Evaluating the Effect of University Grants on Student Dropout: Evidence from a Regression Discontinuity Design Using Bayesian Principal Stratification Analysis

Fan Li        Alessandra Mattei        Fabrizia Mealli

August 6, 2014

ABSTRACT

Regression discontinuity (RD) designs are often interpreted as local randomized experiments: a RD design can be considered as a randomized experiment for units with a realized value of a so-called forcing variable falling around a pre-fixed threshold. Motivated by the evaluation of Italian university grants, we consider a fuzzy RD design where the receipt of the treatment is based on both eligibility criteria and a voluntary application status. Resting on the fact that grant application and grant receipt statuses are post-assignment (eligibility) intermediate variables, we use the principal stratification framework to define causal estimands within the Rubin Causal Model. We then propose a probabilistic formulation of the assignment mechanism underlying RD designs, by re-formulating the Stable Unit Treatment Value Assumption (SUTVA) and making an explicit local overlap assumption for a subpopulation around the threshold. A local randomization assumption is invoked instead of standard continuity assumptions. We also develop a model-based Bayesian approach to select the target subpopulation(s) with adjustment for multiple comparisons, and to draw inference for the target causal estimands in this framework. Applying the method to the data from two Italian universities, we find evidence that university grants are effective in preventing students from low-income families from dropping out of higher education.

KEY WORDS: Bayesian, causal effects, instrumental variable, principal stratification, randomization, regression discontinuity, university grants

1Fan Li is assistant professor, Department of Statistical Science, Duke University, Durham, NC (email: fli@stat.duke.edu); Alessandra Mattei is assistant professor (email: mattei@disia.unifi.it) and Fabrizia Mealli is professor (email: mealli@disia.unifi.it), Department of Statistics, Computer Science and Applications, University of Florence, Italy. The authors are grateful to Quanli Wang for computing support, to Guido Imbens and Sid Chib for helpful comments. Li’s research is partially funded by NSF-SES grant 11-55697. Part of this paper was written when Mattei was a senior fellow of the Uncertainty Quantification program of the U.S. Statistical and Applied Mathematical Sciences Institute (SAMSI).
1 Introduction

The regression discontinuity (RD) design is a quasi-experimental design, often arising in economics and other fields, which can be exploited to identify and estimate causal effects of interventions. Originally introduced in the psychology literature by Thistlethwaite and Campbell (1960), RD designs did not attract much attention until recently (see Cook, 2008, for a historical perspective). Since the late 1990s there has been a large number of studies, mainly in economics, applying and extending RD methods. Recent surveys can be found in Imbens and Lemieux (2008); van der Klaauw (2008); Lee and Lemieux (2010).

There are two general setups in RDs, the sharp and the fuzzy RDs. In the sharp RD design, the original form of the design, the treatment status is assumed to be a deterministic step function of a so-called forcing variable or running variable. All units with a realized value of the forcing variable on one side of a pre-fixed threshold are assigned to one regime and all units on the other side are assigned to the other regime. The basic idea underlying a RD analysis is that one can compare units with very similar values for the forcing variable, but different levels of treatment, to draw inference on the causal effect of the treatment at the threshold. Examples of sharp RD designs can be found, among others, in Berk and Rauma (1983); Berk and de Leuuw (1999); Lee et al. (2004); Lee (2008); Mealli and Rampichini (2012). In the fuzzy RD design, the realized value of the forcing variable does not alone determine the receipt of the treatment, although a value of the forcing variable falling above or below the threshold acts as an encouragement or incentive to participate in the treatment. In those cases, the receipt of the treatment depends also on individual choices, which may confound treatment receipt. Hahn et al. (2001) establish a connection between fuzzy RD designs and the instrumental variable (IV) settings, and show that in a fuzzy RD setting one can identify the local average treatment effect (Imbens and Angrist, 1994) for a subpopulation of compliers at the threshold. Examples of fuzzy RD designs can be found, among others, in van der Klaauw (2002); Battistin and Rettore (2008); Garibaldi et al. (2012).

Our study is motivated by the evaluation of Italian university grants. Italian State Universities offer financial aid every year to a limited number of eligible freshmen. The main objective of
this intervention is to give equal opportunity to achieve higher education to motivated students irrespective of their economic background. Dropout from university is a relevant phenomenon in Italy: indeed, the low rate of university graduates among Italian youths is mainly due to the high dropout rate rather than to a low enrollment rate. Amid the recent economic crisis in Europe, there has been a heated debate on how to arrange student financial support, especially in terms of the instruments used, e.g., loans, grants, tuition waiver. Accurate evaluation of the effectiveness of the existing financial aid systems is crucial for providing information to policy makers to choose between different instruments. Therefore, our main substantive goal in this paper is to investigate the effects of Italian university grants on preventing students belonging to low-income families from dropping out of higher education, using data on first-year enrollees from two universities.

In the Italian university system, allocation of the grant is based on eligibility criteria and a voluntary application status: only students who both meet the eligibility criteria and apply for a grant can receive the grant. The eligibility status depends on an economic measurement of the student’s family income and assets (forcing variable) falling below or above a pre-determined threshold. This rule defines a fuzzy RD design in the sense that not all the eligible students, who have a value of the forcing variable falling below the threshold, get a grant: eligible students must apply to receive a grant. The application is voluntary, and also ineligible students may apply, even if they will not receive any grant. Ineligible students typically apply either because they are not fully aware of their eligibility status, or because they hope that their application will be still considered because of extra funding or other considerations. Comparing to standard fuzzy RD designs where only assignment (eligibility) and receipt of the treatment (grant) are available, the additional data on application status in this study can provide valuable information with important policy implications. In this article, we will show how to capitalize on the application data to draw inference on causal effects. In particular, we will frame the RD design in the context of the Rubin Causal Model (RCM) (Rubin, 1974, 1978) using potential outcomes. Resting on the fact that grant application and grant receipt statuses are post-assignment (eligibility) intermediate variables, we adopt the principal stratification framework (Frangakis and Rubin, 2002)—a
generalization of the IV approach to noncompliance (Angrist et al., 1996; Imbens and Rubin, 1997)—to define causal estimands and lay the basis for inference.

In the literature, causal inference in RD designs is usually based on comparisons of units with close but distinct values of the forcing variable and relies on smoothness assumptions about the relationship between outcomes and the forcing variable around the threshold, which imply randomization at the single threshold value. For example, in fuzzy RDs, estimands are usually specified as ratio of differences of regression functions at the threshold, and inference generally relies on asymptotic approximations (e.g. Imbens and Lemieux, 2008). In real applications, the large-sample approximations might be unreliable due to the small sample size, and exact inference would be preferable. RD designs have been often described as designs that lead to locally randomized experiments around the threshold (Lee, 2008; Lee and Lemieux, 2010; Dinardo and Lee, 2011). Expanding on this interpretation, a recent strand of the literature (e.g. Cattaneo et al., 2014; Sales and Hansen, 2014) is moving towards a formal and well-structured definition of the conditions under which RD designs are equivalent to local randomized experiments, aiming at developing randomization-based approaches to inference that may circumvent potential small sample concerns.

We further develop the idea of local randomization; our goal is to provide a formal definition of the hypothetical experiment underlying RD designs, based on a description of the assignment mechanism, i.e., the process that describes why some units got assigned to different treatments, formalized as a unit-exchangeable stochastic function of covariates and potential outcomes. Our approach presents subtle but important differences with the methodological framework proposed by Cattaneo et al. (2014) and Sales and Hansen (2014). In particular, we develop a framework for RD analysis that is fully consistent with the RCM, by separating and defining the critical assumptions within the framework of principal stratification. We first define treatments and potential outcomes, by clearly re-formulating SUTVA to make our representation of potential outcomes adequate. Then, we introduce a probabilistic formulation of the assignment mechanism for general RD settings. The core of our framework is to assume there exists at least one subpopulation around the threshold where a local overlap assumption holds. For this subpopulation we explic-
itly introduce a *local randomization* assumption. Finally we invoke some structural assumptions, which may help to sharpen inference, and discuss their implications and plausibility.

Alternative to the randomization-based approach used in Cattaneo et al. (2014) and Sales and Hansen (2014), in this article we adopt the Bayesian approach to select the unknown subpopulation(s) and to draw inference on causal effects, for the following reasons. First, causal inference in RD designs usually involves complex observational data, with multiple sources of uncertainties, including the missing potential outcomes; the Bayesian approach is particularly useful for accounting for uncertainties and for pooling information from the data in such complex settings. Second, RD analysis usually relies on a sample of units with values of the forcing variable close to a single point, the size of which may be small; Bayesian methods, not relying on asymptotic approximations, are attractive in dealing with small samples. Third, in the Bayesian paradigm, the missing potential outcomes are treated as random variables, and all inferences are based on the posterior distributions of causal estimands, which are functions of potential outcomes. Thus inference about finite-sample and super-population estimands can be drawn using the same inferential procedures. Finally, pre-treatment variables can be easily incorporated in the Bayesian approach, which may improve efficiency of the analysis, i.e., reduce posterior variability.

Despite these advantages and the existing literature in Bayesian causal inference (e.g., Rubin, 1978; Imbens and Rubin, 1997; Barnard et al., 2003; Li et al., 2009; Elliott et al., 2010; Schwartz et al., 2011; Mattei et al., 2013), Bayesian methods have rarely been discussed in the RD literature (exceptions are Chib and Jacobi, 2011; Chib and Greenberg, 2014). Therefore, another main thrust of the paper is to develop a Bayesian approach to conduct exact inference under the fuzzy RD design considered here, which can be easily adapted to other RD settings. In addition, we develop a Bayesian hierarchical modeling approach to adjust for multiple comparisons in selecting the target subpopulation(s). Systematic model evaluation via posterior predictive checks is also carried out.

In Section 2, we introduce the basic setup and the causal estimands within the principal stratification framework. In Section 3, we propose a probabilistic formulation of the assignment mechanism for general RD designs, explicitly formulating the key assumptions, and elaborate it
for the particular RD design used in the Italian university grants. Selection of the subpopulations
where these assumptions hold is discussed. A Bayesian approach for inferring causal effects in
RD designs is developed in Section 4. The proposed approach is then applied to evaluate causal
effects of Italian university grants on student dropout in Section 5. Section 6 concludes.

2 Causal Estimands

2.1 Basic Setup

We introduce the notation in the context of Italian university grants. Our goal is to provide a
framework within principal stratification that allows us to specify well-defined causal estimands,
and to properly interpret them as causal effects of the grant. This formulation also makes it clear
which information can be drawn from the data about the causal estimands of interest, allowing
us to explicitly define the critical assumptions later.

Let $Z$ be the eligibility status, which is the initial assignment and plays the role of an “in-
strument” or an “encouragement” as in randomized experiments with noncompliance. Consider
a sample or population of $N$ units; each can be either eligible to receive a treatment, $z = 1$, or
ineligible, $z = 0$. In the Italian grants system, eligibility depends on the value of a combined
measurement of one’s assets including income and properties, adjusted for family size, denoted
by $S$. If a student, satisfying preliminary grade criteria, has a value of $S$ falling below a pre-
determined threshold, e.g. $s_0 = 15\,000$ euro, he/she is eligible, and not otherwise. That is, the
eligibility status $Z_i$ for student $i$ is a deterministic function of $S$: $Z_i = 1(S_i \leq s_0)$, where $1(\cdot)$ is
the indicator function. Using the terminology in RD designs, $S$ is the forcing variable.

All variables measured after each unit $i$ is assigned eligibility $Z_i$, namely, the application
status, the receipt of the grant and the dropout status, are post-assignment variables, and, in
principle, eligibility may affect them. In the RCM, potential outcomes need to be defined for
each post-assignment variable as functions of the assignment in order to define causal estimands.
Specifically, for each student $i$ ($i = 1, \ldots, N$), given eligibility status $z$ ($z = 0, 1$), let $A_i(z)$
be an indicator for the potential grant application status (equal to 1 if student $i$ applies for a
grant and 0 otherwise), $W_i(z)$ be an indicator for the potential treatment received (equal to 1 if student $i$ receives a grant and 0 otherwise), and $Y_i(z)$ be the potential indicator for dropout (1 if student $i$ drops out of university, 0 otherwise). Therefore, in our study, for each student $i$ there are six potential outcomes $A_i(0), A_i(1), W_i(0), W_i(1), Y_i(0), Y_i(1)$. This representation of potential outcomes requires assumptions for it to be adequate, and these assumptions need to account for the eligibility rule.

A key feature of RD designs in general, and our RD design in particular, is that eligibility is a deterministic function of the forcing variable, $S$. Therefore the above representation of potential outcomes is adequate only under a modified ‘Stable Unit Treatment Value Assumption’ (SUTVA, Rubin, 1980), which rules out interference between units and assumes that potential outcomes depend on the forcing variable only through eligibility indicators. Formally,

**Assumption 1** (RD-SUTVA). Consider two eligibility statuses $Z'_i = 1(S'_i \leq s_0)$ and $Z''_i = 1(S''_i \leq s_0)$, with possibly $S'_i \neq S''_i$. If $Z'_i = Z''_i$, that is, if either $S'_i \leq s_0$ and $S''_i \leq s_0$, or $S'_i > s_0$ and $S''_i > s_0$, then $A_i(Z') = A_i(Z'')$, $W_i(Z') = W_i(Z'')$, and $Y_i(Z') = Y_i(Z'')$.

The no interference aspect of RD-SUTVA states that potential outcomes for a student cannot be affected by the eligibility status of other students. RD-SUTVA also assumes that there are no levels of the eligibility status other than zero and one. This component of RD-SUTVA is subtle. It amounts to assuming that for each unit $i$, $A_i(z), W_i(z)$ and $Y_i(z)$ are stable in the sense that they do not depend on the realized value of the forcing variable. RD-SUTVA implies that values of the forcing variable leading to the same eligibility status cannot alter potential outcomes for any unit, thus allows us to avoid defining potential outcomes as functions of the forcing variable. Therefore under RD-SUTVA for each unit there exist only two potential outcomes for each post-assignment variable, that is, potential outcomes for a given post-assignment variable are the value of that post-assignment variable if the realized value of the forcing variable fell below the threshold and its value if the realized value of the forcing variable fell above the threshold.

RD-SUTVA may not be plausible for the whole study population and its plausibility depends on the support of $S$ for each unit. However, it appears to be plausible for subpopulations of units who have a relatively large probability that the realized values of $S$ fall in a neighborhood around
For each unit, \(i\), given the observed eligibility status \(Z_i\), the following variables are observed: \(A_{i}^{obs} = A_i(Z_i)\), the observed application status; \(W_{i}^{obs} = W_i(Z_i)\), the observed treatment received; and \(Y_{i}^{obs} = Y_i(Z_i)\), the observed dropout status. Potential outcomes under the eligibility status not realized, \(1 - Z_i\), are missing: \(A_{i}^{mis} = A_i(1 - Z_i)\), \(W_{i}^{mis} = W_i(1 - Z_i)\), and \(Y_{i}^{mis} = Y_i(1 - Z_i)\). A vector of \(p\) pre-treatment variables, \(X_i\), is also observed for each unit. We use boldface upper-case letters to denote the vector of a variable of all units from hereon. For example, 
\[
Z = (Z_1, \ldots, Z_N)' , \quad A^{obs} = (A^{obs}_1, \ldots, A^{obs}_N)' . \]

The \(N \times p\) matrix \(X\) has \(i\)th row equal to \(X_i\).

### 2.2 The Role of Principal Stratification for Causal Inference in RD Designs

In the RCM, a causal effect is defined as a comparison of the potential outcomes \(Y_i(1)\) and \(Y_i(0)\), e.g., \(E[Y_i(1) - Y_i(0)]\), for a common set of units. Obviously, in our study, such comparisons between potential dropout statuses only measure the effect of the eligibility status. To draw inference about the causal effect of receiving a grant, the main target of the study, additional structures and assumptions are required. Moreover, both the application status and receipt of the grant are post-assignment intermediate variables, lying on the causal pathway between the assignment and the outcome. It is crucial to properly adjust for these intermediate variables in the causal analysis, for which we adopt Principal Stratification (PS, Frangakis and Rubin, 2002)—a main framework for post-assignment variable adjustment in the RCM.

For each intermediate variable, Principal Stratification defines a cross-classification of subjects by the joint potential values of that intermediate variable under each of the assignments being compared, i.e., principal strata. In our study, based on the application status \(A\), subjects are classified into four (latent) principal strata, \(G_i \equiv (A_i(0), A_i(1))\): compliant-applicants \(G_i = (0, 1) = CA\), students who would not apply if ineligible, but would apply if eligible; always-applicants \(G_i = (1, 1) = AA\), students who would apply irrespective of their eligibility status; never-applicants \(G_i = (0, 0) = NA\), students who would not apply irrespective of their eligibility status; and defiant-applicants \(G_i = (1, 0) = DA\), students who would not apply if eligible, but would apply if ineligible. Because principal strata are not affected by assignment,
we can define \textit{population-average} causal effects conditional on the principal strata, known as principal causal effects:

\[
\tau_g \equiv \mathbb{E}[Y_i(1) - Y_i(0) | G_i = g],
\]

for \( g = AA, CA, NA, DA \). Then the average causal effect of eligibility on dropout is a weighted average of these principal causal effects:

\[
\mathbb{E}[Y_i(1) - Y_i(0)] = \sum_{g=AA,CA,NA,DA} \pi_g \tau_g,
\]

where \( \pi_g \) is the proportion of units in principal stratum \( g \).

Never-applicants and defiant-applicants never receive a grant, so for them we always observe the outcome in the absence of the grant. By contrast, for always-applicants and compliant-applicants we can observe \( Y_i(1) \) for some eligible students who receive a grant and \( Y_i(0) \) for some other ineligible students who do not receive a grant. Therefore, always-applicants and compliant-applicants are the only groups where we can learn information about the effect of receiving a grant in this study, and thus the corresponding principal causal effects, \( \tau_{AA} \) and \( \tau_{CA} \), are the causal estimands of primary interest.

In the standard IV approach to noncompliance (Angrist et al., 1996; Imbens and Rubin, 1997) as well as standard setting of fuzzy RD designs (e.g., Imbens and Lemieux, 2008), data on application status is not utilized, either because it is not available or because it is ignored. Instead, the analysis is based on the principal strata formed by the intermediate variable of grant receipt status. Specifically, there are four principal strata based on the joint potential grant receipt statuses, \( R_i = (W_i(0), W_i(1)) \): compliers \( R_i = (0, 1) \), students who would receive the grant if eligible and would not receive if ineligible; always-takers \( R_i = (1, 1) \), student would receive the grant regardless of eligibility; never-takers \( R_i = (0, 0) \), student would not receive the grant regardless of eligibility; and defiers \( R_i = (1, 0) \), students who would not receive the grant if eligible and would receive if ineligible. And the focus is generally on the causal effect for compliers:

\[
\tau \equiv \mathbb{E}[Y_i(1) - Y_i(0) \mid R_i = (0, 1)].
\]

We now establish the connection between these two sets of principal strata. The Italian grant assignment rule implies that \( W_i(0) = 0 \) for all \( i \), as ineligible units have no access to a grant,
and \( W_i(1) = 0 \) if \( A_i(1) = 0 \), as eligible units need to apply for a grant to receive a grant. Therefore, by design, there are no always-takers or defiers, and the remaining principal strata \( R' \)'s can be expressed as unions of principal strata \( G' \)'s: never-takers comprise never-applicants and defiant-applicants, and compliers comprise always-applicants and compliant-applicants. As such, \( \tau \) can be rewritten as the weighted average of the causal effects for always-applicants and compliant-applicants:

\[
\tau = \mathbb{E}[Y_i(1) - Y_i(0) | G_i \in \{AA, CA\}] = \frac{\pi_{AA} \tau_{AA} + \pi_{CA} \tau_{CA}}{\pi_{AA} + \pi_{CA}}.
\]

This illustrates that principal strata defined by the application status in fact provide a finer partition of the units than principal strata defined by the grant-receipt status. It is clear that the standard IV causal estimand—the causal effect for compliers \( \tau \), which ignores application behavior, provides information on a ‘marginal’ (with respect to application behavior) causal effect. If causal effects are homogeneous, marginalizing over application behavior does not critically affect the evaluation analysis. Conversely, if causal effects are heterogeneous, as we can reasonably expect in this study, ignoring application behavior represents a loss of useful information with potentially important policy implications. For example, if the grants are found out to have a higher positive effect on always-applicants than compliant-applicants, then it would be useful and cost-effective to study the characteristics of ineligible applicants and include those into the eligibility rule to allocate additional resources.

The estimands \( \tau_{AA}, \tau_{CA} \) and \( \tau \) represent effects of eligibility, rather than effects of the receipt of a grant. However, “the receipt of a grant” is completely confounded with “the eligibility status”: \( W(z) = z \times A(z) = z \) for always-applicants and compliant-applicants. To attribute these effects to “the receipt of a grant”, below we can make an exclusion restriction assumption for compliant-applicants and always-applicants, following the established literature in the IV approach to noncompliance (e.g., Angrist et al., 1996; Imbens and Rubin, 1997):

**Assumption 2 (Exclusion Restriction for Compliant-Applicants and Always-Applicants).** For all units with \( G_i \in \{AA, CA\} \), or equivalently \( R_i = (0, 1) \), the effect of eligibility is only through the receipt of the grant.

\[2\] This exclusion restriction assumption could be formalized by introducing potential outcomes of the form
The notation required to formalize Assumption 2 is in Footnote 2. Assumption 2 is essentially an attribution of the intention-to-treat effect for compliers to the causal effect of the receipt of grant, rather than to its assignment (eligibility). This type of exclusion restriction is routinely made, often implicitly, in randomized experiments with full compliance (Imbens and Rubin, Forthcoming, Assumption 23.4).

In the application of Italian grants evaluation, the sample-average counterpart of the population-average estimands may also be of interest:

$$\tau_g^s \equiv \frac{1}{N_g} \sum_{i:G_i=g} [Y_i(1) - Y_i(0)],$$

(3)

where \(g = AA, CA, \{AA, CA\}\) and \(N_g\) is the number of units in stratum \(g\). Usually the sample-average effects can be estimated more precisely than their population-average counterparts. The subtle difference between them in Bayesian inference is further explained in Section 4. More details can be found, for example, in Rubin (1978); Imbens and Rubin (1997) and Imbens (2004). For simplicity of notation, we do not make the distinction in the methodological discussion, but will present both estimates in the application.

\(Y_i(z, a, w),\) that is, potential outcomes for \(Y\) if eligibility status \(Z\) were set to \(z\), application status \(A\) were set to \(a\), and grant status \(W\) were set to \(w\); \(z, a, w \in \{0, 1\}\). Specifically, in our setting, the (stochastic) exclusion restriction assumption for compliant-applicants and always-applicants would require that for each \(a’, a” \in \{0, 1\}\),

\[\Pr(Y_i(0, a’, w)|G_i \in \{AA, CA\}) = \Pr(Y_i(1, a”, w)|G_i \in \{AA, CA\}), w = 0, 1.\]

This exclusion restriction assumes that for compliant-applicants and always-applicants, i.e., compliers, the potential outcome that would realize if they were eligible and received a grant (did not receive) is equal to the potential outcome that would realize if they were ineligible and received (did not receive) a grant, irrespective of their application status. The potential outcomes \(Y_i(z, a, w)\) are a priori counterfactuals for units who exhibit a value of the application status, \(A_i^{obs}\), and a value of the grant status, \(W_i^{obs}\), under treatment \(z\) not equal to \(a\) and \(w\), respectively, because in one specific experiment, they can be never observed for such type of units. Here we prefer to avoid potential outcomes of the form \(Y_i(z, a, w)\), by focusing on observable potential outcomes \(Y_i(z)\).
3 The Basis for Inference

3.1 Probabilistic Treatment Assignment Mechanism in RD designs

The relatively complex selection process in Italian university grants system implies that the mechanism governing the receipt of the grant, which depends on both institutional and individual choices, is not ignorable. In the analysis of observational studies, an important preliminary step is to evaluate carefully the possibility and plausibility to conceptualize the observational data as having arisen from a complex randomized experiment, where the treatment assignment mechanism has been lost and must be reconstructed. Below we formally introduce a probabilistic assignment mechanism underlying the RD design considered here, which is also applicable to general RD settings with minor modifications.

The key to our proposal lies in reconstructing the hypothetical experiment underlying a RD design by viewing the forcing variable as a random variable with a probability distribution. We now first define the assignment mechanism, which is a row-exchangeable function that assigns probabilities to all \(2^N\) possible \(N\)-dimensional vectors of assignments \(Z\), as a row-exchangeable function that assigns probabilities to all possible \(N\)-dimensional vectors of realizations of the forcing variable, \(S\), above or below the threshold value, \(s_0\). Formally,

\[
\Pr (Z = z | A(0), A(1), W(0), W(1), Y(0), Y(1), X) = \Pr (S \in \Lambda | A(0), A(1), W(0), W(1), Y(0), Y(1), X),
\]

where \(z \in \{0, 1\}^n\) and \(\Lambda \in \left\{(-\infty, s_0]^n, (-\infty, s_0]^{n-1} \times (s_0, \infty), (s_0, \infty) \times (-\infty, s_0]^{n-1}, \ldots, (-\infty, s_0] \times (s_0, \infty)^{n-1}, (-\infty, s_0]^{n-1} \times (s_0, \infty), (s_0, \infty)^n \right\}\). Recall that eligibility status \(Z\) is a deterministic function of \(S\), therefore assumptions on the assignment mechanism can be formulated with respect to either \(Z\) or \(S\). Here we prefer to define the assignment mechanism based on \(S\), because \(S\) is the underlying random variable that describes the reasons for the missing and observed values of potential outcomes: a value of \(S\) is assigned, which in turn determines a value for \(Z\).

We assume that the assignment mechanism is locally independent, that is, the assignment mechanism is separable in the unit assignment probabilities, which do not depend on pre-assignment
variables or potential outcomes of other units. Formally,

\[
\Pr(S \in A|A(0), A(1), W(0), W(1), Y(0), Y(1), X) = \prod_{i=1}^{n} \Pr(S_i \leq s_0|A_i(0), A_i(1), W_i(0), W_i(1), Y_i(0), Y_i(1), X_i)
\]

Statistical inference for causal effects requires additional assumptions on the assignment mechanism, specifically assumptions that allow us to describe RD settings as classical randomized experiments around the threshold. In principle, for each student, each value of the forcing variable is possible \textit{a priori}, that is, before one realization of the forcing variable occurs. It is thus reasonable to assume that the assignment mechanism before the realization of the forcing variable is probabilistic, which implies that for each unit, \( i \), both events \( S_i \leq s_0 \) and \( S_i > s_0 \) have \textit{a priori} a non-zero probability of occurring. Also, we focus on ignorable assignment mechanisms that are known functions of their arguments. Ignorable and locally independent assignment mechanisms are unconfounded, that is, free of dependence of any potential outcomes. Assignment mechanisms satisfying these conditions define classical randomized experiments.

The particular assignment rules underlying RD designs suggest that these assumptions, as well as RD-SUTVA, are more reasonable for a subpopulation of units who have a relatively large probability that the realized values of the forcing variable fall in a neighborhood around the threshold, \( s_0 \). For this subpopulation, we can reasonably assume that the distribution of the forcing variable is unrelated to observed and unobserved characteristics of students. On the other hand students with a very small (close to zero) probability that \( S_i \leq s_0 \), as well as students with a very large (close to one) probability that \( S_i \leq s_0 \) are plausibly systematically different in both observed and unobserved characteristics from other students. For example, potential outcomes observed for very rich students, who do not receive any grant, are plausibly different from potential outcomes for poor students with a realized value of \( S \) around the threshold, who do not receive a grant, and vice versa. Therefore we focus on a subpopulation of students who have a probability that \( S_i \leq s_0 \) strictly between zero and one, and sufficiently far away from zero and one. The following assumption guarantees that at least such a subpopulation of units exists.

\textbf{Assumption 3} (Local overlap). Let \( \mathcal{U} \) be the random sample (or population) of units in the
study. There exists a subset of units, $U_{s_0}$, such that for each $i \in U_{s_0}$, $\Pr(S_i \leq s_0) > \epsilon$ and $\Pr(S_i > s_0) > \epsilon$ for some sufficiently large $\epsilon > 0$.

Assumption 3 is essentially a local overlap assumption, in that it assumes that there exists a subpopulation of units, each of whom has a non-zero probability of being assigned to either treatment levels. This represents a main distinction between our framework and the existing RD literature that often describes RD designs as settings where the overlap assumption is violated. Also, we can now restrict RD-SUTVA within the subpopulation $U_{s_0}$, leading to a more plausible local RD-SUTVA:

**Assumption 4 (Local RD-SUTVA).** For each $i \in U_{s_0}$, consider two eligibility statuses $Z_i' = 1(S_i' \leq s_0)$ and $Z_i'' = 1(S_i'' \leq s_0)$, with possibly $S_i' \neq S_i''$. If $Z_i' = Z_i''$, that is, if either $S_i' \leq s_0$ and $S_i'' \leq s_0$, or $S_i' > s_0$ and $S_i'' > s_0$, then $A_i(Z') = A_i(Z'')$, $W_i(Z') = W_i(Z'')$, and $Y_i(Z') = Y_i(Z'')$.

The next assumption helps to construct the subpopulation in practice, stating that for students belonging to $U_{s_0}$, the realized value of the forcing variable $S$ falls in a neighborhood of the threshold $s_0$.

**Assumption 5** There exists $h > 0$ such that for each $\epsilon > 0$, $\Pr(s_0 - h \leq S_i \leq s_0 + h) > 1 - \epsilon$, for each $i \in U_{s_0}$.

Assumption 5 describes the shape of the subpopulation $U_{s_0}$, implying that $U_{s_0}$ comprises units with a value of the forcing variable $S$ in a symmetric interval of the real line with respect to $s_0$. This allows us to focus on the specific subsets of symmetric intervals among all neighborhoods of different shape around the threshold, $s_0$. Note that under Assumption 5 there may exist more than one such interval around the threshold. In other words, Assumptions 3 and 5 do not imply that $U_{s_0}$ is unique. They only require that there exists at least one subpopulation, $U_{s_0}$. Consistently, we are not interested in finding the largest $h$, but we only aim at determining plausible values for $h$. Thus, in this framework, $h$ is not viewed as a random variable.

Finally, we need to formalize the concept of RD design as local randomized experiment: in
a neighborhood of the threshold $s_0$ the forcing variable does not depend on either the potential outcomes or pre-treatment variables. Formally, we have:

**Assumption 6 (Local randomization).** For each $i \in U_{s_0}$,

$$\Pr(S_i|A_i(0), A_i(1), W_i(0), W_i(1), Y_i(0), Y_i(1), X_i) = \Pr(S_i).$$

Assumption 6 states that within the subpopulation $U_{s_0}$ a Bernoulli trial has been conducted, with individual assignment probabilities, that is, the probabilities of being eligible to receive a grant, depending on the distribution of the forcing variable: $\Pr(Z_i = 1) = \Pr(S_i \leq s_0)$. This assumption is crucial in justifying the key idea underlying any RD design. It implies that values of the forcing variable above or below the threshold $s_0$, and thus the eligibility statuses, are randomly assigned in some small neighborhood, $U_{s_0}$, around $s_0$.

Previous approaches to RD designs, which provided insightful methodological contributions to the literature and led to successfully answer important research questions, generally view the forcing variable as a pre-assignment covariate rather than a random variable. As a consequence, the overlap assumption, which requires that there are both treated and control units for all values of the covariates including the forcing variable, is violated. Violation of the overlap assumption implies that the conditional independence assumption, which trivially holds in RD settings, cannot be exploited directly. Therefore, some kind of extrapolation is required, and in order to avoid that estimates heavily rely on extrapolation, standard analyses focus on causal effects of the treatment for the whole population in sharp RDD or a specific sub-population—the compliers—in fuzzy RDD at the threshold. Smoothness assumptions are required to draw inference on those causal effects. Typically, continuity of conditional regression functions (or the conditional distribution functions) of potential outcomes given the forcing variable is assumed.

From a mathematical perspective continuity assumptions at $s_0$ imply randomization at the threshold, $s_0$, but they do not imply randomization in a neighborhood. On the other hand, local randomization implies that for each unit in $U_{s_0}$, the conditional distributions of potential outcomes given the realized value of the forcing variable $S$ are constant, and thus continuous, in $s$. However, from a substantive perspective continuity assumptions and our local randomiza-
tion assumption are different assumptions: they lead to identification and estimation of different causal estimands and so we cannot directly compare them. Under continuity assumptions units with a realized value of the forcing variable around the threshold are used to draw inference on causal effects at the threshold. In our framework we also use subjects with a realized value of the forcing variable around the threshold, but under an alternative set of assumptions, which allows us to draw inference on causal effects for all (or a subset of) the units belonging to the selected sub-population, not just on causal effects at the threshold.

Assumption 6 may not always be plausible. For instance, when the forcing variable is a deterministic variable, which conceptually cannot be interpreted as a random variable with a non-degenerate probability distribution (such as time), the underlying design cannot, in general, be interpreted as a local randomized experiment (Section 6.3 in Lee and Lemieux, 2010, pp 347).

There are subtle but substantive differences between local RD-SUTVA and local randomization. Both assumptions focus on the relationship between potential outcomes and forcing variable, but their interpretation and their implications are different. Local RD-SUTVA is an exclusion restriction assumption and it is required to make the representation of potential outcomes as functions of the eligibility status adequate. Local randomization is an independence assumption and it is crucial to make inference. RD-SUTVA is different from independence assumptions: it does not imply that the probability that we observe a value of the forcing variable above or below the threshold does not depend on potential outcomes. RD-SUTVA simply implies that the exposure to assignment level \( z \) specifies well-defined potential outcomes, for all unit \( i \) and assignment levels \( z \). In other words, considering potential outcomes as random variables, RD-SUTVA does not imply that potential outcomes have the same distribution for each value of the forcing variable. In order to make the forcing variable independent of potential outcomes, we need to introduce additional assumptions, such as Assumption 6.

Following Assumption 3, we can define a local version of the targeted estimands within \( U_{s_0} \):

\[
\tau_{g,s_0} \equiv \mathbb{E} [Y_i(1) - Y_i(0) \mid G_i = g, i \in U_{s_0}],
\]

(5)
for \( g = AA, CA, \{ AA, CA \} \) and their \textit{finite-sample} counterparts, and we have:

\[
\tau_{\{AA,CA\},s_0} \equiv \tau_{s_0} = \frac{\tau_{AA,s_0} \pi_{AA,s_0} + \tau_{CA,s_0} \pi_{CA,s_0}}{\pi_{AA,s_0} + \pi_{CA,s_0}},
\]

where \( \pi_{g,s_0} = \Pr(G_i = g | i \in \mathcal{U}_{s_0}) \) for \( g = AA, CA, NA, DA \), denote the proportion of principal strata in the subpopulation \( \mathcal{U}_{s_0} \). A special case of \( \mathcal{U}_{s_0} \) contains the subpopulation of units with a realized value of the forcing variable \textit{exactly} equal to the threshold value, \( s_0 \). Similarly as before, these local-version estimands represents causal effects of the eligibility status, rather than causal effects of the receipt of the grant.

It is worth noting that for the subpopulation of students belonging to \( \mathcal{U}_{s_0} \), Assumption 6 implies that

\[
\mathbb{E} \left[ Y_i(1) - Y_i(0) \mid G_i = g, i \in \mathcal{U}_{s_0} \right] = \mathbb{E} \left[ Y_i(1) - Y_i(0) \mid Z_i = 1, G_i = g, i \in \mathcal{U}_{s_0} \right].
\]

Under the allocation rule of the Italian university grants, \( Z_i = W_{i}^{\text{obs}} \) for always-applicants and compliant-applicants. Therefore, the local randomization assumption allows the estimands \( \tau_{AA,s_0}, \tau_{CA,s_0}, \) and \( \tau_{s_0} \) to be interpreted as causal effects of receiving a grant for subpopulations of students who actually receive a grant, analogous to the notion of average treatment effect for the treated.

### 3.2 Two Additional Assumptions

We formulate two additional assumptions that are often plausible in practice and facilitate non-parametric identification of the causal effects. Although in principle they are not necessary for Bayesian inference, these assumptions help sharpen inference.

**Assumption 7 Monotonicity of Application Status:**

\[
A_i(1) \geq A_i(0), \quad \text{for all } i \in \mathcal{U}_{s_0}.
\]

**Assumption 8 Stochastic Exclusion Restriction for Never-Applicants:**

\[
\Pr(Y_i(1) | G_i = NA, i \in \mathcal{U}_{s_0}) = \Pr(Y_i(0) | G_i = NA, i \in \mathcal{U}_{s_0}).
\]
Monotonicity rules out the existence of defiant-applicants, and it appears to be plausible in our study. The exclusion restriction assumption rules out direct effects of eligibility for never-applicants. Never-applicants are students who would never apply for a grant irrespective of their eligibility status. These students would not receive the grant in any case. Never-applicants might comprise motivated students with lower economic needs, and if these students regarded education as an important objective to exclude dropout, they might not consider the opportunity to get a grant. Therefore it is reasonable to assume that these students were completely unaffected by their eligibility status. Note that Assumption 8 is of very different nature from the exclusion restriction for compliant-applicants and always-applicants (Assumption 2): the former has implications for inference but not for interpretation, whereas the latter is made solely for interpreting the causal effects of assignment on the outcome attributable to the causal effects of treatment on the outcome. More discussions on the difference can be found in Imbens and Rubin (Forthcoming)(Section 23) and Mealli and Pacini (2013).

3.3 Selection of the Subpopulations

An important issue in practice is the selection of the subpopulation, \( U_{s_0} \), that is, the choice of a bandwidth \( h \) defining an interval around the threshold, \( s_0 \), where our RD assumptions (Assumptions 3 through 6) hold. We propose a Bayesian approach to select plausible values for \( h \) that exploits the fact that Assumption 6 is a local randomization assumption, in the sense that it holds for a subset of units, but may not hold in general for other units. As such, under Assumption 6, in the subpopulation \( U_{s_0} \), pre-treatment variables should be well balanced in the two subsamples defined by assignment, and thus any test of the null hypothesis of no effect of assignment on pre-treatment covariates should fail to reject the null.

Assessing balance in the observed covariates raises problems of multiple comparisons, which may lead to a much higher than planned type I error if they are ignored (e.g., Benjamini and Hochberg, 1995). We account for multiplicities using a Bayesian hierarchical mixed model, which provides an explicit method for borrowing information across covariates (e.g., Berry and Berry, 2004; Scott and Berger, 2006). Following Berry and Berry (2004), we use a mixture for
the prior distribution of the eligibility parameters by assigning a point mass on equality of the means of the covariates between eligible and ineligible units. This Bayesian procedure provides a measure of the risk that a chosen interval around the threshold, \( s_0 \), defines a subpopulation of units that does not exactly matches the underlying true \( U_{s_0} \), including subjects for which our RD assumptions do not hold (more details are given in Section 5).

The idea to exploit balance tests of pre-assignment variables to select a subpopulation of units is also used in Cattaneo et al. (2014), who propose to apply randomization test to select the widest interval around the threshold where the null hypothesis that the assignment has no effect on pre-assignment variables is not rejected. But their approach aims at selecting the largest subpopulation and does not account for multiple comparisons.

Our approach parallels more conventional RD approaches based on local polynomial regression, which also involve bandwidth selection, but for a very different objective, namely finding an optimal balance between precision and bias at the threshold for local polynomials (e.g., Ludwig and Miller, 2007; Lee and Lemieux, 2010; Imbens and Kalyanaraman, 2012), whereas the objective in our framework is to find a subpopulation where our RD assumptions are plausible and the selected subpopulation defines the target population.

4 Bayesian Inference

Our development of the Bayesian approach builds on the seminal works of Rubin (1978) and Imbens and Rubin (1997). Throughout the discussion, we use \( \Pr(\cdot | \cdot) \) and \( \theta | \cdot \) to denote generic conditional distributions and the corresponding parameters, respectively.

Nine quantities are associated with each unit: \( Y_i(0), Y_i(1), W_i(0), W_i(1), A_i(0), A_i(1), X_i, Z_i, S_i \). Among these, \( S_i \) completely determines \( Z_i \); the principal stratum \( (A_i(0), A_i(1)) \) and \( S_i \) completely determine \( (W_i(0), W_i(1)) \). Therefore, inference for causal effects involves only \( Y_i(0), Y_i(1), A_i(0), A_i(1), X_i, S_i \), of which four are observed: \( S_i, X_i, A_i^{\text{obs}} = A_i(Z_i), Y_i^{\text{obs}} = Y_i(Z_i) \), and two are unobserved: \( A_i^{\text{mis}} = A_i(1 - Z_i), Y_i^{\text{mis}} = Y_i(1 - Z_i) \).

Bayesian inference considers the observed values to be realizations of random variables and
the unobserved values to be unobserved random variables. We focus on the subpopulation of $U_{s_0}$ and let $\Pr(Y(0), Y(1), A(0), A(1), X, S; U_{s_0})$ denote the joint probability (density) function of these random variables of all units in $U_{s_0}$. We assume this distribution is unit-exchangeable, that is, it is invariant under a permutation of the unit indices. Then, with essentially no loss of generality, by appealing to de Finetti’s theorem (de Finetti, 1963), we can assume that there exists an unknown parameter vector $\theta$, which is itself a random variable having a known prior distribution $p(\theta)$ such that:

$$\Pr(Y(0), Y(1), A(0), A(1), X, S; U_{s_0}) = \int \prod_{i \in U_{s_0}} \Pr(Y_i(0), Y_i(1), A_i(0), A_i(1), X_i, S_i; \theta) p(\theta) d\theta.$$  

Within the subpopulation of units in $U_{s_0}$, Bayesian inference of the causal estimands, which are functions of $Y_i(z)$’s and $A_i(z)$’s, centers around deriving the posterior distribution for the parameter vector of their distribution, denoted by $\theta_{Y,G}$. Assumption 6 implies that the assignment mechanism does not depend on potential outcomes or pre-treatment variables. Thus, assuming the parameters governing the distributions of the covariates, $X$, the forcing variable, $S$, and the potential outcomes, $Y_i(z)$ and $A_i(z)$, are a priori distinct and independent from each other, the posterior distribution of $\theta_{Y,G}$ can be written as follows:

$$\Pr(\theta_{Y,G}|Y_{obs}, A_{obs}, X, S; U_{s_0})$$

$$\propto p(\theta_{Y,G}) \times \prod_{i \in U_{s_0}} \left[ \int \int \Pr(Y_i(0), Y_i(1), G_i|X_i; \theta_{Y,G}) dY_i^{mis} dA_i^{mis} \right]$$

$$\propto p(\theta_{Y|G}) \times p(\theta_G) \times \prod_{i \in U_{s_0}} \left[ \int \int \Pr(Y_i(0), Y_i(1)|G_i, X_i; \theta_{Y|G}) \Pr(G_i|X_i; \theta_G) dY_i^{mis} dA_i^{mis} \right].$$

The above decomposition suggests that two models need to be specified for model-based inference: (1) the model for potential outcomes conditional on principal strata and covariates, and (2) the model for principal strata conditional on covariates, as well as the prior distribution for the parameters, $p(\theta_{Y,G})$, with $\theta_{Y,G} = (\theta_G, \theta_{Y|G})$.

Let $\pi_{i,g} = \Pr(G_i = g|X_i; \theta_G)$ and $f_{i,gz} = \Pr(Y_i(z)|G_i = g, X_i; \theta_{Y|G})$. Then the posterior
distribution of $\theta_{Y,G}$ given the observed data can be written as follows:

$$
\Pr\left( \theta_{Y,G} | Y_{obs}, A_{obs}, X, S; U_{s0} \right)
\propto p(\theta_{Y,G}) \times \prod_{i \in U_{s0}: S_i > s_0, A_{iobs} = 0} \left( \pi_{i,CA} f_{i,CA} + \pi_{i,NA} f_{i,NA} \right) \times \prod_{i \in U_{s0}: S_i > s_0, A_{iobs} = 1} \pi_{i,AA} f_{i,AA},
$$

where $f_{i,NA} = f_{i,NA,0} = f_{i,NA,1}$ by the exclusion restriction (Assumption 8). The likelihood function, specified by the four products, does not depend on the association between the potential outcomes $Y_i(0)$ and $Y_i(1)$. Therefore the posterior distribution of the association parameters equal their prior distribution as long as the association parameters are a priori independent of the other parameters, as we assume henceforth. The population-average causal estimands $\tau_{AA,s0}$, $\tau_{CA,s0}$, and $\tau_{s0}$ are functions of the parameter vector $\theta_{Y,G}$, which is free of the association parameters, therefore inference for them does not involve the association parameters (also see discussion in Imbens and Rubin, 1997). Inference for sample-average causal estimands for the units in the study, on the other hand, do generally involve the association parameters. In our application inference for sample-average causal estimands is drawn under the conservative assumption that for each unit $i$, potential outcomes, $Y_i(0)$ and $Y_i(1)$, are independent conditional on $X_i$ and $\theta$.

Posterior inference of $\theta_{Y,G}$ can be obtained using Gibbs sampling with a data augmentation step to impute the missing $A_{i mis}$, and thus the missing $G_i$. Specifically, posterior computation is obtained by iteratively drawing from the two posterior predictive distributions:

$$
\Pr\left( \theta_{Y,G} | Y_{obs}, A_{obs}, A_{mis}, X, S; U_{s0} \right) \quad \text{and} \quad \Pr\left( A_{mis} | Y_{obs}, A_{obs}, X, S, \theta_{Y,S}; U_{s0} \right).
$$

The first term, the posterior distribution of $\theta_{Y,G}$ given the complete intermediate data (principal strata of application), has the following form:

$$
\Pr\left( \theta_{Y,G} | Y_{obs}, A_{obs}, A_{mis}, X, S; U_{s0} \right)
\propto p(\theta_{Y,G}) \times \prod_{i \in U_{s0}: S_i > s_0, G_i = CA} \left( \pi_{i,CA} f_{i,CA} + \pi_{i,NA} f_{i,NA} \right) \times \prod_{i \in U_{s0}: S_i > s_0, G_i = AA} \pi_{i,AA} f_{i,AA},
$$

$$
\times \prod_{i \in U_{s0}: G_i = NA} \pi_{i,NA} f_{i,NA} \times \prod_{i \in U_{s0}: S_i \leq s_0, G_i = AA} \pi_{i,AA} f_{i,AA,1} \times \prod_{i \in U_{s0}: S_i \leq s_0, G_i = CA} \pi_{i,CA} f_{i,CA,1}.
$$
Specification of $\pi_{i,g}, f_{i,gz}$ and corresponding prior to posterior computation depends on the specific application. Details of the models and computation in our application will be provided in Section 5. As a general guideline, we recommend to specify $\pi_{i,g}$ and $f_{i,gz}$ conditional on both covariates $X$ and the forcing variable $S$, even though Equation (6) suggests conditioning on $S$ is not required.

Indeed, if the true subpopulations $U_{s_0}$ were known, in theory, we would not need to adjust for $S$, because local randomization guarantees that for units in $U_{s_0}$ values of the forcing variable falling above or below the threshold are independent of the potential outcomes. However, in practice, the true subpopulations $U_{s_0}$ are usually unknown and once a subpopulation has been selected, that is, once a value for $h$, say $h^*$, has been chosen, there may be some units with a realized value of $S$ between $s_0 - h^*$ and $s_0 + h^*$ who do not belong to $U_{s_0}$. For these units there may be a relationship between the forcing variable and potential outcomes, and these potential dependences need to be modeled. Specifically, systematic differences in the forcing variable $S$ that, by definition, occur between eligible and ineligible units, may affect inference in the presence of students who do not belong to $U_{s_0}$.

5 Application to the Evaluation of Italian University Grants

5.1 Data and Construction of Subpopulation

Data. We apply the proposed method to the data from the cohort of first-year students enrolled in 2004 to 2006 at University of Pisa and University of Florence. For each student, information on grant application status ($A_{i,obs}$), grant receipt status ($W_{i,obs}$) at the beginning of the academic year, dropout status at the end of the academic year, and covariates ($X_i$) is obtained from education ministry and university administrative records. The forcing variable $S$ is a combined economic measure of each student, calculated from one’s income tax return and property adjusted for family size based on a formula that is typically not fully known to the students. In all three years, the threshold of eligibility is the combined economic measure of a student below 15000 euros. Thus, the eligibility status ($Z$) is also observed. Typically, students need support from fiscal experts to
compute their value of $S$, and the income revenue authority conducts random inspections to verify that the official tax return were reported. These factors make extremely difficult, if not impossible, for students or students’ families to manipulate the value of $S$ in order to end up on the right side of the threshold. Therefore we argue that our local randomization assumption is reasonable here.

Covariates include sex, high school grade, high school type (4 categories), major in university (6 categories), indicator of year of enrollment (2004, 2005, 2006) and indicator of university (Pisa vs. Florence). Note that the data only include students who had a high school grade of at least 70/100 and applied either for a grant or for a reduction of tuition fee. Summary statistics of important variables for the students with the combined economic measure $S$ within 1000 euros of the threshold are given in Table 1. An unadjusted comparison would suggest that the applicants have higher high-school grades, which is an important indicator of a student’s academic performance, but also higher dropout rate regardless of their eligibility status.

Application rate and dropout rate as a function of $S$ among the students are given in Figure 1. The overall dropout rate is high, consistently between 30% to 50% regardless of the economic measure. From the fitted lines using local logistic polynomial models with order 3 on the two sides of the threshold, discontinuity is clearly visible in both application rate and dropout rate at the threshold. As the economic measure increases, application rate steadily decreases, while the trend in dropout rate has a concave change at the threshold, increasing on the left of the threshold and decreasing on the right.

Selection of the subpopulation. We applied the Bayesian approach to multiple testing discussed in Section 3.3 to find subpopulations of units where our RD assumptions hold. Specifically we use a hierarchical Bayesian model for assessing the balance of the covariates between eligibility groups. We specify probit models for binary variables; conditional probit models for categorical variables and Gaussian models for continuous variables. Formally, we assume that $X_j \sim N(\gamma_{0j} + \gamma_{1j}Z_i, \sigma_j^2)$ if $X_j$ is continuous, and $\Pr(X_{ij} = 1) = \Pr(X_{ij} > 0)$ with $X_{ij}^* \sim N(\gamma_{0j} + \gamma_{1j}Z_i, 1)$, if $X_j$ is binary. If $X_j$ is a categorical variable taking on $K$ values we assume that $\Pr(X_{ij} = 1) = \Pr(X_{ij}^{(1)} \leq 0)$, and $\Pr(X_{ij} = k) = \Pr\left(\cap_{\ell=1}^{k-1}\{X_{ij}^{(\ell)} > 0\} \cap X_{ij}^{(k)} \leq 0\right)$
Table 1: Summary statistics of the first-year students enrolled in 2004 – 2006 at Universities of Pisa and Florence, for the students with \( S_i \in (14\,000, 16\,000) \) euros (i.e., \( h = 1000, s_0 = 15\,000 \)).

<table>
<thead>
<tr>
<th>Variable</th>
<th>( A_{\text{obs}} = 0 )</th>
<th>( A_{\text{obs}} = 1 )</th>
<th>( A_{\text{obs}} = 0 )</th>
<th>( A_{\text{obs}} = 1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample Size</td>
<td>657</td>
<td>304</td>
<td>703</td>
<td>444</td>
</tr>
<tr>
<td>Dropout</td>
<td>0.36</td>
<td>0.50</td>
<td>0.35</td>
<td>0.36</td>
</tr>
<tr>
<td>( S ) (euros)</td>
<td>15495</td>
<td>15509</td>
<td>14504</td>
<td>14499</td>
</tr>
<tr>
<td>Female</td>
<td>0.59</td>
<td>0.61</td>
<td>0.60</td>
<td>0.55</td>
</tr>
<tr>
<td>HS Grade</td>
<td>80.80</td>
<td>84.35</td>
<td>80.17</td>
<td>84.47</td>
</tr>
<tr>
<td>Univ (Pisa)</td>
<td>0.37</td>
<td>0.51</td>
<td>0.37</td>
<td>0.51</td>
</tr>
</tbody>
</table>

for \( k = 2, \ldots, K - 1 \), where \( X^{*k}_{ij} \sim N(\gamma_0 j^k + \gamma_1 j^k Z_i, 1) \), \( k = 1, \ldots, K - 1 \), independently. Let \( \gamma_0 j = (\gamma_0^{(1)} j, \ldots, \gamma_0^{(K-1)} j)' \) and \( \gamma_1 j = (\gamma_1^{(1)} j, \ldots, \gamma_1^{(K-1)} j)' \).

We specify the following prior distributions for the model parameters. The variances of the continuous variables have an inverse-Gamma distribution: \( \sigma^2_j \sim IG(a, b) \). The \( \gamma_0 \)'s have Gaussian prior distributions: for continuous and binary variables, \( \gamma_0 j \sim N(\mu_{\gamma_0}, \sigma^2_{\gamma_0}) \), and for categorical variables, \( \gamma_0 j \sim N(\mu_{\gamma_0}, \sigma^2_{\gamma_0} I_{K-1}) \) with \( I_{K-1} \) being the identity matrix of order \( K - 1 \). Further, for continuous and binary variables, parameters \( \gamma_1 j \) are the difference between means/proportions for eligible and ineligible units. If \( \gamma_1 j = 0 \) then \( X_j \) has the same distribution for eligible and ineligible units. For a categorical variable taking on \( K \) values, the proportion of units in each category is the same for eligible and ineligible units if and only if \( \gamma_0^{(k)} j = 0 \) for each \( k = 1, \ldots, K - 1 \). We assign positive probability to these possibilities using the following mixture prior distributions:

\[
\gamma_1 j \sim \pi_{\gamma_1} \delta_0(\gamma_1 j) + (1 - \pi_{\gamma_1})N(\mu_{\gamma_1}, \sigma^2_{\gamma_1})
\]
Figure 1: Application rate (a) and dropout rate (b) as a function of the forcing variable for the first year students in Universities of Florence and Pisa in 2004—2006. The smoothed lines are estimated using polynomial logistic regression models (of order 3) on each side of the threshold separately; each point are calculated from the units within a binwidth of 50 euros.

\begin{align*}
\gamma_{1j} \sim \prod_{k=1}^{K-1} \big[ \pi_{\gamma_1} \delta_0(\gamma_{1j}^{(k)}) + (1 - \pi_{\gamma_1})N(\mu_{\gamma_1}, \sigma_{\gamma_1}^2) \big],
\end{align*}

where $\delta_0(\cdot)$ is the Dirac delta distribution.

For the hyperparameters, we assign the following prior distributions: $\mu_{\gamma_0} \sim N(\mu_{\gamma_0}, \sigma_{\gamma_0}^2)$; $\sigma_{\gamma_0}^2 \sim IG(a_{\gamma_0}, b_{\gamma_0})$; $\mu_{\gamma_1} \sim N(\mu_{\gamma_1}, \sigma_{\gamma_1}^2)$; $\sigma_{\gamma_1}^2 \sim IG(a_{\gamma_1}, b_{\gamma_1})$; and $\pi_{\gamma_1} \sim Beta(a_\pi, b_\pi)$.

We implement the Bayesian model for assessing the balance of covariates on the two sides of the threshold for various subpopulations defined by different values of $h$. Details on the Monte Carlo Markov Chain (MCMC) for the posterior computation are relegated to Web Supplementary Material\(^3\). Table 2 shows the posterior probabilities that the covariates have the same distribution between eligible and ineligible students for the subpopulations defined by $h = 250, 500, 750, 1000, 1500, 2000, 2500, 3000, 4000, 5000$. These probabilities provides a measure of the risk that a chosen interval around the threshold, $s_0$, defines a subpopulation that includes units for which our RD assumptions do not hold. The posterior probabilities that the pre-assignment

\(^3\text{Available at: http://www.stat.duke.edu/~f135/RDD/suppl1.pdf}\)
variables are well balanced are quite high (greater than 55%) for subpopulations defined by values of $h$ strictly lower than 1500. For larger subpopulations some covariates, such as the “indicator of university,” are clearly unbalanced.

Given that the risk that a chosen interval around the threshold defines a subpopulation that includes units not belonging to the target subpopulation, $\mathcal{U}_{s_0}$, is not zero, in order to account for the presence of these units, we conduct the subsequent analyses conditioning on both covariates and the realized values on the forcing variable. Also we evaluate the robustness of our results conducting analyses using various values of $h$ ($h = 500, 1000, 1500$).

5.2 Models and Computation

Parametric models. For the units within the selected subpopulation $\mathcal{U}_{s_0}$, we assume parametric models for the outcome ($f_{g,z}$) and principal strata ($\pi_g$). Alternative models, such as Student-$t$ models (Chib and Jacobi, 2011) and Bayesian nonparametric models (Schwartz et al., 2011), can be considered. Note that although we are using parametric models, identification does not rely on parametric assumptions. The model for the principal strata of application consists of two conditional probit models:

$$
\begin{align*}
\Pr(G_i = AA) &= \Pr(G_i^{*}(AA) \leq 0), \\
\Pr(G_i = NA) &= \Pr(G_i^{*}(AA) > 0 \text{ and } G_i^{*}(NA) \leq 0), \\
\Pr(G_i = CA) &= 1 - \Pr(G_i = NA) - \Pr(G_i = AA),
\end{align*}
$$

where

$$
G_i^{*}(AA) = \alpha_{AA,0} + \alpha_{AA}^{(S)} S_i^{*} + X_i' \alpha_{AA}^{(X)} + \epsilon_{AA,i}, \quad G_i^{*}(NA) = \alpha_{NA,0} + \alpha_{NA}^{(S)} S_i^{*} + X_i' \alpha_{NA}^{(X)} + \epsilon_{NA,i},
$$

with $\epsilon_{AA,i} \sim N(0,1)$, $\epsilon_{NA,i} \sim N(0,1)$ independently, and $S_i^{*} = (S_i - s_0)/1000$.

Dropout, the primary outcome in our application, is binary. Therefore, we assume the following generalized linear outcome model with a probit link (Albert and Chib, 1993):

$$
\Pr(Y_i(z) = 1 | G_i = g, S_i, X_i) = \Phi \left( \beta_{0,g,z} + \beta_{g,z}^{(S)} S_i^{*} + X_i' \beta_{g,z}^{(X)} \right).
$$
Define $\alpha_g = [\alpha_{g,0}, \alpha_g^{(S)}, \alpha_g^{(X)}]$, $g = AA, NA$, and $\beta_{g,z} = [\beta_{0,g,z}, \beta_{g,z}^{(S)}]$, $g = AA, CA, NA$; $z = 0, 1$. By Assumption 8, $\beta_{NA,0} = \beta_{NA,1}$. We assume that parameters are a priori independent and use multivariate normal prior distributions:

$$\alpha_g \sim N \left( \mu_{\alpha_g}; \sigma_{\alpha_g}^2 I \right), \quad \beta_{g,z} \sim N \left( \mu_{\beta_{g,z}}; \sigma_{\beta_{g,z}}^2 I \right), \quad \beta^{(X)} \sim N \left( \mu_{\beta}; \sigma_{\beta}^2 I \right)$$

where $I$ is the identity matrix. We specify flat priors setting the hyper-parameters as follows: setting $\mu_{\alpha_g}, \mu_{\beta_{g,z}}, \mu_{\beta}$ to be null vectors; and setting large prior variances $\sigma_{\alpha_g}^2 = 10, \sigma_{\beta_{g,z}}^2 = 10, \sigma_{\beta}^2 = 10$ for $g = AA, CA, NA$; $z = 0, 1$.

**Posterior Computation.** Details of the MCMC algorithm for the posterior computation based on the outline in Section 4 are given in Web Supplementary Material. Upon obtaining the posterior draws of the parameters, we calculate three estimates for each causal estimand: population-average effect within $U_{s_0}$ and at $s_0$, and sample-average effect within $U_{s_0}$. The population-average effects within $U_{s_0}$ are calculated averaging the model-based dropout proportions over the empirical distribution of the pre-assignment variables and the forcing variable:

$$\frac{\sum_{i \in U_{s_0}} \pi_{i,g} \Phi \left( \beta_{0,g,1} + \beta_{g,1}^{(S)} S_i^* + X_i \beta_{g,1}^{(X)} \right)}{\sum_{i \in U_{s_0}} \pi_{i,g}} - \frac{\sum_{i \in U_{s_0}} \pi_{i,g} \Phi \left( \beta_{0,g,0} + \beta_{g,0}^{(S)} S_i^* + X_i \beta_{g,0}^{(X)} \right)}{\sum_{i \in U_{s_0}} \pi_{i,g}}$$

for $g = AA, CA, \{AA, CA\}$. The population-average effects at $s_0$ are calculated in a similar way setting $S_i^* = 0$ (i.e., $S_i = s_0$) for each $i$. To obtain the sample-average estimates, we compute the posterior predictive distributions of the potential outcomes for each student $i$ in $U_{s_0}$, based on which the sample average is calculated.

### 5.3 Results

We conducted Bayesian analysis using $h = 500, 1000, 1500$. Three parallel MCMC chains of 125,000 iterations with different starting values were run for each $h$, with the first 25,000 as burn-in. Mixing of the chains was deemed to be adequate based on the Gelman-Rubin statistic (Gelman and Rubin, 1992) and visual inspection of the traceplots of the causal parameters (functions of model parameters). All chains led to similar posterior summary statistics. Posterior
summaries were based on 5 000 draws from the posterior distributions (thinning per 20 iterations)

Table 3 shows posterior medians and 95% credible intervals of the principal strata proportions under monotonicity and the three estimates for causal parameters $\tau_{AA,s_0}$, $\tau_{CA,s_0}$, $\tau_{s_0}$, for bandwidths ranging from 500 to 1500 euros. The results are robust across different bandwidths. The estimated proportions of the principal strata are very similar across different $h$: there are more than 61% never-applicants, more than 32% always-applicants and less than 6.5% compliant-applicants. The three estimates for the same causal parameter are also similar. The posterior distributions of the causal effect for always-applicants, $\tau_{AA,s_0}$, and the union of always-applicants and compliant-applicants, $\tau_{s_0}$, are centered on negative values, and the 95% credible intervals do not cover 0, irrespective of the choice of the bandwidth.

For instance, consider the finite-sample causal effects for the subpopulation within $h = 1000$ euros around the threshold (middle block of columns in Table 3). The estimated $\tau_{s_0}$ suggests an at least 13.9% (95% CI: (3.4%; 24.7%)) reduction in dropout rate for the students who receive the grants. The estimated $\tau_{AA,s_0}$ suggests an even stronger positive effect among the always-applicants: a 16.1% (95% CI: (5%; 27%)) reduction in dropout rate. In fact, $\tau_{s_0}$, which is a weighted average of the effects for always-applicants and compliant-applicants, appears to be diluted by the somewhat surprising negative effects among the compliant-applicant (a 3.1% increase in dropout rate). However, the data do not seem to contain much information on compliant-applicants (the estimated proportion of compliant-applicants is very small, less than 5%), and the effects were estimated with large uncertainties.

These results suggest that the current Italian university grants are effective in reducing dropout from universities among students from families with annual economic measure around 15 000.

---

4 We also conducted Bayesian analysis using alternative models with different order polynomials in $S$ as well as models conditioning only on $S$ (without using the pretreatment variables) and null models, conditioning on neither $S$ nor the pre-treatment covariates. Consistently to results found in Mealli and Rampichini (2012), higher order polynomials do not lead to substantial inferential benefits, and posterior distributions of the causal effects of interest did not substantially change with the alternative models, so here we only show the results based on models conditioning on both $S$ and the pre-treatment covariates.
euros. Our analysis also reveals some additional information for policy making. Specifically, always-applicants and compliant-applicants are found to be heterogeneous with respect to the effect of the grants. The causal effect for compliers, $\tau_{s0}$, usually estimated in a standard IV analysis that ignores the application information, is attenuated by the small proportion of compliant-applicants. From a cost-effective perspective, it appears more beneficial for education administrations to lower the eligibility criteria (i.e., decrease the threshold $s_0$) to allow more applicants to get the grant, than to increase the amount of the grant to awardees. The combination of low percentage of compliant-applicants and high percentage of always-applicants suggests that for most students with the economic measure being around the threshold who intend to apply for the grants would apply irrespective of their eligibility. From a policy perspective, this implies that education administrations should better explain the rule of eligibility to potential applicants to discourage ineligible students from applying, and thus reduce unnecessary efforts from these students and the administration (for processing these applications).

**Posterior Predictive Model Checking.** Assessing the plausibility of model assumptions is critical in model-based approaches. Model checking here is not as crucial as in other model-based approaches thanks to the randomization assumption, but it is still prudent to check the model fit since there are uncertainties in the selection of $\mathcal{U}_{s0}$. We adopt Bayesian posterior predictive checks (Gelman et al., 1996) to assess goodness-of-fit of our models in the application. Posterior predictive checks evaluate goodness-of-fit of models by measuring the discrepancy between the observed data and replicated data simulated from its posterior predictive distribution. The particular procedure adopted here is similar to that in Mattei et al. (2013, Section 6). Specifically, we consider three discrepancy measures aim at assessing whether the model can preserve broad features of signal, noise and signal-to-noise ratio (SNR) in the drop-out status distribution for compliant-applicants, always-applicants and the union of these two principal strata, and calculate posterior predictive $p$–values (PPPVs) to summarize discrepancies between the observed data and replicated data. Extreme (close to 0 or 1) PPPVs can be interpreted as evidence of lack-of-fit of the model in, at least some aspects of, the observed data. Further details of the procedure are relegated in Web Supplementary Material.
Table 4 shows the PPPVs for the model-fit to the subpopulation with bandwidth of 500, 1000 and 1500 euros, respectively. The PPPVs suggest good model-fit for all bandwidths, except for a slight under-fit for always-applicants in the subpopulation with \( h = 500 \), which is possibly due to the small sample size. We have also calculated the less conservative sampled posterior predictive \( p \)-values (Johnson, 2007; Gosselin, 2011) and obtained similar conclusions.

6 Conclusion

Motivated from the evaluation of Italian university grants, we propose a probabilistic formulation of the assignment mechanism for regression discontinuity designs and develop a full Bayesian approach to draw causal inference within the framework of principal stratification. In particular, we illustrate how to utilize information on application status to gain additional insights in program evaluation. Applying the method to the data from two Italian universities, we find university grants reduce dropping out of higher education for students from low-income families and the effect size is especially pronounced for motivated students (always-applicants).

In the evaluation of Italian university grants, other than dropout, student’s academic performance (measured by total credits taken or passing rate of exams) is also of great interest in policy. As illustrated by Mattei et al. (2013) and Mercatanti et al. (2012) jointly modeling two outcomes, dropout and academic performance in this case, would be worthwhile for both practical and inferential purposes, and it is at the top of our research agenda.

After the first year the Italian university grant assignment rule combines sequential and RD designs (Cellini et al., 2010): grants are allocated both on the basis of students family economic indicator and on the ground of their academic performance (exam scores above a certain threshold). Such complex assignment mechanisms pose challenges to causal inference, requiring new structures and assumptions; meanwhile, they also present great opportunities for extending the existing framework to more general RD settings. One specific direction of our future research is to develop methods that combine Bayesian tools for RDs and dynamic treatment regimes (Murphy, 2003; Zajonc, 2012) in the presence of multiple forcing variables (Imbens and Zajonc, 30).
References


Table 2: Posterior probabilities that the covariates have the same distribution between eligible and ineligible students for various subpopulation

<table>
<thead>
<tr>
<th>Variable</th>
<th>$h=250$</th>
<th>$h=500$</th>
<th>$h=750$</th>
<th>$h=1000$</th>
<th>$h=1500$</th>
<th>$h=2000$</th>
<th>$h=2500$</th>
<th>$h=3000$</th>
<th>$h=4000$</th>
<th>$h=5000$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$(n=528)$</td>
<td>$(n=1042)$</td>
<td>$(n=1577)$</td>
<td>$(n=2108)$</td>
<td>$(n=3166)$</td>
<td>$(n=4197)$</td>
<td>$(n=5159)$</td>
<td>$(n=6113)$</td>
<td>$(n=8061)$</td>
<td>$(n=9846)$</td>
</tr>
<tr>
<td>Sex</td>
<td>.955</td>
<td>.950</td>
<td>.960</td>
<td>.962</td>
<td>.990</td>
<td>.972</td>
<td>.970</td>
<td>.960</td>
<td>.972</td>
<td>.735</td>
</tr>
<tr>
<td>High School Type (Baseline: Other)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Humanity</td>
<td>.951</td>
<td>.952</td>
<td>.949</td>
<td>.955</td>
<td>.977</td>
<td>.986</td>
<td>.973</td>
<td>.986</td>
<td>.969</td>
<td>.957</td>
</tr>
<tr>
<td>Science</td>
<td>.894</td>
<td>.905</td>
<td>.926</td>
<td>.927</td>
<td>.950</td>
<td>.897</td>
<td>.919</td>
<td>.926</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>Tech</td>
<td>.790</td>
<td>.807</td>
<td>.790</td>
<td>.808</td>
<td>.834</td>
<td>.740</td>
<td>.734</td>
<td>.793</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>HS Grade</td>
<td>.955</td>
<td>.958</td>
<td>.972</td>
<td>.978</td>
<td>.984</td>
<td>.975</td>
<td>.980</td>
<td>.990</td>
<td>.975</td>
<td>.977</td>
</tr>
<tr>
<td>Year (Baseline: 2004)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td>.932</td>
<td>.964</td>
<td>.954</td>
<td>.926</td>
<td>.979</td>
<td>.982</td>
<td>.981</td>
<td>.983</td>
<td>.850</td>
<td>.874</td>
</tr>
<tr>
<td>2006</td>
<td>.883</td>
<td>.918</td>
<td>.914</td>
<td>.909</td>
<td>.939</td>
<td>.944</td>
<td>.959</td>
<td>.970</td>
<td>.817</td>
<td>.823</td>
</tr>
<tr>
<td>University (Pisa)</td>
<td>.950</td>
<td>.916</td>
<td>.971</td>
<td>.983</td>
<td>.648</td>
<td>.271</td>
<td>.410</td>
<td>.300</td>
<td>.000</td>
<td>.017</td>
</tr>
<tr>
<td>Major in University (Baseline: Other)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Humanity</td>
<td>.946</td>
<td>.899</td>
<td>.689</td>
<td>.797</td>
<td>.892</td>
<td>.919</td>
<td>.960</td>
<td>.990</td>
<td>.976</td>
<td>.972</td>
</tr>
<tr>
<td>Science</td>
<td>.894</td>
<td>.857</td>
<td>.660</td>
<td>.751</td>
<td>.863</td>
<td>.878</td>
<td>.945</td>
<td>.966</td>
<td>.933</td>
<td>.943</td>
</tr>
<tr>
<td>Social Science</td>
<td>.728</td>
<td>.776</td>
<td>.604</td>
<td>.677</td>
<td>.812</td>
<td>.813</td>
<td>.906</td>
<td>.926</td>
<td>.832</td>
<td>.812</td>
</tr>
<tr>
<td>Tech</td>
<td>.632</td>
<td>.634</td>
<td>.565</td>
<td>.624</td>
<td>.764</td>
<td>.769</td>
<td>.871</td>
<td>.876</td>
<td>.755</td>
<td>.434</td>
</tr>
</tbody>
</table>
Table 3: Posterior median and 95% credible intervals of principal strata proportion and super-population and finite-sample causal effects on dropout for always-applicants (τ_{AA,s_0}), compliant-applicants (τ_{CA,s_0}), and their union (τ_{s_0}), for the subpopulation within different bandwidths h around the threshold.

<table>
<thead>
<tr>
<th>h</th>
<th>Population-average</th>
<th></th>
<th>Sample-average</th>
<th></th>
<th>Population-average at s_0</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median</td>
<td>95% CI</td>
<td>Median</td>
<td>95% CI</td>
<td>Median</td>
<td>95% CI</td>
</tr>
</tbody>
</table>
| h = 500 | Pr(G_i = AA)       | .323  | (.294; .355)   | .322  | (.309; .336)              | .320  | (.291; .352)
|         | Pr(G_i = CA)       | .060  | (.031; .105)   | .041  | (.021; .090)              | .058  | (.030; .094)
|         | Pr(G_i = NA)       | .616  | (.570; .650)   | .637  | (.590; .651)              | .621  | (.583; .654)
|         | τ_{AA,s_0}         | -.153 | (-.313; -.030) | -.152 | (-.307; -.038)            | -.154 | (-.298; -.030)
|         | τ_{CA,s_0}         | .045  | (-.170; .497)  | .074  | (-.256; .545)            | .039  | (-.169; .474)
|         | τ_{s_0}            | -.116 | (-.253; -.005) | -.120 | (-.265; -.009)           | -.120 | (-.245; -.012)
| h = 1000| Pr(G_i = AA)       | .336  | (.312; .365)   | .333  | (.318; .354)              | .335  | (.311; .363)
|         | Pr(G_i = CA)       | .043  | (.002; .086)   | .027  | (.002; .075)              | .043  | (.001; .075)
|         | Pr(G_i = NA)       | .623  | (.584; .652)   | .640  | (.599; .645)              | .625  | (.594; .656)
|         | τ_{AA,s_0}         | -.161 | (-.273; -.052) | -.161 | (-.270; -.057)            | -.154 | (-.259; -.052)
|         | τ_{CA,s_0}         | .028  | (-.745; .828)  | .031  | (-.778; .871)            | .010  | (-.918; .933)
|         | τ_{s_0}            | -.132 | (-.242; -.021) | -.139 | (-.247; -.034)           | -.128 | (-.229; -.020)
| h = 1500| Pr(G_i = AA)       | .332  | (.315; .349)   | .332  | (.326; .337)              | .329  | (.312; .346)
|         | Pr(G_i = CA)       | .042  | (.035; .077)   | .027  | (.020; .066)              | .042  | (.036; .062)
|         | Pr(G_i = NA)       | .625  | (.591; .642)   | .642  | (.605; .644)              | .628  | (.606; .646)
|         | τ_{AA,s_0}         | -.183 | (-.286; -.077) | -.187 | (-.291; -.085)            | -.153 | (-.247; -.063)
|         | τ_{CA,s_0}         | .010  | (-.304; .797)  | .011  | (-.207; .928)            | .000  | (-.154; .951)
|         | τ_{s_0}            | -.153 | (-.256; -.040) | -.165 | (-.266; -.057)           | -.130 | (-.217; -.019)
Table 4: Bayesian $p$—values of signal, noise and SNR under different $h$ for the model used in the application to Italian university grants.

<table>
<thead>
<tr>
<th>$h$</th>
<th>principal strata</th>
<th>Signal</th>
<th>Noise</th>
<th>SNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>${CA}$</td>
<td>.095</td>
<td>.630</td>
<td>.094</td>
</tr>
<tr>
<td></td>
<td>${AA}$</td>
<td>.254</td>
<td>.325</td>
<td>.254</td>
</tr>
<tr>
<td></td>
<td>${AA,CA}$</td>
<td>.338</td>
<td>.273</td>
<td>.370</td>
</tr>
<tr>
<td>1000</td>
<td>${CA}$</td>
<td>.411</td>
<td>.425</td>
<td>.419</td>
</tr>
<tr>
<td></td>
<td>${AA}$</td>
<td>.400</td>
<td>.444</td>
<td>.473</td>
</tr>
<tr>
<td></td>
<td>${AA,CA}$</td>
<td>.493</td>
<td>.335</td>
<td>.518</td>
</tr>
<tr>
<td>1500</td>
<td>${CA}$</td>
<td>.208</td>
<td>.444</td>
<td>.210</td>
</tr>
<tr>
<td></td>
<td>${AA}$</td>
<td>.372</td>
<td>.400</td>
<td>.261</td>
</tr>
<tr>
<td></td>
<td>${AA,CA}$</td>
<td>.455</td>
<td>.337</td>
<td>.470</td>
</tr>
</tbody>
</table>