Perform Better, or Else: Academic Probation, Public Praise and Students Decision-Making*

Nicholas A. Wright†

February 3, 2018

Abstract

This paper examines how college initiatives that ascribe public recognition or written reprimand to a set standard of academic performance impact students decision-making. Many colleges utilize programs such as the Deans list and academic probation policies as mediums to encourage student success. These policies impose a future cost on affected students, either through the loss of acquired benefits or the threat of expulsion if they fail to perform above an established standard in future semesters. As such, to meet these standards, treated students may be induced into increasing the effort they subsequently exert. In addition, they have an incentive to manipulate their behavior along non-effort dimensions, such as through the courses and/or instructors they select. Using the regression discontinuity design, I provide convincing evidence that the students that are treated by either the Deans list or academic probation policy improve their academic performance in subsequent semesters. However, increased effort may be only one of the mechanisms through which students change their behavior following treatment. In particular, there is evidence that the Deans list policy induces treated students to select courses and instructors that are more likely to award higher grades and have a lower failure rate. Similarly, the results suggest that the academic probation policy causes students to improve their match quality by switching majors and to employ a maximin strategy for expected grades when choosing courses. This result does not change when the probation policy becomes more restrictive by an increase in the GPA treatment threshold.

Keywords: Public Praise, Academic Probation, Effort, Course Selection, Major, GPA
JEL Codes: D84, I20, I21, I23, J01

*I thank Barry Hirsch, Dan Kriesman, Tom Mroz, Ross Rubenstein, Jonathan Smith and the numerous participants at seminars and conferences for helpful comments and suggestions. This project has also benefited from the support of Lloyd Barker, Garvin Gordon, Tamara Mathews, Kevin Tai and administrators for logistic, institutional and data support. All remaining errors are my own. This project was made possible by the generous financial support of the Bernard B. and Eugenia A. Ramsey Chair of Private Enterprise.

†Department of Economics, Andrew Young School of Policy Studies, Georgia State University, Atlanta, GA. e-mail: nwright16@student.gsu.edu.

1 Introduction

A college student receives varying administrative feedback from their academic institution that may be pertinent to their subsequent postsecondary decisions and outcomes. In some cases, this feedback informs students that they are performing below the acceptable academic standard expected by the institution and that failure to meet this standard in subsequent assessments will result in actions being taken against them. At other times, the institution may choose to celebrate, levy praise or publicly recognize the academic achievement of high-performing students. In both of these cases, the student potentially receives a more salient signal of the quality of their academic performance. Furthermore, this feedback may change the student's status quo by introducing a punishment they may face or a benefit they may lose in future periods (Baumeister et al., 1990).

Despite the frequency with which students receive these feedback about their performance, the empirical evidence of their impact on students' subsequent academic decisions and outcomes is relatively sparse in the literature. This is because it is difficult to empirically assess the impact of praise and reprimands in a non-experimental setting, owing to the heterogeneity between those who are given these treatment and those who are not. Additionally, quality data on the various performance feedback that postsecondary students receive and their academic records are not widely available. In this paper, I utilize administrative data from a large publicly funded university in Jamaica to estimate the causal impact of receiving a warning letter or being recognized on the Dean's list on students' subsequent academic decisions and outcomes. These initiatives are employed by most colleges and is used as a medium to encourage student success.

The Dean's list initiative is one medium that colleges utilize to recognize the academic performance of high-achieving students. This list is comprised of the names of students whose grade point average (GPA) exceeds some high performance standard. They also receive a combination of public recognition, praise and awards for achieving this accomplishment. This list is dynamically updated at a fixed time interval and new students who have passed this performance threshold are added, while others that have fallen below are dropped from the list. In contrast, the academic probation policy targets poor performing students whose academic performance falls below some minimum standard expected by the institution. The students who satisfy this criteria receive a warning letter that explains that they have failed to meet the academic standard expected by the board of examiners. Once warned, the letter informs the student that if this performance is repeated in the following semester then they will be asked to withdraw from their studies. Both policies provide students with a signal about the quality of their academic performance and imposes a future cost on academic performance below an established standard.

There are four main channels through which these policies can potentially impact a student's subsequent behavior. First, since being treated imposes a cost on future performance below the acceptable standard, affected students may be induced into increasing the level of effort they exert. A student that is affected by these policies enter a new state of nature where he is at risk of facing a penalty or losing an acquired (endowment) benefit. This may incentivize him to more closely align his actions with the institution's preferences(Baumeister et al., 1990). Lindo et al. (2010) found evidence of this mechanism when they examined how the academic probation policy impact students academic outcomes. Second, It is also plausible that these policies could inform about a student's place in the GPA distribution and could cause students to update their beliefs about their academic ability (Weiner et al., 1972). Several studies show that students update their beliefs about their academic ability while in college and make decisions about persisting in a logical manner. This shows that students behave like producers in evaluating their current performance and decide how much effort to invest in subsequent periods (Stinebrickner and Stinebrickner, 2011; Lin, 2013; Elsner and Isphording, 2017). Third, receiving positive or negative information about one's academic performance may alter a student's self-image, confidence, and motivation. This argument has been made in the psychology literature, where it has been posited that goal-setting, personal expectation and mental effort investment may be influenced by the perception of success or failure (Venables and Fairclough, 2009; Bénabou and Tirole, 2002; Seaver and Quarton, 1973; Feather, 1966). Similarly, Baumeister et al. (1990) finds evidence that praise may improve an individuals subsequent performance on tasks that depend primarily on effort. Finally, these initiatives could have an impact on a student's course selection strategy and persistence in their chosen major of study. Since these policies require that students exceed a set standard of academic performance, it may incentivize them to change their behavior along non-effort margins to achieve this objective[1]. I also explore some of these behavioral changes in this paper.

This paper shares some similarity to the works of Lindo et al. (2010), Seaver and Quarton (1973) and Thistlethwaite and Campbell (1960). Lindo et al. (2010) is a recent paper that focused on the impact of performance standards on students subsequent academic outcomes[2]. I contribute to this new literature in two respects. First, I provide extensive results on the mechanisms through which students adjust their behavior after receiving a warning letter. Secondly, this paper offers some insights into how the results may diverge when key institutional details - such as the generosity of the performance standard- are altered. In the psychology literature, Seaver and Quarton (1973) and Thistlethwaite and Campbell (1960) examined the impact of public recognition on students' subsequent attitude and academic performance. Seaver and Quarton (1973) argued that Dean's list recognition serves as a potentially powerful social reinforcer of the behaviors leading to academic achievement, and it should improve students' self-esteem and expectation for future performance relative to others. Similarly, Thistlethwaite and Campbell (1960) found that there is a motivational effect of public recognition which improves students' subsequent academic aspirations and outcomes. To augment the findings in these studies, I examine the mechanism through which students adjust their behavior after being included on the the Dean's list and assess the long-term impact of the program on student success. Further, I exploit across-faculty variations in the design of the Dean's list initiative to explore the importance of the frequency of treatment and degree of public recognition

[1]In psychology, Baumeister et al. (1990) argues that performance-based praise may convey an implicit demand for continued good performance. The warning letter also states this expectation explicitly. In such cases, the student may be pressured into exceedings a certain level of performance. Consequently, if affected students are pressure in this manner, they have an incentive to adjust their behavior in less costly ways to achieve this objective. This may include selecting easier courses, more generous instructors or switching major. These mechanisms may prove less costly than increasing one's effort.

[2]There has been a few papers subsequently published. An example of such a study is Fletcher and Tokmouline (2017).

in students response to the program[3]. Finally, while the scope of both papers are quite novel, the main estimates in this study benefits from several significant improvements in the RD technology that was unavailable when they were written. This is the first paper to examine the impact of the Dean's list policy in the economics literature[4].

To establish the causal effect of each program, the regression discontinuity design is employed. This is done by exploiting the discontinuities in treatment eligibility that is created by institutional policies. In particular, the identifying assumption requires that students who are immediately to the right of the performance standard threshold for academic probation and were not eligible for receiving the warning letter have similar observed and unobserved characteristics to those who are just to the left of the threshold and were treated by receiving the written reprimand. Similarly, we assume that students who are narrowly eligible for a spot on the Dean's list provide a good counterfactual to students whose grade point average (GPA) fell just below the Dean's list academic standard. While this assumption is usually violated by students' manipulation of their eligibility for treatment, this is unlikely to be the case in this paper due to the policies governing new students at the institution being examined.

The results finds that both policies have a positive effect on the academic performance of treated students in subsequent semesters. For the dean's list policy, there is evidence suggesting that the intensity of the effect varies with the frequency of the intervention and the degree of public recognition students receive. The results also suggest that the academic probation policy causes some treated students to exit the university and continuing students improve their match quality by transfering across faculties or switching their major of study. In responding to either policies, I found that treated students are incentivized to engage in strategic course taking behavior, choosing courses that are more generous along dimensions we would expect. However, there is no evidence that either policy had a negative impact on students long-term outcomes, such as their prospects for graduation.

These results provides mixed support for existing findings in the literature and offers several insights on the mechanisms students utilize to meet these standards once treated. In general, the main conclusions are applicable in most contexts where an agent is reprimanded or recognized for their past performance in an attempt to encourage a certain level of performance in future periods. For example, these results could shed light on employees behavioral response to workplace policies of performance appraisal, written reprimands for unacceptable behavior and recognition for outstanding performance.

The remainder of this paper is organized as follows: Section 2 outlines the data and institutional framework specific to the policies evaluated in this study. Section 3 describes the empirical model and major assumption of the regression discontinuity design, section 4 presents the main results, and

[3]faculty is used equivalently to school. Each represents a collection of academic departments. For example, the social sciences faculty contains several departments including government, economics, business, psychology and sociology.

[4]To the best of my knowledge, Seaver and Quarton (1973) is the only paper that has empirically assessed the impact of the Dean's list policy. likewise, the literature on public praise is very sparse. This is not due to a lack of interest in these programs, since most students are affected by this polciy and most colleges utilize them. Further, these topics are extensively examined using different methods in the psychology literature. As such, the scarcity of empirical evidence on the effect of this program is likely due to significant restrictions on the availability of college student-level academic and programs data.

section 5 provides the main the concluding arguments.

2 Data and Institutional Framework

The data employed in this study are taken from the administrative records of a large publicly-funded university in Jamaica. Together, these records form a student-level panel that is comprised of observations spanning the period 2008 to 2016. For each semester a student is enrolled, the panel contains data on their demographic characteristics, course details, major and degree information, academic performance, financial records, academic standing and living arrangements. This is ideal because it comes directly from the confidential records held by the university, and as such, it contains minimal measurement error. Furthermore, the level of detail contained in these records allow us to examine the impact of various academic programs on a wide array of academic decisions and outcomes of postsecondary students.

Though the institution mainly caters to domestic students, approximately 12% of all admitted students are international students coming from 98 countries worldwide. In any given semester, the enrolled student population is about 16000, of which 21% are enrolled as part-time students. The university offers a wide array of programs choices, with students enrolled in about 200 majors over the sample period. Each student's final grades is governed by a university-wide grade scale that is converted into a standardized grade point average (GPA), ranging from a low of 0 to a high of 4.3.

After registration, each student is given a handbook which details the grading criteria, university policies, and provides information about course requirements for their chosen program of study. The majority of the undergraduate degrees offered by the university requires that students complete six semesters (3 years) with a recommended course load of fifteen (15) credits per semester. Each academic year, courses are offered in a three-semester system, corresponding to the fall, spring, and summer framework. All students entering the university must satisfactorily complete a set of core first year courses to meet the pre-requisite requirements for their second year courses. Upon completing their first year, students have the option of changing their major, pursuing a double major, or adding a minor to their current program of study.

To satisfy the requirements of their major, a student must complete a certain number of courses within their area of interest. Each school within the university has its own regulations regarding the number of courses students must complete to satisfy their major requirements, the number of core courses that are mandatory for all students in a given major and the flexibility students have in choosing free elective courses or courses outside of their departments[5].

In analysing each program, the maximum GPA bandwidth choice considered is 0.6. The optimal bandwidth size for the baseline estimates are selected using a data driven approach which ensures that the observed covariates are balanced across treatment groups. This strategy ensures that the results are comparable with previous studies, while satisfying the key assumptions of the RD design. Nonetheless, the results are robust to the bandwidth selection procedures. This is extensively

[5]the set of core courses varies across faculty.

discussed below in section 4.4.

Each faculty independently determines how they will recognize and reward high performing students. As a result, there is variation at the faculty-level in how awards programs are implemented and the criteria for qualifying for such awards. Given the heterogeneity in student award programs across faculties, the Dean's list analysis is divided into two sub-samples. To form a sub-sample, faculties are grouped based on the similarity of their awards program in terms of the criteria for selection, the recognition students receive and the frequency with which students are awarded. The first sub-sample contain students in the restricted neighborhood around the eligibility cutoff in the faculty of Social Sciences. This sample is comprised of 7098 students that are observed across five semesters following treatment. The sample tracks students that are treated in their first semester, starting with the cohort entering in 2009 and ending with the cohort entering 2014. Since students are followed for five semesters, the last cohort are followed through to the second semester in 2016. The second sub-sample include students in the restricted neighborhood around the eligibility cutoff in the Faculty of Medical Sciences and the Faculty of Pure and Applied Sciences. This sub-sample contains 4354 students, observed over four semesters following their first treatment[6]. The sample begins with students treated in their first year of 2008 and ends with the cohort of students entering in the 2014 academic year.

In contrast, the minimum acceptable performance standards for low achieving students and the GPA scheme for the class of degree students obtain at graduation are university-wide policies. These policies are established and administered by the university's examinations board. As such, this sample contain students in the maximum neighborhood around the treatment cutoff, irrespective of their faculty. This sample contains 15,428 students. These students are treated at the end of their first semester and are followed for five semesters thereafter.

The sample is further restricted to newly admitted full-time students enrolled in an undergraduate degree program. These students must be between 17 and 25 years of age and have no major gaps in their academic records. As such, if a student has a previous record with the university, decides to dropout but re-enters in a new major at a later date, only their initial records are included in the analysis. These restrictions are necessary and ensures that the sample includes students that are most likely to comply with the mandatory course requirements governing first year students and are unlikely to have manipulated their qualifying GPA. Table 1 shows the descriptive statistics for the entering cohort of students over the sample period 2008 to 2014. Among these students, 32% are males, the typical age at entry is about 19.26 years and 69% commute to the university from off-campus housing. On average, they are were enrolled in 15.52 credits, paid $1733.30USD in tuition and 12% financed this tuition through student loan agreements.

The outcomes of interest include subsequent GPA and cummulative GPA, the likelihood of using student loan, likelihood of transferring or changing the registration status. Additionally, the analysis examines the likelihood that students who are treated by a given program in their first year, qualify for the same treatment in future semesters. To explore the mechanism through which student's

[6]Both faculties in the second sub-sasmple generates a new Dean's list on a yearly basis. This is outlined in more details in section 2.2below.

adjust their performance following treatment by either policy, I explore whether treated students select into taking less credits or easier course. However, due to data availability limitations, students course choice selection can only be examined for the social sciences[7]. As such, students response to each policy is examined within the context of this limitation.

In measuring the level of difficulty that students select into, two main questions are asked. First, across all courses available to the student, do we observe the treated students enrolling in courses that are more generous in awarding higher grades in the past year? Secondly, if there are several instructors offering a given course at the same time, to what extent are treated students more likely to choose the stream with the professor that was more generous in awarding good grades for that course in past periods.

To explore the first question, for a given courses, we need to calculate the fraction of grades awarded as A , B, C, D and the pass rate at time $t-1$. We then assign these rates to students taking the course at time t. Since a student takes multiple courses in each semester, we average these rates within each student to get an index of the expected difficulty of the courses they take. For example, consider two courses X and Y. If 30% of students pass course X and 80% pass course Y in 2008, then a student enrolled in both X and Y in 2009 would be assigned an expected pass rate of 55%[8]. A tantamount argument can be made for the share of grades that are awarded as A, B, C and D for each course. To interpret these measures as proxies for the level of difficult, the assumption is that courses with a higher pass rate or that awarded more A's or B's are easier on average[9]. As such, if students observe the past grade distribution with enough signal, they can change their performance through this mechanism.

To assess the second question, we exploit the fact that a large cross-section of courses have multiple streams offered simultaneously and independently by various instructors. As such, within these multiple stream courses, each professor is assigned a rank that is based on the level of diffiulty of their stream(s) using the measures outlined above. A student is matched to their professor and a dummy variable is created that takes the value 1 if the student chooses the instructor with the lowest level of difficulty and 0 otherwise. For example, consider two instructors W and Z who both teaches different streams of the same course in 2008. If instructor W has the highest pass rate for this course, then any student selecting the stream taught by W would be assigned 1 and those selecting Z would be assigned 0. Similarly, this measure is averaged across courses for which students are enrolled[10].

As such, if the answer to either of these questions are in the affirmative, that would be an indication that treated students are engaging in strategic course enrolment behavior by choosing courses or professors that had a more generous grading policy for past cohorts.

[7]The data sharing agreement restricted course information to students in this faculty

[8]Weighting each course by the number of credits does not change the results because the majority of courses are for three credits.

[9]This is a good approximation if there is little turnover in the instructors and little changes in the ability distribution of incoming students. Previous studies have focused on a measure of averaged GPA to proxy for course difficulty. This measure is significantly less detailed since averages offer little information on the shape of the GPA distribution.

[10]Students enrolled in courses with only one instructor would automatically be assigned a value of 1 for those courses. From the data, 64% of courses only had a single instructor. Among courses with multiple instructors, 21% had 2 instructors and 10% had 3-4 instructors.

2.1 Dean's List: Social Sciences

The Dean's list initiative was introduced in the social sciences faculty in 2009. The impetus for this policy was to create another medium through which the faculty could recognize the outstanding academic performance of deserving undergraduate students. This list is published on the university's notice boards and website, though eligible students do not receive any other form of recognition, awards, or prizes. To be eligible for inclusion on the list each semester, students are generally required to (i) meet or exceed a semester GPA of 3.60, (ii) register for a minimum of nine credit hours, and (iii) pass all courses that are not taken on a pass/fail basis. As such, the the assignment variable this study exploits is the GPA each student receives in their first semester at the university. Within a bandwidth of 0.6, about 35% of full-time students recieved treatment by this policy at the end of their first semester over the period of interest.

The list is usually compiled and published within four weeks following the official deadline for the posting of final grades for the applicable semester. Each student is notified of his inclusion on the list by an electronic letter from the dean of the faculty. As such, students who have made the list are likely to be aware that they have achieved this accomplishment.

2.2 Dean's List: Medical, Pure and Applied Sciences

The Dean's list initiative has been longstanding in the Medical Sciences and the Pure and Applied Sciences faculties (MPAS). This program is utilized to celebrate and acknowledge the academic achievement of high-performing students.

To be eligible for the Deans list in the Medical Sciences college, students' must (i) carry a full course load and have no incomplete or failing grades (ii) have no disciplinary actions taken or pending against them, and (ii) maintain a weighted GPA above 3.6 (3.7 for specific degrees) for the previous academic year. The students who have achieved this distinction are awarded a framed certificate at an annual ceremony that is held at the end of the previous academic year. Each student is notified about his inclusion on the list about 3 weeks prior to the award ceremony by letter and email.

Similarly, to be eligible for the Deans list in the Science and Technology college, students must (i) register for a minimum of three courses that directly relate to their program of study (ii) have exam grades that are B+ and above (iii) meet or exceed a GPA above 3.6 in each semester of a given academic year, and (iv) pass all courses in both semesters or face disqualification. As such, the policy suggests that students should be treated in a given academic year if:

$$T_i = \begin{cases} 1, & \min\left(GPA_{it}, GPA_{i,t-1}\right) \geq 3.6 \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

The students meeting these criteria are invited to an annual award ceremony where they are presented with a trophy for achieving this distinction. Across both faculties, approximately 19% of students attain this honor at the end of their first academic year, within a bandwidth of 0.6 around the qualifying GPA cutoff.

The award ceremonies for both of these colleges are not open to the general student body or public. Generally, each student is allowed to invite two guests to the award ceremony to help them celebrate this accomplishment. However, unlike in the Social Sciences College, the list is not usually made public or advertised on either the university's website or faculty notice boards. As such, while the awardees are likely to be aware that they have attained this recognition, the general student body and teaching staff may be unaware of the students who have achieved this level of excellence.

For the Medical Sciences faculty, the assignment variable utilized in this study is each student weighted semester GPA at the end of their first academic year. Similarly, given the assignment to treatment equation depicted above, the running variable this study utilizes for the Science and Technology College is the minimum semester GPA a student recieves between semester 1 and 2 of their first year.

2.3 Academic Probation Policy

The university employs an academic probation policy to ensure that students performance is above the level necessary for the successful completion of their degree. This policy applies uniformly to all students enrolled in undergraduate program at the institution.

Each semester, all enrolled students are required to perform above a minimum academic standard to remain in good academic standing with the university. This standard was set to a GPA of 0.75 over the period 2008 to 2011 and then subsequently increased to a GPA of 1.00 during the years 2012 to 2013[11]. Those who failed to achieve this standard are placed on academic probation, received an official warning letter from the institution and their academic standing is changed from *GOOD* to *WARNING* on their online student transcript. About 13% of students are affected by this policy in their first semester.

This letter informs the student that his academic performance is deemed to be unsatisfactory and that failure to meet this standard in the subsequent semester would lead to him being asked to withdraw from the university. The student is also advised to take it seriously and seek counselling from his academic advisor or by accessing the resources of the academic support unit. If a student is required to withdraw as a result of the conditions outlined in this policy, he is allowed to return to his studies one year after withdrawal. Alternatively, he may apply for a withdrawal waiver from the Dean of his faculty. Consequently, the extent to which the university's threat of expulsion is viewed as credible can be questioned. In a bandwidth of 0.6 around each of the cutoffs, about 33% of students are treated by this policy at the end of their first semester and 11.5% are required to withdraw at the end of their first year. Among those who were required to withdraw, only 46% actually exited the university. As such, a large share of students are allowed to continue among those that were expected to withdraw.

Nontheless, there is an incentive to avoid being placed on academic probation. The information about the criteria for this policy is outlined in the regulations handbook provided to students at

[11]The analysis is done for the full sample period (2008-13), with a dummy variable included to control for the shift in the warning threshold from 0.75 to 1.0. The results are robust to the inclusion of this control.

registration. However, due to the size of the handbook (over 200 pages), most students may be unaware of the policy until they are bounded by it. Additionally, the university has strict course taking requirements for first year students. The extent to which students can manipulate their treatment status is explored in more details in the following sections. A copy of the warning letter treated students recieve is presented in the appendix.

3 Empirical Design

The regression discontinuity design is the main empirical method utilized in this study. This approach enables us to derive causal estimates that are locally valid and unbiased. Further, it complements the benefits derived from using administrative data and reduces the likelihood that the estimates outlined in this paper are affected by the typical causes of endogeneity such as measurement error or selection bias.

As summarized in Imbens and Lemieux (2008), McCrary (2008), and Lee (2008), there are three key assumptions that must be satisfied for causal identification using the RD approach. The first is that approaching the cutoff from either direction, there is a discontinuous change in the likelihood of being assigned to treatment once the threshold is crossed. For the warning letter program, the eligibility criteria for any given semester is solely dependent on that semester's GPA and is uniformly applied across all faculties. However, students GPA do not perfectly predict their assignment to treatment. This may be due to student's applying for grade reviews, other cases of pending grades or intervening program specific regulations. While administrators robustly enforce this policy, the relatively few cases where students GPA do not predict their treated eligibility merits the use of a fuzzy RD design. Similarly, to become qualified for Dean's list recognition, the student must satisfies a GPA requirement and several other criteria that are outlined above. Conditional on meeting the GPA requirement, the other criteria are enough to disqualify some students from making the list. As such, student's GPA do not perfectly predict their recognition on the Dean's list. To uncover the average treatment effect of the policy, the fuzzy RD design is also employed in evaluating this program. The second RD assumption requires that in a small neighborhood around the treatment eligibility GPA cutoff, those who are untreated provide a good counterfactual for those who narrowly satisfied the eligibility criteria. As such, any differences in observable student attributes within the neighborhood of the discontinuity are not systematic and are solely due to differences in the assignment variable (Bettinger et al. (2016); Leeds and DesJardins (2015)). To evaluate these assumptions, for each program, the next section provides evidence on the probability of treatment across the GPA distribution and the degree to which students characteristics are balanced in the optimal neighborhood around each program's eligibility cutoff.

The third important criteria for local randomization is that students are not able to manipulate their GPA in order to be eligible for assignment or lack thereof. This is an assumption that is frequently of concern in RD studies that uses GPA as the assignment variable. For example, Carruthers and Özek (2016) find evidence of bunching around the GPA cutoff determining eligibility for scholarship renewal. However, as Bettinger et al. (2016), Leeds and DesJardins (2015), and Cohodes and

Goodman (2014) show, when students are unaware of the GPA threshold or do not have sufficient control over changes in the assignment variable, this criterion is often not violated. Similarly, we intuitively expect to find no evidence of manipulation of the running variable in this paper due to four main reasons: (a) at this university, students in their first year are required to complete a set of core mandatory prerequisite courses which they must pass in order to qualify for entry into second year courses; (b) the university requires that each full-time student take a minimum amount of credits each year or withdraw from the university; (c) all mid-term and final examinations are graded using at least a single-blind process. That is, the student is always required to use his university identification number in place of his name and the grading of undergraduate papers is usually done by graduate students and (d) the university policy stipulates that all exam scripts ought to be independently reviewed by both a primary and secondary grader before final grades can be certified by the university registrar. In support of this theory, the nonparametric local polynomial density plots and statistics recommended in Cattaneo et al. (2016a) are presented below. The procedure conduct a hypothesis test to determine if the density of the running variable is continuous around the cutoff[12]. That is, we are interested in evaluating the null hypotheses that $H_0 : \lim_{x\uparrow\bar{x}} f(x) = \lim_{x\downarrow\bar{x}} f(x)$, signaling no manipulation at the cutoff[13].

Given that treatment by the program is only partly determined by the established GPA cutoff, an instrumental variables approach is necessary to estimate the treatment effect. The first stage is given by:

$$D_i = \gamma_0 + \sum_{j=1}^{k}[\gamma_j PriGPA_i^j \times T_i] + \gamma_3 T_i + \sum_{j=1}^{k}[\rho_j PriGPA_i^j] + \gamma_r + \varepsilon_i$$

For each policy of interest, T_i measures the treatment status of student i that is implied by their GPA and is defined by the indicator function $1(PriGPA_i > 0)$; D_i is a dummy variable that takes the value 1 if the student was treated and zero otherwise, and $PriGPA$ denotes each student's first semester GPA measured in deviation from the established eligibility threshold[14]. From the first stage regression, we can assess the extent to which assignment to treatment using the running variable predict students observed treatment status. The second stage can then outlined as follows:

[12]This procedure is very similar to McCrary (2008). However, McCrary (2008) requires pre-binning of the data and introduces additional tuning parameters.

[13]However, some pundits may argue that given that in a given semester, each student takes a finite number of courses and receives a set number of possible grades, then an assignment variable based solely on GPA is discrete. In the context of this study, there is some merit to this argument, given that within a GPA bandwidth of 0.6 around the cutoff, there are only about 100 unique values that the qualifying GPA takes on for each policy. Frandsen (2017) argues that when the running variable is discrete, then the existing RD manipulation tests are inconsistent. For such case, he proposes a test that is consistent and locally unbiased relying only on support points at and immediately adjacent to the RD threshold. Given this criticism, while this study presents the manipulation test proposed in Cattaneo et al. (2016a), the procedure outlined in Frandsen (2014) are conducted for robustness. The evidence in support of the null hypothesis of no manipulation is strong when the manipulation test proposed in both Cattaneo et al. (2016a) and Frandsen (2017) are utilized.

[14]For either policy, a student with a positive PriGPA has met the GPA requirement for treatment. This ensures that a positive α_3 can be interpreted as the policy having a positive treatment effect. To ensure this ease of interpretability, in evaluating the warning letter policy $PriGPA = -(GPA - c)$, since a student is treated by this policy when they receive a GPA below c.

$$Y_{is} = \alpha_0 + \sum_{j=1}^{k}[\alpha_j PriGPA_i^j \times D_i] + \alpha_3 D_i + \sum_{j=1}^{k}[\pi_j PriGPA_i^j] + \nu_s + \nu_r + u_{is}$$

Where Y_{is} is the academic decision or outcome variable for student i in semester s and k is the degree of the local polynomial. The parameter of interest is α_3. All the model specifications presented in this paper includes controls for semester and grading regulations fixed effects[15].This parameter provides an estimate of the causal impact of each policy on students' subsequent behavior. These estimates are presented as the conventional RD estimates in the results section.

As recommended by Gelman and Imbens (2014), we refrain from the use of higher order polynomials in the implementation of the RD design in this study. In particular, all baseline estimation and bandwidth selection in this paper use a linear polynomial specification ($k = 1$). However, the results are robust to the choice of k. Following Fan and Gijbels (1996) calculation that the triangular kernel is optimal for estimating local linear regressions at the boundary, all estimation procedures in this paper employs a triangular kernel, such that the highest weights are given to observations closest to the cutoff and the weight given to each observation decreases linearly as $PriGPA$ tends to the bandwidth size h. The consensus in the RD literature is that in practice, the choice of kernel is usually inconsequential (Lee and Lemieux (2010))[16]. This study also employs the approach recommended by Lee and Card (2008), clustering the standard errors by each value of normalized first semester GPA. This clustering is necessary because we have repeated observations with a single $PriGPA$ and because changes in running variable occurs at a discrete rate of 0.01.

Calonico et al. (2014a) argue that the conventional nonparametric local polynomial estimates of the treatment effect have confidence intervals that may be biased, and as such, it may substantially over-reject the null hypothesis of the non-existence of a treatment effect. This is because in estimating the conventional treatment effect, an approximation of the regression function is required on both sides of the threshold using weighted polynomial regressions. These weights are generated by means of a kernel function which is based on the distance of each observation's running variable relative to the eligibility cutoff and the optimal bandwidth selected. They posit that the usual optimal bandwidth selectors which balance the squared-bias and variance of the RD estimator, yield bandwidth choices that are too large to meet invoked distributional approximations. They propose the RD confidence intervals for treatment effects that are robust to bandwidth choices such as those obtained from selection using cross-validation or mean-squared error minimization. In the first procedure, they directly bias-correct the estimator by proposing an MSE-optimal bandwidth choice as an estimate of the asymptotic bias. The proposed robust confidence interval is based on the bias-corrected RD treatment effect, but it utilizes a different variance structure that accounts for the additional variability introduced by the biased estimate.

In selecting the optimal bandwidth around the threshold, the balance of covariates test and

[15]The grading policy governing enrolled students were reformed in 2012 and again in 2014. Including fixed effect for these changes may improve the standard error of the estimates. However, they had no practical impact on the estimated parameters

[16]The choice of the kernel has no impact on the estimated treatment effects in this study.

bandwidth selection procedure proposed by Cattaneo et al. (2016b) and Calonico et al. (2014b) are utilized. The first uses finite-sample methods to select the optimal neighborhood around the eligibility threshold where the randomization assumption is most plausibly satisfied. The bandwidth selected using this approach ensures that students' age, gender, credits attempted, commuting and loan status are balanced on both sides of the threshold. The second is a data driven bandwidth method which provides an MSE optimal bandwidth choice.

4 Empirical Results

This section presents formal test statistics and graphical evidence supporting the hypothesis that the qualifying GPA of students are not manipulated within the optimal bandwidth around the policy cutoffs examined. Further, evidence is shown that the observed covariates between the treated and untreated groups are balanced. Sections 5.2 and 5.3 then discusses the estimates from the main specification for the Dean's list and warning letter policies respectively.

4.1 Satisfying the RD Assumptions

Potential subjects' manipulation of their eligibility status is the main concern of applied researchers when using the regression discontinuity design. In the context of this study, this becomes an issue if students who are aware of the treatment thresholds are incentivized to change their behavior to influence their treatment eligibility status. This is especially true for the marginal students who expect their performance to be close enough to the treatment threshold. As such, these students may be motivated to exert a higher level of effort or enroll in what they perceive to be easier classes in order to avoid getting a warning letter or to become qualified for Dean's list recognition. Additionally, such students may lower the number of credits they pursue in a given semester so that they may devote more time and effort per enrolled course. While these concerns are usually merited, the course enrollment and grading regulations at the examined university makes a number of the usual methods of GPA manipulation unlikely in this study. In their first year, all students at this university are required to complete a mandatory set of core courses if they are to satisfy prerequisite requirements for the second year (level II) courses. Furthermore, there is a minimum number of credits that students must successfully complete in any given academic year or they are mandatorily required to withdraw from the university. The student is also prevented from manipulating the threshold by appealing for leniency from professors since students are required to only provide their student identification number on examination documents and each script is independently evaluated by at least two graders[17].

If students manipulate their GPA and non-randomly sort themselves to treatment, this would lead to a violation of the local randomization assumption. As such, we would observe a discontinuous change in the density of the running variable at the eligibility cutoff, and the distribution of observ-

[17]It is also reasonable to expect that the majority of students are unaware of the academic warning policy in their first semester

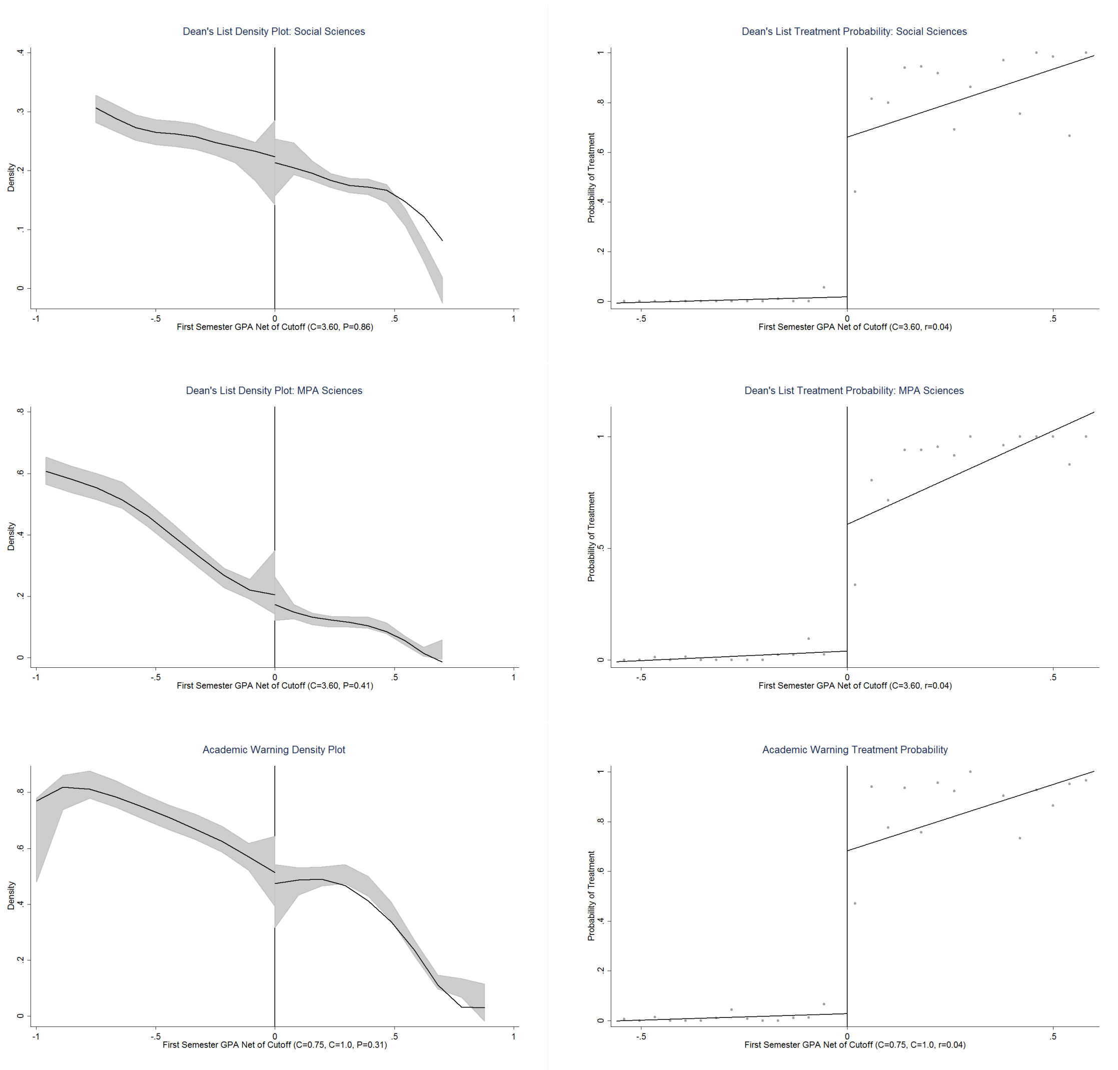

Figure 1: Regression Discontinuity Assumptions: Manipulation, Discontinous Jump
Where C is the program cutoff, P is the p-value of the manipulation test and r is the bin width size.
Each dot is the mean treatment within an interval of r around the point.

able characteristics on both sides of the cutoff within the optimal bandwidth would be significantly different. Using the density test outlined in Cattaneo et al. (2016a), we fail to reject the null hypothesis of no manipulation for both policies and each sub-sample. This result is further supported by the density plots presented in figure 1, showing that students' qualifying GPA is continuous across each program's cutoff. This implies that within this university environment, students are dissuaded from changing their behavior to be treated by these policies.

To further support this conclusion, we can examine evidence on the extent to which the observed characteristics between the treated and untreated groups are balanced. Table 2 show that across policies and sub-samples, the treated and untreated groups had no significant statistical differences in their gender, age, loan status, attempted credits and commuting. Together, these results provides strong evidence that students are unable to manipulate their qualifying GPA. As such, it is reasonable to expect that there should be no systematic differences in the unobserved attributes of students, and that receiving a warning letter or being recognized on the Dean's list in the can be considered random in the neighborhood around the eligibility cutoff.

Figures 1 also provide evidence that that the probability of treatment discontinuously increases once a student's GPA exceeds the eligibility threshold. This supports the notion that the programs administrators largely adhered to the GPA eligibility criteria as outline in the institution's policy documents. However, a student's treatment status is not a perfect function of their GPA. The last column in Table 2 shows the first stage results of the fuzzy regression discontinuity for each policy. For both policy, the treatment status implied by a student's GPA largely predicts their observed treatment status.

4.2 Dean's List Results

For the faculty of Social Sciences, the estimates presented in columns 1 and 2 of table 3 indicate that students who make the Dean's list in their first semester experience a better academic trajectory than the students' who were narrowly ineligible. On average, these students were found to subsequently obtain a higher GPA and have a greater likelihood of making the dean's list in future semesters relative to the untreated students. In this faculty, the impact of the dean's list were found to be immediate and persistent. The students who are eligible in their first semester experienced a substantial improvement in their GPA in semesters 2 to 3, and this difference persists until the time when the student is expected to graduate. In particular, across all future semesters, the treated students obtained a semester GPA that is 0.37 points higher relative to that of the untreated students. This represent a substantial improvement for the treated students, accounting for approximately 12% of the average GPA in the sample. Further, the treated maintain a degree GPA that is roughly 0.36 points higher than similarly students that were not treated.

For the MPAS sample, the Dean's list have a less intensive impact on students' subsequent behavior. Similar to the Social Sciences, the results presented in table 4 indicate that treated students experience a higher GPA in subsequent semesters. On average, those who were treated obtained a GPA that was about 0.16 points higher, a difference that roughly accounts for 4.8% of the average GPA in the sample. However, there is a neglible impact on students degree GPA and no persistent effect on their likelihood of making the Dean's list in future semesters.

The differences in the implementation of the Dean's list initiative across colleges may account for the observed differences in the impact of the program on student behavior[18]. The method utilized to recognize student's accomplishment and the awards students receive varies profusely across both samples. In particular, the social sciences faculty publicly displays a physical copy of the Dean's list on the university website and college notice boards. However, students who meet the requirements and recieves this distinction receives no additional benefits, awards or prizes. In contrast, the MPA sciences do not publicly display a physical copy of the Dean's list and students who achieve this

[18]There may be other unobservable differences between high achieving students in both faculties that could plausibly explain the differences in the results

recognition are invited to a ceremony where they presented with certificates or trophies for this achievement. Another key distinction between the two sets of policies is that while the social sciences creates a new list at the end of each semester, the MPA sciences recognizes students at the end of each academic year.

Therefore, the mechanism which accounts for the differences in the results may be the degree of public recognition and/or the frequency of the intervention. In both cases, students receive some positive benefits from having their accomplishment recognized by their institution. However, while recognition in one college is done in an award ceremony that is closed to the general public, the recognition students receive in the other college mainly offers public adulation. As such, students on the Dean's list in the Social Sciences will have a greater incentive to maintain this achievement if students are very sensitive to public humiliation from failing to consistently make the list.

Together, the estimated parameters across both samples provides supporting evidence that the Dean's list program induces students into performing better in future semesters, though the magnitude and persistence of the impact may vary depending on how the program is designed. There are several mechanisms through which treated students may attain this benefit in subsequent semesters. On the one hand, this policy may induce students into exerting a higher level of effort to maintain the recognition they have attained. For example, Seaver and Quarton (1973) notes that the Dean's list may be a powerful social reinforcer of the behaviors leading to academic achievement and it should improve students' self-esteem and their expectation for future performance. Similarly, Thistlethwaite and Campbell (1960) posits that public recognition has a motivational effect that ought to improve students' subsequent academic aspirations and outcomes. If this is the mechanism at work, we would expect treated students to invest in no less credits and choose courses that are equally difficult relative to those who were marginally ineligible for treatment. This would align with the administrators vision of the program, since the goal of the program is to recognize those that have distinguished themselves as high-performing and model students. On the other hand, being publicly recognized for academic performance could induce students into making myopic perverse choices in an attempt to continue making the list. This is because if they fail to make the list again, the student would lose the positive social reinforcement they have attained or potentially even face public humiliated. If we observe treated students reducing their credit load or opting for easier courses relative to the control group, this can be taken as evidence in support of this mechanism. These action lead to long-term adverse consequencies on the quality of education the student receives and would offset some of the gains of the improved performance that is observed.

The estimates in column 4 of tables 3 and 4 support the theory that the recognition treated student's receive increases the level of effort they subsequently exert. In the Social Sciences, treated students increase the number of credits they attempted in semesters 2 and 3, though by a very small margin. However, the Dean's list has no total impact on students credit taking behavior across all subsequent semesters following treatment to the expected date of graduation. For the MPAS, students who are treated by the Dean's list are observed enrolling in more credits than the control group. These results suggest that the improved performance treated students receive cannot be explained through their credit taking behavior. In fact, treated students are induced into accumulate more credits compared to the control group.

Due to data availability restrictions, the analysis regarding course selection focuses primarily on the social sciences. As such, we can explore how the Dean's list policy affect students course selection behavior in this faculty. The result for this analysis is presented in row 1 of table 6.[19] Columns 1-6 show the difference in the difficulty of courses that the treated and control groups take. From this measure, it can be seen that students who make the Dean's list are induced into selecting courses that are easier relative to those taken by the control group. The treated group is observed enrolling in courses that awards 3.7% less D's, has a 4.9% higher pass rate and awards about 5% more B's and 2% more A's than those who did not make the list in their first semester. This implies that treated students becomes more aversed to selecting courses with a high failure rate and are incentivize to choose courses that are more likely to award grades that are B and above. For courses with multiple instructors, each instructor is assigned a ranked based on the percent of A's and B's they award and the pass rate of the stream(s) they teach. The results show that students that are treated by the Dean's list are 15% more likely to select the instructor that previously awarded the highest fraction of A's and B's and are 21% more likely to select instructors that has the highest pass rate, among the available within-course choice set of professors.

Jointly, these results indicate that treated students are able to ascertain enough signal about the level of difficulty of courses and instructors based on the distribution of past grade. They utilize these signals to engage in strategic course taking behaviors that could partially account for their improved performance in future semesters. However, given that the treated group are enrolled in courses that awards only slighly more A's and B's, it is implausible to expect that the improved performance is being overwhelmingly driven by course selection. As such, it is likely that treated students are induced into both increasing their level of effort as well as engaging in strategic course-taking behavior to maintain a place on the Dean's list in subsequent semesters.

Since it is not clear if students are trading off the content they learn by choosing easier courses or instructors, it is difficult to assess the long term implications of these actions. However, the results do suggest that there are tangible benefits that treated students receive in the medium term not restricted to academic performance. For example, there is some evidence that the treated students in the social sciences may have been more likely to benefit from financial assistantships that reduced their likelihood of using a loan by about 29% in their third year. Further, these students were more likely to graduate on time or graduate at no worse durations relative to those who were not treated.

4.3 Academic Probation Results

The academic probation policy is another mechanism through which university administrators provide feedback to students about the quality of their academic performance relative to expected university standards. The students who fail to meet or exceed this level of academic performance

[19]row 2 of table 6 provide estimates for the MPAS sample. However, these are not very precise because they are based on the out of faculty courses taken by MPAS students. Nonetheless, they show the same general trend.

receive a letter that formally places them on academic warning. If this poor performance persist in the subsequent semester, the regulation requires that the student be asked to withdraw from the institution. As such, after students are warned in their first semester, they have to make key decisions about the courses they select, whether to exit and if they should persist in the major they have selected when they started the university. The results below shows the effect of the academic probation policy on students subsequent academic performance and explores the mechanism through which students adjust their behavior following treatment.

Among the students that continue their studies beyond the first year, the estimates in columns 1 and 2 of table 5 shows a modest improvement in their subsequent academic performance. The overall improvement in degree GPA of about 0.16 accounts for roughly 9% of the control group's mean degree GPA over the sample period. However, the estimated parameters are insignicant at conventional levels. This may be partly explained by students' heterogeneous reponse to the policy, where the average effect masks substantial heterogeneity across groups, as argued in Lindo et al. (2010). If we restrict our focus to students in the social sciences as is done in panel E, the estimated parameters suggests a much larger improvement in their subsequent degree GPA. These estimates are directly comparable in direction and magnitude to those presented in Lindo et al. (2010). However, the average response suggests that these strong positive effects for some students are offset by non-positive performance response for other segments of the student population. The impact of being warned in the first semester appears to become effective during the second year, with students response dissipating as they progress. This provides an opportunity for innovation, where the potential benefits of this policy can be greater exploited through sustained engagement with low-performing students[20].

The estimates presented in columns 5-7 indicate that this policy improves the match quality between the university and the enrolled student population. The students who receive a warning letter in their first semester are on average 9% more likely to exit the university at the end of the first academic year relative to the untreated students. This decision to exit is driven mainly by the withdrawal requirement of the policy, rather than being the voluntary decision of students. This is supported by fact that there is no impact on the decision to exit at the end of the first semester or at the end of the second year[21]. The results also shows that the students who are placed on academic probation in their first semester are 9% more likely to switch to a new major and 8% more likely to transfer to a new faculty in their second year relative to students in the control group. This suggests that changes in the program of study is one mechanism through which students adjust after being placed on academic warning. [22] Together, these findings suggests this policy induces weaker treated

[20]Several universities are now applying data analytics to identify weak performing students and provide academic support to help them progress towards graduation. These universities have seen significant improvements in their graduation rates

[21]The exit column captures students decision as of the end of the previous or the beginning of the current semester. For example, panel A captures the decision to exit before at the end of semester 1 or panel B captures the decision to exit at the end of semester 2. This explains the non-conformity in the observations with other outcomes.

[22]The university administrators view transfers and major change as the main response of students that perform poorly in their first year. They argue that poor performing students from the pure and applied sciences faculty switch to degree programs in the social sciences faculties at the end of their first year. Citing anecdotal evidence, they suggest that these transfers can be explained by the general perception helb by pure and applied sciences students that social sciences courses are easier on average. This trend is confirmed by the data with 85% of all transfers within the optimal

students to either exit the university or change their program of study. As such, students that are not a good fit for the university have their matches terminated and others are allowed to make new matches with majors that may better fit their academic ability. However, with a mere 9% differential in the likelihood of students switching majors, the impact of these decisions on students subsequent performance may not be large, but certainly not negligible.

The results suggest that treated students also adjust their behavior along other dimensions not restricted to effort or match quality. The result in column 4 suggests a small reduction in the number of credits treated students take in their second year by 0.8 or 5% of the sample mean for the control group. This is consistent with the magnitude presented in previous studies. However, while previous studies did not find any evidence suggesting that being placed on probation causes continuing students to enroll in easier courses, the estimates in row 3 of table 6 contradicts this result. Lindo et al. (2010) posits that using average course grades as the measure of difficulty, treated students do not enrol in easier courses. Similarly, the results suggests that treated students do not select into courses that has a higher pass rate[23]. However, when the level of difficulty is measured as the share of the grade distribution allocated as A's, B's, C's and D's, a completely different picture arises. Treated students appear to follow a maximin strategy when choosing their courses. As such, they select into courses that have higher density in the lower tail and lower density in the upper tail, awarding more C's and D's but less A's. On average, these students enroll in courses that awards about 6.8% more C's and D's relative to the courses taken by the control group. However, there is no evidence that when treated students are given the option of choosing among several instructors teaching the same course, they were relatively more selective to choosing those the professors that were generous in passing students or awarding higher grades[24]. It is not clear if this maximin selection strategy helps the treated students to perform better in the short term, given that there is no difference in the pass rate and they are less likely to get grades in the upper tail of the distribution. However, there are no evidence that their behavioral changes has any long term impacts on their graduation rates. Treated students had an equal likelihood of graduating and were also equally likely to graduate in three years or four years.

The results established for the academic probation policy in this paper offers mixed support for those established in the literature, while offering more robust clarification for the mechanisms at work. For example, Lindo et. al. (2010) argues that this policy induces students to voluntarily exit the university and exert a higher level of effort in future semesters. Additionally, they do not find any evidence that students adjust their behavior along other dimensions unrelated to their effort choice. In contrast, the results presented above shows that the probation policy influences the match quality between students and the university through exits, transfers and major switches. It also induces the treated students that continue to change their course taking behavior along several dimensions. As such, not much can be said about the impact of the policy on student's subsequent

bandwidth consisting of flows from the pure and applied sciences to the social sciences faculty. This could be an effective strategy if students have enough signal belief about courses they switch are correct.

[23]This study was also able to confirm their result using average course grade as the measure of difficulty

[24]The same results are obtained when instructors are ranking using the percent of C's and D's they award

effort decisions. It is not clear the extent to which key institutional differences across studies may account for these contraditions[25]. There is much that future studies need to examine for us to garner a fulsome understanding of the impact of these policies on students behavior.

4.3.1 What is the Impact of Making the Probation Policy More Restrictive?

In the 2014-2015 academic year, the university administrators implemented several changes to the institution's grading and degree granting regulations. These changes increased the cutoffs for each letter grade and also changed the warning letter threshold from 1.0 to 2.0. This policy change is a natural experiment that be exploited to test the sensitivity of the main results to the generosity of the warning letter policy. That is, it allows us to examine how the impact of the policy varies as the performance threshold becomes more restrictive and binding on students that are higher up in the grade distribution[26].

The estimated impact of the 2.0 probation policy presented in table 9 largely mirrors that of 1.0 policy presented in tables 5 and 8. The probation policy leads to a small improvement in degree GPA of about 0.1-0.13 or about 5-6% of the control group's mean degree GPA over the sample period. This is similar to the 9% that was found for the 1.0 policy. Similarly, the results indicate that treated students were more likely to switch their majors and faculty of study at rates that are similar in magnitude to those previously estimated. However, these parameters only become significant as the bandwidth size increases. Finally, the impact on involuntary exits reduces in magnitude and is no longer significant[27].

As such, the conclusion remains the same as reported for the baseline estimates. That is, the evidence implies that the small improvement in students subsequent performance may not be mainly driven by an increase in the level of effort they exert. The key insight of this result is that when the probation environment becomes more restrictive and ability of the marginally treated increases, the effect of the policy remains unchanged and the channels through which students adjust their behavior is largely unaffected.

4.4 Falsification Tests

While great care is taken in selecting the optimal bandwidth, some may rightfully be skeptical about the robustness of the results to the bandwidth choice. As such, the estimates from the main RD specifications are evaluated using a falsification tests. The sensitivity of the estimated parameters are assessed by examinng the impact of changing the optimal qualifying GPA bandwidth within the interval [0.2, 0.6], for each stepwise unit of 0.1[28].

[25]The most notable institutional difference is that at the institution being examined, approximately one-half of the students that are required to withdraw receive a waiver of this requirement. Additionally, the probation policy cutoff examined in this paper is 1.00 and 2.00, compared to 1.60 in Lindo et al. (2010)

[26]With the revised grading regulation, approximately 30% of students are affected by the policy at the end of their first semester. This compares to 13% being affected under the prior threshold

[27]There is still no evidence of an effect on voluntary exits.

[28]We arrive at similar conclusions when the falsification test of shifting the cutoff is utilized.

Table 7 shows the falsification results for the Dean's list policy and both the social sciences and MPA sciences samples. Across all of the bandwidth choices, the estimated effect of the policy is largely consistent in direction. As the bandwidth increases, there are no major improvements in the efficiency of the estimated parameters, but the estimated impact on students' subsequent performance decays quickly and monotonically as the bandwidth interval tends to 0.6. This suggests a downward bias in the estimated parameters further away from the cutoff. It also indicates that the standard errors estimates are reliabile for bandwidth choices relatively close to zero. From this exercise, it becomes more convincing that the policy has a positive impact on students' subsequent performance in the social sciences, but a negligible impact in the medical, pure and applied sciences faculties.

Table 8 presents the results for the academic probation policy. For all potential bandwidth sizes within the interval of interest, the estimated parameters are consistent in both direction and magnitude. As the bandwidth size tends to the upper limit of the interval of interest, all the estimates remain relatively stable. As such, the estimates and conclusions from table 5 are very robust to the bandwidth choice.

5 Conclusion

This study examines the extent to which college students are incentivized to change their behavior when they receive administrative feedback that either reprimand or reward them for their past academic performance. To explore the effects of praising, rewarding or reprimanding students for their past performance, this paper examines the impact of the academic probation and Dean's list policies on college students' academic decision-making. Further, the mechanisms through which students treated by these policies subsequently alter their behavior is extensively documented and discussed. The main conclusions outlined in this study are applicable in most context where an agent's past performance is used as a basis for reprimand or recognition, in an attempt to influence their future performance. For example, in the service industry, workers performance are frequently appraised and those performing exceptionally are often publicly recognized. Similarly, several HR departments utilize a warning letter system to discourage employees from continuously engaging in unproductive behaviors[29]. Employees who exhibit these behavior receive a warning letter for their first offense and are terminated if this behavior persists.

Consistent with the few studies in the literature, I find that the Dean's list policy improves the academic performance of treated students in subsequent semesters. However, the intensity of the effect seem to vary by the design and implementation of the program. In particular, the frequency of the intervention and the degree of public recognition seem to be important policy conditions that influence how students respond to being treated. The result also found mixed support for the mechanism through which treated students adjust their subsequent behavior. The evidence suggests that treated students are incentivized to engage in strategic course taking behavior. However, the

[29] Such behavior includes showing up to work late or failing to meet some set targets among other things

degree of this adjustment does not seem large enough to fully explain the estimated effect of the policy on academic performance. As such, treated students likely improve their subsequent performance through both their effort and course taking decisions.

The results for the academic probation policy offers mixed support for findings in the literature. This study finds that treated students had a modest improvement in their subsequent academic performance. However, this improvement masks significant heterogeniety in how students responded to the policy. Consistent with the literature, this study finds that the policy causes low-ability students to exit the university at a relatively high rate. However, unlike previous studies, the decision to exit is exclusively involuntary. The evidence suggests that treated students are incentivized to improve their match quality by transferring across faculties or switching their major of study. There is also evidence that the policy induces students to adjust their credit taking and course selection behavior. Students are observed following a maximin strategy when selecting courses, by choosing courses that have higher density in the lower tail and less density in the upper tail. As such, this study provides evidence that refutes the claim that improvements in student's subsequent GPA can be explained primarily by increased effort. It also suggests that the data utilized in this study does not fit the simple theoretical performance standards model presented in previous papers.

These results demonstrate that a university administrator's action of reprimanding or rewarding students for their past academic performance may induce significant behavioral changes. While there is no evidence that either policy have resulted in any long-run negative outcomes, the results suggests that the positive gains to student academic performance may not be solely driven by increased effort. The positive impact of being warned in the first semester appears to become effective during the second year, with students response dissipating as they progress. This provides an opportunity for innovation, where the potential benefits of this policy can be greater exploited through university's sustained engagement with low-performing students. This is already being utilized by some universities that uses a data-analytics driven approach to targeting low-performing students and offering them greater assistance that improve their likelihood of success. This approach is now feasible due to significant improvements in technology and it may yet show greater promise in promoting student success than a static warning letter system. The results also offer insights that policymakers must consider when rewarding high performing students. One recommendation is for administrators to augment such programs with academic advising on course selection, academic support and stress management facilities so that students can effectively manage the expectation of continued good performance that such programs may induce.

6 Appendix: Results

Table 1: Descriptive Statistics

	Mean	Std Deviation	Obs.
Student Characteristics			
Male	0.32	0.46	20580
Age	19.26	1.50	20580
Commute	0.69	0.46	20580
Tuition ($US)	1733.3	2189.8	20580
Outcomes			
Credits	15.52	2.59	20580
GPA	2.09	1.12	20580
Loan	0.12	0.32	20580
Treatment w/n bw of 0.6			
Warning Letter 1st Semester	0.33	0.47	2175
Dean's List 1st Sem: SS	.35	0.48	1551
Dean's List 1st Year: MPAS	0.19	0.39	1545

Table 2: Local Randomization Assumption

	Male	Age	Credits	Commute	Loan	Treatment
Dean's List						
Social Sciences	-0.077	-0.22	-0.04	-0.019	0.008	0.57***
	(0.047)	(0.31)	(0.33)	(0.098)	(0.059)	(0.15)
Mean	0.26	19.18	14.77	0.76	0.085	0.42
Obs	625	625	625	625	625	625
MPA Sciences	-0.0351	0.166	0.314	0.111	0.0552	0.612***
	(0.118)	(0.466)	(2.070)	(0.0955)	(0.0684)	(0.107)
Mean	0.34	19.26	17.72	0.67	0.10	0.26
Obs	636	636	636	636	636	636
Warning Letter						
$C = 0.75$, $C = 1.0$	0.009	-0.096	0.44	-0.01	-0.07	0.96***
	(0.0755)	(0.0896)	(1.039)	(0.0851)	(0.0524)	(0.0502)
Mean	0.37	18.89	15.68	0.68	0.09	0.36
Obs	1013	1013	1013	1013	1013	1013

Table 3: Impact of the Dean's List Policy: Social Sciences

	GPA	CGPA	Dean List	Credits	Loan
Panel A: Sem 2 to 4					
Robust	0.50***	0.44***	0.29***	-0.30	-0.26***
	(0.11)	(0.07)	(0.08)	(0.25)	(0.08)
Conventional	0.24*	0.19*	0.13	-0.19	-0.064
	(0.13)	(0.11)	(0.08)	(0.16)	(0.12)
Obs	1738	1738	1738	1738	1738
Panel B: Sem 5 to 6					
Robust	0.22	0.26**	0.11	0.34	-0.32***
	(0.14)	(0.11)	(0.10)	(0.36)	(0.066)
Conventional	0.15*	0.16*	0.16**	-0.14	-0.11
	(0.085)	(0.089)	(0.079)	(0.30)	(0.10)
Obs	985	985	985	985	985
Panel B: Sem 2 to 6					
Robust	0.37***	0.36***	0.22***	-0.08	-0.29***
	(0.087)	(0.075)	(0.079)	(0.25)	(0.07)
Conventional	0.19**	0.17*	0.14**	-0.20	-0.09
	(0.09)	(0.096)	(0.067)	(0.17)	(0.11)
Obs	2723	2723	2723	2723	2723
Mean	3.18	3.27	0.28	14.83	0.11
SD	0.67	0.35	0.45	2.24	0.32

Table 4: Impact of the Dean's List Policy: Medical, Pure and Applied Sciences

	GPA	CGPA	Dean List	Credits	Loan
Panel A: Sem 3 & 4					
Robust	0.08	-0.04	0.0006	1.83	0.013
	(0.23)	(0.07)	(0.179)	(1.899)	(0.147)
Conventional	0.23	0.04	0.181	1.58	-0.037
	(0.161)	(0.05)	(0.129)	(1.246)	(0.103)
Obs	1096	1096	1096	1096	1096
Panel B: Sem 5 to 6					
Robust	0.23	0.03	-0.130	3.9*	-0.049
	(0.20)	(0.08)	(0.197)	(2.117)	(0.124)
Conventional	0.20	0.10*	-0.0596	3.9***	-0.07
	(0.13)	(0.06)	(0.135)	(1.488)	(0.086)
Obs	1061	1061	1061	1061	1061
Panel B: Sem 3 to 6					
Robust	0.16	-0.005	-0.0581	2.75*	-0.023
	(0.17)	(0.07)	(0.143)	(1.673)	(0.124)
Conventional	0.22*	0.07	0.06	2.70**	-0.052
	(0.12)	(0.05)	(0.10)	(1.16)	(0.085)
Obs	2157	2157	2157	2157	2157
Mean	3.33	3.42	0.32	18.96	0.16
SD	0.56	0.26	0.47	4.97	0.37

Table 5: Impact of the Academic Warning Policy: Threshold of 0.75 & 1.0

	GPA	CGPA	Warning	Credits	Major	Transfer	Exit
Panel A: Sem 2							
Robust	-0.25	-0.07	0.20	-0.18	0	0	-0.004
	(0.192)	(0.116)	(0.17)	(1.977)			(0.0400)
Conventional	-0.003	-0.03	0.06	0.08	0	0	-0.003
	(0.158)	(0.0834)	(0.0985)	(1.254)			(0.0271)
Obs	980	980	980	980	980	980	1013
Panel B: Sem 3 & 4							
Robust	0.15	0.16	0.009	-0.79*	0.094**	0.077**	0.11**
	(0.237)	(0.138)	(0.0796)	(0.465)	(0.0469)	(0.0387)	(0.0477)
Conventional	0.18	0.13	0.02	-0.41	0.056	0.057*	0.10***
	(0.16)	(0.0923)	(0.0586)	(0.379)	(0.0389)	(0.0312)	(0.0341)
Obs	1716	1716	1716	1716	1716	1716	990
Panel C: Sem 5 & 6							
Robust	-0.11	0.16	0.02	-0.10	0.02	0.075*	0.07
	(0.254)	(0.164)	(0.0650)	(0.772)	(0.0759)	(0.0448)	(0.0751)
Conventional	-0.07	0.13	0.024	0.15	0.04	0.076**	0.07
	(0.171)	(0.103)	(0.0468)	(0.462)	(0.0574)	(0.0358)	(0.0504)
Obs	1494	1494	1494	1494	1494	1494	849
Panel D: Sem 3 to 6							
Robust	0.03	0.16	0.01	-0.47	0.06	0.076*	0.09**
	(0.224)	(0.141)	(0.0616)	(0.572)	(0.0578)	(0.0410)	(0.0392)
Conventional	0.07	0.13	0.02	-0.15	0.05	0.066**	0.088***
	(0.156)	(0.0919)	(0.0473)	(0.389)	(0.0441)	(0.0332)	(0.0277)
Obs	3210	3210	3210	3210	3210	3210	1839
Panel E: Sub-Sample							
Robust	0.17	0.28***	-0.098**	-0.64**	0.095***	0.14***	0.06
	(0.189)	(0.0976)	(0.0472)	(0.261)	(0.0309)	(0.00897)	(0.0530)
Conventional	0.11	0.21***	-0.003	-0.27	0.11**	0.083***	0.06
	(0.136)	(0.0756)	(0.0621)	(0.280)	(0.0511)	(0.0296)	(0.0402)
Obs	1620	1620	1620	1620	1574	875	

Table 6: Policies Impact on Course Selection Behavior

	A-rate	B-rate	C-rate	D-rate	P-rate	Easy (A,B)	Easy(P-rate)
SS	0.02	0.05***	0.012	-0.037***	0.049***	0.15***	0.21***
	(0.0186)	(0.0187)	(0.00820)	(0.0101)	(0.0146)	(0.024)	(0.018)
MPAS	-0.03	0.07	-0.06	-0.07	-0.09	0.11	-0.01
	(0.0832)	(0.0954)	(0.0789)	(0.0698)	(0.108)	(0.247)	(0.18)
Warn, $c = 1.0$	-0.058***	-0.012	0.03	0.038***	-0.004	-0.02	-0.01
	(0.00981)	(0.0197)	(0.0197)	(0.00885)	(0.0306)	(0.0419)	(0.034)
Warn, $c = 2.0$	0.002	0.019	0.005	-0.003	0.02**	0.05	0.09**
	(0.0134)	(0.0127)	(0.0189)	(0.00465)	(0.00885)	(0.0482)	(0.0366)

Table 7: Robustness of RD Estimates: Dean's List Policy

	GPA	CGPA	Dean List	Credits	Loan	Obs
Panel A: Social Sciences						
BW of 0.2	0.40***	0.39***	0.20**	-0.14	-0.36***	2436
	(0.10)	(0.085)	(0.089)	(0.28)	(0.07)	
BW of 0.3	0.33***	0.30***	0.21***	-0.18	-0.21**	3485
	(0.10)	(0.086)	(0.072)	(0.23)	(0.10)	
BW of 0.4	0.25***	0.22**	0.15**	-0.24	-0.14	4523
	(0.095)	(0.092)	(0.075)	(0.19)	(0.10)	
BW of 0.5	0.13	0.12	0.11	-0.22	-0.06	5944
	(0.093)	(0.093)	(0.070)	(0.17)	(0.098)	
BW of 0.6	0.09	0.09	0.09	-0.14	-0.05	7098
	(0.093)	(0.09)	(0.069)	(0.15)	(0.089)	
Panel B: MPA Sciences						
BW of 0.2	0.33	0.0723	0.0300	2.76	-0.047	1356
	(0.216)	(0.0825)	(0.187)	(2.084)	(0.154)	
BW of 0.3	0.16	-0.005	-0.058	2.75*	-0.023	2145
	(0.166)	(0.0679)	(0.143)	(1.673)	(0.124)	
BW of 0.4	0.14	-0.006	-0.036	2.97*	0.0013	3084
	(0.160)	(0.0681)	(0.134)	(1.573)	(0.116)	
BW of 0.5	0.14	-0.006	-0.031	3.04*	0.008	4116
	(0.160)	(0.0692)	(0.133)	(1.558)	(0.115)	
BW of 0.6	0.14	-0.005	-0.028	3.08**	0.01	5271
	(0.161)	(0.0699)	(0.133)	(1.553)	(0.114)	

Table 8: Robustness of RD Estimates: Academic Warning Policy

	GPA	CGPA	Warn	Credits	Major	Transfer	Exit	Obs
$C = 0.75, C = 1.0$								
BW of 0.2	0.03	0.14	0.007	-0.74*	0.03	0.079**	0.092**	2313
	(0.18)	(0.11)	(0.06)	(0.44)	(0.053)	(0.033)	(0.039)	
BW of 0.3	0.025	0.16	0.02	-0.51	0.06	0.076*	0.09**	3210
	(0.23)	(0.14)	(0.06)	(0.59)	(0.054)	(0.04)	(0.039)	
BW of 0.4	0.05	0.16	0.03	-0.36	0.08	0.079*	0.084**	4576
	(0.21)	(0.14)	(0.06)	(0.57)	(0.055)	(0.04)	(0.038)	
BW of 0.5	0.08	0.17	0.03	-0.24	0.06	0.082**	0.099***	5845
	(0.19)	(0.12)	(0.06)	(0.51)	(0.05)	(0.038)	(0.035)	
BW of 0.6	0.06	0.16	0.03	-0.22	0.05	0.071**	0.11***	6977
	(0.17)	(0.10)	(0.05)	(0.45)	(0.046)	(0.035)	(0.034)	

Table 9: Robustness of RD Estimates: Academic Warning Policy- GPA Threshold of 2.0

	GPA	CGPA	Warn	Credits	Major	Transfer	Exit	Obs
$C = 2.0$, Sem 3-6								
BW of 0.2	0.26	0.09	-0.02	-0.166	-0.0322	0.043	0.018	1746
	(0.16)	(0.087)	(0.099)	(0.69)	(0.095)	(0.09)	(0.09)	
BW of 0.3	0.23*	0.13*	-0.03	0.0226	0.0437	0.063	0.056	2600
	(0.13)	(0.08)	(0.08)	(0.56)	(0.07)	(0.07)	(0.064)	
BW of 0.6	0.17	0.11**	-0.005	0.363	0.081*	0.069*	0.058	5332
	(0.11)	(0.056)	(0.056)	(0.38)	(0.04)	(0.036)	(0.045)	
BW of 1.0	0.15*	0.095**	0.001	0.255	0.0405	0.0368	0.084**	8451
	(0.089)	(0.046)	(0.04)	(0.30)	(0.035)	(0.026)	(0.038)	

Table 10: The Impact on Long-Term Outcomes

	Graduate in 3 yrs or less	Duration in Months
Dean's: SS	0.09	-0.89
	(0.091)	(1.28)
Dean's: MPAS	-0.075	5.94
	(0.313)	(6.513)
Warn	-0.06	0.096
	(0.105)	(2.624)

7 Appendix: Warning Letter

Dear { Name } ,

This is to inform you that the Board of Examiners in accordance with Faulty Regulations and the University Examination Regulations, considered your examination results for the first Semester of the 2016/2017 academic year and concluded that your performance was unsatisfactory and may indicate serious weakness in your academic preparation. The Grade Point Average system (GPA), Regulation xxx indicates that a student whose GPA for a given semester is less than 2.00 shall be deemed to be performing unsatisfactorily, and shall be placed on WARNING. A student on warning whose GPA for the succeeding semester is less than or equal to 1.99, will be required to withdraw.

You are hereby notified that as a consequence of your academic performance, you are on now on WARNING and this is reflected on your academic record which may be viewed online. Please be aware that such performance, if repeated, would lead to your being asked to withdraw from the Faculty.

In order not to jeopardize your academic future, I am strongly recommending that if there are any significant academic or personal problems preventing you from functioning to the best of your ability, that you seek assistance by contacting your academic advisor, accessing the study skills programme offered by the Academic Support Unit, or utilizing appropriate counseling programme offered at the Health Centre.

We trust that you will take the necessary steps to rectify the situation and wish for you all the best in your future studies.

Yours sincerely,
{ Name }
Campus Registrar

References

Baumeister, R. F., Hutton, D. G., and Cairns, K. J. (1990). Negative effects of praise on skilled performance. *Basic and applied social psychology*, 11(2):131–148.

Bénabou, R. and Tirole, J. (2002). Self-confidence and personal motivation. *The Quarterly Journal of Economics*, 117(3):871–915.

Bettinger, E., Gurantz, O., Kawano, L., and Sacerdote, B. (2016). The long run impacts of merit aid: Evidence from californias cal grant. Technical report, National Bureau of Economic Research.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014a). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326.

Calonico, S., Cattaneo, M. D., Titiunik, R., et al. (2014b). Robust data-driven inference in the regression-discontinuity design. *Stata Journal*, 14(4):909–946.

Carruthers, C. K. and Özek, U. (2016). Losing hope: Financial aid and the line between college and work. *Economics of education review*, 53:1–15.

Cattaneo, M. D., Jansson, M., and Ma, X. (2016a). rddensity: Manipulation testing based on density discontinuity. *The Stata Journal (ii)*, pages 1–18.

Cattaneo, M. D., Titiunik, R., and Vazquez-Bare, G. (2016b). Inference in regression discontinuity designs under local randomization. *Stata Journal*, 16(2):331–367.

Cohodes, S. R. and Goodman, J. S. (2014). Merit aid, college quality, and college completion: Massachusetts' adams scholarship as an in-kind subsidy. *American Economic Journal: Applied Economics*, 6(4):251–285.

Elsner, B. and Isphording, I. E. (2017). A big fish in a small pond: Ability rank and human capital investment. *Journal of Labor Economics*, 35(3):787–828.

Fan, J. and Gijbels, I. (1996). *Local polynomial modelling and its applications: monographs on statistics and applied probability 66*, volume 66. CRC Press.

Feather, N. T. (1966). Effects of prior success and failure on expectations of success and subsequent performance. *Journal of personality and social psychology*, 3(3):287.

Fletcher, J. M. and Tokmouline, M. (2017). The effects of academic probation on college success: Regression discontinuity evidence from four texas universities.

Frandsen, B. R. (2017). Party bias in union representation elections: Testing for manipulation in the regression discontinuity design when the running variable is discrete. In *Regression Discontinuity Designs: Theory and Applications*, pages 281–315. Emerald Publishing Limited.

Gelman, A. and Imbens, G. (2014). Why high-order polynomials should not be used in regression discontinuity designs. Technical report, National Bureau of Economic Research.

Imbens, G. W. and Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of econometrics*, 142(2):615–635.

Lee, D. S. and Card, D. (2008). Regression discontinuity inference with specification error. *Journal of Econometrics*, 142(2):655–674.

Lee, D. S. and Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of economic literature*, 48(2):281–355.

Leeds, D. M. and DesJardins, S. L. (2015). The effect of merit aid on enrollment: A regression discontinuity analysis of iowas national scholars award. *Research in Higher Education*, 56(5):471–495.

Lin, T.-C. (2013). Student performance and subsequent effort increment investment: do students behave like producers? *International Journal of Education Economics and Development*, 4(3):219–232.

Lindo, J. M., Sanders, N. J., and Oreopoulos, P. (2010). Ability, gender, and performance standards: Evidence from academic probation. *American Economic Journal: Applied Economics*, 2(2):95–117.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of econometrics*, 142(2):698–714.

Seaver, W. B. and Quarton, R. J. (1973). Social reinforcement of excellence: Dean's list and academic achievement.

Stinebrickner, T. R. and Stinebrickner, R. (2011). Math or science? using longitudinal expectations data to examine the process of choosing a college major. Technical report, National Bureau of Economic Research.

Thistlethwaite, D. L. and Campbell, D. T. (1960). Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational psychology*, 51(6):309.

Venables, L. and Fairclough, S. H. (2009). The influence of performance feedback on goal-setting and mental effort regulation. *Motivation and Emotion*, 33(1):63–74.

Weiner, B., Heckhausen, H., and Meyer, W.-U. (1972). Causal ascriptions and achievement behavior: a conceptual analysis of effort and reanalysis of locus of control. *Journal of personality and social psychology*, 21(2):239.