheses for explaining this finding. First, our results suggest that completely losing financial support may have decreased college persistence of lower-ability students who did not meet the merit requirements. Second, the reform also reduced their GPA as they had to accept lower grades to fulfill the merit requirements (in the Italian HE system students can retake the exams if they are not satisfied with the grades received). Given that lower grades would have affected the final graduation score, lower-ability students might have left PoliMI to enroll in a

men (statistically nonsignificant). However, the results are reversed when we consider the probability of graduating in 3 or 4 years, for which we observe an about 11 pp increase for men but a nil effect for women.

[16] Although assessing causality in this context is difficult, longer graduation times may simply signal lower graduates' skills to employers (Aina and Pastore, 2020).

less selective HE institution. Unfortunately, with only PoliMI data at hand, we do not know whether this is true dropping out from HE or deciding to enroll in another HE institution. In the best-case scenario, the reform is likely to have worsened the academic performance in the student careers of lower-ability students or forced them to leave a selective institution, raising equity concerns. In the worst-case scenario, the policy may have induced lower-ability students to drop out from higher education, with even greater individual and societal economic losses.

Although lowering participation of less-able students might be the intended goal of tightening merit requirements, and some could perhaps argue that we should not care too much about it, we argue that this is not the case. First, this runs against equity, since students with similar ability levels from more advantaged backgrounds were not subjected to the same performance requirements. It is therefore key not to set the bar of "merit"—which only applies to disadvantaged students, i.e. those receiving financial aid—too high. Secondly, our measure of student ability is based on performance in a university entry exam. This may depend not only on innate ability or "merit" (e.g. student effort) but also on inputs that are likely to be affected by socio-economic backgrounds, such as secondary school quality or out-of-school inputs, e.g. private tutoring classes (Sianou-Kyrgiou, 2008; Guimarães and Sampaio, 2013). Thus, our results may be interpreted also as stricter merit requirements penalizing more students who were disadvantaged from start, i.e. when entering higher education. Finally, as we argue below, finding negative effects on disadvantaged students enrolled in highly selective institutions, in which the student intake is likely to be positively selected, and for STEM degrees, which offer very high labor market returns, may be particularly worrying from an equity perspective. In other words, too high merit requirements may push disadvantaged students out of the most prestigious higher education institutions and from STEM degrees.

By leveraging quasi-experimental variation, our analysis shows some clean and robust evidence on the effect of tightening merit-based requirements in a specific institutional setting. However, our results might not be easily generalizable to all contexts. On the one hand, the university considered in our analysis is located in the major town of Northern Italy (Milan), where students face very high costs of living. Thus, losing financial aid may be particularly harmful and some students may drop out from the university or transfer to universities in other cities to continue their tertiary education studies. This could make our results on lower-ability students an "upper-bound" (in absolute value) estimate of the reform's effect. On the other hand, the university we consider in our analysis is a highly selective institution offering STEM degrees, which command very high returns in the labor market and attract highly motivated and abler students compared to the average Italian university. This probably makes the positive results on higher- and medium-ability student performance an "upper-bound" too, but a "lower bound" for the probability of not enrolling in the second year of lower-ability students. Indeed, in view of the high expected returns, the latter may be more likely to persist in their studies even after losing student aid compared to those enrolled in less prestigious and selective HE institutions for which the higher costs induced by similar reforms might greatly exceed the returns of continuing in their studies. These specific features of the student population under analysis might question the external validity of the results presented in the paper in that they might not apply homogeneously to different HE institutions in Italy and elsewhere. Thus, as a suggestion for future research, it would be useful to carry out similar analyses using

administrative data from less selective HE institutions, from other cities, regions or on the whole Italian (or other countries') higher education system(s). Unfortunately accessing these data is not easy, at least in the Italian context, mainly for data confidentiality restrictions.

A final important conclusion of our study is that as student aid programs are generally designed as "all or nothing" packages when it comes to meeting merit requirements, policy makers should consider second chances for students who lose funding because they do not meet these requirements the first time around. As suggested by Todd et al. (2020), a redesign of merit-based renewal requirements under the form of a sliding scale, in which students only partially lose financial aid depending on their performance may more effectively increase aid efficiency without exacerbating inequality.

References

Aghion, P., Dewatripont, M., Hoxby, C., Mas-Colell, A., Sapir, A. (2010). The governance and performance of universities: evidence from Europe and the US. Economic Policy, 25(61), 7-59.

Altonji, J. G., Kahn, L. B., Speer, J. D. (2014). Trends in earnings differentials across college majors and the changing task composition of jobs. American Economic Review, 104(5), 387–393.

Aina, C., Baici, E., Casalone, G. (2011). Time to degree: Students' abilities, university characteristics or something else? Evidence from Italy. Education Economics, 19(3), 311–325.

Aina, C., Casalone, G., (2020). Early labor market outcomes of university graduates: Does time to degree matter?, Socio-Economic Planning Sciences, 71(C), 100822.

Aina, C., Pastore, F. (2020), Delayed graduation and overeducation in Italy: A test of the human capital model versus the screening hypothesis, Social Indicators Research, 152(2), 533-553.

Anderson, D. M., Broton, K. M., Goldrick-Rab, S. Kelchen, R. (2020). Experimental Evidence on the Impacts of Need-Based Financial Aid: Longitudinal Assessment of the Wisconsin Scholars Grant. Journal of Policy Analysis and Management, 39(3), 720-739.

Angrist, J., Oreopoulos, P., Williams, T. (2014). When opportunity knocks, who answers? New evidence on college achievement awards. Journal of Human Resources, 49(3), 572-610.

Belot, M., Canton, E., Webbink, D. (2007). Does reducing student support affect scholastic performance? Evidence from a Dutch reform. Empirical Economics, 32(2-3), 261-275.

Carrell, S. E., Hoekstra, M., Kuka, E. (2018). The long-run effects of disruptive peers, American Economic Review, 108(11), 3377-3415.

Civera, A., Cattaneo, M., Meoli, M., Paleari, S., Seeber, M. (2020). Universities' responses to crises: The influence of competition and reputation on tuition fees. Higher Education. https://doi.org/10.1007/s10734-020-00622-2.

Dynarski, S., Scott-Clayton, J. (2013). Financial aid policy: Lessons from research The Future of Children, 23(1), 67-91.

European Commission/EACEA/Eurydice (2018). National Student Fee and Support Systems in European Higher Education – 2018/19. Eurydice – Facts and Figures. Luxembourg: Publications Office of the European Union.

Feng, A., Graetz, G. (2017). A question of degree: The effects of degree class on labor market outcomes, Economics of Education Review, 61(C), 140-161.

Garibaldi, P., Giavazzi, F., Ichino, A., Rettore, E. (2012). College cost and time to complete a degree: Evidence from tuition discontinuities. Review of Economics and Statistics, 94(3), 699–711.

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

Guimarães, J., Sampaio, B. (2013). Family background and students' achievement on a university entrance exam in Brazil. Education Economics, 21(1), 38-59.

Leuven, E., Oosterbeek, H., van der Klaauw, B. (2010). The effect of financial rewards on students' achievement: Evidence from a randomized experiment. Journal of the European Economic Association, 8(6), 1243-1265.

Mealli, F., C. Rampichini (2012). Evaluating the effects of university grants by using regression discontinuity designs, Journal of the Royal Statistical Society: Series A (Statistics in Society), 175(3), 775–798.

Modena, F., Rettore, E., Tanzi, G. M. (2020). The effect of grants on university dropout rates: Evidence from the Italian case. Journal of Human Capital, 14(3), 343-370.

Montalbán, J. (2019). Countering moral hazard in higher education: The role of performance incentives in need-based grants. ffhalshs-02160365v2f.

Naylor, R., Smith, J., Telhaj, S. (2016). Graduate returns, degree class premia and higher education expansion in the UK, Oxford Economic Papers, 68(2), 525-545.

OECD (2020), "Introduction", in Education at a Glance 2020: OECD Indicators, OECD Publishing, Paris,https://doi.org/10.1787/acc7a871-en.

Schudde, L., Scott-Clayton, J. (2016). Pell Grants as performance-based scholarships? An examination of satisfactory academic progress requirements in the nation's largest need-based aid program, Research in Higher Education, 57(8), 943–967.

Scott-Clayton, J., Schudde, L. (2020). The consequences of performance standards in need-based aid: Evidence from Community Colleges, Journal of Human Resources, 55(4), 1105-1136.

Scott-Clayton, J. (2011). On money and motivation a quasi-experimental analysis of financial incentives for college achievement. Journal of Human Resources, 46(3), 614-646.

Sianou-Kyrgiou, E. (2008) Social class and access to higher education in Greece: supportive preparation lessons and success in national exams, International Studies in Sociology of Education, 18:3-4, 173-183.

Todd R. J., Kreisman, D., Rubenstein, R., Searcy, C., Bhatt, R. (2020). The Effects of Financial Aid Loss on Persistence and Graduation: A Multi-Dimensional Regression Discontinuity Approach. IZA Discussion Papers 13849, Institute of Labor Economics (IZA), Bonn.

Villa, J.M. (2016). Diff: Simplifying the Estimation of Difference-in-differences Treatment Effects. The Stata Journal, 16(1), 52-71.

Table 1. Means of student characteristics and student outcomes by DSU aid status, before vs. after the reform

Variables	Mean before		Mean after		DID
	a	b	c	d	(d-c)-(b-a)
	Non-DSU	DSU	Non-DSU	DSU	
"Short-term" outcomes sample					
Background characteristics					
female	0.332	0.348	0.318	0.280	-0.054*
admission test score	0.064	0.249	0.049	0.384	0.149**
age at entry	19.264	19.123	19.305	19.089	-0.075
Short-term outcomes					
Credits earned 1st year	39.752	50.595	40.136	53.383	2.404*
Credits earned Aug 10	34.559	45.652	34.358	49.149	3.698***
Credits earned Nov 30	40.188	50.680	40.271	53.474	2.711*
GPA 1st year	24.091	24.549	24.055	24.823	0.309*
GPA Aug 10	24.212	24.597	24.182	24.904	0.336*
GPA Nov 30	24.094	24.553	24.057	24.828	0.313*
Enrolled 2nd year	0.803	0.966	0.799	0.954	-0.008
"Long-term" outcomes sample					
Background characteristics					
female	0.344	0.360	0.318	0.280	-0.054
admission test score	0.049	0.276	0.049	0.384	0.108
age at entry	19.205	19.061	19.305	19.089	-0.072
Longer-term outcomes					
Enrolled in the 3rd year	0.712	0.907	0.721	0.954	0.038
Graduated in 3yrs	0.325	0.444	0.331	0.534	0.084**
Graduated in 3/4 yrs	0.586	0.757	0.582	0.820	0.067
Graduation mark 3 yrs	100.449	101.832	100.747	101.230	-0.899
Graduation mark 3/4 yrs	97.817	99.136	97.991	98.509	-0.801

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses.

Note. This table reports sample means of some student characteristics and student outcomes of students from Milan Polytechnic for the cohorts before the reform (2008/9, 2009/10, 2010/11 in the "short-term sample" and 2008/9 in the "long-term sample") and after the reform (2011/12 and 2012/13) by student aid status (DSU vs. non-DSU). We further select non mature (i.e. aged 17-25) and Italian students. The DID column shows the differences between DSU and non-DSU students, after vs. before the reform. Admission test scores are standardized by field of study and student cohort (to have zero mean and unit standard deviation).

Table 2. Evidence on the impact of compositional changes in the student cohorts on the student outcomes

Short-term outcomes	Coeff.	S.E.	N.	Longer-term outcomes	Coeff.	S.E.	N.
Credits earned 1st year	0.408	(0.520)	33,274	Enrolled in the 3rd year	-0.002	(0.010)	20,175
Credits earned Aug 10	0.491	(0.521)	33,274	Graduated in 3yrs	0.000	(0.014)	20,175
Credits earned Nov 30	0.365	(0.527)	33,274	Graduated in 3/4 yrs	-0.002	(0.013)	20,175
GPA 1st year	0.056	(0.083)	29,407	Graduation mark 3 yrs	-0.243	(0.375)	6,739
GPA Aug 10	0.055	(0.082)	29,113	Graduation mark 3/4 yrs	-0.160	(0.375)	11,886
GPA Nov 30	0.056	(0.083)	29,410				
Enrolled in the 2nd year	0.011	(0.007)	33,274				
Control variables		Yes				Yes	

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses.

Note. The Table shows the DID estimates (i.e. the $DSU_i * After_i$ coefficient) of specifications in which the dependent variables are the predicted student outcomes from a regression of the observed outcomes on the control variables (gender, age, entry test score, income brackets, residence, field of study). See Carrell et al. (2018) for more details. N. is the number of observations. The number of observations changes across student outcomes as GPA is only available for exams that attribute a grade (so exams assessed as "pass" or "fail" are excluded). Among the longer-term outcomes, graduation marks are conditional on graduation, and the analysis only includes the 2008 pre-reform cohort.

Table 3. DID estimates of the effect of the reform

Short-term outcomes	Coeff.	S.E.	N.	Longer-term outcomes	Coeff.	S.E.	N.
Credits earned 1st year	1.627**	(0.784)	33,274	Enrolled in the 3rd year	0.037	(0.024)	20,175
Credits earned Aug 10	2.887***	(0.798)	33,274	Graduated in 3yrs	0.089**	(0.042)	20,175
Credits earned Nov 30	2.026***	(0.781)	33,274	Graduated in 3/4 yrs	0.070*	(0.036)	20,175
GPA 1st year	0.234	(0.159)	29,407	Graduation mark 3 yrs	-0.649	(0.818)	6,739
GPA Aug 10	0.263	(0.161)	29,113	Graduation mark 3/4 yrs	-0.647	(0.684)	11,886
GPA Nov 30	0.240	(0.159)	29,410				
Enrolled in the 2nd year	-0.020	(0.013)	33,274				
Control variables		Yes				Yes	

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses.

Note. The Table shows the DID estimates (i.e. the $DSU_i * After_i$ coefficient) of equation (1). N. is the number of observations. The number of observations changes across student outcomes as GPA is only available for exams that attribute a grade (so exams assessed as "pass" or "fail" are excluded). Among the longer-term outcomes, Graduation marks are conditional on graduation, and the analysis only includes the 2008 pre-reform cohort. Control variables include gender, age, entry test score, income brackets, residence, field of study.

Table 4. DID estimates of the effect of the reform (with a third-degree polynomial in student entry test scores)

Short-term outcomes	Coeff.	S.E.	N.	Longer-term outcomes	Coeff.	S.E.	N.
Credits earned 1st year	1.758**	(0.773)	33,274	Enrolled in the 3rd year	0.039	(0.024)	20,175
Credits earned Aug 10	2.979***	(0.793)	33,274	Graduated in 3yrs	0.087**	(0.042)	20,175
Credits earned Nov 30	2.158***	(0.770)	33,274	Graduated in 3/4 yrs	0.071**	(0.036)	20,175
GPA 1st year	0.240	(0.158)	29,407	Graduation mark 3 yrs	-0.666	(0.821)	6,739
GPA Aug 10	0.271*	(0.160)	29,113	Graduation mark 3/4 yrs	-0.645	(0.686)	11,886
GPA Nov 30	0.246	(0.158)	29,410				
Enrolled in the 2nd year	-0.016	(0.012)	33,274				
Control variables		Yes				Yes	

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses.

Note. The Table shows the DID estimates (i.e. the $DSU_i * After_i$ coefficient) of equation (1). N. is the number of observations. Unlike in Table 3, the entry test score enters with a third-degree polynomial instead of linearly. The number of observations changes across student outcomes as GPA is only available for exams that attribute a grade (so exams assessed as "pass" or "fail" are excluded). Among the longer-term outcomes, Graduation marks are conditional on graduation, and the analysis only includes the 2008 pre-reform cohort. Control variables include gender, age, entry test score, income brackets, residence, field of study.

Table 5. Sensitivity of DID estimates to changing student composition by family income (i.e. fee brackets)

Short-term outcomes	Fee brackets	Coeff.	S.E.	N.	Longer-term outcomes	Fee brackets	Coeff.	S.E.	N.
Credits earned 1st year	4	1.647*	(0.875)	8,538	Enrolled in the 3rd year	4	0.025	(0.028)	5,140
	5	1.855**	(0.834)	12,354		5	0.039	(0.026)	7,476
	6	1.632**	(0.805)	18,341		6	0.037	(0.025)	11,086
	All	1.627**	(0.784)	33,274		All	0.037	(0.024)	20,175
Credits earned Aug 10	4	2.815***	(0.874)	8,538	Graduated in 3yrs	4	0.075*	(0.043)	5,140
	5	3.009***	(0.838)	12,354		5	0.084**	(0.043)	7,476
	6	2.881***	(0.813)	18,341		6	0.079*	(0.042)	11,086
	All	2.887***	(0.798)	33,274		All	0.089**	(0.042)	20,175
Credits earned Nov 30	4	2.011**	(0.872)	8,538	Graduated in 3/4 yrs	4	0.059	(0.038)	5,140
	5	2.278***	(0.831)	12,354		5	0.072*	(0.037)	7,476
	6	2.026**	(0.802)	18,341		6	0.062*	(0.037)	11,086
	All	2.026***	(0.781)	33,274		All	0.070*	(0.036)	20,175
GPA 1st year	4	0.220	(0.168)	7,606	Graduation mark 3 yrs	4	-1.050	(0.894)	1,523
	5	0.289*	(0.164)	11,038		5	-0.792	(0.860)	2,261
	6	0.270*	(0.161)	16,502		6	-0.865	(0.838)	3,498
	All	0.234	(0.159)	29,407		All	-0.649	(0.818)	6,739
GPA Aug 10	4	0.283*	(0.170)	7,527	Graduation mark 3/4 yrs	4	-0.813	(0.730)	2,859
	5	0.334**	(0.166)	10,921		5	-0.559	(0.709)	4,220
	6	0.311*	(0.163)	16,337		6	-0.735	(0.696)	6,402
	All	0.263	(0.161)	29,113		All	-0.647	(0.684)	11,886

continue

Table 5. (continue)

Short-term outcomes	Fee brackets	Coeff.	S.E.	N.	Longer-term outcomes	Fee brackets	Coeff.	S.E.	N.
GPA Nov 30	4	0.224	(0.168)	7,606					
	5	0.294*	(0.164)	11,039					
	6	0.276*	(0.161)	16,503					
	All	0.240	(0.159)	29,410					
Enrolled in the 2nd year	4	-0.017	(0.015)	8,538					
	5	-0.008	(0.014)	12,354					
	6	-0.016	(0.013)	18,341					
	All	-0.020	(0.013)	33,274					
Control variables	Yes					Yes			

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses.

Note. The Table shows the DID estimates (i.e. the $DSU_i * After_i$ coefficient) of equation (1). N. is the number of observations. The estimates are reported in samples including the first four, five, six and all fee brackets. The number of observations changes across student outcomes as GPA is only available for exams that attribute a grade (so exams assessed as "pass" or "fail" are excluded). Among the longer-term outcomes, Graduation marks are conditional on graduation, and the analysis only includes the 2008 pre-reform cohort. Control variables include gender, age, entry test score, income brackets, residence, field of study.

Table 6. Matching-DID estimates of the effect of the reform

Short-term outcomes	Coeff.	S.E.	N.	Longer-term outcomes	Coeff.	S.E.	N.
Credits earned 1st year	1.349*	(0.785)	33,105	Enrolled in the 3rd year	0.039	(0.025)	20,004
Credits earned Aug 10	2.408***	(0.830)	33,105	Graduated in 3yrs	0.080*	(0.045)	20,004
Credits earned Nov 30	1.642**	(0.782)	33,105	Graduated in 3/4 yrs	0.064*	(0.038)	20,004
GPA 1st year	0.168	(0.183)	29,353	Graduation mark 3 yrs	-0.631	(0.901)	6,498
GPA Aug 10	0.222	(0.185)	29,075	Graduation mark 3/4 yrs	-0.810	(0.795)	11,847
GPA Nov 30	0.170	(0.183)	29,356				
Enrolled in the 2nd year	-0.020*	(0.012)	33,105				
Control variables		Yes				Yes	

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses.

Note. The Table shows the matching-DID estimates (i.e. the $DSU_i * After_i$ coefficient) for equation (1). Matching is performed using kernel matching and estimates are reported on the common support (so the number of observations differs from Table 2). The estimates have been used using the Stata package `diff` (Villa, 2016). N. is the number of observations. The number of observations changes across student outcomes as GPA is only available for exams that attribute a grade (so exams assessed as "pass" or "fail" are excluded). Among the longer-term outcomes, Graduation marks are conditional on graduation, and the analysis only includes the 2008 pre-reform cohort. Control variables include gender, age, entry test score, income brackets, residence, field of study. Matching is based on the control variables except for income brackets (since higher income students do not have good matches with DSU-students, not being eligible for student aid).

Table 7. Heterogeneous effects by student ability (i.e. entry test score) on short-term outcomes

Short-term outcomes	(1) CFU first year	(2) CFUs Aug 10 (1st year)	(3) CFUs Nov 30 (1st year)	(4) GPA 1st year	(5) GPA Aug 10 (1st year)	(6) GPA Nov 30 (1st year)	(7) Enrolled in the 2nd year
DSU	7.126	7.845*	6.431	0.902	0.653	0.894	-0.509***
	(4.533)	(4.232)	(4.517)	(0.799)	(0.807)	(0.800)	(0.042)
DSU * high ability	-7.979***	-7.249***	-7.707***	-0.283	-0.293	-0.262	-0.206***
	(1.479)	(1.701)	(1.473)	(0.267)	(0.281)	(0.267)	(0.016)
DSU * medium ability	-4.281***	-4.500***	-4.151***	-0.740***	-0.669***	-0.735***	-0.125***
	(1.300)	(1.505)	(1.300)	(0.238)	(0.250)	(0.238)	(0.014)
DSU * post (a)	1.074	0.984	1.567	-0.694**	-0.724**	-0.689**	-0.103***
	(2.243)	(2.316)	(2.243)	(0.338)	(0.343)	(0.338)	(0.038)
DSU * high ability * post (b)	-0.139	1.735	-0.500	1.015**	1.222***	1.001**	0.095**
	(2.508)	(2.634)	(2.509)	(0.435)	(0.439)	(0.435)	(0.041)
DSU * medium ability * post (c)	1.899	3.097	1.972	1.082***	1.048**	1.092***	0.113***
	(2.452)	(2.515)	(2.448)	(0.403)	(0.410)	(0.403)	(0.042)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N.	33,274	33,274	33,274	29,407	29,113	29,410	33,274
(a) + (b)	0.934	2.719**	1.067	0.320	0.497*	0.311*	-0.008
p-value	0.430	0.037	0.368	0.246	0.0721	0.0739	0.568
(a) + (c)	2.972***	4.081***	3.539***	0.387*	0.324	0.403	0.010
p-value	0.004	0.000	0.000	0.0862	0.158	0.259	0.613

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses.

Note. DSU students are students receiving aid. Post-reform cohorts (post) are those enrolled after 2010/2011. The number of observations changes across student outcomes as GPA is only available for exams that attribute a grade (so exams assessed as "pass" or "fail" are excluded). Control variables include gender, age, entry test score, income brackets, residence, field of study. Lower-ability, medium-ability and higher-ability students are those belonging to the 1st, 2nd -3rd and 4th quartiles of the entry-test score distribution, respectively. N. is the number of observations.

Table 8. Heterogeneous effects by student ability (i.e. entry test score) on longer-term outcomes

	(1)	(2)	(3)	(4)	(5)
Longer-term outcomes	Enrolled in the 3rd year	Graduated in 3 yrs	Graduated in 3/4 yrs	Graduation mark 3 yrs	Graduation mark 3/4 yrs
DSU	0.156	-0.369***	-0.013	2.587*	-0.313
	(0.178)	(0.103)	(0.176)	(1.464)	(1.473)
DSU * high ability	-0.153***	0.074	-0.219**	-2.048	0.217
	(0.045)	(0.095)	(0.085)	(1.729)	(1.441)
DSU * medium ability	-0.155***	0.085	-0.141*	0.162	1.557
	(0.050)	(0.085)	(0.079)	(1.725)	(1.342)
DSU * post (a)	-0.065	0.084	-0.098	-4.168*	-1.764
	(0.059)	(0.096)	(0.095)	(2.237)	(1.612)
DSU * high ability * post (b)	0.066	0.000	0.206*	4.811*	1.845
	(0.065)	(0.120)	(0.111)	(2.517)	(1.959)
DSU * medium ability * post (c)	0.168**	0.011	0.206*	3.330	0.930
	(0.069)	(0.112)	(0.107)	(2.524)	(1.872)
Control variables	Yes	Yes	Yes	Yes	Yes
N.	20,175	20,175	20,175	6,739	11,886
(a) + (b)	0.001	0.084	0.109**	0.644	0.081
p-value	0.979	0.249	0.031	0.585	0.943
(a) + (c)	0.103***	0.095	0.109*	-0.837	-0.834
p-value	0.006	0.110	0.063	0.481	0.393

*** p<0.01, ** p<0.05, * p<0.1

Note. DSU students are students receiving aid. Post-reform cohorts (post) are those enrolled after 2010/2011. The number of observations changes across student outcomes as GPA is only available for exams that attribute a grade (so exams assessed as "pass" or "fail" are excluded). Among the longer-term outcomes, graduation marks are conditional on graduation, and the analysis only includes the 2008 pre-reform cohort. Control variables include gender, age, entry test score, income brackets, residence, field of study. N. is the number of observations.

Online Appendix to

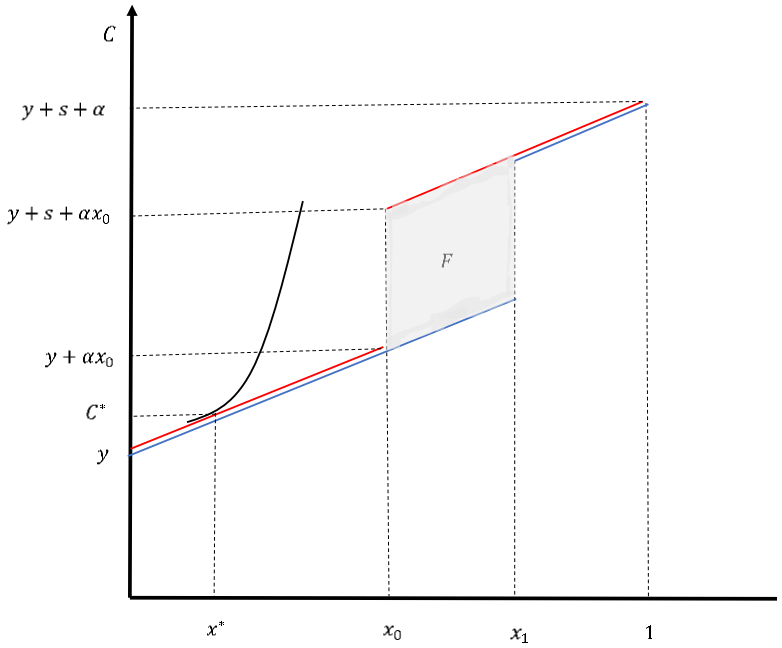
"When the need meets merit: The effect of increasing merit requirements in need-based student aid"

by Tommaso Agasisti, Massimiliano Bratti and Veronica Minaya

APPENDIX A. Additional tables and figures

Figure A1. Merit based student aid, student effort and academic performance: the effect of increasing the merit-based requirements

(a) Lower ability student: no effect



(b) Lower ability student increases performance but loses the scholarship

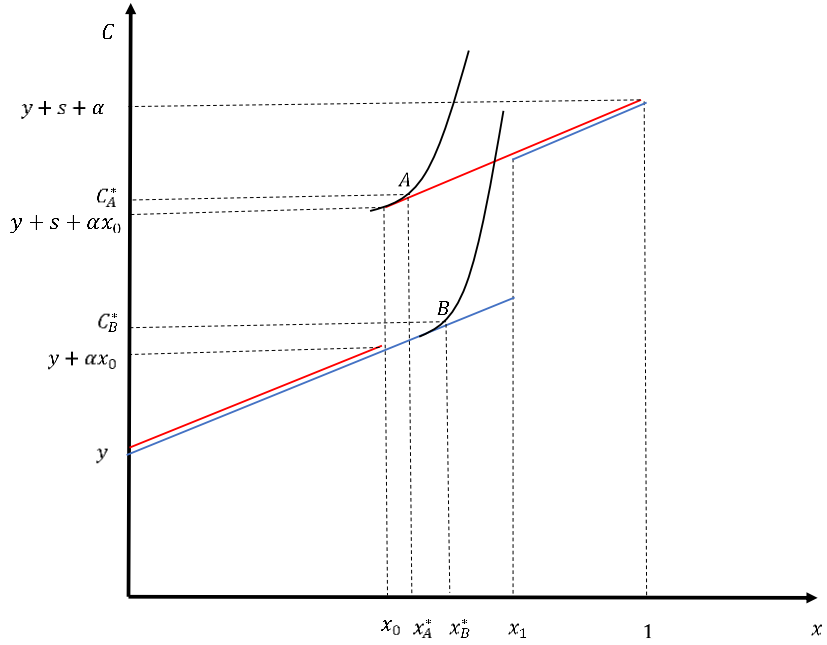
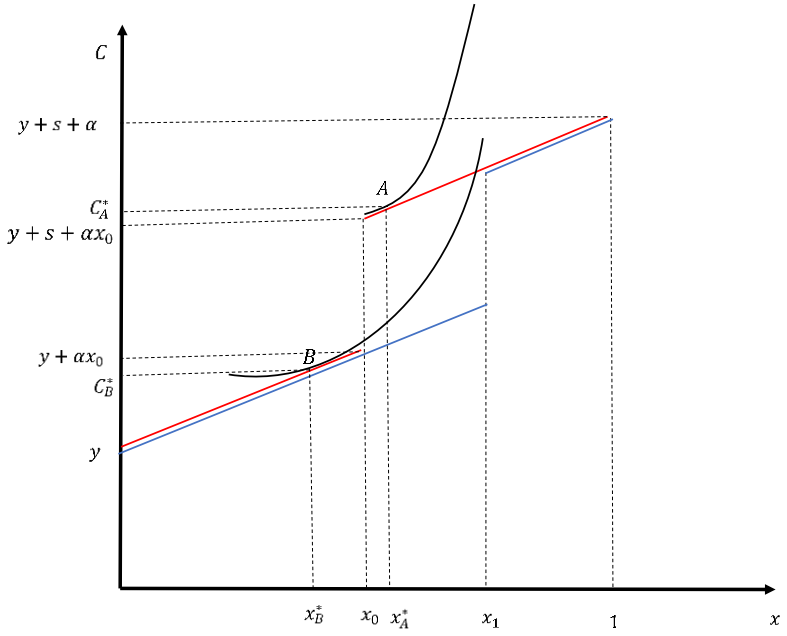


Figure A1. continue

(c) Lower ability student decreases performance and loses the scholarship



(d) Higher ability student: no effect

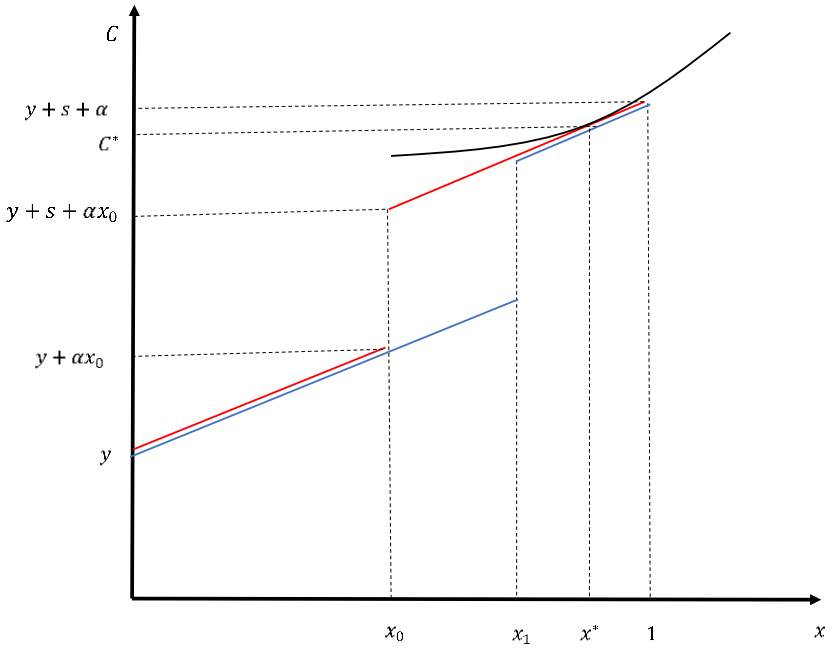
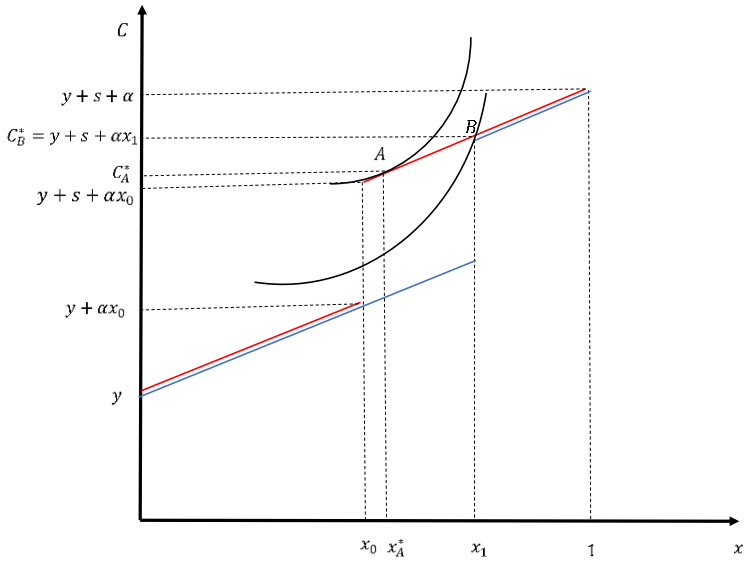


Figure A1. continue

(e) Intermediate-ability students increase student performance and maintain student aid



Note. Consumption (C) is reported on the vertical axis and academic performance (equal to student effort $x \in [0,1]$ by assumption) on the horizontal axis. x_A^* and x_B^* are the pre-reform and the post-reform levels of performance and C_A^* and C_B^*, the corresponding levels of consumption. Students receive a scholarship equal to s if their academic performance is above the merit cutoff, x_0 and x_1 in the pre- and post-reform periods, respectively. y is the individual exogenous income (i.e. parental transfers) and αx is labour income (proportional to academic performance). Utility depends on the choice variables, with flatter indifference curves for high ability students.

Table A1. Effects of the reform by gender --- Short-term outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Short-term outcomes	CFU first year	CFU Aug 10 (1st year)	CFU Nov 30 (1st year)	GPA 1st year	GPA Aug 10 (1st year)	GPA Nov 30 (1st year)	Enrolled in the 2nd year
DSU	5.083	5.473	4.551	0.408	0.164	0.408	-0.592***
	(5.021)	(4.472)	(5.000)	(0.810)	(0.817)	(0.810)	(0.048)
DSU * post (a)	1.912**	2.758***	2.325***	0.206	0.237	0.211	-0.019
	(0.905)	(0.930)	(0.902)	(0.197)	(0.197)	(0.197)	(0.014)
DSU * female	-1.691*	-2.361**	-1.755*	-0.257	-0.262	-0.264	-0.024
	(1.028)	(1.097)	(1.013)	(0.180)	(0.189)	(0.180)	(0.016)
DSU * female * post (b)	-1.412	-0.091	-1.478	0.042	0.033	0.041	-0.009
	(1.687)	(1.701)	(1.686)	(0.323)	(0.330)	(0.323)	(0.028)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N.	33,274	33,274	33,274	29,407	29,113	29,410	33,274
(a) + (b)	0.500	2.668***	0.847	0.248	0.269	0.253	-0.028
p-value	0.733	0.068	0.563	0.341	0.317	0.331	0.266

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses.

Note. DSU students are students receiving aid. Post-reform cohorts (post) are those enrolled after 2010/2011. The number of observations changes across student outcomes as GPA is only available for exams that attribute a grade (so exams assessed as "pass" or "fail" are excluded). Control variables include: gender, age, entry test score, income brackets, residence, field of study. N. is the number of observations.

Table A2. Effects of the reform by gender --- Longer-term outcomes

	(1)	(2)	(3)	(4)	(5)
Longer-term outcomes	Enrolled in 3rd year	Graduated in 3 years	Graduated in 3/4 years	Final mark 3 yrs	Final mark 3/4 yrs
DSU	0.086	-0.302***	-0.132	2.160**	0.514
	(0.175)	(0.077)	(0.162)	(0.951)	(0.948)
DSU * post (a)	0.055*	0.071	0.112**	-0.777	-1.371
	(0.032)	(0.051)	(0.047)	(1.113)	(0.897)
DSU * female	-0.019	-0.059	0.025	-0.598	-1.764*
	(0.038)	(0.068)	(0.055)	(1.205)	(1.011)
DSU * female * post (b)	-0.070	0.046	-0.143**	0.269	1.796
	(0.046)	(0.086)	(0.070)	(1.563)	(1.329)
Control variables	Yes	Yes	Yes	Yes	Yes
Observations	20,175	20,175	20,175	6,739	11,886
(a) + (b)	-0.015	0.117*	-0.031	-0.508	0.424
p-value	0.674	0.095	0.552	0.651	0.672

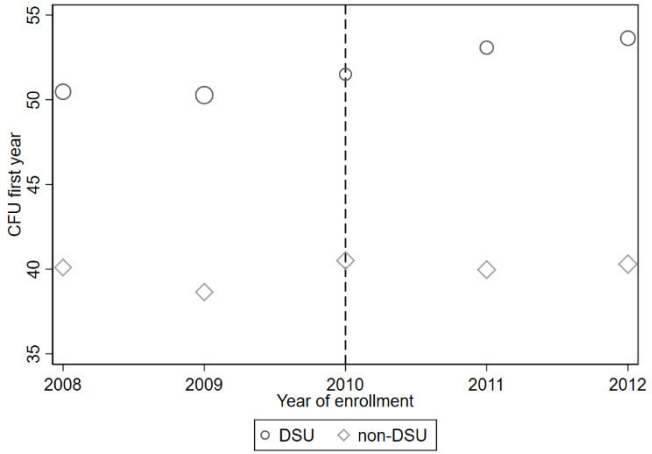
*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses.

Note. DSU students are students receiving aid. Post-reform cohorts (post) are those enrolled after 2010/2011. The number of observations changes across student outcomes as GPA is only available for exams that attribute a grade (so exams assessed as "pass" or "fail" are excluded). Control variables include: gender, age, entry test score, income brackets, residence, field of study. N. is the number of observations. In order to avoid treatment contamination only the 2008 pre-reform cohort is retained in the estimation.

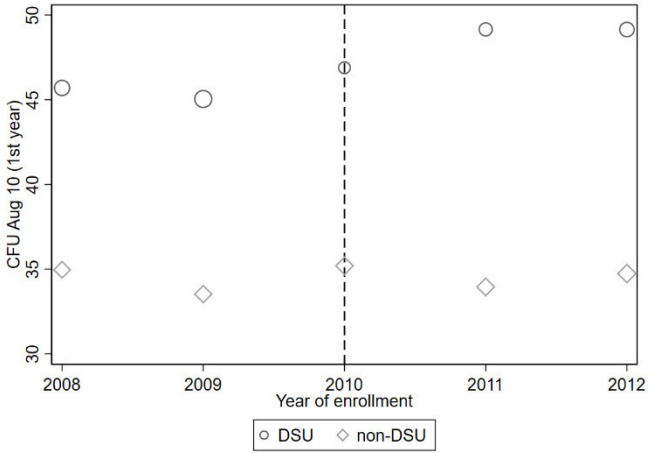
APPENDIX B. Graphs on parallel trend assumption

Figure B1. Parallel trend assumption – Short-term outcomes

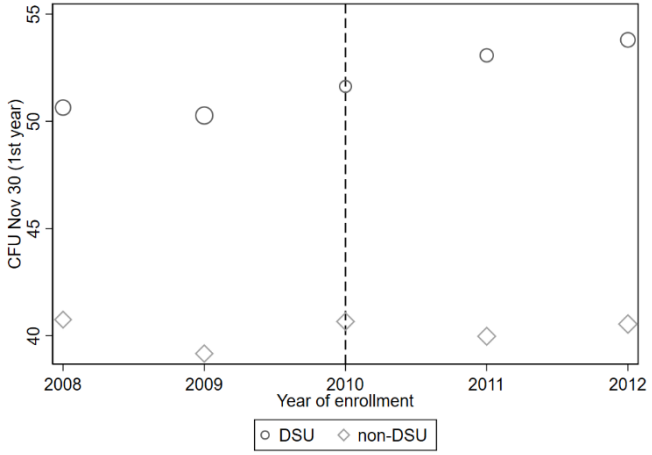
(a) Credits (CFUs) earned in the 1st year



(b) Credits (CFUs) earned by 10th August 1st year



(c) Credits attempted by 30th November 1st year



(d) GPA in the 1st year

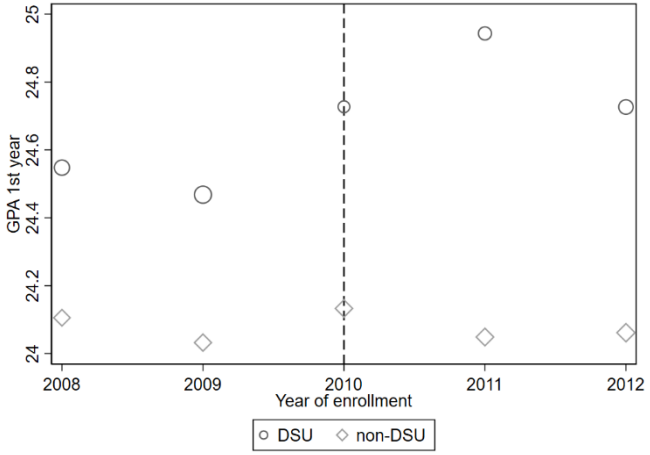
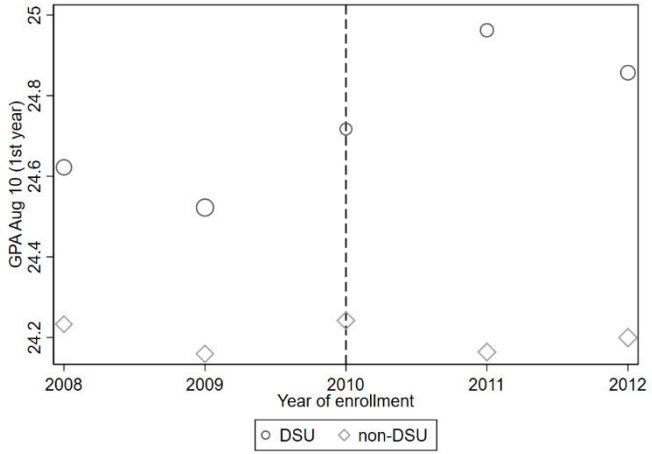
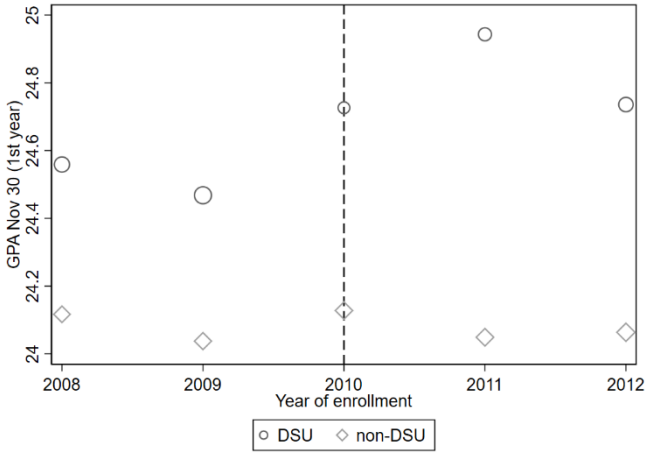


Figure B1. continue

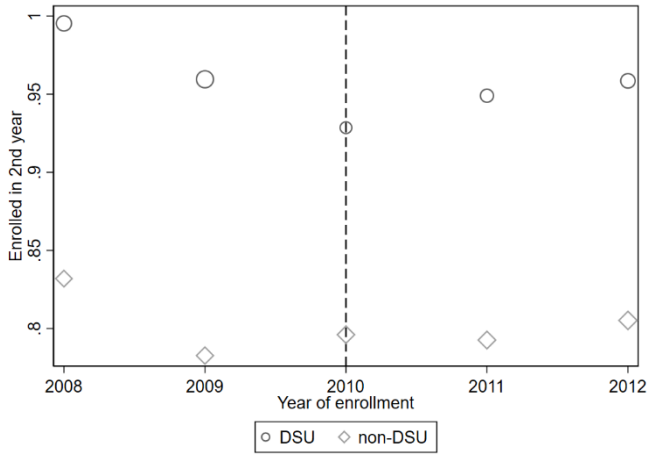
(e) GPA at 10th August 1st year



(f) GPA at 30th November 1st year



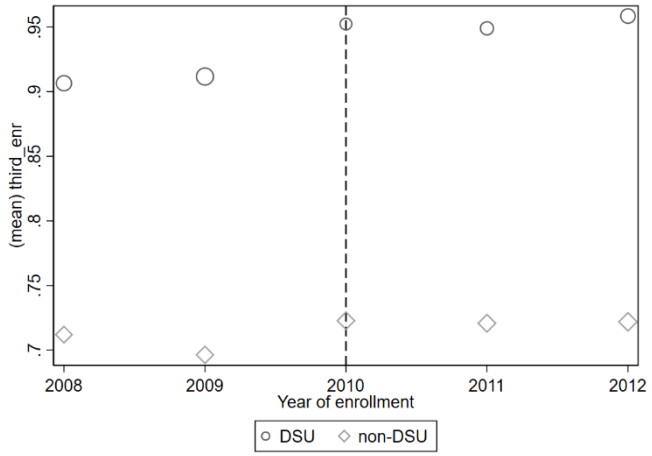
(g) 2nd year enrolment



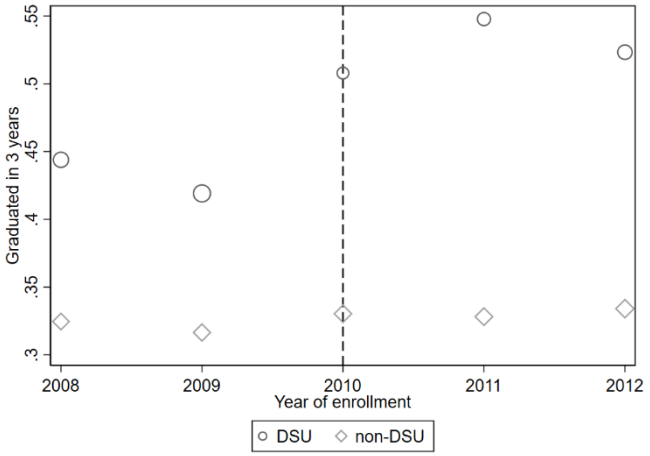
Note. The dashed vertical line indicates the last pre-reform cohort. Plots are computed on PoliMI data including Italian students aged 19-25 at entry.

Figure B2. Parallel trend assumption - Longer term outcomes

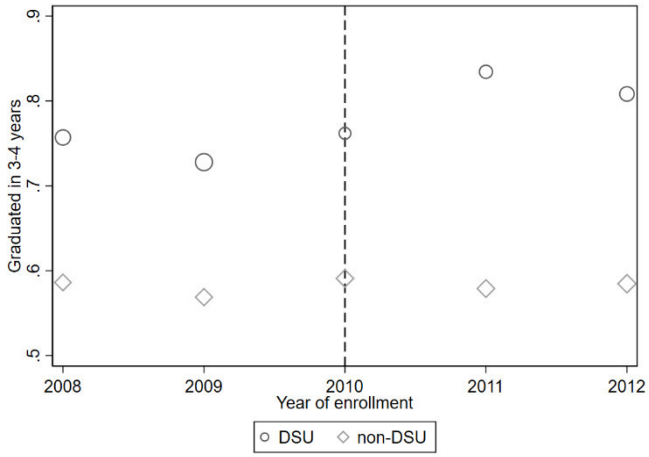
(a) 3rd year enrollment



(b) Probability of graduation in 3 years



(c) Probability of graduation in 3



(d) Final graduation mark, 3 or 4 years

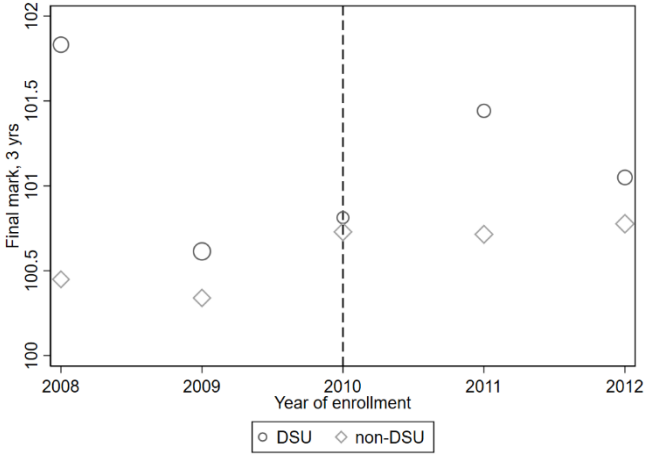
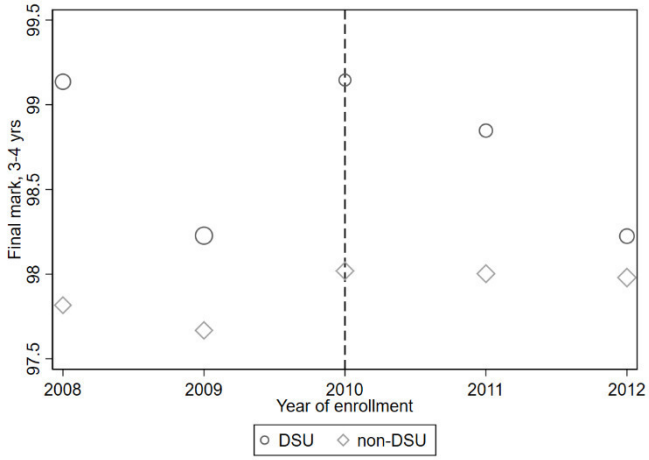


Figure B2. Parallel trend assumption - Longer term outcomes

(e) Final graduation mark, 3 or 4 years



Note. The dashed vertical line indicates the last pre-reform cohort. Plots are computed on PoliMI data including Italian students aged 19-25 at entry.