On the Design of Grant Assignment Rules[*]

Francesca Modena[†1], Santiago Pereda-Fernández[‡1], and Giulia Martina Tanzi[§1]

[1]Banca d'Italia

August 24, 2020

Abstract

The effect of grants on several academic outcomes has been widely studied, finding that they reduce the probability of dropping out of college. In this paper we assess how different rules affect dropout rates in the first year of enrollment, as well as the characteristics of the recipients of the grants and the dropouts. The analysis uses administrative data from all Italian universities in the period 2003-13. We find that awarding the grant to all the eligible students would significantly increase public expenses with only a slightly reduction in the dropout rate. It is therefore important to target the policy, also because we find that the grants have a heterogeneous effect according to students' characteristics. Targeting high-performing students minimizes dropout rates amongst recipients, even if this rule would slightly increase the overall dropout since it excludes students with a higher ex-ante dropout probability. On the other hand, an assignment that targets those who benefit the most achieves the maximum reduction in dropout rates at the cost of increasing the number of grant recipients who drop-out.

Keywords: Grants, Treatment Assignment

JEL classification: C25, I21, I22

[*]All remaining errors are our own. The views presented in this paper do not necessarily reflect those of the Banca d'Italia.

[†]Banca d'Italia, Piazza Alessandro Vittoria 6, 38122 Trento, Italy. Email: francesca.modena@bancaditalia.it

[‡]Corresponding author. Banca d'Italia, Via Nazionale 91, 00184 Roma, Italy. I can be reached via email at santiago.pereda@.bancaditalia.it.

[§]Banca d'Italia, Milano, Italy. Email: giulia-martina.tanzi@bancaditalia.it

1 Introduction

Students in higher education incur into substantial costs, including fees, housing or the opportunity cost of not working for several years. Grants constitute a way to mitigate these costs, inducing more students to enroll at university and completing a degree. They are particularly relevant for students whose families are not wealthy, as they are often credit constrained and are forced to abandon their studies before they obtain a university degree (Stinebrickner and Stinebrickner, 2008; Deming and Dynarski, 2009). Moreover, they have also been used to induce students to put more effort and to make sure that enrolled students are those who can make the most out of the education. Consequently, grants have typically been assigned on financial grounds, aimed at students with low socio-economic status (*e.g.* Fack and Grenet, 2015), or merit, aimed at high-performing students (*e.g.* Schudde and Scott-Clayton, 2016).

Funds available for grants are limited. Furthermore, different assignment rules target different students. Consequently, it is crucial to identify the effect that awarding a grant has on students' performance, as each assignment rule will have a different impact on dropout rates for the population as a whole, for grant recipients, and for different demographic groups. A policy maker should take all these aspects into consideration to appropriately design the assignment rule.

In this setting, this paper makes the following contributions. First, we study the effect of need-based grants on student's performance in Italy, focusing in particular on the drop out rate. Using administrative data, we estimate the causal effect of grants, comparing students' predicted performance depending on whether they are assigned a grant or not. Using this result, we can compute the effect of an increase in the number of grants awarded by the government on students' dropout rate.

Second, we execute some counterfactual exercises through which we can asses, for a given level of government's expenditure, the impact of different grant assignment rules on the performance of the whole population of students and of the different subpopulations of grant recipients.

The assignment of need-based grants to first year students does not depend on students' past performance. After they are enrolled, students are ranked according to two indices that reflect income and assets of their family. Those below a certain threshold can apply for a grant. When funds are not sufficient to cover all eligible students, grants are awarded according to the economic condition within each university. This constitutes the basis for identification, which allows us to compare the outcomes of grant recipients to those of who were eligible but were not awarded a grant. Most importantly, we focus on the probability that they do not drop out during their first year.

The potential outcomes framework is the most natural choice to evaluate the effect of a grant on the probability of dropping out. The population of students can be split into three types: those who would graduate or enroll the following year without a grant, those who would drop out even with a grant, and those who would graduate or enroll the following year only if they are awarded a grant. The latter conforms the group that benefits from the grant, so awarding grants to these students would be an effective use of funds. On the other hand, the first two types represent two different kinds of ineffective uses of grants funds, since recipients' outcomes are not affected by the grant.

Although it is not possible to determine whether a student would benefit from the grant, it is possible to estimate the probability that they do not drop out when they are a grant recipient and when they are not. The difference between these two probabilities represents the expected increase in the probability of enrolling the year after or graduating due to the grant. This information allows policy makers to weigh in the efficiency of the grant against making grants available for students with little financial means.

The effects of student financial aid on different outcomes have been thoroughly studied in the literature.[1] Many works have studied the effect of a grant on the probability of dropping out (Singell, 2004; Bettinger, 2004; Bettinger et al., 2012; Mealli and Rampichini, 2012; Castleman and Long, 2016; Denning, 2019). Moreover, several other papers have estimated the effect of obtaining a grant on other outcomes, such as enrollment (Lauer, 2002; Kane,

---

[1]See e.g., Dynarski and Scott-Clayton (2013) for a review on how such programs work.

2003; Baumgartner and Steiner, 2006; Cornwell et al., 2006; Goodman, 2008; Deming and Dynarski, 2009; Nielsen et al., 2010; Steiner and Wrohlich, 2012; Vergolini and Zanini, 2015), grades (Cappelli and Won, 2016), or time to finish the university degree (Glocker, 2011; Garibaldi et al., 2012; Denning et al., 2019). However, none of these papers have addressed the question of how to assign grants appropriately. As far as we know, we are the first to do it by considering the effect of different assignment rules on dropout rates for the overall population and specific subpopulations, as well as other measures that a policy maker could be interested in. Exploring the effects of alternative grant assignment rules can provide interesting insights for understanding the potential role of government intervention. The only similar work is Hanushek et al. (2014) who developed a dynamic general equilibrium model that explores the impact of alternative ways to subsidize higher education (need based aid, merit based aid, loans), from both an equity and an efficiency perspective.

In this paper we focus on the Italian case, which is a particularly relevant case study because the country has the lowest percentage of university graduates among the European Union countries, due to both a low enrollment rate and to high dropout rates. Also the availability of public grants is very limited compared to neighboring countries. Thus, understanding the most appropriate mechanism to award grants turns out to be crucial.

Previous works undertaken in the Italian context found that students' performance and completion rates are strongly and positively affected by grants. Modena et al. (2020) found that around one third of the low income students would have left university in the first year in absence of the grants. Other works (Mealli and Rampichini, 2012; Sneyers et al., 2016) also found positive strong effects of the grants, but they relied on small samples of students in selected universities and academic years.

In contrast, we measure the impact of need-based aid by using student-level administrative data over the period 2003-2013 that cover the entire population of Italian university students. The data follow students from their enrollment to graduation/dropout and provide several items of information on their academic career and educational background. The use of administrative data over a very long time span constitutes a major advantage of this paper

4

with respect to previous works. We focus on the first year of university. Significant numbers of dropouts occur during the first year of study (Zotti (2015); Gitto et al. (2015); Mealli and Rampichini (2012)):

The results obtained are crucial from a policy point of view. First, we found that an assignment of the grants to all the eligible students would reduce the drop out rate by only 0.7 percentage points relative to the case with the observed number of awarded grants. This type of policy interventions, however, would almost double the number of awarded students, strongly increase the required funds. Moreover, we found that the effect on the probability of dropping out is very heterogeneous across population groups: the reduction in the drop out rate related to the grant is stronger for males, foreign students, those students that are residents in the South of Italy and for those with a relatively low high school score.

Then, we moved to the theoretical counterfactual exercises in which we assess the effect of different theoretical assignment rules, for a given level of grants. These can be split into roughly two categories, those based on efficiency and those based on merit. The first type targets students who would be more likely to benefit from the grant, those with an ex ante higher probability of dropout. Applying this rule allows the policy maker to achieve the maximum reduction of dropout rates due to the grants, although a large number of grant recipients would drop out. On the other hand, the rules based on merit target students who would be more likely to enroll the following year or graduate without the grant. These would minimize dropout rates amongst grant recipients, but most of them would not have dropped out even without the grant. This highlights a tradeoff in grant assignment: decreasing overall dropout rates tends to increase the number of recipients who drop out. Finally, setting quotas affects the population of grant recipients but have a more modest effect on dropout rates. This counterfactual exercise is particularly useful for understanding the differences in the effectiveness of different types of assignment rules. On the one hand, the policy maker may be interested in maximizing the efficiency of the rule in terms of achieving an higher number of graduates. On the other hand, he may be interested in maximizing only the number of graduate students who deserve it, because of their ex ante level of commitment. Of course,

5

the implementation of such rules may not be feasible from a legal standpoint, as the rules could be targeted to individuals with certain characteristics, such as gender, birthplace or high school grades.

The rest of the paper is organized as follows. The institutional details of the Italian system or higher education and the data used in this paper are respectively described in Sections 2 and 3. The empirical strategy is discussed in Section 4, whereas the results of the effect on grant on the probability of obtaining a college degree are presented in Section 5. On the other hand, Section 6 compares the outcomes of the different assignment rules. Finally, Section 7 concludes and enumerates the following steps to be taken.

2 Institutional Details

Students in the Italian university system can apply for a grant during their first year. The eligibility criterion is based exclusively on the student's family economic condition[2]. Applicants are ranked according to two indices, the Equivalent Financial Situation Indicator (ISEE) and Equivalent Asset Situation Indicator (ISPE), which depend on family's yearly income and assets, respectively. A maximum threshold for these two indices is set at the national level, which guarantees that only students from low income and low assets families are eligible. Because the eligibility threshold is low, students are comparable in terms of financial conditions regardless of whether the eligible student's indices lie close or far from the thresholds.[3] Note also that the eligibility threshold varies slightly over time and between universities: it is updated every year on the basis of the general ISTAT index of consumer prices and the national law establishes a range within which each region can slightly vary the threshold. Moreover, it tends to be lower in the South, which is the poorest area of Italy.

---

[2]The second payment of the grant is conditional on the achievement of a minimum level of credits. This number is set by the regions after consulting the universities, up to a maximum of 20 credits (Prime Ministerial Decree, April 9, 2001).

[3]*E.g.*, in 2008 the ISEE cutoff for eligibility was around €19,000. For a household with both spouses and one child, with zero assets, this is equivalent to a after-tax yearly income as large as €27,000, which in turn is approximately equivalent to 77% of the average Italian yearly income that year for a household of that type.

The application for a grant is submitted to the regional agency where the university is located after the students are enrolled, and notice of acceptance is communicated a few months after enrollment. However, funds are not sufficient to award a grant to every eligible student. Within each university, eligible students are awarded a grant according to the time of application, until the available funds are exhausted. Hence, in some universities in certain years some eligible students are not awarded a grant. Consequently, we define the treatment group as those students who were awarded a grant, and the control group as those who were eligible but were not grant recipients.

The amount of the grant depends on whether students are resident in the city where the university is located, whether they are daily commuters or out-of-site students. Every year the Ministry of Education sets the minimum amount for a grant, but the differences over time are very small. For example, in 2013 the minimum yearly amounts for the three categories of students were, respectively, €1,904, €2,785 and €5,053; the average amount was about €3,035.[4] Even if not all the eligible students are awarded the grant, these students are all exempted from the payment of tuition fees, which on average equal about €1000.

3 Data

We exploit the Anagrafe Nazionale Studenti (ANS), a dataset that contains administrative records on enrollments, students' school background and their academic careers in Italian universities. The main advantage of our database is that it covers the entire population of university students in Italy over a long spell of time. We focus on students aged between 18 and 20 who enrolled at an Italian university for the first time during the period 2003-13. This avoids comparability problems between students who started university immediately after completing high school and those who did it at a later stage.

The database contains a grant eligibility indicator and another for actually obtaining it. This allows us to generate the working sample: first-year grant-recipients conform the

---

[4]Source: *Osservatorio Regionale per l'Università e il Diritto allo studio universitario del Piemonte.*

treatment group, while eligible students that were not awarded the grant belong to the control group.[5] Unfortunately, we do not have access to neither the ISEE nor the ISPE. The sample size equals an average of 31,000 students per year.

Given that the grant eligibility depends on students' previous performance, the empirical strategy is appropriate to study the effect on first-year outcomes. The most important one is whether students dropped-out or not, although we also consider the average grades at the end of the first year and the number of passed credits. Moreover, because arguably an important goal of the policy maker is to increase the number of students with a degree, we also consider whether they finished in time.

Some descriptive statistics of the sample are shown in Table 1. We define dropout students as those enrolled as first year students in the academic year t who did not enroll at any university in the following academic year $t+1$ (ANVUR, 2016; Modena et al., 2020). The dropout rate was, on average, close to 7.5%, with a downward trend. Grant recipients represented slightly over 50% of all eligible students, and they had a higher chance of successfully finishing the first year and of graduating in time. Moreover, students with a grant passed on average 2 extra credits during the first year (the mean value was equal to 34.7). In contrast, the average score in the first year was about the same regardless of being awarded a grant.

Almost two thirds of them are female, and only a tiny fraction has foreign origin. The majority of them come from the South of Italy and they tend to study in the region where they are resident, although there is a persistent positive flow of students from the South to the Center-North of Italy (De Angelis et al., 2017). Over a half of them accessed university with a high school degree, and almost a third with a technical school degree. Finally, less than a half of them come from a urban area. Grant assignment was balanced across genders, given the percentages of male and female students. However, foreign students, residents in either then North or the Center of Italy, studying in a different area, with a relatively low score, and from a rural area received a grant disproportionally often.[6]

---

[5]For completeness, we provide some analysis on the non-eligible students in Appendix B.2
[6]Italy is split into four different areas: North-West, North-East, Center, and South & Isles.

Table 1: Descriptive statistics

	Total	Treated	Untreated
Continued into second year	92.6	93.0	92.1
First year average score	25.1	25.0	25.2
Number of credits	34.7	35.7	33.6
Graduated in time	36.1	39.2	33.0
Female	64.2	64.6	63.9
Male	35.8	35.4	36.1
Foreign	2.4	3.3	1.5
Native	97.6	96.7	98.5
North	25.6	33.6	17.3
Center	12.6	15.3	9.8
South	61.8	51.0	72.8
Dif reg	14.1	18.9	9.2
Same reg	85.9	81.1	90.8
HS deg	58.6	56.2	61.0
TS deg	28.4	29.6	27.1
Ot deg	13.1	14.2	11.9
Score 1	13.4	15.0	11.7
Score 2	21.3	23.3	19.3
Score 3	23.3	24.7	21.8
Score 4	20.2	19.6	20.8
Score 5	21.8	17.4	26.4
Urban	42.6	39.6	45.6
Rural	57.4	60.4	54.4
N	340205	173070	167135

Notes: North, Center and South denote the region of residence of the student; Different region (Same region) denotes that student region of residence and the region of the university are different (the same); HS degree, TS degree, and Other degree respectively denote the possession of a high school degree, a technical degree and other type of degree; Scores 1-5 indicate that the average high school score was in each of the 5 following intervals: 0-69, 70-79, 80-89, 90-99, 100.

4 Empirical Strategy

Let $Y_{it\tau}$ be a dummy variable that takes value 1 when student i successfully completes the τ-th year of a university degree at time t and 0 if the student drops out, and $D_{it\tau}$ be another dummy variable that equals 1 when student i is awarded an grant. This relation can be modeled as

$$Y_{it\tau} = q\left(D_{it\tau}, X_{it\tau}, \epsilon_{it\tau}\right) \tag{1}$$

where $X_{it\tau}$ is the vector of observed characteristics, including time and university effects, and $\epsilon_{it\tau}$ be an idiosyncratic disturbance term. The assignment of grants outlined in Section 3 points at grants rationing as the key source of variation to achieve identification. Our methodology therefore consists in comparing, grant beneficiaries (the treated group) with eligible non-beneficiaries (the control group).

The estimation presents the challenge of isolating the effect of the grant from other factors that affect college success. In particular, family financial conditions determine the access to aid and may also be directly associated with student outcomes. In our setting beneficiaries and eligible students had very similar family characteristics: eligibility required not to exceed certain family's yearly income and assets thresholds. Moreover, grants rationing varies across years and universities. These issues are addressed by controlling for a set of student, household and university characteristics that could be correlated to the ISEE and student's performance. Namely, we include gender, nationality, area of residence, a dummy for studying in a macro area different from the area of residence, high school type and grade, and a dummy for the local urban labor system of residence (Adamopoulou and Tanzi, 2017; Di Pietro, 2004).

Note that timing of the grant assignment is not a concern for the selection of students into enrolling, since the assignment is not known beforehand, both grant recipients and non-recipients decide to enroll regardless of whether they will be awarded the grant or not. Also, another potential cause for endogeneity would be if grant assignment was merit-based.

However, grant-assignment for first-year students is exclusively based on their households' financial situation. Thus, beneficiaries should not be ex-ante different in terms of a student's merit and abilities.

Finally, although in principle there would be room for students to strategically self-select into universities with higher coverage ratio, in practice this is precluded because these ratios are not public information due to the delayed notice of acceptance. Students' strategic behavior could be based on past information, but the coverage rates vary widely over time because they depend on the availability of public funds and on political choices. Moreover, in our setting this selection would have been a concern only if beneficiaries and eligible students had had a different set of information about coverage rates, *i.e.* if students' strategic behavior had been correlated with household income and financial wealth. Note, however, that even if our strategy is not enough to net out the differences between the two groups with respect to financial conditions, the resulting estimate is likely to be biased towards zero, *i.e.*, against finding an impact of grant college retention.

Formally, our identification is based on selection on observables, *i.e.* $Y_{it1}(1), Y_{it1}(0) \perp D_{it1}|X_{it\tau}$, where $Y_{it\tau}(d)$ denotes the potential outcome when $D_{it\tau} = d$. In this setting, it is possible to estimate the effect of interest with different estimators, each of which models the relation between the outcome variable and the treatment in a different way.

For example, if the observed variables enter linearly in Equation 1, one could estimate the model by OLS. A more flexible alternative would be blocking with regression (Imbens, 2015), which splits the population according to their value of the propensity score and runs a linear regression within each cell. Moreover, if the dependent variable is binary it would be possible to model Equation 1 with a latent variable and use limited dependent variable methods such as logit.[7] On the other hand, if the outcome is continuous one could use distribution regression with different link functions. Rather than arguing which model is the

---

[7]Note that in some universities all eligible students were awarded a grant. For the binary choice estimators with university fixed effects this implies that we run into the perfect separation problem (Albert and Anderson, 1984), so there is no well defined MLE. Rather than estimating the model with a penalization (e.g. as in Firth, 1993), we constrain the maximum (absolute) value that the coefficients can take to 10. This creates an attenuation bias for the coefficients of these estimators, which is small in terms of the predicted probabilities. See Appendix B.1 for further details.

most appropriate, we consider several of them and we use 10-fold cross validation, choosing the estimator that minimizes the following Mean Squared Error (MSE) function:

$$MSE = \sum_{t=1}^{T} \sum_{i=1}^{N_t} \left(y_{it\tau} - \hat{y}_{it\tau} \left(d_{it\tau}, x_{it\tau}\right)\right)^2 \tag{2}$$

where $\hat{y}_{it\tau} \left(d_{it\tau}, x_{it\tau}\right)$ is the estimated probability that student i does not drop out, given the observables $(d_{it\tau}, x'_{it\tau})'$.

We consider different combinations of time and university dummies, as well as interactions between the grant and individual characteristics. For the selection of covariates, we run a separate 10-fold cross validation Lasso procedure that is explained in Appendix A. This model and covariates selection mechanism presents some advantages. First, it can be used to compare methods that are based on different objective functions, such as OLS or MLE. Second, it can be easily adapted to outcome variables that are not binary, as we do for some secondary outcomes. Third, for some of the estimation methods, the selection of covariates is based on both the outcome and the treatment equation, making it more robust against misspecification (Belloni et al., 2014b).

Define the probability that student i successfully completes the academic year τ at time t as $P_{it\tau}\left(d_{it\tau}\right) \equiv \mathbb{P}\left(Y_{it\tau} = 1 | D = d_{it\tau}, X = x_{it\tau}\right)$, for $d_{it\tau} = 0, 1$. The following three quantities are relevant for the assessment of the assignment of grants by the policy maker:

$$P_{it\tau}\left(0\right) \tag{3}$$

$$P_{it\tau}\left(1\right) \tag{4}$$

$$\Delta P_{it\tau} \equiv P_{it\tau}\left(1\right) - P_{it\tau}\left(0\right) \tag{5}$$

Equation 3 denotes the probability that student i successfully completes year τ without a grant. This quantity also represents the probability that awarding the grant on student i would be wasteful, since the student would not need it to continue to the next year. We refer to this as an ineffective grant of type I. Similarly, Equation 4 denotes the probability

that student i successfully completes year τ with a grant. As a result, awarding the grant to student i would be also wasteful with probability $1 - P_{it\tau}(1)$, which is the probability that student i would dropout from university regardless of whether he is assigned a grant or not. We refer to this as an ineffective grant of type II. Finally, the difference between these two probabilities (Equation 5) captures the probability that student i completes year τ only if he is awarded a grant. We refer to this as an effective grant.

Knowledge of these probabilities allows the policy maker to allocate grants according to some rules that maximize some social welfare measure. For example, the policy maker could be interested in maximizing the absolute number of successful graduates, a measure that captures the efficiency of the allocation mechanism. On the other hand, the policy maker could be concerned with inequality of outcomes across different subpopulations, such as socio-economic status or gender.

Consequently, the assignment rule can discriminate based on merit or belonging to some subpopulation. In this analysis, **after considering a baseline simulation when grants are allocated to all eligible students**, we consider the following assignment rules:

1. Effectiveness-based assignment: grants are awarded to students with the highest decrease in the probability of dropping off when they receive a grant (highest probability of the grant being effective).

2. Merit-based assignment: grants are awarded to students with the lowest probability of dropping when they receive a grant.

3. Effectiveness-based assignment with regional quotas: each region awards the same number of grants as in the actual data; grants are awarded according to assignment rule (1) within regions.

4. Merit-based assignment with regional quotas: each region awards the same number of grants as in the actual data; grants are awarded according to assignment rule (2) within regions.

5. Effectiveness-based assignment with gender quotas: half of the grants are awarded to students of each gender according to assignment rule (1).

6. Merit-based assignment with gender quotas: half of the grants are awarded to students of each gender according to assignment rule (2).

7. Score-based assignment: grants are awarded to students with the highest high school scores.

These rules can be classified into two main categories: those that target students who are expected to benefit the most from the grant[8], and those that target high-performing students, either focusing on past performance, or predicted future performance. Moreover, we also consider variations that first set some quotas based on regions or gender. The former follows from the fact that grants are administered at the regional level, and regional agencies could be reluctant to transfer money to students who study at universities in other regions. The latter could be relevant for the equality of opportunity debate, as some grants and scholarships are targeted for specific subpopulations, often those of a particular gender.

5   Results

Table 2 reports the cross validated MSE of all estimators for the outcomes considered. All binary choice estimators have a slightly smaller MSE than the OLS and blocking with regression estimators for both binary outcomes. Specifically, the logit that interacts the treatment with the selected covariates and includes university and period effects (specification 6) minimizes the MSE for the probability of continuing into the second year, whereas using that set of covariates with the probit (specification 8) minimizes the MSE for the probability of graduating in time. On the other hand, distributional regression methods do a poorer job with the average first year grades and the number of credits. For these two, the specification with the minimum MSE is 2, *i.e.* OLS with interactions between the treatment and the

---

[8]The efficiency criteria might be unfeasible due to legal issues and it is a theoretical benchmark.

selected covariates plus university-period dummies attains the minimum MSE.

Table 2: 10-Fold Cross Validated MSE

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$Y(1)$	21975	21957	22042	22022	21915	21908	21917	21909	21979	21985
$Y(2)$	17592	17562	17614	17605	18472	18461	18138	18129	17584	17601
$Y(3)$	10141	10108	10249	10225	18057	17874	13283	13217	10135	10134
$Y(4)$	67123	67014	64687	64660	63909	63879	63904	63875	64397	64402
Estimator	OLS	OLS	OLS	OLS	Logit	Logit	Probit	Probit	BR	BR
TIC		✓		✓		✓		✓		✓
UPE	✓	✓								
UE			✓	✓	✓	✓	✓	✓	✓	✓
YE			✓	✓	✓	✓	✓	✓	✓	✓

Notes: $Y(1)$, $Y(2)$, $Y(3)$, and $Y(4)$ respectively denote the probability of continuing into the second year, the average score, the number of credits, and the probability of graduating in time; BR, TIC, UPE, UE, and YE stand for blocking with regression, treatment interacted with covariates, university times period effect, university effect and year effect; the MSE of the average score and the number of credits is respectively scaled by 100 and 1000.

The main results of the estimation with the selected model are presented in Table 3.[9] Dropout rates greatly vary for students with different characteristics. In particular, male students, residents in the Center of Italy, those who study at a university inside their region of residence, those have a technical or vocational school degree, scored low in high school, or come from a urban area, have a higher probability of dropping out during their first year at college.

If every eligible student was awarded grant, it would increase the amount of students who successfully complete the first year by about 1.1 percentage points relative to the case in which no additional grants are awarded, *i.e.* a decrease of almost 15% of the first year dropouts, and 0.7 percentage points with respect to the case with the observed number of grants. Consequently, this would strongly increase grants expenditure for the policy maker. The gain tends to be larger for those students who would have a higher probability of dropping out without a grant. In fact, the correlation coefficient between the estimated gain and the estimated probability without a scholarship equals 0.6. Hence, the gain is over twice

---

[9]The results for all 10 estimation models are shown in Appendix B.3.

Table 3: Main results

	%	Continued into second year			First year average score			Number of credits			Graduated in time		
		$P_{it\tau}(1)$	$P_{it\tau}(0)$	$\Delta P_{it\tau}$	$P_{it\tau}(1)$	$P_{it\tau}(0)$	$\Delta P_{it\tau}$	$P_{it\tau}(1)$	$P_{it\tau}(0)$	$\Delta P_{it\tau}$	$P_{it\tau}(1)$	$P_{it\tau}(0)$	$\Delta P_{it\tau}$
Total	100.0	93.2	92.1	1.1	25.1	25.1	0.0	35.9	33.4	2.5	36.6	35.6	1.0
Female	64.2	94.0	93.1	0.8	25.4	25.4	0.0	36.5	34.6	1.9	38.2	37.7	0.5
Male	35.8	92.0	90.3	1.7	24.6	24.5	0.1	34.6	31.2	3.5	33.6	31.7	2.0
Foreign	2.4	94.0	91.5	2.4	24.3	24.0	0.2	36.4	32.7	3.7	34.3	33.6	0.7
Native	97.6	93.2	92.1	1.1	25.1	25.1	0.0	35.8	33.4	2.4	36.6	35.6	1.0
North	25.6	93.0	93.8	-0.7	25.2	25.2	0.0	39.8	38.3	1.4	47.6	47.9	-0.3
Center	12.6	92.0	90.9	1.1	25.5	25.6	-0.1	36.1	33.3	2.8	38.3	37.4	0.9
South	61.8	93.6	91.7	1.9	25.0	25.0	0.0	34.2	31.4	2.8	31.6	30.1	1.6
Dif reg	14.1	96.3	95.8	0.5	24.8	24.9	-0.1	39.7	36.7	2.9	44.7	42.3	2.5
Same reg	85.9	92.7	91.5	1.2	25.2	25.1	0.0	35.2	32.8	2.4	35.2	34.4	0.8
HS deg	58.6	95.9	95.3	0.5	25.5	25.5	-0.1	37.8	35.8	2.0	39.3	38.7	0.6
TS deg	28.4	90.0	87.6	2.4	24.5	24.4	0.1	33.1	29.7	3.3	32.3	30.4	1.9
Ot deg	13.1	88.6	87.5	1.1	24.6	24.6	0.0	33.3	30.5	2.8	33.6	32.6	1.0
Score 1	13.4	86.5	84.7	1.8	23.7	23.6	0.0	26.3	22.7	3.6	20.1	19.0	1.1
Score 2	21.3	90.9	89.7	1.2	24.2	24.2	0.0	31.9	28.9	3.0	28.7	27.1	1.6
Score 3	23.3	93.9	92.8	1.1	24.8	24.8	0.0	35.9	33.6	2.3	35.3	34.4	0.9
Score 4	20.2	95.4	94.3	1.1	25.5	25.4	0.0	38.9	36.5	2.5	41.9	40.5	1.4
Score 5	21.8	96.9	96.2	0.7	26.5	26.6	-0.1	42.8	41.3	1.5	50.7	50.6	0.1
Urban	42.6	93.0	91.8	1.2	25.3	25.3	0.0	36.2	33.8	2.4	37.2	36.4	0.7
Rural	57.4	93.4	92.3	1.1	25.0	25.0	0.0	35.6	33.1	2.5	36.1	34.9	1.2

Notes: North, Center and South denote the region of residence of the student; Different region (Same region) denotes that student region of residence and the region of the university are different (the same); HS degree, TS degree, and Other degree respectively denote the possession of a high school degree, a technical degree and other type of degree; Scores 1-5 indicate that the average high school score was in each of the 5 following intervals: 0-69, 70-79, 80-89, 90-99, 100.

as large for male than for female students, or for foreign rather than native students. In regional terms, residents in the South of Italy would benefit the most, and the effect for students in the North is on average negative. Moreover, students who moved to another region to attend university are less likely to dropout, so the gain is also smaller for them. Interestingly, although high school score has a positive predictive effect on finishing the first year, the gain is smaller as the score increases. This suggests that awarding grants based on good past scores leads to low dropout rates for those who are awarded the scholarship, but also a relatively small decrease in overall dropout rates.

As for the probability of graduating in time, the effect is larger again for male students who are resident in the South. However, movers benefit about three times as much as non-movers, suggesting that those who do not study in their region of residence are more likely to increase their effort if they receive a grant. Furthermore, although high school grades are a good predictor of the probability of graduating in time, it is no longer the case that those who benefit are those with the lower grades. Students with a grant pass approximately 2.5 extra credits on average during their first year. The gain is larger for males, foreign students from the Center or South of Italy and, interestingly, students with low high school scores. Unexpectedly, the treatment effect on the average score is zero up to the first decimal point for students of all demographic classes.

This degree of heterogeneity in the estimated effect is reflected in the distribution of the estimated probability of continuing into the second year of university. Table 4 shows this distribution when no student receives a grant, when all of them do, and the distribution of the difference of these two cases. The first two distributions show that most enrolled students have a low probability of dropping out during their first year, even without a grant. Moreover, awarding a grant to every student would be preferable to not awarding them, as the distribution when all students have a grant dominates the distribution when no student is awarded one. This does not imply that the effect is positive for every individual. However, almost 75% of the students are estimated to benefit from the grant. Moreover, for 23.6% of the students this increase is larger than 2 percentage points. These distributions are shown

graphically in Figure 1.

Table 4: Unconditional distributions

		quantile					Mean
		0.1	0.25	0.5	0.75	0.9	
	NONE	82.0	88.9	94.6	97.5	98.7	92.1
$Y(1)$	ALL	84.8	90.7	95.2	97.7	98.8	93.2
	DIF	-1.0	-0.1	0.6	1.9	4.3	1.1
	NONE	23.4	24.1	25.0	26.0	27.0	25.1
$Y(2)$	ALL	23.5	24.2	25.1	26.0	26.8	25.1
	DIF	-0.2	-0.1	0.0	0.1	0.2	0.0
	NONE	20.7	26.7	33.6	40.2	45.7	33.4
$Y(3)$	ALL	24.1	29.7	36.1	42.1	47.0	35.9
	DIF	0.8	1.2	2.5	3.6	4.8	2.5
	NONE	7.6	20.2	35.1	50.1	64.0	35.6
$Y(4)$	ALL	9.5	22.2	36.2	50.8	64.2	36.6
	DIF	-2,0	-0,7	0,9	2,7	4,4	1,0

Notes: $Y(1)$, $Y(2)$, $Y(3)$, and $Y(4)$ respectively denote the probability of continuing into the second year, the average score, the number of credits, and the probability of graduating in time; NONE, ALL, and DIF respectively show the distribution of the estimated probability of continuing into the second year at the university when no student was awarded a grant (Equation 3), when all were awarded a grant (Equation 4), and the distribution of the difference between these two probabilities (Equation 5). All probabilities are expressed in percentage points.

Regarding the average score, the distribution when all students are treated is very similar to the distribution when none is, whereas for the distribution of the number of credits, the gain is relatively uniformly distributed between 0 and 5. Finally, the distribution of the probability of graduating in time shows a jump for students in the lower tail. This contrasts with those who have a higher probability in the absence of grant, for whom the increase would be small on average. As in the probability of continuing into the second year, the effect would be negative for a sizable percentage of the students.

Figure 1: Distribution of probability of continuing into second year
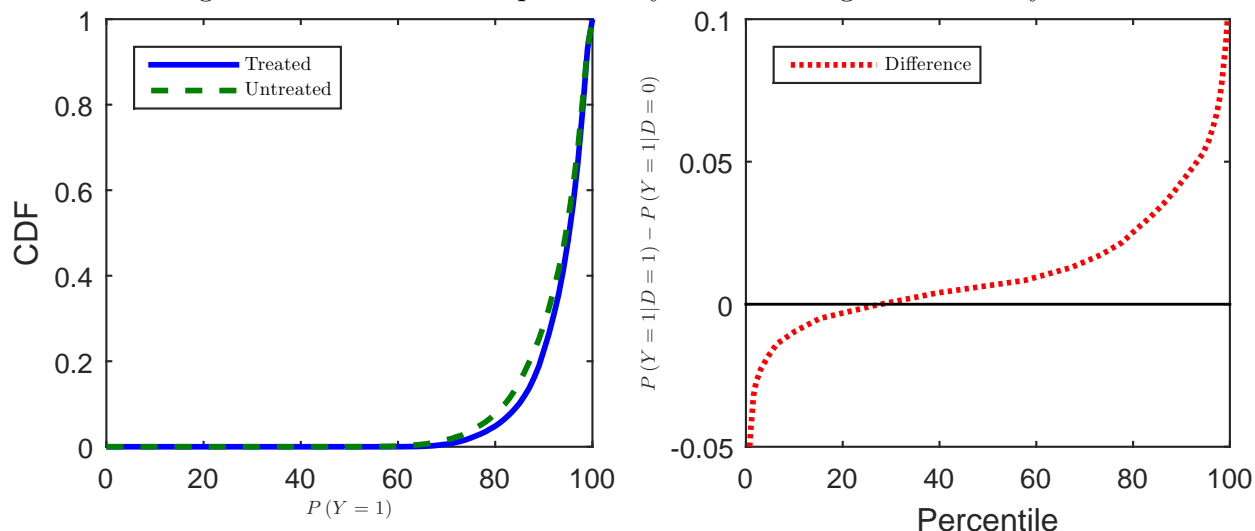
Notes: The left panel represents the cdf of the estimated probability of enrolling during the second year if every student was either a grant recipient or not; the right panel represents the quantile function of the estimated gain in the probability of graduating. All probabilities are expressed in percentage points.

6   Assignment of Grants

In our data, the number of students with a grant equals 173,070, 50.9% of the sample. Table 5 shows the characteristics of grant recipients under different counterfactuals. In the baseline scenario, foreign students, residents in either then North or the Center of Italy, studying in a different region, with a relatively low score and a rural area are more likely to receive a grant.

The effectiveness assignment rule targets those with the largest gain, meaning that it would increase the number of grant recipients who are male, native, come from the South and study in the same region of residence, have a technical school degree, have an even relatively lower score, and come from a urban area. The most striking change is the sharp increase of grants awarded to students in Southern Italy, of over 30 percentage points. On the other hand, the merit-based assignment rule would target those with low dropout rates, meaning that it would increase the number of grant recipients who are female, foreign, come from the South and have moved to a different region, have a high school degree, were high achievers in high school, and live in an urban area. Under this rule, the increase in the South

19

would not be as big, but still quite substantial. On the other hand, targeting high-achieving students means that their gain would be relatively small.

Table 5: Characteristics of grant recipients under different counterfactuals

	BL	CF1	CF2	CF3	CF4	CF5	CF6	CF7
Female	64.6	59.9	69.0	56.5	68.0	50.0	50.0	67.9
Male	35.4	40.1	31.0	43.5	32.0	50.0	50.0	32.1
Foreign	3.3	2.1	2.6	3.3	3.1	2.3	2.5	1.7
Native	96.7	97.9	97.4	96.7	96.9	97.7	97.5	98.3
North	33.6	4.3	25.9	33.6	33.6	6.6	26.0	22.5
Center	15.3	11.5	10.8	15.3	15.3	11.7	11.2	12.9
South	51.0	84.1	63.3	51.0	51.0	81.6	62.7	64.6
Different region	18.9	9.7	20.9	6.7	19.5	11.4	21.0	14.5
Same region	81.1	90.3	79.1	93.3	80.5	88.6	79.0	85.5
HS degree	56.2	47.1	80.6	40.1	77.8	45.1	74.3	58.9
TS degree	29.6	41.7	14.0	47.2	15.8	43.5	19.3	28.1
Other degree	14.2	11.2	5.5	12.8	6.4	11.4	6.4	13.0
Score 1	15.0	16.4	3.2	17.4	4.6	17.6	5.3	0.0
Score 2	23.3	25.0	11.8	24.4	12.2	25.1	14.0	0.0
Score 3	24.7	23.7	23.9	22.1	22.8	21.6	21.6	17.4
Score 4	19.6	18.8	26.9	19.6	25.9	18.6	24.8	39.7
Score 5	17.4	16.0	34.2	16.6	34.4	17.0	34.4	42.9
Urban	39.6	42.4	42.6	47.4	43.9	43.8	43.0	43.3
Rural	60.4	57.6	57.4	52.6	56.1	56.2	57.0	56.7

Notes: BL and CFX stand for baseline and counterfactual #X, as described in Section 4; North, Center and South denote the region of residence of the student; Different region (Same region) denotes that student region of residence and the region of the university are different (the same); HS degree, TS degree, and Other degree respectively denote the possession of a high school degree, a technical degree and other type of degree; Scores 1-5 indicate that the average high school score was in each of the 5 following intervals: 0-69, 70-79, 80-89, 90-99, 100.

Combining either the effectiveness or the merit-based assignment rules with regional or gender quotas mostly changes the intended quotas, leaving the remaining variables balanced as in the quota-free rules. Finally, the scored-based assignment rule offers a picture similar to the merit-based assignment, with a slightly smaller proportion of grants awarded to those not studying in their regions of residence and those who have a high school degree.

The counterfactual results are reported in Table 6. In the baseline scenario, almost 7.4% of students would drop out after a year at college. Effectiveness-based assignment

rules (CF1, CF3, and CF5) would attain a reduction of the first-year dropout rate, and in the absence of quotas, only 6.5% of all students would drop out. On the other hand, merit-based assignment rules (CF2, CF4, and CF6) would actually increase the dropout rate during the first year. A similar story can be told about the probability of graduating in time: merit-based assignments would reduce the percentage of students who finish in time, whereas effectiveness-based assignments would increase it. Comparing the counterfactuals without quotas, it would mean an increase of 0.7% of the student body finishing in time. Similarly to the merit-based rules, the score-based one would slightly reduce the percentage of students who graduate in time.

Regarding the average number of passed credits, different rules would have a modest impact: effectiveness-based assignment rules would increase the average number of passed credits, whereas merit-based assignment rules would reduce it. This can be reconciled with the results in Table 4: the largest difference between the distribution in which everyone and noone receives a grant takes place at the bottom half, so giving grants to relatively low-performing students would increase the average number of passed credits. Lastly, the impact on the average scores during the first year would be negligible for the different rules.

Table 6: Mean counterfactual outcomes

	BL	CF1	CF2	CF3	CF4	CF5	CF6	CF7
$Y(1)$	92.6	93.5	92.3	93.3	92.2	93.4	92.4	92.6
$Y(2)$	25.1	25.1	25.1	25.1	25.1	25.2	25.1	25.1
$Y(3)$	34.7	35.1	34.3	34.9	34.3	35.1	34.5	34.4
$Y(4)$	36.1	36.5	35.8	36.4	35.8	36.6	36.0	35.9

Notes: $Y(1)$, $Y(2)$, $Y(3)$, and $Y(4)$ respectively denote the probability of continuing into the second year, the average score, the number of credits, and the probability of graduating in time; BL and CFX stand for baseline and counterfactual #X, as described in Section 4.

However, to get a grasp of how beneficial grants were, it is convenient to look at the percentage of efficient grants, as defined in Section 4. Under the baseline assignment, 0.9% of the students who would be awarded a grant would not drop out thanks to the grant. This means that 99% of the grants would not affect the outcome of their recipients: 92% of them

would have successfully finished the first year even without a grant (type I inefficiency), and 7% would have dropped out with the grant (type II inefficiency).

Table 7: Counterfactual probabilities of continuing into second year

	BL	CF1	CF2	CF3	CF4	CF5	CF6	CF7
Dropout rate	7.4	6.5	7.7	6.7	7.8	6.6	7.6	7.4
Efficient grants	0.9	2.7	0.4	2.3	0.3	2.6	0.6	0.9
Type I inefficient grants	92.1	88.8	97.2	89.3	97.1	88.8	96.3	94.9
Type II inefficient grants	7.0	8.6	2.4	8.4	2.7	8.6	3.0	4.2

Notes: BL and CFX stand for baseline and counterfactual #X, as described in Section 4.

The effectiveness-based assignment rule would minimize the dropout rate, which would amount to 6.5%, thanks to tripling the fraction of efficient grants relative to the baseline scenario. On the other hand, type II inefficiency would be maximized under this assignment rule, and over 8% of the grant recipients would drop out. This inefficiency would be minimized under the merit-based assignment rule, but targeting students that are the least likely to drop out if they were given a grant implies that the percentage of efficient grants would be halved with respect to the baseline scenario, and as such, the dropout rate would increase. Introducing the quotas slightly mitigate the results towards the baseline scenario, and the fraction of efficient grants are quite close to those under the quota-free assignment rule. Finally, the results under score-based assignment rule are remarkably similar to those in the baseline scenario.

These results point at the existence of a clear trade-off between efficiency and type I inefficient grants. On the one hand, increasing the number of graduates requires targeting those who gain the most out of the scholarship, who tend to be those with a low probability of finishing the first year. This implies that a relatively large fraction of the grants will be given to students who will not enroll in the subsequent year. On the other hand, targeting those with the smallest probability of dropping out translates into a modest increase in the number of students who successfully complete their first year. Depending on the objective function of the policy maker, this trade-off would suggest that either an effectiveness or merit-based rule is more appropriate.

Note also that some of the rules could set perverse incentives. For example, with the effectiveness-based assignment rule, students who got a relatively small score at high school would have a higher chance of getting a scholarship than those who got a relatively high score. Hence, students in the subsequent years would have an incentive to put less effort in high school to have a higher chance to get the grant.[10] Similarly, if students who moved away from their region of residence had a lower probability of receiving the grant, the rule would reduce mobility. Finally, the current analysis takes enrollment as granted, and a change in the assignment rule could affect the decision to enroll, which would change the population of students and change the estimated effects.

7   Conclusions

This paper studied how the design of grant assignment rules affects dropout rates in Italy. We estimated the predicted probability that each student continues into the second year with and without the grant. This allowed us to obtain the following results. First, we found that awarding the grant to all the eligible students would double the government expenditure on grants, but this would only slightly reduce the dropout rate. It is therefore important to target the policy, also because we found that the grants have a heterogeneous effect on the probability of dropping out according to students' characteristics. This means that the design of the assignment rule may have a first order impact on the number of dropouts. We then tried different assignment rules, maintaining the number of grants constant but changing the characteristics of the grant recipients. It is important to notice that in all the exercises, due to the data availability, we abstract from the possible effect that such changes could have on the enrollment of new students. There are two main types of rules: those based on effectiveness, *i.e.* those that target students with an expected larger reduction of the dropout probability due to the grant[11], and those based on merit, *i.e.* those that

---

[10]In any case, this perverse incentive could be offset if access to the desired university degree requires a minimum high school score.

[11]The efficiency criteria is unfeasible in practice and it is only a theoretical benchmark.

target students with the smallest probability of dropping out if they are awarded a grant. The first type achieves the minimum number of dropouts, about 2000 students per year or 6.5% of the student population. On the other hand, the merit-based assignment rule minimizes the number of dropouts amongst grant recipients (2.4%, relative to 8.6% under the efficiency-based assignment rule), at the cost of increasing the overall number of dropouts to over 2350 students, about 7.7% of the student population.

The results highlight the existence of a trade-off between efficiency, measured as the total number of students who do not drop-out, and merit, *i.e.* targeting students who are more likely to continue studying if the receive a grant. Conventional wisdom would suggest focusing on the drop-out rate amongst grant recipients to insure that the grants are not wasted on dropout. However, this would increase the number of dropouts amongst non-recipients, increasing the overall number of dropouts.

To get a better understanding of the required effort to reduce the incidence of dropouts, every 88 awarded grants, 1 would be given to a student who would have otherwise dropped out, 81 to students who would not have dropped out even without it, and the remaining 6 would have been given to students who would have dropped out regardless.

References

Adamopoulou, E. and G. M. Tanzi (2017). Academic drop-out and the great recession. *Journal of Human Capital 11*(1), 35–71.

Albert, A. and J. A. Anderson (1984). On the existence of maximum likelihood estimates in logistic regression models. *Biometrika 71*(1), 1–10.

ANVUR (2016). Rapporto biennale sullo stato del sistema universitario e della ricerca.

Baumgartner, H. J. and V. Steiner (2006). Does more generous student aid increase enrolment rates into higher education? evaluating the german student aid reform of 2001. *IZA Discussino Paper*.

Belloni, A., V. Chernozhukov, and C. Hansen (2014a). High-dimensional methods and inference on structural and treatment effects. *The Journal of Economic Perspectives 28*(2), 29–50.

Belloni, A., V. Chernozhukov, and C. Hansen (2014b). Inference on treatment effects

after selection among high-dimensional controls. *The Review of Economic Studies 81*(2), 608–650.

Bettinger, E. (2004). How financial aid affects persistence. In *College choices: The economics of where to go, when to go, and how to pay for it*, pp. 207–238. University of Chicago Press.

Bettinger, E. P., B. T. Long, P. Oreopoulos, and L. Sanbonmatsu (2012). The role of application assistance and information in college decisions: Results from the h&r block fafsa experiment. *The Quarterly Journal of Economics 127*(3), 1205–1242.

Cappelli, P. and S. Won (2016). How you pay affects how you do: financial aid type and student performance in college. Technical report, National Bureau of Economic Research.

Castleman, B. L. and B. T. Long (2016). Looking beyond enrollment: The causal effect of need-based grants on college access, persistence, and graduation. *Journal of Labor Economics 34*(4), 1023–1073.

Cornwell, C., D. B. Mustard, and D. J. Sridhar (2006). The enrollment effects of merit-based financial aid: Evidence from georgia's hope program. *Journal of Labor Economics 24*(4), 761–786.

De Angelis, I., V. Mariani, and R. Torrini (2017). New evidence on interregional mobility of students in tertiary education: the case of italy. *Politica economica 33*(1), 73–96.

Deming, D. and S. Dynarski (2009). Into college, out of poverty? policies to increase the postsecondary attainment of the poor. Technical report, NBER Working Paper.

Denning, J. T. (2019). Born under a lucky star financial aid, college completion, labor supply, and credit constraints. *Journal of Human Resources 54*(3), 760–784.

Denning, J. T., B. M. Marx, and L. J. Turner (2019). Propelled: The effects of grants on graduation, earnings, and welfare. *American Economic Journal: Applied Economics 3*(11), 193–224.

Di Pietro, G. (2004). The determinants of university dropout in italy: a bivariate probability model with sample selection. *Applied Economics Letters 11*(3), 187–191.

Dynarski, S. and J. Scott-Clayton (2013). Financial aid policy: Lessons from research. Technical report, National Bureau of Economic Research.

Fack, G. and J. Grenet (2015). Improving college access and success for low-income students: Evidence from a large need-based grant program. *American Economic Journal: Applied Economics 7*(2), 1–34.

Firth, D. (1993). Bias reduction of maximum likelihood estimates. *Biometrika 80*(1), 27–38.

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

Gitto, L., L. F. Minervini, and L. Monaco (2015). University dropout rates in italy. In *Youth and the Crisis*, pp. 89–102. Routledge.

Glocker, D. (2011). The effect of student aid on the duration of study. *Economics of Education Review 30*(1), 177–190.

Goodman, J. (2008). Who merits financial aid?: Massachusetts' adams scholarship. *Journal of public Economics 92*(10-11), 2121–2131.

Hanushek, E. A., C. K. Y. Leung, and K. Yilmaz (2014). Borrowing constraints, college aid, and intergenerational mobility. *Journal of Human Capital 8*(1), 1–41.

Imbens, G. W. (2015). Matching methods in practice: Three examples. *Journal of Human Resources 50*(2), 373–419.

Kane, T. J. (2003). A quasi-experimental estimate of the impact of financial aid on college-going. Technical report, National Bureau of Economic Research.

Lauer, C. (2002). Enrolments in higher education: do economic incentives matter? *Education+ Training 44*(4/5), 179–185.

Mealli, F. and 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., E. Rettore, and G. Tanzi (2020). The effect of grants on university drop-out rates: Evidence on the italian case. *Journal of Human Capital forthcoming.*

Nielsen, H. S., T. Sørensen, and C. Taber (2010). Estimating the effect of student aid on college enrollment: Evidence from a government grant policy reform. *American Economic Journal: Economic Policy 2*(2), 185–215.

Schudde, L. and J. Scott-Clayton (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.

Singell, L. D. (2004). Come and stay a while: does financial aid effect retention conditioned on enrollment at a large public university? *Economics of Education review 23*(5), 459–471.

Sneyers, E., T. Agasisti, K. De Witte, and G. Graziosi (2016). The impact of scholarships on students' performance: a study on five italian universities. Technical Report 1, DEAMS Research paper.

Steiner, V. and K. Wrohlich (2012). Financial student aid and enrollment in higher education: New evidence from germany. *The Scandinavian Journal of Economics 114*(1), 124–147.

Stinebrickner, R. and T. Stinebrickner (2008). The effect of credit constraints on the college drop-out decision: A direct approach using a new panel study. *American Economic Review 98*(5), 2163–84.

Vergolini, L. and N. Zanini (2015). Away, but not too far from home. the effects of financial aid on university enrolment decisions. *Economics of Education Review 49*, 91–109.

Zotti, R. (2015). Should i stay or should i go? dropping out from university: an empirical analysis of students' performances. In *Youth and the Crisis*, pp. 71–88. Routledge.

A   Selection of Covariates

The covariates used in this paper are the following: female, foreign, region of residence (north, center, and south and islands), region of the university different from region of origin of the student, type of high school degree (regular, technical, and others), high school score split into 5 intervals, and whether the local labor system is urban. We consider these variables, as well as all their interactions, together with a set of fixed effects. Moreover, in some of the models the treatment variable is interacted with all covariates. The selection is done using the following 10-fold cross validation Lasso for each model:

- For OLS and binary choice models with no interactions between the covariates and the grant indicator, the set of selected covariates (as well as the fixed effects) is the union of the non-zero covariates of the Lasso of the first stage equation (receiving a grant as a function of the covariates) and of the second stage equation.

- For OLS and binary choice models with interactions between the covariates and the grant indicator, the set of selected covariates (as well as the fixed effects) is the union of the non-zero covariates of the Lasso of the first stage equation (receiving a grant as a function of the covariates) and of the second stage equation separately for those who received the grant, and those who did not.

- For blocking with regression, the set of selected covariates (as well as the fixed effects) is the set of non-zero covariates of the Lasso of the first stage equation.

The rationale for these choices mimics that of Belloni et al. (2014a,b). In the first case, some variables could be important for either the treatment or the outcome but not the other. The double selection mechanism safeguards against the possibility of selecting only one of the two sets of relevant variables. In the second case, given the interactions between the treatment and the covariates, the selection is differently applied for the outcome when treated and when not. To the best of our knowledge there is no theoretical result on the selection of variables in the third case, so we use only the first stage equation using the

Lasso for the binary choice model. Finally, given the selected set of covariates, we run the regression without penalization for each model.

B   Additional results

B.1   Restricted sample analysis

The identification of the model is based on selection on observables. However, the internal validity of the estimates shown in Section 5 could be threatened if the common support assumption is not satisfied. In some years in some of the universities all students who were eligible for a grant were awarded it. Because of the fixed effects included in the regression, this implies that the probability of being awarded a grant equals 1, violating this assumption.

The following analysis restricts the sample to those students for which the common support assumption and in each university and period at there were both students with and without a grant. In Table 8 I compare the fit of the estimates with the full and the restricted samples, for the population of the restricted sample. The order of fit of the estimators is largely unchanged, with the exception of the blocking with regression estimators. The fit of these two is now better than that of the OLS specifications, but worse than that of the binary choice estimators. As for the estimated ATE for the restricted sample, it is smaller for all estimators, and the variability across them remains large. For the best specification, it falls from 1.1% to 0.8%.

B.2   Analysis for non-eligible students

Students who were not eligible to apply for a grant were excluded from the main analysis. Their dropout rate is displayed in Table 9. Overall, 11.5% of non-eligible students did not manage to successfully complete the first year, almost 4 percentage points higher than eligible students. The difference in the dropout rate was positive for all demographic groups, but it was particularly larger for foreign students and those with a technical degree. Moreover,

Table 8: Restricted sample analysis

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
MSE	Full	19771	19744	19845	19817	19717	19699	19722	19704	19762	19748
	Restricted	19770	19742	19844	19814	19715	19696	19719	19701	19719	19723
ATE	Full	0.65	0.85	0.85	1.05	0.95	1.11	0.97	1.17	1.66	1.57
	Restricted	0.24	0.42	0.54	0.72	0.68	0.84	0.68	0.88	0.74	0.81
	Estimator	OLS	OLS	OLS	OLS	Logit	Logit	Probit	Probit	BR	BR
	Treatment interactions		✓		✓		✓		✓		✓
	University×period effect	✓	✓								
	University&period effect			✓	✓	✓	✓	✓	✓	✓	✓

Notes: BR, TIC, UPE, UE, and YE stand for blocking with regression, treatment interacted with covariates, university times period effect, university effect and year effect; the MSE of the average score and the number of credits is respectively scaled by 100 and 1000; the restricted sample consists only of students that where in universities in years in which some of the students who were eligible for a grant were awarded it, and some did not.

the lower the high school scores, the largest the difference. Using the estimates for the eligible sample with the selected model, we can fit the average probability of continuing into the second year for each group. The average probability is overestimated for every group considered, with the exception of students with the highest high school scores. However, for several demographic groups the difference between the actual value and the estimate is smaller than the difference between the actual value for eligible and non-eligible students without a grant. This evidence suggests that non-eligible and eligible students are different in terms of their unobservable characteristics.

	Non-eligible		Eligible without a grant	
	$\overline{y}$	$\overline{\hat{y}}$	$\overline{y}$	$\overline{\hat{y}}$
Total	88.5	92.1	92.1	92.1
Female	90.1	93.2	93.2	93.2
Male	86.4	90.7	90.2	90.2
Foreign	84.5	91.3	93.0	93.0
Native	88.5	92.1	92.1	92.1
North	89.9	94.3	93.5	93.4
Center	89.0	90.1	93.5	93.5
South	86.8	90.9	91.6	91.6
Different region	92.8	95.6	96.7	96.7
Same region	88.0	91.7	91.6	91.6
HS degree	92.8	95.0	95.4	95.4
TS degree	80.4	86.9	87.2	87.2
Other degree	82.5	87.8	86.3	86.3
Score 1	80.8	87.2	83.6	83.6
Score 2	87.1	91.2	89.3	89.3
Score 3	91.1	94.0	92.5	92.5
Score 4	93.8	95.3	94.0	94.0
Score 5	96.5	97.2	96.1	96.1
Urban	88.8	92.2	91.9	91.9
Rural	88.1	92.0	92.2	92.2

Table 9:

Notes: North, Center and South denote the region of residence of the student; Different region (Same region) denotes that student region of residence and the region of the university are different (the same); HS degree, TS degree, and Other degree respectively denote the possession of a high school degree, a technical degree and other type of degree; Scores 1-5 indicate that the average high school score was in each of the 5 following intervals: 0-69, 70-79, 80-89, 90-99, 100.

B.3 Extra Tables

Table 10: Mean probability of continuing into second year without a grant, all models

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Total	92.2	92.2	92.1	92.2	92.1	92.1	92.1	92.1	91.0	91.1
Female	93.1	93.3	93.0	93.2	93.0	93.1	93.0	93.1	91.9	91.9
Male	90.6	90.3	90.5	90.3	90.4	90.3	90.4	90.2	89.5	89.5
Foreign	93.2	92.1	93.1	92.1	93.0	91.5	93.0	91.6	91.2	91.3
Native	92.2	92.2	92.1	92.2	92.1	92.1	92.0	92.1	91.0	91.1
North	92.9	93.8	92.7	93.9	92.7	93.8	92.7	93.7	91.1	91.2
Center	91.4	90.9	91.3	91.0	91.1	90.9	91.1	90.9	90.2	90.3
South	92.1	91.8	92.0	91.7	92.0	91.7	92.0	91.7	91.1	91.2
Different region	95.7	95.7	95.6	96.0	95.7	95.8	95.7	95.8	93.8	93.9
Same region	91.7	91.6	91.6	91.6	91.5	91.5	91.5	91.5	90.6	90.6
HS degree	95.2	95.5	95.1	95.4	95.2	95.3	95.2	95.3	94.1	94.1
TS degree	88.3	87.6	88.2	87.7	88.0	87.6	88.0	87.6	87.1	87.1
Other degree	87.3	87.5	87.2	87.6	86.9	87.5	87.0	87.5	86.0	86.1
Score 1	85.1	84.7	85.0	84.6	84.5	84.7	84.6	84.6	83.7	83.8
Score 2	89.9	89.8	89.8	89.8	89.6	89.7	89.6	89.7	88.5	88.6
Score 3	92.9	92.9	92.8	92.8	92.8	92.8	92.8	92.8	91.6	91.7
Score 4	94.3	94.3	94.2	94.3	94.3	94.3	94.2	94.2	93.1	93.2
Score 5	96.2	96.5	96.1	96.5	96.3	96.2	96.3	96.2	95.3	95.4
Urban	91.9	92.0	91.8	91.9	91.7	91.8	91.7	91.8	90.8	90.8
Rural	92.5	92.4	92.4	92.4	92.3	92.3	92.3	92.3	91.2	91.3

Notes: BR, TIC, UPE, UE, and YE stand for blocking with regression, treatment interacted with covariates, university times period effect, university effect and year effect; North, Center and South denote the region of residence of the student; Different region (Same region) denotes that student region of residence and the region of the university are different (the same); HS degree, TS degree, and Other degree respectively denote the possession of a high school degree, a technical degree and other type of degree; Scores 1-5 indicate that the average high school score was in each of the 5 following intervals: 0-69, 70-79, 80-89, 90-99, 100.

Table 11: Mean probability of continuing into second year with a grant, all models

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Total	92.9	93.1	93.0	93.2	93.0	93.2	93.0	93.3	92.7	92.7
Female	93.8	93.7	93.9	93.9	93.9	94.0	93.9	94.0	93.6	93.5
Male	91.3	91.8	91.4	92.0	91.5	92.0	91.5	92.0	91.1	91.1
Foreign	93.8	94.0	93.9	94.0	93.9	94.0	93.9	94.0	93.8	93.7
Native	92.9	93.0	93.0	93.2	93.0	93.2	93.0	93.3	92.6	92.6
North	93.5	92.9	93.6	92.9	93.6	93.0	93.6	93.1	93.4	93.4
Center	92.1	92.1	92.2	92.0	92.2	92.0	92.2	92.0	92.1	92.1
South	92.8	93.3	92.9	93.6	93.0	93.6	93.0	93.6	92.5	92.5
Different region	96.4	96.4	96.4	96.3	96.3	96.3	96.3	96.3	96.3	96.3
Same region	92.3	92.5	92.4	92.7	92.5	92.7	92.5	92.8	92.1	92.1
HS degree	95.9	95.6	96.0	95.8	95.8	95.9	95.9	95.9	95.6	95.6
TS degree	89.0	89.8	89.1	89.9	89.4	90.0	89.4	90.0	88.8	88.8
Other degree	88.0	88.5	88.1	88.6	88.4	88.6	88.4	88.7	87.8	87.8
Score 1	85.7	86.4	85.8	86.5	86.2	86.5	86.2	86.5	85.6	85.6
Score 2	90.5	90.7	90.6	90.9	90.8	90.9	90.8	91.0	90.4	90.4
Score 3	93.6	93.7	93.7	93.9	93.7	93.9	93.8	94.0	93.4	93.4
Score 4	95.0	95.3	95.1	95.4	95.0	95.4	95.0	95.4	94.7	94.7
Score 5	96.9	96.7	97.0	96.8	96.8	96.9	96.8	96.9	96.6	96.5
Urban	92.5	92.7	92.6	92.9	92.7	93.0	92.7	93.0	92.3	92.3
Rural	93.1	93.3	93.2	93.5	93.3	93.4	93.3	93.5	93.0	92.9

Notes: BR, TIC, UPE, UE, and YE stand for blocking with regression, treatment interacted with covariates, university times period effect, university effect and year effect; North, Center and South denote the region of residence of the student; Different region (Same region) denotes that student region of residence and the region of the university are different (the same); HS degree, TS degree, and Other degree respectively denote the possession of a high school degree, a technical degree and other type of degree; Scores 1-5 indicate that the average high school score was in each of the 5 following intervals: 0-69, 70-79, 80-89, 90-99, 100.

Table 12: Mean gain in probability of continuing into second year from a grant, all models

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Total	0.7	0.9	0.9	1.0	1.0	1.1	1.0	1.2	1.7	1.6
Female	0.7	0.4	0.9	0.7	0.9	0.8	0.9	0.9	1.7	1.6
Male	0.7	1.6	0.9	1.7	1.1	1.7	1.1	1.8	1.6	1.5
Foreign	0.7	1.9	0.9	1.9	0.9	2.4	0.9	2.4	2.6	2.5
Native	0.7	0.8	0.9	1.0	1.0	1.1	1.0	1.2	1.6	1.5
North	0.7	-0.8	0.9	-1.0	0.9	-0.7	0.9	-0.6	2.3	2.2
Center	0.7	1.2	0.9	1.0	1.1	1.1	1.1	1.1	1.8	1.8
South	0.7	1.5	0.9	1.9	1.0	1.9	1.0	2.0	1.3	1.3
Different region	0.7	0.7	0.9	0.3	0.5	0.5	0.6	0.4	2.5	2.4
Same region	0.7	0.9	0.9	1.2	1.0	1.2	1.0	1.3	1.5	1.4
HS degree	0.7	0.2	0.9	0.5	0.6	0.5	0.7	0.6	1.6	1.5
TS degree	0.7	2.2	0.9	2.3	1.4	2.4	1.4	2.5	1.8	1.7
Other degree	0.7	1.0	0.9	1.0	1.5	1.1	1.5	1.2	1.8	1.7
Score 1	0.7	1.7	0.9	1.8	1.8	1.8	1.7	1.9	1.9	1.8
Score 2	0.7	0.9	0.9	1.1	1.3	1.2	1.2	1.3	1.8	1.7
Score 3	0.7	0.9	0.9	1.1	0.9	1.1	0.9	1.1	1.8	1.7
Score 4	0.7	1.0	0.9	1.2	0.7	1.1	0.8	1.2	1.6	1.5
Score 5	0.7	0.2	0.9	0.3	0.5	0.7	0.5	0.7	1.2	1.1
Urban	0.7	0.8	0.9	1.0	1.0	1.2	1.0	1.2	1.5	1.4
Rural	0.7	0.9	0.9	1.1	0.9	1.1	1.0	1.2	1.8	1.7

Notes: BR, TIC, UPE, UE, and YE stand for blocking with regression, treatment interacted with covariates, university times period effect, university effect and year effect; North, Center and South denote the region of residence of the student; Different region (Same region) denotes that student region of residence and the region of the university are different (the same); HS degree, TS degree, and Other degree respectively denote the possession of a high school degree, a technical degree and other type of degree; Scores 1-5 indicate that the average high school score was in each of the 5 following intervals: 0-69, 70-79, 80-89, 90-99, 100.

Table 13: Mean first year grades without a grant, all models

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Total	25.1	25.1	25.1	25.1	25.7	25.7	25.5	25.5	25.1	25.1
Female	25.4	25.4	25.4	25.4	25.9	25.9	25.8	25.8	25.4	25.4
Male	24.6	24.5	24.6	24.6	25.2	25.1	25.0	25.0	24.6	24.6
Foreign	24.2	24.0	24.2	24.1	24.7	24.4	24.6	24.4	24.2	24.2
Native	25.1	25.1	25.1	25.2	25.7	25.7	25.6	25.6	25.1	25.1
North	25.2	25.2	25.2	25.3	25.8	25.8	25.6	25.7	25.2	25.2
Center	25.6	25.6	25.6	25.6	26.2	26.2	26.1	26.0	25.6	25.6
South	25.0	25.0	25.0	25.0	25.5	25.5	25.4	25.4	25.0	25.0
Different region	24.8	24.9	24.8	24.9	25.4	25.4	25.3	25.3	24.8	24.8
Same region	25.2	25.1	25.2	25.2	25.7	25.7	25.6	25.6	25.2	25.2
HS degree	25.5	25.5	25.5	25.5	26.1	26.1	26.0	26.0	25.5	25.5
TS degree	24.5	24.4	24.5	24.4	25.0	24.9	24.8	24.8	24.5	24.5
Other degree	24.6	24.6	24.6	24.6	25.1	25.1	25.0	25.0	24.6	24.6
Score 1	23.7	23.6	23.7	23.7	24.1	24.1	24.0	24.0	23.7	23.7
Score 2	24.2	24.2	24.2	24.2	24.7	24.6	24.6	24.6	24.2	24.2
Score 3	24.8	24.8	24.9	24.9	25.4	25.4	25.2	25.2	24.8	24.9
Score 4	25.5	25.4	25.5	25.5	26.0	26.0	25.9	25.9	25.5	25.5
Score 5	26.6	26.6	26.6	26.6	27.3	27.4	27.1	27.2	26.6	26.6
Urban	25.3	25.3	25.3	25.3	25.9	25.8	25.7	25.7	25.3	25.3
Rural	25.0	25.0	25.0	25.0	25.5	25.5	25.4	25.4	25.0	25.0

Notes: BR, TIC, UPE, UE, and YE stand for blocking with regression, treatment interacted with covariates, university times period effect, university effect and year effect; North, Center and South denote the region of residence of the student; Different region (Same region) denotes that student region of residence and the region of the university are different (the same); HS degree, TS degree, and Other degree respectively denote the possession of a high school degree, a technical degree and other type of degree; Scores 1-5 indicate that the average high school score was in each of the 5 following intervals: 0-69, 70-79, 80-89, 90-99, 100.

Table 14: Mean first year grades with a grant, all models

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Total	25.1	25.1	25.1	25.1	25.6	25.6	25.5	25.5	25.1	25.1
Female	25.4	25.4	25.4	25.4	25.8	25.9	25.7	25.7	25.4	25.4
Male	24.6	24.6	24.6	24.6	25.1	25.1	24.9	25.0	24.6	24.6
Foreign	24.2	24.3	24.2	24.2	24.6	24.7	24.5	24.6	24.2	24.2
Native	25.1	25.1	25.1	25.1	25.6	25.6	25.5	25.5	25.1	25.1
North	25.2	25.2	25.2	25.2	25.7	25.7	25.6	25.6	25.2	25.2
Center	25.6	25.5	25.6	25.5	26.1	26.1	26.0	26.0	25.6	25.6
South	25.0	25.0	24.9	25.0	25.4	25.5	25.3	25.3	24.9	24.9
Different region	24.8	24.8	24.8	24.8	25.3	25.2	25.2	25.1	24.8	24.8
Same region	25.1	25.2	25.1	25.1	25.6	25.7	25.5	25.5	25.1	25.1
HS degree	25.5	25.5	25.5	25.5	26.0	26.0	25.9	25.9	25.5	25.5
TS degree	24.4	24.5	24.4	24.5	24.8	24.9	24.7	24.8	24.4	24.4
Other degree	24.6	24.6	24.6	24.6	25.0	25.1	24.9	25.0	24.6	24.6
Score 1	23.7	23.7	23.6	23.7	24.0	24.1	23.9	24.0	23.6	23.6
Score 2	24.2	24.2	24.2	24.2	24.6	24.6	24.5	24.5	24.2	24.2
Score 3	24.8	24.8	24.8	24.8	25.3	25.3	25.2	25.2	24.8	24.8
Score 4	25.4	25.5	25.4	25.5	25.9	26.0	25.8	25.9	25.4	25.4
Score 5	26.6	26.5	26.5	26.5	27.2	27.2	27.1	27.0	26.6	26.6
Urban	25.3	25.3	25.2	25.3	25.8	25.8	25.6	25.7	25.2	25.2
Rural	25.0	25.0	25.0	25.0	25.4	25.4	25.3	25.3	25.0	25.0

Notes: BR, TIC, UPE, UE, and YE stand for blocking with regression, treatment interacted with covariates, university times period effect, university effect and year effect; North, Center and South denote the region of residence of the student; Different region (Same region) denotes that student region of residence and the region of the university are different (the same); HS degree, TS degree, and Other degree respectively denote the possession of a high school degree, a technical degree and other type of degree; Scores 1-5 indicate that the average high school score was in each of the 5 following intervals: 0-69, 70-79, 80-89, 90-99, 100.

Table 15: Mean gain in first year grades from a grant, all models

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Total	0.0	0.0	0.0	0.0	-0.1	-0.1	-0.1	-0.1	0.0	0.0
Female	0.0	0.0	0.0	-0.1	-0.1	-0.1	-0.1	-0.1	0.0	0.0
Male	0.0	0.1	0.0	0.0	-0.1	0.0	-0.1	0.0	0.0	0.0
Foreign	0.0	0.2	0.0	0.2	-0.1	0.3	-0.1	0.2	0.0	-0.1
Native	0.0	0.0	0.0	0.0	-0.1	-0.1	-0.1	-0.1	0.0	0.0
North	0.0	0.0	0.0	-0.1	-0.1	-0.1	-0.1	-0.1	0.0	0.0
Center	0.0	-0.1	0.0	0.0	-0.1	-0.1	-0.1	-0.1	0.0	0.0
South	0.0	0.0	0.0	0.0	-0.1	0.0	-0.1	0.0	0.0	0.0
Different region	0.0	-0.1	0.0	-0.1	-0.1	-0.2	-0.1	-0.2	0.0	0.0
Same region	0.0	0.0	0.0	0.0	-0.1	0.0	-0.1	0.0	0.0	0.0
HS degree	0.0	-0.1	0.0	-0.1	-0.1	-0.1	-0.1	-0.1	0.0	0.0
TS degree	0.0	0.1	0.0	0.1	-0.1	0.0	-0.1	0.1	0.0	0.0
Other degree	0.0	0.0	0.0	0.0	-0.1	0.0	-0.1	0.0	0.0	0.0
Score 1	0.0	0.0	0.0	0.0	-0.1	0.0	-0.1	0.0	0.0	0.0
Score 2	0.0	0.0	0.0	0.0	-0.1	0.0	-0.1	0.0	0.0	0.0
Score 3	0.0	0.0	0.0	0.0	-0.1	-0.1	-0.1	-0.1	0.0	0.0
Score 4	0.0	0.0	0.0	0.0	-0.1	0.0	-0.1	0.0	0.0	0.0
Score 5	0.0	-0.1	0.0	-0.1	-0.1	-0.1	-0.1	-0.1	0.0	0.0
Urban	0.0	0.0	0.0	0.0	-0.1	0.0	-0.1	0.0	0.0	0.0
Rural	0.0	0.0	0.0	-0.1	-0.1	-0.1	-0.1	-0.1	0.0	0.0

Notes: BR, TIC, UPE, UE, and YE stand for blocking with regression, treatment interacted with covariates, university times period effect, university effect and year effect; North, Center and South denote the region of residence of the student; Different region (Same region) denotes that student region of residence and the region of the university are different (the same); HS degree, TS degree, and Other degree respectively denote the possession of a high school degree, a technical degree and other type of degree; Scores 1-5 indicate that the average high school score was in each of the 5 following intervals: 0-69, 70-79, 80-89, 90-99, 100.

Table 16: Mean first year credits without a grant, all models

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Total	33.5	33.4	33.4	33.4	18.3	18.9	24.1	24.4	31.1	31.1
Female	34.5	34.6	34.4	34.6	19.7	20.7	25.6	26.3	32.0	32.0
Male	31.7	31.2	31.6	31.2	15.9	15.7	21.6	21.1	29.5	29.5
Foreign	33.8	32.7	33.7	32.7	18.5	17.4	24.3	23.0	29.5	29.6
Native	33.5	33.4	33.4	33.4	18.3	18.9	24.1	24.4	31.1	31.1
North	37.9	38.3	37.8	38.6	23.4	25.8	29.4	31.2	34.2	34.3
Center	33.8	33.3	33.7	33.2	17.8	17.3	23.8	23.1	31.7	31.8
South	31.6	31.4	31.5	31.2	16.4	16.4	22.1	21.9	29.6	29.7
Different region	37.2	36.7	37.1	36.9	24.4	24.4	29.7	29.7	33.3	33.4
Same region	32.9	32.8	32.8	32.8	17.3	18.0	23.2	23.5	30.7	30.7
HS degree	35.7	35.8	35.6	35.7	21.8	22.6	27.6	28.0	33.4	33.5
TS degree	30.2	29.7	30.1	29.8	13.5	13.5	19.3	19.1	27.7	27.7
Other degree	30.6	30.5	30.5	30.6	13.1	14.0	19.2	19.7	27.9	28.0
Score 1	23.4	22.7	23.3	22.8	5.6	5.6	10.4	10.1	20.6	20.7
Score 2	29.3	28.9	29.2	28.9	12.1	12.1	18.1	17.9	26.6	26.6
Score 3	33.7	33.6	33.6	33.6	18.3	18.6	24.4	24.5	31.0	31.0
Score 4	36.4	36.5	36.3	36.5	22.2	23.0	28.3	28.7	34.0	34.0
Score 5	41.0	41.3	40.9	41.3	28.8	30.4	34.5	35.5	39.3	39.4
Urban	33.8	33.8	33.7	33.8	18.6	19.3	24.4	24.8	31.6	31.6
Rural	33.3	33.1	33.2	33.1	18.1	18.6	23.9	24.1	30.7	30.7

Notes: BR, TIC, UPE, UE, and YE stand for blocking with regression, treatment interacted with covariates, university times period effect, university effect and year effect; North, Center and South denote the region of residence of the student; Different region (Same region) denotes that student region of residence and the region of the university are different (the same); HS degree, TS degree, and Other degree respectively denote the possession of a high school degree, a technical degree and other type of degree; Scores 1-5 indicate that the average high school score was in each of the 5 following intervals: 0-69, 70-79, 80-89, 90-99, 100.

Table 17: Mean first year credits with a grant, all models

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Total	35.8	35.9	35.9	36.0	25.1	24.9	29.5	29.5	35.1	35.1
Female	36.8	36.5	36.9	36.7	26.9	25.9	31.1	30.5	36.1	36.1
Male	34.0	34.6	34.1	34.8	21.9	23.1	26.5	27.7	33.4	33.3
Foreign	36.1	36.4	36.2	36.5	25.2	26.0	29.7	30.2	35.8	35.8
Native	35.8	35.8	35.9	36.0	25.1	24.9	29.5	29.5	35.1	35.1
North	40.2	39.8	40.2	39.6	30.9	29.1	35.1	33.8	39.8	39.8
Center	36.1	36.1	36.2	36.1	24.3	24.0	29.0	28.8	35.7	35.7
South	33.9	34.2	34.0	34.5	22.9	23.4	27.3	27.9	33.1	33.1
Different region	39.5	39.7	39.6	39.6	32.0	32.0	35.4	35.4	39.1	39.1
Same region	35.2	35.2	35.3	35.4	24.0	23.7	28.5	28.6	34.5	34.5
HS degree	38.0	37.8	38.1	38.0	29.4	28.6	33.3	32.8	37.3	37.3
TS degree	32.5	33.1	32.6	33.1	19.1	19.9	24.2	25.0	32.0	31.9
Other degree	32.9	33.3	33.0	33.3	18.7	19.0	24.1	24.6	32.3	32.3
Score 1	25.7	26.3	25.7	26.3	8.8	9.3	14.0	14.8	25.2	25.2
Score 2	31.6	31.9	31.7	32.0	17.7	18.1	23.1	23.6	31.0	31.0
Score 3	36.0	35.9	36.0	36.1	25.4	25.2	30.0	29.9	35.4	35.4
Score 4	38.7	38.9	38.8	39.1	30.1	30.4	34.1	34.5	38.0	38.0
Score 5	43.3	42.8	43.4	43.0	37.6	35.9	40.6	39.5	42.4	42.3
Urban	36.1	36.2	36.2	36.3	25.4	25.2	29.7	29.8	35.3	35.3
Rural	35.6	35.6	35.6	35.8	24.9	24.7	29.3	29.3	35.0	35.0

Notes: BR, TIC, UPE, UE, and YE stand for blocking with regression, treatment interacted with covariates, university times period effect, university effect and year effect; North, Center and South denote the region of residence of the student; Different region (Same region) denotes that student region of residence and the region of the university are different (the same); HS degree, TS degree, and Other degree respectively denote the possession of a high school degree, a technical degree and other type of degree; Scores 1-5 indicate that the average high school score was in each of the 5 following intervals: 0-69, 70-79, 80-89, 90-99, 100.

Table 18: Mean gain in first year credits from a grant, all models

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Total	2.3	2.5	2.5	2.6	6.8	6.0	5.3	5.1	4.1	4.0
Female	2.3	1.9	2.5	2.1	7.2	5.2	5.6	4.3	4.2	4.1
Male	2.3	3.5	2.5	3.7	6.0	7.4	5.0	6.6	3.9	3.8
Foreign	2.3	3.7	2.5	3.8	6.7	8.5	5.4	7.3	6.3	6.2
Native	2.3	2.4	2.5	2.6	6.8	5.9	5.3	5.1	4.0	4.0
North	2.3	1.4	2.5	1.0	7.5	3.3	5.6	2.6	5.6	5.5
Center	2.3	2.8	2.5	2.8	6.5	6.6	5.3	5.7	4.0	3.9
South	2.3	2.8	2.5	3.3	6.5	7.0	5.3	6.0	3.4	3.4
Different region	2.3	2.9	2.5	2.7	7.6	7.6	5.7	5.7	5.8	5.7
Same region	2.3	2.4	2.5	2.6	6.6	5.7	5.3	5.0	3.8	3.7
HS degree	2.3	2.0	2.5	2.3	7.6	6.0	5.7	4.8	3.9	3.8
TS degree	2.3	3.3	2.5	3.4	5.7	6.4	4.9	5.9	4.3	4.2
Other degree	2.3	2.8	2.5	2.7	5.6	5.0	4.9	4.8	4.4	4.3
Score 1	2.3	3.6	2.5	3.6	3.2	3.7	3.7	4.7	4.6	4.5
Score 2	2.3	3.0	2.5	3.1	5.6	6.0	5.0	5.7	4.5	4.4
Score 3	2.3	2.3	2.5	2.5	7.2	6.6	5.6	5.4	4.4	4.3
Score 4	2.3	2.5	2.5	2.6	7.9	7.4	5.8	5.8	4.0	4.0
Score 5	2.3	1.5	2.5	1.6	8.8	5.5	6.0	4.0	3.0	3.0
Urban	2.3	2.4	2.5	2.6	6.7	5.8	5.3	5.0	3.8	3.7
Rural	2.3	2.5	2.5	2.7	6.8	6.1	5.4	5.2	4.3	4.2

Notes: BR, TIC, UPE, UE, and YE stand for blocking with regression, treatment interacted with covariates, university times period effect, university effect and year effect; North, Center and South denote the region of residence of the student; Different region (Same region) denotes that student region of residence and the region of the university are different (the same); HS degree, TS degree, and Other degree respectively denote the possession of a high school degree, a technical degree and other type of degree; Scores 1-5 indicate that the average high school score was in each of the 5 following intervals: 0-69, 70-79, 80-89, 90-99, 100.

Table 19: Mean probability of graduating in time without a grant, all models

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Total	34.9	34.5	35.6	35.3	35.8	35.6	35.7	35.6	31.8	31.9
Female	36.8	36.6	37.4	37.4	37.6	37.8	37.6	37.7	33.5	33.6
Male	31.5	30.6	32.2	31.5	32.4	31.7	32.3	31.7	28.7	28.8
Foreign	32.4	31.5	33.3	33.3	33.6	33.8	33.6	33.6	26.4	26.5
Native	35.0	34.5	35.6	35.4	35.8	35.6	35.8	35.6	31.9	32.0
North	46.3	45.5	47.2	47.0	47.4	48.0	47.3	47.9	41.3	41.4
Center	36.7	37.1	37.5	37.1	37.7	37.5	37.7	37.4	34.6	34.7
South	29.8	29.4	30.4	30.1	30.5	30.1	30.5	30.1	27.3	27.3
Different region	42.4	39.6	43.3	41.5	43.5	42.4	43.5	42.3	37.3	37.4
Same region	33.7	33.6	34.3	34.3	34.5	34.5	34.4	34.4	30.9	31.0
HS degree	38.0	37.7	38.6	38.4	38.8	38.7	38.7	38.7	35.0	35.1
TS degree	30.1	29.3	30.8	30.3	31.0	30.4	31.0	30.4	26.8	26.9
Other degree	31.6	31.3	32.3	32.4	32.5	32.6	32.4	32.6	28.1	28.2
Score 1	18.1	18.4	18.9	19.2	19.2	19.0	19.1	19.0	14.6	14.7
Score 2	26.7	26.2	27.4	27.1	27.7	27.1	27.6	27.1	23.2	23.3
Score 3	33.5	33.0	34.3	34.1	34.4	34.4	34.4	34.4	30.1	30.1
Score 4	39.9	39.0	40.5	40.0	40.7	40.5	40.7	40.5	36.8	36.8
Score 5	50.0	49.7	50.5	50.2	50.6	50.7	50.6	50.6	47.9	48.0
Urban	35.5	35.3	36.1	36.1	36.3	36.5	36.3	36.4	32.7	32.8
Rural	34.4	33.9	35.1	34.7	35.3	34.9	35.3	34.9	31.1	31.2

Notes: BR, TIC, UPE, UE, and YE stand for blocking with regression, treatment interacted with covariates, university times period effect, university effect and year effect; North, Center and South denote the region of residence of the student; Different region (Same region) denotes that student region of residence and the region of the university are different (the same); HS degree, TS degree, and Other degree respectively denote the possession of a high school degree, a technical degree and other type of degree; Scores 1-5 indicate that the average high school score was in each of the 5 following intervals: 0-69, 70-79, 80-89, 90-99, 100.

Table 20: Mean probability of graduating in time with a grant, all models

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Total	37.3	37.3	36.7	36.7	36.5	36.6	36.5	36.6	36.1	36.1
Female	39.2	38.9	38.6	38.3	38.4	38.2	38.4	38.2	38.0	38.0
Male	33.9	34.4	33.3	33.7	33.1	33.7	33.1	33.6	32.8	32.7
Foreign	34.8	35.3	34.4	34.4	34.3	34.3	34.3	34.3	34.2	34.2
Native	37.4	37.4	36.8	36.7	36.5	36.6	36.5	36.6	36.2	36.2
North	48.7	48.8	48.3	48.1	48.2	47.6	48.1	47.6	48.0	48.0
Center	39.1	38.7	38.6	38.7	38.5	38.4	38.4	38.3	38.4	38.4
South	32.2	32.2	31.5	31.5	31.2	31.7	31.2	31.6	30.7	30.7
Different region	44.8	46.1	44.4	45.2	44.3	44.7	44.3	44.7	44.2	44.2
Same region	36.1	35.9	35.4	35.3	35.2	35.3	35.2	35.2	34.8	34.8
HS degree	40.4	40.0	39.7	39.5	39.5	39.3	39.5	39.3	39.2	39.1
TS degree	32.5	33.0	31.9	32.3	31.7	32.3	31.7	32.3	31.4	31.4
Other degree	34.0	34.4	33.4	33.6	33.2	33.6	33.2	33.6	32.9	32.9
Score 1	20.6	20.4	20.0	19.9	19.8	20.2	19.7	20.1	19.6	19.6
Score 2	29.2	29.3	28.6	28.7	28.3	28.8	28.3	28.7	28.1	28.1
Score 3	36.0	36.1	35.4	35.4	35.2	35.3	35.2	35.3	34.9	34.9
Score 4	42.3	42.8	41.7	42.0	41.5	41.9	41.5	41.9	41.1	41.0
Score 5	52.4	51.7	51.6	51.2	51.4	50.7	51.4	50.7	50.9	50.8
Urban	38.0	38.0	37.3	37.3	37.1	37.2	37.1	37.2	36.7	36.7
Rural	36.9	36.8	36.3	36.2	36.1	36.2	36.0	36.1	35.7	35.7

Notes: BR, TIC, UPE, UE, and YE stand for blocking with regression, treatment interacted with covariates, university times period effect, university effect and year effect; North, Center and South denote the region of residence of the student; Different region (Same region) denotes that student region of residence and the region of the university are different (the same); HS degree, TS degree, and Other degree respectively denote the possession of a high school degree, a technical degree and other type of degree; Scores 1-5 indicate that the average high school score was in each of the 5 following intervals: 0-69, 70-79, 80-89, 90-99, 100.

Table 21: Mean gain in probability of graduating in time from a grant, all models

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Total	2.4	2.8	1.1	1.4	0.7	1.0	0.8	1.0	4.3	4.3
Female	2.4	2.3	1.1	0.9	0.8	0.5	0.8	0.5	4.5	4.4
Male	2.4	3.8	1.1	2.1	0.7	2.0	0.7	2.0	4.1	4.0
Foreign	2.4	3.8	1.1	1.1	0.7	0.5	0.7	0.7	7.8	7.7
Native	2.4	2.8	1.1	1.4	0.7	1.0	0.8	1.0	4.3	4.2
North	2.4	3.4	1.1	1.1	0.8	-0.3	0.8	-0.3	6.8	6.6
Center	2.4	1.7	1.1	1.6	0.8	0.9	0.8	0.9	3.8	3.7
South	2.4	2.8	1.1	1.4	0.7	1.6	0.7	1.6	3.5	3.4
Different region	2.4	6.4	1.1	3.7	0.8	2.4	0.8	2.5	6.9	6.8
Same region	2.4	2.2	1.1	1.0	0.7	0.8	0.8	0.8	3.9	3.8
HS degree	2.4	2.3	1.1	1.1	0.8	0.6	0.8	0.6	4.1	4.0
TS degree	2.4	3.8	1.1	2.0	0.7	1.9	0.7	1.9	4.6	4.5
Other degree	2.4	3.2	1.1	1.2	0.7	1.0	0.7	1.0	4.8	4.7
Score 1	2.4	2.0	1.1	0.6	0.6	1.2	0.6	1.1	5.0	4.9
Score 2	2.4	3.0	1.1	1.6	0.7	1.7	0.7	1.6	4.9	4.8
Score 3	2.4	3.1	1.1	1.3	0.8	0.9	0.8	0.9	4.8	4.7
Score 4	2.4	3.8	1.1	2.1	0.8	1.3	0.8	1.4	4.3	4.2
Score 5	2.4	2.0	1.1	0.9	0.8	0.1	0.8	0.1	2.9	2.9
Urban	2.4	2.7	1.1	1.2	0.7	0.7	0.8	0.7	4.0	3.9
Rural	2.4	2.9	1.1	1.5	0.7	1.2	0.8	1.2	4.6	4.5

Notes: BR, TIC, UPE, UE, and YE stand for blocking with regression, treatment interacted with covariates, university times period effect, university effect and year effect; North, Center and South denote the region of residence of the student; Different region (Same region) denotes that student region of residence and the region of the university are different (the same); HS degree, TS degree, and Other degree respectively denote the possession of a high school degree, a technical degree and other type of degree; Scores 1-5 indicate that the average high school score was in each of the 5 following intervals: 0-69, 70-79, 80-89, 90-99, 100.